Technische Universität München

Fakultät für Chemie

Biosystemchemie

Methods For The Isolation And Modification Of Novel Polycyclic Tetramate Macrolactams (PoTeMs)

Anna Glöckle

Vollständiger Abdruck der von der Fakultät für Chemie der Technischen Universität München zur Erlangung des akademischen Grades eines Doktors der Naturwissenschaften (Dr. rer. nat.) genehmigten Dissertation.

Vorsitzender: Prof. Dr. Michael Groll

Prüfer der Dissertation:

1. Prof. Dr. Tobias A.M. Gulder
2. Prof. Dr. Kathrin Lang

Diese Dissertation wurde am 28.09.2020 bei der Technischen Universität München eingereicht und durch die Fakultät für Chemie am 17.11.2020 angenommen.

List of Publications

Greunke, C.; **Glöckle, A.**; Antosch, J.; Gulder, T.A.M. Biocatalytic Total Synthesis of Ikarugamycin. *Angew. Chem. Int. Ed.* **2017**, *56*, 4351-4355, DOI: 10.1002/anie.201611063

Greunke, C.; Duell, E.R.; D'Agostino, P.M.; **Glöckle, A.**; Lamm, K.; Gulder, T.A.M. Direct Pathway Cloning (DiPaC) to unlock natural product biosynthetic potential. Metab. Eng. **2018**, *47*, 334-34, DOI: 10.1016/j.ymbem.2018.03.010

Glöckle, A.; Gulder, T.A.M. A Pericyclic Reaction Cascade in Leporin Biosynthesis. *Angew.Chem. Int. Ed.* **2018**, 130, 2802-2804, DOI: 10.1002/anie.201800629

B1.1

Abstract

In times of increasing antibiotic resistance, the discovery of novel natural products (NPs) has evolved an emerging significance as they represent a great majority of currently FDA-approved new chemical entities and thus comprise the potential for future drug leads. Progress in whole genome sequencing techniques has revolutionized NP discovery by transitioning from simple cultivation of bacterial strains and isolating compounds to a targeted search of biosynthetic gene cluster (BGCs) within the sequenced genomes of microorganisms. This offers the possibility to get access to novel NPs from minor amounts of genetic material by capturing the clusters of interest and heterologously expressing target compounds. However, currently used methods for gene cluster capturing are highly labor- and time-intensive.

This thesis aims to contribute to a novel, time-effective gene cluster capturing method called 'Direct Pathway Cloning' (DiPaC) which is based on long amplicon PCR amplification of the BGC of interest with subsequent insertion into desired expression vectors by Gibson assembly. Furthermore, using this method, NPs belonging to the class of polycyclic tetramate macrolactams (PoTeMs) are to be heterologously expressed in *Streptomyces* to identify and modify novel molecules. This class of NPs exhibit great pharmacological value, as the scope of biological activities ranges from antibacterial to antifungal or even cytotoxic. Furthermore, the biosynthesis of structurally diverse PoTeMs is based on a single, common precursor. This enables a broad range of heterologous expression and engineering possibilities.

Previous results identified the *ika* BGC to be responsible for ikarugamycin (**39**) biosynthesis. However, a heterologous expression attempt in *E. coli* gave insufficient yields for further applications. In this study, it was possible to efficiently clone the *ika* BGC and optimize heterologous expression in *Streptomyces* to obtain a final isolated yield of 25 mg/L. Furthermore, the heterologous expression of the two biosynthetic intermediates of **39**, **53** and **54**, was accomplished. Unfortunately, it was not possible to express novel PoTeMs by capturing and expressing the native BGCs *spi* (*S. spinosa*), *deg* (*S. degradans*), *fer* (*P. fermentans*), *Tü6314* (*Streptomyces* sp. Tü6314), and *ros* (*S. roseosporus*).

The failure in heterologously expressing novel PoTeMs led to the design of a plug-and-play expression system based on *ikaA*, the gene responsible for the production of the common PoTeM precursor **53**. This approach contained an expression vector set of plasmids based on pSET152_ermE::*ikaA* with a second promoter downstream of *ikaA*. For this study, nine different constitutive promoters were selected. Additional restriction sites facilitated the flexible integration of single PoTeM BGC parts. The system was established using the *ika* BGC and modifying enzymes DegA, PtmD, and IkaD, and successfully applied using CftA. As proof of

principle, this study enabled the determination of enzyme functions and the expression of derivatives of **39** in decent amounts.

The established plug-and-play system was further used for the heterologous expression of the three additional clusters *spi*, *deg*, and *fer*. The production of a novel compound was detectable for all the three BGCs integrated into the plug-and-play system. However, amount and stability of the products were not sufficient for structure elucidation. Nevertheless, this study allowed understanding of promoter use for heterologous expression in *Streptomyces*. Promoters *gapdhP*(EL) and *rpsLP*(XC) were identified to be highly suitable for the production of stable compounds in *S. albus*, whereas *kasOP** and *SF14P* promoters showed the best results for non-stable and more complex compounds both in *S. albus* as well as *S. coelicolor* strains. *S. lividans* was determined as not suitable for PoTeM heterologous expression. Summing up, before the heterologous expression of an NP, genetic difference between original producer and heterologous host, as well as possible stability and biosynthetic timing of the product must be considered.

Zusammenfassung

In Zeiten eines erhöhten Aufkommens an Antibiotikaresistenzen hat die Entdeckung neuer Naturstoffe (NP) an Bedeutung zugenommen, da NPs die Mehrzahl aktueller, neuer FDA-zugelassener chemischer Strukturen darstellt. Fortschritte in Sequenzierungsmethoden haben die Entdeckung neuer NPs revolutioniert, indem die einfache Kultivierung von Mikroorganismen und die Isolierung der vorhandenen NPs zu einer gezielten Suche nach Biosynthesegencluster (BGC) weiterentwickelt wurde. Dies eröffnet die Möglichkeit, Zugriff auf NPs auch aus minimalem Mengen genetischen Materials zu erhalten, indem die BGC kloniert werden und anschließend die jeweiligen Moleküle heterolog exprimiert werden können. Die derzeitigen Methoden zur Klonierung von BGCs sind jedoch noch sehr arbeits- und zeitintensiv.

Diese Arbeit soll zur Entwicklung einer neuen, zeiteffizienten Cluster-Klonierungsmethode, dem sogenannten ‚Direct Pathway Cloning' (DiPaC), beitragen. Diese Methode basiert auf einer PCR Amplifikation der BGCs in einem Fragment und einer anschließenden Integration in einen Expressionsvektor durch Gibson Assembly. Unter Verwendung dieser Methode sollen neue und modifizierte Stoffe der Naturstoffklasse polyzyklischer tetratmsäurehaltiger Makrolaktame (PoTeMs) heterolog in *Streptomyceten* exprimiert werden. Die Naturstoffklasse der PoTeMs ist pharmakologisch interessant, da die einzelnen Vertreter ein großes Spektrum an biologischen Aktivitäten aufweisen. Diese reichen von antibakteriell über antifugal bis hin zu zytotoxisch. Zudem zeichnet sich die Biosynthese der strukturell vielfältigen PoTeMs durch ein gemeinsames Vorläufermolekül aus. Diese Eigenschaft ermöglicht die unterschiedlichsten Herangehensweisen für die heterologe Expression und gentechnische Modifizierung.

In vorangegangen Studien konnte gezeigt werden, dass das *ika* BGC für die Biosynthese von Ikarugamycin (**39**) verantwortlich ist. Die Ausbeute einer ersteren heterologen Expression in *E. coli* war jedoch für weitere Anwendungen unzureichend. Im Rahmen der vorliegenden Arbeit war es möglich das *ika* BGC effizient zu klonieren und ein heterologes Expressionssystem in *Streptomyceten* zu etablieren, sodass sich eine finale isolierte Ausbeute von 25 mg/L ergab. Zudem konnte die heterologe Expression der biosynthetischen Intermediate von **39**, **53** und **54**, erarbeitet werden. Es war jedoch nicht möglich, die Naturstoffe der BGC *spi* (*S. spinosa*), *deg* (*S. degradans*), *fer* (*P. fermentans*), *Tü6314* (*Streptomyces* sp. Tü6314), and *ros* (*S. roseosporus*) nach einer erfolgreichen Klonierung funktional zu exprimieren.

Das Scheitern der heterologen Expression von neuen PoTeM-Molekülen hat zum Design eines neuartigen plug-and-play Expressionssystems geführt. Dieses basiert auf *ikaA*, dem Gen, das für die Biosynthese des gemeinsamen Vorläufermoleküls (**53**) zuständig ist. Diese

Herangehensweise beinhaltet neun Expressionsvektor die auf pSET152_ermE::*ikaA* mit einem nachgeschalteten Promoter beruhen. Hierfür wurden neun unterschiedliche konstitutiv-aktive Promotoren ausgewählt. Zusätzlich wurden Restriktionsschnittstellen eingefügt, die eine flexible Integration einzelner BGC-Fragmente ermöglichen. Eine Etablierung diese Systems wurde mit dem *ika* BGC und den modifizierenden Enzymen DegA, PtmD und IkaD durchgeführt und erfolgreich mit CftA angewandt. Hiermit konnte bestätigt werden, dass die Anwendung des etablierten Systems zur Darstellung modifizierter PoTeMs geeignet ist. Zudem wurden Enzymaktivitäten charakterisiert und Derivate von **39** produziert.

Nachfolgend konnte das neue Expressionssystem für die heterologe Expression der zusätzlichen BGCs *spi*, *deg* und *fer* verwendet werden. Hierbei war die Detektion neuen Stoffe durch die Integration der Cluster in das plug-and-play System möglich. Jedoch war weder die Menge noch die Stabilität der neuen Moleküle ausreichend für eine Strukturaufklärung. Nichts destotrotz ermöglichte die Studie ein besseres Verständnis für die Nutzung von Promotoren in *Streptomyceten*. Die Promotoren *gapdhP*(EL) und *rpsLP*(XC) haben sich als sehr geeignet für die Produktion von stabilen Molekülen in *S. albus* erwiesen, wohingegen die Promotoren *kasOP** und *SF14P* die besten Ergebnisse für instabile oder komplexe Moleküle in *S. albus* und *S. coelicolor* erzielt haben. Der heterologe Wirt *S. lividans* erwies sich als ungeeignet für die Produktion von PoTeMs. Zusammenfassend lässt sich sagen, dass vor der heterologen Expression von NPs der genetische Abstand zwischen dem natürlichen Produzenten und dem heterologen Host beachtet werden muss. Zudem beeinflussen die voraussichtliche Stabilität und das biosynthetische Timing die Produktionsmengen und dementsprechend die zu wählenden heterologen Expressionsbedingungen.

Table of contents

1 Introduction

1.1 Natural Products in Drug Discovery

Secondary metabolites are complex molecules produced in a great variety of living organisms in addition to their primary metabolites. While primary metabolism is essential for energy supply and reproduction, secondary metabolites are mainly used for communication between different organisms or diverse cells in one organism, conferring an evolutionary benefit for the producer by, e.g., defense against predators or by attracting symbionts.[1] The biological effects of some of these natural compounds have been used by humans for centuries to improve their lives, without neither the knowledge of the compounds' structure nor their mechanisms of action. Some prominent examples comprise the natural compound derived acetylsalicylic acid (**1**), morphine (**2**), or tubocurarine (**3**) amongst many others (Figure 1). Compound **1**, also known as Aspirin,® is a derivative of the analgesic and antipyretic compound contained in the bark of *Salix alba* that was formally administered as a plant extract, as mentioned in the *Corpus Hippocraticum*.[2] The first isolated active compound was identified as salicin and is nowadays marketed in its oxidized and acetylated form Aspirin® by Bayer.[3],[4] Morphine (**2**) has been used for several decades for its analgesic effect, originally applied as a crude extract from dried latex of *Paper somniferum* seedpods. Around 1805 Sertürner was able to isolate **2** from the opiate extract[5] making this discovery a milestone to one of the most valuable drugs in modern medicine, as **2** and its derivatives are the most frequently applied drugs in chronic pain patients and also recommended to treat severe pain by the WHO.[6] Compound **3** is one of the active compounds isolated from the arrow poison 'Curare', used by ancient South Americans for the efficient killing of their prey, and was isolated from *Chondrodendron tomentosum*.[7] The chemical property of 'Curare' to be poisonous when directly injected into the blood, but not passing through mucosal membranes, rendered it useful for hunting. Once the active compound was isolated, its muscle relaxing function could be applied pharmacologically, e.g. during surgery.[8],[9] These few molecules already depict the great impact natural products (NPs) had and still have on human well-being.

There are many more molecules isolated from natural sources that are either applied directly or modified prior to application as pharmaceuticals today.[10] However, the process of discovery has gone through some transformation over the past decades. The first isolated natural compounds were extracted from the biological material that has been applied for decades or even centuries and whose effects have been identified through long-lasting human experience.[11] However, through the accidental discovery of penicillin G (**4**) by A. Fleming in 1928[12] the general approach to find new bioactive natural products changed. A. Fleming observed an inhibiting areola in a *Staphylococcus* culture contaminated with *Penicillium*

notatum spores, leading to the assumption that an antibiotic compound was produced by the contaminant. After considerable amount of isolation efforts, it was possible to characterize **4**, and after further labor-intensive work for the optimization of fermentation, this compound did pioneer work in the treatment of infectious diseases.[13] Additionally, the discovery of an anti-infective drug by applying a possible natural compound containing extract to a pathogen revolutionized NP research and drug discovery. In the 1940s the first selective screenings for new antibiotics from nature were started. In this context, the Waksman group accomplished cutting-edge isolation efforts to identify streptomycin (**5**) from *Streptomyces griseus*.[14] Amongst all the other newly discovered compounds by this group, this molecule showed strong activity against gram-positive and gram-negative bacterial pathogens and still belongs to the most valuable drugs in the history of tuberculosis treatment.[15] These first achievements promoted pharmaceutical companies as well as academia to collect many samples, preferably soil samples with more recently extension to marine sources, from all over the world to test the contained microorganisms for the production of antibacterial compounds.[16] However, the effort for identifying a single new molecule is quite substantial, as primary hits in inhibitory assays need product upscaling to allow purification of the single extract components and subsequent structure elucidation of the active molecule.[17] Nevertheless, during this period, also called the 'golden age', many new molecules with completely new molecular architectures were identified.

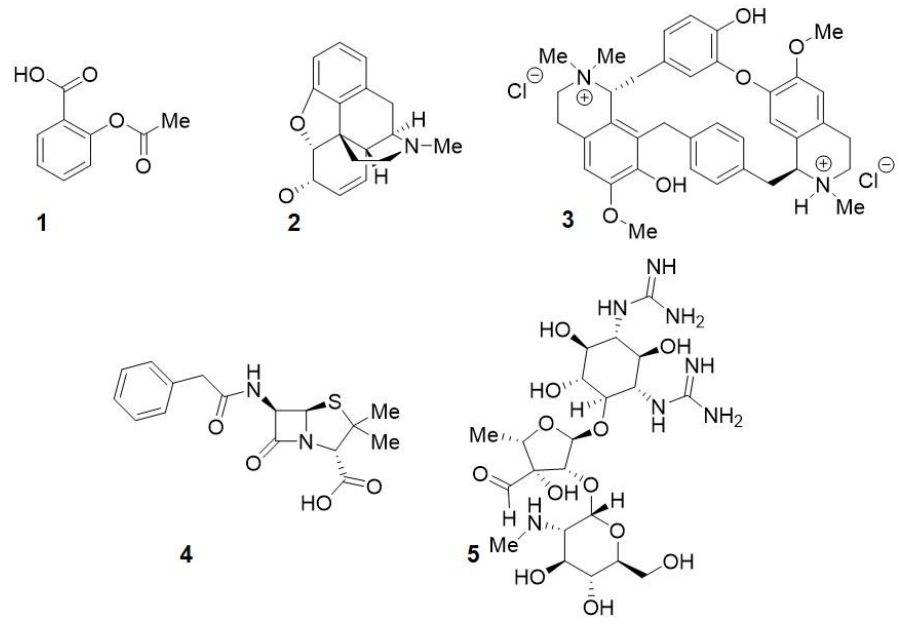


Figure 1. Examples of natural (derived) compounds with historic interest: acetylsalicylic acid (**1**), morphine (**2**), tubocurarine (**3**), penicillin G (**4**), and streptomycin (**5**).

During initial screenings, which mainly focused on the identification of anti-infective drugs, molecules like the glycopeptide vancomycin (**6**), the macrolide tylosin (**7**), or the lipopeptide

daptomycin (**8**) were discovered (Figure 2). These three examples depict NPs with sophisticating chemical frameworks that were all first isolated from microorganisms found in soil samples during phenotypic screenings, **6** from *Amycolatopsis orientalis*,[18],[19],[20] **7** from *Streptomyces fradiae*,[21],[22] and **8** from *Streptomyces roseosporus*.[23] Furthermore, all these compounds are still valuable drugs in human and veterinary medicine. Whereas **6** has been mostly overlooked after its initial application in 1958, it regained recognition after the appearance of methicillin-resistant bacteria and today it belongs to one of the antibiotics of last resort to fight multi resistant germs.[24],[25] Compound **7** is applied as an antibiotic to treat infections in animals[26],[27],[28] and **8** is an effective antibiotic to fight skin infections.[29],[30],[31] The growing knowledge in diseases also enlarged the knowledge about possible drug targets, leading to more ambitious phenotypic screenings. These included enzyme-based assays to find inhibitors or activators for specific signal transduction- or biosynthesis pathways. A prominent example of a NP found in an advanced screening is lovastatin (**9**). It was isolated by two different groups almost simultaneously from *Monascus ruber*[32] and *Aspergillus terreus*.[33] By inhibiting the HMG-CoA-reductase it functions as a cholesterol lowering agent. The examples depict the great success of NPs in drug discovery until the 1980s, with numerous impressive new structures identified. Until today, NPs and NP derivatives generate a large quantity of clinically approved drugs.[34]

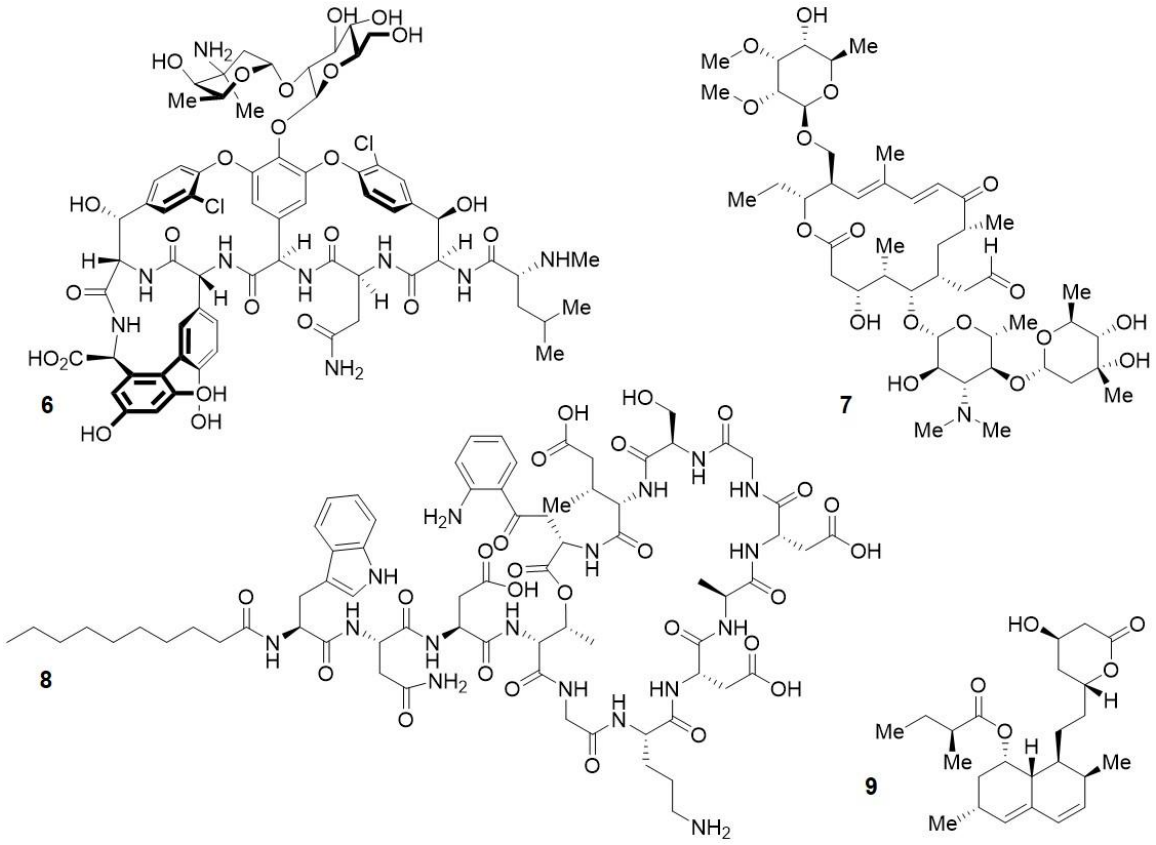


Figure 2. Examples of natural products isolated in the golden age of NP discovery: vancomycin (**6**), tylosin (**7**), daptomycin (**8**), and lovastatin (**9**).

In the beginning of 1990, the improvements in combinatorial chemistry along with the introduction of high throughput screenings (HTS) facilitated the analysis of large chemical libraries to find new drug-leads.[16] Also due to the upcoming problem of reisolating already known NPs, chemical libraries thus replaced the extract collections in pharmaceutical companies.[11] Current perceptions, however, show that the discovery of new chemical entities (NCE) based on these libraries is rather unsatisfactory. For a long time, sorafenib (**10**) was the only example of a drug comprising a NCE. With vemurafenib (**11**) and ataluren (**12**), this list could be expanded in 2011 and 2014, respectively.[35] Additionally, substantial screenings have failed to promote the stagnating discovery of new antibiotics, a necessity in times of exceeding antibiotic resistance.[36] These findings showed researchers that pure combinatorial chemistry libraries contain a too small chemical diversity, for example considering variable cyclic structures or stereochemistry, when compared to complex NP collections.[37] Redirecting drug research to libraries consisting of compounds which reassembled NP entities, a concept called ‘diversity-oriented synthesis’ (DOS), was established to solve this problem. These natural product derived libraries were constructed using three different approaches to reassemble NP diversity: (a) they exhibited the structural skeleton of an active compound, (b) they included a specific feature of a NP class or (c) they were constructed to mimic common features found in NPs.[38] Including the advancements in organic chemistry, DOS seems to offer a helpful tool combining NPs and combinatorial chemistry to find new lead structures and even determine protein interactions,[39] however this approach is based on already isolated compounds and does not serve the need of NCE discovery to fight severe infectious diseases.

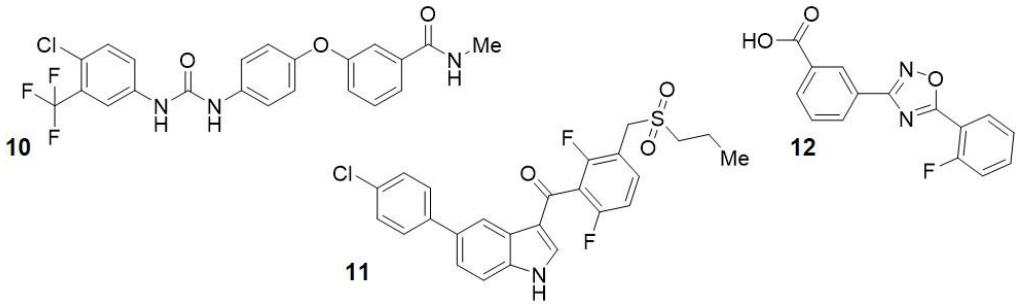


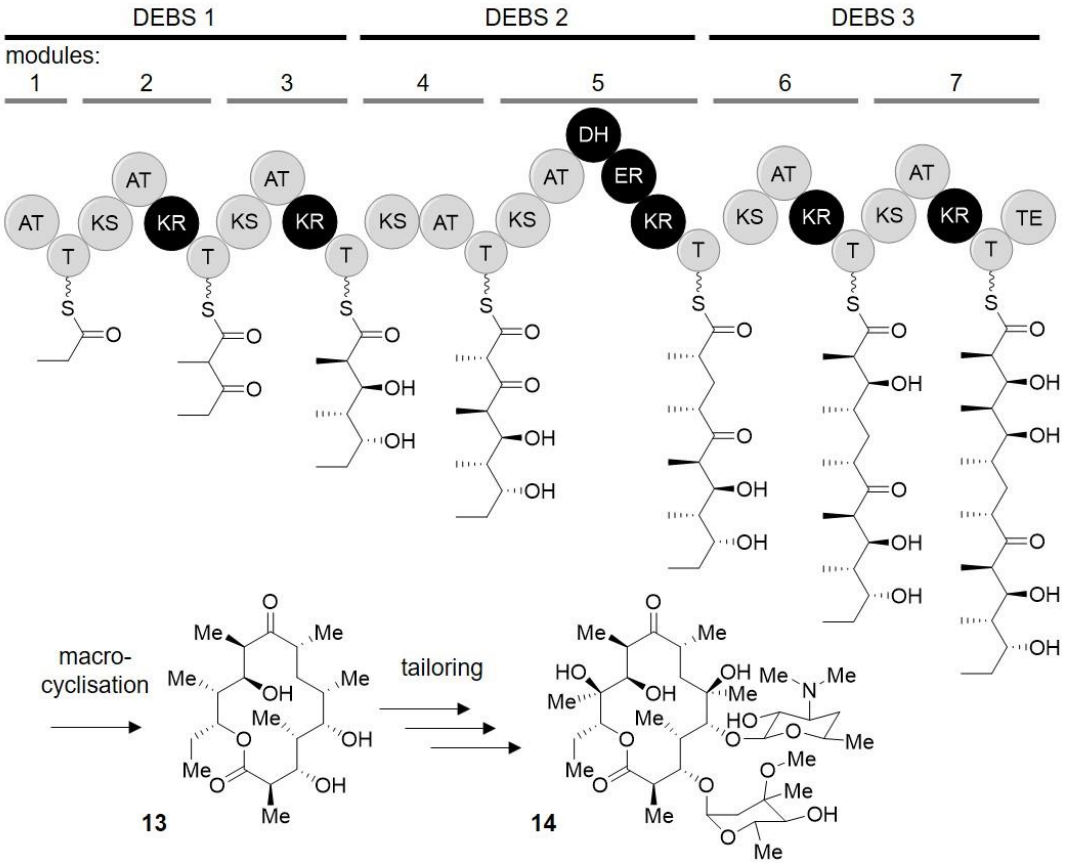
Figure 3. Examples of approved drugs with NCE derived from combinatorial libraries: sorafenib (**10**), vemurafenib (**11**), and ataluren (**12**).

1.2 Biosynthesis of Natural Products

While the identification of new NPs stagnated, there was an extensive gain of knowledge in understanding the biochemical mechanisms underlying NP biosynthesis. A crucial finding in NP research was the fact that genes responsible for the biosynthesis of NPs cluster in close proximity to each other in the genome of the producer, especially in bacteria. However, this is

not usually the case in plants, rendering biosynthetic elucidations in plants highly complex.[40] Amongst the different NP classes some are constructed in an assembly-line manner, such as fatty acid synthases (FAS), including polyketides (PKs), terpenes or non-ribosomal peptides (NRPs).[41] One of the first elucidated biosynthetic pathways to prove this kind of assembly was the biosynthesis of 6-deoxyerytronolide B (6-DEB) (**13**), the polyketide backbone of erythromycin A (**14**), a commonly used antibiotic.[42],[43],[44],[45] Two independent groups were able to prove that large multi-modular enzyme complexes construct **13** in an assembly-line fashion, while stepwise incorporating single C_3 units. [46],[47] These groundbreaking findings and subsequent research enabled a quite distinct understanding of the enzymatic reactions underling PK- as well as NRP biosynthesis. The biosynthesis of those pathways is performed while the growing PK- or NRP chain is covalently bound to the multi modular enzyme complex, proceeding from module to module. Building blocks for polyketide synthases (PKSs) are quite limited, including the activated thioesters acetyl-CoA (**15**), malonyl-CoA (**16**) or methyl malonyl-CoA (**17**). However, the incorporation of more diverse building blocks has been proven but requires additional biosynthetic genes to access these units. An example is chloroethylmalonly-CoA in salinosporamid biosynthesis.[48] The PKS machinery is composed of several modules with each module containing several domains. The four obligatory domains (depicted in light grey, Scheme 1) are the acyltransferase (AT), the thiolation (T), the keto synthase (KS), and the final thioesterase (TE) domain. The AT domain is responsible for substrate specificity and attachment of the substrate to the T domain. The latter domain serves as an anchor during chain elongation, whereas the KS domain catalyzes the elongation of the PK chain. The last obligatory domain, the TE domain, is required to release the final product. The first module usually consists only of an AT- and a T-domain, the last module contains the additional TE domain for product release. Essential for the anchoring of the substrate to the T-domain is a previous incorporation of a phosphopantetheinyl prosthetic group to a conserved serine residue in the T-domain active site that can form the thioester to retain the substrate building block. This transformation to the holo-form of the domain is catalyzed by a phosphopantetheinyl transferase, such as Sfp.[49] In addition to the obligatory domains, there are several optional tailoring domains (depicted in black, see Scheme 1) responsible for the great diversity of polyketide natural products that are otherwise strongly restricted in building block diversity. Examples are the reductive domains β-keto reductase (KR), β-hydroxy acyl dehydratase (DH) or α,β-enoyl reductases (ER) amongst others. In Scheme 1 the exemplary biosynthetic pathway to produce **13** shows the stepwise elongation catalyzed by the different domains and the final cleavage and cyclization to obtain **13**, which will subsequently be modified via hydroxylation and glycosylation catalyzed by tailoring enzymes to produce **14**. The PKS is structured in three synthases (DEBS1-3) containing a total of seven modules, each with 2-6 catalytically active domains. Propionyl-CoA is the starter unit that gets elongated with

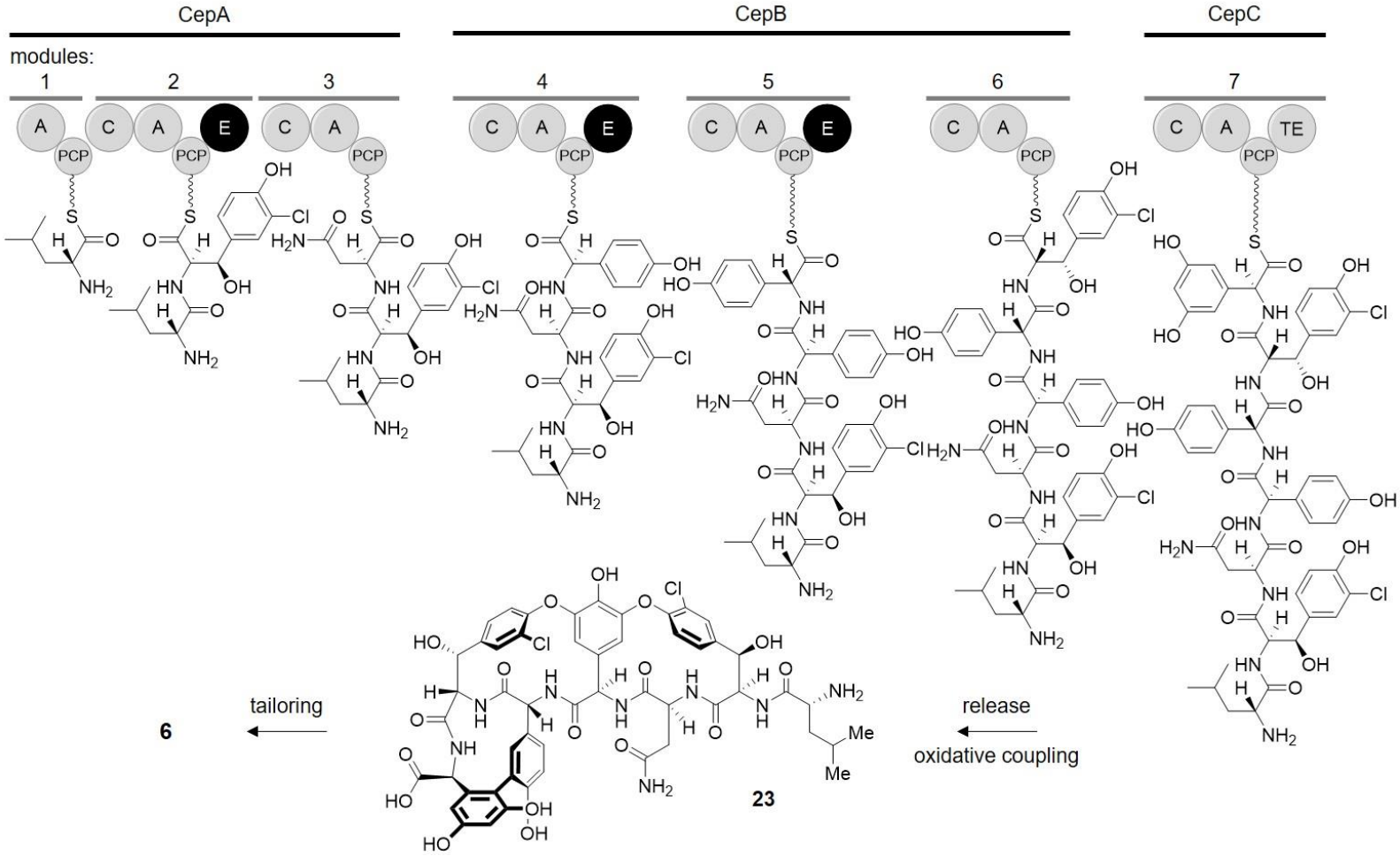
six methyl malonyl-CoAs to incorporate an α-branched C_3 at each module. Further, the newly incorporated entity is either unmodified (module 4), reduced to the β-hydroxy function (modules 2, 3, 6, 7) or completely reduced (module 5). The final TE-domain in module 7 catalyzes the macrocyclization of the hydroxy function introduced in the first biosynthetic step with the carbonyl carbon of the last incorporated methyl malonyl unit to release the final product **13**.



Scheme 1. Biosynthesis of 6-DEB (**13**) as an example for PK biosynthesis. The PKS is composed of three different synthases (DEBS1-3) with a total of 7 modules, each responsible for the incorporation of a C_3 unit and subsequent reduction steps. The final macrocyclization is catalyzed by the TE domain to form **13**, which is further modified by different tailoring enzymes to synthesize the product erythromycin A (**14**). Obligatory PKS domains for any PKS are depicted in light grey, optional domains, however, mandatory for this specific biosynthetic pathway, are shown in black. AT - acyltransferase, T - thiolation, KS - keto synthase, KR - β-keto reductase, DH - β-hydroxy acyl dehydratase, ER - α,β-enoyl reductase, TE - thioesterase.

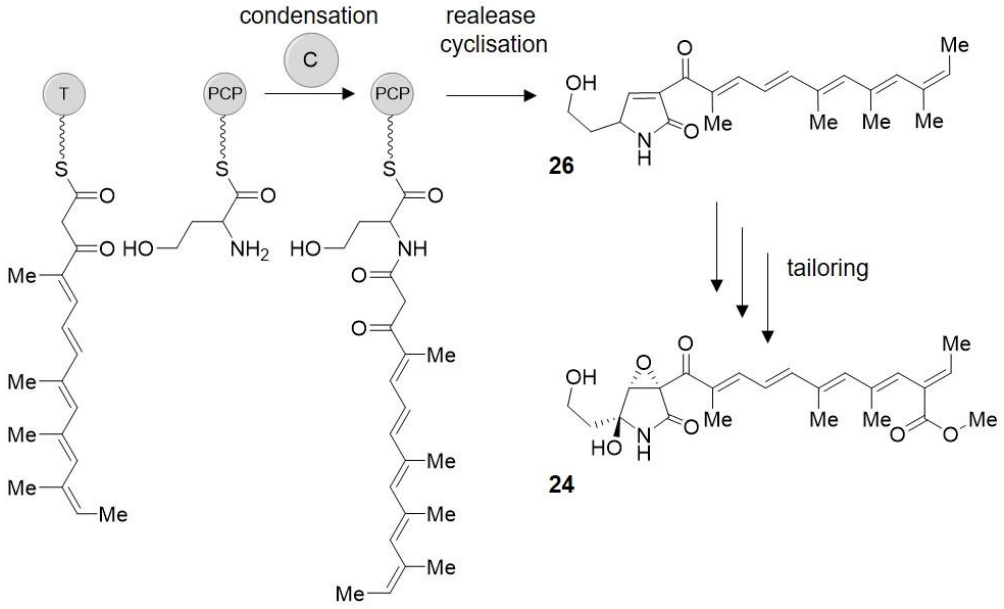
The biosynthesis of NRPs is highly similar to PKs, the assembly being facilitated by the multi-modular enzyme complex called non-ribosomal peptide synthetase (NRPS). As in PKS, the synthetase is arranged in a modular fashion, again comprising four mandatory domains (depicted in light grey, Scheme 2). First, the adenylation (A) domain is responsible for substrate selection, activation and subsequent transfer to the peptidyl carrier protein (PCP), the second mandatory domain. To enable binding of the substrate to the PCP-domain, it must also be posttranslational converted to its holo-form by the transfer of a phosphopantetheinyl group to a specific active-site serine residue. The peptide bond is established between two neighboring

amino acids (AAs) bound to PCP_n and $PCP_{(n-1)}$ catalyzed by the condensation (C) domain C_n. Identical to the PKS system, the first NRPS module usually consists only of an A- and a PCP-domain, whereas the last module contains the last mandatory domain, the thioesterase (TE) domain, which is responsible for the release of the compound.[41] Additional optional domains (depicted in black, Scheme 2) can further modify the peptide chain. Amongst others there are the epimerization (E) domains changing the natural L-configuration of the respective AA to a D-configuration, or the methyltransferase (MT) domains catalyzing AA *N*-methylation. However, there is a huge difference compared to PKs concerning the possible building blocks that can be selected by the A-domain of NRPSs. Whereas there are mainly three elongation units for PKSs, there are more than 500 (non-) proteinogenic AAs that can be incorporated by NRPSs, with the biosynthetic genes for these amino acids usually being found in the biosynthetic gene cluster (BGC) producing the compound.[50] A well understood NRPS pathway is the biosynthesis of **6** that incorporates seven AAs of which five are non-proteinogenic AAs with dedicated biosynthetic pathways (β-hydroxychlorotyrosine (β-OH-Cl-Tyr) (**18**), L-4-hydroxyphenylglycine (HPG) (**19**), and L-3,4-dihydroxyphenylglycine (DPG) (**20**)) and four AAs are configured in the unnatural D-configuration (D-Leucin (Leu) (**21**), D-β-OH-Cl-Tyr, and D-HPG). The non-proteinogenic HPG (**19**) depicts a crucial structural feature of vancomycin-type antibiotics and therefore its biosynthesis is well studied. Starting with labeling experiments to prove its origin from tyrosine,[51],[52] a biosynthetic pathway including the transformation to *p*-hydroxyphenylpyruvate building a catalytic cycle to HPG was established.[53] The presence of homologues of the enzymes proposed to catalyze HPG synthesis, such as HmaS, Hmo and HpgT, in different organisms producing vancomycin-like antibiotics strongly supports the proposed biosynthetic route.[54] The assembly of the AAs constructing **6** has been nicely shown for the chloroeremomycin (**22**) BGC, whose main difference is the subsequent glycosylation pattern.[55] The synthetase consists of three enzymes with a total of seven different modules, with 2-4 catalytically active domains. Whereas the first incorporated AA, D-Leu, is already incorporated in the unnatural configuration, modules 2, 4, and 5 contain E-domains to change the AA configuration from L to D for β-OH-Cl-Tyr and HPG. The cyclic structure is established by two biaryl-/ biarylester-bond formations, obtaining the rigid structure of the peptide backbone of this NP that is crucial for its antibiotic activity and subsequently the NRPS product **23** is released from the enzyme complex.[56] Furthermore, tailoring enzymes add an N-methylation[57] and the glycosylation pattern[58] to finish the biosynthesis of **6**.



Scheme 2. Biosynthesis of vancomycin (**6**) as an example of NRPS. The NRPS consists of three enzymes (CepA-C) containing seven modules with 2-4 catalytically active domains. Incorporation of the AAs D-Leu, L-β-OH-Cl-Tyr, L-Asn, L-HPG, L-HPG, L-β-OH-Cl-Tyr, L-DPG, with epimerization domains in modules 2,4,5 to convert the L-AA into D-AA. Biarylester bonds are introduced by P450 enzymes to the PCP-bound intermediate. Final release of the product through hydrolysis by the TE-domain. Domains mandatory for any NRPS biosynthetic pathway are depicted in light grey, optional domains, however, mandatory for this specific biosynthetic pathway, are colored in black. A - adenylation, PCP - peptide carrier protein, C - condensation, E - epimerization, and TE - thioesterase.

Utilizing the identical mode of action as well as the general composition of PKS and NRPS there are additional NPs that combine both machineries, so called hybrid PKS-NRPS.[41] Crucial for this machinery to work is the ability of a C-domain directly following the PKS system to accept (methyl-)malonyl acyl groups as a substrate and to establish an amide bond as well as for the KS-domain, subsequent to a NRPS, to utilize a peptidyl chain as substrate.[41],[59] Amongst the first NPs to be identified as a product of a hybrid PKS-NRPS was furasin C (**24**), a potent mycotoxin, isolated from various species such as *F. moniliforme*[60] and *F. venenatum*.[61],[62] The identified hybrid PKS-NRPS, *fusA*, builds a polyketide chain from one unit of **15** and six units of **16**, which is subsequently coupled to L-homoserine (**25**) and released from the enzyme by the formation of a tetramic actid moiety to form prefusarin C (**26**). Further modifications lead to the final product **24** (Scheme 3).



Scheme 3. Crucial steps in fusarin C (**24**) biosynthesis exemplarily for a hybrid PKS-NRPS. Depicted is the amide bond formation between the polyketide precursor and L-homoserine (**25**) catalyzed by the C-domain to obtain the precursor **26**, as the connecting step between PKS and NRPS, with a final release from the enzyme complex, a cyclisation and additional tailoring to receive the product **24**.

An interesting biosynthesis which might be considered between PKS and hybrid PKS-NRPS is that of lovastatin (**9**). Two PKSs make up the polyketide backbone, LovB and LovF,[63] where LovB shows high analogy to known hybrid PKS-NRPS systems, with a condensation domain of a function yet to be determined.[64] These examples show the great opportunities arising with the assembly line biosynthetic logic, creating all kinds of chemical entities, with many more structural features and biosynthetic particularities still to be determined.

1.3 Natural products in the genomic area

Besides the extended gain of understanding biosynthetic pathways producing NPs, great improvements were accomplished in sequencing techniques, rendering sequencing of whole genomes possible. In 1995, the first bacterial genome was sequenced[65] and with accumulating numbers of sequenced genomes it became obvious that microorganisms contain a much greater number of possible BGCs than NPs found under laboratory conditions.[66] Apart from the great number of possible BGCs in cultured microorganisms, it is proposed that up to 99% of bacteria and 95% of fungi have not yet been cultured,[17] leaving an enormous potential to find more NPs with NCE or pharmacological activities.

To use this newly available genomic information about BGCs in microorganisms, several approaches were developed to isolate new NPs. The 'One Strain - Many Compounds' (OSMAC) approach relies on the fact that many BGCs are not active under standard, quite artificial, laboratory cultivation conditions and that simple variations in cultivation procedures might enable biosynthesis of formerly silenced clusters.[67] This assumption derived from the idea that microorganisms need to change their metabolism according to different environmental influences, considering, for example, starvation, communication or defense mechanisms. In addition, it has already been shown that certain components in cultivation media can increase or decrease the production of different secondary metabolites.[68],[69] Therefore, this method utilizes systematic diversification of culturing conditions regarding general requirements like pH value, temperature, carbon, trace element, or oxygen supply as well as physical characteristics of the cultivation tank. The identification and production optimization of hexacyclinic acid (**27**) from *Streptomyces cellulosae* (strain S1013) by changing cultivation tanks, alkali halide concentrations and carbon sources[70] and the conversion of the NP profile of *Streptomyces* sp. (strain A1) from mainly rubromycin (**28**) to streptazoline-like compounds (**29**) when media composition was changed,[71],[72] are only two examples proving the great scope of applications to identify new compounds or optimize production titers using OSMAC. However, as the precise mechanisms of biosynthetic pathway activation underlying this approach are not yet determined and likely different from case to case the activation of pathways to new compounds is quite random on a labor-intensive trial and error basis.

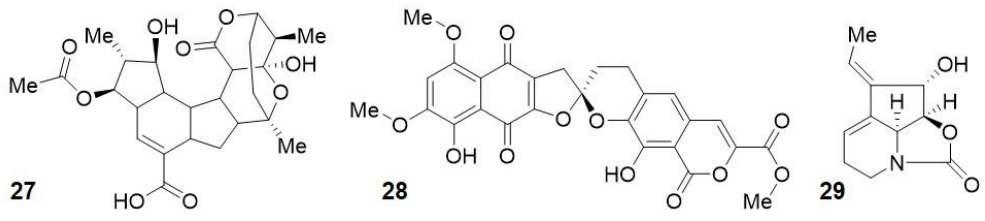


Figure 4. Examples of NPs discovered using the OSMAC strategy. Hexacyclinic acid (**27**), rubromycin (**28**) and streptazolin (**29**).

A different strategy to use the new knowledge of NP biosynthesis is combinatorial biochemistry, which utilizes single steps of biosynthetic pathways to engineer the production of new compounds. As the modularity of PKS and NRPS seems to enable direct prediction of the structures of the resulting NPs, the substitution of single modules could lead to an incorporation of a different building block, the inactivation of single domains could change reduction states or stereochemistry, and the replacement of different tailoring enzymes could result in new 'unnatural' NPs.[73] The first example of such an approach was performed even before whole genome sequencing was used to identify possible NP biosynthetic pathways. Genes from the captured actinorhodin (**30**) biosynthetic pathway were transferred into producers of different NPs, amongst others *Streptomyces* sp. AM-7161, the producer of medermycin (**31**), resulting in a new medermycin derivative containing an actinorhodin-like hydroxylation (**32**) (Figure 5).[74] More recent approaches were able to change glycosylation patterns in glycopeptides[75] or even the fatty acid chain present in lipopeptides, to produce whole libraries of novel lipopeptide compounds.[76],[77] In a different study, the inactivation of unique domains in the biosynthetic pathway of **13** established a great variety of precursors to produce erythromycin analogues.[78] These few examples underline the enormous opportunities of combinatorial biosynthesis to detect or engineer NPs. One issue of this approach is the substrate promiscuity of tailoring enzymes to make further modification of precursors possible. Another great burden for complete engineering of NP biosynthetic pathways by combining domains of different pathways is the uncertainty of how much interaction occurs between different domains and modules and how interfaces need to be designed to be recognized by a non-cognate module.[79]

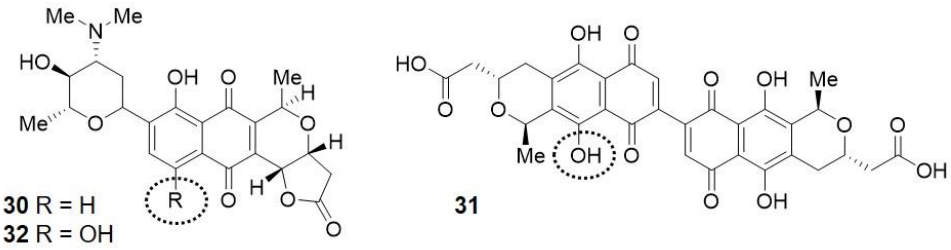


Figure 5. Examples of NPs discovered or modified using combinatorial biochemistry. Actinorhodin (**30**), medermycin (**31**) and the modified compound actinorhodin containing the medermycin-like modification (**32**).

Besides the approaches to activate or engineer BGCs in the original producer, heterologous expression of the whole native or refactored cluster in a suitable heterologous host represents an emerging possibility to get access to new NPs. An impressive example of this method is the discovery of the methylarcyriarubin (**33**) BGC *mar* from an environmental DNA (eDNA) soil sample library.[80] The isolation and structure elucidation of this molecule depicts that a tiny

amount of DNA is sufficient to express BGCs without requiring cultivation of the natural producer.

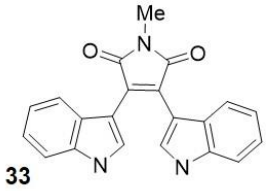
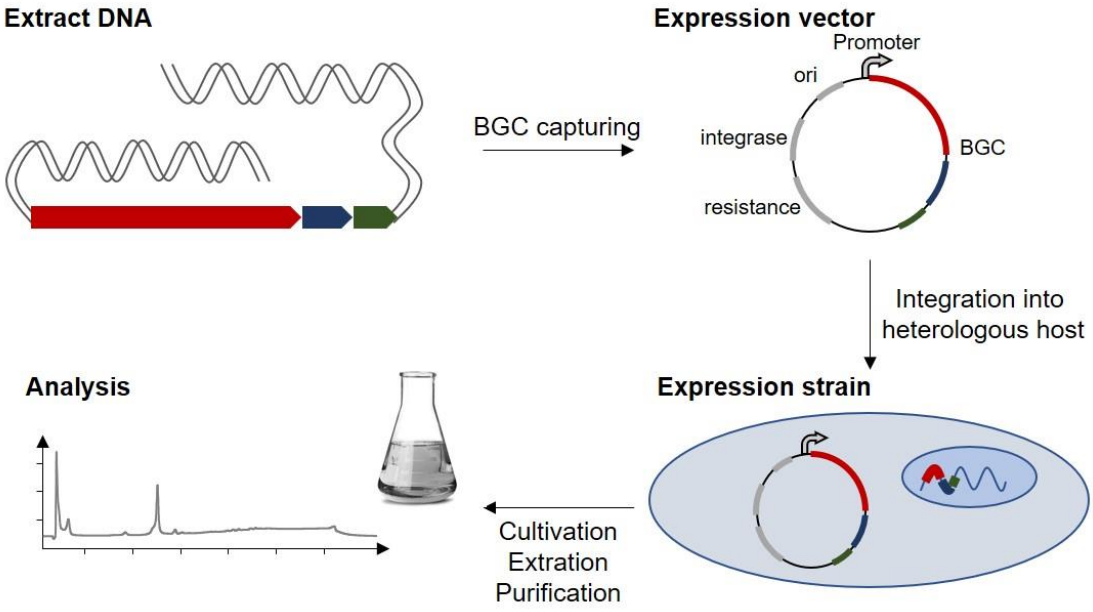


Figure 6. Molecule obtained by heterologous expression: methylarcyriarubin (**33**).

However, the expression of a BGC in a heterologous host implies several considerations for a successful production of a NP. The most important aspect is the capturing of the BGC from the genome of the natural producer. This has probably been one of the most intensively studied factors in heterologous expression with various methods that arose over the last few decades. Additionally, the choice of the right heterologous host, the activation of silent genes or the repression of inhibitory factors as well as the right expression conditions are further parameters to be determined for an efficient NP production in a heterologous host (Scheme 4).



Scheme 4. Schematic workflow of a heterologous NP production. The starting point of a heterologous expression is extracted DNA either from environmental samples or directly of microbial origin. The presence of a BGC is predicted through bioinformatic analysis. Using a capturing method, the BGC can be integrated into a suitable expression vector for the chosen heterologous host. The expression vector containing the BGC can be transferred into the heterologous host, where it is either present on a plasmid or can be integrated into the host genome. After cultivation, extraction and purification, the new NP can be analyzed using LC-MS and NMR to elucidate the structure.

The first broadly used method to identify and capture BGCs, even before whole genome sequencing was a standard application, was the production of genomic libraries. The method is based on the incorporation of large gDNA fragments into fosmids, cosmids as well as bacterial artificial chromosomes (BACs) or P1 phage derived artificial chromosomes (PACs). These vectors are capable of incorporating fragments of up to 300 kb as in the case of PACs. The gDNA needs to be cautiously purified and ligated into the chosen vector type, the vectors can then be transferred into *E. coli* and screening of the obtained clones is used to identify the clone harboring the BGC on interest. A recent example for the successful gene cluster capturing and subsequent expression of NPs is the PAC library containing the 83.5 kb FK506 (**34**) BCG from *Streptomyces tsukubaensis* NRRL 18488, thereby enabling downstream optimization of the production titer of this immunosuppressant drug.[81] Despite a large number of successfully cloned BGCs and expressed NPs, this method is highly labor intensive due to the extensive DNA library screening efforts. Additionally, the incorporation of large BGCs in fosmids/cosmids/BACs/PACs is very challenging and furthermore the incorporation of a DNA fragment harboring the whole BGC is not ensured, implicating further screening if only a partial cluster can be identified and possibly necessitating assembly of several cluster parts.

In addition to cloning using restriction endonucleases and ligases, homologous recombination (HR) mediated by truncated RecE and full length RecT, so called ‘ET cloning’, was established as a useful tool to manipulate DNA independent of restriction sites.[82] In this case, the desired vector is flanked by a 20 to 60 bp homology region that is also present on the genes of interest. The linear gene of interest and the circular vector are transformed into *E. coli* containing this enzyme pair and recombination can be observed at a high frequency. The preconditions for vector and insert gave this method the term ‘linear plus circular homologous recombination’ (LCHR). Recently it was shown that full length RecE and RecT facilitate HR of a linear gene of interest and a linearized cloning vector (LLHR) very efficiently, especially for large genetic regions, such as BGCs.[83] With its remarkably high recombination stability, HR facilities a useful alternative for gene cluster capturing compared to genomic libraries, especially considering the highly reduced screening effort. The potency of HR for the capturing of BCGs and heterologous expression was nicely shown for the 48.4 kb fostriecin (**35**) BGC,[84] which was captured in one piece. Impressively, the full potential of HR was proven by capturing the 106 kb salinomycin (**36**) gene cluster, which was divided into three pieces and reassembled into one intact BGC via LLHR that could prove to express **36**.[85] These two examples demonstrate the efficiency of HR for capturing gene clusters up to 50 kb, however, reassembling of three large DNA pieces needs a huge amount of additional cloning expertise.

Besides HR in *E. coli*, which relies on the heterologous expression of RecET or different HR facilitating enzymes, yeast is able to conduct HR using endogenous enzymes. Based on this ability, transformation-associated recombination (TAR) cloning was established to capture

large BGCs in one step. For this matter, capturing vectors containing homology regions at the end of the cluster of interest are constructed and together with the appropriate gDNA that is digested with restriction enzymes whose sites are found near the cluster ends, they are transformed into yeast spheroplasts. Using HR this method allows capturing of BGCs with a size up to 250 kb in one recombination step.[86] Amongst many successful examples, the isolation of taromycin A (**37**) by heterologously expressing the 67 kb orphan *tar* BGC depicts the great possibility of this method to efficiently capture gene clusters.[87] Like the previous capturing method, this method contains burdens, such as the high purity of large gDNA fragments needed for transformation and the required skilled handling of yeast spheroplasts for efficient incorporation of foreign DNA.

As most of these methods show significant burdens to capture large gene clusters, the development for faster and more efficient strategies is still ongoing. Cas9-Assiciated Targeting of Chromosomal segments (CATCH) depicts a new methodology combining RNA-guided cleavage of designated genomic regions with assembly into the expression vector by Gibson assembly.[88],[89] CATCH seems to be a very promising method although it contains various DNA manipulating techniques that need a lot of laboratory know-how. Nevertheless, it shows the creativity that is used to develop new and sophisticated tools to make this high amount of bioinformatic data accessible for NP discovery and in the long term for pharmacological use.

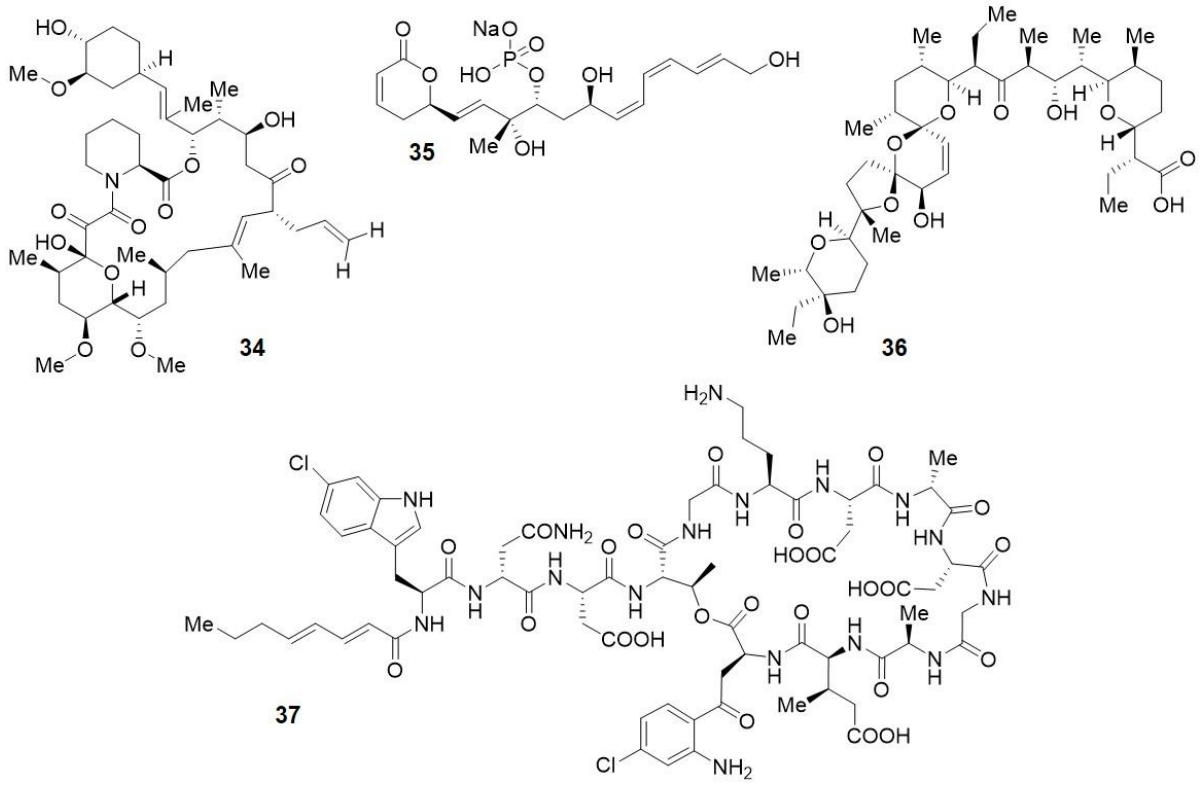


Figure 7. Examples of NPs obtained by heterologous expression. Gene cluster capturing was conducted utilizing genomic libraries: FK506 (**34**), homologous recombination (HR): fostriecin (**35**) and salinomycin (**36**) or transfomation-associated recombination (TAR) cloning: taromycin A (**37**).

1.4 Polycyclic tetramate macrolactams

The natural product class of polycyclic tetramate macrolactams (PoTeMs) consists of complex molecules harboring a similar carbon backbone (**38**) and exhibiting a broad variety of pharmacologically interesting properties. As shown in Figure 8, the compounds consist of a tetramic acid moiety (E) embedded into a macrolactam ring system (D) that is connected to a sequence of smaller carbocycles comprising either a 5-5, a 5-5-6, a 5-6-5 or even a 5-4-6 cyclization pattern (A-C). The great structural variety of the diverse PoTeMs arises from the different arrangement of the smaller carbocycles, additional functional groups attached to the backbones, or different reduction states of carbon atoms.

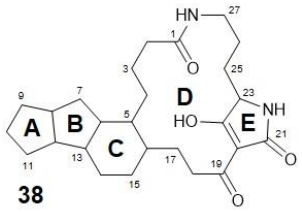


Figure 8. Schematic representation of the PoTeM backbone.

The first isolated compound of this class was ikarugamycin (**39**) in 1972, an antiprotozoal and antibiotic substance discovered in cell extracts of *Streptomyces* No. 8603, which was later renamed to *Streptomyces phaeochromogenes* var. *ikaruganensis*.[90] The additional property of **39** to inhibit the uptake of oxidized low-density lipoproteins[91] and its interference in clathrin-dependent endocytosis renders it a broadly applied tool to study cellular uptake mechanism nowadays.[92],[93],[94],[95],[96] In the 1990s, this natural product class began to grow with the incoherent discovery of several new compounds. Initially, discodermide (**40**) isolated from the marine sponge *Discodermia dissoluta* was discovered in 1991, possessing antifungal and cytotoxic properties.[97] Subsequently, the discovery of the cytotoxic compound alteramide A (**41**) found in a sponge symbiont *Altermonas* sp.[98] and another cytotoxic compound cylindramide (**42**) isolated from the sponge *Halichondria cylinakata*[99] but potentially also produced by a sponge symbiont was made. Later in the 1990s, the antifungal maltophilin (**43**) was discovered in a *Stenotvophomonas maltophilia* R3089 extract,[100] a possible precursor of the subsequently isolated dihydromaltophilin (**44**).[101] It was first discovered in a *Streptomyces* sp. and was later identified as the antifungal compound in the biocontrol agent *Lysobacter enzymogenes* C3 exhibiting a yet undescribed mode of antifungal activity by interfering with the synthesis of sphingolipids and thereby impairing polarized growth of filamentous fungi.[102] This discovery also gave the compound its common name ‘heat-stable antifungal factor’ (HSAF) and today it belongs to the best studied PoTeMs.[103] More structures

were identified, including aburatubilactam A,[104] geodin A,[105] xanthobaccin A,[106] or ripromycin.[107] However, none of these compounds were discovered using a targeted approach for PoTeMs, they were rather just randomly isolated natural compounds from different microorganisms.

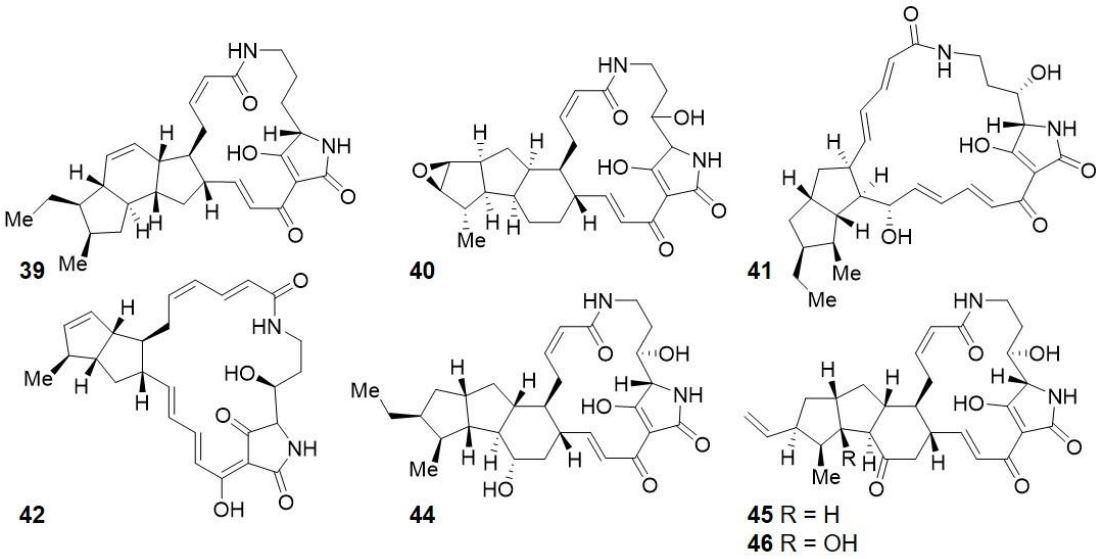
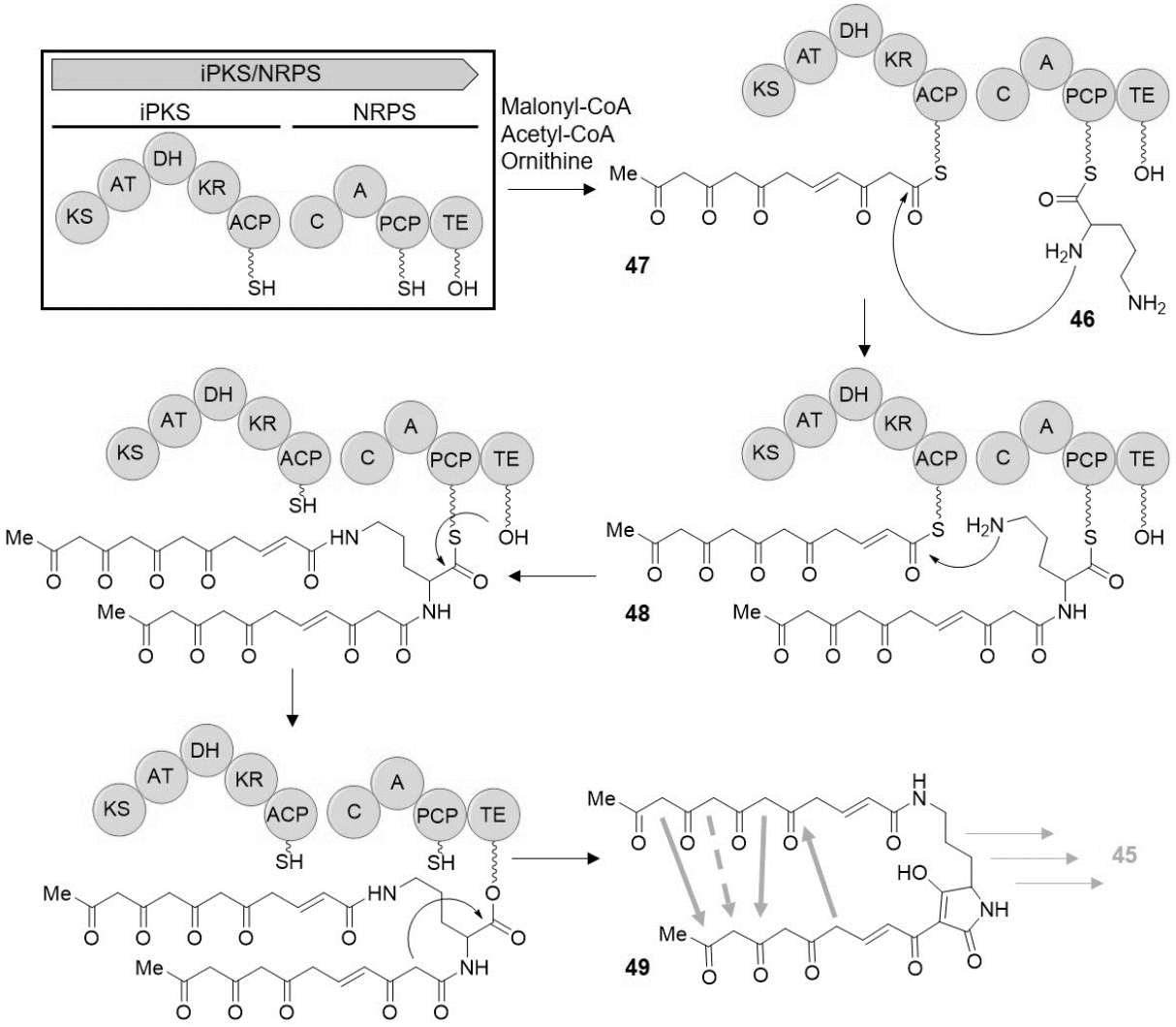


Figure 9. Selected PoTeM structures: ikarugamycin (**39**), discodermide (**40**), alteramide A (**41**), cylindramide (**42**), HSAF (**44**), and frontalamide A (**45**) and B (**46**).

The advances in whole genome sequencing offered the possibility to transition from coincidentally isolating natural compounds to specifically searching for distinct natural products in sequenced organisms. Two fundamental discoveries for PoTeM biosynthetic gene clusters by Clardy *et al.*[108] and Du *et al*. [109] generated the basis of modern PoTeM discovery. By isolating frontalamide A (**45**) and frontalamide B (**46**) and using genome mining and gene deletion approaches to identify the biosynthetic pathway of **45**, an uncommon iterative PKS/NRPS gene was found to be involved in PoTeM assembly.[108] Such iterative PKSs (iPKS) are typical for fungi but are rarely found in bacteria.[110] Additionally, it was shown that diverse microorganisms contain such an iPKS/NRPS gene in their genome, with genes arranged around it similar to the frontalamide gene cluster.[108] Having this information in hand, a first theoretical biosynthetic assumption for **45** was expressed (Scheme 5). Starting with the proposal that the FtdB (the iPKS/NRPS) produces two polyketide chains (**48** and **49**) that are linked to ornithine (**47**) to build the tetramic acid containing precursor **50**. However, the subsequent biosynthetic route suggested, based on the proposed precursor **50**, does not fit chemical requirements for the accomplishment of the required cyclization reactions. As depicted in Scheme 5, three cyclization reactions (bold arrows) would be possible as in each case a nucleophile would attack a electrophile in typical PKS-type cyclization biochemistry, whereas the last cyclization (dotted arrow) would underly a bond formation between two

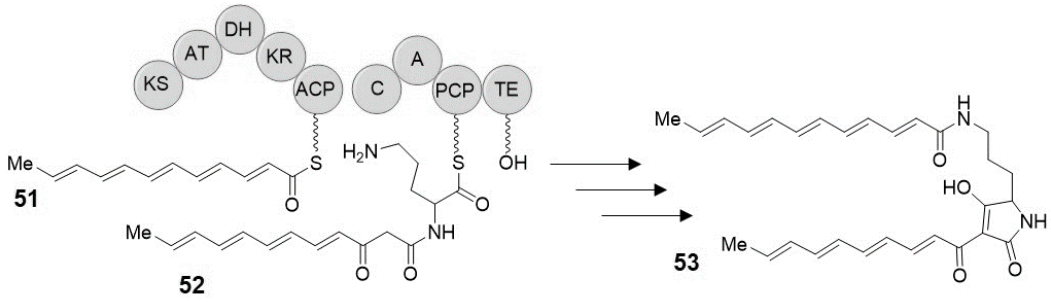
nucleophiles, rendering this proposed biosynthesis impossible. Moreover, the common precursor **50** with the reduction states as proposed by Clardy *et al.* cannot serve as the precursor for all different PoTeM. However, the exact mechanism of the downstream cyclization sequence remained unsolved.



Scheme 5. Proposed mechanism of the PoTeM precursor **50** biosynthesis. The iPKS/NRPS enzyme complex consists of a PKS containing a KS-, AT-, DH-, KR-, and ACP domain that are responsible for synthesizing the polyene chains (**48** and **49**), and a NPRS containing a C-, A-, PCP-, and TE-domain that catalyze the attachment of **48** and **49** to ornithine (**47**) and releasing the final precursor **50** by tetramic acid formation. The further proposal of the biosynthetic pathway to **45** with the proposed cyclization depicted with arrows is indicated in lighter grey.

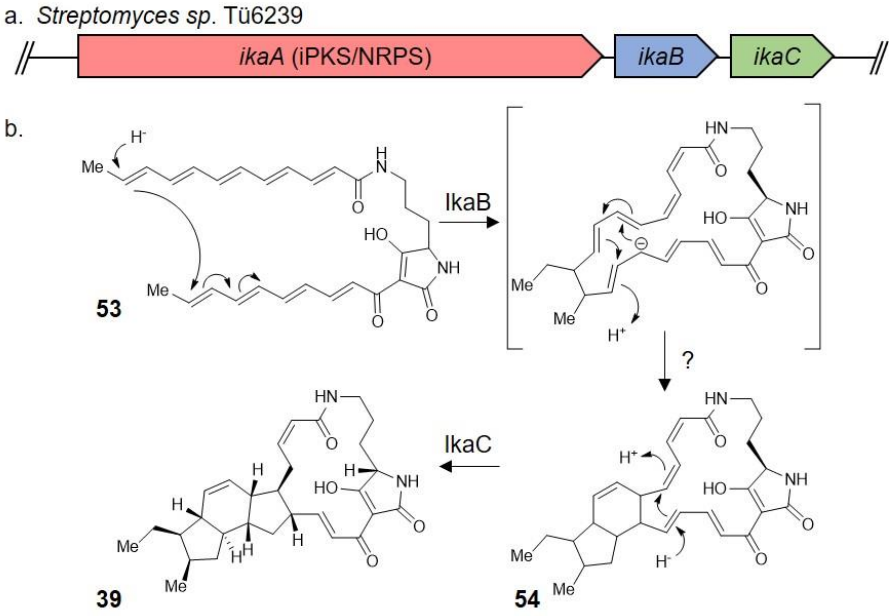
More proof for this pathway to the common precursor was needed. Studies on the NRPS module of the iPKS/NRPS complex found in the biosynthetic pathway of HSAF (**44**), verified that **47** is activated by the A-domain. Furthermore, the reaction of **47** bound to the A-domain of the NRPS module with acyl chains bound to ACP-domains result in a tetramic acid containing precursor bound to the two acyl chains. Thereby, the theoretical assumption of PoTeM iPKS/NRPS system to be capable of establishing a peptide bond between a polyene-chain and the terminal amino-function of **47** as well as coupling of a second polyene-chain to

the α-amino group of **47** and releasing the product by TE-catalyzed cyclisation to establish the tetramic acid moiety was proven.[109] These findings could later be supported by verifying the assembly of the actual polyene-chains (**51** and **52**) necessary for PoTeM biosynthesis with ornithine leading to the common precursor **53** (Scheme 6).[111] Understanding these basic principles of PoTeMs biosynthesis, it was possible to perform a targeted screening for PoTeM clusters and new molecules.



Scheme 6. Proposed biosynthesis of the common precursor **53**. The two polyene chains **51** and **52** are coupled to **47** and released by tetramic acid formation, for precise mechanism see Scheme 5.

One approach for cluster identification is a selective PCR-based screening for the specific TE-domain of PoTeM clusters, as performed in *Streptomyces* sp. Tü6239. This strain was proven to produce ikarugamycin (**39**) but had not been sequenced yet.[107] Degenerated primers were used to screen a fosmid-library to identify a clone containing the whole cluster encoding **39**. Heterologous expression of the cluster-containing fosmid in *E. coli* BAP1 confirmed the production of **39**. Thereby, it was demonstrated that only the three genes *ikaABC* are required for the biosynthesis of **39**. Additional deletion studies indicated the production of intermediate **54** when deleting IkaC, thereby enabling the first postulation of a biosynthetic route to **39**.[112] This proposed pathway is a further proof of **53** holding the right reduction state of the polyene chains. Starting with this polyene-ornithine-polyene intermediate a possible hydride-induced cyclisation is catalyzed by the first reductive enzyme to establish the outer bond. This step is followed by a possible non-enzymatic Diels-Alder reaction establishing the two inner connections with the formation of the inner bond catalyzed by the second reductive enzyme as a final step (Scheme 7). Additional studies by Zhang *et al*. confirmed these findings and were able to prove that IkaC is indeed responsible for the cyclisation of the inner ring by *in vitro* enzymatic assays.[113]



Scheme 7. Biosynthetic pathway of **39**. a. schematic representation of the ikarugamycin BGC and b. the proposed biosynthetic pathway underlying the production of **39**.

For the identification of novel PoTeM entities in a non-sequenced microorganism a similar strategy was used, resulting in the discover of clifednamide A (**55**) and B (**56**). Several environmental isolates were PCR-analyzed and screened for the occurrence of a conserved region in the beginning of the iPKS/NRPS-gene. Culturing *Streptomyces* sp. JV178, which was proven to contain this fragment, enabled the isolation of the new compounds.[114]

By means of whole genome sequencing heterologous expression of PoTeMs seemed to become the method of choice for identifying new compounds as most of the PoTeM cluster harboring strains were never identified to produce PoTeMs. Therefore, Zhao *et al*. cloned a putative PoTeM BGC from *Streptomyces griseus* and heterologously expressed this construct in *Streptomyces lividans*. To ensure gene transcription, promoters were cloned in front of each gene of the biosynthetic gene cluster and subsequently assembled into an expression construct. Thereby, the group was able to switch on the expression of a silent gene cluster and isolate the produced PoTeM, compound A, a molecule structurally closely related to **41**. The design of the expression system and experimental set-up enabled simple construction of expression plasmids missing single genes, facilitating the isolation of biosynthetic intermediates compounds B-D, with compound D (**57**) comprising the yet undescribed 5-4-6 carbocyclic pattern.[115] In a different approach, the whole genome of *Streptomyces albus* J1074 was bioinformatically analyzed for natural product BGCs. The analysis revealed the presence of several clusters with one containing the typical PoTeM iPKS/NRPS. Using homologous recombination, promoters were integrated into the genome, either in front of the whole cluster or in front of the iPKS/NRPS gene. These engineered strains

produced 6-epi-altermide A (**58**), which was only produced in very low quantities in the wildtype strain, and additionally 6-epi-altermide B (**59**), not detected at all in the wildtype stain. Both compounds represent new PoTeM entities.[116]

Further understanding of PoTeM biosynthesis was achieved by the analysis of single proteins found in the corresponding BGCs. The highly abundant sterol desaturase, arranged in front of the iPKS/NRPS, is a straightforward example. According to bioinformatic analysis and intermediate abundance, it was already assigned to hydroxylating the C-25 position in the PoTeM backbone. However, it has never been proven to accomplish this hydroxylation. Shen *et al*. could demonstrate that a *Lysobacter enzymogenes* mutant lacking the sterol desaturase would only produce 3-deOH-HSAF, giving experimental proof for the assumed function of the sterol desaturase.[117] Furthermore, this enzyme was demonstrated to work promiscuously as it also hydroxylated **39** when added to the heterologous *E. coli* expression strain *in vivo* or when **39** was enzymatically converted *in vitro*.[118] The resulting butremycin (**60**) was a recently discovered new PoTeM.[119]

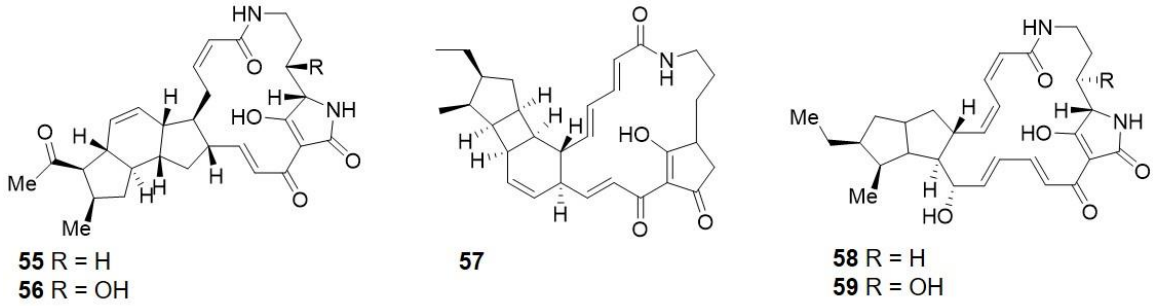


Figure 10. Selected PoTeMs discovered using targeted genomic approaches. Clifednamide A (**55**) and B (**56**), compound D (**57**) and 6-epi-altermide A (**58**) and B (**59**).

The targeted identification of new PoTeM molecules and clusters underlines the impact of the first biosynthetic pathway elucidations. They enabled significant progress concerning the understanding of PoTeM BGCs and the corresponding proteins. It also emphasized that similarly organized PoTeM BGCs often result in quite different natural compounds, either varying in stereochemistry, in cyclization pattern, or in the addition of functional groups. This leaves a great fundus of possible new compounds and a large pool of proteins with yet undescribed function, which might also work promiscuously and could possibly produce bioengineered new compounds.

1.5 Aims of this thesis

The progression of methods to isolate NPs with biological activities is crucial for developing new drugs in the fight of the current pharmacological crisis in particular in the field of antibiotics. As PoTeMs exhibit such a large range of pharmacologically interesting effects and their iPKS/NRPS gene is unique and easy to detect by bioinformatic analysis, this class of secondary metabolites constitutes a good target for NP research. Accordingly, this thesis aims to develop new methods to heterologously express PoTeM BGCs in *Streptomyces* to identify new molecules as well as designing approaches to modify known PoTeMs to detect promiscuity and mechanisms of PoTeM biosynthesis.

Heterologous expression of ikarugamycin based molecules

In a first step, the published BGC of **39** is used as a model system to establish a robust *Streptomyces* expression system (Figure 11). Therefore, the cluster captured in a fosmid is transferred into four different expression vectors (Table 1), all suitable for heterologous expression in *Streptomyces*. This step will already help to optimize a new PCR and Gibson assembly[89] based gene cluster capturing method. The obtained plasmids will then be integrated into three different *Streptomyces* hosts, *S. albus* DSM40313,[120] *S. lividans* TK24,[121] and *S. coelicolor* M1154,[122] all exhibiting different advantages considering secondary metabolite expression, to find an optimal host for PoTeM expression. In a last step, culturing conditions and purification methods will be established to obtain NMR data of the pure compound.

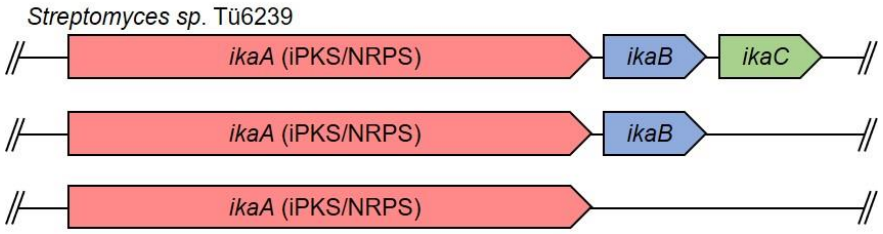


Figure 11. BGCs used for heterologous expression of **39** and its intermediates **53** and **54**.

Furthermore, the expression of truncated ikarugamycin clusters will be used to elucidate the full biosynthetic route of **39**. Therefore, the optimal expression conditions identified previously for **39** will be used and optimized to express vectors containing only *ikaA* or *ikaAB*. Again, culturing and purification protocols will be established to elucidate the exact structure of the intermediates and gain further insights into the biosynthesis of PoTeMs. As *S. albus* DSM40313 encodes a PoTeM BGC in its genome, a knockout strain will be set up

and used for heterologous expression to circumvent modification of the intermediates by host strain specific PoTeM enzymes.

Table 1. Summary of utilized expression vectors.

Vector	Promoter	Transfer into host
pSET152_*ermE*	Constitutive (*ermE*)	Conjugation
pUWL201PW	Constitutive (*ermE*)	Protoplast transformation
pWHM4*	Constitutive (*ermE*-like)	Protoplast transformation
pWHM1120	Inducible (*tipA*)	Protoplast transformation

Heterologous expression of novel PoTeM clusters

Using the protocol set up and validated in the first part, heterologous expression of five different PoTeM clusters will be optimized to identify the novel produced molecules (Figure 12). For this part, the capturing of the BGCs will be based on a new PCR- and Gibson assembly-based method. However, the multiple cloning site (MCS) of the expression vectors will be adjusted to enable the production of only one PCR product that can be integrated into all the different expression vectors. In a second step, the location of the promoter will be changed considering knowledge gained in heterologous expression of PoTeMs. Finally, the distance between ribosomal binding site (RBS) and start codon will be considered, withdrawing the advantages by having only one MCS. All different expression constructs will be integrated into the three different *Streptomyces* hosts and culturing conditions as well as purification methods will be established to obtain pure compounds. Utilizing NMR and mass spectrometry, the structure of the novel products will be elucidated.

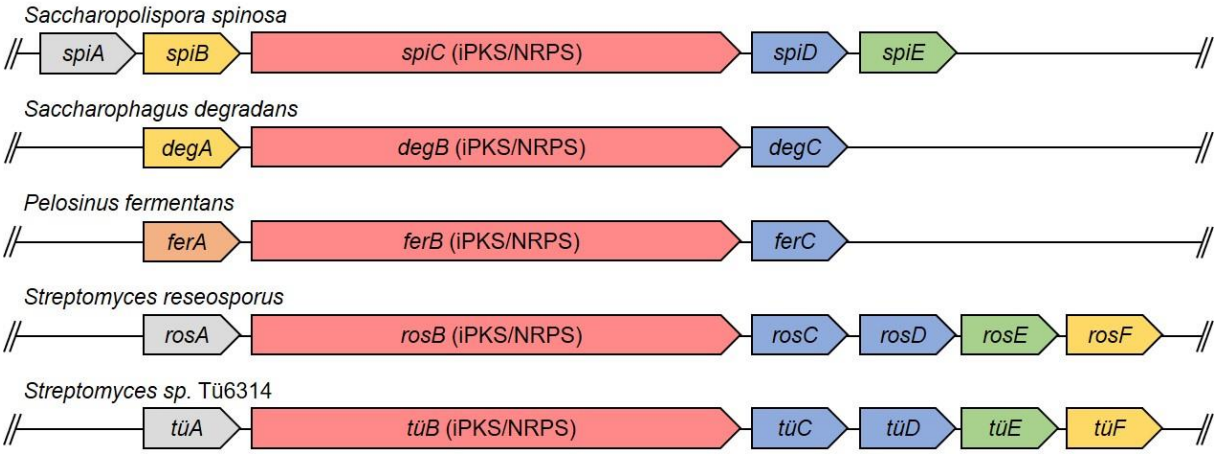


Figure 12 Novel PoTeM clusters utilized in this study. The PoTeM BGCs from the five different organisms *S. spinosa*, *S. degradans*, *P. fermentans*, *S. roseosporus*, and *Streptomyces* sp. Tü6314 are schematically depicted, not scaled to kbp. Red genes: iPKS/NRPS; blue: FAD-dependent oxidoreductases; green: alcohol dehydrogenase; yellow: cytochrome P450 enzymes; grey: sterol desaturase/ PoTeM hydroxylase, orange: acyltransferase.

Establishing a plug-and-play system for PoTeM expression

Regarding the difficulties expressing PoTeMs, a new heterologous plug-and-play system will be generated. In this case, the characteristic of one common precursor **53** of PoTeM biosynthesis is utilized to establish an expression system based on the iPKS/NRPS *ikaA.* Formerly gained results suggest that the expression of **53** might be the crucial step in heterologously expressing PoTeMs. Therefore, a basic expression plasmid only containing *ikaA* already proven to produce **53** will be utilized and modified with unique restriction sites and promoters (Figure 13a). To prove the potential of this system, the missing genes of the **39** BGC (*ikaBC*) will be added at the StuI restriction site by Gibson assembly to produce **39** (Figure 13b). Furthermore, different modifying enzymes from different PoTeM BGCs will be integrated at the XbaI restriction site using conventional restriction cloning (Figure 13c). All different expression constructs will be integrated into the three *Streptomyces* hosts and fermented using the protocol established in the beginning of the thesis. The modified products will be isolated and analyzed by NMR and mass spectrometry to identify the structure of the modified products. This study will help understand the mode of action of different PoTeM modifying enzymes and additionally it will serve as a platform to express further yet unknown PoTeMs.

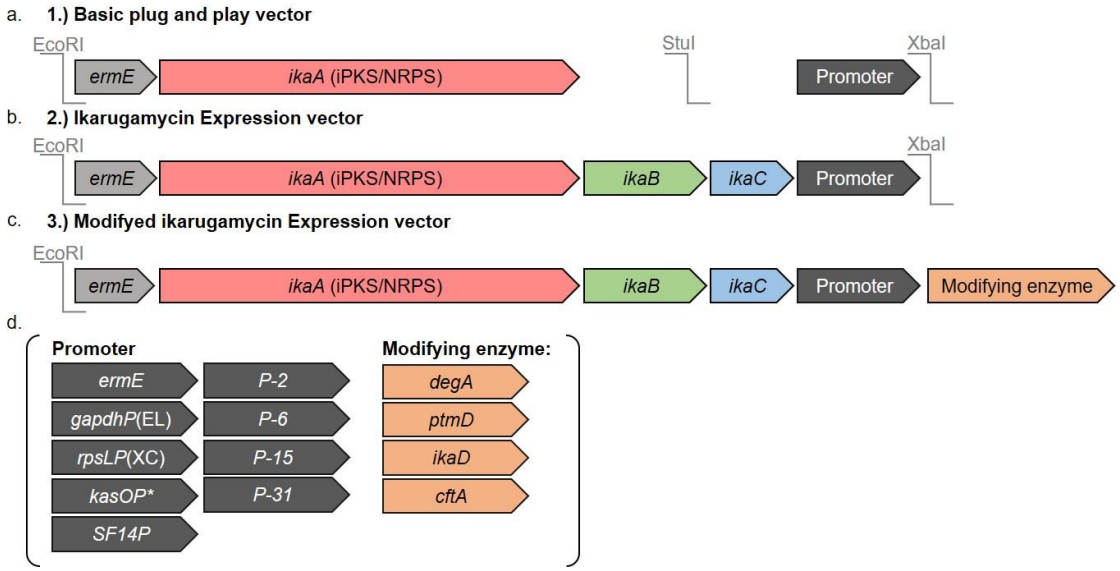


Figure 13. Plug-and-play system for PoTeM expression. a. Basic expression system containing *ikaA* and the different promoters and the unique restriction sites. b. Plug-and-play expression system for the expression of **39**. c. Plug-and-play expression system for modified **39**. d. Different promoters and modifying enzymes used in the study.

Considering the results of this part, a condensed version of starting vectors (Figure 13a) and expression hosts will be used to express novel PoTeMs from clusters already used in the second part of this thesis (Figure 12). In this case, the modifying gene of the PoTeM BGC found in *S. spinosa* arranged around *ikaA* as shown in Figure 14, either depicting the order found in the natural cluster (Figure 14a 1) or rearranged to utilize the different promoters present in the plug-and-play system (Figure 14a 2). Expression constructs containing the same arrangement of genes will be cloned for the PoTeM cluster present in the *S. degradans* genome as well as the *P. fermentans* genome. Production of novel PoTeMs will be analyzed for these clusters.

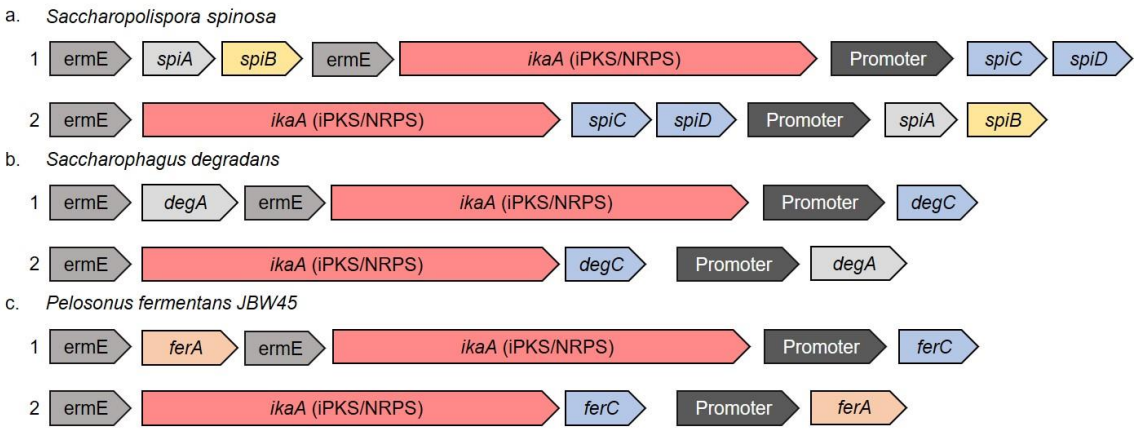


Figure 14. Schematic map of plug-and-play expression constructs with new PoTeM BCG. Depicted are the rearranged PoTeM clusters found in the organisms a. *S. spinosa*, b. *S. degradans* and c. *P. fermentans* either used in their natural order (1) or modified (2).

Crystallization of IkaC

The last step in the biosynthesis of **39** is the inner ring cyclization catalyzed by IkaC. To further understand this process, the crystallization of IkaC will be conducted in this study. To identify the molecular mechanisms of the cyclization reaction, a co-crystallization of IkaC containing intermediate **54**, its substrate, and a reduced NADH will be attempted. Furthermore, mutants will be established to elucidate the mechanism of IkaC.

2 Results and Discussion

2.1 Heterologous expression of ikarugamycin

The *ika* BCG encoding for **39** was identified by Antosch *et al*.[112] and a first heterologous expression in *E. coli* BAP1 was established. However, the yields of this heterologous expression system were only enough to detected **39** by mass spectrometry but not for NMR analysis nor further downstream applications. To improve the yields and establish a robust heterologous expression platform, the *ika* cluster from *Streptomyces* sp. Tü6239 was selected for heterologous expression in *Streptomyces*.

2.1.1 Preparing expression vectors

The four expression plasmids pSET152_ermE (Figure A1), pUWL201PW (Figure A2), pWHM4* (Figure A3), and pWHM1120 (Figure A4) were chosen as expression vectors for this study. pSET152 did not contain a promoter and therefore the commonly used *ermE* promoter from the erythromycin BGC[123] was chosen to be integrated into the vector to generate pSET152_ermE. To accomplish the integration of the promoter into the expression vector, a sequence was developed containing the promoter (*ermEP2* and *ermEP1**), a ribosomal binding site (RBS), a spacer, and restriction sites for the integration into the cloning vector pBluescript (Figure 15a, Figure A5). The overlap extension PCR was conducted (chapter 4.2.13), the PCR product and pBluescript were digested with EcoRI (chapter 4.2.5), ligated (chapter 4.2.7) and successful integration of the promoter into the vector was confirmed by Sanger sequencing (Figure 15b).

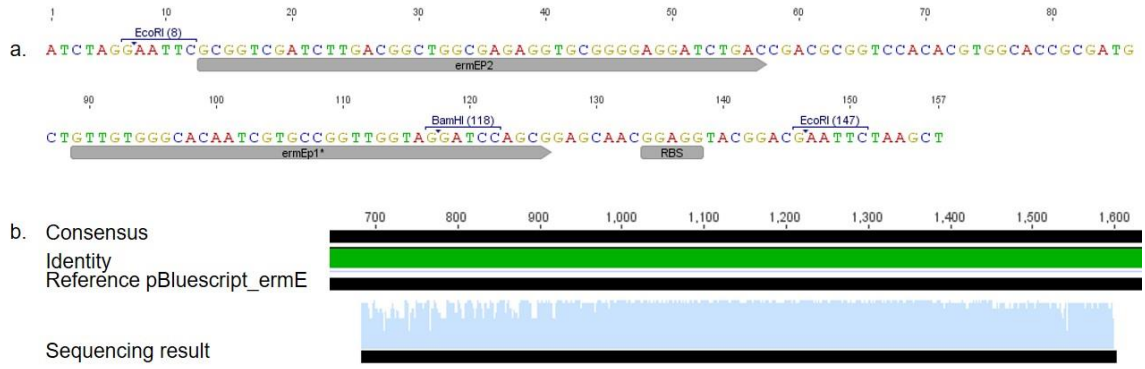


Figure 15. Capturing the *ermE* promoter into pBluescript. a. designed sequence of the *ermE* promoter used for this study, with promoter sequence (*ermEP1** and *ermEP2*), RBS, restriction sites for EcoRI and overhangs to enable digestion by the enzyme EcoRI. b. Sequencing result of *ermE* integration into pBluescript, proving a 100% identity (green) to the reference sequence.

To integrate the captured *ermE* promoter into pSET152, the promoter was PCR amplified (chapter 4.2.13; Figure 16a) using primers containing homology regions for Gibson assembly, which also contained an additional StuI restriction site, thereby integrating this unique restriction site behind the promoter for integration of prospective clusters. The vector was digested using EcoRI and XbaI (chapter 4.2.5), dephosphorylated (chapter 4.2.6) and Gibson assembly was conducted with the purified PCR product and vector (chapter 4.2.7), the mixture was transformed into chemically competent *E. coli* DH5α (chapter 4.2.10). Positive cones were screened by colony PCR (chapter 4.2.13) and verified using Sanger sequencing (chapter 4.2.14). Clone 1 was determined to have successfully integrated the promoter and utilized for downstream experiments.

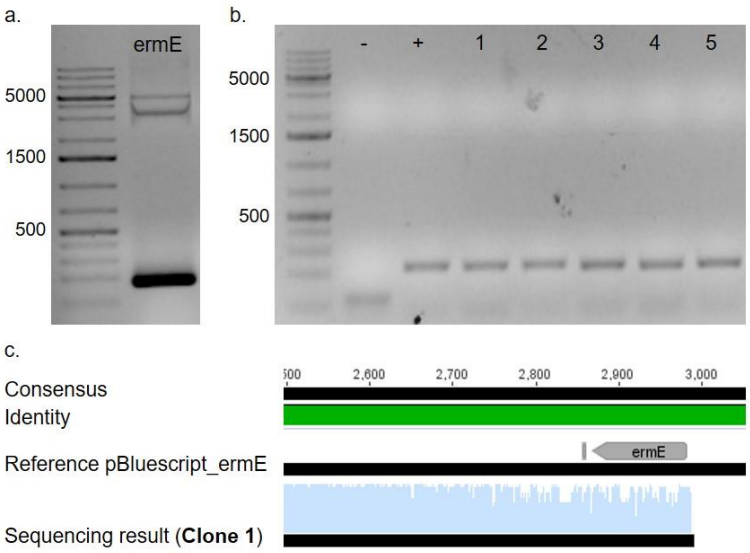


Figure 16. Assembly of pSET152_ermE expression vector. a. Amplification of the *ermE* promoter with a size of 183 bp, b. Colony PCR of clones obtained after transformation screened with M13_fwd and Lac-Promoter pSET152 primers, with expected bands for positive integration at 266 bp and without integration at 151 bp. c. Clone 1 was sent for sequencing and showed 100% identity with the reference sequence (green).

2.1.2 Preparation of ikarugamycin expression constructs

In a first approach the capturing of the *ika* BGC was aimed to be conducted by HR using a novel designed capturing fosmid, which was kindly provided by the working group of Dr. L. Kaysser. The new vector was constructed to facilitate HR with pCC1FOS to directly recombine BGCs found within genomic libraries with an expression vector suitable for *Streptomyces*.[124] Figure 17 shows the XbaI-linearized pLK01 vector, which contains all important features for *Streptomyces* conjugation, integration into the *Streptomyces* genomes, an antibiotic selection marker, and furthermore homology sequences with pCC1FOS to enable HR with the expression vector of interest.

Figure 17. Linearized capturing and expression vector pLK01. The vector was constructed containing features for conjugation into *Streptomyces* (OriT and TraJ), integration into a *Streptomyces* genome (attP and integrase), a resistance marker (NeoR) and additionally it contains homology overhangs to the library fosmid pCC1FOS depicted in purple and magenta.

Using this approach to integrate the BGC of **39** into an expression vector in a first step, the fosmid containing the cluster was transformed into *E. coli* BW25113 pKD46 and successful transformation was verified by colony PCR (chapter 4.2.13). In parallel, the capturing vector was digested with XbaI (chapter 4.2.5) and purified (chapter 4.2.4). The linearized pLK01 vector was transformed into *E. coli* BW25113 containing the fosmid pKD46 and HR was conducted (chapter 4.2.8). Screening of the obtained colonies (chapter 4.2.13) showed successful HR, therefore the purified plasmid was transformed into *E. coli* ET12567. However, it was never possible to find a positive colony containing pLK01::*ika*-fosmid11. Accordingly, it was assumed that the purified plasmid was a mixture of recombined and not recombined vector, of which the smaller, in this case not recombined vector showed a higher transformation efficiency, which led to the inability to find *E. coli* ET12567 colonies harboring the recombined expression vector. A trial to transform DNA picked from an agarose gel, showing the length of the recombined expression vector, into *E. coli* ET12567 also failed. A possibility to separate plasmid mixtures is to transform the mixture into *E. coli* cells, which are usually only able to take up one plasmid. Accordingly, a second transformation and screening could have identified colonies harboring only the recombined plasmid. While eventually successful, we decided to move on with the novel BGC capturing method to heterologously express the *ika* cluster.

The newly developed method combining a long amplicon PCR and Gibson assembly was established for capturing the cluster responsible for **39** biosynthesis. Therefore, four inserts were amplified by PCR with primers containing homology arms for the respective expression vector and the *ika* cluster to enable subsequent Gibson assembly. As a DNA template for the PCR the published fosmid containing the *ika* cluster was used.[112] As observed in Figure 18, the entire *ika* cluster was successfully amplified as a single product by PCR. These results show that it is possible to amplify a 12.3 kbp cluster with 72% GC-content as a single PCR product. However, it is visible that all of the PCRs show unspecific amplification of lower size DNA fragments. Therefore, the PCR product was purified using gel extraction (chapter 4.2.4) for subsequent subcloning. The expression vectors were digested (chapter 4.2.5) and dephosphorylated (chapter 4.2.6) with the respective restriction endonucleases (Table 2) and subsequently purified using the PCR purification kit (chapter4.2.4).

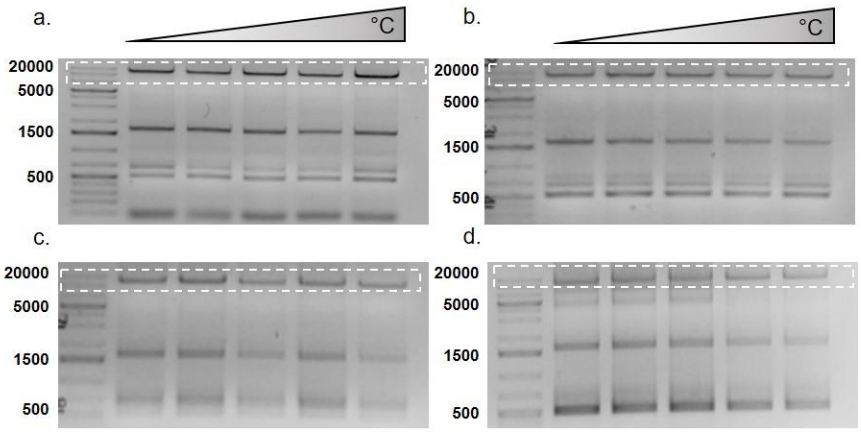


Figure 18. Amplification of the *ika* BGC. Temperature gradient PCR to produce the whole cluster with homology arms for a. pSET152_ermE, b. pUWL201PW, c. pWHM4*, and d. pWHM1120.

Table 2. Expression vectors with restriction sites for cloning.

Expression vector	5′ restriction site	3′ restriction site
pSET152_ermE	StuI	StuI
pUWL201PW	EcoRI	BamHI
pWHM4*	EcoRI	EcoRI
pWHM1120	XbaI	BamHI

The purified DNA fragments were assembled in a Gibson assembly reaction (chapter 4.2.7) and transformed into chemically competent *E. coli* DH5α (chapter 4.2.10). Grown colonies were analyzed for successful integration of the cluster by colony PCR (chapter 4.2.13) and plasmid DNA was prepared from possibly positive clones to verify cluster integration by analytical restriction digestion (chapter 4.2.5) and Sanger sequencing (chapter 4.2.14). The chosen clones can be seen in Table 3, the figures of colony PCR, restriction digest, and sequencing results can be found in the respective reference to the appendix.

Less than 40 clones had to be screened to find a clone with a successful integration of the cluster into one of the four different expression vectors, underlining the great opportunity gained by this new PCR and Gibson assembly-based method. The screening effort is highly reduced when compared to the classical methods for capturing a gene cluster, such as large-insert genomic DNA library construction. Additionally, the main laboratory work is similar to normal restriction cloning, providing results in less than a week, which represents a significant time reduction to other capturing techniques, such as methods based on homologous recombination. Proving that this method exhibits great potential for the capturing of small to mid-size BGC with GC-contents above 70%, whereas the capturing by HR was no possible.

Table 3. Result summary of ikarugamycin expression constructs. The Table shows the chosen clones for the four different expression vectors with a reference to the respective figure in the appendix for result description.

Expression vector	Chosen Clone	Reference to results
pSET152_ermE::*ika*	Clone 2	Figure A14
pUWL201PW::*ika*	Clone 4	Figure A15
pWHM4*::*ika*	Clone 22	Figure A16
pWHM1120::*ika*	Clone 15	Figure A17

2.1.3 Optimizing heterologous conditions for **39** production

The cloned expression constructs were integrated into the two expression hosts *S. albus* DSM40313 and *S. lividans* TK24 either using conjugation (chapter 4.2.12) for pSET152_ermE::*ika* or by protoplast transformation (chapter 4.2.11) for pUWL201PW::*ika*, pWHM4*::*ika* and pWHM1120::*ika*. Additionally, pSET152_ermE::*ika* was transferred into *S. coelicolor* M1154 using conjugation. A transfer of the other constructs into *S. coelicolor* M1154 was not performed as it was not possible to prepare competent protoplast for this strain. Single *Streptomyces* colonies were picked and cultivated in GYM or CASO medium for two to three days. A main expression culture of 50 ml Zhang Medium[113] was inoculated 1:10 with the preculture and incubated at 28 °C for five to seven days. Cells and supernatant were extracted (chapter 4.2.19) and analyzed using a LCQ fleet LC-MS system (chapter 4.2.20) with the method depicted in Table 35. The analysis showed an expression of **39** for all the different expression vectors, as detected by MS. The first expression conditions to detect **39** in the UV absorption of the LCQ Fleet was pSET152_ermE::*ika* in *S. albus* DMS40313 incubated 7d in Zhang Medium. At this point the analysis was changed to the Jasco-Advion LC-MS system (chapter 4.2.20) with the gradient shown in Table 36, as the MS of this system is less sensitive and a UV detector has increased sensitivity. The employed Jasco-Advion LC-MS system will now be referred to as LC-MS. A 4 L culture was inoculated to increase yields and analyzed by LC-MS (Figure 19), which suggested that *S. albus* seems to be a suitable host for heterologous expression of PoTeMs and pSET152_ermE expression vector results in decent yields of the product.

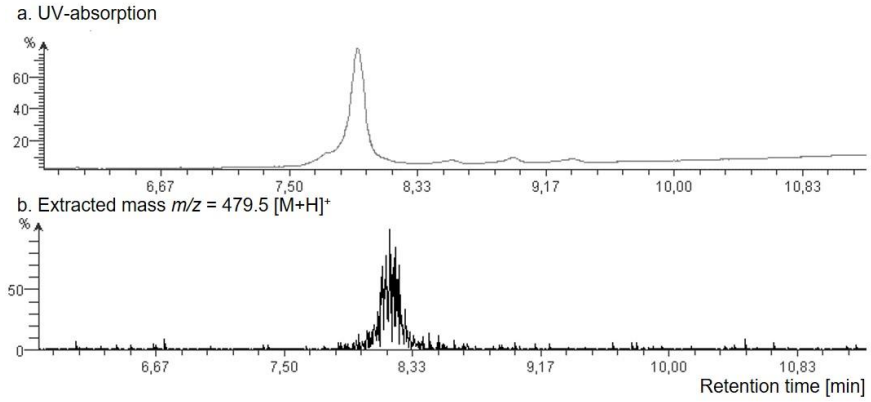


Figure 19. LC-MS analysis of extracts from heterologously expressed **39**. Results of extract obtained from 4 L supernatant of *S. albus* DSM40313 harboring pSET152_ermE::*ika* incubated in Zhang Medium for 7 d. a. UV absorption of the extract and b. extracted mass results for *m/z* = 479.5 $[M+H]^+$, the mass of **39**.

Unfortunately, the NP amounts were not yet satisfactory for the establishment of a robust PoTeM expression system. Therefore, further optimization was conducted. To optimize the expression, two different inoculation ratios (cells:medium) were used, 1:50 and 1:10, showing an overall better expression titer by inoculating with a ratio of 1:10 (Figure 20b). Additionally, three different cultivation media were used, the previous Zhang medium, YMG medium and ISP-4 medium, with all three used with (+) or without (-) a wire spiral to ensure homogenous culturing and enabling better cell growth. To ensure reproducibility of the results, each heterologous expression condition was performed in triplicates. Based on the results depicted in Figure 20, optimal production was obtained with *S. albus* containing the pSET152_ermE::*ika* expression construct inoculated with a ration of 1:10 at 28 °C for 7 d in ISP-4 medium without a wire spiral (Figure 20b). Furthermore a correlation between dry cell weight and production of **39** was determined (Figure 20c), further promoting these culturing conditions, as significantly more **39** was produced compared to total cell mass of the extracts. With these optimized culturing conditions in hand, the additional application of amberlites was attempted.[125] Amberlite beads can capture small molecules with different chemical properties from the supernatant to stabilize the molecules and remove free NP from the culture, thereby minimizing effects of potential negative feedback loops that may reduce NP production by the recombinant host. As **39** is a small, hydrophobic molecule XAD7HP was selected. The resin was added to a 50 ml culture either 0.5 g or 1 g directly or 3 d after inoculation, resulting in an enhanced production of **39** when adding 0.5 g XAD7HP directly after inoculation (see Figure 21). Accordingly, the host strain, the expression construct and the culturing conditions for heterologously expressing **39** were determined. As increased culture volumes seemed to inhibit **39** production, upscaling was achieved by fermenting several 50 ml cultures in 250 ml Erlenmeyer flasks.

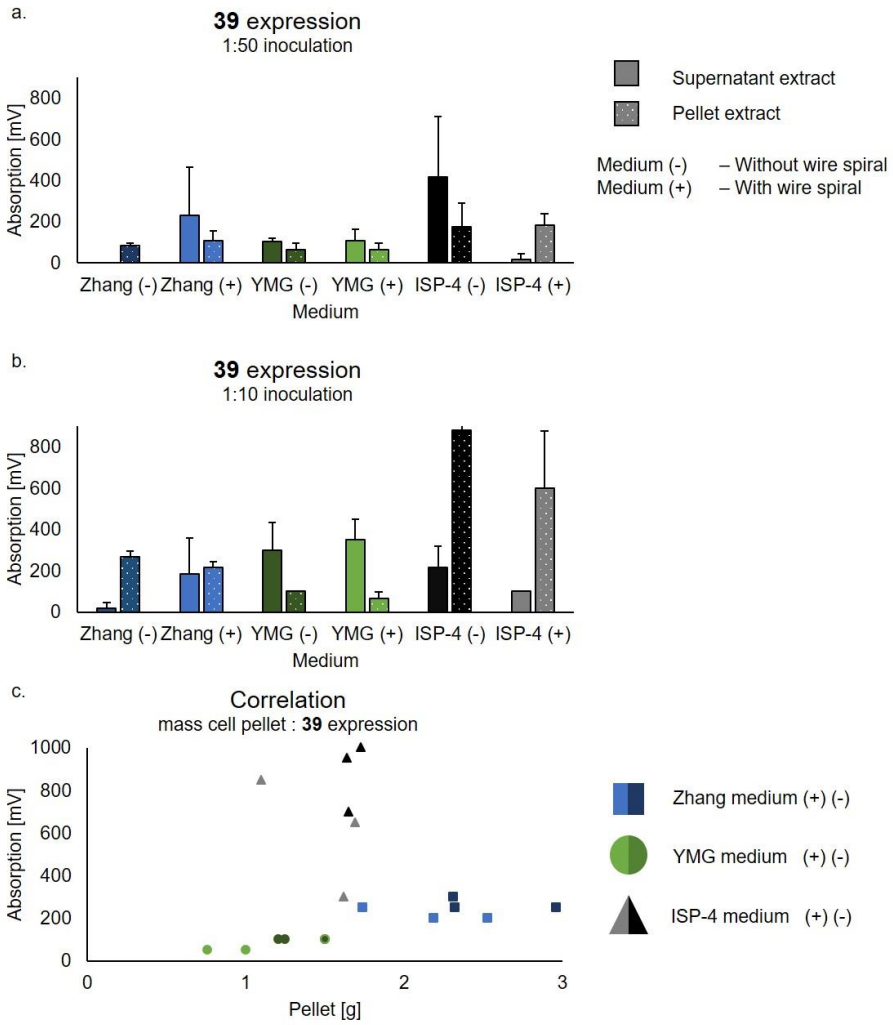


Figure 20. Optimization of **39** production. The best producer of **39**, *S. albus* DSM40313 harboring pSET152_ermE::*ika* was fermented in two different inoculation densities (1:50 (a.) and 1:10 (b.)), three different media (Zhang in blue, YMG in green, and ISP-4 in black) either with (+) or without (-) wire spiral for 7 d at 28 °C and 200 rpm. Each condition was conducted in triplicates, bars show mean and standard deviation as error bars. c. correlation between cell weight of the cell pellet and **39** detected in the extract of 1:10 inoculated cultures, lighter colors depict cultures with wire spiral, darker colors cultures without wire spiral.

After optimizing heterologous expression conditions, **39** needed to be purified from the crude extract, therefore, a preparative HPLC method was established using a C8 column (see chapter 4.2.20) and the gradient shown in Table 38. Having the pure compound in hand, a final isolated yield of **39** of 25 mg/L was determined (Figure A105). However, the poor solubility of **39** as a pure compound as well as in a crude extract enabled only an approximate determination of NP production. As a result, our optimized expression system was better than the recently reported isolation of **39** from the natural producer *Streptomyces* sp. ZJ306. In this study Zhang *et al.* were able to purify 103 mg **39** out of a 20 L fermentation broth.[113] This results in an expression titer of a little more than 5 mg/L, a titer that is 5 times less than the amount gained by the heterologous expression platform established in this thesis. Unfortunately, the optimized heterologous production of **39** was still lower than the reported

80 mg/L from the native producer *S. phaeochromogenes* var. *ikaruganensis* No. 8603.[90] However, as visible in Figure 21, the heterologous expression culture optimized for ikarugamycin production produced almost pure **39**, a fact that is highly advantageous for the subsequent compound purification.

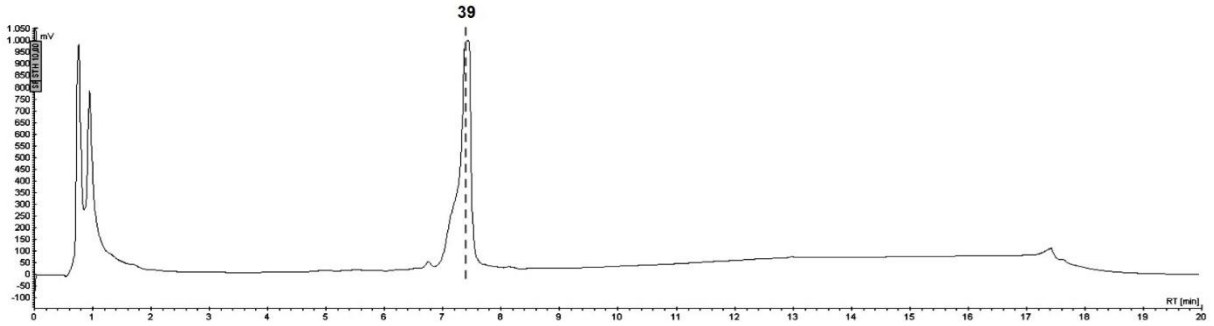


Figure 21. HPLC analysis of optimized **39** production. The chromatogram shows production of **39** analyzed from a 50 ml culture containing the optimized heterologous expression conditions, when additional 0.5 g XAD7HP were added at the beginning of the inoculation. Analytical method: Eurosphere_IkaFAST_CG.Meth.

The results underline that the heterologous expression of **39** in *S. albus* DSM40313 could greatly improve the yield when compared to the heterologous expression in *E. coli* BAP1. The expression titer did not yet reach the 80 mg/L published in the 1970s.[90] However, such a high titer of NPs is quite uncommon. The high amount of NP in this study might have also been the reason why **39** was the first characterized PoTeM, as a lot of material was still needed for structural elucidations at that time. More common are titers found in the study by Zhang *et al*.[113], around 5 mg/L. Comparing the 25 mg/L achieved by the heterologous expression system of this study with the natural producer shows heterologous expression offers an opportunity for increased yields of NP of interest. And furthermore, the purity of the compound prior to purification is a great benefit present in this heterologous expression system. More importantly, the yields were enough for an extensive NMR analysis. As a different study in our group addressed the total biosynthesis of **39**, and NMR spectra of such complex molecules are elaborate, having a comparable standard compound greatly simplifies the analysis. Accordingly, the purified compound from this study could serve as a standard material for NMR analysis in the study of the one pot biosynthesis of **39** (Figure S26/27 of the reference paper).[126]

It was possible to prove the impact of the newly developed long amplicon PCR and Gibson assembly-based method to be very efficient in cloning small to mid-size gene clusters even with high GC-contents. As depicted above, the screening of less than 40 clones for each construct greatly decreases the screening effort compared to formerly used capturing methods, such as genomic libraries, where screening of about 1000 clones for a cluster of interest is common. Additionally, it enables the direct cloning into an expression vector of

choice - and in this case even in four different expression vectors - to optimize the subsequent heterologous NP production. An advantage that is not possible with other methods, where either a cosmid/fosmid/BAC or one specific capturing vector is needed. This method facilitates gene cluster capturing as easy as conventional restriction cloning.

2.2 Modifying expression vectors for efficient cluster capturing

The newly established gene cluster capturing method requires a linearized vector and a PCR product with homology overhangs to the expression vector to allow Gibson assembly. As the first trial to heterologously express **39** in *Streptomyces* was successful for all four expression vectors (see Table 1), but a different PCR product needed to be amplified for each vector, it was decided to change the MCS of each vector to be identical in the final templates. Accordingly, only one PCR product needed to be amplified and could be integrated into all four expression vectors.

The first step for this goal was to identify a restriction endonuclease that does not cut in any of the four vectors, to facilitate linearization of the vectors. In this case, NdeI was chosen for this strategy, however, pWHM4* had an internal NdeI site and was not used to continue with this strategy. In a second step, a 20 bp overhang at each side of the NdeI site was designed, suiting the optimal condition for Gibson assembly and Q5-PCR amplification with a melting temperature of 60 °C (Figure 22). Accordingly, primers were designed, annealing to the MCSs of the three different expression vectors and additionally containing the new MCS with NdeI sites at each primer. To enable restriction digest, six random bases were also added to the 5′-end of each primer. Using the vector as a template, a Q5-PCR was conducted to amplify the new expression vector (chapter 4.2.13), the PCR products were purified (chapter 4.2.4) and subsequently digested with NdeI (chapter 4.2.5) to render ligation of the linear fragment possible (Figure 23). Ligation was performed at 16 °C overnight (chapter 4.2.7) and the inactivated reaction was transformed into *E. coli* DH5α by heat shock (chapter 4.2.10).

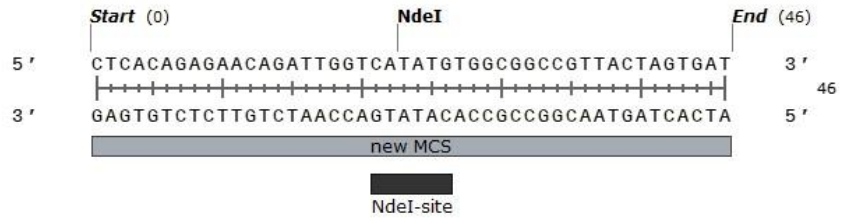


Figure 22. Sequence of new MCS for the expression vectors. The sequence contains a NdeI site that is unique for all three expression constructs, and two overhangs that fulfill optimal requirements for Gibson assembly and Q5 PCR amplification.

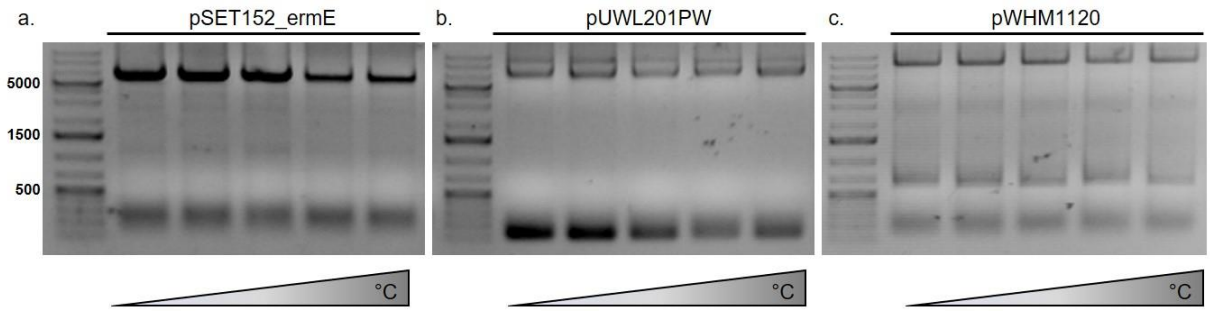


Figure 23. Temperature gradient PCR of new MCS vector amplification. The agarose gel pictures of the PCR results for the amplification of the three expression vectors pSET152_ermE with a size of 5876 (a.), pUWL201PW with a size of 6903 (b.), and pWHM1120 with a size of 9810 (c.) are shown with increasing annealing temperatures used in the PCR cycling (50-70 °C).

Possible positive clones were detected using colony PCR (chapter 4.2.13) and the plasmids of these colonies were isolated (chapter 4.2.4). Plasmids that showed the correct restriction pattern after analytical restriction digest (chapter 4.2.5) were submitted for Sanger sequencing (chapter 4.2.14). Within the first five clones, there was a correct clone for each construct (Table 4).

Table 4. Result summary of expression constructs with a new MCS. The table shows the chosen clones for the three different expression vectors with a reference to the respective figure in the appendix for result description.

Expression vector	Chosen clone	Reference to results
pSET152_ermE_new_MCS	Clone 1	Figure A9
pUWL201PW_new_MCS	Clone 3	Figure A10
pWHM1120_new_MCS	Clone 2	Figure A11

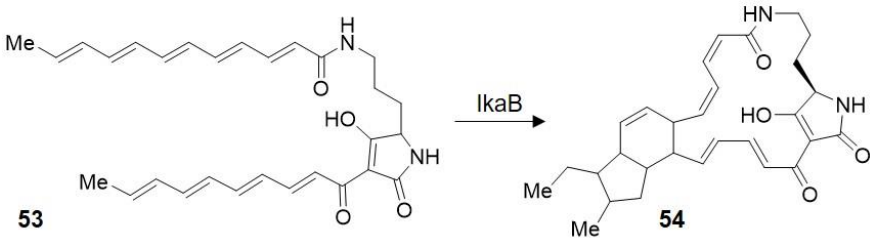
The successful cloning of these three constructs is a further proof that long amplicon PCRs are a powerful tool when manipulating DNA. In this case it was possible to construct three expression plasmids that will simplify the capturing of BGCs, as only one PCR product is needed that can be integrated into all three expression vectors, thereby saving time and resources. Having these constructs cloned, they were applied in all subsequent work of this thesis.

However, when starting the cloning with these constructs it became obvious that the reverse orientation of the *ermE* promoter in pSET152_ermE was not considered when cloning the constructs containing the new MCS. This results in a wrong orientation of the insert in pSET152_ermE_new_MCS when using the same primers as for pUWL201PW and pWHM1120. Results of incorrectly oriented clusters into pSET152_ermE_new_MCS will not be discussed in this thesis, as the resulting constructs could not be used for heterologous

expression of NPs. Accordingly, for most of the constructs, primers were designed for pSET152_ermE_new_MCS and additional primers for pUWL201PW_new_MCS and pWHM1120_new_MCS.

2.3 Heterologous expression of ikarugamycin intermediates

For the better understanding of the mechanisms underlying the transformation of **53** to **54** (Scheme 8) during ikarugamycin biosynthesis, one goal of the thesis is the expression, purification and structure elucidation of the two intermediates **53** and **54**. By now it is postulated that IkaB catalyzes the first ring closure of the common intermediate **53**, followed by a possible non enzymatic Diels-Alder reaction. However, if IkaB would also catalyze the Diels-Alder reaction, this would be one of the first examples of an enzyme catalyzing this kind of reaction.



Scheme 8. Transformation of **53** to **54** catalyzed by IkaB.

To get access to the intermediates, incomplete versions of the *ika* BGC, lacking *ikaB/C* or *ikaC*, were cloned into the pSET152_ermE_new MCS vector, as this vector showed the best results when expressing **39**. Using the long amplicon PCR strategy, primers were designed, flanking the clusters and additionally containing the overhang, newly integrated into the pSET152_ermE_new_MCS vector. The reverse primer could be used for both incomplete clusters *ikaA* and *ikaAB*, however the forward primer was unique. Using the pSET152_ermE::*ika* plasmid as a template, both inserts could be easily amplified using long amplicon PCR (Figure 24).

Cloning by Gibson assembly (chapter 4.2.7) was conducted, linking the NdeI digested (chapter 4.2.5) and dephosphorylated (chapter 4.2.6) vector to the purified (chapter 4.2.4) DNA inserts. To verify the cloning success, colony PCR was conducted (chapter 4.2.13) by picking some of the new clones. If a positive band was detected, the plasmid DNA was purified (chapter 4.2.4) and analyzed with an analytical restriction digest (chapter 4.2.5). The results are shown in Figure 25. Clones with positive restriction digest results were sent to be verified by Sanger sequencing (chapter 4.2.14) (Figure A21).

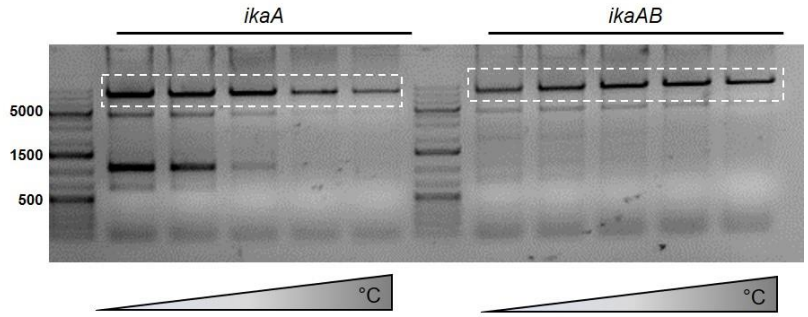


Figure 24. Long amplicon PCR for amplifying *ikaA* and *ikaAB*. An agarose gel with the results of a temperature gradient PCR (50-70 °C) is depicted for the amplification of the two incomplete versions of the *ika* BGC. The product is suitable for Gibson assembly with pSET152_ermE_new_MCS.

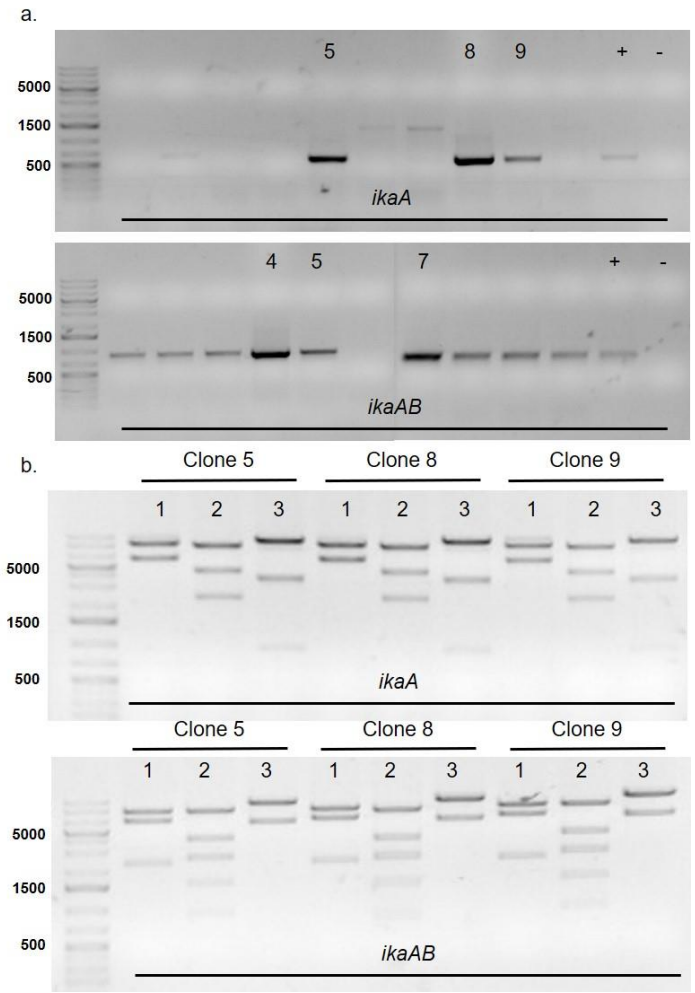


Figure 25. Analysis to detect positive clones for *ikaA* and *ikaAB* in pSET152_ermE_new_MCS. a. Results of the colony PCR for *ikaA* (upper lane) and *ikaAB* (lower lane) to detect possible correct clones, negative control conducted with water, positive control 1 µl of Gibson assembly reaction. b. Results of analytical restriction digest with all clones showing the predicted restriction pattern.

The cloning results underline again the efficiency of the new cloning strategy. With only 10 clones screened for each construct, at least one positive clone could be identified. Compared

to other gene clusters, the size of 9375 bp (*ikaA*) and 11252 bp (*ikaAB*) is not large. However, the GC-content for both is over 70%, usually a very high burden for PCR amplification. The high performance of the Q5 polymerase combined with the efficient ligation by Gibson assembly lead to the successful cloning with this new long amplicon method.

In parallel, the cloning of both *ikaA* and *ikaAB* into pUWL201PW and pWHM1120 were conducted. The two clusters were amplified in one PCR reaction (chapter 4.2.13) and purified by Gel extraction (chapter 4.2.4). Simultaneously, the vectors were digested with NdeI (chapter 4.2.5) and dephosphorylated (chapter 4.2.6). Vector and insert were linked using Gibson assembly (chapter 4.2.7) and transformed into *E. coli* DH5α using heat shock transformation (chapter 4.2.10).

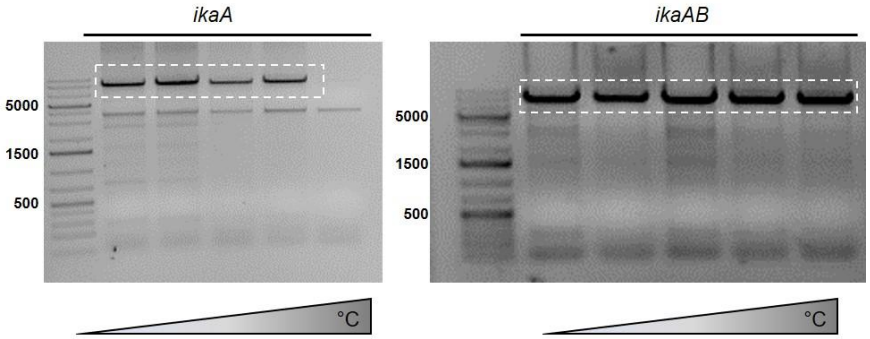


Figure 26. Long amplicon PCR for *ikaA* and *ikaAB*. Agarose gels with the results of a temperature gradient PCR (50-70 °C) are depicted for the amplification of the two truncated versions of the *ika* BGC, the product is suitable for Gibson assembly into pUWL201PW_new_MCS and pWHM1120_new_MCS.

As it was not possible to obtain colonies for pWHM1120_new_MCS::*ikaA* nor pWHM1120_new_MCS::*ikaAB*, the cloning of these constructs was not continued. Possible correct colonies for pUWL201PW were identified by colony PCR (chapter 4.2.13), plasmid DNA was extracted (chapter 4.2.4), analyzed with an analytical restriction digest (chapter 4.2.5) and DNA showing the right sizes was submitted to sequencing (chapter 4.2.14).

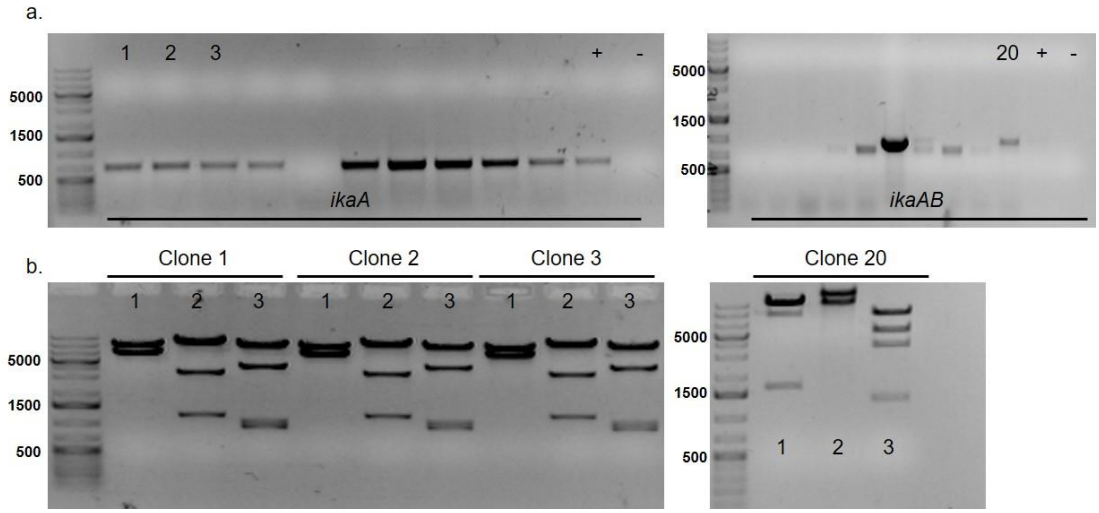


Figure 27. Analysis to detect positive clones for *ikaA* and *ikaAB* in pUWL201PW_new_MCS. a. Results of the colony PCR for *ikaA* (left) and *ikaAB* (right) to detect possible correct clones, negative control conducted with water, positive control 1 µl of Gibson assembly reaction. b. Results of analytical restriction digest with all clones showing the predicted restriction pattern.

Sequencing results show that pUWL201PW_new_MCS::*ikaA* did not contain any mutations, however, pUWL201PW_new_MCS::*ikaAB* had an insertion of the three bases ATG at the beginning of *ikaA*. As this was the only correct clone detected after a few screening steps it was decided to continue with this clone and see if the additional methionine would interfere with PoTeM production.

After cloning the expression constructs, the vectors needed to be transferred into *Streptomyces*. For conjugation, pSET152_ermE_new_MCS::*ikaA* and pSET152_ermE_new_MCS::*ikaAB,* were transformed by electroporation into the donor strain *E. coli* ET12567 pUZ8002 (chapter 4.2.10). PCR-verified clones (chapter 4.2.13) were used for conjugation into the three *Streptomyces* strains *S. albus*, *S. lividans*, and *S. coelicolor* (chapter 4.2.12). Exconjugants were again verified by PCR and one clone was used to inoculate a preculture in CASO. This culture was used to inoculate a 50 ml ISP-4 culture (1:10). Using the best expression conditions gained in the previous experiments, the cultures were grown for seven days at 28 °C at 200 rpm. To extract the NP produced in the culture, cell pellet and supernatant were separated using centrifugation, cells were extracted with methanol and acetone, and the supernatant was extracted with ethyl acetate (chapter 4.2.19). Extracts were dissolved in methanol, filtered through 0.02 µm syringe filters and measured utilizing LC-MS (chapter 4.2.20) with the method shown in Table 36. However, none of the extracts contained a peak with the corresponding mass detectable. Accordingly, cultivation was repeated, this time shorter culturing periods were used in case of instability of the compounds **53** and **54**. But again, the desired product was not detectable in the HPLC chromatogram, indicating that the

expression construct might not be suitable for heterologous expression of NPs in *Streptomyces.*

In parallel pUWL201PW_new_MCS::*ikaA* and pUWL201PW_new_MCS::*ikaAB* were tried to be integrated into *Streptomyces* using protoplast transformation (chapter 4.2.11). However, none of the transformation trials resulted in exconjugants, suggesting that the protoplasts were not able to take up DNA. After several new protoplast isolation and transformation trials it was not possible to integrate the two expression constructs into *Streptomyces*. Therefore expression using pUWL201PW_new_MCS::*ikaA* and pUWL201PW_new_MCS::*ikaAB* was discontinued.

As the heterologous expression of **39** worked for all constructs with very high yields, it was quite surprising that it would not work for the intermediates. Therefore, we went back and examined the difference between the two new constructs compared to the first pSET152_ermE::*ika*. The only difference between the plasmids was the newly integrated MCS. However, this newly integrated sequence also enlarged the distance between the RBS (CCTCC) and the start codon (ATG, reverse complement TAC) (Figure 28). As *Streptomyces* are able to tolerate up to 50 bp distance between RBS and start codon, the normal distance should be around 5 to 15 bp.[127] In this case the distance of 16 bp as in the pSET152_ermE::*ika* expression construct was tolerated whereas 41 bp as in the expression constructs pSET152_ermE_new_MCS::*ikaA* and pSET152_ermE_new_MCS::*ikaAB* possibly resulted in lack of expression. Therefore, it was not possible to detect **53** or **54** in the heterologous cultures.

Consequently, new primers were designed suitable for *ikaA* and *ikaAB* amplification with overhangs exactly matching the pSET152_ermE::*ika* construct. The reverse primer used to amplify the whole *ika* BGC could be recycled and only new forward primers were needed.

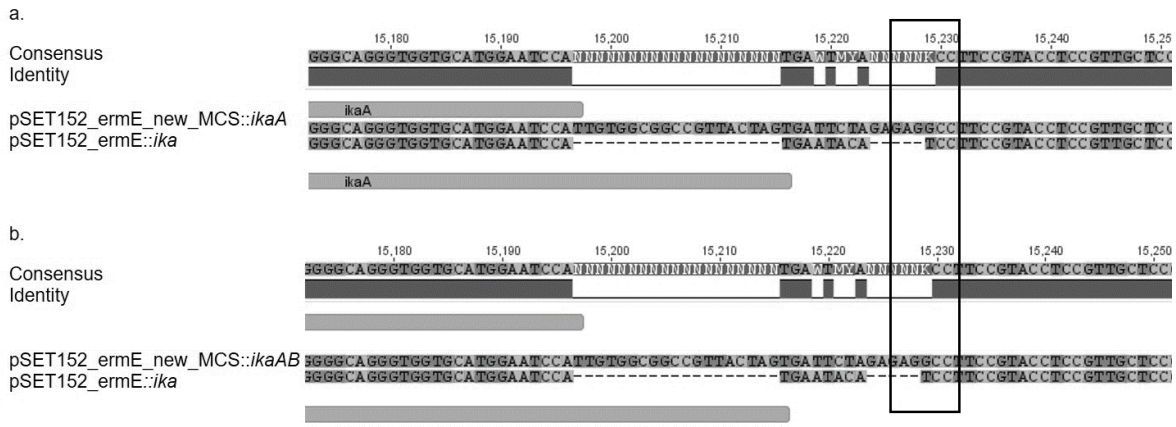


Figure 28. Alignment of pSET152_ermE_new_MCS::*ikaA* and ::*ikaAB* to pSET152_ermE::*ika*. The alignment of *ikaA* (a.) and *ikaAB* (b.) in the expression construct pSET152_ermE_new_MCS aligned to the expression construct pSET152_ermE::*ika* depict the difference obtained by the integration of the new MCS. The distance between RBS (see black box) and the start codon was changed from 16 bp to 41 bp.

The new long amplicon PCR (chapter 4.2.13) resulted in good amplification of the two incomplete versions of the *ika* BGC (Figure 29). As previously, the vector was digested with StuI (chapter 4.2.5) and dephosphorylated (chapter 4.2.6). Subsequently the purified PCR product (see chapter 4.2.4) and the linearized vector were ligated by Gibson assembly (chapter 4.2.7) and transformed into *E. coli* DH5α using heat shock (chapter 4.2.10).

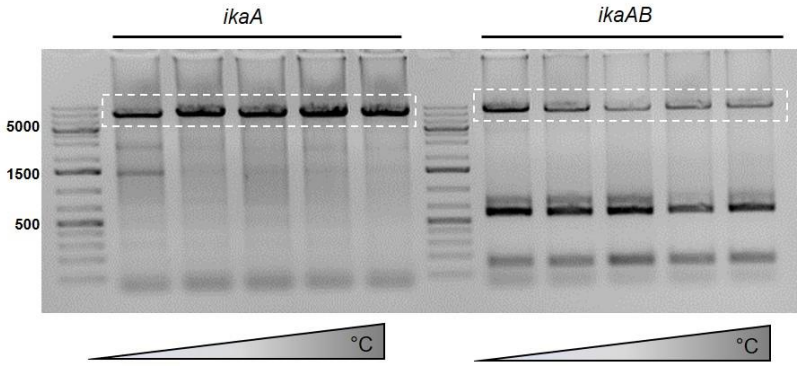


Figure 29. Long amplicon PCR for *ikaA* and *ikaAB*. An agarose gel with the results of a temperature gradient PCR (50-70 °C) is depicted for the amplification of the two truncated versions of the *ika* BGC, the product is suitable for Gibson assembly into pSET152_ermE.

Colonies were picked and analyzed concerning the possible integration of the incomplete clusters by colony PCR (chapter 4.2.13). Clones showing a band at the right size were grown overnight and plasmid DNA was extracted (chapter 4.2.4). The plasmids were analyzed with an analytical restriction digest (chapter 4.2.5) (Figure 30) and one possible positive clone for each construct was sent for sequencing (chapter 4.2.14). Construct pSET152_ermE::*ikaA* was found without mutations in clone 8 and pSET152_ermE::*ikaAB* was found in clone 3 without mutations (Figure A23). Accordingly, these plasmids were chosen for conjugation into *Streptomyces*.

The purified plasmids were transformed into *E. coli* ET12567 pUZ8002 using electroporation (chapter 4.2.10) and integration of the plasmid was verified utilizing colony PCR (chapter 4.2.13). Positive clones were grown to conduct conjugation with *S. albus*, *S. lividans*, and *S. coelicolor* (chapter 4.2.12). Exconjugants were verified using colony PCR (chapter 4.2.13) and five different clones were used to inoculate precultures. It is published that the position of cluster integration into the host genome can greatly influence the amount of transcribed mRNA which subsequently would significantly influence the amount of NP produced.[128] Therefore, the number of screened clones was enlarged for 1 to 5 to ensure that the absence of NP is not be due to the location of the cluster integration into the genome.

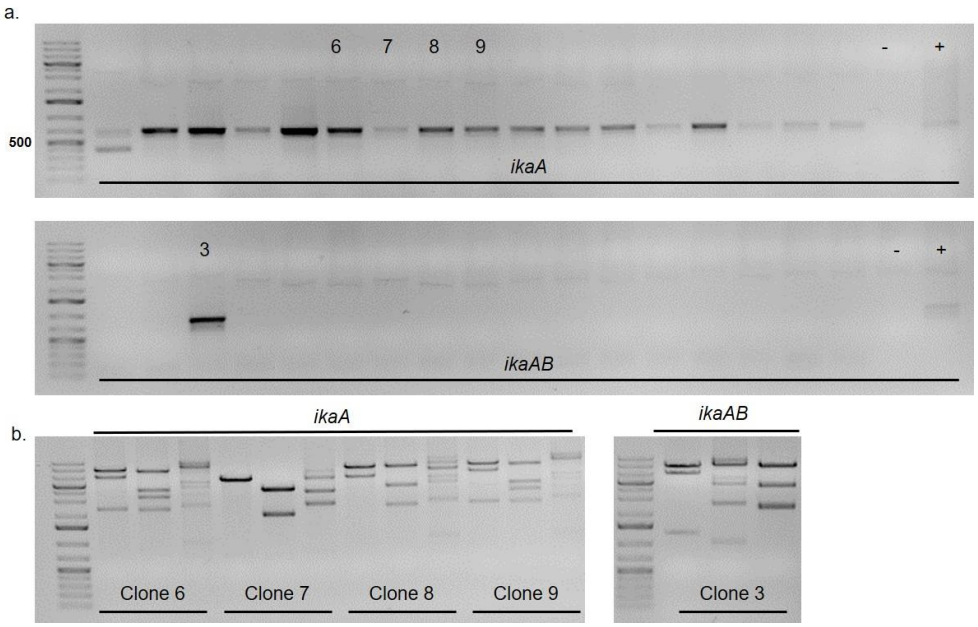


Figure 30. Analysis to detect positive clones for pSET152_ermE::*ikaA* and *::ikaAB*. a. Results of the colony PCR for *ikaA* (upper lane) and *ikaAB* (lower lane) to detect possible correct clones, negative control conducted with water, positive control 1 µl of Gibson assembly reaction. b. Results of analytical restriction digest with clone 8 for *ikaA* and clone 3 for *ikaAB* showing the predicted restriction pattern.

2.3.1 Heterologous expression of **53**

To get access to the heterologously expressed NPs in the cultured *Streptomyces* harboring the pSET152_ermE::*ikaA* expression vector, pellet and supernatant of the cultures were separated by centrifugation and extracted with organic solvents (chapter 4.2.19). The extracts were analyzed utilizing LC-MS using the method shown in Table 36. However, after a cultivation of seven days, no peak with the corresponding mass of m/z = 475.2597 $[M+H]^+$ was observed (Figure 31).

As depicted in Figure 31, all the three utilized heterologous expression host did not show any significant peaks besides the polar molecules with retention times between 0 to 2 min. Furthermore, *S. lividans* and *S. coelicolor* pellet extracts shown two peaks around a retention time of 8 min, those peaks are also present in wildtype cultures (data not shown), accordingly they do not result from the expression plasmid. A beneficial feature of *S. albus* as a heterologous host for PoTeM production compared to *S. lividans* and *S. coelicolor* is the presence of an endogenous PoTeM BGC. Considering the common precursor for any molecule of this NP class, *S. albus* should be able to provide general cellular biosynthesis platform for PoTeMs. Consequently, continuative experiments for **53** expression were conducted with *S. albus* as the only heterologous host.

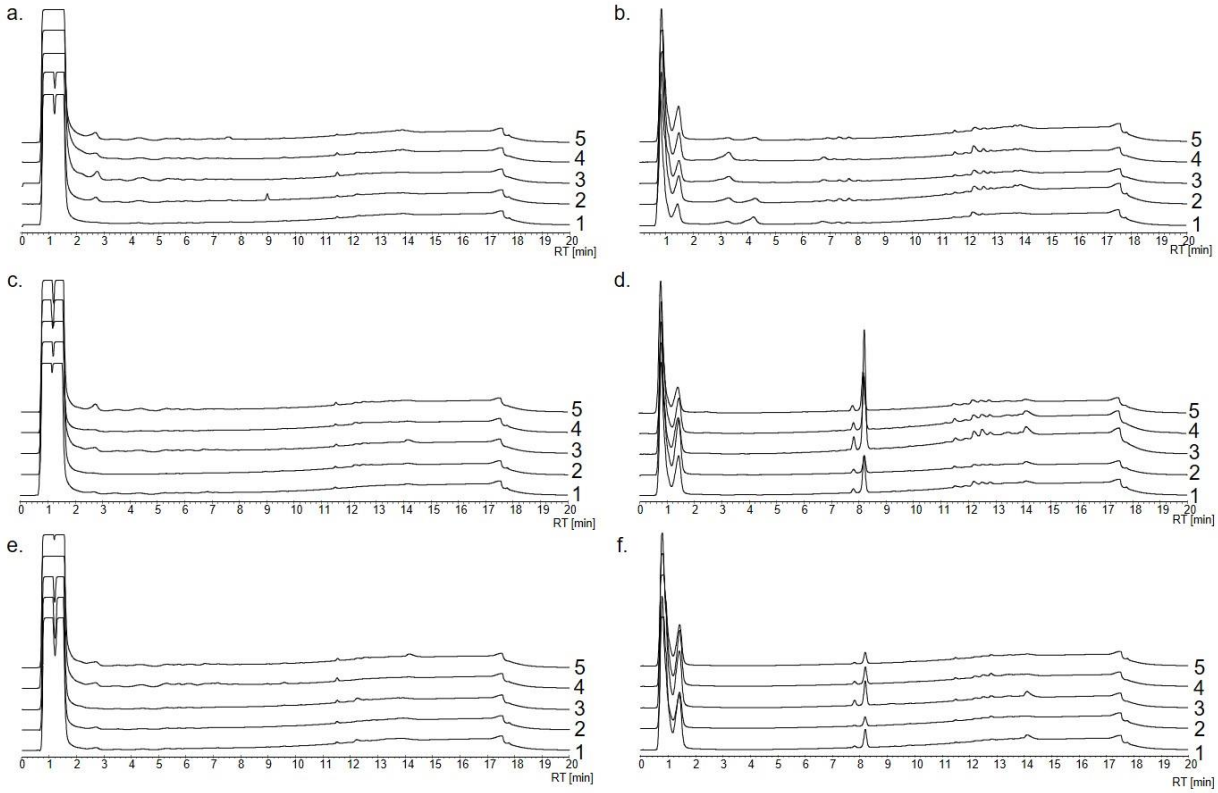


Figure 31. HPLC results for expression of pSET152_ermE::*ikaA* in *Streptomyces*. Depicted are the HPLC chromatographs obtained when measuring the extracts form samples a. *S. albus* supernatant, b. *S. albus* pellet, c. *S. lividans* supernatant, d. *S. lividans* pellet, e. *S. coelicolor* supernatant and f. *S. coelicolor* pellet. For each sample five clones were picked, numbered with 1 to 5. Analytical method: Eurosphere_IkaFAST_CG.Meth.

Taking the possible instability of the intermediate into account, an expression with a shorter incubation time was tried for the clones of the *Streptomyces* strain *S. albus*. However, no product could be detected after three nor five days.

Two possible reasons for failure of expressing **53** in *S. albus* are feasible: first, the constitutive activity of the *ermE* promoter impairing expression control and second the endogenous PoTeM BGC present in the *S. albus* genome, possibly encoding a cluster for **41**. As the characteristic of PoTeMs is the common biosynthetic precursor that is the same for each compound, the endogenous cluster might consume the heterologously expressed compound, which can therefore not be found in the extracts anymore. Support for the assumption is the small peak visible in the *S. albus* pSET152_ermE::*ikaA* cell pellet extracts (Figure 31b) at 3 min that shows a mass of *m/z* = 511.2 $[M+H]^+$, the same mass as **58**. As a consequence, two approaches to improve **53** expression were conducted: (1) cloning of an expression vector harboring an inducible promoter and (2) a knockout *S. albus* strain needed to be established, to diminish consumption of **53** and thereby increasing the yields of **53** to enable structure elucidation.

The expression vector pAEM35 (Figure A6) with the inducible promoter *PnitA* was selected.[129] This vector does not contain an OriT and TraJ element to be suitable for *Streptomyces*

conjugation. First trials to integrate this vector into *Streptomyces* using protoplast transformation failed. It was assumed that the storage of the protoplast at -80 °C for a longer period decreased their ability to integrate plasmids. Therefore, new protoplasts were prepared (chapter 4.2.3). However, probably due to an incomplete digestion of the cell wall protoplast transformation was no longer possible. To facilitate *Streptomyces* conjugation, the plasmid needed to be modified with an OriT and TraJ element. For this modification, primers were designed to amplify these two features from pSET152 (chapter 4.2.13), and they were integrated into a PciI digested (chapter 4.2.5) and dephosphorylated (chapter 4.2.6) pAEM35 vector using ligation cloning (chapter 4.2.7). This approach did not result in a positive clone. Therefore, the cloning strategy was changed to instead incorporate the PnitA promoter into pSET152. Again, primers were designed, flanking the NitA and NitA-resistance region of pAEM35 (chapter 4.2.13), to be integrated into NotI digested (chapter 4.2.5) and dephosphorylated (chapter 4.2.6) pSET152 by Gibson assembly (chapter 4.2.7). Similar to the modification of pAEM35 with OriT, no colony containing the right construct could be identified. This might be due to the presence of two identical NitA regions in both constructs, as internal repeat sequences usually strongly interfere with cloning efficiency. Additionally, the targeted PCR amplification of the NitA and NitA resistance region always results in unspecific amplification, as any designed primer will bind twice in the target region. Even after gel extraction it cannot be excluded that the smaller fragment contaminates the Gibson assembly reaction. Another critical point for this cloning is the requirement of homology regions for Gibson assembly, which will be disturbed by two 103 bp regions that are completely identical. Accordingly, due to the inability to obtain positive clones, the expression of **53** with an inducible promoter was discontinued.

For the establishment of a **58**-knockout strain, an intermediate plasmid needed to be generated harboring the 5' and 3' homology ends to the original cluster encoding for **58** and a selection marker. Therefore, pSET152 was chosen as a suitable vector. Usually up to 5000 bp of the cluster flanking region are chosen to serve as the homology arms. However, the strain *S. albus* DSM40313 was not sequenced. Therefore the homology arms were designed based on the published PoTeM gene cluster for *S. albus* J1074.[116] 702 bp of the 5' sterol desaturase and 1000 bp of the 3' cytochrome P450 were chosen as homology arms, these sizes would ensure inactive enzymes in the final knock out strain. As a selection maker, a thiostreptone resistance cassette was integrated in between the two homology regions. The cloning strategy was a four fragment Gibson assembly with the two homology regions amplified from *S. albus* DSM40313 gDNA, the resistance cassette amplified from pNHD_Duet_Rab1b_DrrA_mGEF, kindly provided by the Lang Lab, and BamHI linearized pSET152.

To clone the knockout plasmid, pSET152 was digested with BamHI (chapter 4.2.5) and dephosphorylated (chapter 4.2.6). In parallel, the three insert DNA fragments were PCR amplified (chapter 4.2.13), as shown in Figure 32, and gel extracted (chapter 4.2.4). The four fragments were ligated in a four fragment Gibson assembly reaction (chapter 4.2.7) and transformed into *E. coli* DH5α by heat shock. The transformation was plated on LB agar plates containing apramycin and clones were picked to perform colony PCR (chapter 4.2.13) to detect possible positive clones. Restriction analysis and Sanger sequencing were performed to verify the cloning success (Figure A24). However, when culturing *E. coli* harboring the possible knockout plasmid in LB containing thiostreptone, there was no growth detectable, indicating an unfunctional ThioR cassette used for this plasmid. Additionally, it became obvious that a pSET152-based knockout plasmid is not a suitable choice, as pSET152 contains a high copy OriT, causing a high number of plasmid copies in the cell. For the homologous recombination that is needed for the knockout, a low copy plasmid would be more efficient.

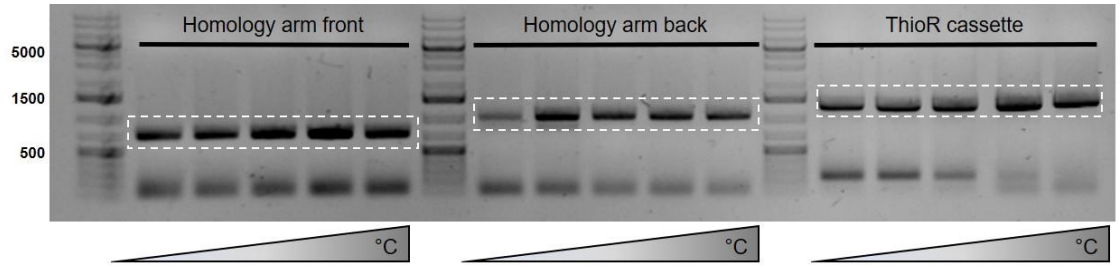


Figure 32. Temperature gradient PCR results for pSET152-based knockout plasmid. Gel picture shows DNA detected after temperature gradient PCR (50-70 °C) for the three PCR fragments of the pSET152-based knockout plasmid. Homology arm front: 735 bp sterol desaturase; Homology arm back: 1034 bp cytochrome P450 and ThioR cassette: 1296 bp thiostreptone resistance with promoter.

Based on the insufficient thiostreptone resistance and the high copy number of pSET152, a new knockout plasmid was designed based on low copy pCC1FOS vector. In this case a four fragment Gibson assembly approach was designed. Again, the two homology regions were amplified from *S. albus* DSM40313 gDNA. However, this time a 3000 bp region up and downstream of the PoTeM BGC was chosen, based on the genome data of *S. albus* J1074. Hoping that the flanking regions of the cluster in the two strains are similar, both homology regions contained a 50 bp fragment of the cluster genes, sterol desaturase and cytochrome P450. The pCC1FOS backbone was amplified from a cluster containing pCC1FOS plasmid (pDELcluster06MluI) kindly provided from Zhengyi Qian and the ThioR cassette was amplified from pUWL201PW.

The four fragments were amplified in a Q5-PCR (chapter 4.2.13), shown in Figure 33. The PCR products were impure and needed to be purified using gel extraction (chapter 4.2.4). The isolated fragments were subsequently ligated in a Gibson assembly reaction (chapter 4.2.7).

This reaction mixture was transformed into chemically competent *E. coli* DH5α. Colonies grown on LB agar plates containing Ampicillin were screened for successful assembly using colony PCR (chapter 4.2.13). As it was possible to identify correctly assembled plasmid using analytical restriction digestion (chapter 4.2.5) clone 12 was send for Sanger sequencing (chapter 4.2.14) and was proven to be correctly assembled (Figure A25).

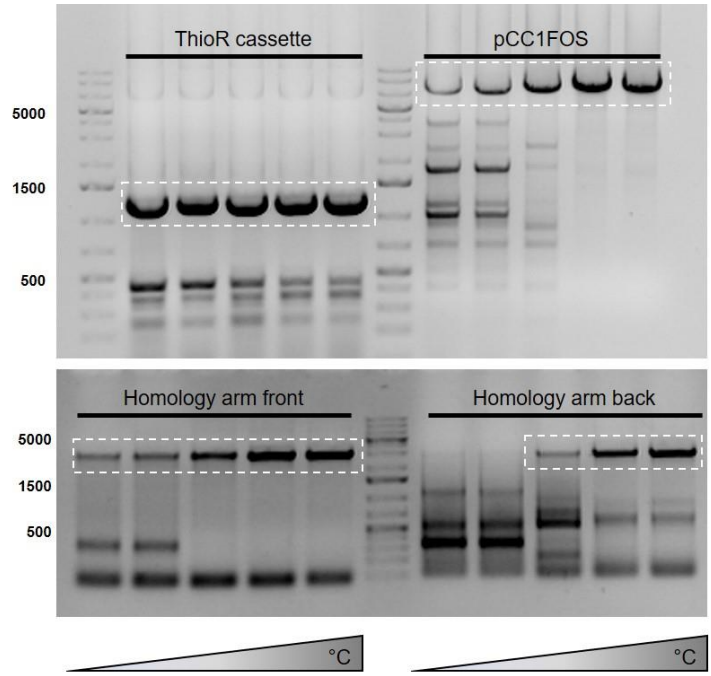


Figure 33. Temperature gradient PCR results of pCC1FOS-based knockout plasmid. Gel picture shows DNA detected after temperature gradient PCR (50-70 °C) for the four PCR fragments of the pCC1FOS-based knockout plasmid. ThioR cassette: 1176 bp thiostreptone resistance with promoter; pCC1FOS: 7205 bp pCC1FOS backbone; Homology arm front: 3030 bp sterol desaturase; Homology arm back: 3030 bp cytochrome P450.

The plasmid was transformed into *E. coli* ET12567 pUZ8002. Colonies were checked for successful plasmid uptake by colony PCR (chapter 4.2.13) and used for conjugation into *S. albus* DSM40313 (chapter 4.2.12). Conjugation was performed on MS-agar without antibiotic and after 20 h incubation the plates were overlaid with nalidixic acid and thiostreptone. Cells were incubated at 30 °C for 7 d to allow homologous recombination. Five exconjugants were picked, grown in CASO and gDNA was extracted (chapter 4.2.4). To verify the knockout, two different colony PCRs were performed using gDNA of *S. albus* DSM40313 as a control. The first screening PCR was done using a forward primer binding at the end of the front homology arm and a reverse primer binding at the beginning of the back homology arm (Figure 34 Screening 1). Exconjugants accomplishing the knockout would show a band at 1876 bp, as the sequence between the primers is cut out and only the thiostreptone resistance would be amplified. In case of the gDNA, no amplicon is expected, as the elongation time of the PCR is too short to accomplish amplification of the whole cluster. To exclude a positive

result in this screening due to plasmid contamination, a second screening was performed using primers binding inside the knocked-out cluster region. In this case an amplicon is expected for the *S. albus* DSM40313 gDNA. However, no band should be visible for the exconjugants. As visible in Figure 34 Screening 2, exconjugant 1, 4 and the gDNA show a band, however exconjugant 2, 3, and 5 do not. Thereby these three exconjugants proved to have replaced the *S. albus* PoTeM cluster by a thiostreptone resistance cassette.

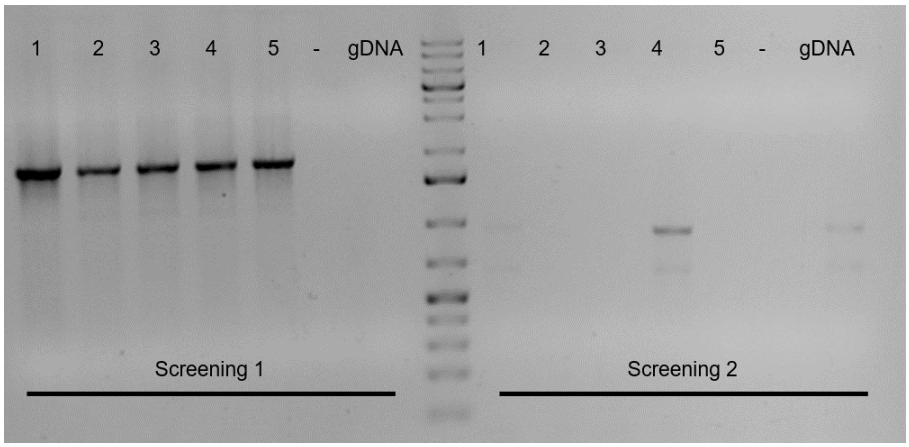


Figure 34. Screening *S. albus* DSM40313 for cluster (**58**) knockout. Five exconjugants were examined according to accomplished homologous recombination to knockout the internal PoTeM cluster. In screening 1 primers SEQ_Primer_albus_KO_fwd/rev were used, for successful knockout a 1876 bp fragment was expected, in the two controls negative: H_2O and *S. albus* DSM40313 gDNA no amplicon was expected. In screening 2 primers SEQ_Primer_albus_KO_fwd and Cyp_fwd were used, this negative screening shows no band in case of successful knockout and a band of around 1000 bp in case of a not accomplished knockout.

Having the knockout *S. albus* (named *S. albus* ΔPoTeM) strain, the pSET152_ermE::*ikaA* expression plasmid could be integrated again to determine whether **53** expression is possible without the endogenous PoTeM cluster. Accordingly, conjugation with all three positive knockout exconjugants was performed (chapter 4.2.12). Successful integration of five exconjugants of each knockout clone was verified using colony PCR (chapter 4.2.13) and grown in CASO. A 50 ml ISP-4 culture was inoculated 1:10 with a dense preculture and incubated at 28 °C, 200 rpm for 7 d. NPs were extracted using organic solvents (chapter 4.2.19) and analyzed using LC-MS with the method shown in Table 36.

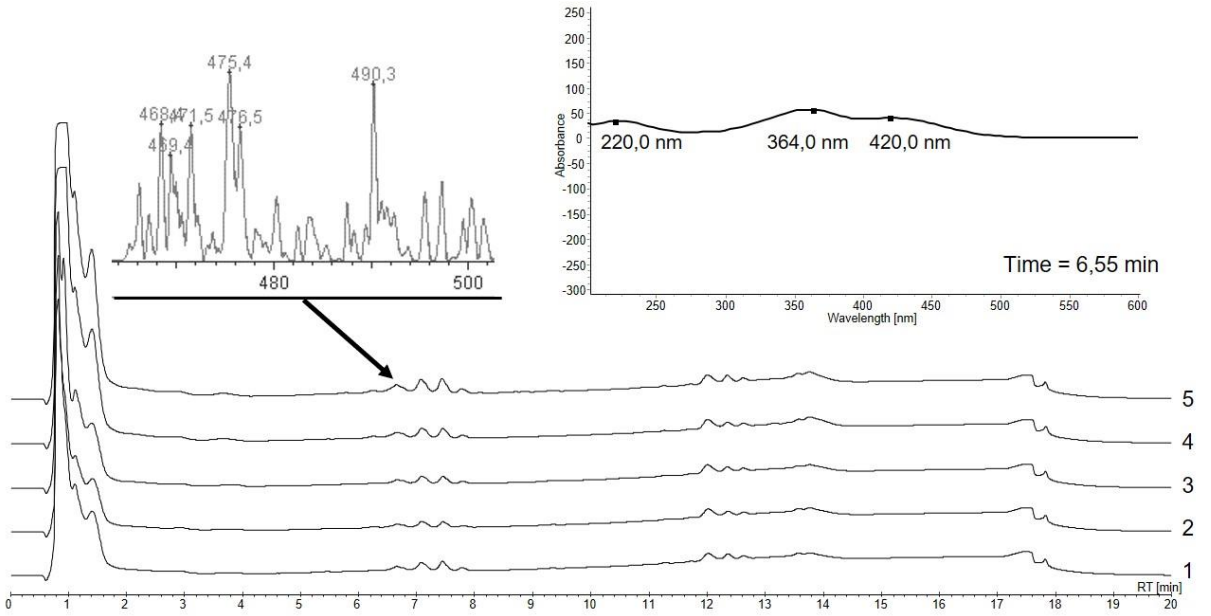


Figure 35. LC-MS analysis of **53** expression in *S. albus* ΔPoTeM clone 5. Depicted are the HPLC chromatograms of the five *S. albus* ΔPoTeM exconjugants harboring pSET152_ermE::*ikaA*. In the peak at 6.75 min a mass of *m/z* = 475.4 $[M+H]^+$ suitable for **53** can be detected. The UV-spectrum of the peak shows the published spectrum of **53**.[113] Analytical method: Eurosphere_IkaFAST_CG.Meth.

For the first time, a peak with the corresponding mass of **53** (*m/z* = 475.4 $[M+H]^+$) was detected (Figure 35). Furthermore, the UV-absorption spectrum of the peak at 6.55 min was in accordance with the published spectrum of **53**.[113] However, the amount was not enough for purification and structural elucidation utilizing NMR analysis. Therefore, the culturing conditions were optimized to increase the amount of **53**. The best results were detected using clone 5 of the new *S. albus* ΔPoTeM strain harboring pSET152_ermE::*ikaA* for Clone 5 (Figure 35). Therefore, this clone was used for further optimization. As **53** is an unstable compound, a cultivation on agar plates and extraction with organic solvents from agar was attempted (chapter 4.2.19). As production on agar could be observed (Figure 36a and b), a time course for optimal **53** production was measured, suggesting an incubation of the agar plates at 30 °C for 5 d (Figure 36c).

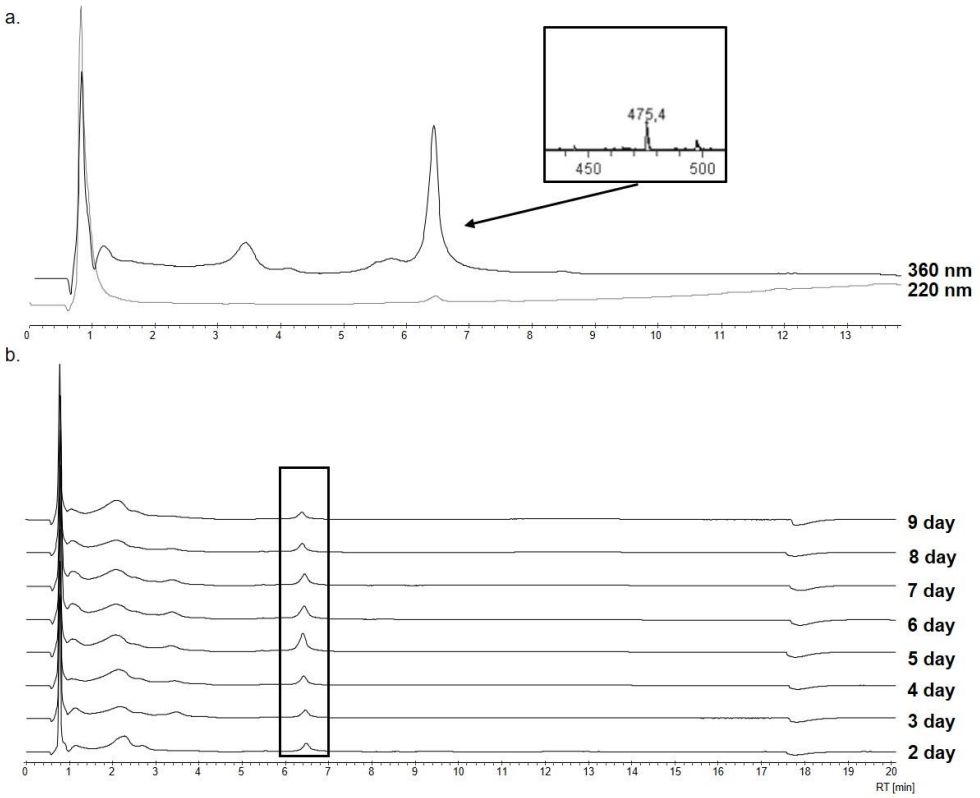


Figure 36. Optimization of **53** expression. *S. albus* ΔPoTeM pSET152_ermE::*ikaA* was grown on ISP-4 agar and extracted with organic solvents, a. results of HPLC analysis of cell extracts measured at 220 nm (light grey, common wavelength for NPs) and 360 nm (dark grey, absorption maximum of **53**) and the corresponding mass results at 6.5 min. b. HPLC chromatograms at 360 nm of the time course measured using ISP-4 agar plate extracted after incubation of the assigned days. Analytical method: Eurosphere_IkaFAST_CG.Meth.

Taken together, this thesis established a method to overexpress **53**, the common intermediate in PoTeM biosynthesis. In this case, a new knockout strain needed to be constructed to minimize interference of endogenous clusters with the unstable compound. In contrast to **39**, this compound is produced in higher quantities on ISP-4 agar plates with an optimal incubation of 5 d. A longer incubation causes a decrease in detectable NP. This could be due to a decreased production by the heterologous host but is additionally caused by the instability of the compound, which is most likely degraded during longer incubation times. Based on the established expression protocol, elucidation of the structure on **53** by NMR will be possible for the first time, leading to a better understanding of the biosynthesis of PoTeMs. By now the exact mechanisms underlying the transformation from **53** to **54** are not well understood. Therefore, access to **53** in sufficient amounts will help to characterize IkaB and the transformations necessary for **39** biosynthesis. One possibility to use the heterologously expressed compound would be a co crystallization of **53** with IkaB. Thereby the structural arrangement of **53** in IkaB can be determined and further information for the mode of action of IkaB can be obtained. Furthermore, higher amounts of **53** could support the elucidation of the differences needed to retain a different cyclization pattern in different PoTeMs. Presumably

the first ring closure is crucial for further modifications and thereby determines the final cyclization pattern. Accordingly, this part of the thesis has built a promising basis for the analysis of PoTeM biosynthesis by making the common intermediate accessible.

2.3.2 Heterologous expression of **54**

Besides the expression of the first intermediate **53**, a second goal was the expression and purification of the second intermediate **54**. As described above, five exconjugants of *S. albus* DSM40313 pSET152_ermE::*ikaAB* were grown in ISP-4 for 7 d, pellet and supernatant were extracted (chapter 4.2.19) and analyzed by HPLC using the method shown in Table 36.

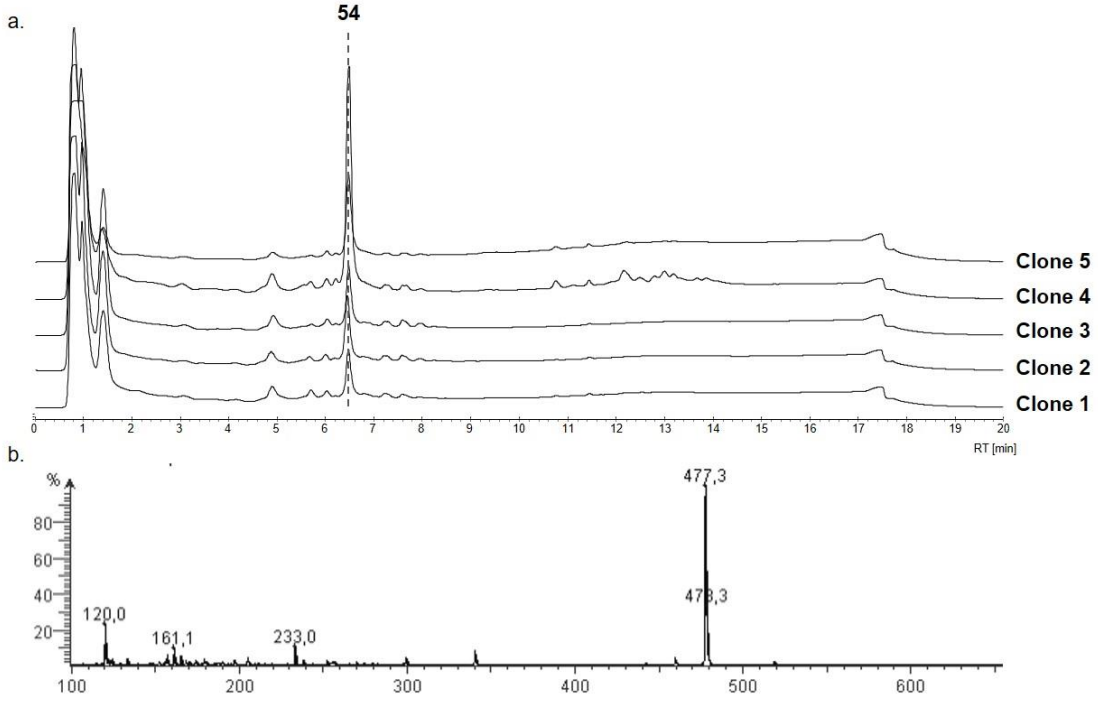


Figure 37. LC-MS analysis of *S. albus* pSET152_ermE::*ikaAB* pellet extracts. a. HPLC chromatograms of pellet extracts from *S. albus* pSET152_ermE::*ikaAB* cultures, b. mass analysis at a retention time of 6.5 min depicted for clone 4 that showed the best **54** expression, the published mass for **54** *m/z* = 477.2678 $[M+H]^+$ is present with almost no impurities.[113] Analytical method: Eurosphere_IkaFAST_CG.Meth.

As visible in Figure 37 a very strong and pure production of **54** could be detected in the pellet extracts of the heterologous expression cultures. These results prove the discussion that the newly integrated MCS in the expression vector was the disturbing factor for successful NP production, as this intermediate could be easily expressed in high amounts. Accordingly, a cryo culture was stored at -80 °C to save this expression strain. However, when reviving this culture, no more production of **54** could be observed. As it is generally proposed that

Streptomyces loose heterologous NP expression efficiency after some time of incubation, for example by methylation of the DNA, a new conjugation (chapter 4.2.12) was performed and again five clones were picked, cultured and extracted (chapter 4.2.19). The extracts were examined using LC-MS with the method shown in Table 36. The results showed a strongly reduced expression of **54** and a much higher metabolic complexity (Figure 38a), unsatisfactory when compared to the results gained in the first expression (Figure 37).

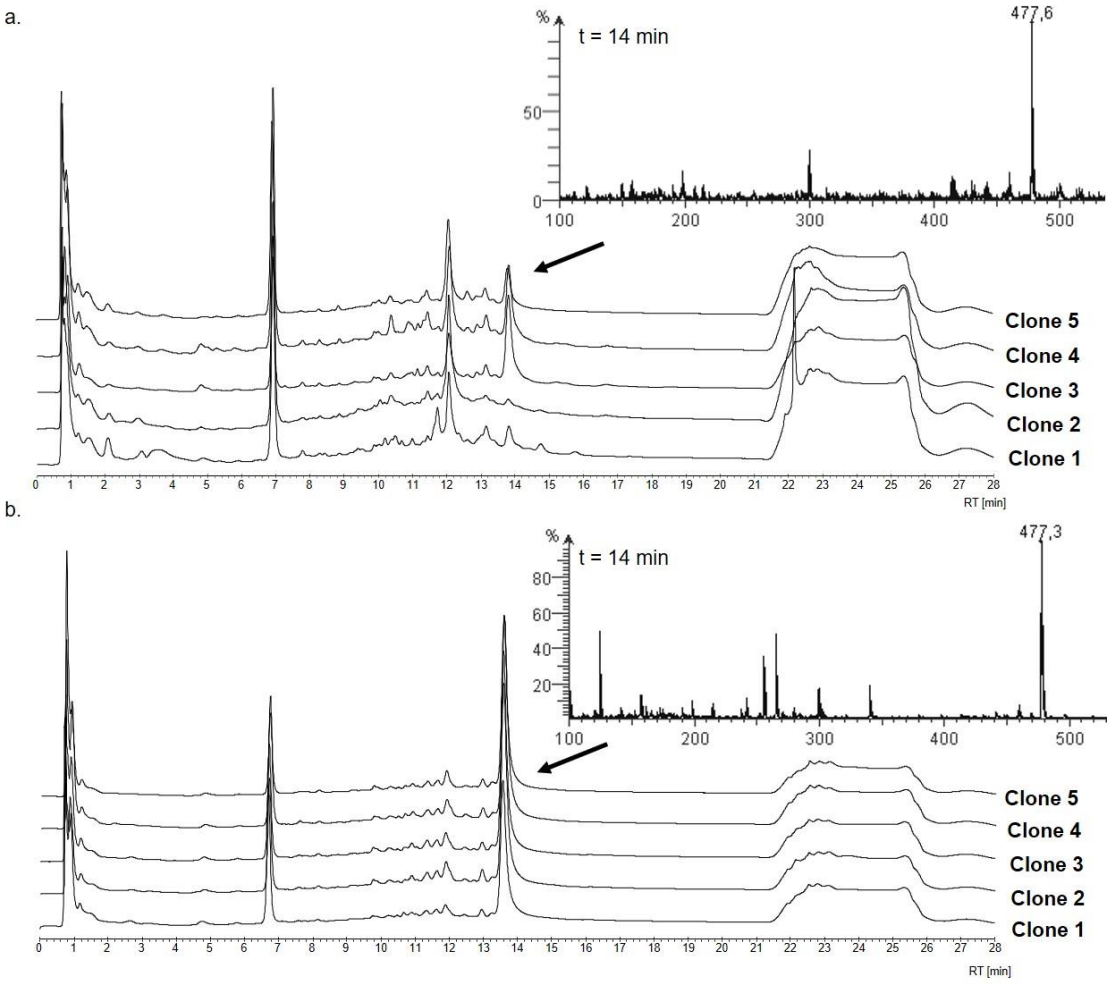


Figure 38. Optimization of **54** expression. a. HPLC chromatograms of *S. albus* pSET152_ermE::*ikaAB* pellet extracts after new conjugation, with mass data for clone 3 at t=14 min showing the best production of **54**. b. HPLC chromatograms of *S. albus* ΔPoTeM pSET152_ermE::*ikaAB* pellet extracts with mass data for clone 5 at t=14 min showing the best production of **54**. Analytical method: Eurosphere_Ika_lange Methode_AG.Meth.

As a consequence, the expression construct pSET152_ermE::*ikaAB* was conjugated into *S. albus* ΔPoTeM (chapter 4.2.12), as this showed promising results for the production of **53**. All five picked exconjugants showed a good expression of **54**, as depicted in Figure 38b. Using the knockout strain, a reliable expression was thus established. These results prove the assumption that *Streptomyces* change the heterologous expression profile after longer storage times. Accordingly, a new conjugation needs to be performed, when the NP production of an active clone decreases. Furthermore, these results are a proof for the fact that PoTeM

intermediates are not stably expressed in *S. albus* DSM40313, as the endogenous PoTeM cluster can use up and modify the intermediates. However, *S. albus* DSM40313 proved to be the best heterologous host for **39** (chapter 2.1.3), which might also be caused by the internal PoTeM cluster, as precursor pathways are already established for the own PoTeM. With the establishment of *S. albus* ΔPoTeM, an optimized PoTeM producing strain has now been generated.

In a next step, the purification of the intermediate was developed. The cultivation was performed in multiple 50 ml cultures, as is was detected before that larger volumes decrease NP production. The cell pellet was extracted, the extract was dissolved in methanol and filtered through 0.02 µm syringe filters. If the solution contained too many particles, it was centrifuged before filtering. Up to 500 µl of the solution were injected into the semi preparative HPLC (chapter 4.2.20) and purified using method shown in Table 38, optimized for the purification of **54**.

A reliable expression and purification method for the accessibility of **54** was thereby established. Having this intermediate in hands, further biosynthetic elucidation of the biosynthetic pathway of **39** can now be performed. On the one hand it can serve as a standard compound for assays of **53** with recombinant IkaB, as **54** is the product of this reaction. This understanding of the reaction would produce a large gain of knowledge of the factors influencing cyclization patterns in different PoTeMs. Additionally, co crystallization studies of **54** with IkaC can be performed to structurally analyze the molecular mode of action of the alcohol dehydrogenase IkaC.

2.4 Heterologous expression of novel PoTeMs

After the reliable establishment of a gene cluster capturing method and a heterologous expression system using **39** as the model compound, the gained expertise was transferred to the predicted novel PoTeM clusters (Table 5). At the beginning of this thesis, the PoTeM cluster from *S. spinosa*, *S. degradans* and *Streptomyces* sp. Tü6314 were already captured from the gDNA of the producing organisms and cloned into the pET-28b-SUMO cloning vector by C. Greunke. The quality of the *P. fermentans* gDNA was not sufficient to capture the whole PoTeM gene cluster. However, the acyltransferase and FAD-dependent oxidoreductase had been cloned into the expression plasmids pHis8-TEV.

Table 5. PoTeM BGCs used in this study.

Original producer	Size [bp]	GC-content [%]	Genes: predicted function	Internal name
S. spinosa	15,976	69	spiA: Alcohol dehydrogenase *spiB*: Cytochrome P450 *spiC*: iPKS/NRPS *spiD*: FAD-dep. oxidoreductase *spiE*: FAD-dep. oxidoreductase	*spi*
S. degradans	12,326	47	*degA*: Sterol desaturase *degB*: iPKS/NRPS *degC*: FAD-dep. oxidoreductase	*deg*
P. fermentans JBW45	12,027	41	*ferA*: Acyl transferase *ferB*: iPKS/NRPS *ferC*: FAD-dep. oxidoreductase	*fer*
S. roseosporus	16,166	72	*rosA*: Sterol desaturase *rosB*: iPKS/NRPS *rosC*: FAD-dep. oxidoreductase *rosD*: FAD-dep. oxidoreductase *rosE*: Alcohol dehydrogenase *rosF*: Cytochrome P450	*ros*
Streptomyces sp. Tü6314	16,128	72	*TüA*: Sterol desaturase *TüB*: iPKS/NRPS *TüC*: FAD-dep. oxidoreductase *TüD*: FAD-dep. oxidoreductase *TüE*: Alcohol dehydrogenase *TüF*: Cytochrome P450	*Tü6314*

The first step in this thesis was the cloning of the different PoTeM clusters in the newly established *Streptomyces* expression vectors pSET152_ermE_new_MCS, pUWL201PW_new_MCS. Accordingly, Gibson primers were designed, one pair for the integration into pSET152_ermE_new_MCS and one pair for pUWL201PW_new_MCS. For the amplification of the clusters, if existent, the pET-28b-Sumo constructs were utilized as Q5 PCR templates. If not, gDNA was extracted from the producing organism (chapter 4.2.4) and used for the PCR (chapter 4.2.13). Results of the long amplicon PCR are depicted in Figure 39, proving that a one piece amplification of clusters up to 16,000 bp can be reproducibly accomplished.

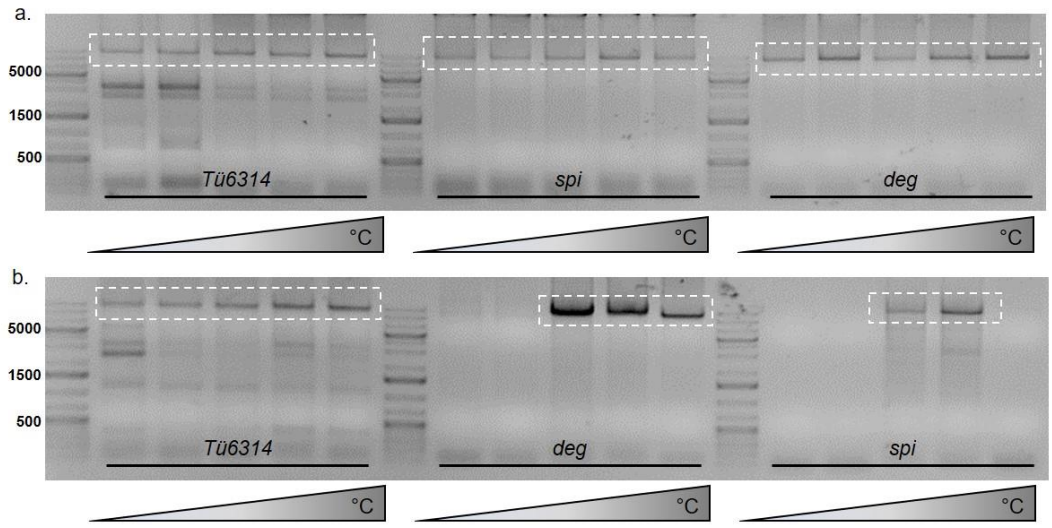


Figure 39. Temperature gradient PCR to amplify PoTeM cluster. a. Results of the long amplicon PCR amplifying the clusters *Tü6134* (16,176 bp), *spi* (16,080 bp) and *deg* (12,376 bp) for integration into pSET152_ermE_new_MCS, b. results of the long amplicon PCR amplifying the clusters from *Tü6134* (16,176 bp), *deg* (12,076 bp) and *spi* (16,076 bp) for the integration into pUWL201PW_new_MCS. Amplicon size includes homologous regions added by Gibson primers.

The PCR products were purified by gel extraction (chapter 4.2.4) and integrated into NdeI digested (chapter 4.2.5) and dephosphorylated (chapter 4.2.6) expression vectors using Gibson assembly (chapter 4.2.7). The mixture was transformed into *E. coli* DH5α using heat shock transformation (chapter 4.2.10). Grown colonies were analyzed by colony PCR (chapter 4.2.13) according to possible integration of the clusters. Positive clones were grown overnight, plasmid DNA was extracted (chapter 4.2.4), analyzed using analytical restriction digest (chapter 4.2.5) and clones showing the correct restriction pattern were sent for sequencing (chapter 4.2.14). The results show (Table 6) that correct clones were detected for all three new clusters in the two expression vectors pSET152_ermE_new_MCS and pUWL201PW_new_MCS. Again, less than 30 clones were needed to be analyzed to identify one correct clone. These results give further support for the efficiency of the new long amplicon PCR and Gibson assembly-based method. In contrast to the *ika* cluster, *spi* and *Tü6314* clusters are 4 kb larger in size and could be equally easily captured. Furthermore, the GC content of the two larger clusters are comparably high as the *ika* cluster with 69% in the *spi* cluster and 72% in the *Tü6314* cluster. In contrast to these clusters, the *deg* cluster has only a GC-content of 42%, underlining that this method is suitable for high as well as low GC clusters.

Table 6. Result summary of expression constructs with new clusters. The table shows the chosen clones for the three different new clusters (*spi*, *Tü6314*, and *deg*) in the expression vectors pSET152_ermE_new MCS and pUWL201PW_new MCS with a reference to the respective figure in the appendix for result description.

Expression vector	Correct Clone	Reference to results
pSET152_ermE_new_MCS::*spi*	Clone 28	Figure A26
pUWL201PW_new_MCS::*spi*	Clone 1	Figure A27
pSET152_ermE_new_MCS::*Tü6314*	Clone 13	Figure A28
pUWL201PW_new MCS::*Tü6314*	Clone 6	Figure A29
pSET152_ermE_new_MCS::*deg*	Clone 7	Figure A30
pUWL201PW_new_MCS::*deg*	Clone 3	Figure A31

In parallel, a PoTeM cluster found in *S. roseosporus* (*ros*) was captured from gDNA using the same approach as described above. As the capturing from gDNA is less efficient, two different gDNA concentrations were tried in the temperature gradient PCR (Figure 40). As it is clearly detectable, the yields from amplifying whole clusters from gDNA are much lower. However, the samples obtained at the different temperatures were pooled and purified by gel extraction (chapter 4.2.4). These samples were used to conduct Gibson assembly with NdeI digested (chapter 4.2.5) and dephosphorylated (chapter 4.2.6) vectors.

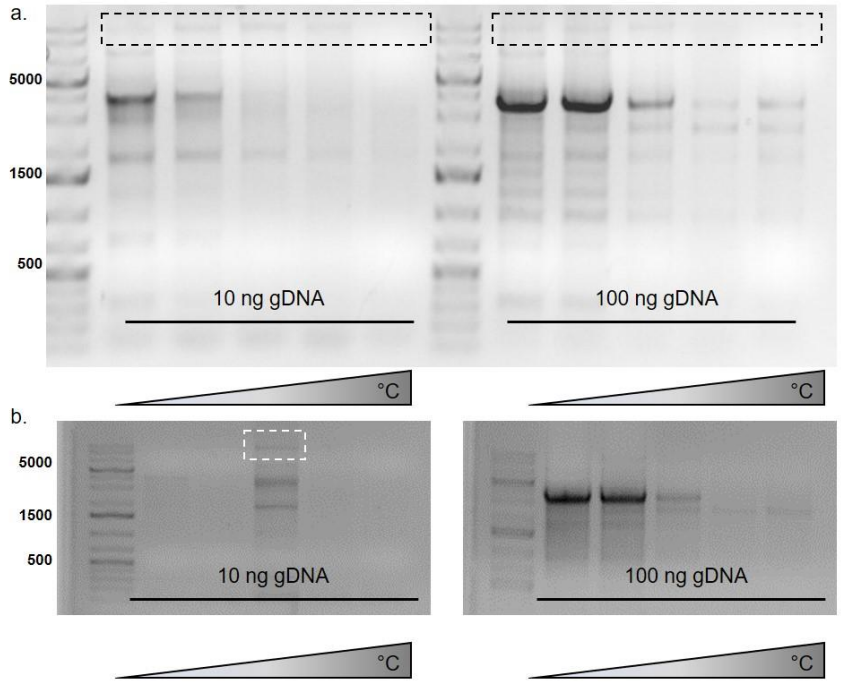


Figure 40. Temperature gradient PCR to amplify the *ros* cluster. a. Amplification for Gibson assembly into pSET152_ermE_new_MCS and b. amplification for Gibson assembly into pUWL201PW_new_MCS. Two different gDNA concentrations were tired with a temperature gradient from 50-70 °C to amplify the PoTeM cluster in one piece (16,214 bp).

After heat shock transformation (chapter 4.2.10), plasmid purification (chapter 4.2.4), analytical restriction digest (chapter 4.2.5) and sequencing (chapter 4.2.14) it was obvious that it is

possible to clone such a large cluster from gDNA using the Gibson assembly-based method (Figure A32). However, even after extended screening efforts, it was not possible to find a correct clone for the pUWL201PW_new_MCS::*ros* construct. As sequencing results show, there was always just a partial fragment of the cluster integrated (Figure 41).

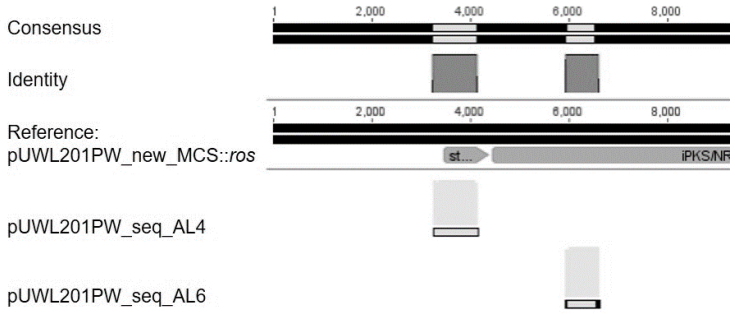


Figure 41. Sequencing results of pUWL201PW_new_MCS::*ros*. Primers binding on the vector backbone were used for sequencing, one contig aligned to the beginning of the cluster found in *S. roseosporus* and the other aligned inside the iPKS/NRPS of the *ros* cluster.

The size of the integrated fragment was about 3400 bp, which is about the size that can be found as a strong impurity of the long amplicon PCR (Figure 40b). This leads to the conclusion that even a gel extraction PCR cannot exclude impurities from DNA fragments of different sizes. Thereby a first drawback of the new long amplicon PCR and Gibson assembly-based gene cluster capturing method can be defined. The reliability on a pure PCR product is crucial for cloning success, with the fact that small DNA fragment are more likely amplified in PCRs, impurities can cause a high burden for successful cloning.

Downstream applications using the pSET152_ermE_new_MCS::*ros* construct by C. Greunke revealed a mutation in the first FAD dependent oxidoreductase of the *ros* PoTeM cluster (Figure 42). In this case a cytosine is mutated to a thymidine, which in this case leads to a silent mutation, however, this is a second drawback of the capturing method. High fidelity polymerases have such an efficient proofreading function that the error rate measured by NEB for the Q5 is assigned to 1 error in 1,000,000 amplified bases. (https://international.neb.com/tools-and-resources/feature-articles/polymerase-fidelity-what-is-it-and-what-does-it-mean-for-your-pcr). Accordingly, when amplifying a 10,000 bp fragment, there is statistically one mistake in every 100 amplicons. Furthermore, the high GC content of the DNA template increases the error rate, leading to a lower fidelity of the enzyme when the additional GC-buffer is required. Accordingly, high GC content DNA templates can increase the error rate of the novel Gibson assembly-based gene cluster capturing method.

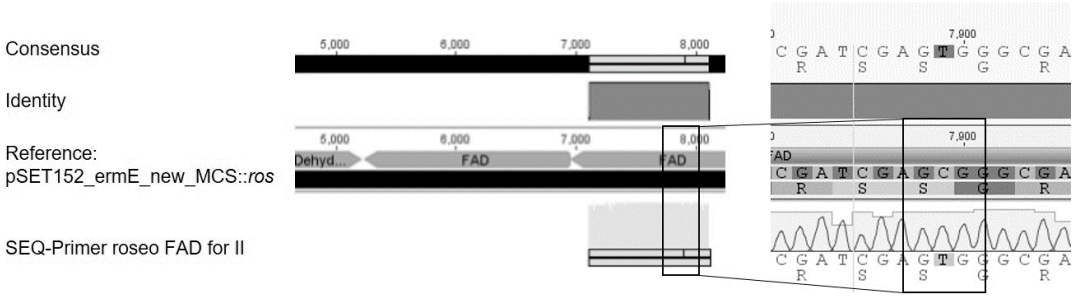


Figure 42. Sequencing results of the pSET152_ermE_new_MCS::*ros*. Results using the SEQ-Primer_roseo_FAD_for_II depict a mutation in the first FAD dependent oxidoreductase, a C is changed to a T, causing a silent mutation.

The efficiency of the long amplicon PCR and Gibson assembly-based method was published with its new assigned name 'Direct Pathway Cloning' (DiPaC).[130] This thesis has added a lot of understanding of practical procedures, improvements for primer designs and also the drawbacks to this method, thereby giving a great input for the results published as described above.

The last cluster for this study is found in *P. fermentans* JBW45. However, it has not been possible to amplify the whole gene cluster from the gDNA. Only the two accessory genes were captured in *E. coli* expression vectors. Accordingly, for the cloning of this expression construct, a different cloning approach was needed. In this case the feature of the common intermediate was used and an expression construct replacing the endogenous iPKS/NRPS by *ikaA* was designed. This was accomplished by integrating the acyltransferase at the EcoRI restriction site in the pSET152_ermE_new_MCS::*ikaA* expression construct. In a second step, the FAD dependent oxidoreductase was integrated behind *ikaA* at the NdeI restriction site. Both was accomplished by amplifying the two genes from pHis8-TEV vectors with primers suitable for Gibson assembly. After two cloning steps, the final expression construct, pSET152_ermE_new_MCS::*ikaA_JBW45* (referred to as fer_{MOD}) was accessible (Figure A33 and Figure A34).

The expression constructs containing the new PoTeM clusters *spi*, *Tü6314*, *deg* and fer_{MOD} were used for conjugation into the three *Streptomyces* hosts (chapter 4.2.12). The exconjugants were picked and cultured in ISP-4 for 7 d at 200 rpm and 28 °C, as this has proven to be the optimal culturing conditions for **39**. Cells and supernatant were separated using centrifugation and both were extracted with organic solvents (chapter 4.2.19). The extracts were dissolved in methanol and analyzed using LC-MS (chapter 4.2.20). However, none of the extracts showed any new peak that could correspond to a PoTeM mass (data not shown). Even the results after cultivating the exconjugants for 5 d did not show any difference to the wild type strains (data not shown).

The general absence of any new compound indicated a general mistake in the heterologous expression strategy. Comparing successful heterologous expression publications for PoTeMs[115-116] suggested that a promoter is needed in front of the iPKS/NRPS gene. As the BGC for **39** does not contain a PoTeM hydroxylase in the beginning, this was not any issue for the expression of this compound. In contrast all the clusters of the new compounds have a modifying gene in front of the iPKS/NRPS gene.

According to the findings, new cloning strategies for the *spi* and *deg* clusters were developed. As the *Tü6314* cluster was proposed to encode for **41**, an already known compound, this cluster was excluded from this extended cloning strategy. The clusters were divided into two parts, a front part, which was located before the iPKS/NRPS and a back part that started with the iPKS/NRPS gene and contained all the following modifying enzymes (Figure 43a). The front parts were amplified from the former expression constructs, containing the *ermE* promoter, and were integrated at the EcoRI restriction site of pSET152_ermE_new_MCS. Subsequently, the back part was also amplified from the former expression constructs and was integrated behind the *ermE* promoter of pSET152_ermE_new_MCS at the NdeI restriction site (Figure 43b and c).

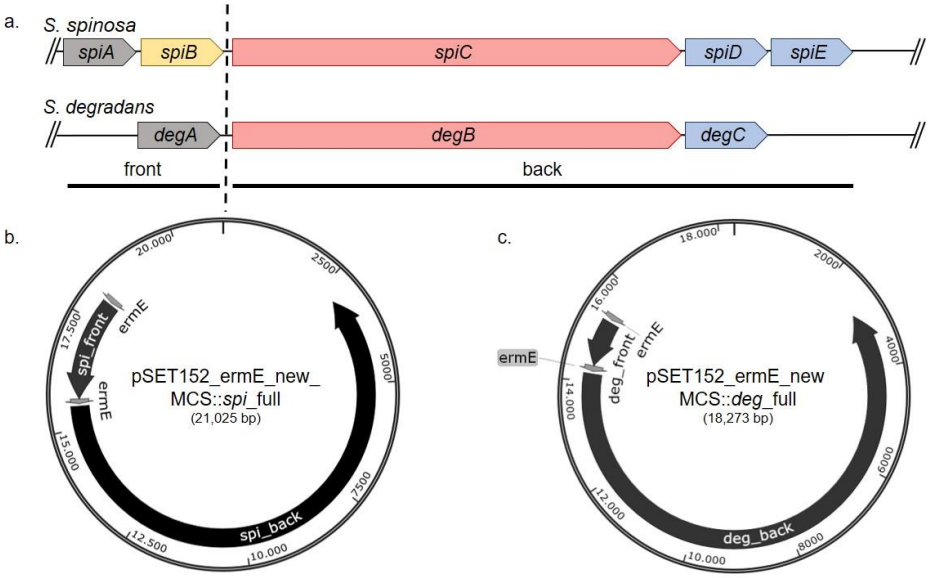


Figure 43. New expression constructs for novel PoTeM clusters. a. The *spi* cluster and *deg* cluster divided into a front and a back part. Each part was integrated subsequently into pSET152_ermE_new_MCS, final constructs are shown for the *spi* cluster (b.) and *deg* cluster (c.) Thereby two *ermE* promoters are present, with on in front of the whole cluster and one in front of the iPKS/NRPS.

As the newly published cloning strategy DiPaC was utilized for this cloning, primers suitable for Gibson assembly were designed and used for long amplicon PCR to amplify the cluster parts. The expression vector pSET152_ermE_new_MCS was digested at the EcoRI site (chapter 4.2.5) and dephosphorylated (chapter 4.2.6). The purified PCR products (Figure A35a

and c) were integrated into the vectors using Gibson assembly (chapter 4.2.7) and transformed into *E. coli* DH5α (chapter 4.2.10). Possible positive clones were selected using colony PCR (chapter 4.2.13) and analytical restriction digest (chapter 4.2.5), with one positive clone each that was verified using Sanger sequencing (chapter 4.2.14). The results are depicted in Figure A36 and Figure A39. The newly cloned constructs were digested with NdeI (chapter 4.2.5) to subsequently integrate the clusters back parts (Figure A35b and d) by Gibson assembly into the expression vectors to obtain the full cluster expression vectors. Again a verification was conducted using colony PCR (chapter 4.2.13), analytical restriction digest (chapter 4.2.5) and Sanger sequencing (chapter 4.2.14). As depicted in Figure A37 and Figure A40, it was possible to obtain colonies containing the expression vector with the presence of two *ermE* promoters. In addition to the full length constructs, a second set of expression plasmids was cloned only containing the back parts of the cluster, by omitting the first cloning step (Figure A38 and Figure A41). So far it has been published that accessory genes in front of the iPKS/NRPS are only needed to subsequently modify the PoTeM but are not crucial for backbone assembly. Therefore, it was hoped that these constructs could serve for a better understanding of the mechanisms underlying PoTeM biosynthesis.

Table 7. Result summary of expression constructs with new cluster structure. The table shows the chosen clones for the four different new expression constructs harboring the full or the back part of the cluster with an *ermE* promoter upstream of the iPKS/NRPS gene as well as in front of the whole cluster. References to the respective figure in the appendix for result description are assigned.

Expression vector	Correct clone	Reference to results
pSET152_ermE_new_MCS::*spi*_full	Clone 16	Figure A37
pSET152_ermE_new_MCS::*spi*_back	Clone 2	Figure A38
pSET152_ermE_new_MCS::*deg*_full	Clone 17	Figure A40
pSET152_ermE_new_MCS::*deg*_back	Clone 16	Figure A41

These four constructs (Table 7) were utilized for conjugation into the *S. albus* host strains (chapter 4.2.12). Exconjugants for each construct were picked and analyzed by colony PCR (chapter 4.2.13) for successful integration of the expression vector. Positive clones were cultured in CASO medium and a 50 ml YMG or ISP-4 main culture was inoculated 1:10 with a dense CASO preculture, incubated at 28 °C, 200 rpm for 7 d and extracted with organic solvents (chapter 4.2.19). The dried extracts were dissolved in methanol, filtered through 0.02 µm syringe filters and analyzed by LCQ-Fleet using the method shown in Table 35.

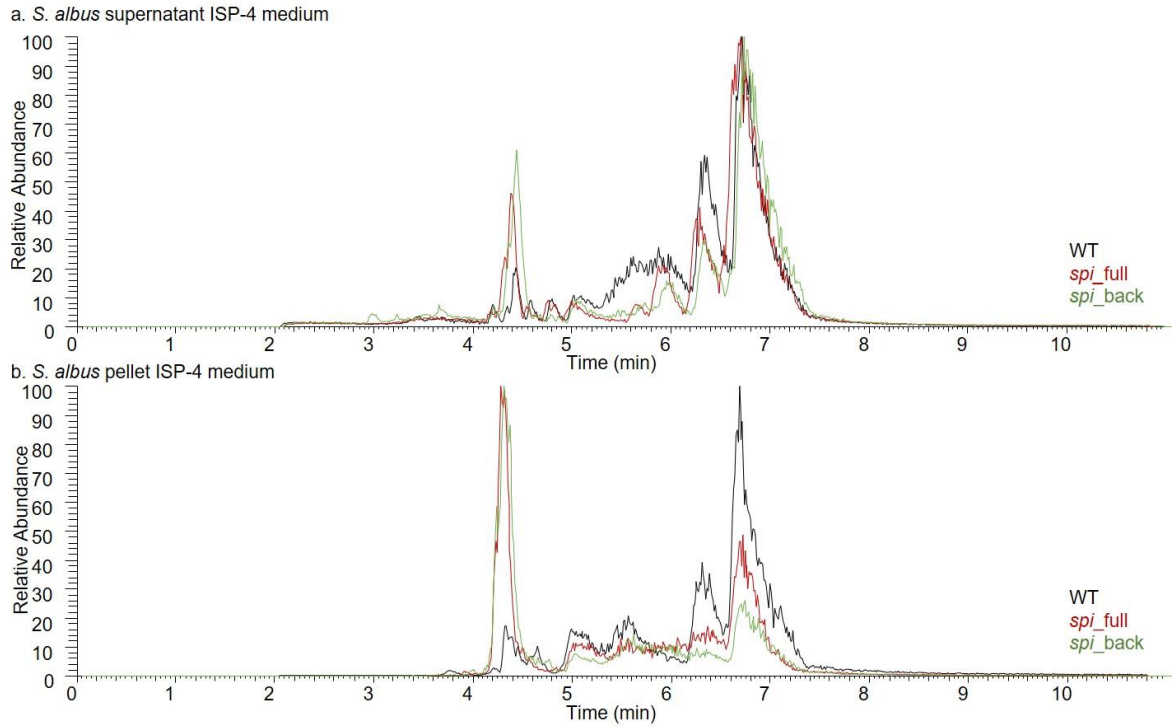


Figure 44. LC-MS results *S. albus* pSET152_ermE_new_MCS::*spi*_full and *::spi*_back. Depicted are the base-peak results measured from *S. albus* supernatant (a.) or pellet (b.) extracts containing either no (black) expression vector or the pSET152_ermE_new_MCS::*spi*_full (red) or pSET152_ermE_new_MCS::*spi*_back (green) expression vectors cultured in ISP-4 medium. Analytical method: LCQ-Fleet.Meth.

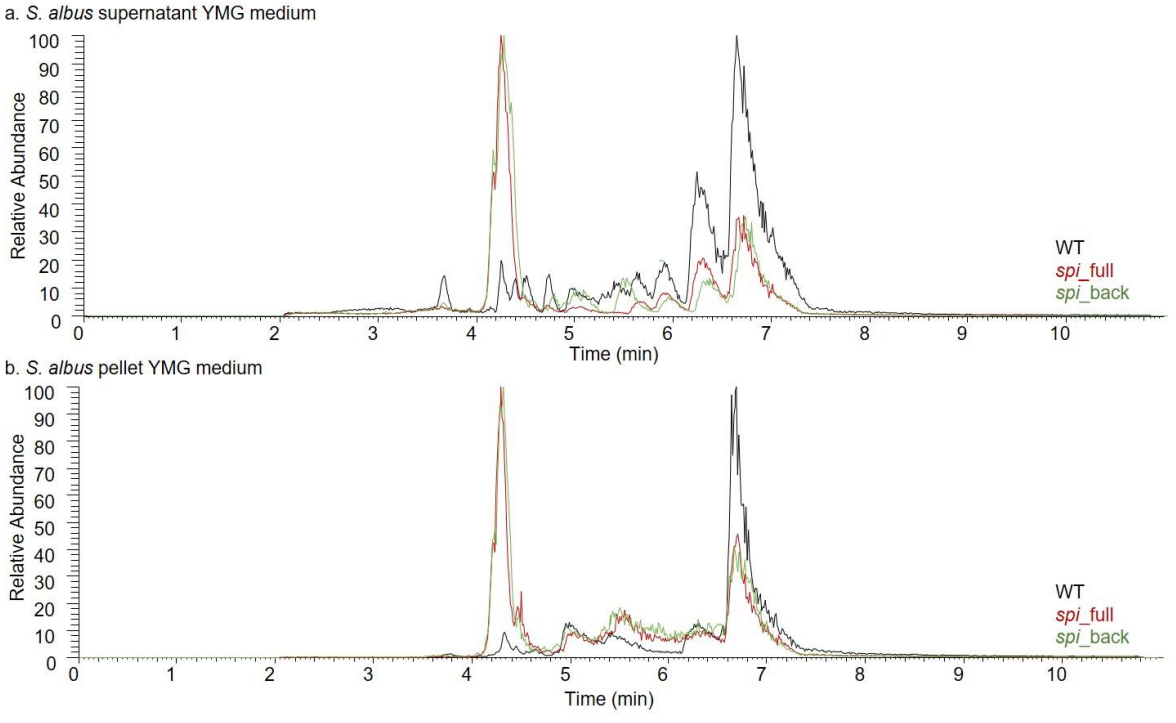


Figure 45. LC-MS results *S. albus* pSET152_ermE_new_MCS::*spi*_full and *::spi*_back. Depicted are the base-peak results measured from *S. albus* supernatant (a.) or pellet (b.) extracts containing either no (black) expression vector or the pSET152_ermE_new_MCS::*spi*_full (red) or pSET152_ermE_new_MCS::*spi*_back (green) expression vectors cultured in YMG medium. Analytical method: LCQ-Fleet.Meth.

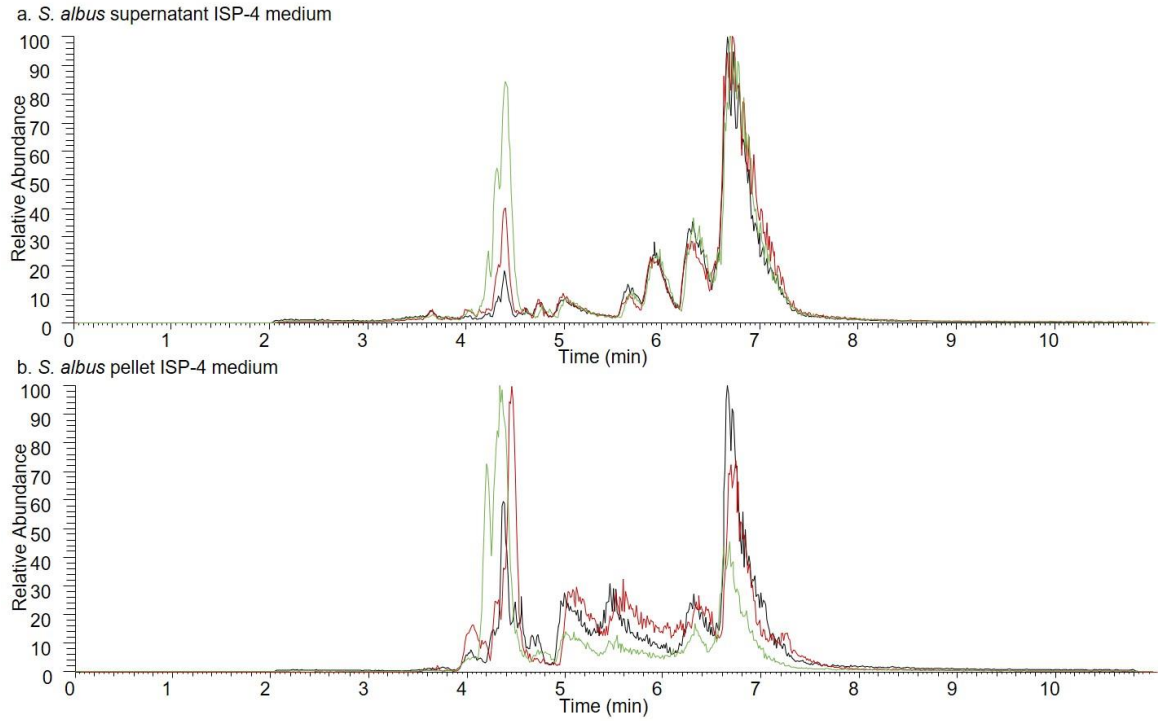


Figure 46. LC-MS results *S. albus* pSET152_ermE_new_MCS::*deg*_full and ::deg_back. Depicted are the base-peak results measured from *S. albus* supernatant (a.) or pellet (b.) extracts containing either no (black) expression vector or the pSET152_ermE_new_MCS::*deg*_full (red) or pSET152_ermE_new_MCS::*deg*_back (green) expression vectors cultured in ISP-4 medium. Analytical method: LCQ-Fleet.Meth.

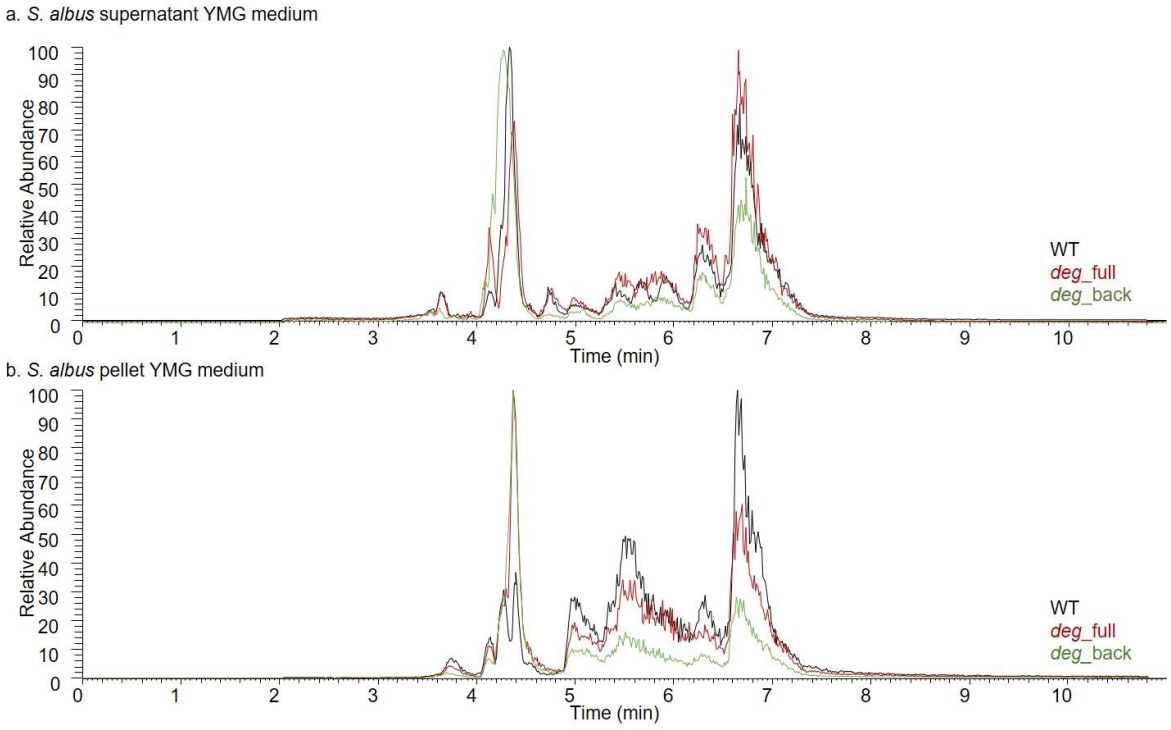


Figure 47. LC-MS results *S. albus* pSET152_ermE_new_MCS::*deg*_full and *::deg*_back. Depicted are the base-peak results measured from *S. albus* supernatant (a.) or pellet (b.) extracts containing either no (black) expression vector or the pSET152_ermE_new_MCS::*deg*_full (red) or pSET152_ermE_new_MCS::*deg*_back (green) expression vectors cultured in YMG medium. Analytical method: LCQ-Fleet.Meth.

The comparison of the extracts harboring the expression constructs and the wildtype (wt) strain clearly shows that the only differing peak is present in the *S. albus* pSET152_ermE_new_MCS::*spi*_full and pSET152_ermE_new_MCS::*spi*_back extracts at about 4.5 min, retention time. However, the most abundant mass in this peak is *m/z* = 912,78 $[M+H]^+$, a mass that does not fit to a potential PoTeM mass, which should be around *m/z* = 451 to 551 $[M+H]^+$. As the peak can also be detected in the wt extract, however in a much smaller quantity, it might be an endogenous molecule, whose expression is increased when the expression constructs are present in the cell.

As all the optimization trials did not result in a heterologous expression of the new PoTeMs, the same analysis as for the ikarugamycin intermediate production was conducted. Comparing the working ikarugamycin expression construct with the not working constructs of the new PoTeMs, the only difference can be found in the distance between RBS and start codon. Whereas there are 17 bases between RBS and start codon in the ikarugamycin expression construct (Figure 48a), there is a total of 41 bases between RBS and start codon in new expression plasmids containing the new MCS (Figure 48b), which might again be the determining factor for the negative expression results.

Figure 48. Comparison of expression constructs. a. Distance of RBS and i*kaA* start codon of pSET152_ermE::*ika* (17 bases), b. distance between RBS and iPKS/NRPS of pSET152_ermE_new_MCS::*spi*_full (41 bases).

According to these assumptions and considering the position of the promoter, three new constructs were designed for the novel PoTeM BGC *spi* and *deg*. Just as described before, one construct containing the *ermE* promoter in front of the whole cluster, one construct containing two *ermE* promoter, one in front of the whole cluster and one in front of the iPKS/NRPS, and one construct only containing the back part with an *ermE* promoter in front

of the iPKS/NRPS. As the pSET152_ermE expression vector was used for these designs, restriction sites EcoRI in front of the *ermE* promoter and StuI behind the *ermE* promoter were utilized. The Gibson primers were designed accordingly and clusters as well as cluster fragments were amplified (Figure A42) and successfully integrated into the expression vector (Table 8).

Table 8. Result summary of expression constructs with clusters in pSET152_ermE. The table shows the correct clones for the six new expression vectors containing the BGCs *spi* and *deg* in normal and reorganized structure. References to the respective figure in the appendix for result description are assigned.

Expression vector	Correct Clone	Reference to results
pSET152_ermE::*spi*_new	Clone 61	Figure A43
pSET152_ermE::*spi*_full_new	Clone 7	Figure A45
pSET152_ermE::*spi*_back_new	Clone 2	Figure A46
pSET152_ermE::*deg*_new	Clone 3	Figure A47
pSET152_ermE::*deg*_full_new	Clone 13	Figure A48
pSET152_ermE::*deg*_back_new	Clone 3	Figure A49

For heterologous expression of the novel compounds, the expression plasmids were integrated into the three *Streptomyces* host strains *S. albus, S. lividans* and *S. coelicolor* utilizing conjugation (chapter 4.2.12). Exconjugants were picked, plasmid integration was verified by colony PCR (chapter 4.2.13) and main cultures were inoculated. Cultures were incubated for 7 d at 28 °C and 200 rpm. Subsequently cell pellet and supernatant were separated by centrifugation and extracted with organic solvents (chapter 4.2.19). The extracts were dissolved in methanol, filtered through syringe filters and analyzed by LC-MS using the method shown in Table 36.

Supernatant and pellet extracts of all three *Streptomyces* strains harboring the pSET152_ermE::*spi*_new constructs as well as pSET152_ermE::*deg*_new constructs did not show any new peak that could indicate the production of a new PoTeM, as depicted for the *spi* cluster in Figure 49 and for the *deg* cluster in Figure 50. These results underline the assumption that a promoter is needed directly in front of the iPKS/NRPS, as these constructs only exhibit one *ermE* promoter in front of the whole cluster. Accordingly, one promoter is not sufficient to activate transcription of the large open reading frame (ORF) for the iPKS/NRPS. Therefore, the extractions of *Streptomyces* cultures containing the additional constructs pSET152_ermE::*spi*_full_new, pSET152_ermE::*spi*_back_new, pSET152_ermE::*deg*_full_new, and pSET152_ermE::*deg*_back_new were analyzed.

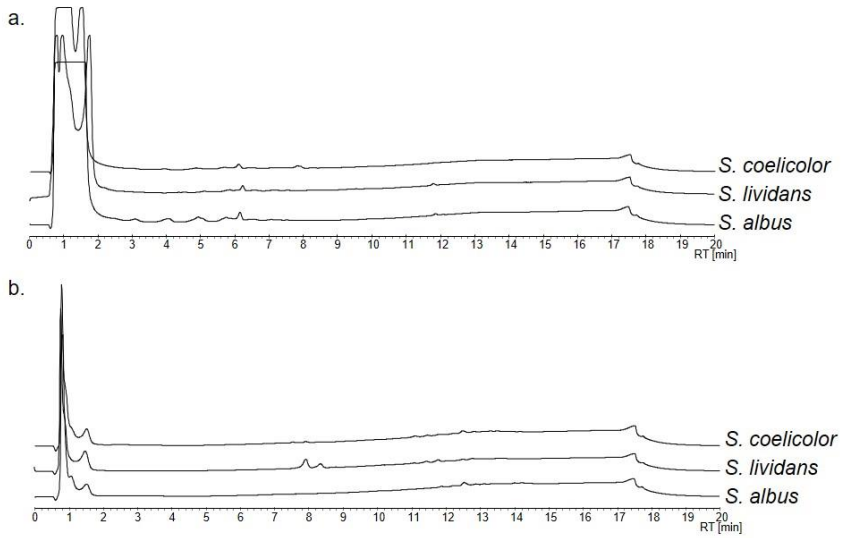


Figure 49. LC-MS analysis of extracts harboring pSET152_ermE::*spi*_new. HPLC chromatograms of a. supernatant and b. pellet extracts of expression hosts *S. albus*, *S. lividans, S. coelicolor* harboring pSET152_ermE::*spi*_new cultured in ISP-4 medium. Analytical method: Eurosphere_IkaFast_CG.Meth.

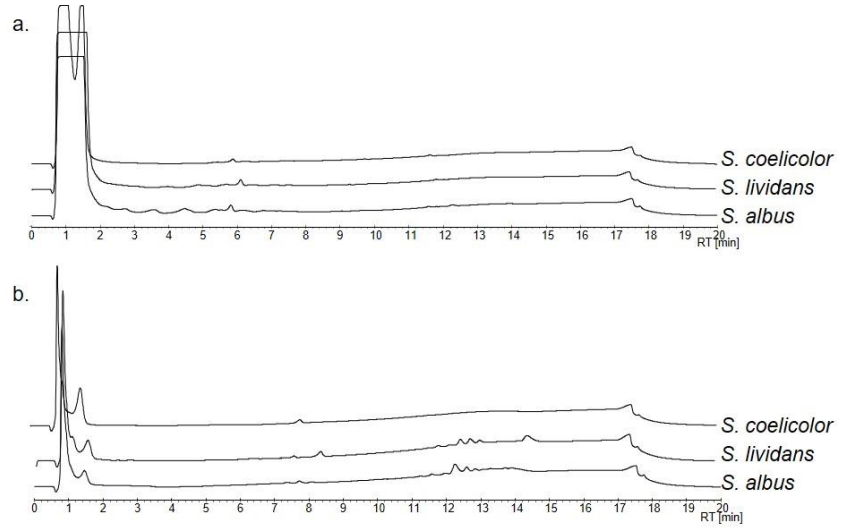


Figure 50. LC-MS analysis of extracts harboring pSET152_ermE::*deg*_new. HPLC chromatograms of a. supernatant and b. pellet extracts of expression hosts *S. albus*, *S. lividans, S. coelicolor* harboring pSET152_ermE::*deg*_new cultured in ISP-4 medium. Analytical method: Eurosphere_IkaFast_CG.Meth.

The first results were obtained for the extracts containing *Streptomyces* with pSET152_ermE::*spi*_back_new and pSET152_ermE::*deg*_back_new. As depicted in Figure 51 for the *spi* cluster and in Figure 52 for the *deg* cluster, there is almost no new peak detectable in the extracts from cultures containing the expression vector. However, there are two minor peaks arising at a retention time of about 6.25 and 6.75 min in the *S. albus* pellet extract containing pSET152_ermE::*spi*_back_new. When analyzing the mass of these two peaks, compound **spi-1** at 6.25 min *m/z* = 495.2 $[M+H]^+$ (Figure A107a) and compound **spi-2** at 6.75 min *m/z* = 479.2 $[M+H]^+$ can be detected (Figure A107b). Both masses would fit the typical PoTeM range, which lead to the assumption that a new PoTeM might have been

expressed in this experimental set up. To ensure that the new compound is not **39**, as the mass of *m/z* = 479.2 $[M+H]^+$ is identical, the UV spectrum of the two peaks was analyzed (Figure A107c and d), showing a great difference to the published spectrum for **39**, underlining the possibility that this expression strain produces a new PoTeM.

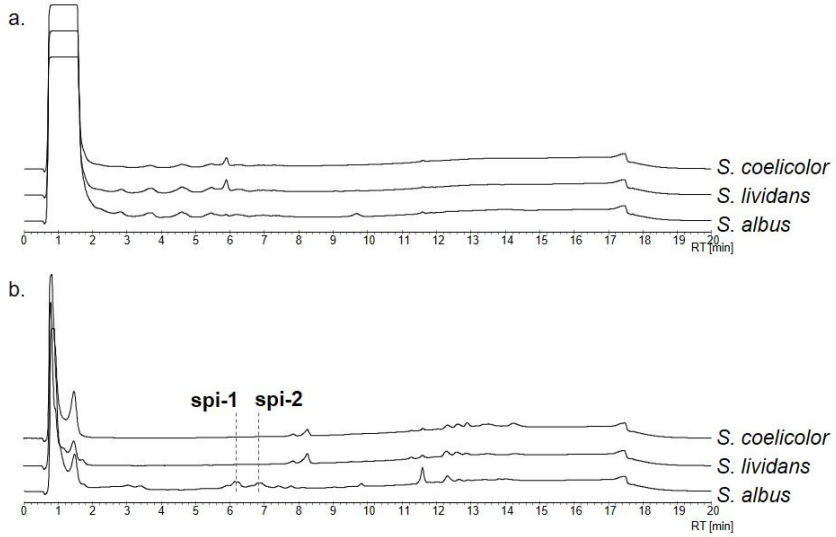


Figure 51. LC-MS analysis of extracts harboring pSET152_ermE::*spi*_back_new. HPLC chromatograms of a. supernatant and b. pellet extracts of expression hosts *S. albus*, *S. lividans, S. coelicolor* harboring pSET152_ermE::*spi*_back_new cultured in ISP-4 medium. Analytical method: Eurosphere_IkaFast_CG.Meth.

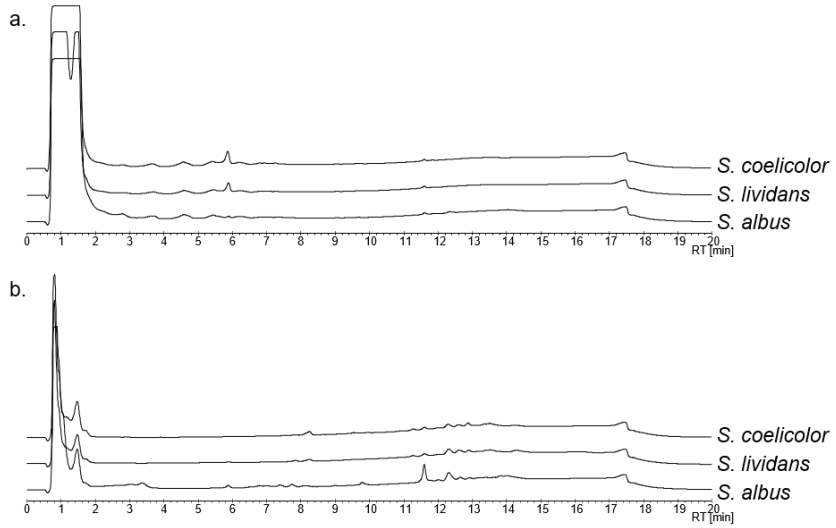


Figure 52. LC-MS analysis of extracts harboring pSET152_ermE::*deg*_back_new. HPLC chromatograms of a. supernatant and b. pellet extracts of expression hosts *S. albus*, *S. lividans, S. coelicolor* harboring pSET152_ermE::*deg*_back_new cultured in ISP-4 medium. Analytical method: Eurosphere_IkaFast_CG.Meth.

The last examined extracts were obtained from cultures containing the expression plasmids pSET152_ermE::*spi*_full_new and pSET152_ermE::*deg*_full_new. Similar to the previous results, there was no visible new peak in the extracts harboring the *deg* cluster (Figure 54). However, the same peaks visible in the extracts from *S. albus* containing only the back part of the *spi* cluster (pSET152_ermE::*spi*_back_new) are also visible in the same extracts

containing the full cluster (pSET152_ermE::*spi*_full_new), Figure 53, with the same masses detectable in the new peaks at a retention time about 6.25 min and 6.75 min, Figure A108.

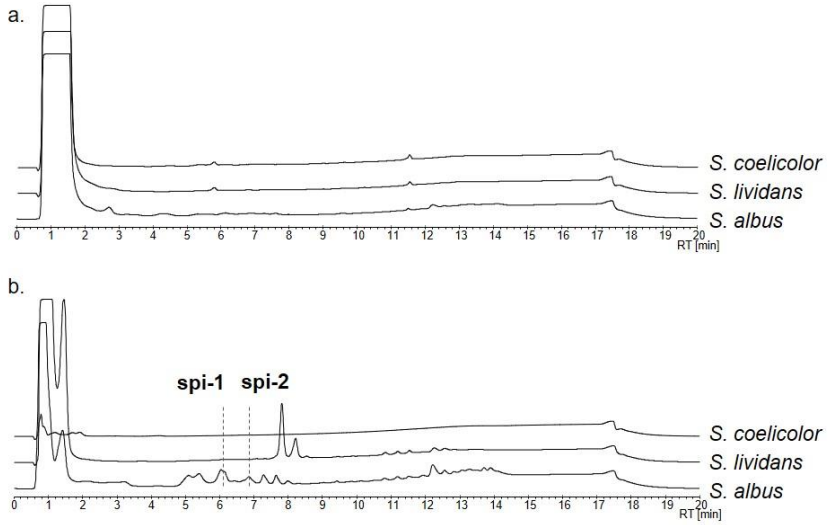


Figure 53. LC-MS analysis of extracts harboring pSET152_ermE::*spi*_full_new. HPLC chromatograms of a. supernatant and b. pellet extracts of expression hosts *S. albus*, *S. lividans, S. coelicolor* harboring pSET152_ermE::*spi*_full_new cultured in ISP-4 medium. Analytical method: Eurosphere_IkaFast_CG.Meth.

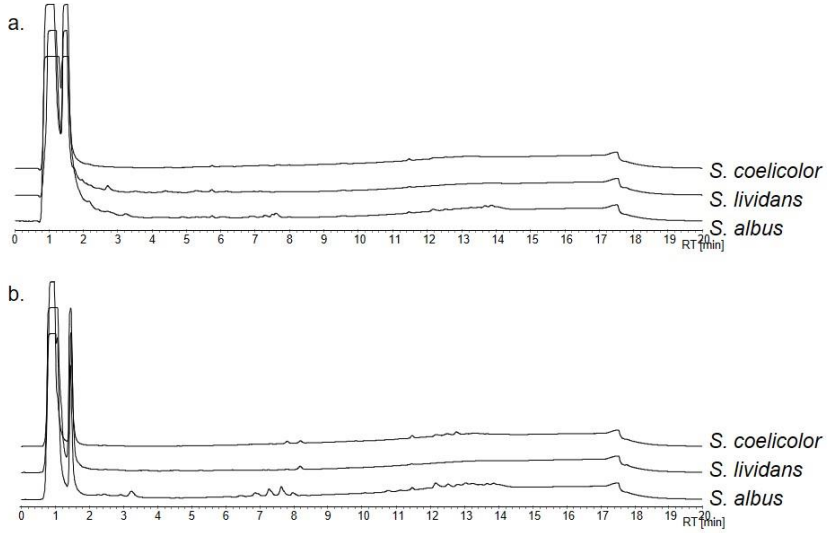


Figure 54. LC-MS analysis of extracts harboring pSET152_ermE::*deg*_full_new. HPLC chromatograms of a. supernatant and b. pellet extracts of expression hosts *S. albus*, *S. lividans, S. coelicolor* cultured in ISP-4 medium. Analytical method: Eurosphere_IkaFast_CG.Meth.

According to the reoccurrence of the new peak, the next aim was to maximize to amount of this new PoTeM that could be produced. To do so, more *S. albus* exconjugants were picked and analyzed for their PoTeM production titers. The cultivation time and medium were altered and amberlites (XAD7HP) were added to the cultures to maximize NP production, the most promising results are shown in Figure 55. Accordingly, an incubation period of 4 d in ISP-4 medium using the heterologous host strain *S. albus* containing the expression plasmid pSET152_ermE::*spi*_full_new enables the best production of novel compound. However, the

yields are much lower when compared to **39**. As a shorter incubation time causes a higher yield of NP, it can be assumed that the molecule is quite unstable and produced at an early time point of cultivation. Accordingly, further purification and structure elucidation trials need to consider the instability of the compound.

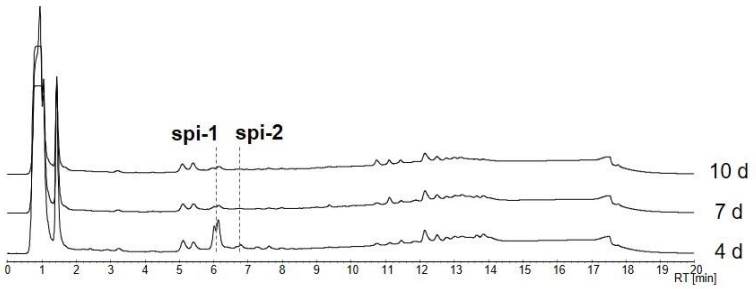


Figure 55. Optimizing heterologous expression of the *spi* cluster. HPLC chromatograms of three different cultivation times of pellet extracts from *S. albus* harboring pSET152_ermE::*spi*_full_new in ISP-4 at 28 °C and 200 rpm are shown. Analytical method: Eurosphere_IkaFast_CG.Meth.

For large scale purification, twenty 50 ml cultures were incubated under the optimized conditions. After cultivation, only the cell pellets were extracted. The crude extract was dissolved in the lowest possible amount of methanol and used for semi preparative HPLC purification. An optimal semi preparative HPLC method was established (Table 40) and used for purification of the crude extract, an exemplary HPLC chromatogram is depicted in Figure 56. The peak at 7.5-8.0 min retention time and the peak at 9.5-10.0 min retention time were collected separately. The peak with a retention time of 7.25 min until 8.25 min was collected (Figure 56, black square). 0.5 mg of a yellowish powder were dissolved in deuterated DMSO and measured using an NMR.

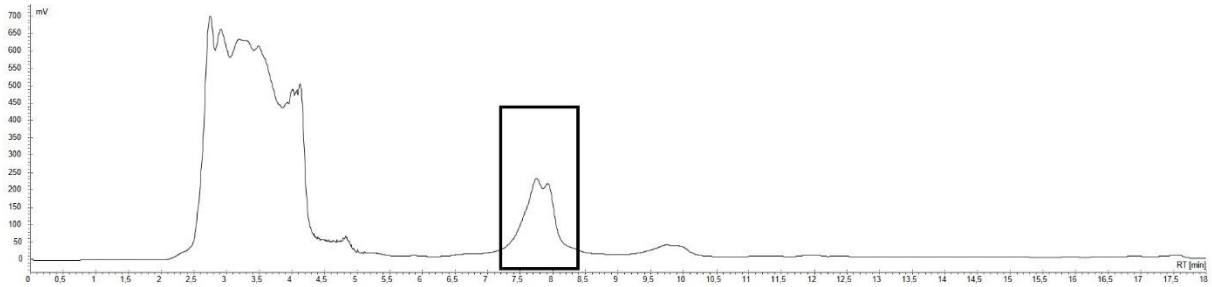


Figure 56. Semi preparative HPLC chromatogram of *spi* PoTeM purification. Analytical method: Eurosphere_IkaFAST_CG.Meth.

The preliminary NMR result in Figure A109 clearly shows a peak at 7.35 ppm, which is characteristic for the hydrogen atom at the tetramic acid of PoTeMs. Accordingly, it was assumed that the purified compound represents a novel PoTeM. However, the complexity of the compound and the low amount of purified product renders a structural elucidation with

these NMR data impossible. A higher yielding method for heterologously expressing the compound encoded in the *spi* BGC needs to be established for a complete structural elucidation.

2.5 Establishing a plug-and-play system for PoTeMs

As a consequence of the unsatisfactory heterologous expression of novel PoTeMs, a new approach to construct a universal expression plasmid for PoTeM expression was established. The crucial step to render heterologous expression of PoTeMs possible was identified to be the sufficient expression of the iPKS/NRPS (IkaA), and thereby the production of adequate amounts of precursor material (**53**) by the heterologous host. Considering the success of heterologously expressing **39** (chapter 2.1.3) and the failure in expressing the *deg* PoTeM cluster with the minor expression of the *spi* cluster, the iPKS/NRPS IkaA represents a suitable candidate for PoTeM heterologous expression. Furthermore, it was already proven that a IkaA recombinant protein expression system in *E. coli* BL21 is possible.[126]

According to previous results (chapter 2.3.1), a new plug-and-play expression system was constructed. Based on the pSET152_ermE expression vector, an expression plasmid was designed, utilizing the iPKS/NRPS gene, *ikaA*, downstream of the ermE promoter in addition to a subsequent set of different *Streptomyces* promoters downstream of *ikaA*. The presence of restriction sites up- and downstream of *ikaA* and downstream of the second promoter facilitates the flexible integration of accessory genes from different PoTeM clusters (Figure 57). This system enables a platform to efficiently clone PoTeM clusters with the iPKS/NRPS exchanged to *ikaA*, and furthermore enables the rearrangement and interchanging of genes from different clusters.

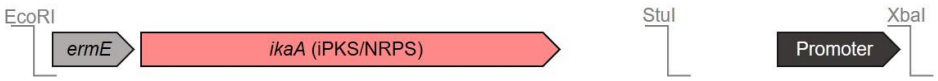


Figure 57. Schematic map of the basic plug-and-play expression vector. The new heterologous PoTeM expression vector is based on the pSET152_ermE expression vector, with *ikaA* and a second promoter present, additionally harboring the restriction sites for EcoRI, StuI, and XbaI for the incorporation of different PoTeM BGCs and accessory genes.

In addition to the main goal of expressing novel PoTeMs, this plug-and-play system was used for a comparative study of different promoters active in *Streptomyces*, as extensive studies on this topic were not present in literature.[131] Accordingly, nine different promoters from varying studies were chosen for this project (Table 42). The first chosen promoter was the *ermE* promoter already proven to be suitable for heterologous expression in *Streptomyces*

(chapter 2.1.3). The second pair of promoters, *gapdhP*(EL) and *rpsLP*(XC) were identified in a study that compared possible promoter regions upstream of house-keeping genes in *S. grieseus*.[132] The promoters in front of the *gapdh* operon (*gapdhP*) and an operon coding for the 30s ribosomal proteins S12 and S7 as well as elongation factors (*rpsLP*) were identified as strongly promoting regions. Subsequently an examination of the same regions from 18 different *Streptomyces* species to compare activity was conducted and proved *gapdhP* from *Eggerthella lenta* (EL) and *rpsLP* from Xylanimonas cellulosilytica (XC) to be the most activating promoters in their reporter assay.[132] The fourth promoter *kasOP** was generated by a study that used a random library of the promoter found upstream of *kasO*, a transcriptional regulator of a PKS, in *S. coelicolor*. While screening the engineered promoters, *kasOP** could be identified to upregulate expression of an reporter gene up to five times compared to *ermE*.[133] The fifth promoter, *SF14P*, was published as a promising promoter found in the genome of phage *I19* which could be isolated from a *Streptomyces* strain.[134] The last four promoters *P-2*, *P-6*, *P-15*, and *P-31* were discovered in a large study that compared non-transcribed intergenic regions in *S. albus* J1074 and could prove these four (and additional) regions to have a significant activating effect in the *xylE* reporter assay.[135]

To gain access to the different promoters with the two unique restriction sites, the *ermE* promoter was amplified with primers containing the restriction sites. However, all the other promoters were ordered as gene synthesis. To ensure activity, an RBS equal to the one used for the *ermE* promoter with a small spacer were added to the promoter sequences of *gapdhP*(EL), *rpsLP*(XC), *kasOP**, and *SF14P*; as the other promoters depict the whole region between two ORFs in *S. albus* J1074 it was assumed that an RBS was already present in these sequences (Table 42). Furthermore, the two restriction sites were added to the sequence for ordering.

To establish the basic vectors for the plug-and-play system, the *ermE* promoter was amplified using Q5-PCR (chapter 4.2.13, for result Figure A50a), the PCR product of *ermE* and the amplified vectors containing the eight additional synthesized promoters were digested with StuI and XbaI (chapter 4.2.5). The promoters were gel extracted (chapter 4.2.4) and ligated (chapter 4.2.7) to the digested (chapter 4.2.5) and dephosphorylated (see chapter 4.2.6) pSET152_ermE::*ikaA*. The ligation mixture was inactivated and transformed into chemically competent *E. coli* DH5α (chapter 4.2.10). Clones were picked and screened using colony PCR (chapter 4.2.13), DNA from possible positive colonies was extracted (chapter 4.2.4) and analyzed using analytical restriction digestion (chapter 4.2.5) subsequent verification of successful cloning was done by Sanger sequencing (chapter 4.2.14). The successful cloning of the nine constructs builds the basis of the new heterologous PoTeM expression system based on the iPKS/NRPS enzyme IkaA.

Table 9. Result summary of expression constructs pSET152_ermE::*ikaA*_Promoter. The table shows the correct clones for the nine new expression vectors for the plug-and-play system. References to the respective figure in the appendix for result description are assigned.

Expression construct	Correct Clone	Reference to results
pSET152_ermE::*ikaA*_ermE	Clone 1	Figure A50
pSET152_ermE::*ikaA*_gapdhP(EL)	Clone 1	Figure A51
pSET152_ermE::*ikaA*_rpsLP(XC)	Clone 1	Figure A52
pSET152_ermE::*ikaA*_kasOP*	Clone 2	Figure A53
pSET152_ermE::*ikaA*_SF14P	Clone 1	Figure A54
pSET152_ermE::*ikaA*_P-2	Clone 1	Figure A55
pSET152_ermE::*ikaA*_P-6	Clone 1	Figure A56
pSET152_ermE::*ikaA*_P-15	Clone 1	Figure A57
pSET152_ermE::*ikaA*_P-31	Clone 2	Figure A58

2.5.1 Establishing the plug-and-play system for derivatives of **39**

In a first step, the potential of the new heterologous plug-and-play system should be examined by expressing derivatives of **39**. Therefore, the two additional genes of the *ika* cluster (*ikaB* and *ikaC*) were integrated at the StuI restriction site upstream of the second promoter. Subsequently three PoTeM modifying genes were chosen to be integrated downstream of the second promoter at the XbaI restriction site to generate expression vectors for the heterologous expression of **39** derivatives (Figure 58).

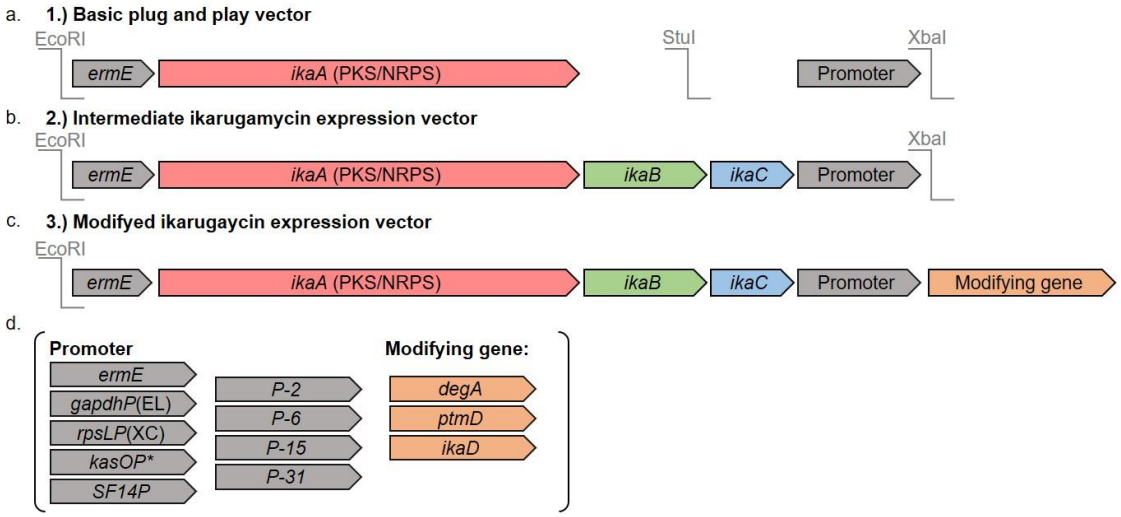
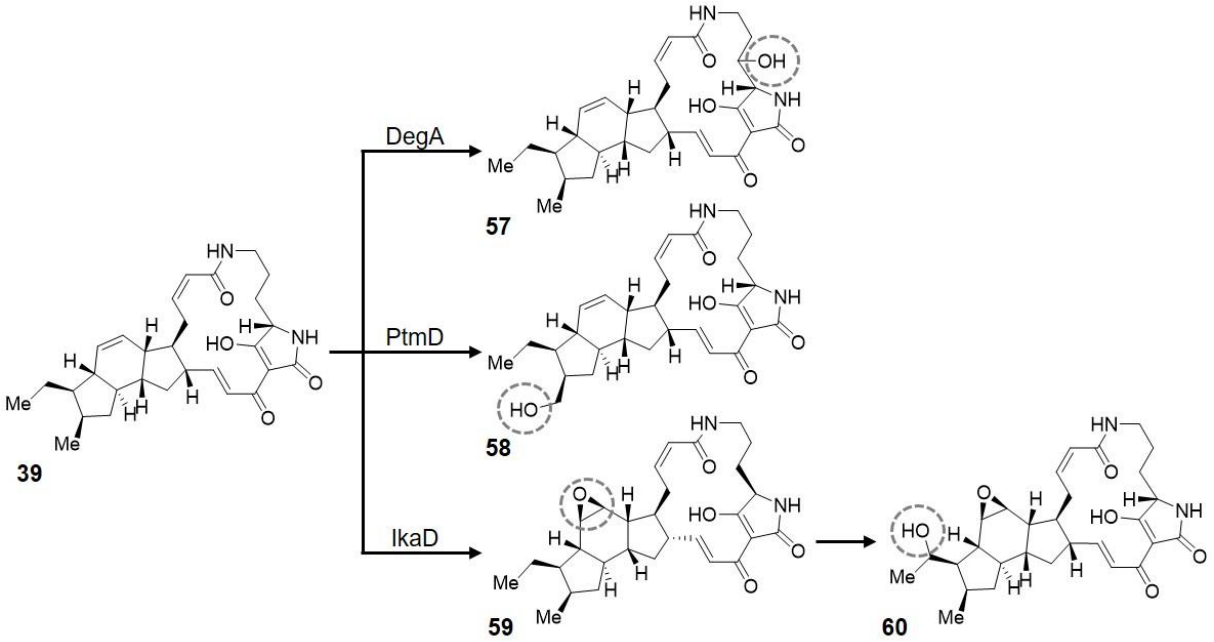


Figure 58. Plug-and-play strategy for heterologous production of derivatives of **39** a. Basic plug-and-play vector (pSET152_ermE::*ikaA*_Promoter). b. Intermediate **39** expression vector (pSET152_ermE::*ikaABC*_Promoter) c. Expression vector for derivatives of **39** based on the plug-and-play vector (pSET152_ermE::*ikaABC*_Promoter_mod.gene). d. Promoters and modifying genes used for the derivatization.

The first chosen modifying enzyme (PoTeM hydroxylase (DegA)) originates form the *deg* cluster. Enzymes of this class were already postulated to hydroxylate PoTeMs at C-25 (Figure 8),[108] a modification frequently found amongst different PoTeMs. Furthermore, our group could already prove a PoTeM hydroxylase homologous to DegA to accept **39** as a substrate for the conversion into **60**.[118] However, earlier heterologous expression or *in vitro* biosynthesis trials have not achieved sufficient yields for downstream applications. The second modifying enzyme was published in the newly discovered pactamide (**61**) BGC and was postulated to hydroxylate at C-29 (**62**), a new modification found in pactamide B and D.[136] The last modifying enzyme, IkaD, is present in close proximity to the *ika* BGC. However, in the natural producer it is mapped on a different ORF with a reverse orientation. The enzyme IkaD belongs to a new group of P450 monooxygenases and was be proven to modify **39** by introduction of an epoxide at the C7/C8-double bond to result in the ikarugamycin epoxide (**63**) and furthermore by hydroxylation at C-30 (**64**).[137] As especially the epoxide is interesting for downstream application using further chemical modification, such as introduction of an alkyne to facilitate click chemistry in target cells, purifying this compound could improve the knowledge about **39**'s cellular target and its mode of action. Possible modification products are depicted in Scheme 9.



Scheme 9. Postulated products of heterologously expressed **39** derivatives. Possible modification of **39** by DegA to be hydroxylated at C-25 (**60**), by PtmD to be hydroxylated at C-29 (**62**) and by IkaD to be hydroxylated at C-30 and introduction of an epoxide at C7-C8 (**63**, **64**). Modifications are emphasized by dotted circles.

The first step to get access to the modified PoTeM products was cloning the different expression constructs. The cloning was started with the introduction of the missing *ika* BGC part, *ikaBC*, by PCR amplification of the BGC fragment for each plug-and-play vector (chapter

4.2.13, PCR results Figure A59) and integration into the StuI digested (chapter 4.2.5) and dephosphorylated (chapter 4.2.6) basic plug-and-play vectors using Gibson assembly (see chapter 4.2.7). The verification of positive clones was conducted by colony PCR (chapter 4.2.13), analytical restriction digest (chapter 4.2.5) and finally Sanger sequencing of the plasmids (chapter 4.2.14). The cloning results of the first cloning step can be seen in Table 10.

Table 10. Result summary of expression constructs pSET152_ermE::*ikaABC*_Promoter. The Table shows the correct clones for the nine new ikarugamycin expression constructs based on the new plug-and-play system. References to the respective figure in the appendix for result description are assigned.

Expression construct	Correct clone	Reference to results
pSET152_ermE::*ikaABC*_ermE	Clone 16	Figure A60
pSET152_ermE::*ikaABC*_gapdhP(EL)	Clone 2	Figure A61
pSET152_ermE::*ikaABC*_rpsLP(XC)	Clone 3	Figure A62
pSET152_ermE::*ikaABC*_kasOP*	Clone 3	Figure A63
pSET152_ermE::*ikaABC*_SF14P	Clone 1	Figure A64
pSET152_ermE::*ikaABC*_P-2	Clone 1	Figure A65
pSET152_ermE::*ikaABC*_P-6	Clone 20	Figure A66
pSET152_ermE::*ikaABC*_P-15	Clone 1	Figure A67
pSET152_ermE::*ikaABC*_P-31	Clone 1	Figure A68

The cloning of the intermediate constructs was successful for all the selected promoters. However, when culturing *E. coli* cells harboring those expression constructs, it was observed that the cells grew significantly slower, which was probably due to the complexity of the expression plasmids with two different constitutive promoters that possibly interfered with the cellular protein biosynthesis. A second point to consider with the novel **39** expression plasmids was the arrangement of *ikaA* and *ikaB*. In the natural *ika* BGC the sequence of *ikaA* and *ikaB* overlap at the ends, whereas due to the plug-and-play systems composition, these two genes do not overlap anymore in the newly cloned vectors (Figure 59). To ensure that this fact would not decrease **39** production, one construct (pSET152_ermE::*ikaABC*_SF14P) was conjugated into *S. albus* DSM40313 (chapter 4.2.12), and a culture was extracted to check for **39** production. As there was a good production of **39** visible (Figure A106), the integration of modifying enzymes was started.

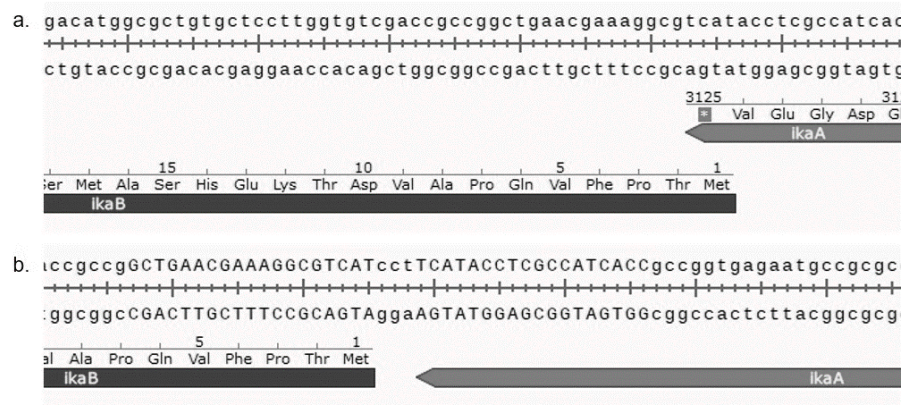


Figure 59. ORFs of *ikaA* and *ikaB* in original and rearranged cluster. a. Structure of the original **39** BGC with *ikaA* and *ikaB* overlapping. b. Rearranged position of *ikaA* and *ikaB* in the plug-and-play vectors with no overlap of *ikaA* and *ikaB*.

The cloning strategy for integrating the modifying enzymes was based on restriction cloning to avoid the need of several primers generating overhangs to the different promoters as required for Gibson assembly-based cloning. *DegA* was amplified from *S. degradans* gDNA, *ikaD* was amplified from *Streptomyces sp*. Tü6239 gDNA and *ptmD* was ordered as a synthetic gene and was amplified from the obtained synthetic plasmid (Figure A69). The PCR products were digested using XbaI (chapter 4.2.5) and ligated (chapter 4.2.7) into the XbaI digested (chapter 4.2.5) and dephosphorylated (chapter 4.2.6) plug-and-play vectors containing *ikaABC* and the respective promoter. Cloning success was verified using colony PCR (chapter 4.2.13) and analytical restriction digest (chapter 4.2.5) (data not shown) and finally proven by Sanger sequencing (chapter 4.2.14). Detailed results for the cloning are depicted in Table 11. It is obvious that correct clones for most of the constructs were found in less than 10 screened colonies. However, it was not possible to clone the expression vector that contain two *ermE* promoters. This is most likely because the XbaI restriction enzyme is sensitive to Dam methylation and can thereby not cut the vector produced in *E. coli* DH5α. A trial to passage the intermediate construct through *E. coli* ET12567, a strain that does not perform Dam methylation, and subsequent cloning also failed in obtaining the correct expression plasmid. Accordingly, pSET152_ermE::*ikaA*_ermE was excluded from the list of basic plug-and-play vectors. The second promoter that showed some difficulties during cloning was the *P-6* promoter. As shown in Table 11, it took screening of 83 colonies to identify a correct clone. This is caused by complexity of the *P-6* promoter with a GC content of 73%, including GC-islands of up to 90% GC content, almost impossible to be amplified by any polymerase. Furthermore, it includes secondary structures such as hairpins, that most likely interfere when the linearized plasmid should be ligated with the modifying enzyme. This represent a general cloning obstacle that should be considered when using the basic plug-and-play vector pSET152_ermE::*ikaA*_P-6.

All 24 expression plasmids were conjugated into *S. albus* DSM40313, *S. lividans* TK24 and *S. coelicolor* M1154 (chapter 4.2.12), successful conjugation was PCR verified (chapter 4.2.13) and subsequently 3 clones of each expression construct were cultivated at 28 °C, 200 rpm for 7 d. The cultures were extracted (chapter 4.2.19) and analyzed by LC-MS (chapter 4.2.20) using the method shown in Table 37.

Table 11. Result summary of plug-and-play vectors containing modifying enzymes. The table shows the correct clones for the new modified ikarugamycin expression constructs based on the new plug-and-play system. References to the respective figure in the appendix for result description are assigned.

Expression construct	Correct clone	Reference to results
	degA	
pSET152_ermE::*ikaABC*_ermE_*degA*	-	-
pSET152_ermE::*ikaABC*_gapdhP(EL)_*degA*	Clone 16	Figure A70
pSET152_ermE::*ikaABC*_rpsLP(XC)_*degA*	Clone 2	Figure A71
pSET152_ermE::*ikaABC*_kasOP*_*degA*	Clone 4	Figure A72
pSET152_ermE::*ikaABC*_SF14P_*degA*	Clone 11	Figure A73
pSET152_ermE::*ikaABC*_P-2_*degA*	Clone 10	Figure A74
pSET152_ermE::*ikaABC*_P-6_*degA*	Clone 2	Figure A75
pSET152_ermE::*ikaABC*_P-15_*degA*	Clone 1	Figure A76
pSET152_ermE::*ikaABC*_P-31_*degA*	Clone 6	Figure A77
	ikaD	
pSET152_ermE::*ikaABC*_ermE_*ikaD*	-	-
pSET152_ermE::*ikaABC*_gapdhP(EL)_*ikaD*	Clone 9	Figure A78
pSET152_ermE::*ikaABC*_rpsLP(XC)_*ikaD*	Clone 11	Figure A79
pSET152_ermE::*ikaABC*_kasOP*_*ikaD*	Clone 10	Figure A80
pSET152_ermE::*ikaABC*_SF14P_*ikaD*	Clone 8	Figure A81
pSET152_ermE::*ikaABC*_P-2_*ikaD*	Clone 3	Figure A82
pSET152_ermE::*ikaABC*_P-6_*ikaD*	Clone 83	Figure A83
pSET152_ermE::*ikaABC*_P-15_*ikaD*	Clone 18	Figure A84
pSET152_ermE::*ikaABC*_P-31_*ikaD*	Clone 3	Figure A85
	ptmD	
pSET152_ermE::*ikaABC*_ermE_*ptmD*	-	-
pSET152_ermE::*ikaABC*_gapdhP(EL)_*ptmD*	Clone 7	Figure A86
pSET152_ermE::*ikaABC*_rpsLP(XC)_*ptmD*	Clone 7	Figure A87
pSET152_ermE::*ikaABC*_kasOP*_*ptmD*	Clone 3	Figure A88
pSET152_ermE::*ikaABC*_SF14P_*ptmD*	Clone 4	Figure A89
pSET152_ermE::*ikaABC*_P-2_*ptmD*	Clone 7	Figure A90
pSET152_ermE::*ikaABC*_P-6_*ptmD*	Clone 16	Figure A91
pSET152_ermE::*ikaABC*_P-15_*ptmD*	Clone 5	Figure A92
pSET152_ermE::*ikaABC*_P-31_*ptmD*	Clone 1	Figure A93

The HPLC-MS analysis of cultures containing the modifying enzyme DegA showed a partial conversion of **39** to a new peak, assumed to be **60** (Figure 60). The peak at the retention time of 15.5 min was proved to be **39** by the respective mass and UV spectrum. The new peak at 14.5 min showed a mass of *m/z* = 495.1 $[M+H]^+$, corresponding to **60**, additionally the UV spectra was identical to the ikarugamycin spectra underlying that the new peak is a derivative of **39**.

Figure 60 depicts an exemplary HPLC chromatogram of a culture containing a plug-and-play expression plasmid encoding the PoTeM hydroxylase modifying enzyme DegA. For comparison, the total absorption value of each extraction was determined and is represented in Figure 61 as bar diagrams. The results of three different extracts from cultures of each of the eight expression plasmids in the three different heterologous host are matched. Already evident in the HPLC chromatograms, the conversion is insufficient, with maximal absorptions for **60** around 120 mV, a value about ten times less than measured for **39** (chapter 2.1.3).

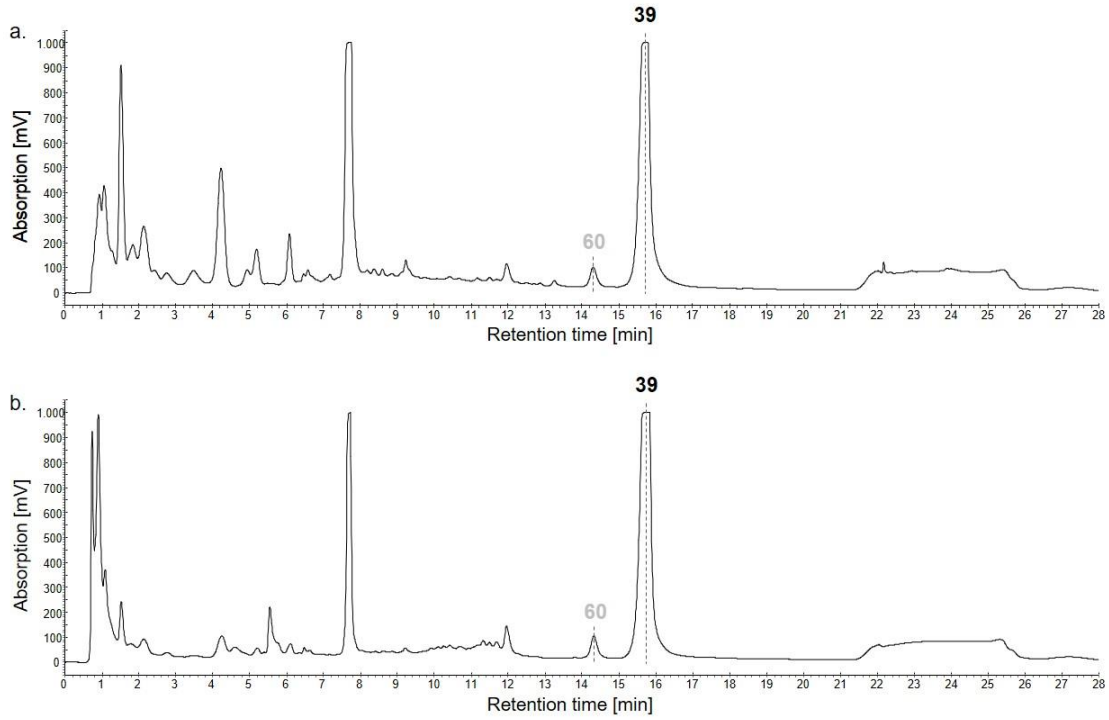


Figure 60. HPLC result of extracts containing the modifying enzyme DegA. Exemplary HPLC chromatogram of *S. albus* supernatant (a.) and pellet (b.) extracts containing the expression construct pSET152_ermE::*ikaABC*_rpsLC(XC)_*degA* are shown. The predicted **39** derivative is depicted in light grey. Analytical method: Eurosphere_Ika_lange Methode_AG.Meth.

When comparing the different heterologous hosts, the product **60** is primarily detected in *S. albus* (Figure 61a and b). As discussed earlier, *S. albus* has its own PoTeM BGC, thereby already supplying the precursor needed for PoTeM biosynthesis. Furthermore, this leads to a higher amount of substrate to be converted by PoTeM hydroxylase to **60**. Moreover, the

culturing conditions used for this comparative study were optimized for **39** production in *S. albus*. However, growing characteristics of the three strains differ drastically, which can be already observed by the morphology of the cultures after 7 d incubation. When harvesting the expression cultures, *S. albus* cultures were dense and cloudy, whereas *S. lividans* and *S. coelicolor* cultures consisted of cellular spheres and clear culturing broth. The dry cell weight of *S. albus* cultures usually ranged around less than 2 g, whereas *S. lividans* and *S. coelicolor* cultures weighed around 2-5 g. The spherical cultures allow cell growth, however, only the outer layer has an optimal nutrition supply, which is crucial for NP production. Culturing *S. lividans* and *S. coelicolor* with wire spirals is one possibility to inhibit sphere formation and might induce NP production. Nevertheless, in this study the conditions were predetermined as experimentally descried in chapter 2.1.3, additional comparison of culturing condition would have exceeded to scope of this study.

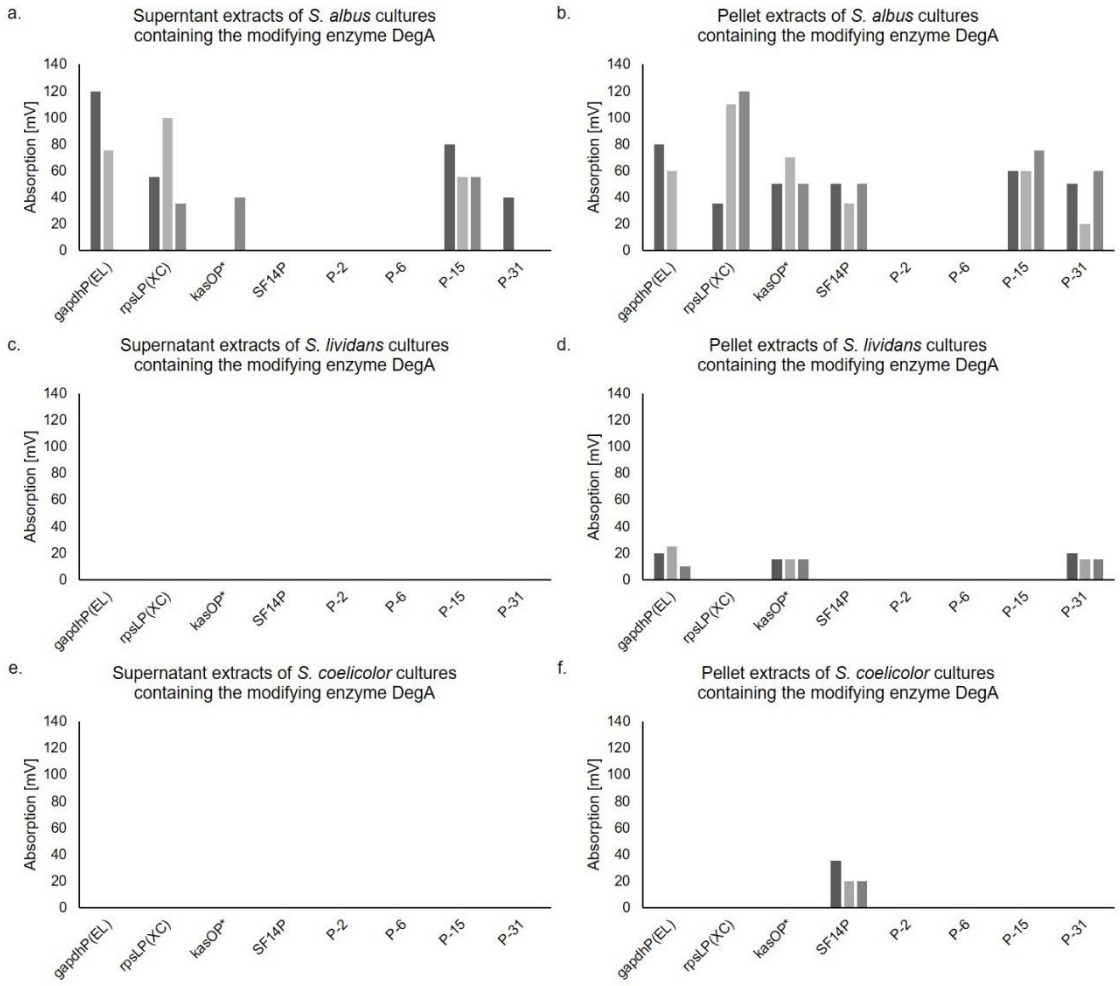


Figure 61. Bar diagrams of HPLC absorption for DegA containing expression cultures. Depicted are the absolute absorption values of the newly observed peak in the heterologous expression extracts containing the modifying enzyme DegA for *S. albus* supernatant (a.) and pellet (b.), *S. lividans* supernatant (c.) and pellet (d.) and *S. coelicolor* cultures supernatant (e.) and pellet (f.). The mass detected for the new product peak has an *m/z* = 495.1 $[M+H]^+$, the mass predicted for **60**. The three different clones extracted are represented in different shades of gray.

Comparison of the different promoters showed a stable production of **60** in cultures of *S. albus* containing the expression plasmids with one of the six promoters *gapdhP*(EL), *rspLP*(XC), *kasOP**, *SF14P*, *P-15*, and *P-31*, with marginally higher **39** production in cultures containing the *gapdhP*(EL) and *rpsLP*(XC) promoters.

From the extracts of cultures containing the promoters *P-2* and *P-6*, **60** could not be detected. However, in these cultures **39** could also not be detected, which might indicate a systematic error during the heterologous expression. As the absence of both NPs can be seen in all the extracts of the different strains containing the expression plasmids pSET152_ermE::*ikaABC*_P-2_*degA* and pSET152_ermE::*ikaABC*_P-6_*degA*, mutations in the expression plasmids or failed conjugations might be responsible. As later experiments of culture extracts containing pSET152_ermE::*ikaABC*_P-2 and pSET152_ermE::*ikaABC*_P-6 expression vectors with different modifying enzymes showed production of at least **39**, a mutation in the intermediate plasmid responsible for disrupting biosynthesis can be excluded. As for each promoter the last cloning step was proceeded with the same intermediate plasmid, at least an expression of **39** should be detectable in the extracts. Accordingly, a failed conjugation, resulting in *Streptomyces* capable of growing in apramycin supplemented media, is most likely the reason for the absence of NPs. The difficulties during conjugation, as the complex expression vectors containing a complete BGC with two promoters, lead to several passages of the *E. coli* ET12567 pUZ8002 in media containing apramycin, which possibly induced a resistance of these crucial cells for conjugation. This could result in rearrangements of the expression vectors that might have been conjugated into the *Streptomyces* strains, making the unsusceptible for apramycin. As a result, these two promoters cannot be accounted with the modifying enzyme DegA in comparison to the other promoters.

Only a minor difference in promoter activity can be detected using the conversion of **39** to **60** as a parameter. The most promising culture *S. albus* pSET152_ermE::*ikaABC*_rpsLP(XC)_*degA* clone 2 was chosen for large scale expression to purify **60**. The large-scale purification was started with the inoculated of a preculture with frozen spores, obtained during the comparison study of the different promoters. This preculture was slowly scaled up to receive enough volume to inoculate twenty 50 ml IPS-4 cultures. These cultures were incubated, merged, and extracted as described above (chapter 4.2.1, chapter 4.2.19). However, no natural product was obtained from precultures inoculated with frozen spores. According to the previous findings, a novel conjugation of the plasmid was conducted into *S. albus* DSM40313 (chapter 4.2.12). Without previous screening for NP production, a large-scale fermentation was started,1 L ISP-4 split in twenty 50 ml cultures. The small cultures were combined and extracted as described in chapter 4.2.19. A minor production of **39** and subsequent conversion to the new PoTeM was detected using analytical HPLC chromatography. As depicted in Figure A110, the amount of NP present in the supernatant

was insufficient. Accordingly, this extract was excluded from further purification. The cell pellet extract did not contain yields of **39** as detected before. However, a conversion of about 10-15% to the novel derivative was detected, rendering it suitable for purification. The extract was prepared for semi-preparative HPLC purification and isolated using the Eurosphere C8 column (12GE084E2J, 125 x 8 mm, 100-5 C8 A) and the method shown in Table 39. The fractions containing the derivative of **39** were combined, ACN was removed *in vacuo* and the residual water was lyophilized. Nevertheless, the yield of NP was less than 0.1 mg, an insufficient amount for NMR analysis.

It was already proven that the accessory genes for PoTeM BGCs annotated as sterol desaturase hydroxylate PoTeMs at position C-25, and therefore be renamed to PoTeM hydroxylases.[118] The PoTeM hydroxylase chosen for this study was utilized from the *deg* BGC, the 939 bp gene has an GC content of 44%, far less than typically for *Streptomyces*. The insufficient conversion of **39** to **60** was possibly caused by the genetic difference between heterologous host and original producer. Usage of such different GC contents influences the codon usage of the desired recombinant protein. Accordingly, codons with very low to zero usage in *Streptomyces* used for heterologous expression are present in the chosen modifying gene. A low amount of synthesized protein likely leads to the low conversion of **39** to **60** in the heterologous host. To circumvent this, the utilized gene can be codon optimized and ordered as a synthetic gene. Cloning this gene into one of the expression vectors might increase the amount of protein translated in the cells and thereby also increase the amount of the desired derivative. The primary study of promiscuous PoTeM modifying enzymes was conducted using the PoTeM hydroxylase present in the *L. capsici* PoTeM BGC. This 1005 bp gene has a GC-content of 65%, a GC-content more common for *Streptomyces*, which might also explain the elevated conversion measured by Greunke *et al*.[118] However, in addition to the proof of principle that PoTeM hydroxylases are promiscuous, this study showed that even genetically diverse genes can be combined to produce a hybrid PoTeM. It was shown, for a different PoTeM hydroxylase, from a BGCs coding for a yet unknown NP to hydroxylate **39.** Furthermore, the great difference of the *ika* cluster compared to the *deg* cluster emphasizes the potential of the novel PoTeM heterologous expression system, as it is possible to utilize PoTeM BGC genes from almost any organism and integrate it into the expression vectors to obtain a rearranged, hybrid or novel PoTeM in a plug-and-play fashion.

The second selected modification enzyme was PtmD, an enzyme proposed to integrate a hydroxy functionality at C-29.[136] The eight different expression constructs were again conjugated into the three different heterologous hosts and the extracts were analyzed using LC-MS. An exemplary HPLC chromatogram of extracts obtained from a *S. albus* culture harboring pSET152_ermE::*ikaABC*_gapdhP(EL)_*ptmD* is depicted in Figure 62. This proves the presence of **39** with a retention time of around 15 min and a second peak arising at a

slightly shorter retention time of around 14 min, supporting the higher polarity of the new compound, which is expected to be caused by the introduction of a hydroxyl group. Furthermore, the measured mass of the new metabolite is *m/z* = 495.1 $[M+H]^+$, which corresponds to the proposed product **62**.

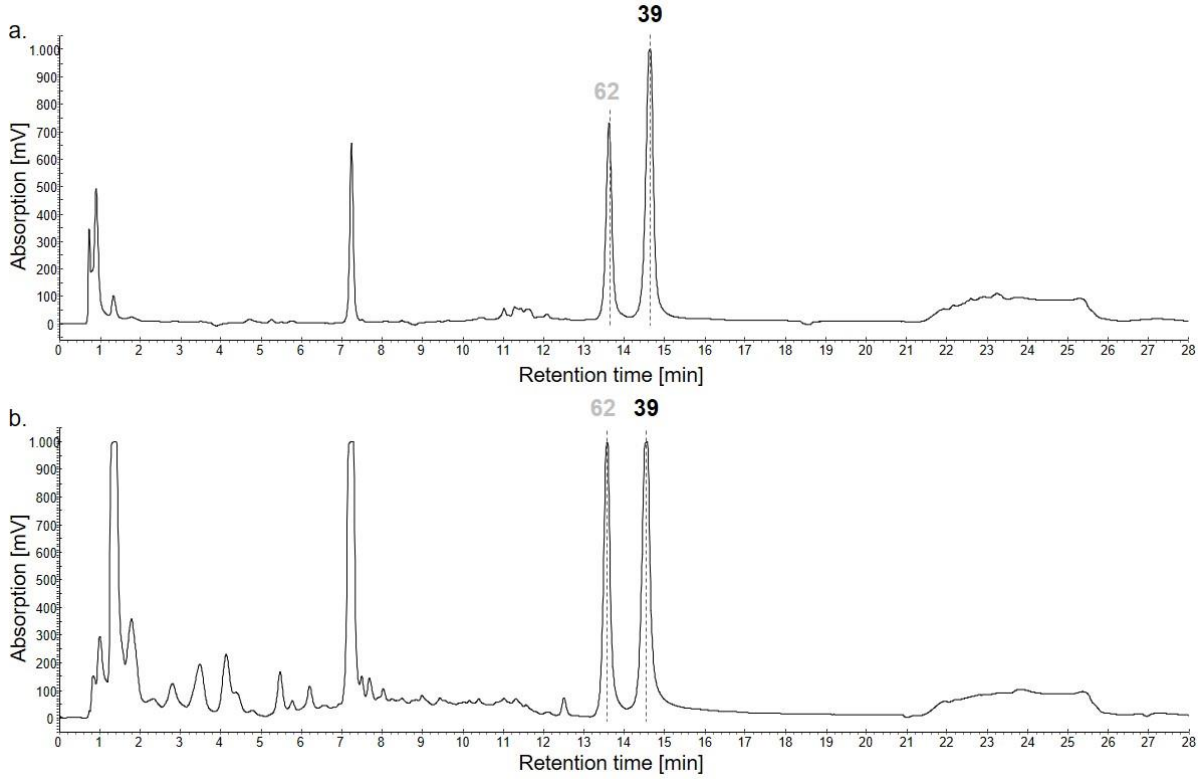


Figure 62. HPLC result of extracts containing the modifying enzyme PtmD. Exemplary HPLC chromatogram of *S. albus* supernatant (a.) and pellet (b.) extracts containing the expression constructs pSET152_ermE::*ikaABC*_gapdhP(EL)_*ptmD*. Predicted derivative of **39** is depicted in light grey (**62**). Analytical method: Eurosphere_Ika_lange Methode_AG.Meth.

The presentation of the absolute absorption values for the extracts containing the modifying enzyme PtmD (Figure 63) showed similarities to the measurements for the modifying enzyme DegA. The highest levels of NP are detected in extracts from *S. albus*, with minor expression in some *S. lividans* or *S. coelicolor* pellet extracts (Figure 63d and f). Additionally, the comparison of the promoters showed a similar result when compared to the analysis of DegA containing extracts. The promoters *gapdhP*(EL) and *rpsLP*(XC) effectuate a high production of the possible product **62** in the supernatant as well as in the pellet. Extracts of cell pellets containing *kasOP**, *SF14P*, *P-2* and *P-31* show similarly high levels of possible **62**. However, the level of the product is highly reduced in the supernatant extracts when compared to the results for *gapdhP*(EL) and *rpsLP*(XC). For the extracts containing the promoter *P-6* and *P-15*, neither in the pellet nor in the supernatant satisfactory amounts of the product were detected. These results justify the decision to use clone 3 harboring the expression construct

pSET152_ermE::*ikaABC*_gapdhP(EL)_*ptmD* for large scale production of the modified **39** product.

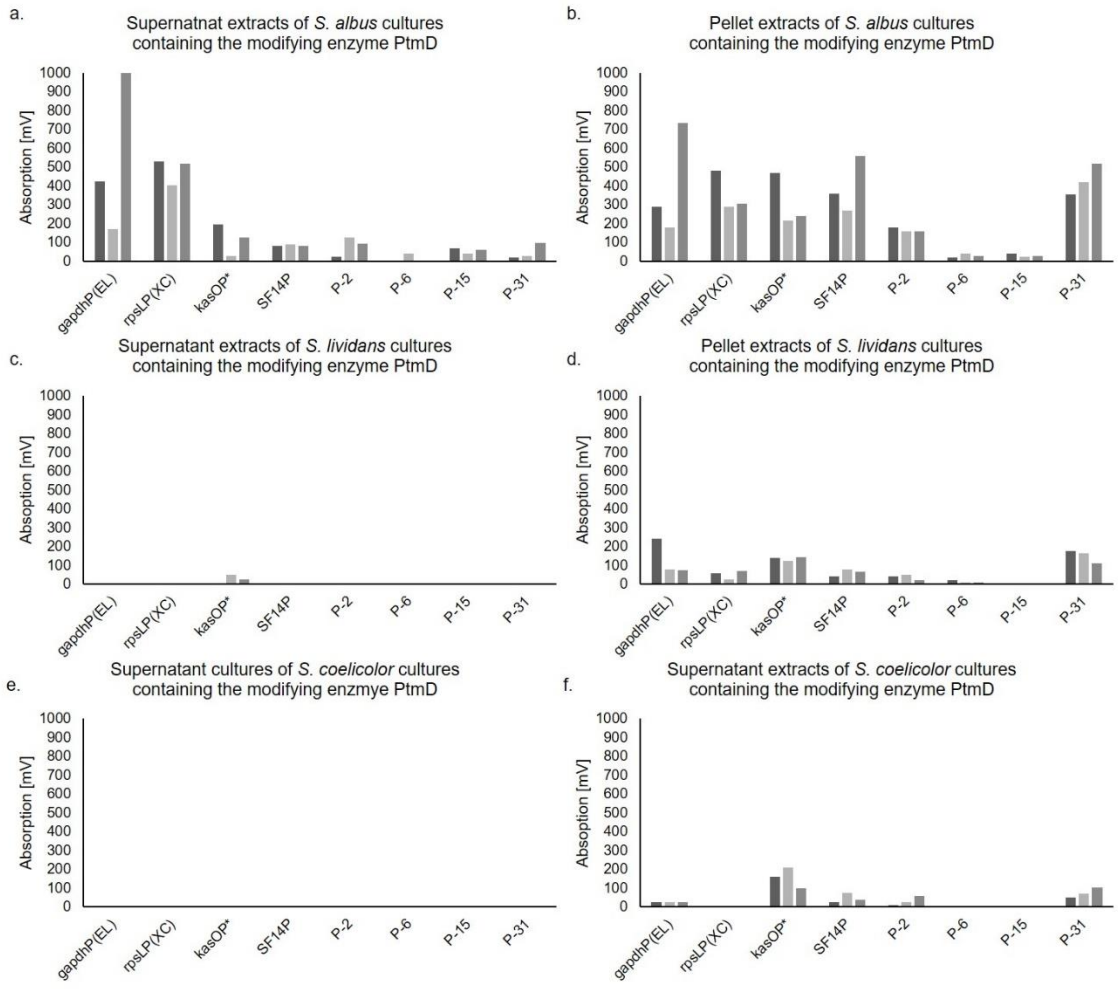


Figure 63. Bar diagrams of HPLC absorption for PtmD containing expression cultures. Depicted are the absolute absorption values of the newly observed peak in the heterologous expression extracts containing the modifying enzyme PtmD for *S. albus* supernatant (a.) and pellet (b.), *S. lividans* supernatant (c.) and pellet (d.) and *S. coelicolor* cultures supernatant (e.) and pellet (f.). The mass detected in the peak is *m/z* = 495.1 $[M+H]^+$, the mass predicted for **62**. The three different clones extracted are represented in different shades of gray.

A preculture was inoculated with frozen spores of the clone and passaged until enough volume was reached for the inoculation of 1 L total culture volume. Cells were incubated and extracted as described in chapter 4.2.1 and chapter 4.2.19 and analytical HPCL analysis of the extracts showed a much lower expression titer than measured in the previous experiments. After attempting different purification optimizations such as a previous sephadex column, changing for the small (12GE084E2J, 125 x 8 mm, 100-5 C8 A) to the bigger semi preparative HPLC column (25GE084E2J, 250 x 8 mm, 100-5 C8 A), it became obvious that the error were the frozen spores for culture inoculation. Therefore, a new conjugation was conducted, with exconjugants that were used for large scale purification without further analysis of the NP

production. After semi preparative HPLC purification, merging of the fractions and removal of the solvents, 1.8 mg of yellow powder was obtained and analyzed by ^{1}H-NMR analysis.

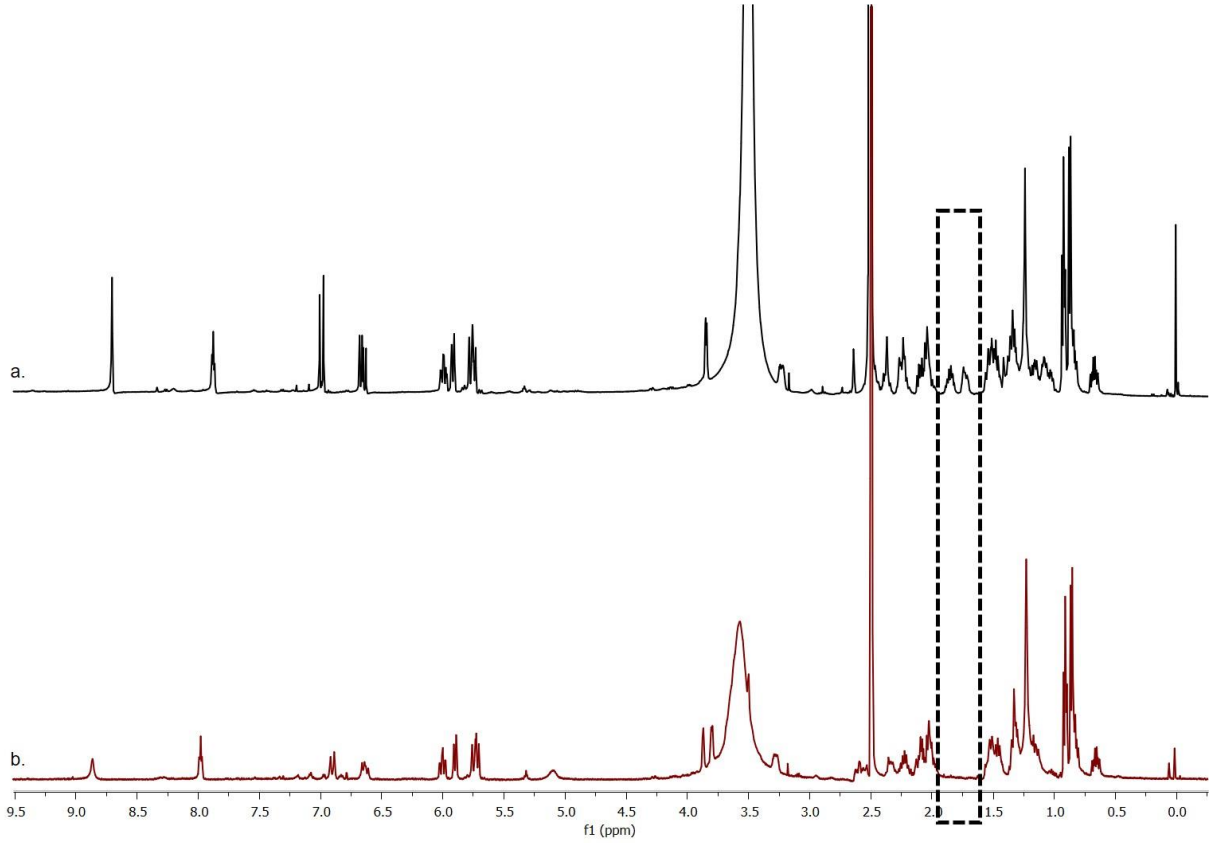


Figure 64. Comparing NMR spectra of **39** and modification catalyzed by PtmD. a. NMR spectrum of **39**, b. NMR spectrum of compound purified from PtmD containing cultures. Both compounds were dissolved in DMSO-D6 and measured at the AHV500 spectrometer (Bruker).

The comparison of the two ^{1}H NMR spectra of **39** (Figure 64a) and the modified compound (Figure 64b) received simultaneously during the purification gives a precise hint of the position that is modified by PtmD. The main difference between the two spectra is the absence of two signals between 1.5 and 2.0 ppm of the modified compound (Figure 64, dotted square). These two signals were assigned to possibly correspond to the two hydrogen atoms present at position C-25.[113] Furthermore, the published NMR spectrum for **60** lacks these peaks as well.[119] This result combined with the LC-MS data leads to the conclusion that PtmD is another PoTeM hydroxylase incorporating a hydroxy functionality at position C-25, resulting in **60** as the modified derivative in contrast to the proposed compound **62**. This data strongly suggests that the postulate that PtmD to hydroxylates at position C-29, as published by Saha *et al.* is incorrect,[136] and there must be a different modifying enzyme responsible for hydroxylation at C-29 found in the pactamides. This assumption was further proven by merging extracts from cultures containing *degA* and *ptmD*. As depicted in Figure A111 there is only one peak of a

derivative of **39** detectable, strongly suggesting **60** to be the product of **39** derivatized by DegA as well as PtmD.

The result of PtmD most likely belonging to the PoTeM hydroxylases emphasizes the potential of the strategy for PoTeM heterologous expression. This plug-and-play system enabled the investigation of the function of a new PoTeM modifying in a fast and efficient toolbox fashion.

The last analyzed modifying enzyme was the monooxygenase IkaD. HPLC chromatograms of the heterologous expression extracts showed a greatly differing course compared to the PoTeM hydroxylase DegA. Figure 65 shows a representative expression of *S. albus* DMS40313 harboring pSET152_ermE::*ikaABC*_kasOP*_*ikaD*. The substrate **39** is detected at a retention time of 12.5 min, with almost no substrate remaining. The supernatant extracts show a large range of several new peaks arising between 8 to 12 min. The major peaks can be detected at retention times of 10 min and 7.5 min, depicting a conversion into **63** and **64** respectively (Figure 65a). The peaks between 8 to 12 min show a **39**-like UV-spectrum suggesting that further modification with the epoxide occurred, either in the cell, or in the culture broth. As an epoxide is a very reactive functional group, further modification likely has happened without enzymatic participation in the culture broth. The pellet extract shows less unspecific modifications and almost exclusively harbors the two major peaks **63** and **64**, and a minor amount of **39** is also remaining (Figure 65b). The presence of both NPs was further corroborated by MS and UV analysis.

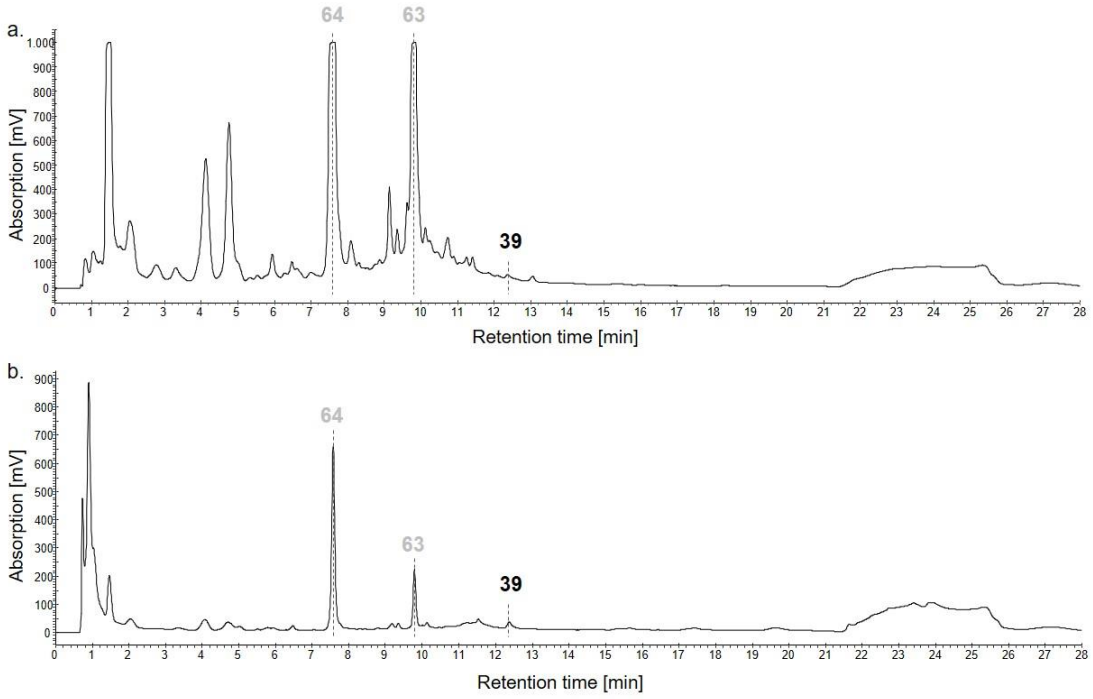


Figure 65. HPLC result of extracts containing the modifying enzyme IkaD. Exemplary, HPLC chromatogram of *S. albus* supernatant (a.) and pellet (b.) extracts containing the expression constructs pSET152_ermE::*ikaABC*_kasOP*_*ikaD* are shown. Predicted derivatives of **39** are labeled in light grey (**63** and **64**). Analytical method: Eurosphere_Ika_lange Methode_AG.Meth.

The total absorption values for the different extractions conducted with heterologous expression cultures containing IkaD as a modifying enzyme are depicted in Figure 66 and Figure 67. When starting to analyze the modification of **39** catalyzed by IkaD, the epoxide **63** was the mainly expected product with a mass of *m/z* = 495.1 $[M+H]^+$ as shown in Figure 66. None of the three different heterologous hosts showed satisfying titers of the expected product. The promoters *P-2* and *P-15* represent positive exceptions in the heterologous host *S. albus*. In this case the absorption was around 400 mV. However, the product was only present in the cell pellet, supporting the assumption that the epoxide was further modified in the culture broth without enzyme catalysis. Furthermore, *S. albus* extracts showed impurities with a similar retention time as **63**, hindering a straightforward large-scale purification for NMR analysis. Comparison of *S. albus* and *S. coelicolor* revealed that the same promoters, *kasOP** as well as the promoters from *S. albus* J1024 (*P-2*, *P-6*, *P-15*, and *P-31*) enabled the heterologous expression of **63** (Figure 66b and f). In contrast to the *S. albus* extracts, *S. coelicolor* extracts did not show impurities, rendering the use of *S. coelicolor* for the heterologous expression of **63** possible. The minor presence of **63** in the supernatant fractions of *S. coelicolor* cultures harboring the *kasOP** promoters indicates a potential transport of the NP out of the cell, usually suggesting a high NP production titer inside the cell, a fact that encouraged the decision for the pSET152_ermE::*ikaABC*_kasOP*_*ikaD* construct as the choice for large scale expression.

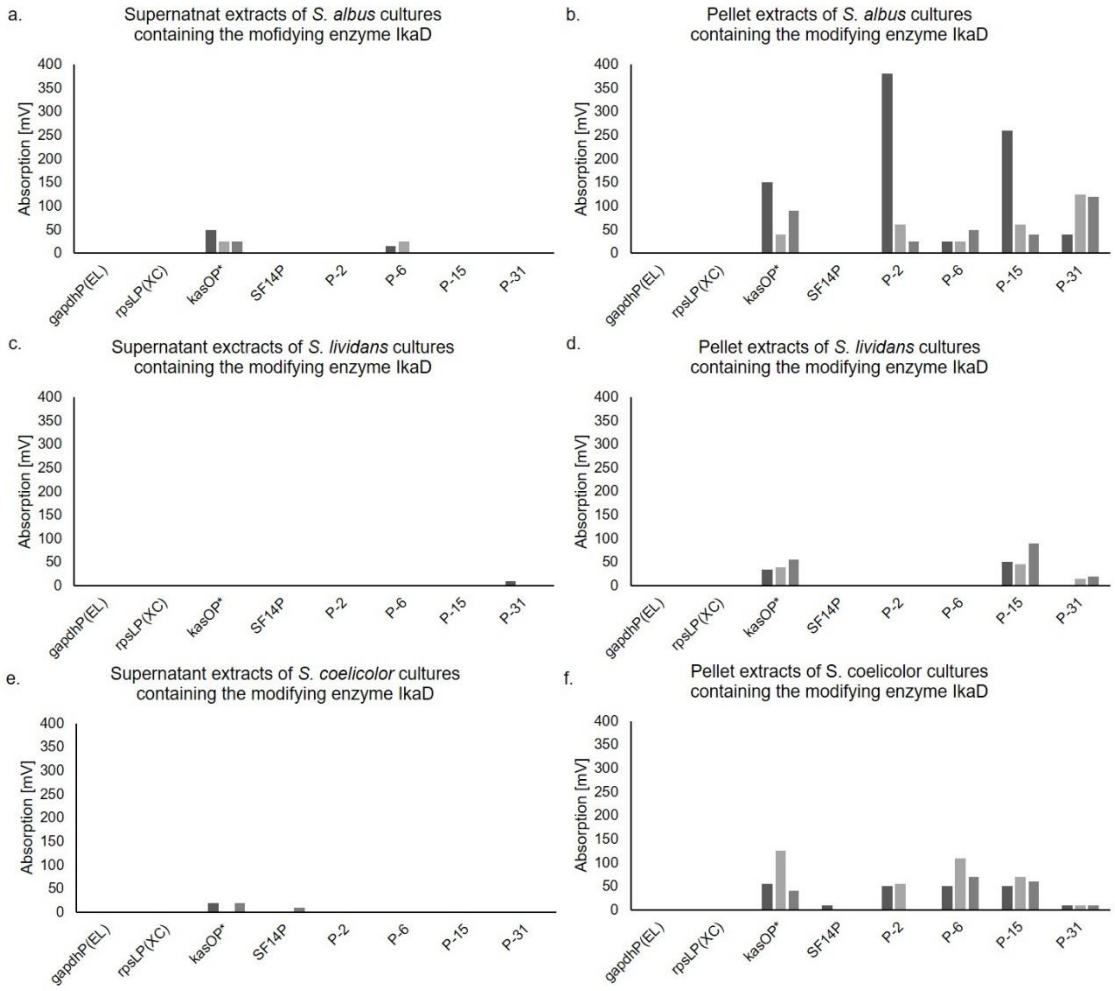


Figure 66. Bar diagrams of HPLC absorption for IkaD containing expression cultures. Depicted are the absolute absorption values of the newly observed peak corresponding to the ikarugamycin-epoxide (**63**) in the heterologous expression extracts containing the modifying enzyme IkaD for *S. albus* supernatant (a.) and pellet (b.), *S. lividans* supernatant (c.) and pellet (d.) and *S. coelicolor* cultures supernatant (e.) and pellet (f.). The mass detected in the peak is *m/z* = 495.1 $[M+H]^+$, the mass predicted for **63**. The three different clones extracted are represented in different shades of gray.

Literature research descried IkaD to have a secondary functionality.[137] Additionally to forming an epoxide, a hydroxylation at C-30 can be conducted by IkaD. Accordingly, the HPLC chromatograms were analyzed considering the double-modified product **64** (Figure 67). Consequential to **63** being an intermediate product, the amounts of **64** detected in the extracts is much higher. As visible in Figure 67, a high amount of possible **64** can be detected in the supernatant, which might be caused by a higher stability of the final product in the supernatant, or by an increased export of **64** from the cell into the culture broth.

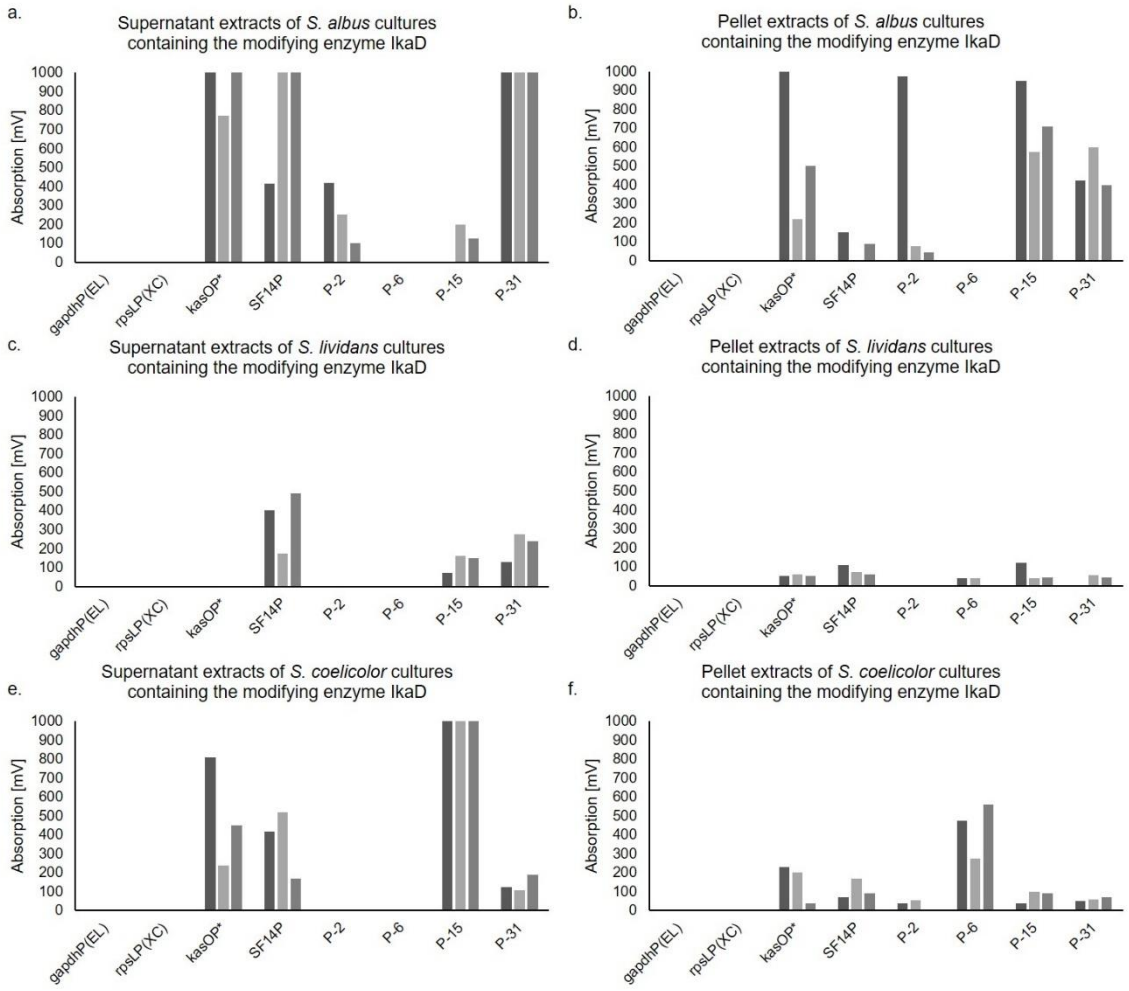


Figure 67. Bar diagrams of HPLC absorption for IkaD containing expression cultures. Depicted are the absolute absorption values of the newly observed peak corresponding to the hydroxylated ikarugamycin-epoxide (**64**) in the heterologous expression extracts containing the modifying enzyme IkaD for *S. albus* supernatant (a.) and pellet (b.), *S. lividans* supernatant (c.) and pellet (d.) and *S. coelicolor* cultures supernatant (e.) and pellet (f.). The mass detected in the peak is *m/z* = 511.5 $[M+H]^+$, the mass predicted for **64**. The three different clones extracted are represented in different shades of gray.

For all three hosts, promoters exist that enable production of **64**. However, *S. albus* and *S. coelicolor* represent the more promising hosts for isolation of both NPs because of increased yields. While the four promoters derived from *S. albus* J1024 reliably promoted the conversion of **39** to **63**, the additional conversion to **64** shows large deviation. In one extract containing *P-2* in *S. albus,* a high titer of **63** can be observed. However, none of the other *P-2* containing extracts show similar production enhancement. Extracts containing *P-6* show almost no conversion to **64** with only the *S. coelicolor* cell pellet extracts depicting a satisfactory production. *P-15* and *P-31* seem to enhance **64** production greatly. However, for *S. albus* the titers are higher in the pellet extracts and *S. coelicolor* cultures containing these two promoters show higher titers in the supernatant extracts. Compound **64** measured in extracts harboring the *kasOP** promoter showed constantly high expression as detected for **63** in the different heterologous hosts. Overall, these observations lead to the assumption that not only the

promoter, but also the respective heterologous host as well as the selected culturing conditions influence the production of the NP expressed using the plug-and-play expression construct. It must be kept in mind that the optimized expression conditions used in this study were exerted from the previous **39** heterologous expression in *S. albus* (chapter 2.1.3), neglecting the different growth characteristics of the two other hosts *S. lividans* and *S. coelicolor*. Additionally, *SF14P* containing constructs showed good **64** levels, especially in *S. lividans*, whereas there was no detection of **63** in any of the three hosts possible.

In the extracts harboring the two promoters *gapdhP*(EL) and *rpsLP*(XC), neither **39** nor one of the products **63** or **64** were detected, considering the results from the modification with the PoTeM hydroxylase, a systematic error can be assumed. Most likely, the conjugation did not work successfully as explained earlier, the impaired growth of *E. coli* ET12567 containing the complex expression vector was a high burden for conjugation. Accordingly, it might be possible that exconjugants did not harbor the full expression plasmid but are no longer sensitive to apramycin for the reason described above. A mutation in the expression vector is unlikely the reason for NP absence, as the same intermediate expression constructs were used to integrate *degA*, *ikaD* and *ptmD*. As already proven for the PoTeM hydroxylase, **39** should have been detected using these expression systems, which is absent in the pSET152_ermE::*ikaABC*_gapdhP(EL)_*ikaD* and pSET152_ermE::*ikaABC*_rpsLP(XC)_*ikaD* containing cultures.

For large scale production the amount of heterologously expressed compound, **63** and **64**, as well as the simplicity of the isolation due to less complex overall metabolome as shown in the HPLC chromatograms was considered to enable a straightforward purification. Accordingly, *S. coelicolor* harboring pSET152_ermE::*ikaABC*_kasOP*_*ikaD* clone 1 and *S. albus* harboring pSET152_ermE::*ikaABC*_P-2_*ikaD* clone 1 were selected.

Small precultures were inoculated with frozen spores of the corresponding clone and passaged into larger volumes until a sufficient amount for the inoculation of twenty 50 ml cultures were reached. The cultures were incubated as described (chapter 4.2.1) and cell pellet and supernatant were extracted with ethyl acetate (chapter 4.2.19). Extracts were prepared for semi preparative HPLC purification. However, the amount of detected NP was much lower and only detectable in the pellet fraction. As the different modifying enzymes were handled in parallel, the reason of failed heterologous NP production from frozen spores without new conjugation was not identified. Accordingly, the presence of TFA in the organic solvents of the HPLC was assumed to be responsible for the destruction of **63** and **64**. As a result, a purification without TFA was attempted, again resulting in impure and insufficient amounts of purified NP. With a novel conjugation of pSET152_ermE::*ikaABC*_P-2_*ikaD* into *S. albus*, the amounts of **63** and **64** in analytical HPLC increased, leading to a new trial of purification using

semipreparative HPLC with the Eurosphere C8 column (25GE084E2J, 250 x 8 mm, 100-5 C8 A) and the gradient represented in Table 41. The fractions were collected, directly diluted with water to reduce the TFA concentration, combined and ACN was removed *in vacuo*. The remainder was lyophilized, leading to yellowish powder of up to 3 mg of compound **64**. However, additional work will be needed to accomplish a complete structural elucidation by NMR.

In future studies, it might be interesting to further optimize IkaD expression. As it was shown that *gapdhP*(EL) and *rpsLP*(XC) increase the modification of **39** by the other accessory enzymes, it might be interesting to see if a higher NP titer can be obtained with these expression constructs. Furthermore, the significant amount of IkaD modified NP in *S. coelicolor*, without the impurities and unspecific modifications found in *S. albus*, might support the optimization of IkaD derivatization for this strain. This would mainly include a time curve, as this strain growth is different to *S. albus* and maybe further studies on the usage of different culturing media or wire spirals.

The two promoters originating from a comparable study of house-keeping genes in *S. griseus*, *gapdhP*(EL) and *rpsLP*(XC), reliably result in a high titer of **39** derivative, corresponding to an adequate production of the corresponding enzyme. It is still yet to determine the result of these to promoters for IkaD production, as in the current study no results were obtained. When using these promoters it has to be kept in mind that they are very strong constitutive promoters, with *rpsLP*(XC) even been described as an early onset promoter.[133] This must be considered for heterologous expression of complex or unstable NPs.

Cultures containing *kasOP** expression constructs exhibit solid production levels. However, the titers of converted product were mostly lower compared to *gapdhP*(EL) and *rpsLP*(XC). It was obvious, especially for IkaD, that *S. coelicolor* cultures harboring expression constructs with *kasOP** manifest a significant compound conversion. As *kasOP** originates from a promoter found in a BGC of a *S. coelicolor* strain, this supports the assumption that *kasOP** is most effective for heterologous expression in *S. coelicolor*. If further experiments, as described above, are conducted for optimizing **63**/**64** production in *S. coelicolor*, it might reveal the potential of this promoter for *S. coelicolor*. As this study was optimized for *S. albus* considering incubation time, culture media and usage of a wire spirals, the full potential of this promoter might be even higher.

Similar to *kasOP**, a considerable conversion of **39** was detected in cultures harboring *SF14P* containing expression constructs. Again, lower titers of the desired derivatives were observed when compared to *gapdhP*(EL) and *rpsLP*(XC). For this study, it was not sufficient to be used for large scale production. However, a lower amount of recombinant protein in the cells might be an advantage for labile compounds in slower growing heterologous hosts.

The four promoters identified from untranslated regions in *S. albus* J1074, *P-2*, *P-6*, *P-15*, and *P-31*, underlined the potential of the study conducted by Luo *et al*.,[135] as cultures containing expression vectors with these promoters show considerable conversions of **39**. However, the poor evidence of the actual promoting bases in those untranslated regions lead to large gap of noncoding bases in the final plug-and-play expression vector. All the four promoters, especially *P-6*, include secondary structures in their sequence that rendered cloning difficult to almost impossible. Accordingly, a more precise study of the actual promoter regions inside the untranslated regions would further improve their usability for heterologous expression in *Streptomyces.*

2.5.2 Applying the plug-and-play system for derivatives of **39**

The reproducibility of the gained expertise using the new plug-and-play system was proven by the integration of a fourth modifying enzyme behind the ikarugamycin BGC (Figure 68). For this experimental setup, CftA was chosen as the modifying enzyme of interest. CftA is a cytochrome P450 enzyme encoded in the BGC of the clifednamides.[114]

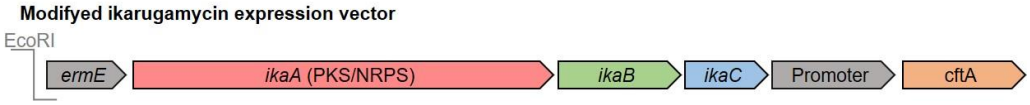


Figure 68. Expression vector for heterologous production of derivatives of **39** modified by CftA.

CftA was already demonstrated to modify **39** by inserting a keto functionality at C-30.**[138]** The expected product is depicted in Figure 69. With the knowledge of the possible product (**55**), the modification of **39** using CftA with our newly established plug-and-play system represents a straightforward proof of concept.

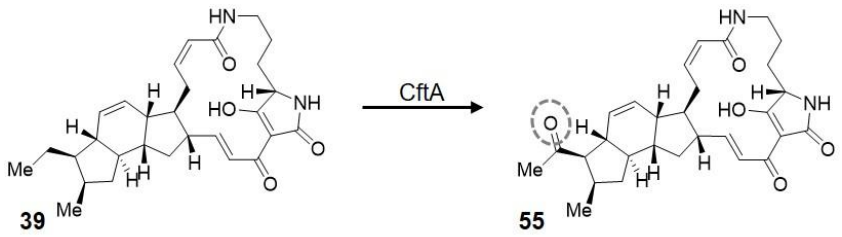


Figure 69. Modification of **39** by CftA. The expected product **55** of the heterologous expression cultures using expression vectors containing the modifying gene *cftA* is depicted.

The previous fermentations revealed the optimized expression vector composition and culturing conditions, as described above. Therefore, the new modifying enzyme CftA was integrated into pSET152_ermE::*ikaABC*_gapdhP(EL). The gene was ordered as a synthetic

construct in a standard vector pUC57. For isolation of the gene, the plasmid was amplified in *E. coli* DH5α, the isolated vector was digested using XbaI (chapter 4.2.5), and the gene was purified using gel extraction (chapter 4.2.4). The isolated DNA and the XbaI digested vector were ligated (chapter 4.2.7) and transformed into *E. coli* DH5α (chapter 4.2.10). Positive colonies were determined using colony screening PCR (chapter 4.2.13), verified by analytical restriction digest (chapter 4.2.5) and Sanger sequencing (chapter 4.2.14). The results of the successful cloning are represented in Figure A94. The final expression construct was transformed into *E. coli* ET12567 and underwent conjugation into *S. albus* DSM40313 (chapter 4.2.12). Three exconjugants were picked and utilized for a 7 d culture in ISP-4 medium (chapter 4.2.1). After fermentation, the cells were harvested and both supernatant and pellet were extracted using organic solvents (chapter 4.2.19). LC-MS analysis of the extracts was performed (chapter 4.2.20) using the gradient depicted in Table 37. The results of the fermentation showed that CftA, encoded in the new expression plasmid, is capable of modifying **39**, with an almost complete conversion into the derivatives **55** and **56**. As depicted in Figure 70, two new major peaks arise at a retention time of 9.5 min and 10.5 min. Both are detectable in the supernatant and pellet fraction, with a significantly higher amount present in the supernatant fraction and only traces of NP in the cell pellet. Both peaks show the characteristic UV-spectrum of **39** (data not shown).

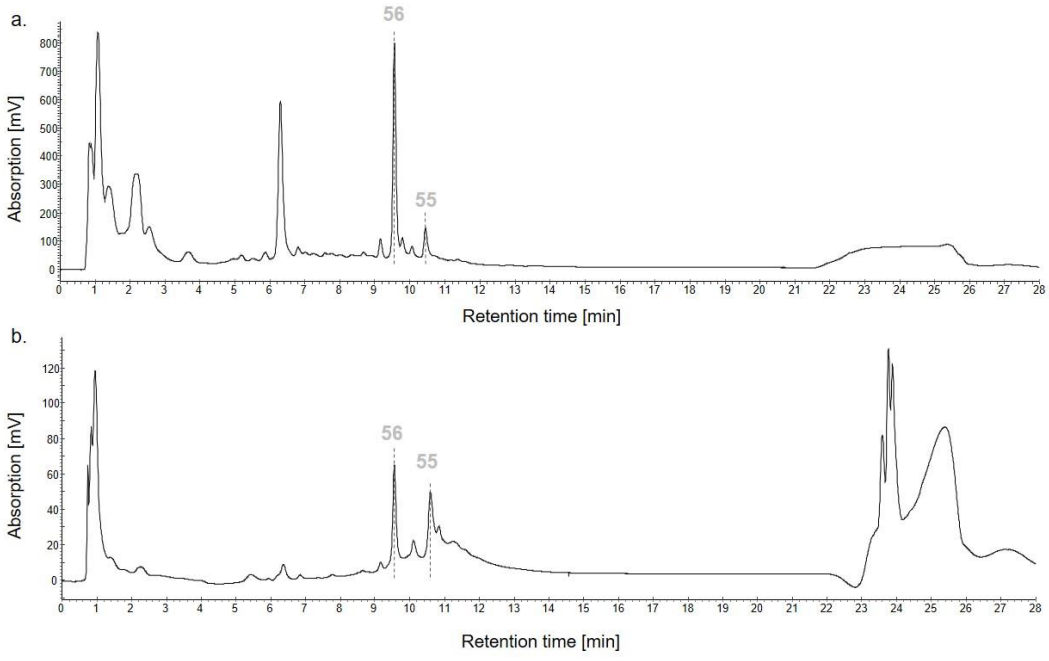


Figure 70. HPLC result of extracts containing the modifying enzyme CftA. a.) Exemplary HPLC chromatogram of *S. albus* clone 3 supernatant and b.) pellet extracts containing the expression construct pSET152_ermE::*ikaABC*_gapdhP(EL)_*cftA*. Predicted derivative of **39** is depicted in light grey. Analytical method: Eurosphere_Ika_lange Methode_AG.Meth.

Evaluation of the mass data indicated that the minor peak at 10.5 min should correspond to clifednamide A (**55**), as the expected mass of *m/z* = 493.3 $[M+H]^+$ was detected (Figure 71a). The second, more prominent peak at a retention time of 9.5 min showed a mass of *m/z* = 509.3 $[M+H]^+$ (Figure 71b). This mass would correspond to a further modification of **55** by hydroxylation to result in **56**, a second compound, clifednamide B, present in the natural producer.[114]

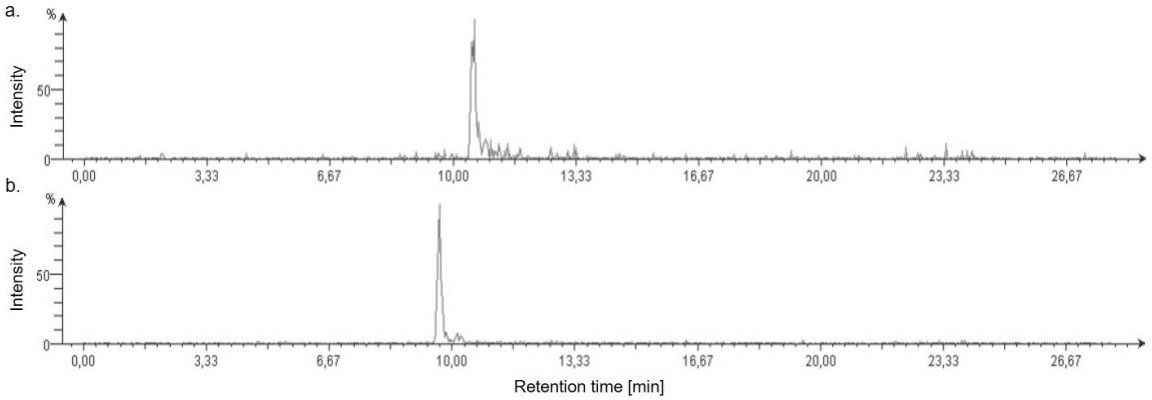


Figure 71. Mass analysis of *S. albus* pSET152_ermE::*ikaABC*_gapdhP(EL)_*cftA*. a.) Depicted is the extracted mass at *m/z* = 493.3 $[M+H]^+$, found in the HPLC chromatogram of the supernatant at 10.5 min, corresponding to **55**, the keto-functionalized **39**. b.) Depicted is the extracted mass of *m/z* = 509.3 $[M+H]^+$ present in the HPLC chromatogram at 9.5 min, the mass corresponds to a further hydroxylated derivative of **39**. Analytical method: Eurosphere_Ika_lange Methode_AG.Meth.

For structural elucidation and further proof that *cftA* added to the *ika* BGC results in the production of **55** and **56**, a large-scale fermentation of the exconjugant was performed. One-liter fermentation culture, separated in 20 times 50 ml cultures, were incubated for 7 d and the supernatant was extracted with organic solvents, as the majority of NP was present in the supernatant fraction during the small-scale fermentation. The extract was dissolved in methanol, filtered and used for purification by semi preparative HPLC (chapter 4.2.20) using the method depicted in Table 41. The collected fractions of the compound with a retention time of 9.5 min were combined, ACN was removed *in vacuo* and the watery fraction was lyophilized. The final dry NP yielded 5 mg. In this case, the time period between conjugation and large-scale fermentation was minimized, without the additional freezing of spores, further underlining the importance of a fresh conjugation for maximal NP yield.

Preliminary NMR data supports the double modified compound, **56**, as the major product of the heterologous expression. However, more material is needed to conduct a complete structure elucidation of the heterologous expression products.

2.5.3 Applying the plug-and-play system for novel PoTeMs

The established PoTeM plug-and-play system showed promising results to produce derivatives of **39**. As it was not possible to obtain novel PoTeMs in sufficient amounts using the general heterologous expression approach, the novel plug-and-play system was next to be examined for the PoTeM clusters *spi*, *deg* and *fer*.

The PoTeM BGC present in the *S. spinosa* genome, *spi*, was already proven to produce a novel compound that could possibly be a novel PoTeM (chapter 2.4). Accordingly, this cluster was chosen to be the first novel PoTeM BGC heterologously expressed utilizing the new plug-and-play system. In contrast to the cluster responsible for the biosynthesis of **39**, this cluster contains genes positioned upstream of the iPKS/NRPS. Correspondingly, different organizations of the genes can be tested with the novel plug-and-play system to yield maximal compound expression. The two orders of the rearranged cluster tested in this study are depicted in Figure 72b and c. The order displayed in Figure 72b, corresponds to that found in the wildtype cluster and was hence named as *spi*$_{WT}$, whereas the order in Figure 72c shows a modified arrangement therefore termed *spi*$_{MOD}$. The two PCR products for each possibility are indicated by the bold lines, with the cloning step specified on top of the line.

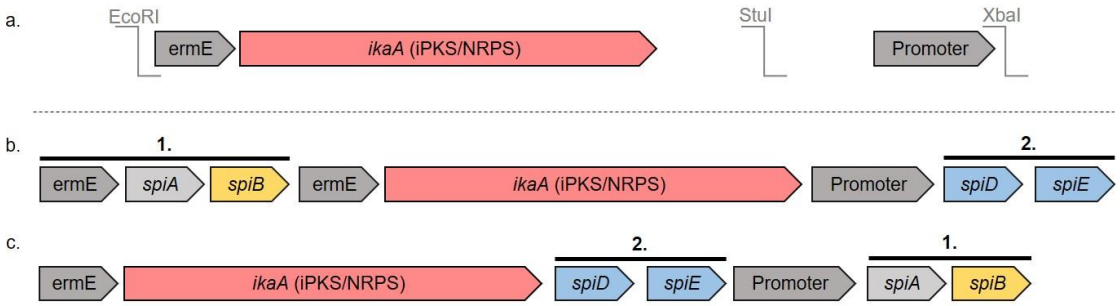


Figure 72. Plug-and-play gene sequences of PoTeM cluster *spi*. a. Basic plug-and-play expression vector. b./c. Possibilities to rearrange *spi* PoTeM cluster to fit into novel PoTeM plug-and-play system. PCR products for cloning are indicated by bold lines, numbers above indicating cloning order.

Cloning of the *spi*$_{WT}$ constructs was achieved by PCR amplifying the front part, *spiAB*, of the *spi* cluster (*spi*_front, *sf*) from an expression construct obtained in chapter 2.4 (pSET152_ermE::*spi*_full_new). To ensure transcription of the genes, the additional *ermE* promoter was included on the PCR product (ermE_*sf*). The basic plug-and-play vectors were EcoRI digested (chapter 4.2.5) and ligation of the two pieces was conducted using the SLIC method (chapter 4.2.7). Plasmid preparation (chapter 4.2.4) was performed for possible positive clones and cloning success was verified using analytical restriction digest (chapter 4.2.5) and Sanger sequencing (chapter 4.2.14). As depicted in Table 12, all seven constructs were easily cloned using the new SLIC based DiPaC method.[139] In this case, the method was

used for an insert approximately 2,500 bp in size and a vector of more than 15,000 bp and for none of the constructs more than ten clones were needed to be screened to find a positive clone. As described above, the two promoters *ermE* and *P-6* interfered negatively in cloning efficiency and were therefore excluded from further studies.

Table 12. Result summary of pSET152_ermE_*sf*_ermE_*ikaA*_P cloning. The table shows the correct clones for the first cloning step to construct *spi*$_{WT}$ expression vectors. References to the respective figure in the appendix for result description are assigned.

Expression construct	Correct clone	Reference to results
pSET152_ermE::*sf*_ermE_*ikaA*_gapdhP(EL)	Clone 1	Figure A95
pSET152_ermE::*sf*_ermE_*ikaA*_rpsLP(XC)	Clone 5	Figure A95
pSET152_ermE::*sf*_ermE_*ikaA*_kasOP*	Clone 1	Figure A95
pSET152_ermE::*sf*_ermE_*ikaA*_SF14P	Clone 1	Figure A95
pSET152_ermE::*sf*_ermE_*ikaA*_P-2	Clone 6	Figure A95
pSET152_ermE::*sf*_ermE_*ikaA*_P-15	Clone 2	Figure A95
pSET152_ermE::*sf*_ermE_*ikaA*_P-31	Clone 3	Figure A95

For the second cloning step, conventional ligation cloning was the method of choice, as this again enabled one PCR product to be integrated into all the seven different vectors. The insert, the back part, *spiDE*, of the *spi* BGC (*spi*_back, *sb*), was PCR amplified (chapter 4.2.13) and both, vector and PCR product were digested with XbaI (chapter 4.2.5), the vector was additionally dephosphorylated (chapter 4.2.6). Ligation (chapter 4.2.7), transformation into chemically competent *E. coli* DH5α (chapter 4.2.10) and plasmid preparation (chapter 4.2.4) were conducted. Successful integration of the insert was verified using analytical restriction digest (chapter 4.2.5) and Sanger sequencing (chapter 4.2.14).

Table 13. Result summary of pSET152_ermE::*sf*_ermE_*ikaA*_P_*sb* cloning. The table shows the correct clones for the second cloning step to finalize *spi*$_{WT}$ expression vectors. References to the respective figure in the appendix for result description are assigned.

Expression construct	Correct clone	Reference to results
pSET152_ermE::*sf*_ermE_*ikaA*_gapdhP(EL)_*sb*	Clone 16	Figure A96
pSET152_ermE::*sf*_ermE_*ikaA*_rpsLP(XC)_*sb*	Clone 28	Figure A96
pSET152_ermE::*sf*_ermE_*ikaA*_kasOP*_*sb*	Clone 17	Figure A96
pSET152_ermE::*sf*_ermE_*ikaA*_SF14P_*sb*	Clone 18	Figure A96
pSET152_ermE::*sf*_ermE_*ikaA*_P-2_*sb*	Clone 15	Figure A96
pSET152_ermE::*sf*_ermE_*ikaA*_P-15_*sb*	Clone 21	Figure A96
pSET152_ermE::*sf*_ermE_*ikaA*_P-31_*sb*	Clone 3	Figure A96

The modified expression constructs for the *spi* BGC, spi_{MOD}, was also obtained in two separate cloning steps. First the front part (*sf*) was PCR amplified (chapter 4.2.13). In this case, the PCR product did not contain the additional *ermE* promoter. Subsequently it was introduced into the basic plug-and-play vector utilizing conventional ligation cloning. This cloning strategy enabled one PCR product to be integrated into all the different plug-and-play vectors. The vectors were digested with XbaI (chapter 4.2.5) and dephosphorylated (chapter 4.2.6), subsequently the XbaI digested insert was ligated to the digested vectors (chapter 4.2.7) and transformed into chemically competent *E. coli* DH5α (chapter 4.2.10). Cloning success was verified using analytical restriction digest (chapter 4.2.5) and Sanger sequencing (chapter 4.2.14).

Table 14. Result summary of pSET152_ermE::*ikaA*_P_*sf* cloning. The table shows the correct clones for the first cloning step to obtain spi_{MOD} expression vectors. References to the respective figure in the appendix for result description are assigned.

Expression construct	Correct clone	Reference to results
pSET152_ermE::*ikaA*_gapdhP(EL)_*sf*	Clone 4	Figure A97
pSET152_ermE::*ikaA*_rpsLP(XC)_*sf*	Clone 3	Figure A97
pSET152_ermE::*ikaA*_kasOP*_*sf*	Clone 2	Figure A97
pSET152_ermE::*ikaA*_SF14P_*sf*	Clone 1	Figure A97
pSET152_ermE::*ikaA*_P-2_*sf*	Clone 10	Figure A97
pSET152_ermE::*ikaA*_P-15_*sf*	Clone 7	Figure A97
pSET152_ermE::*ikaA*_P-31_*sf*	Clone 1	Figure A97

The final spi_{MOD} constructs were obtained by a second cloning step. SLIC was chosen as the cloning method, although a single PCR product was needed for each promoter. However, the presence of two StuI restriction sites inside *spiDE* (*sb*) prohibited a conventional ligation cloning strategy. Accordingly, the PCR products were amplified (chapter 4.2.13) and the vectors were digested with StuI (chapter 4.2.5) and dephosphorylated (chapter 4.2.6). SLIC was performed (see chapter 4.2.7) and plasmid DNA was prepared (chapter 4.2.4) from possible positive clones determined by colony screening PCR (chapter 4.2.13). Cloning success was verified conducting analytical restriction digest (chapter 4.2.5) and Sanger sequencing (chapter 4.2.14).

Table 15. Result summary of pSET152_ermE::*ikaA_sb*_P_*sf* cloning. The table shows the correct clones for the final cloning step to obtain *spi*$_{MOD}$ expression vectors. References to the respective figure in the appendix for result description are assigned.

Expression construct	Correct clone	Reference to results
pSET152_ermE::*ikaA_sb*_gapdhP(EL)_*sf*	Clone 1	Figure A98
pSET152_ermE::*ikaA_sb*_rpsLP(XC)_*sf*	Clone 4	Figure A98
pSET152_ermE::*ikaA_sb*_kasOP*_*sf*	Clone 4	Figure A98
pSET152_ermE::*ikaA_sb*_SF14P_*sf*	Clone 1	Figure A98
pSET152_ermE::*ikaA_sb*_P-2_*sf*	Clone 1	Figure A98
pSET152_ermE::*ikaA_sb*_P-15_*sf*	Clone 3	Figure A98
pSET152_ermE::*ikaA_sb*_P-31_*sf*	Clone 15	Figure A98

The finalized expression plasmids containing *spi*$_{WT/MOD}$ were conjugated into *S. albus* DSM40313 (chapter 4.2.12). *S. albus* was exclusively chosen as the only expression host, as this host generally represented the most favorable host for PoTeM expression (chapter 2.5.1). Three exconjugants for each construct were PCR verified (chapter 4.2.13) and used for inoculation of a preculture. As determined in previous studies, the main cultures were incubated for 3 d at 28 °C and 200 rpm (chapter 4.2.1). Subsequent to the incubation, cell pellets and supernatants were separated and extracted with organic solvent (chapter 4.2.19). Dried cell extracts were dissolved in methanol and analyzed by LC-MS (chapter 4.2.20) using the gradient depicted in Table 37. As described in the primary study, two new peaks with a mass of *m/z* = 479 $[M+H]^+$, at a retention time of 14.5 min and *m/z* = 495 $[M+H]^+$, at a retention time of 13.5 min were visible. In Figure 73 an exemplary HPLC chromatogram of a supernatant and pellet extract is shown, which indicates that the novel NP is mainly found in the pellet fraction. From the identical UV-spectrum of the two peaks and the difference in mass, it can be proposed that the peak at 13.5 min depicts a hydroxylated NP version of the compound found at a retention time of 14.5 min.

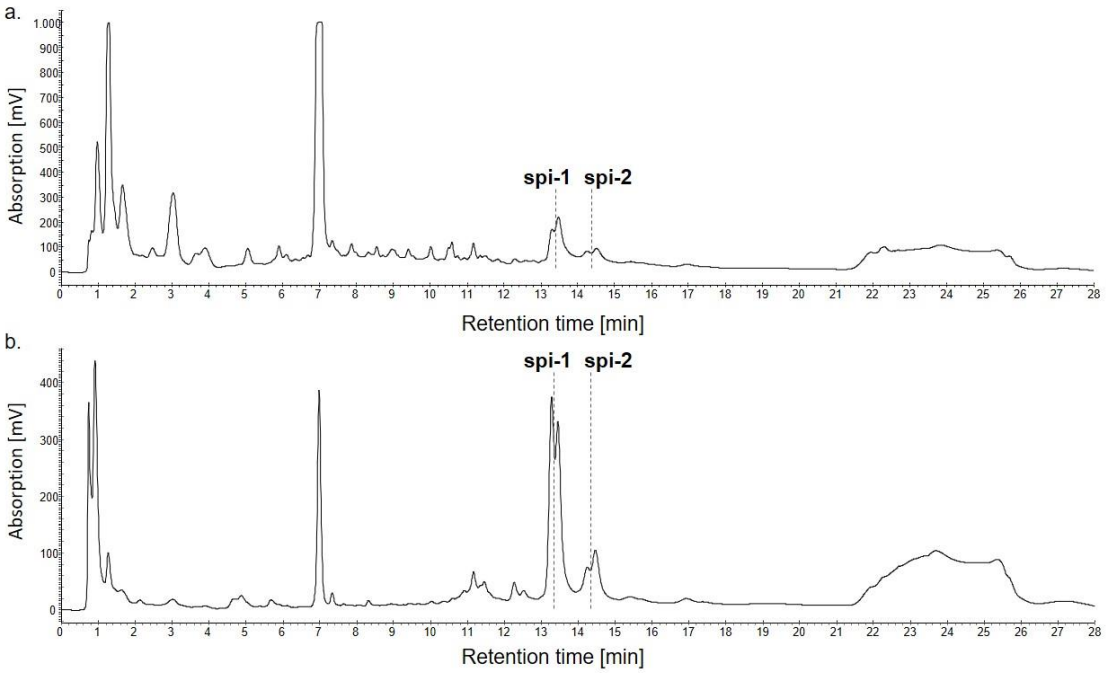


Figure 73. HPLC results of *spi*$_{WT}$ containing extracts. Exemplary HPLC chromatogram of *S. albus* supernatant (a.) and pellet (b.) extracts containing the expression constructs *spi*$_{WT}$. Analytical method: Eurosphere_Ika_lange Methode_AG.Meth.

Absolute production levels of the novel NPs are represented in bar diagrams in Figure 74 for *spi*$_{WT}$ constructs and Figure 75 for *spi*$_{MOD}$ constructs. The results show that a novel NP and also a possibly hydroxylated derivative are biosynthesized in the heterologous host for both arrangements of the *spi* BGC. The expression constructs harboring the *spi* cluster in the arrangement found in the natural producer enables the detection of novel compounds in the pellet as well as in the supernatant fraction (Figure 74), whereas the modified order of genes renders only a detection of novel compounds in the pellet fraction possible (Figure 75). This might be the consequence of a lower total amount of NP produced in heterologous hosts containing the *spi*$_{MOD}$ expression constructs. However, when comparing the total absorption values of the novel NPs it can be detected that the compound (**spi-1**) with *m/z* = 495 $[M+H]^+$ is produced in higher amounts in heterologous host harboring the *spi*$_{WT}$ constructs, whereas the compound (**spi-2**) with *m/z* = 479 $[M+H]^+$ yields higher amounts in heterologous cultures with the *spi*$_{MOD}$ constructs. However, the overall yield obtained with the heterologous expression of *spi*$_{MOD}$ constructs is unsatisfactory compared to *spi*$_{WT}$ cultures, accordingly *spi*$_{WT}$ constructs were chosen for large scale expression. When comparing the yields of novel NP according to the usage of the different promoters, *SF14P* produced the highest yield. The different results for the promoter usage, as well as the differences of cluster arrangement and cultivation time, compared to the ikarugamycin derivatives underlines the discrepancy of the compound according to stability and biosynthesis. As a cultivation time of 3 d represents the optimal period before compound degradation, the stability of the compound is significantly less

compared to **39** which allows an accumulation of NP before extraction of the heterologous culture (chapter 2.1.3).

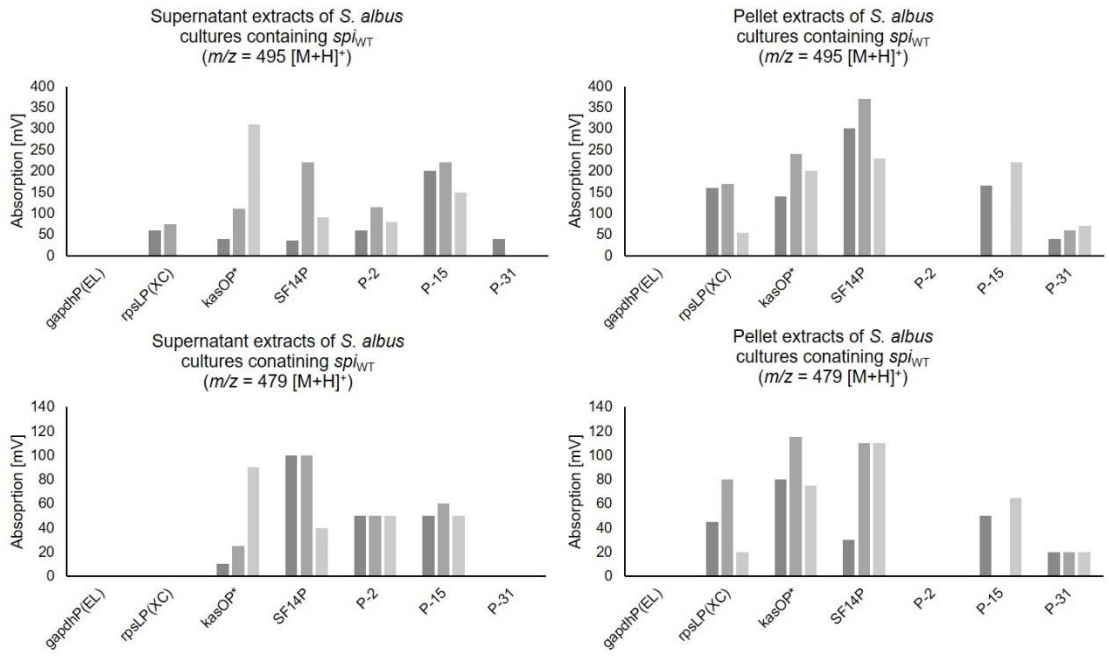


Figure 74. Bar diagrams of HPLC absorption for *spi*$_{WT}$ containing expression cultures. Depicted are the absolute absorption values measured in extracts of *S. albus* strains harboring *spi*$_{WT}$ expression constructs. Absorption of *m/z* = 479 $[M+H]^+$ peak in supernatant (a.) and pellet (b.) and m/z = 495 $[M+H]^+$ peak in supernatant (c.) and pellet (d.).

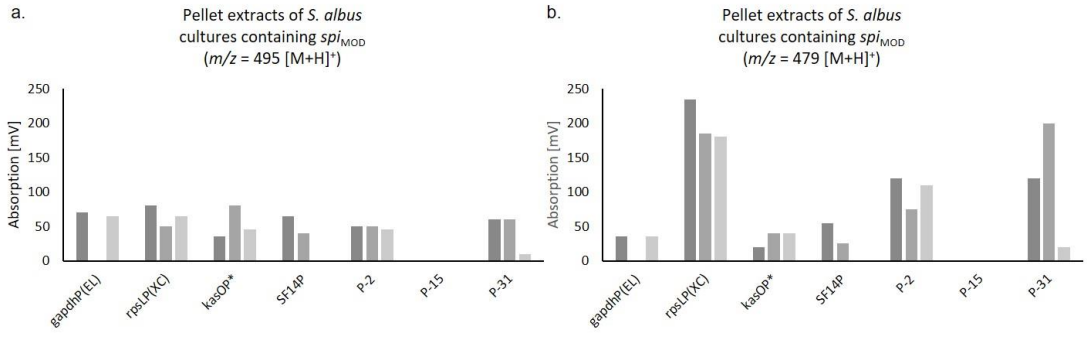


Figure 75. Bar diagrams of HPLC absorption for *spi*$_{MOD}$ containing expression cultures. Depicted are the absolute absorption values measured in extracts of *S. albus* strains harboring *spi*$_{MOD}$ expression constructs. Absorption of *m/z* = 495 $[M+H]^+$ peak in the pellet fraction (a.) and *m/z* = 495 $[M+H]^+$ peak the pellet fraction (b.).

The results represent a novel successful application of the plug-and-play system for PoTeMs, it underlines the efficiency of the cloning strategy with even an additional option for rearranging the cluster. It also proved that the genetic arrangement of the PoTeM BGCs can affect the production titers. Furthermore, it underlines the possibility to do combinatorial biosynthesis with PoTeM BGC to obtain novel compounds. This strategy can be used to shed further light on the single biosynthetic pathways and offers the possibility to design hybrid unnatural compounds.

The next goal was to purify the novel compounds to conduct structural elucidation. As done before for the ikarugamycin derivatives, **55**-**56**, **60**, and **63**-**64**, large scale cultures were inoculated using several 50 ml cultures. Cell pellet and supernatant were extracted, the dried extract was dissolved in methanol and filtered to be injected into the semi preparative HPLC. For isolation the Eurosphere C8 column (Knauer) (12GE084E2J, 125 x 8 mm, 100-5 C8 A) was initially utilized using the method depicted in Table 40. However, the separation of the compounds was not satisfactory. Therefore, the column was changed to the larger Eurosphere C8 column (Knauer) (25GE084E2J, 250 x 8 mm, 100-5 C8 A) and ran with the gradient shown in Table 41. The isolated fractions contained a slightly yellow compound, which could be confirmed to be the desired product by HPLC. However, when removing the acetonitrile, the solution changed the color from yellow to colorless, and a subsequent NMR analysis of the lyophilized white powder underlined the assumption that the compound degraded during acetonitrile removal. Further dilution of the isolated fraction with H_2O, lowering of the temperature during rotary evaporation and removal of TFA did not render isolation of the compound possible.

The future goal to complete this study will be the structure elucidation of the novel compound. Considering the lack of stability of the compound it might be interesting to try a heterologous expression in *S. coelicolor*, as the compound production of **63**, an unstable derivative of **39**, was most reliable in this host. Additionally, a time curve of the NP production in the two hosts *S. albus* and *S. coelicolor* would point out the optimal cultivation period to yield the maximal NP concentration in the extracts. A preliminary purification of the extract using size exclusion chromatography did not result in a better yield, as the compound already degraded on the size exclusion column. However, for compound purification using the semi-preparative HPLC it can be examined whether the exchange of acetonitrile to methanol might improve the stability of the compound during the drying process. When omitting TFA from the solvents during semi-preparative HPLC purification, the peaks widened significantly, rendering a precise separation impossible. In this case, it might be possible to establish a different purification method, which might focus the peaks to give a shape elution profile. If a purification of the compound using standard methods is not possible, it might be tried to do a structural elucidation using HPLC coupled to NMR, a method that only needs minor amounts of the NP. Importantly, removal of the solvents, which in this purification seemed to be the crucial step for degradation, is not necessary for such an approach.

The obtained results confirm the possibility to use the novel plug-and-play system for PoTeM heterologous expression, even for PoTeM BGCs with a different iPKS/NRPS when compared to IkaA. It was shown that several expression vectors with various promoters and different cluster arrangements can be easily cloned using the DiPaC method and heterologous expression can be achieved. Furthermore, this study of a novel PoTeM underlined the

importance of testing different promoters, as the biosynthetic timing and the stability of different PoTeMs can vary and therefore some promoters might be more suitable for certain PoTeMs, whereas others work better for other molecules. It was also shown that the different arrangements of the BGC in the plug-and-play system can result in a different ratio of intermediate to final product, which might already give a hint about the biosynthetic steps involved. Accordingly, the plug-and-play system might also be a useful tool for the elucidation of PoTeM biosynthetic pathways, by omitting single genes or interchanging the order of the different genes.

The two clusters *deg* and *fer*, found in *S. degradans* and *P. fermentans* JBW45, respectively, were cloned into the novel PoTeM heterologous expression system, see Figure 76. The two cluster sequences used for the *spi* cluster were also attempted for these two clusters, leading to expression vectors containing either the wildtype arrangement deg_{WT} and fer_{WT}, Figure 76b and d, or the modified deg_{MOD}, Figure 76c.

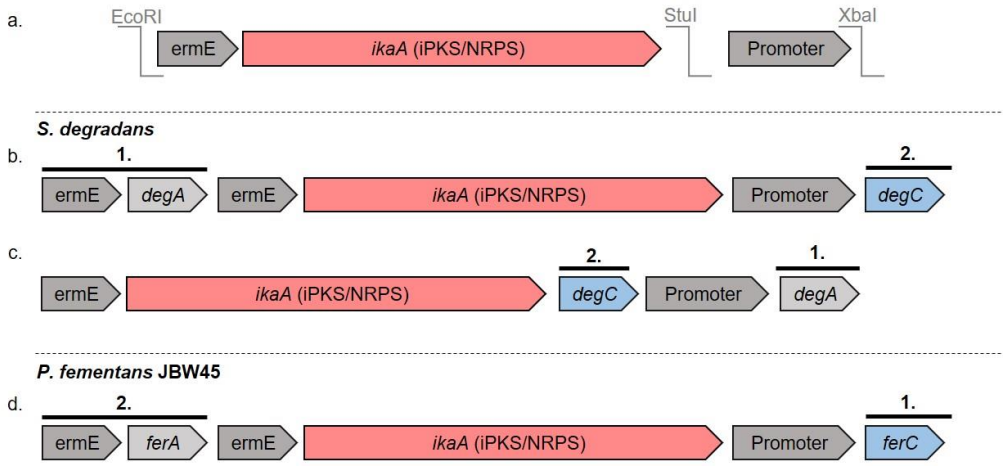


Figure 76. Plug-and-play gene sequences of novel PoTeM cluster. a. Basic plug-and-play expression vector. b./c. Possibilities to rearrange PoTeM cluster *deg* to fit into the novel PoTeM plug-and-play system. d. Possibility used to rearrange PoTeM cluster *fer* to fit into novel PoTeM plug-and-play system. PCR products for cloning are indicated by the bold lines, numbers above indicate cloning order.

Two cloning steps were accomplished to obtain the complete expression vector for each cluster version. Deg_{WT} was created by first integrating the PCR product harboring the *ermE* promoter and *degA* (*deg*_front, *df*) into the EcoRI digested (chapter 4.2.5) and dephosphorylated (chapter 4.2.6) basic plug-and-play vector using SLIC (chapter 4.2.7). Possible positive clones were analyzed using colony PCR (chapter 4.2.13), plasmid DNA was purified (chapter 4.2.4) and cloning success was verified utilizing analytical restriction digestion (chapter 4.2.5) and Sanger sequencing (chapter 4.2.14). Cloning results are depicted in Table 16. The second cloning step to finalize the deg_{WT} expression vectors was conducted using conventional ligation cloning, as this enabled the use of one PCR product to be integrated in all the different expression vectors. The PCR product containing *degC* (*deg*_back, *db*) was

amplified (chapter 4.2.13) and XbaI digested (chapter 4.2.5). Ligation was conducted using the XbaI digested (chapter 4.2.5) and dephosphorylated (chapter 4.2.6) intermediate expression vectors with subsequent heat shock transformation into chemically competent *E. coli* DH5α (chapter 4.2.10). Possible positive clones were screened using colony PCR (chapter 4.2.13), plasmid DNA was purified (chapter 4.2.4) and cloning success was verified by conducting analytical restriction digestion (chapter 4.2.5) and Sanger sequencing (chapter 4.2.14). As displayed in Table 16, cloning of the different deg_{WT} expression vectors could be efficiently accomplished with less than 15 screened clones for any expression construct.

Table 16. Result summary of pSET152_ermE::*df_ikaA*_P(_*db*) cloning. The table shows the correct clones for the two cloning steps to obtain deg_{WT} expression vectors. References to the respective figure in the appendix for result description are assigned.

Expression construct	Correct clone	Reference to results
pSET152_ermE::*df*__ermE_*ikaA*_gapdhP(EL)	Clone 1	Figure A99
pSET152_ermE::*df*_ermE_*ikaA*_gapdhP(EL)_*db*	Clone 2	Figure A100
pSET152_ermE::*df*_ermE_*ikaA*_rpsLP(XC)	Clone 1	Figure A99
pSET152_ermE::*df*_ermE_*ikaA*_rpsLP(XC)_*db*	Clone 1	Figure A100
pSET152_ermE::*df*_ermE_*ikaA*_kasOP*	Clone 1	Figure A99
pSET152_ermE::*df*_ermE_*ikaA*_kasOP*_*db*	Clone 1	Figure A100
pSET152_ermE::*df*_ermE_*ikaA*_SF14P	Clone 2	Figure A99
pSET152_ermE::*df*_ermE_*ikaA*_SF14P_*db*	Clone 1	Figure A100
pSET152_ermE::*df*_ermE_*ikaA*_P-15	Clone 1	Figure A99
pSET152_ermE::*df*_ermE_*ikaA*_P-15_*db*	Clone 14	Figure A100
pSET152_ermE::*df*_ermE_*ikaA*_P-31	Clone 1	Figure A99
pSET152_ermE::*df*_ermE_*ikaA*_P-31_*db*	Clone 14	Figure A100

Access to the deg_{MOD} expression vectors was also generated in a two-step cloning. In a first step *degC* (*db*) was PCR amplified (chapter 4.2.13) and StuI digested (chapter 4.2.5), one single PCR product could be used for conventional ligation cloning with all the six different StuI digested (chapter 4.2.5) and dephosphorylated (chapter 4.2.6) basic plug-and-play vectors. The ligation mixture was transformed into chemically competent *E. coli* DH5α by heat shock (chapter 4.2.10) and possible positive clones were identified using colony PCR (chapter 4.2.13). Plasmid DNA was purified (chapter 4.2.4) and successful cloning was verified using analytical restriction digestion (chapter 4.2.5) and Sanger sequencing (chapter 4.2.14). The second cloning step was conducted using the XbaI digested (chapter 4.2.5) PCR product containing *degA* (*df*) to obtain the finalized deg_{MOD} expression vectors. As shown in Table 17, the expression vector could be cloned for almost all the six different promoters. However, the strategy to integrate two pieces using conventional ligation cloning with only one restriction enzyme for each cloning step significantly reduces the success rate for this cloning approach,

as an integration of the insert in the forward or reverse direction can occur. Less than 10 clones needed to be screened for most of the promoters, whereas up to 33 clones were needed for pSET152_ermE::*ikaA*_*db*_SF14P_*df*. For this study, it was not possible to clone the deg_{MOD} expression vector harboring the *rpsLP*(XC). However, as this was the first proof of principle step to see whether it is possible to obtain a novel PoTeM using the plug-and-play strategy based on the *deg* cluster, sufficient different expression constructs were established for the first heterologous expression experiments.

Table 17. Result summary of pSET152_ermE::*ikaA*_*db*_P(_*df*) cloning. The table shows the correct clones for the two cloning steps to obtain deg_{MOD} expression vectors. References to the respective figure in the appendix for result description are assigned.

Expression construct	Correct clone	Reference to results
pSET152_ermE::*ikaA*_*db*_gapdhP(EL)	Clone 2	Figure A101
pSET152_ermE::*ikaA*_*db*_gapdhP(EL)_*df*	Clone 8	Figure A102
pSET152_ermE::*ikaA*_*db*_rpsLP(XC)	Clone 1	Figure A101
pSET152_ermE::*ikaA*_*db*_rpsLP(XC)_*df*	-	-
pSET152_ermE::*ikaA*_*db*_kasOP*	Clone 1	Figure A101
pSET152_ermE::*ikaA*_*db*_kasOP*_*df*	Clone 23	Figure A102
pSET152_ermE::*ikaA*_*db*_SF14P	Clone 2	Figure A101
pSET152_ermE::*ikaA*_*db*_SF14P_*df*	Clone 33	Figure A102
pSET152_ermE::*ikaA*_*db*_P-15	Clone 1	Figure A101
pSET152_ermE::*ikaA*_*db*_P-15_*df*	Clone 10	Figure A102
pSET152_ermE::*ikaA*_*db*_P-31	Clone 2	Figure A101
pSET152_ermE::*ikaA*_*db*_P-31_*df*	Clone 5	Figure A102

The final constructs were integrated into *S. albus* utilizing horizontal conjugation (chapter 4.2.12) and exconjugants were picked for small scale heterologous expression trials (chapter 4.2.1). Cells were harvested after an incubation of 7 d, supernatant and cell pellet were extracted with organic solvents (chapter 4.2.19) and extracts were dissolved in methanol for LC-MS analysis (chapter 4.2.20). HPLC analysis of the extract depict the formation of new peaks at a retention time of 10 to 12 min (Figure 77a), when extracts were ran with the gradient shown in Table 37. The peak at a retention time of 10.2 min was reproducibly existent in pellet extracts of heterologous cultures harboring the deg_{MOD} expression constructs for all six different promoters. Analysis of the MS data of this peak revealed a mass of m/z = 511.5 $[M+H]^+$ and m/z = 495.5 $[M+H]^+$. Displayed in Figure 77b, the extracted mass of m/z = 511.5 $[M+H]^+$ represents a precise mass peak at this retention time. With m/z = 511.5 $[M+H]^+$ being a possible mass of a PoTeM, this peak might represent the PoTeM encoded by the *deg* cluster. Analysis of the extracts of heterologous expression cultures containing the deg_{WT} expression vectors indicated the same novel peak at a retention time of

10.2 min for the constructs with the promoter *gapdhP*(EL) and *rpsLP*(XC), whereas for the other promoters it was not possible to identify a new peak (data not shown). These results demonstrate that novel PoTeMs-like compounds can be produced for both cluster arrangements of the *deg* cluster.

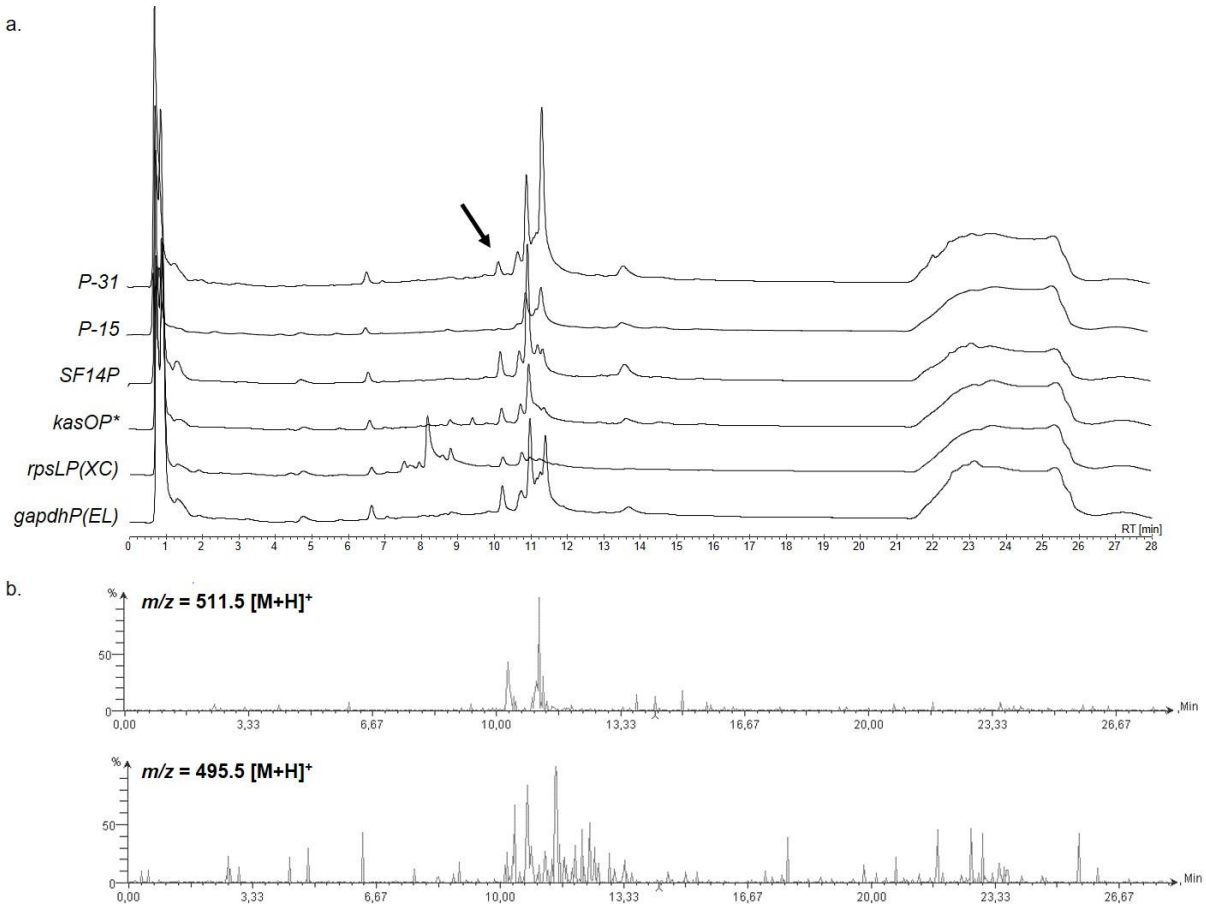


Figure 77. LC-MS analysis of expression cultures harboring *deg*$_{MOD}$ constructs. a. HPCL chromatograms of representative pellet extracts of *S. albus* heterologous cultures containing *deg*$_{MOD}$ expression vectors. b. extracted masses for *m/z* = 511.5 $[M+H]^+$ and *m/z* = 495.5 $[M+H]^+$, both masses are highly abundant in the novel peak at a retention time of 10.2 min. Analysis method: Eurosphere_Ika_lange Methode_AG.Meth.

The yield of detected novel compound was not sufficient for a large-scale production and subsequent purification for structure elucidation. Accordingly, an optimization of heterologous expression conditions was necessary. A first optimization trial was accomplished by generating a time curve using the two expression constructs pSET152_ermE::*ikaA_db*_gapdhP(EL)_*df* and pSET152_ermE::*ikaA_db*_SF14P_*df*. Both time curves were generated by extracting the heterologous expression cultures after 3 d, 5 d, and 7 d. For unknown reasons, it was not possible to detect any NP in the extracts harboring the expression plasmid containing the *gapdhP*(EL) promoter (data not shown). The results for extracts harboring *SF14P* containing plasmids are depicted in Figure 78.

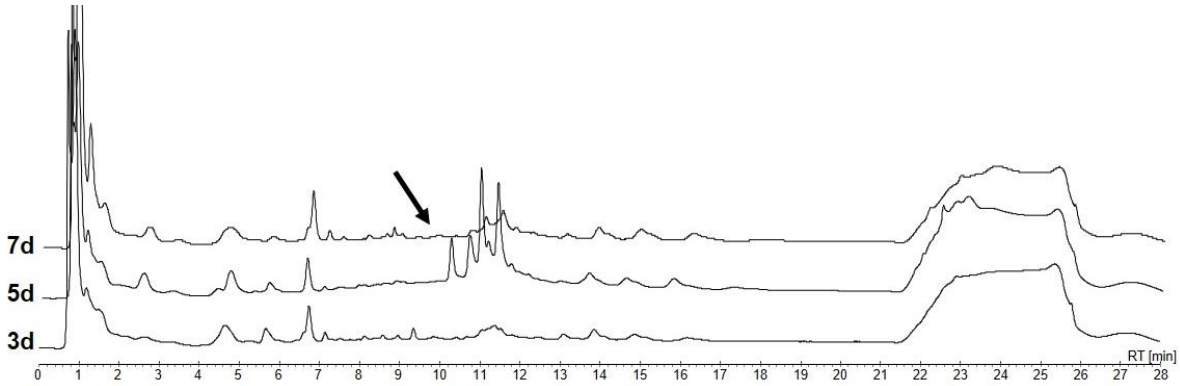


Figure 78. Time course heterologous expression pSET152_ermE::*ikaA*_*db*_SF14P_*df*. Comparison of *S. albus* heterologous expression cultures pellet extracts harboring pSET152_ermE::*ikaA*_*db*_SF14P_*df* for a 3 d, 5 d, and 7 d incubation period. Analytical method: Eurosphere_Ika_lange Methode_AG.Meth.

It was obvious that no NP was detectable after a 3 d and after a 7 d incubation, remarkably, as the first expression trial, results shown in Figure 77, were obtained using the standard incubation time of 7 d. These results imply that the heterologous host might silence the integrated cluster, which would make a novel conjugation of the expression plasmid into the heterologous host necessary before starting an expression culture. Furthermore, a novel conjugation of the plasmid into the heterologous host before culturing the *Streptomyces* was shown to be beneficious for optimal NP production (chapter 2.3.2). However, the low amount of NP detected in the extracts of the first expression might not be compensated by optimizing culturing conditions. As discussed before (chapter 2.4), the genetic origin of *S. albus* and *S. degradans* is rather distant, resulting in a crucial difference in GC content and codon usage, already obvious with the usage of a GUG start codon. Before optimizing the original cluster in a host that does not frequently use the *deg* clusters naturally available codons, it might be interesting to codon-optimize the two original *deg* cluster parts to examine if this increases the amount of produced NP and if this results in reproducible yields of novel compound. If so, general culturing condition optimizations can be conducted to maximize compound production in the heterologous host, considering factors such as incubation time, media, use of a wire spiral or changing the heterologous host to a strain with cluster knockouts. Within the scope of this study, a further optimization was not conducted. However, the preliminary results are promising to enable the identification of the PoTeM encoded in the *deg* cluster.

The results obtained in expressing the previous PoTeM clusters were considered to reduce the cloning effort for constructing the *fer* cluster expression vectors based on the plug-and-play system (Figure 76). As mostly successful, the cluster sequence corresponding to the gene order found in the natural producer was selected as the order for the plug-and-play expression vectors. Furthermore, the promoter selection was reduced to the four promoters *gapdhP*(EL), *rpsLP*(XC), *SF14P*, and *P-2*.

In contrast to the other PoTeM clusters, an expression construct harboring *ferA* (*fer_front*, *ff*) behind the *ermE* promoter was not available, therefore, an intermediate plasmid providing ermE_*ff* as a PCR template for the plug-and-play cloning strategy was designed and cloned. Accordingly, *ferC* (*fer_back*, *fb*) was integrated first into the basic plug-and-play constructs. the back cluster part *fb* was PCR amplified (chapter 4.2.13) with homology overhangs for a SLIC cloning strategy (chapter 4.2.7) and integrated into the XbaI digested (chaper 4.2.5) and dephosphorylated (chapter 4.2.6) basic plug-and-play vectors. The mixture was transformed into chemically competent *E. coli* DH5α cells by heat shock transformation (chapter 4.2.10). Possible positive clones were screened using colony PCR (chapter 4.2.13), plasmid DNA was purified (chapter 4.2.4) and verified using analytical restriction digestions (chapter 4.2.5) and Sanger sequencing (chapter 4.2.14).

For the second cloning step, ermE_*ff* was PCR amplified (chapter 1404.2.13) using the intermediate construct and integrated into the EcoRI digested (chapter 4.2.5) and dephosphorylated (chapter 4.2.6) intermediate plug-and-play expression vector using SLIC (chapter 4.2.7). The ligation mixture was transformed by heat shock into chemically competent *E. coli* DH5α cells (chapter 4.2.10). Possible positive clones were selected using colony PCR (chapter 4.2.13), plasmid DNA was purified and verified using analytical restriction digestion (chapter 4.2.5) and Sanger sequencing (chapter 4.2.14). Results for the cloning can be found in Table 18.

Table 18. Result summary of pSET152_ermE::(*ff*_ermE_)*ikaA*_ P_*fb* cloning. The table shows the correct clones for cloning *fer* cluster plug-and-play expression constructs. References to the respective figure in the appendix for result description are assigned.

Expression construct	Correct clone	Reference to results
pSET152_ermE..*ikaA*_gapdhP(EL)_*fb*	Clone 1	Figure A 103
pSET152_ermE::*ff*_ermE_*ikaA*_gapdhP(EL)_*fb*	Clone 9	Figure A 104
pSET152_ermE::*ikaA*_rpsLP(XC)_fb	Clone 1	Figure A 103
pSET152_ermE::*ff*_ermE_*ikaA*_rpsLP(XC)_*fb*	Clone 5	Figure A 104
pSET152_ermE::*ikaA*_SF14P_*fb*	Clone 1	Figure A 103
pSET152_ermE::*ff*_ermE_*ikaA*_SF14P_*fb*	Clone 3	Figure A 104
pSET152_ermE::*ikaA*_P-2_*fb*	Clone 1	Figure A 103
pSET152_ermE::*ff*_ermE_*ikaA*_P-2_*fb*	Clone 1	Figure A 104

The *fer* cluster expression plasmids were integrated into the heterologous host *S. albus* using *Streptomyces* conjugation (chapter 4.2.12) and incubated in 50 ml cultures to get a first result of the possibility to express the compound encoded by the *fer* cluster. As depicted in Figure 79, the expression yielded a similar result when compared to the *deg* cluster, producing new peaks within a retention time between 10 to 12 min. The same masses, particularly

m/z = 511.5 $[M+H]^+$, can be found, when analyzing the mass data of the LC-MS analysis (data not shown). This might hint at a general mistake that occurred during the expression. However, it might also be possible that the same compound is encoded by both clusters, *deg* and *fer*. The peaks were not present in *S. albus* expression cultures without an expression plasmid, which indicates that they originate from the presence of the plug-and-play expression vectors. As both clusters are similarly small, consisting of only three genes, one encoding for the iPKS/NRPS and one up and one downstream of the iPKS/NRPS, the possibilities to process the precursor **53** are limited. A fact that could argue for the same compound being encoded by the two different PoTeM clusters *deg* and *fer*. However, alignments of the nucleotide sequences or the amino acid sequences did not reveal significant similarities between *degA*/*ferA* of *degC*/*ferC*. Accordingly, a large-scale purification with subsequent structure elucidation of the compound represented in the novel peaks is necessary to confirm the assumption.

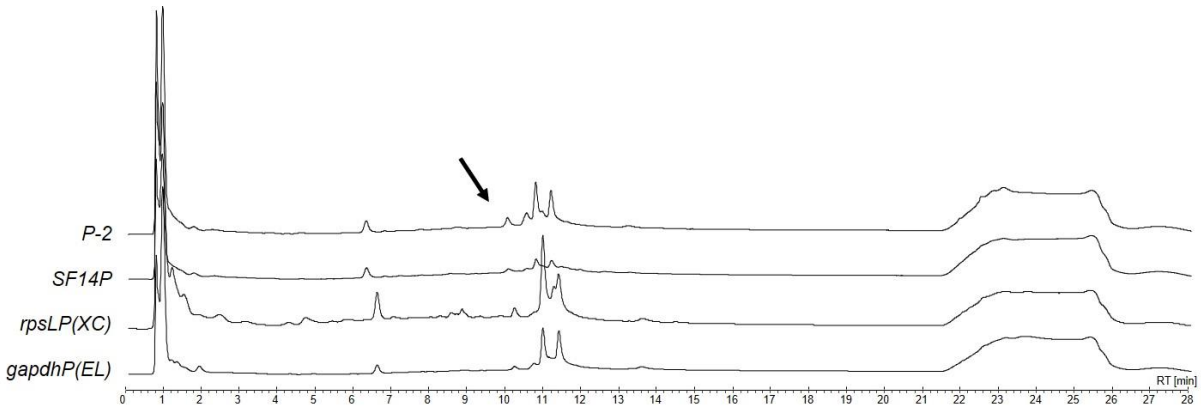


Figure 79. HPLC analysis of expression cultures harboring *fer* cluster. Comparison of representative *S. albus* heterologous expression cultures pellet extracts for the different promoter incubated for 7 d. Analytical method: Eurosphere_Ika_lange Methode_AG.Meth.

The present heterologous expression results of the *fer* cluster preclude large scale purification, as the compound titers were too low. An optimization of the expression is needed to yield compound amounts sufficient for NMR analysis for structure elucidation. DNA sequence analysis of *ferA* and *ferC* highlight the genetic difference between the heterologous host *S. albus* and the original producer *P. fermentans*. The 855 bases comprising *ferA* hold a GC content of 35% and the 1554 bases of *ferC* are with a GC content of 38% far below the average GC content of *S. albus*. Furthermore, the TTA codon is present in both genes, *ferA* and *ferC*, a codon that is rarely used in *Streptomyces*.[140] As the genetic difference between heterologous host and original producer should be kept as small as possible to prevent obstacles in heterologous expression, such as codon usage, heterologous expression of the *fer* cluster might be already hindered by this general fact. Codon optimization of the two genes for expression in *S. albus* would be a first possibility to enhance heterologous expression for

compound identification. Furthermore, an optimization of the culture conditions as extensively conducted in chapter 2.1.3 for **39**, considering culturing time, media, and culture conditions, can be examined. The reduction to the four promoters also depicts a suitable selection, as these all have been demonstrated to enhance PoTeM expression in *Streptomyces*. Promoter *gapdhP*(EL) and *rpsLP*(XC) proved to be suitable candidates for supporting the expression of high compound amounts for stable products. Promoters *P-2* and *SF14P* showed decent NP production enhancement of less stable compounds, whereas *SF14P* showed best results for compounds that are produced at an early culturing state. This holds also true for *kasOP**, a promoter that needs to be considered when using the heterologous host *S. coelicolor* as consistent high yielding results were obtained utilizing *kasOP** in *S. coelicolor*.

The gained results of utilizing the novel plug-and-play system for PoTeM BGC of yet unknown compounds facilitated a large advance towards identifying novel chemistry. It was shown that all three different clusters *spi*, *deg*, and *fer* can easily be cloned and rearranged into the novel expression vector based on *ikaA* and different *Streptomyces* promoters. Heterologous expression of the clusters in *S. albus* resulted in the appearance of novel peaks during LC-MS analysis with suitable masses for PoTeMs. However, it was not yet possible to structurally elucidate the compounds comprising those new peaks. Extensive studies enabled a sufficient production of the *spi* encoded products, with large difficulties during purification caused by the low stability of the product in solvents, an issue that needs to be solved for compound identification. The amount of novel compound present in heterologous expression extracts harboring the *deg* or *fer* cluster are yet too low for compound purification. In this case it might be necessary to codon-optimize the accessory genes of the clusters for the heterologous expression host *S. albus*, as the GC-contents might be too different for successful expression. The comparison of the promoters craved out four promoters most suitable for heterologous expression and cloning strategies in *S. albus*, *gapdhP*(EL), *rpsLP*(XC), *SF14P* and *P-2*, and one further promoter most suitable for expression in *S. coelicolor*, *kasOP**. This set of promoters covers the expression enhancement for most kinds of PoTeM compound characteristics, like stability issues or biosynthetic timing. Accordingly, this novel plug-and-play strategy for PoTeM heterologous expression represents an efficient and flexible approach to clone novel, known, or rearranged PoTeM BGCs and yield compound expression mostly during the first expression trial. However, general heterologous expression issues, as genetic differences and culturing condition, still need to be considered when using the novel expression constructs for PoTeM heterologous expression.

2.6 Crystallization of IkaC

The biosynthesis of ikarugamycin (**39**) was extensively studied, whereby the different intermediates have been elucidated and a biosynthetic route to **39** has been established. A significant advance towards this goal was provided by Greunke *et al*.[126] who proved all enzymes crucial for **39** biosynthesis to be efficiently and soluble expressed in *E. coli*. These findings enable mechanistic analysis of the biosynthetic processes. The last enzymatic reaction in the biosynthesis of **39** is the inner ring closure catalyzed by the NADPH-dependent alcohol dehydrogenase IkaC.[113] As IkaC was easily obtainable in sufficient amounts through recombinant protein purification, we aimed at obtaining insight into its molecular mechanism by structural biology. A first crystal structure was obtained by C. Greunke (protein purification) and E. Duell in collaboration with the working group of S. Schneider prior to this study (Figure 80). Crystals were obtained using sitting drop vapor diffusion conducted by C. Scheidler.

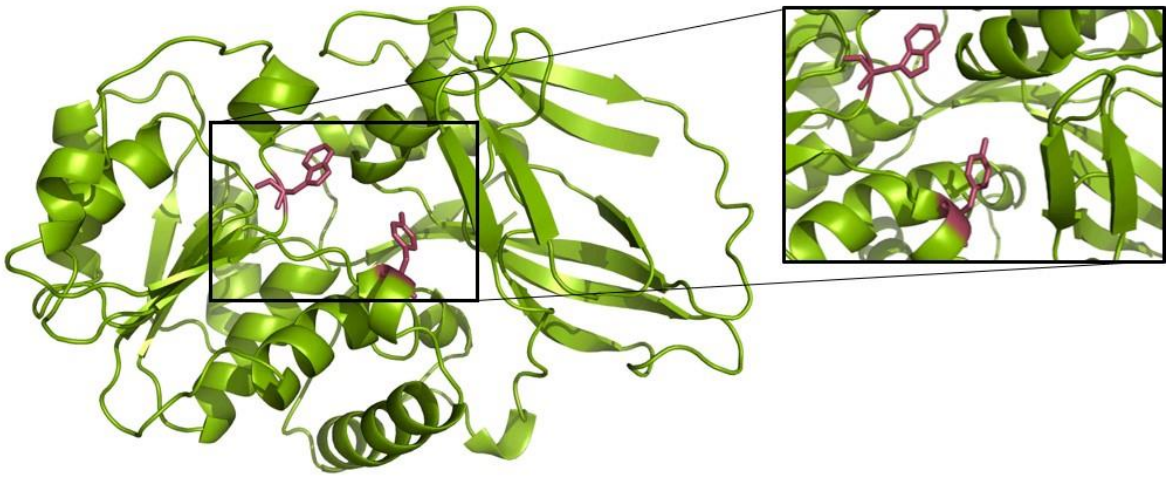


Figure 80. Crystal structure of IkaC. Amino acid residues W260 and Y292 are accentuated in magenta.

According to these primary results, two approaches to gain a greater knowledge of this enzyme were conducted. First, the amino acids that might be crucial for enzymatic reactivity, determined by the already obtained crystal structure (Figure 80, amino acid residues are depicted in magenta), should be mutated to change the chemical environment and thereby amend the biochemical activity of IkaC. Additionally, a co-crystallization of IkaC with **54** to analyze the positioning of the substrate in the enzyme was aimed for. With a greater knowledge on how the substrate is bound in the enzyme and which amino acids are in proximity, a postulation of the mechanism of action for IkaC could be developed. This would for the first time shed light on the molecular mechanism of an accessory enzyme found in a PoTeM

biosynthesis and might render a greater understanding of the molecular processes underlying the biosynthesis.

The enzyme IkaC was expressed in *E. coli* BL21 (chapter 4.2.16) using an expression plasmid generated by C. Greunke (pHis8::SA_AlcD) and purified as published (chapter 4.2.17).[126] In parallel, the substrate (**54**) was heterologously expressed and purified as established in chapter 2.3.2. To prove the functionality of the enzyme, an assay with heterologously expressed **54** was conducted. Therefore, conditions similar to the published enzymatic assay were tested,[113, 126] see Table 34. Due to the minor amount of compound, analysis of the assays was performed using the LCQ-Fleet system (chapter 4.2.20). Retention times of the substrate **54** and the product **39** are depicted in Figure 81. The graphs of the respective mass traces are shown, with *m/z* = 477.5 $[M+H]^+$ for the substrate (**54**) and *m/z* = 479.5 $[M+H]^+$ for the product (**39**).

As visible in Figure 82, **54** is converted to **39** catalyzed by IkaC under any chosen enzymatic assay condition. Both buffers, NaH_2PO_4 and HEPES, as well as both co-factors, NADPH and NADH, show a significant peak in the product mass trace. Accordingly, any of the applied conditions can be used to prove the functionality of the enzyme IkaC. Assays conducted lacking the enzyme showed a peak of the product only in background noise amounts, see Figure A112 and Figure A113.

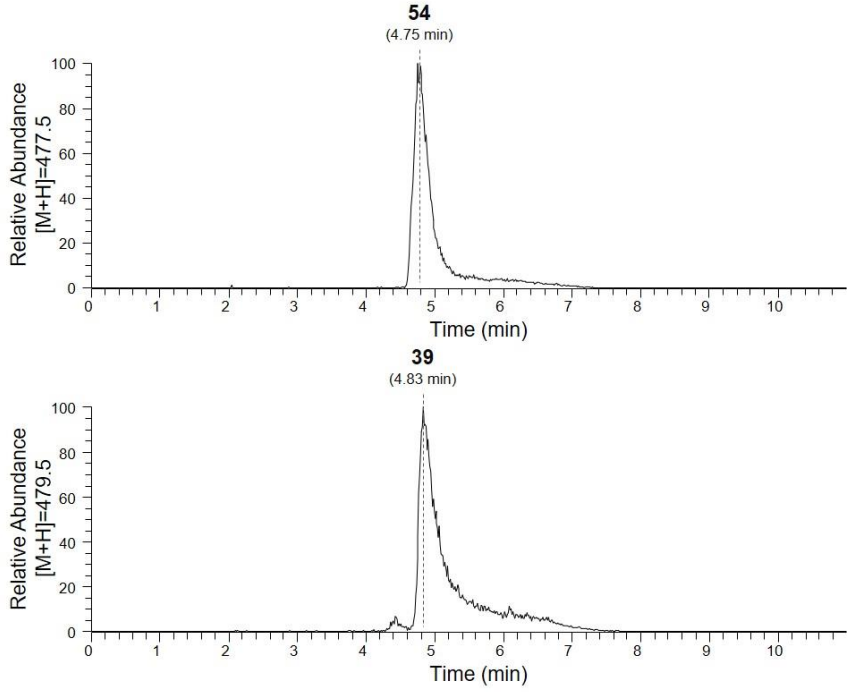


Figure 81. MS chromatograms of substrate and product for IkaC assays. Depicted are the LCQ-fleet chromatograms of the heterologously expressed substrate **54** dissolved in methanol (top) and the heterologously expressed product **39** dissolved in methanol (bottom). Analytical method: LCQ-Fleet.Meth.

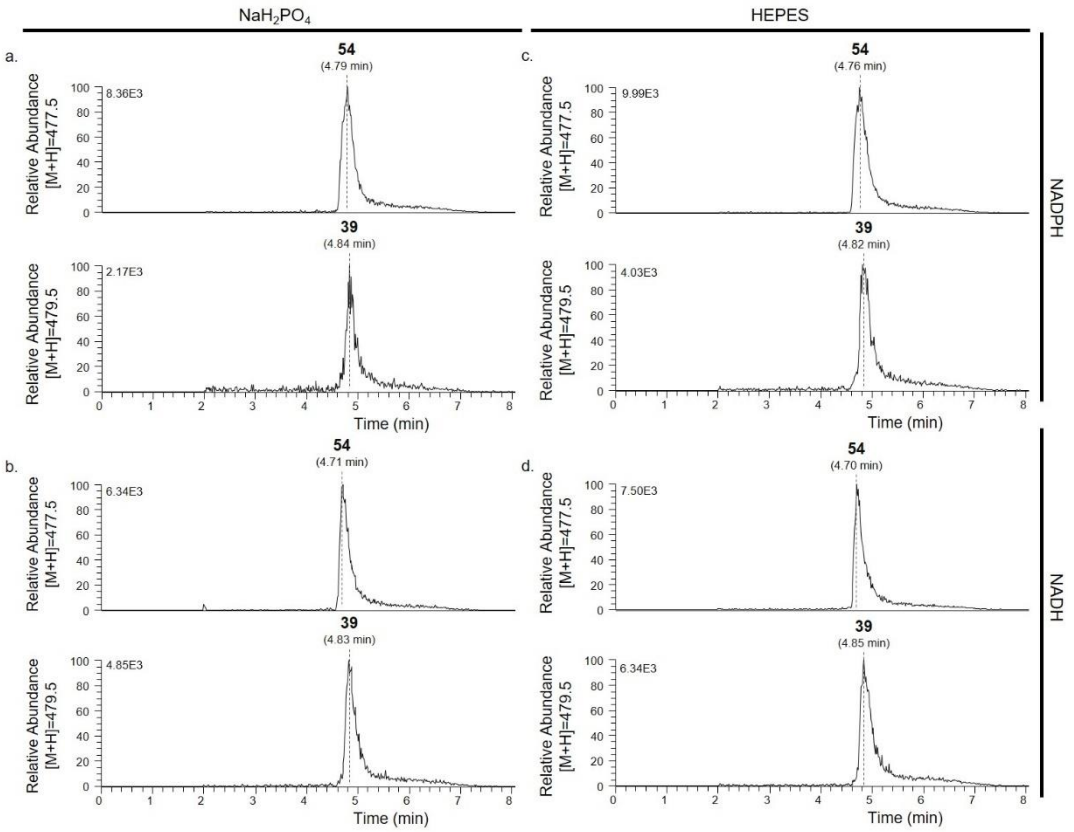


Figure 82. MS results of IkaC assay. a. MS results of IkaC assay performed in NaH_2PO_4 buffer with NADPH, extracted mass for **54** at the top and for **39** at the bottom. b. MS results of IkaC assay performed in NaH_2PO_4 buffer with NADH, extracted mass for **54** at the top and for **39** at the bottom. c. MS results of IkaC assay performed in HEPES buffer with NADPH, extracted mass for **54** at the top and for **39** at the bottom. d. MS results of IkaC assay performed in HEPES buffer with NADH, extracted mass for **54** at the top and for **39** at the bottom. Analytical method: LCQ-Fleet.Meth.

Preparation of IkaC for crystallization was performed by an additional FLPC purification step after recombinant expression, which was conducted by the structural biologist in the group of Dr. Sabine Schneider (TU Munich, now LMU). First crystallization results showed a common alcohol dehydrogenase structure. It included a flexible loop, whose structure could not be determined.

Further work on this project included a trial of co-crystallizing IkaC with the heterologously expressed substrate **54** to enable structure elucidation of the flexible loop. Additionally, two amino acid residues were proposed to be crucial for substrate binding, tryptophan at position 260 (W260) and phenylalanine at position 292 (Y292) (Figure 80, amino acids are depicted in magenta). To better understand the mechanism underlying the last step of the biosynthesis of **39**, both residues were mutated to alanine. Furthermore, Y292 seemed to constitute a gate for substrate binding, therefore this residue was additionally mutated to arginine, a bulky and basic amino acid, to lock the gate and inhibit substrate binding.

The mutations were introduced using site directed mutagenesis (SDM) (chapter 4.2.9). Temperature gradient PCR (chapter 4.2.13) of the three mutant vectors showed strong amplification (Figure 83a), after gel extraction (chapter 4.2.4), KLD reaction (chapter 4.2.9), transformation (chapter 4.2.10) and DNA plasmid preparation (chapter 4.2.4), a positive clone was obtained for all the three mutations, verified using Sanger sequencing (chapter 4.2.14). Results of the sequencing are displayed in Figure 83b.

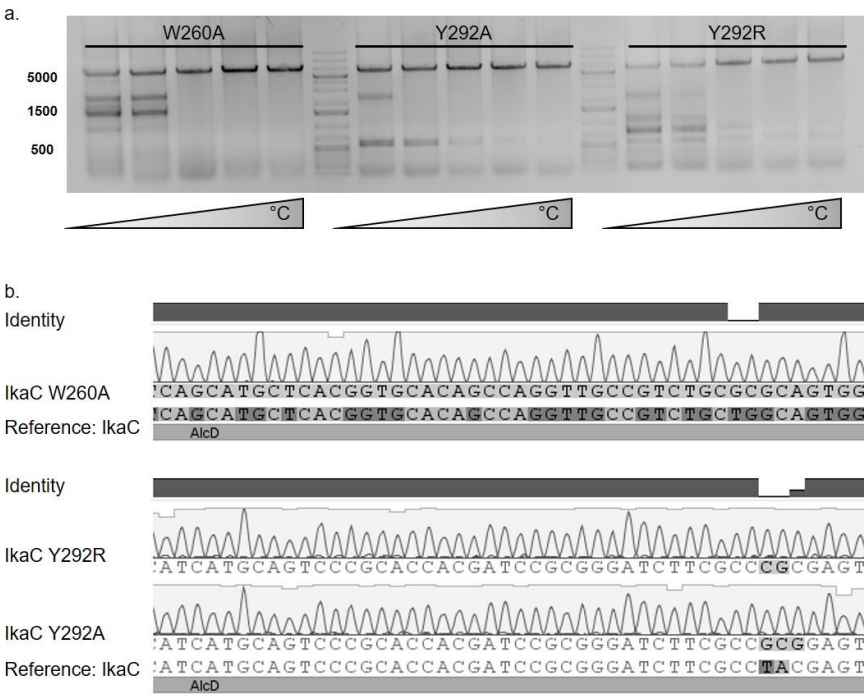


Figure 83. Site directed mutagenesis of IkaC. a. Result of temperature gradient PCR as template for KLD reaction. b. Sequencing results of the three mutant constructs aligned to the wildtype IkaC sequence.

The three mutant enzymes were recombinantly expressed and purified as described above (chapter 4.2.17) to be tested for functionality in an enzymatic assay. Using the experience gained in the first enzymatic assay HEPES buffer and the cheaper co-factor NADH were utilized. The results of the assays are depicted in Figure 84 and confirm a functionality of all mutants. According to this continued functionality of the enzyme, both amino acid residues are not decisive for the enzymatic activity. However, they might still have an important role in the reaction catalyzed by IkaC. It might be possible that the exchange of a single amino acids does not completely abolish the catalytical activity, as the enzymatic assay conducted only allows a qualitative determination of enzyme activity, whereas a quantitative analysis was not possible. As Y292 was proposed as a possible gate to enable substrate specificity and keep the substrate bound in the catalytical pocket, the exchange to the spatially small amino acid alanine might only lead to a reduced specificity and a reduced affinity of the substrate. However, the conducted assay is only a quantitative assay, proving a functional enzyme. To

analyze specificity and substrate affinity, assays with derivatives of **54** and kinetic studies must be conducted. Both options turned out difficult, as the amounts of substrate gained from the heterologous expression are minor. The activity of the Y292R mutant further hints that tyrosine at position 292 does not function as a gate. However, a mutation to proline could give a further hint whether Y292 is involved in the structural constitution of the catalytical pocket, as proline residues greatly interfere in the secondary structures of polypeptide chains. Similar to the Y292A mutation, the role of W260 in substrate binging cannot be disproven by the activity of the W260A mutant. Binding of the substrate might still be enabled by other residues found in the catalytical pocket. Accordingly, double mutations of several residues in the catalytical pocket might be required to identify all amino acid residues necessary for substrate binding.

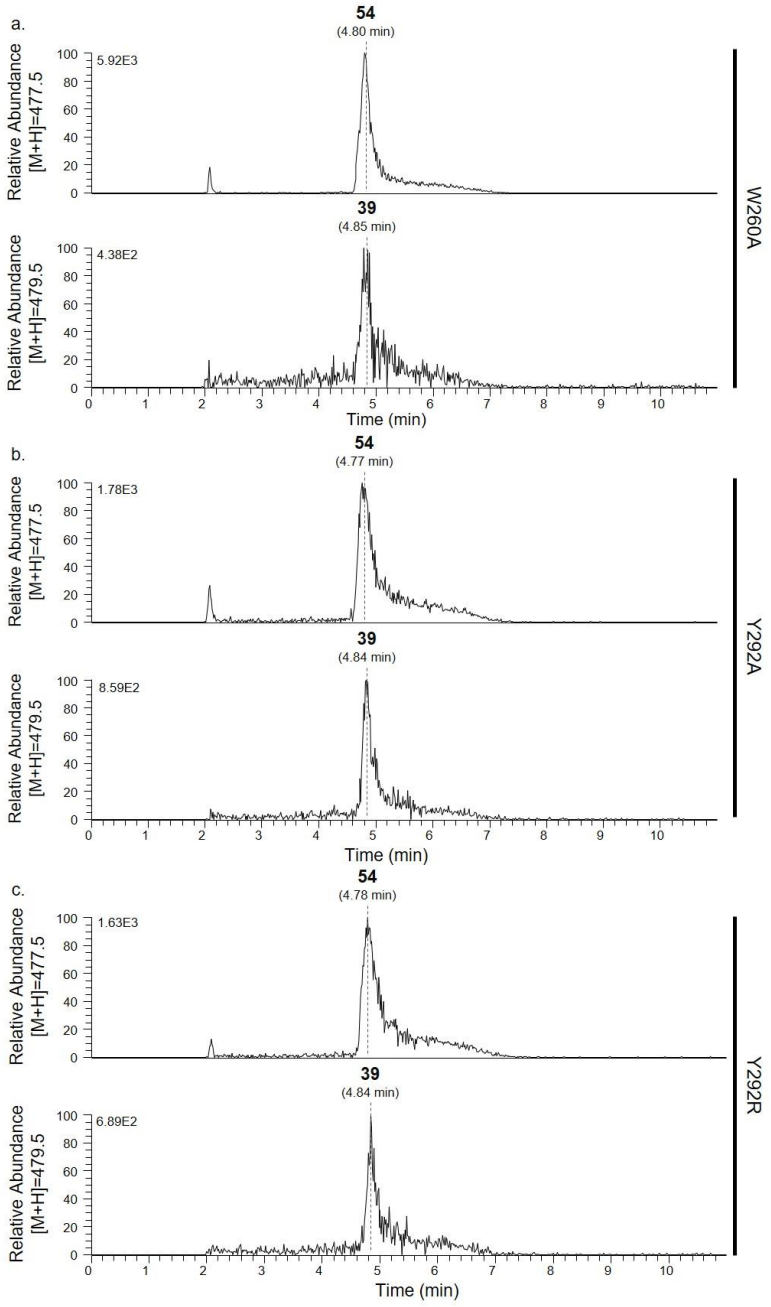


Figure 84. MS results of functional assays using IkaC mutants. MS graph for the extracted mass of the educt **54** (above) and product **39** (below) of functional assay conducted with the IkaC mutant W260A (a.), Y292A (b.), and Y292R (c.). Analytical method: LCQ-Fleet.Meth.

A different possibility to identify amino acids involved in substrate binding and catalysis of IkaC might be a co-crystallization with the substrate and the co-factor. In this case, the enzyme, the substrate, and the co-factor are incubated at the optimal conditions for crystallization. As this might include conditions in which the enzyme is active, a reduced form of NADH, NADH*, was prepared by J. Evers. This reduced co-factor could be proven to abolish catalytical activity of IkaC, see Figure 85. Accordingly, crystallization batches containing all the components in different ratios were prepared by the co-workers in the Schneider group.

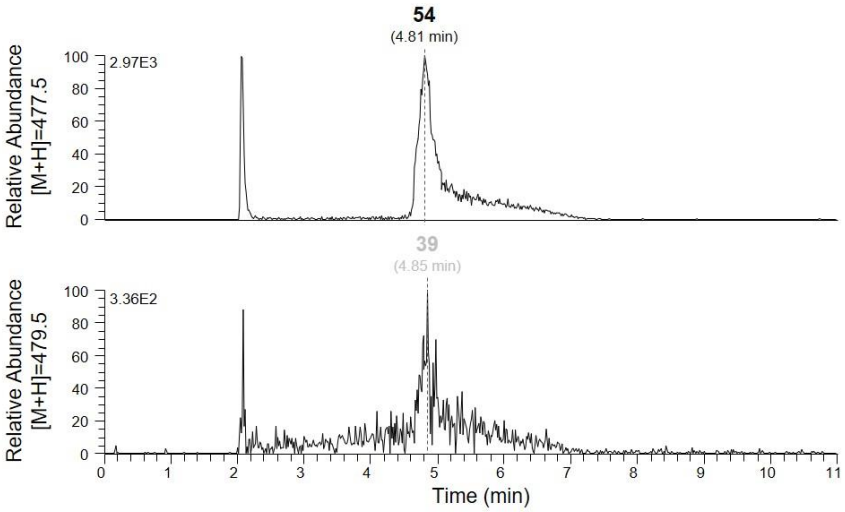


Figure 85. MS results of functional IkaC assay with reduced NADH*. MS graphs for the extracted masses of **54** (above) and **39** (below) are depicted. Analytical method: LCQ-Fleet.Meth.

When analyzing the results of the x-ray data, it became obvious that neither the substrate nor the co-factor was bound in the crystals, only components of the buffers used to support crystallization could be found co-crystallized. Buffer conditions that resulted best crystal growth were 0.1 M potassium thiocyanate, 30% PEG MME 200. To exclude that the amount of **54** and reduced NADH was small minor for co-crystallization, several ratios of the two components were applied to the purified protein. However, the amounts needed to be low enough to ensure solubility of the protein. With these changed conditions, the resulting crystals still showed a co-crystallization of IkaC with a buffer component. Accordingly, it was not possible to obtain a crystal structure of IkaC bound with the substrate **54** to further understand the mechanism underlying the enzymatic reaction.

Following these results, it was not possible using the attempted approaches to gain greater knowledge on the structural organization of IkaC than found in the primary crystallization trials. Further trials might include double mutations of different possibly crucial amino acid residues or changing single amino acids to proline to disrupt α-helical structures. Furthermore, additional crystallization conditions for a co-crystallization with **54** and reduced NADH could be tested, excluding those already proven to co-crystallize with buffer components as mentioned above.

3 Summary

Natural products are secondary metabolites that can increase the fitness of their native producers, such as protection against predators. These natural characteristics can often be used in the pharmacological industry to develop novel drug lead structures. In times of the antibiotic crisis, the discovery of novel natural compounds has therefore gained increasing importance. Classicall methods for the discovery of novel NPs, such as fermentation and extraction with organic solvents was very successful in the past. However, reisolation of known compounds and culturing issues of so called 'unculturable' bacteria has become a great burden to these methods. The progress in sequencing techniques enables a time and cost-efficient sequencing of whole bacterial genomes with small amounts of genomic DNA. Accordingly, novel methods have been developed to capture biosynthetic gene clusters and express the encoded compounds in heterologous hosts.

This thesis is focused on the natural product class of polycyclic tetramate macrolactams (PoTeMs), molecules exhibiting pharmacological properties ranging from antibacterial (**39**)[90] to antifungal (**44**)[102] and even cytotoxic function(**41**).[98] The particular interest in this NP class is furthermore due to the biosynthetic characteristic of an uncommon iPKS/NRPS enzyme that synthesizes one common precursor (**53**) for all known compounds of this class.[108] This enables the possibility to find novel BGCs encoding for PoTeMs by the alignment of the iPKS/NRPS sequence to collections of bacterial genomes. This thesis aimed to establish methods for the discovery of novel PoTeMs, whose BGCs have already been identified by bioinformatic analysis. Furthermore, a novel strategy for gene cluster capturing was to be developed to enable more efficient cloning of expression plasmid harboring complete BGCs.

In a first step, the heterologous production of ikarugamycin (**39**) was optimized. Previous studies have identified the *ika* BGC to encode the biosynthetic enzymes for the biosynthesis of **39**.[112] However, the production titers when performing heterologous expression in *E. coli* was insufficient for any downstream application. Thus, heterologous expression in this study was performed in *Streptomyces*, bacteria which are known to efficiently express secondary metabolites. The *ika* cluster was successfully inserted into several *Streptomyces* expression vectors and integrated in different strains of this heterologous host. Cloning was accomplished using the novel 'Direct Pathway Cloning' (DiPaC) method based on the PCR amplification of the whole BGC with subsequent integration into the desired expression vector using Gibson assembly.[130] pSET152_ermE::*ika* showed best expression results in *S. albus* DSM40313 (Figure 86). Further optimization of the culturing conditions facilitated a final yield of 25 mg/L of heterologously expressed and purified product **39**. The purified compound was utilized as analytical standard in the biocatalytic total synthesis of ikarugamycin.[126]

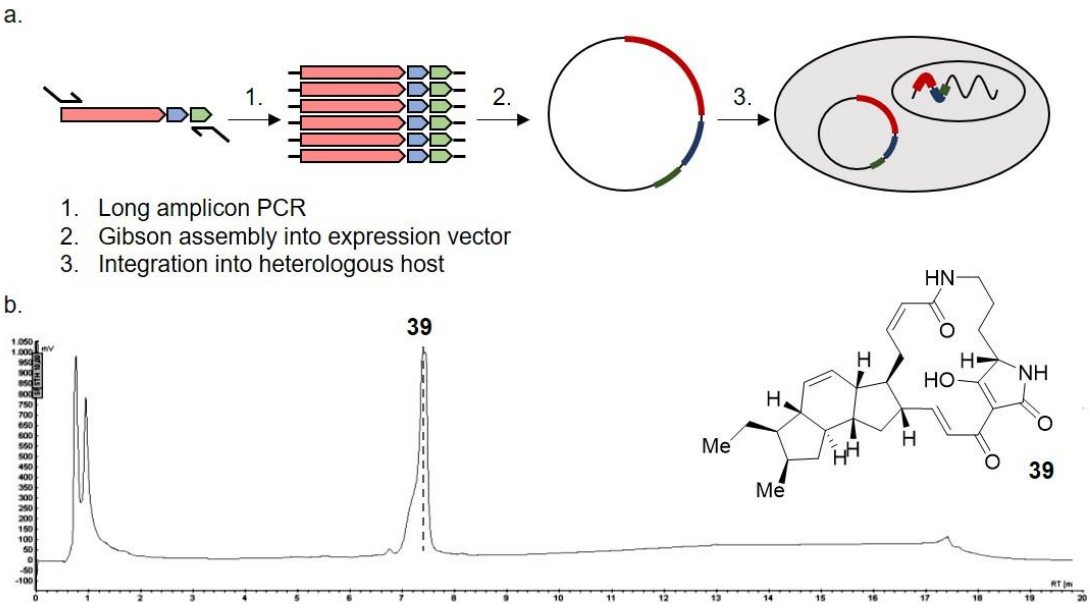


Figure 86. Heterologous expression and production of ikarugamycin (**39**). a. Schematic representation of cloning the *ika* BGC into expression vectors using the DiPaC method and integration into the heterologous host. b. HPLC chromatogram of optimized production of **39**. Analytical method: Eurosphere_IkaFAST_CG.Meth.

The knowledge gained for the expression of **39** was further applied to the heterologous production of the biosynthetic intermediates of **39**, **53** and **54**. In this case, incomplete *ika* cluster parts were integrated into the expression vector pSET152_ermE using DiPaC. Subsequently, the expression constructs were heterologously expressed in *S. albus* ΔPoTeM, a successfully modified strain lacking the internal *S. albus* PoTeM cluster. It was possible to produce both intermediates in sufficient amounts (Figure 87). Furthermore, **54** was isolated and could then be used for downstream applications. The production of the two intermediates builds a novel basis to enable a more precise knowledge of the biosynthetic steps towards **39**.

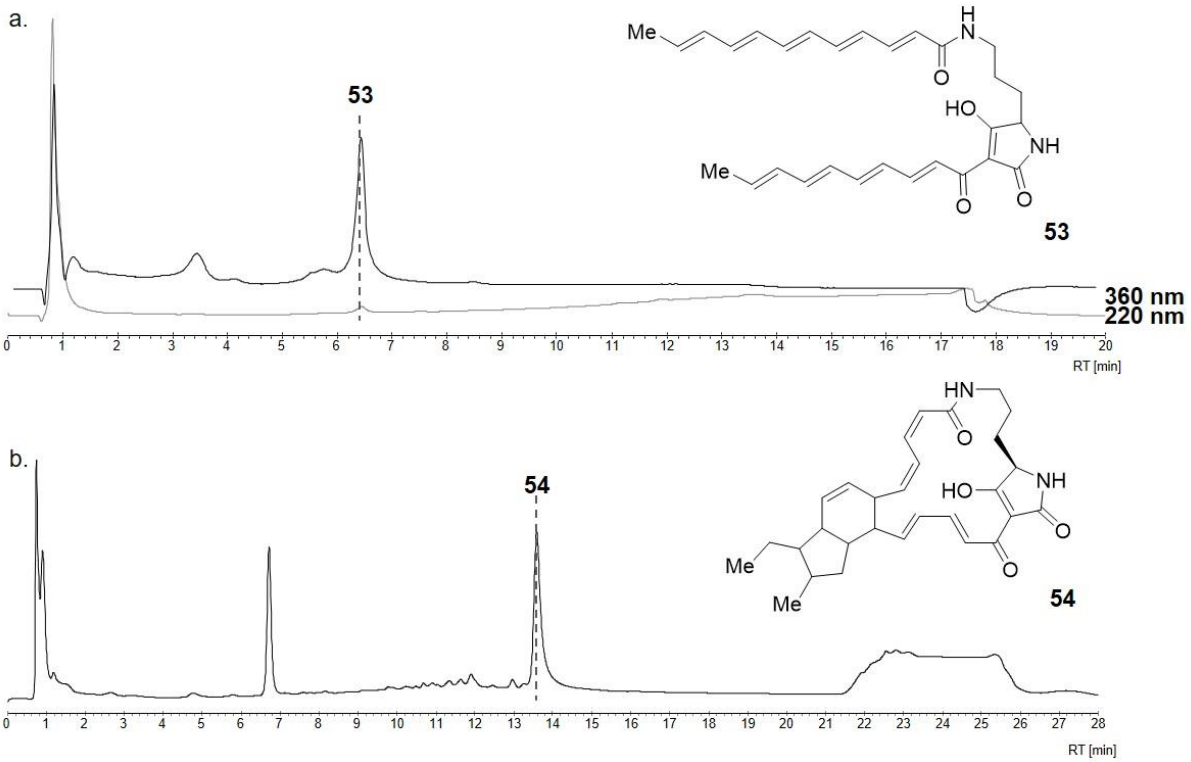


Figure 87. Heterologous expression of intermediates in the biosynthesis of **39**. a. HPLC analysis of the heterologous expression of **53**, chromatograms at 360 nm and 220 nm are depicted. Analytical method: Eurosphere_IkaFAST_CG.Meth. b. HPLC analysis of the heterologous expression of **54**. Analytical method: Eurosphere_Ika_lange Methode_AG.Meth.

It was not possible to heterologously express novel PoTeMs using the respective native BGCs captured by DiPaC. To solve this problem, a novel plug-and-play strategy was designed and developed. The iPKS/NRPS enzyme present in all known PoTeM BGCs was utilized to construct a basic expression vector. This vector contained *ikaA*, the iPKS/NRPS found in the *ika* cluster, as this already proved to produce high amounts of precursor material (**53**) for heterologous NP production. Furthermore, a set of nine promoters was integrated downstream of *ikaA* with additional restriction sites up- and downstream to facilitate an efficient cloning and refactoring of novel PoTeM clusters (Figure 88).

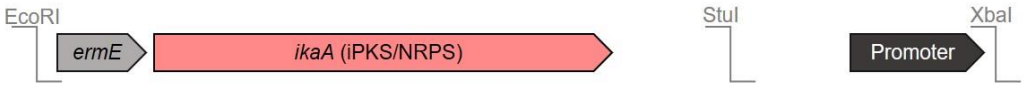


Figure 88. Basic plug-and-play expression vector.

In a first step, this plug-and-play PoTeM expression system was used to modify **39** using the three modifying enzymes DegA, PtmD, and IkaD. In previous biosynthetic studies, it has already been proven that these enzymes modify **39**[118, 136-137] and they were thus used to optimize the expression conditions and analyze the suitability of the different promoters for

PoTeM expression. Using the optimal conditions, a novel modifying enzyme, CftA, was chosen to be integrated into the plug-and-play system to modify **39** (Figure 89).

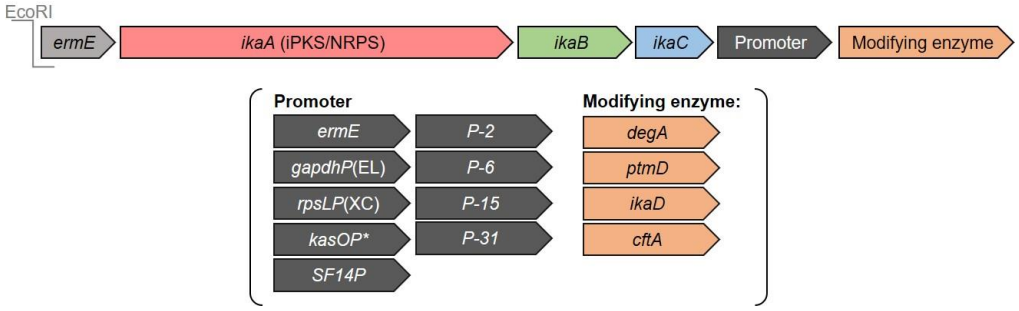


Figure 89. PoTeM plug-and-play expression vector for **39** modification. Depicted is the final expression vector for the modification 39 with the four modifying enzymes DegA, PtmD, IkaD, and CftA and the subset of promoters.

This study proved the functionality of DegA as a PoTeM-hydroxylase to hydroxylate **39** at position C-25, resulting in the PoTeM butremycin (**60**). Furthermore, it was possible to prove that PtmD is another PoTeM-hydroxylase, in contrast to the postulated function to hydroxylate at position C-29. The monooxygenase IkaD was also verified to integrate an epoxide at C7-8 (**59**) as well as hydroxy functionality at C-30 (**60**). Comparison of the expression results considering promoter effectivity already gave a first hint that stable compounds such as **60** are best expressed utilizing strong constitutive promoters like *gapdhP*(EL) and *rpsLP*(XC), whereas highly reactive compounds such as the ikarugamycin epoxide (**59**) show more stable expression titers using promoters such as *kasOP** or *SF14P*. Even a transition from the fast-growing *S. albus* DSM40313 to *S. coelicolor* M1154 was proven to yield a cleaner heterologous expression of the unstable **59**. Utilizing the gained knowledge, pSET152_ermE*::ikaABC*_gapdhP(EL)*_cftA* was expressed in *S. albus* DSM40313, which resulted in the formation of two new compounds. According to mass analysis, the minor product corresponds to clifednamide A (**55**) and the major product is the double-modified clifednamide B (**56**) (Figure 90). Further NMR analysis is needed to firmly validate these results.

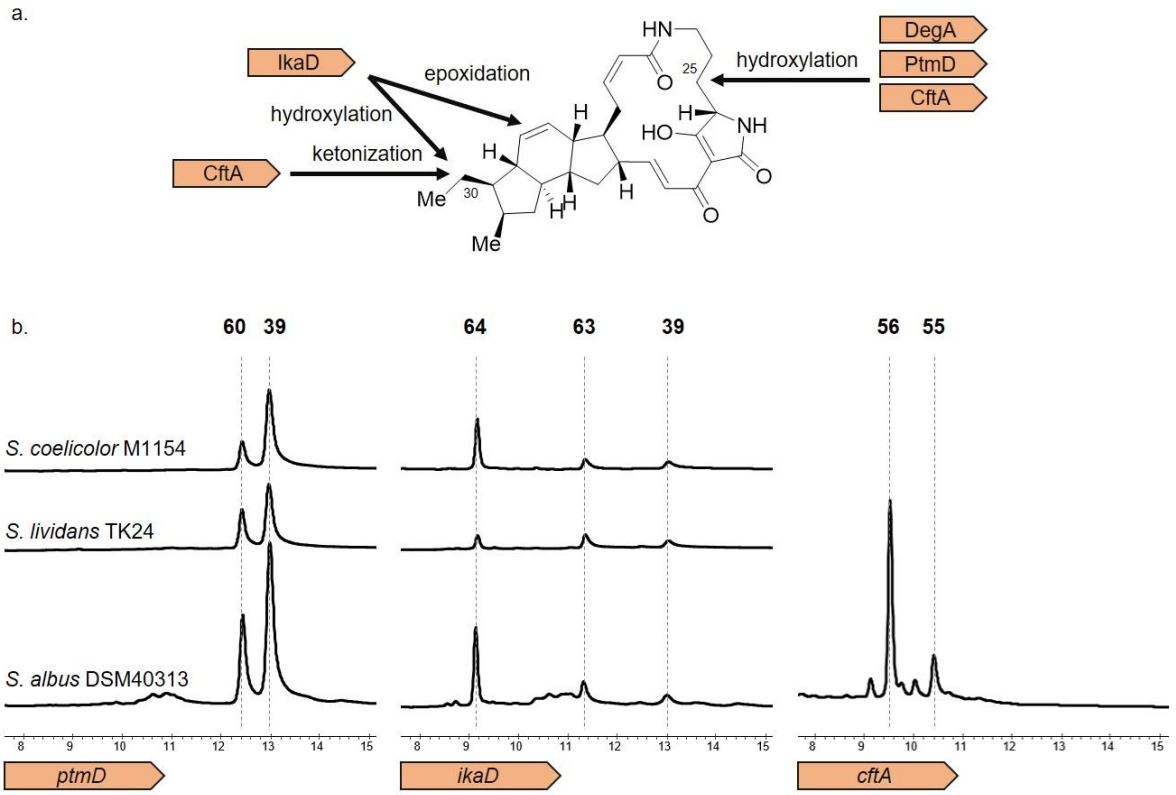


Figure 90. Results of the modification of **39** using the developed plug-and-play system. a. Schematic presentation of modifications conducted by the modifying enzymes DegA, PtmD, IkaD, and CftA. b. Depicted are the results of the modification of **39** with the three modifying enzymes PtmD, IkaD, and CftA in the three heterologous hosts *S. albus* DSM40313, *S. lividans* TK24, and *S. coelicolor* M1154. The results with DegA are not shown, as PtmD catalyzes the same modification with better conversion. Analytical method: Eurosphere_Ika_lange Methode_AG.Meth.

In addition to the modification of **39**, the plug-and-play system was also utilized to access novel PoTeMs. The biosynthetic genes of the three clusters *spi*, *deg*, and *fer* were integrated up- and downstream of *ikaA* into the plug-and-play expression vector and heterologously expressed in *S. albus* to identify the compounds encoded by these clusters (Figure 91). The clusters were successfully cloned and refactored using the DiPaC method.

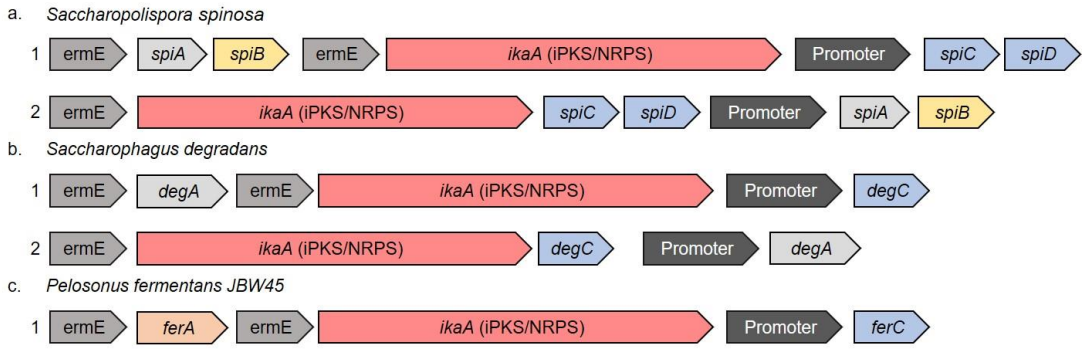


Figure 91. Plug-and-play expression vectors for novel PoTeM BGCs. a. Expression vector containing the wildtype (1) and rearranged (2) *spi* cluster. b. Expression vector containing the wildtype (1) and rearranged (2) *deg* cluster. c. expression vector containing the wildtype (1) *fer* cluster.

The expression of the *spi* cluster resulted in the production of two new compounds, **spi-1** and **spi-2** (Figure 92). MS-analysis revealed that the two compounds differed in the presence of one hydroxy group. In the scope of this study, it was not possible to purify the compounds for NMR analysis. However, the purification trials identified these compounds to be unstable. In accordance to the previous results considering promoter usage, the best expression was detected using *SF14P*, thereby underlining the suitability of *SF14P* to be appropriate for unstable compounds. Preliminary expression results for the two clusters *deg* and *fer* also showed promising production of a novel compound, with a structure to be determined in further studies.

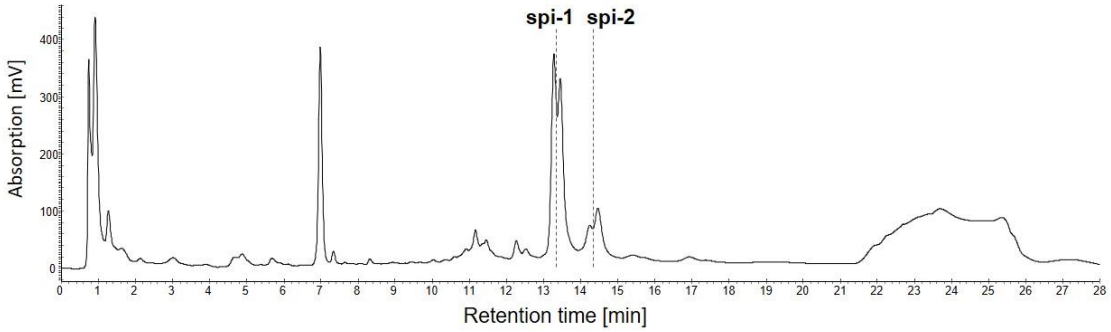


Figure 92. Heterologous expression of the *spi* cluster using the plug-and-play system. Depicted is the expression of pSET152_ermE*::sb_ikaA*_SF-14_*sb* revealing the two new compounds **spi-1** and **spi-2**. Analytical method: Eurosphere_Ika_lange Methode_AG.Meth.

Taken together, this work has generated ample knowledge for the heterologous production of PoTeMs by developing and applying an innovative synthetic biology platform.

4 Material and Methods

4.1 Materials

4.1.1 Primer

Primers used in the study are depicted in Table 19.

Table 19. List of utilized primers. Primer names, sequences and utilized functions are depicted.

Name	Sequence (5′ → 3′)	Function
ermE*_RBS_OE_for	ATCTAGGAATTCGCGGTCGATCTTGACGGCTGG CGAGAGGTGCGGGGAGGATCTGACCGACGCGGT CCACACGTGGCACCGCGATGCTGT	Overlap extension of *ermE* promoter
ermE*_RBS_OE_rev	AGCTTAGAATTCGTCCGTACCTCCGTTGCTCCG CTGGATCCTACCAACCGGCACGATTGTGCCCAC AACAGCATCGCGGTGCCACGTG	Overlap extension of *ermE* promoter
For_pSET152-ermE-rev	GCTGCAGGTCGACTCTAGAGAGGCCTTCCGTAC CTCCGTTGCT	Amplification of *ermE*
Rev_pSET152-ermE-rev	ACAGCTATGACATGATTACGAATTCGCGGTCGA TCTTGACGGC	Amplification of *ermE*
For_pSET152_ermE_rev-IKA	AGGTCGACTCTAGAGAGGCCTCTACAGGGCGAC CAGGACCTTG	*ika* amplification pSET152_ermE
Rev_pSET152_ermE_rev-IKA	CAACGGAGGTACGGAAGGATGTATTCATGGATT CCATGCACCACCCTGC	*ika* amplification pSET152_ermE
For_pULW201PW-Ika	GACGGTATCGATAAGCTTGATATCGAATTCATG GATTCCATGCACCACCCTGC	*ika* amplification pUWL201PW
Rev_pUWL201PW-Ika	TGCAGAGCTTCTAGAACTAGTGGATCCTACAGG GCGACCAGGACCTTG	*ika* amplification pUWL201PW
For_pWHM4*_IKA	GATCCCCGGGTACCGAGCTCGAATTCATGGATT CCATGCACCACC	*ika* amplification pWHM4*
Rev_pWHM4*_IKA_neu	GTAAAACGACGGCCAGTGAATTCTACAGGGCGA CCAGGACCTTG	*ika* amplification pWHM4*
For_pWHM1120-IKA	TTGCATGCCTGCAGGTCGACTCTAGAATGGATT CCATGCACCACCCTGC	*ika* amplification pWHM41120
Rev_pWHM1120-IKA	AGCTCGGTACCCGGGGATCCTCTACAGGGCGAC CAGGACCTTG	*ika* amplification pWHM41120
For_pSET152_ermE_rev new MCS	ATTGGTCATATGTGGCGGCCGTTACTAGTGATT CTAGAGAGGCCTTCCGTACCTCCG	New MCS for pSET152_ermE_rev
Rev_pSET152_ermE_rev new MCS	CCGCCACATATGACCAATCTGTTCTCTGTGAGG TCGACCTGCAGCCCAAGCTTGGCA	New MCS for pSET152_ermE_rev
For_pUWL201PW new MCS	ATTGGTCATATGTGGCGGCCGTTACTAGTGATT ATCGATAAGCTTGATATCGAATTC	New MCS for pUWL201PW
Rev_pUWL201PW new MCS	CCGCCACATATGACCAATCTGTTCTCTGTGAGT ATTTCTAGAACTAGTGGATCGATC	New MCS for pUWL201PW
For_pWHM1120 new MCS	ATTGGTCATATGTGGCGGCCGTTACTAGTGATT GAAGCTTGCATGCCTGCAG	New MCS for pWHM1120
Rev_pWHM1120 new MCS	CCGCCACATATGACCAATCTGTTCTCTGTGAGA TGTCCGCTCCCTTCTCTGA	New MCS for pWHM1120
pSET152_ermE_rev-ikaA_for	CTCACAGAGAACAGATTGGTCATATGTCATACC TCGCCATCACC	*ikaA* amplification
pSET152_ermE_rev-ikaA_rev	ATCACTAGTAACGGCCGCCACAATGGATTCCAT GCACCAC	*ikaA* and *ikaAB amplification*
For_pSET152_ermE_ikaAB	GCAGGTCGACTCTAGAGAGGCCTCTACACGCTC CTCGACGATG	*ikaAB amplification*
pSET152_ermE_ikaA_for	CTCACAGAGAACAGATTGGTCATATGATGGATT CCATGCACCAC	*ikaA* amplification for pUWL201PW
pSET152_ermE_ikaA_rev	ATCACTAGTAACGGCCGCCACATCATACCTCGC CATCACC	*ikaA* amplification for pUWL201PW

pUWL201PW-ikaA/B_rev	ATCACTAGTAACGGCCGCCACACTACACGCTCCTCGACGA	*ikaAB* amplification for pUWL201PW
For_pSET152_ermE_ikaA	GCAGGTCGACTCTAGAGAGGCCTTCATACCTCGCCATCACC	*ikaA* amplification for pSET152_ermE
For_pSET152_ermE_ikaAB	GCAGGTCGACTCTAGAGAGGCCTCTACACGCTCCTCGACGATG	*ikaAB* amplification for pSET152_ermE
OriT_PstI For	ACGAATACATGTCGCTAGAGCTTGCATGC	*oriT* and *traJ* amplification
OriT_PstI Rev	GACTTACATGTCAGCTAGCTAGAGTCGAC	*oriT* and *traJ* amplification
pSET152_nitA_for	TTCTAGCGGCCGCAGGCCTGCGAACTCCCTTATGCG	nitA amplification
pSET152_nitA_rev	TTACTGCGGCCGCCTACGAAACCTCCGTCGGTG	nitA amplification
pSET_albus_back_KO_fwd	GCTGCAGGTCGACTCTAGAGTCACCAGGTGACCGGGAGCC	Knockout plasmid
pSET_albus_back_KO_rev	CTCGACCTGAATCTCGGCCCCTCCCGCTCCCA	Knockout plasmid
pSET_tetR_KO_fwd	GGAGGGGCCGAGATTCAGGTCGAGGTGGCCCG	Knockout plasmid
pSET_tetR_KO_rev	GGCCGGCTCAACTAATTGACAGCTTATCATCGATAAGCTTTAATGCGG	Knockout plasmid
pSET_albus_front_KO_fwd	AAGCTGTCAATTAGTTGAGCCGGCCGATGCGGA	Knockout plasmid
pSET_albus_front_KO_rev	CGATATCGCGCGCGGCCGCGATGTCCCATAGCGATCCGGATTCCG	Knockout plasmid
For_pCC1FOS_oriT_albus_KO	GCGACACACTTGCATCGG	Knockout plasmid
Rev_pCC1FOS_oriT_albus_KO	CAGGCGTAGCAACCAGGC	Knockout plasmid
For_pCC1FOS_albus_KO_back	ATCCGATGCAAGTGTGTCGCGCCAGGGCGAGCTGGTTGG	Knockout plasmid
Rev_pCC1FOS_albus_KO_back	CAACCGATAAGCCCGGCCGGCACCTCAC	Knockout plasmid
For_pCC1FOS_albus_KO_ThioR	CCGGCCGGGCTTATCGGTTGGCCGCGAGATTCCTG	Knockout plasmid
Rev_pCC1FOS_albus_KO_ThioR	ACGGATTCCTATGGCGCGGTCGCGGTCG	Knockout plasmid
For_pCC1FOS_albus_KO_front	ACCGCGCCATAGGAATCCGTCCGGGACAATACC	Knockout plasmid
Back_pCC1FOS_albus_KO_front	ACGCCTGGTTGCTACGCCTGGGCAGCCTGATCCCGCCG	Knockout plasmid
pSET152_ermE_rev-spinosa_for	CTCACAGAGAACAGATTGGTCATATGCTGCCGGTCATGTGGTGA	*spi* amplification
pSET152_ermE_rev-spinosa_rev	ATCACTAGTAACGGCCGCCACAATGCGCATTTGGCTGGTAAC	*spi* amplification
pSET152-spinosa_for	CTCACAGAGAACAGATTGGTCAATGCGCATTTGGCTGGTAAC	*spi* amplification
pSET152-spinoas_rev	ATCACTAGTAACGGCCGCCACACTGCCGGTCATGTGGTGA	*spi* amplification
pSET152_ermE_rev-degradans_for	CTCACAGAGAACAGATTGGTCATATGTTAAGCCGGTTCGGTTTC	*deg* amplification
pSET152_ermE_rev-degradans_rev	ATCACTAGTAACGGCCGCCACAGTGGATACCTTAACTCCCTCAG	*deg* amplification
pSET152_degradans_for	CTCACAGAGAACAGATTGGTCAGTGGATACCTTAACTCCCTCAG	*deg* amplification
pSET152_degradans_rev	ATCACTAGTAACGGCCGCCACATTAAGCCGGTTCGGTTTCGTG	*deg* amplification
pSET152_ermE_rev-Tü6314_for	CTCACAGAGAACAGATTGGTCATATGCTACCGGCTGACGGAGAG	*Tü6314* amplification
pSET152_ermE_rev-Tü6314_rev	ATCACTAGTAACGGCCGCCACAATGGACGAGTGCGAACAC	*Tü6314* amplification
pSET152-Tü6314_for	CTCACAGAGAACAGATTGGTCAATGGACGAGTGCGAACACG	*Tü6314* amplification
pSET152-Tü6314_rev	ATCACTAGTAACGGCCGCCACATATGCTACCGGCTGACGGAGAGC	*Tü6314* amplification
pSET152_ermE_rev-roseosporus_for	CTCACAGAGAACAGATTGGTCATCACCAGCTCACCCGGAG	*ros* amplification
pSET152_ermE_rev-reseosporus_rev	ATCACTAGTAACGGCCGCCACATATGATGGATGAGCGCGAGCTC	*ros* amplification
pUWL201PW-roseosporus_for	CTCACAGAGAACAGATTGGTCAATGGATGAGCGCGAGCTC	*ros* amplification
pUWL201PW-roseosporus_rev	ATCACTAGTAACGGCCGCCACATATGGTCACCAGCTCACCCGGA	*ros* amplification
pSET-ikaA-acJBW45_for	TCCATTGTGGCGGCCGTTACTAATTAAAAATTTCTTCAACTTCCTC	*ferA* amplification
pSET-ikaA-acJBW45_rev	GGCCTCTCTAGAATCACTAGTATGTCAAAGTTAAACGACATTATATATC	*ferA* amplification

pSET-ikaA-fadJBW45_for	AGAGAACAGATTGGTCATATGTTAGTCGGCAAGTTGTGC	*ferC* amplification
pSET-ikaA-fadJBW45_rev	ATGGCGAGGTATGACATTGGAGAAAAATGAAAAATCAATAATTATTATC	*ferC* amplification
pSET152_spinosa-front_for	AGCCGTCAAGATCGACCGCGTCACGTCCAGTCGACCGG	*spi*_front (*spiAB*) amplification
pSET152_spinosa-front_rev	ACAGCTATGACATGATTACGGCGGTCGATCTTGACGGC	*spi*_front (*spiAB*) amplification
pSET152_spinosa-back_for	CACAGAGAACAGATTGGTCATATGCTGCCGGTCATGTGGT	*spi*_back (*spiDE*) amplification
pSET152_spinosa-back_rev	CACTAGTAACGGCCGCCACAATGGTGAACTCGCAGAATCC	*spi*_back (*spiDE*) amplification
pSET152_degradans-front_for	AGCCGTCAAGATCGACCGCGTTACTGGCGGTAGGTAATGTTG	*deg*_front (*degA*) amplification
pSET152_degradans-front_rev	ACAGCTATGACATGATTACGGCGGTCGATCTTGACGGCTG	*deg*_front (*degA*) amplification
pSET152_degradans-back_for	CACAGAGAACAGATTGGTCATTAAGCCGGTTCGGTTTC	*deg*_back (*degC)* amplification
pSET152_degradans-back_rev	CACTAGTAACGGCCGCCACAATGAAAGATTCAAACAAGAGC	*deg*_back (*degC*) amplification
pSET152_ermE_spinosa_for_new	GCAGGTCGACTCTAGAGAGGCTGCCGGTCATGTGGTGA	*spi*_new amplification
pSET152_ermE_spinosa_rev_new	AGCAACGGAGGTACGGAAGGATGTATTCATGCGCATTTGGCTGGTAAC	*spi*_new amplification
For_pSET152_spinosa_front_new	AGCCGTCAAGATCGACCGCGTCACGTCCAGTCGACCGG	*spi*_front_new (*spiAB*) amplification
pSET152_ermE_spinosa_rev_new	AGCAACGGAGGTACGGAAGGATGTATTCATGCGCATTTGGCTGGTAAC	*spi*_front_new (*spiAB*) amplification
pSET152_ermE_spinosa_for_new_I	GCAGGTCGACTCTAGAGAGGCTGCCGGTCATGTGGTGA	*spi*_back_new (*spiDE*) amplification
Rev_pSET152_spinosa_back_new	AGCAACGGAGGTACGGAAGGAGGATGTAATGGTGAACTCGCAGAATCCTG	*spi*_back_new (*spiDE*) amplification
pSET152_ermE_degradans_rev_new	AGCAACGGAGGTACGGAAGGATGTATTCGTGGATACCTTAACTCCC	*deg*_new amplification
pSET152_ermE_degradans_for_new	GCAGGTCGACTCTAGAGAGGTTAAGCCGGTTCGGTTTC	*deg*_new amplification
For_pSET_deg_full_front	CAGCCGTCAAGATCGACCGCGTCAATATTCCTTAACTAAATAATTAATAGGCGTAACG	*deg*_front_new (*degA*) amplification
Rev_pSET_deg_full_front	ACAGCTATGACATGATTACGGCGGTCGATCTTGACGGC	*deg*_front_new (*degA*) amplification
For_pSET_deg_full_back	CGCGGCCGCGCGCGATATCGTTAAGCCGGTTCGGTTTCGTG	*deg*_back_new (*degC*) amplification
Rev_pSET_deg_full_back	TTTAGTTAAGGAATATTGACGCGGTCGATCTTGACGGCTG	*deg*_back_new (*degC*) amplification
For_ermE_Promotoren	TGACTCTAGATCCGTACCTCCGTTGCTC	ermE amplification for plug-and-play
Rev_ermE_Promotoren	AGTCAGGCCTGCGGTCGATCTTGACGGC	ermE amplification for plug-and-play
Rev_ikaA_P_ikaBC	GTGATGGCGAGGTATGAAGGATGACGCCTTTCGTTCAGC	*ikaBC* amplification for plug-and-play
For_ikaA_gapdhP_ikaBC	CCGACCGAAGGAGCAGCAGGCTACAGGGCGACCAGGAC	*ikaBC* amplification for plug-and-play
For_ikaA_rpsLP (XC)_ikaBC	ACTTCCGCCTGCAGGGCAGGCTACAGGGCGACCAGGAC	*ikaBC* amplification for plug-and-play
For_ikaA_kasOP(Sternchen)_ikaBC	CCGTTCGAATGTGAACAAGGCTACAGGGCGACCAGGAC	*ikaBC* amplification for plug-and-play
For_ikaA_SF14_ikaBC	AATATCTCCTGGATAGGAGGCTACAGGGCGACCAGGAC	*ikaBC* amplification for plug-and-play
For_ikaA_P2_ikaBC	GCCGGATATGGCCGGGCAGGCTACAGGGCGACCAGGACCTT	*ikaBC* amplification for plug-and-play
For_ikaA_P6_ikaBC	AGTGTGGTGCGGTCGGCGCCAGGCTACAGGGCGACCAGGACCTT	*ikaBC* amplification for plug-and-play
For_ikaA_P-15_ikaBC	CGGGCCGGCGGCGCGGAAGGCTACAGGGCGACCAGGAC	*ikaBC* amplification for plug-and-play
For_ikaA_P-31_ikaBC	CGTGAGGAGAGGTCCGGAGGCTACAGGGCGACCAGGAC	*ikaBC* amplification for plug-and-play

For_ikaABC_P_PH	GTACTCTAGATTACTGGCGGTAGGTAATGT	*degA* amplification
Rev_ikaABC_P_PH	GCTATCTAGAGTGGATACCTTAACTCCCTC	*degA* amplification
For_ikaABC_P_ikaD	GTACTCTAGACTACCAGGCGACCGGCAGT	*ikaD* amplification
Rev_ikaABC_P_ikaD	GCTATCTAGAATGCCCGGACAGCAGGAACA	*ikaD* amplification
For_ikaABC_P_ptmD	GTACTCTAGACTACGCGGTGTGGGTCGG	*ptmD* amplification
Rev_ikaABC_P_ptmD	GCTATCTAGAATGGAGATTCTCCGCATGGAAG	*ptmD* amplification
For_spi_front_EcoRI (GA)	AGCCGTCAAGATCGACCGCGTCACGTCCAGTCGACCGG	*spi*_front (*spiAB*) amplification (WT)
Rev_spi_front_EcoRI (GA)	ACAGCTATGACATGATTACGGCGGTCGATCTTGACGGCTGG	*spi*_front (*spiAB*) amplification (WT)
Fwd_spi_back_XbaI	AGTCTCTAGATCAAGTCATGGCCCCGG	*spi*_back (*spiDE*) amplification (WT)
Rev_spi_back_XbaI	AGCTTCTAGAATGACGCGTGACAACAGT	*spi*_back (*spiDE*) amplification (WT)
Fwd_spi_front_XbaI	GCTATCTAGATCACGTCCAGTCGACCGG	*spi*_front (*spiAB*) amplification (MOD)
Rev_spi_front_XbaI	AGTCTCTAGAATGCGCATTTGGCTGGT	*spi*_front (*spiAB*) amplification (MOD)
Fwd_spi_back_StuI_gapdhP(EL)GA	CCGACCGAAGGAGCAGCAGGTCAAGTCATGGCCCCGGTGC	*spi*_back (*spiDE*) amplification (MOD)
Fwd_spi_back_StuI_rpsLP(XC)GA	ACTTCCGCCTGCAGGGCAGGTCAAGTCATGGCCCCGGTGCGCT	*spi*_back (*spiDE*) amplification (MOD)
For_spi_back_StuI_kasOP (GA)	CCGTTCGAATGTGAACAAGGTCAAGTCATGGCCCCGGTGC	*spi*_back (*spiDE*) amplification (MOD)
Fwd_spi_back_StuI_SF-14 (GA)	AATATCTCCTGGATAGGAGGTCAAGTCATGGCCCCGGTGC	*spi*_back (*spiDE*) amplification (MOD)
For_spi_back_StuI_P-2 (GA)	GCCGGATATGGCCGGGCAGGTCAAGTCATGGCCCCGGTGC	*spi*_back (*spiDE*) amplification (MOD)
Fwd_spi_back_StuI_P-15 (GA)	CGGGCCGGCGGCGCGGAAGGTCAAGTCATGGCCCCGGTGC	*spi*_back (*spiDE*) amplification (MOD)
For_spi_back_StuI_P-31 (GA)	CGTGAGGAGAGGTCCGGAGGTCAAGTCATGGCCCCGGTGC	*spi*_back (*spiDE*) amplification (MOD)
Rev_spi_back_StuI_P (GA)	GTGATGGCGAGGTATGAAGGATGACGCGTGACAACAGTGC	*spi*_back (*spiDE*) amplification (MOD)
For_ikaA_P_deg_front_EcoRI	AGCCGTCAAGATCGACCGCGTAATTAATAGGCGTAACGATGCCG	*Deg_front* (*degA*) amplification (WT)
Rev_ikaA_P_deg_front_EcoRI	ACAGCTATGACATGATTACGGCGGTCGATCTTGACGGC	*Deg_front* (*degA*) amplification (WT)
For_deg_back_XbaI	CGTATCTAGAATGCAGCCAGCACACAAAC	*Deg_back* (*degC*) amplification (WT)
Rev_deg_back_XbaI	GACTTCTAGATTAAGCCGGTTCGGTTTCG	*Deg_back* (*degC*) amplification (WT)
For_deg_back_StuI	GCATAGGCCTATGCAGCCAGCACACAAAC	*Deg_back* (*degC*) amplification (MOD)
Rev_deg_back_StuI	GTACAGGCCTTTAAGCCGGTTCGGTTTCG	*Deg_back* (*degC*) amplification (MOD)
For_deg_front_XbaI	AGCTTCTAGAGTGGATACCTTAACTCCC	*Deg_front* (*degA*) amplification (MOD)
Rev_deg_front_XbaI	ATCGTCTAGATAATTAATAGGCGTAACG	*Deg_front* (*degA*) amplification (MOD)
Rev_ikaA_P_FAD(JBW45)	GAGCAACGGAGGTACGGACTTTGGAGAAAAATGAAAAATCAATAATTATTATC	*Fer_back* (*ferC*) amplification
For_ikaA_gapdhP(EL)_FAD(JBW45)	GCTTGGGCTGCAGGTCGACTTTAGTCGGCAAGTTGTGC	*Fer_back* (*ferC*) amplification
For_ikaA_rpsLP(XC)_FAD(JBW45)	GCTTGGGCTGCAGGTCGACTTTAGTCGGCAAGTTGTGC	*Fer_back* (*ferC*) amplification
For_ikaA_SF-14_FAD(JBW45)	GCTTGGGCTGCAGGTCGACTTTAGTCGGCAAGTTGTGC	*Fer_back* (*ferC*) amplification
For_ikaA_P-2_FAD(JBW45)	GCTTGGGCTGCAGGTCGACTTTAGTCGGCAAGTTGTGC	*Fer_back* (*ferC*) amplification
For_ermE_AT(JBW45)	GCAGGTCGACTCTAGAGAGGCTAATTAAAAATTTCTTCAACTTCC	*Fer_front* (*ferA*) intermediate vector
Rev_ermE_AT(JBW45)	AGCAACGGAGGTACGGAAGGATGTATTCATGTCAAAGTTAAACGACATTATATATC	*Fer_front* (*ferA*) intermediate vector
For_ikaA_P_FAD_AT(JBW45)	AGCCGTCAAGATCGACCGCGCTAATTAAAAATTTCTTCAACTTCCTCAAGTTTTTTTATATCCTG	*Fer_front* (*ferA*) amplification

Rev_ikaA_P_FAD_AT(JBW45)	ACAGCTATGACATGATTACGGCGGTCGATCTTG ACGGC	*Fer_front* (*ferA*) amplification
260 Try-Ala_for	TGCCGTCTGCgcGCAGTGGGCGA	Mutagenesis of IkaC
260 Try-Ala_rev	ACCTGGCTGTGCACCGTG	Mutagenesis of IkaC
292 Tyr-Ala_for	GATCTTCGCCgcgGAGTGGTTCACCGAGGAGAA CCTCACG	Mutagenesis of IkaC
292 Tyr-Ala_rev	CCGCGGATCGTGGTGCGG	Mutagenesis of IkaC
292-Tyr-Arg_for	GATCTTCGCCcgcGAGTGGTTCACCGAGGAG	Mutagenesis of IkaC
292-Tyr-Arg_rev	CCGCGGATCGTGGTGCGG	Mutagenesis of IkaC
Sequencing Primer		
M13_for	GTAAAACGACGGCCAGT	Sequencing Primer
M13_rev	CAGGAAACAGCTATGAC	Sequencing Primer
Lac-Promoter pSET152	CTTCCGGCTCGTATGTTGT	Sequencing Primer
SEQ-Primer TüAlcD ForII	GGTGTTCACGATGCTCAA	Sequencing Primer
Screening_ikaA_rev	GCAGGATCTTGTAGTCGA	Sequencing Primer
pUWL201PWseq_rev_AL4	GATGTCGGACCGGAGTT	Sequencing Primer
pUWL201PWseq_AL6	CAATACGCAAACCGCCTCT	Sequencing Primer
pXY200_seq_rev_4a	AGGATCTTACCGCTGTTG	Sequencing Primer
TE_Spez for	TTCAACTACCTGATGGGCGA	Sequencing Primer
pCC1FOS_capcluster_seq-F	TCGGTCCTTTTGGTTCATTGCTGCTCGTTACGT CGACCAATTCTCAT	Sequencing Primer
pAEM35_seq_rev_1	TCCCGACGATCTTCACC	Sequencing Primer
SEQ_Primer_albus_KO_fwd	AAGGTCTCGGCCCATTA	Sequencing Primer
SEQ_Primer_albus_KO_rev	AAGAATCACTCGGGCTC	Sequencing Primer
Cyp_fwd	TCTCGTCAGCCTGGTCTGC	Sequencing Primer
SEQ-Primer_spinosa FAD2 ForI	GGTACTTCGCGCTTTTCCT	Sequencing Primer
SEQ Primer S. degradans FAD ForII	GAAGAACAGGTGAGTCGC	Sequencing Primer
SEQ-Primer_deg_SD_for	CGGCCATTTCAGCGTAG	Sequencing Primer
SEQ-Primer_deg_PKS_rev	CAAGCGAGTGCGGCAAT	Sequencing Primer
SEQ_Primer-Tü6314 CYP for	GGAGTTCAGAGCCTTCGG	Sequencing Primer
SEQ-Primer_roseo-SD_revI	GGTGAGCAGGAAGTAGAT	Sequencing Primer
SEQ-Primer JBW45-FAD	CCTGGGAATTAATCGAGATC	Sequencing Primer
SEQ-Primer_spinosa_cyp_for	TCAACCACTTCGGTGCG	Sequencing Primer
SEQ-Primer_spinosa-PKS_rev	GTCCATGTGCTCTGCCT	Sequencing Primer

4.1.2 Antibiotics

Utilized antibiotics and respective concentrations are depicted in Table 20.

Table 20. List of utilized antibiotics. Solvent and final concentration are depicted.

Name of antibiotic (abbreviation)	Working concentration	Solvent
Ampicillin (Amp)	100 µg/ml	50% EtOH
Kanamycin (Kan)	50 µg/ml	rH_2O
Chloramphenicol (Cam)	25 µg/ml	EtOH (abs.)
Apramycin (Apra)	30 µg/ml	rH_20
Nalidixic acid (NA)	25 µg/ml	0.3 M NaOH
Thiostreptone (Thio)	25 µg/ml	DMSO

4.1.3 Medium

Composition of the used media are shown in Table 21.

Table 21. List of utilized cultivation media.

Medium	Components for 1 l	
Luria Bertani (LB)	10 g	Tryptone
	4 g	Yeast extract
	10 g	NaCl
		For agar add:
	15 g	Agar
Terrific broth (TB)	20 g	Tryptone
	24 g	Yeast extract
	4 ml	Glycerol
		Autoclave in 900 ml water and add 100 ml of 10x phosphate buffer by sterile filtering
	0.17 M	KH_2PO_4
	0.72 M	K_2HPO_4
CASO / TSB	17 g	Peptone from Casein
	3 g	Peptone from Soy
	2.5 g	K_2HPO_4
	5 g	NaCl
	2.5 g	Glucose monohydrate
MS agar	10 g	Mannitol
	10 g	Soy flour
	10 g	Agar
GYM	6 g	Glucose
	7 g	Yeast extract
	10 g	Malt extract
		Adjust pH to 7.2 with KOH
		For agar add:
	12 g	Agar
	2 g	$CaCO_3$
ISP-4	10 g	Soluble starch
	1 g	K_2HPO_4
	1 g	$MgSO_4$
	1 g	NaCl
	2 g	$(NH_4)_2SO_4$
	2 g	$CaCO_3$
	1 mg	$FeSO_4 \times 7H_2O$
	1 mg	$MnCl_2 \times 4H_2O$
	1 mg	$ZnSO_4 \times 7H_2O$

2xYT	16 g	Tryptone
	10 g	Yeast extract
	8 g	NaCl
		Adjust pH to 7.0 with NaOH
R5	103 g	Sucrose
	10 g	Glucose
	10.12 g	$MgCl2 \times 6H_2O$
	0.25 g	K_2SO_4
	0.1 g	Difco Casamino acids
	5 g	Yeast extract
	2 ml	Trace element solution
	5.73 g	TES buffer
		Adjust pH to 7.3 with NaOH
		Autoclave and add the following sterile solutions:
	10 ml	0.5% KH_2PO_4
	10 ml	5 M $CaCl_2$
	1.5 ml	20% L-proline
		For agar add:
	22 g	Agar
	Trace element solution for R5 media (per l)	
	40 g	$ZnCl_2$
	200 g	$FeCl_2 \times 6H_2O$
	10 g	$CuCl_2 \times 2H_2O$
	10 g	$MnCl_2$
	10 g	$Na_2B_4O^7 \times 10\ H_2O$
	10 g	$(NH_4)_6Mo_7O_4 \times 4\ H_2O$
SOB	20 g	Tryptone
	5 g	Yeast extract
	0.58 g	NaCl
	0.19 g	KCl
		Autoclave and add the following sterile solutions:
	10 ml	1 M $MgCl_2$
	10 ml	1 M $MgSO_4$
SOC		For transforming SOB into SOC add sterile solution:
	9 ml	40% (v/w) D-glucose
Φ-medium	0.305 g	$MgSO_4$ x $7H_2O$
	0.925 g	$CaCl_2$ x $2H_2O$
	12.5 g	Glucose
	6.25 g	Tryptone
	6.25 g	Yeast extract
	6.25 g	Beef extract
		Set pH to 7.2 with NaOH

YEME	340 g	Sucrose
	10 g	Glucose
	9 g	Tryptone
	10 g	Yeast extract
	11 g	Malt extract
		Dissolve in 973 ml H_2O
		Set pH to 7.0
		Autoclave and add sterile filtered:
	25 ml	20% (v/v) glycerol
	2 ml	2.5 M $MgCl_2$

4.1.4 Buffer

Chemically competent cells

Buffer compositions required for preparation of chemically competent *E. coli* cells are depicted in Table 22.

Table 22. Composition of buffers used to prepare chemically competent cells.

Buffer	Concentration	Component
Transformation buffer	10 mM	HEPES
	15 mM	$CaCl_2$ x $2H_2O$
	250 mM	KCl
		Dissolve in 95 ml H_2O, adjust pH to 6.7
		Add 1.1 M $MnCl_2$ solution that is sterile filtered

Protoplast preparation

Composition of solutions required for the preparation of protoplasts are shown in Table 23, buffer compositions are depicted in Table 24.

Table 23. Solutions for protoplast isolation.

Solutions (prepared first)	
0.9% (w/v)	NaCl-solution
20% (v/v)	Glycine solution
0.5 % (w/v)	KH_2PO_4-solution
3.68% (w/v)	$CaCl_2$-solution
5.73% (w/v)	TES-solution
10.3% (w/v)	Sucrose-solution
2.5% (w/v)	K_2SO_4-solution
2.5 M	$MgCl_2$-solution
0.25 M	$CaCl_2$-solution
Trace element solution (components per l)	
40 mg	$ZnCl_2$
200 mg	$FeCl_3 \times 6H_2O$
10 mg	$CuCl_2 \times 2H_2O$
10 mg	$MnCl_2 \times 2H_2O$
10 mg	$Na_2B_4O_7 \times 10H_2O$
10 mg	$(NH_4)_6Mo_7O_{24} \times 4H_2O$

Table 24. Buffer compositions for protoplast isolation.

Buffer	Concentration	Component
Protoplast buffer	51.5 g	Sucrose
	0.125 g	K_2SO_4
	1.01 g	$MgCl_2 \times 6H_2O$
	1 ml	Trace element solution
	400 ml	H_2O
		Divide into 5x 80 ml aliquots and autoclave
		Add in the following order sterile filtered:
	1 ml	0.5% KH_2PO_4-solution
	10 ml	3.68% $CaCl_2$-solution
	10 ml	5.73% TES-solution
Lysis buffer	100 ml	10.3% Sucrose-solution
	10 ml	5.73% TES-solution
	1 ml	2.4% K_2SO_4-solution
	0.2 ml	Trace element solution
	1 ml	0.5% KH_2PO_4-solution
	4 ml	2.5 M $MgCl_2$-solution
	1 ml	0.25 M $CaCl_2$-solution
		Right before usage: Add 1 mg/ml sterile lysozyme

gDNA extraction

Solutions and buffers needed for gDNA extractions are shown in Table 25 and Table 26.

Table 25. Solutions for gDNA extraction.

Solutions	
10% (w/v)	Sodium dodecyl sulfate polyacrylamide (SDS)-solution
10% (w/v)	Cetyl trimethyl ammonium bromide (CTBA)-solution

Table 26. Buffer composition for gDNA extraction.

Buffer	Concentration	Component
Lysis buffer	25 mM	EDTA
	0.3 M	Sucrose
	25 mM	Tris-HCl
		Set pH to 7.5
TE-buffer	10 mM	Tris-HCl
	1 mM	EDTA

Protein purification

Buffer composition of buffers used for protein purification are shown in Table 27.

Table 27. List of buffers and buffer composition for protein purification.

Buffer	Concentration	Component
Lysis buffer	50 mM	NaH_2PO_4
	300 mM	NaCl
	10 mM	Imidazole
	10% (v/v)	Glycerol
Washing buffer 1	50 mM	NaH_2PO_4
	300 mM	NaCl
	20 mM	Imidazole
	10% (v/v)	Glycerol
Washing buffer 2	50 mM	NaH_2PO_4
	300 mM	NaCl
	40 mM	Imidazole
	10% (v/v)	Glycerol
Elution buffer	50 mM	NaH_2PO_4
	300 mM	NaCl
	250 mM	Imidazole
	10% (v/v)	Glycerol

DNA analysis

Buffers used for DNA analysis are depicted in Table 28.

Table 28. List of buffer composition used for DNA analysis.

Buffer	Concentration	Component
10x Loading Dye	20% (w/v)	Sucrose
	0.15% (v/v)	Orange G (Serva)
	0.05% (v/v)	Xylen cyanol
	0.05% (w/v)	Bromphenolblue

Protein analysis

Buffers used for protein analysis are depicted in Table 29.

Table 29. List of buffer composition used for protein analysis.

Buffer	Concentration	Component
4x separating gel buffer	1.5 M	TrisHCl
	0.8% (w/v)	SDS
		Set pH to 8.8
4x stacking gel buffer	0.5 M	TrisHCl
	0.8% (w/v)	SDS
		Set pH to 6.8
5x SDS loading dye	220 mM	TrisHCl
	5% (v/v)	β-mercaptoethanol
	10% (w/v)	SDS
	0.2% (w/v)	Bromophenolblue
	50% (v/v)	Glycerol
Coomassie staining solution	10% (v/v)	Acetic acid
	30% (v/v)	Methanol
	0.25% (w/v)	Coomassie brilliantblue G-250
Destaining solution	10% (v/v)	Acetic acid
	30% (v/v)	Methanol

Enzymatic Assays

Buffers used for enzymatic assays are depicted in Table 30.

Table 30. List of buffer composition used for enzymatic assays.

Buffer	Concentration	Component
NaH_2PO_4	400 mM	NaH_2PO_4
		Set pH to 8.0
HEPES	1 M	HEPES
	15%	Glycerol
		Set pH to 7.8

4.1.5 Equipment

The equipment utilized for this study is listed in Table 31, the utilized software is shown in Table 32.

Table 31. Utilized equipment for this study. Depicted are the name, manufacturer and function of the respective piece of equipment.

Equipment	Manufacturer	Function
Advion Expression single-quadrupole mass analyzer	Advion	LC-MS
Ultrasound Sonopuls HD2070	Bandelin	Protein purification
Life Eco PCR cycler	Bioer	PCR
Autoclave 3170 ELV	Biomedis	Autoclave
Mini Protean® Tetra System gel chamber	Bio-Rad	SDS-PAGE
MicroPulser	Bio-Rad	Electroporation Transformation
T100 ™ Thermal cycler	Bio-Rad	PCR
AVHD500	Bruker	NMR
AV500-cryo	Bruker	NMR
Alpha 2-4	Christ	Lyophilization
Electrophoresis Power Supply EV243	Consort	Gel electrophoresis
RV12 Vacuum pump	Edwards	LC-MS
ThermoMixer C	Eppendorf	Incubation
LCQ Fleet Ion Trap MS spectrometer	Eppendorf	LC-MS
Centrifuge 5418R	Eppendorf	Centrifuge
Water bath FBH 604	Fischerbrand	Incubation

ÄKTA purifier	GE Healthcare	Protein purification
Nanophotometer P330	Implen	DNA/ protein concentration
Incubation shaker Multitron Standard	Infors HAT	Incubation
Plus Intelligent Prep. pump	Jasco	HPLC/LC-MS
PU-1580 Intelligent HPLC pump	Jasco	HPLC/ LC-MS
MD-2010 Plus Multiwavelength detector	Jasco	HPLC/ LC-MS
DG-2080-53 3-Line Degaser	Jasco	HPLC/ LC-MS
PU-2086 Plus Intelligent Pumps	Jasco	HPLC/ LC-MS
AS-2055 Plus Intelligent Sampler	Jasco	HPLC/ LC-MS
UV-1575 Intelligent UV/VIS detector	Jasco	HPLC/ LC-MS
N118LA nitorgene generator	Peak Scientific	LC-MS
MIKA 1000 dynamic mixing chamber, 1000 µl	Portmann Instruments	HPLC/ LC-MS
1000 µl/ 5000 µl sample loop	Portmann Instruments	HPLC/ LC-MS
Bio-Imaging-System Gene Genius	SynGene	Agarose Gel Analyzation
TC-5000 PCR cycler	Techne	PCR
LCQ Fleet Ion Trap MS spectrometer	Thermo Scientific	LC-MS
UltiMate 3000 HPLC system	Thermo Scientific	LC-MS
Heraeus Multifuge X3R Centrifuge	Thermo Scientific	Centrifuge
Chemistry-Hybrid-Pump-RC6-Pumpe	Vacuumbrand	Lyophilisation
24/6 centrifuge	VWR	Centrifuge

Table 32. Utilized software for this study. Depicted are the name, manufacturer and function of the respective software.

Software	Manufacturer	Function
Data Express Software	Advion	LC-MS
Mass Express Software	Advion	LC-MS
Geneious	Biomatters	DNA analysis
SnapGene	GSL Biotech LLC	DNA analysis
ChromPass Chromatography Data System-Software	Jasco	Analytical HPLC
Galaxie Chromatography software	Jasco	Semi-preparative HPLC
MestReNova	Mestrelab Research S. L.	NMR
PyMOL	Schrödinger, Inc.	Crystal structures
GeneSnap	Syngene	Gel picture analysis
Xcalibur	Thermo Scientific	LC-MS
Xcalibur Qual Browser 2.2 SP1.48	Thermo Scientific	LC-MS

4.2 Methods

4.2.1 Bacterial strains and cultivation

The utilized bacterial strains are listed in Table 33.

Table 33. List of utilized bacterial strains with function and origin.

Strain	Function	Reference
E. coli DH5α	Host for cloning	NEB
E.coli EPI300™	Host for cloning complex DNA fragments	Epicenter
E. coli BL21 (DE3)	Host for heterologous protein expression	NEB
E. coli ET12567 (pUZ8002)	Donor strain for conjugation	[141]
E. coli BW25113 (pKD46)	Strain for homologous recombination	[142],[143]
S. albus DSM40313	Host strain for heterologous NP expression	DSMZ [120]
S. lividans TK24	Host strain for heterologous NP expression	[121]
S. coelicolor M1154	Host strain for heterologous NP expression	[122]

E. coli strains were cultivated in 5 ml LB medium or plated on LB agar pates containing the respective antibiotic. Cultures were incubated at 37 °C and 200 rpm. For long term storage, a fresh overnight culture was mixed with a final concentration of 15% (v/v) glycerol and stored at -80 °C.

Streptomyces were cultivated in CASO bouillon or plated on MS agar plates. Expression cultures of *Streptomyces* were inoculated with a 1:10 or 1:50 of cell in mainly ISP-4 medium but additionally GYM and a cultivation medium published by Zhang *et al*.[113], from now referred to as Zhang, were used. Cells were grown for 3-12 days at 28 °C and 200 rpm. Plates were incubated at 30 °C. For long time storage liquid cultures were 1:1 mixed with *Streptomyces* cryo solution, and spores were resuspended in 20% (v/v) glycerol both were stored at -80 °C.

4.2.2 Preparing competent cells

To enable *E. coli* cells to integrate foreign DNA they need to be made competent. There are two different broadly applied transformation techniques that need either electro competence or chemically competence of the cells.

Chemically competent cells

A 250 ml SOC culture was inoculated with 2.5 ml of a preculture containing the appropriate *E. coli* strain. Cultivation was conducted at 18 °C overnight until the cells reach an OD_{600} between 0.4-0.6. Subsequently cells are cooled on ice for 10 min, harvested at 4000 rpm, 4 °C for 10 min and resuspended in 80 ml transformation buffer (TB buffer). A second 10 min incubation on ice is performed, the cells are harvested at 4000 rpm, 4°C for 10 min and resuspended in 18.6 ml TB buffer and 1.4 ml DMSO. Cells are aliquoted in 100 µl aliquots, frozen in liquid nitrogen and stored at -80 °C.

Electrocompetent cells

A 250 ml LB culture was inoculated with 2.5 ml of the preculture containing the appropriate *E. coli* strain. Cells are cultivated at 37 °C, 200 rpm until an OD_{600} between 0.4-0.5 is reached. Cells will be washed with 10% (v/v) glycerol for several times, to remove all salts that would disturb the electroporation, all washing steps were performed on ice. Firstly, cells were harvested (4000 rpm, 5 min, 4 °C) and resuspended in 40 ml 10% (v/v) glycerol, and secondly cells were harvested and resuspended in 20 ml 10% glycerol. Next, cells were harvested and resuspended in 10 ml 10% (v/v) glycerol, and finally cells were harvested and resuspended in 500 µl 10% (v/v) glycerol and aliquoted in 50 µl aliquots. Aliquots were frozen in liquid nitrogen and stored at -80 °C.

4.2.3 Protoplast preparation

Gram positive cells cannot be transformed with DNA as easy as gram negative cells, one possibility to render gram positive cells susceptible to incorporating foreign DNA is to remove the proteoglycan cell wall, which would result in protoplasts that can be transformed using polyethylene glycol (PEG).

To isolate protoplast, a 5 ml φ-medium preculture was prepared, inoculated with 20 µl spore solution of the strain to be protoplasted, and grown at 28 °C, 200 rpm until dense (about 48 h). A 20 ml YEME-medium main culture was inoculated with 2 ml homogenized preculture and incubate at 28 °C, 200 rpm for 24 h. Cells were harvested at 5000 rpm, 4 °C, 10 min, supernatant was discarded, and cell pellet was resuspend in 15 ml cold 0.9% (w/v) NaCl-solution. Cells were again harvested at 5000 rpm, 4 °C, for 10 min, supernatant was discarded, and cell pellet was resuspended in 15 ml protoplast buffer. One last time, the cells were harvested at 5000 rpm, 4 °C for 10 min, supernatant was discarded, and cell pellet was

resuspended in 5 ml lysis buffer (freshly supplemented with 1 mg/ml lysozyme). Cells were transferred into a 15 ml centrifuge tube and incubate at 28 °C, 200 rpm for 75 min. In this step the proteoglycan cell wall will be digested by lysozyme. After the incubation 10 ml protoplast buffer were added and the mixture was centrifuged at 1000 rpm, 4 °C for 10 min to harvest the mycelia. The supernatant was transferred into a new 15 ml centrifuge tube and centrifuge at 5000 rpm, 4 °C, for 10 min to harvest the protoplasts. The supernatant was discarded, and the protoplasts were carefully resuspended in 15 ml protoplast buffer, the suspension was centrifuge again at 5000 rpm, 4 °C for 10 min and cell pellet was resuspend once more in 10 ml protoplast buffer. Protoplast were aliquoted in 100 µl in 1.5 ml tubes and slowly frozen, slow freezing is important for transformation efficiency. In detail, aliquots were placed in a disposable bag, the bag was placed in a bowl of ice and stored at -80 °C.

4.2.4 DNA purification

gDNA extraction

A phenol:chloroform:isoamylalcohol based purification method for the extraction of high quality gDNA from various microorganisms is applied. Microorganisms were grown until late exponential or early stationary phase, subsequently harvested by centrifugation (10,000x g, 10 min) and washed with 0.9% (w/v) NaCl solution to remove residual medium. In the next step, cells were resuspended in 5 ml lysis buffer, frozen and thawed in liquid nitrogen and 50 °C water bath three times to render cells susceptible to lysis. Then, 1 mg/ml lysozyme and 10 µg/ml RNase was added and incubated at 37 °C for 1 h, if cell lysis was not completed incubation was repeated for 30 min. Subsequently 0.5 mg/ml proteinase K and SDS to a final concentration of 1% (w/v) was added to the suspension, mixed well and incubated first at 37 °C for 30 min and afterwards at 55 °C for 30 min. NaCl (final concentration 1 M) and cetyltrimethylammoniumbromid (CATB) to a final concentration of 1% (w/v) was added, mixed well and incubated at 65 °C for 10 min. An equal amount of phenol:chloroform:isoamylalcohol (25:24:1) solution was added to isolate DNA from other cell compartments. The sample was mixed by inversion until homogenous. An additional incubation step for 30 min of ice followed, where it was important to ensure mixing of the sample. To separate the phases, the solution was centrifuged at 10,000x g for 10 min and the aqueous phase was transferred into a fresh 15 ml centrifuge tube. The aqueous phase was extracted with one volume of phenol:chloroform:isoamylalcohol again inverted until homogenous, centrifuged at 10,000x g for 10 min, and aqueous phase was transferred to a new 15 ml centrifuge tube. The phenol:chloroform:isoamylalcohol step was once more repeated and aqueous phase was divided in 1.2 ml fractions and transferred into 2 ml reaction tubes. A total of 0.6 vol (720 µl) of

Isopropanol were added to the fractions, inverted 15-30 times, and DNA was pelleted by centrifugation for 30 min at top speed. Supernatant was removed, DNA was washed with 1 ml 70% (v/v) EtOH and again pelleted by centrifugation for 15 min at top speed. Finally, EtOH needs to be removed completely and DNA needs to be dried, either air dried of at 60 °C for no longer than 5 min. Extracted gDNA was dissolved in 50-200 µl TE-buffer. Quality of the gDNA can be ensured using Nanodrop-, gel electrophoresis-, or 16S-analysis.

Plasmid extraction

For the purification of plasmid DNA the peqGOLD Plasmid Isolation Kit (Peqlab) was applied according to manufactures instruction. High copy number plasmids were isolated from 5 ml culture, low copy plasmids were isolated from 15 ml culture.

PCR purification

The purification of PCR products was conducted depending on the specificity of amplification during the PCR. If the PCR reaction showed a single specific product, a PCR purification kit (Jena Bioscience) was utilized. If the PCR showed additional unspecific amplification, the PCR mixture was separated by agarose gel electrophoresis and the desired DNA band was excided from the agarose gel. The DNA from the gel was purified by either using the Gel Extraction Kit (Peqlab) or the Monarch DNA Gel extraction kit (NEB) and eluted in the smallest possible volume of dH_2O to ensure maximal DNA concentration. Kits were applied according to manufactures instruction.

4.2.5 Restriction Digestion

Restriction digestion is a method to cut DNA at specific sites, to either linearize vectors for cloning or to analyze new assembled DNA according to their accuracy. Therefore, two different protocols were applied for a preparative restriction digest of an analytical restriction digest.

Preparative restriction digest

Large amount of vector or insert DNA need to be digested before cloning. 1 µg of the DNA, 5 µl of the appropriate buffer (NEB) and 1 µl of each restriction enzyme were combined in a

total volume of 50 µl and incubated at 37 °C for 4 h. Subsequently, restriction enzymes were heat inactivated according to manufactures specifications.

Analytical restriction digest

The correct assembly of newly cloned constructs was ensured by digesting the plasmid with specific restriction enzymes that will result in DNA fragments of predicted size. For this analysis 240 ng of the DNA, 1 µl of the appropriate buffer (NEB) and 0.125 µl of each restriction enzyme were combined in 10 µl total reaction volume. The mixture was at least incubated for 1 h at 37 °C and adjacently analyzed using agarose gel electrophoresis.

4.2.6 Dephosphorylation

Ligation needs a 5'-phosphate group to link the two DNA fragments, to exclude the religation of the vector to itself, the 5'-phosphate moiety is removed. To accomplish this, the heat inactivated preparative restriction digest was mixed with 6 µl Antarctic phosphatase buffer, 2 µl Antarctic phosphatase in a total volume of 60 µl. The reaction was incubated at 37 °C overnight and the Antarctic phosphatase was heat inactivated at 65 °C for 10 min.

4.2.7 Ligation

The assembly of two or more fragments can be accomplished using different methods. In this thesis conventional ligation, Gibson assembly[89] and Sequence and ligation independent cloning (SLIC)[144] were used.

Cloning by ligation

Conventional cloning by ligation was conducted with using different restriction sites. Utilizing only one restriction site rendered the possibility for the insert to integrate in both directions dropping cloning success to 50%. For ligation reaction 0.02 pmol of digested vector and 0.1 pmol insert were combined with 2 µl 10x T4 ligation buffer (Jena bioscience) and 1 µl T4 ligase (Jena bioscience) in a total volume of 20 µl. The mixture was incubated at 16 °C overnight and heat inactivated at 90 °C for 2 min.

Gibson assembly

Gibson assembly bases on an isothermal reaction of a DNA-exonuclease, a DNA polymerase and a DNA ligase using homologous overhangs of vector and insert to assembly the DNA fragments.[89]

The PCR amplification of either the insert or the vector needed to insert a 18-21 bp homologous region at the assembly ends. In the isothermal reaction, the overhangs are converted into single strands by the exonuclease, followed by a reintegration of the missing nucleotides by the polymerase and a final connection of the two fragments by the ligase. For a single reaction batch 0.02 pmol vector and 0.1 pmol insert in a total volume of 10 µl are combined with the 2x HIFI DNA assembly mix (NEB) and incubated at 50 °C for 1 h.

Sequence and ligation independent cloning (SLIC)

SLIC bases on the same principle as Gibson assembly, however replacing the exonuclease and the polymerase by the T4-polymerase that exhibits both functions and using the cellular ligation machinery of *E. coli*.[144]

DNA fragments need to fulfil the same requirements as for Gibson assembly. For a SLIC reaction, 0.02 pmol vector and 0.1 pmol insert were mixed with 1 µl 2.1 buffer (NEB) and 0.5 µl T4-polymerase in a total volume of 10 µl. The mixture was incubated 10 min at room temperate and additionally 10 min on ice. To use the cellular ligase, the whole reaction mixture is transformed into *E. coli* DH5α.

4.2.8 Homologous recombination

Homologous recombination is a method to exchange parts of DNA based on homology regions and recombinases that would rearrange the DNA at the homologous sequences. It can be used for various applications like capturing whole gene clusters from genomic DNA or cosmids/fosmids, or to exchange vector backbones for expression.

The expression plasmid was transformed into electrocompetent *E. coli* BW25113 pKD46 and selected on LB-agar containing Amp (selection for pKD46) and the appropriate antibiotic for the desired plasmid. As pKD46 has a heat sensitive OriT, the bacteria needed to be grown at 30 °C. Integration of the plasmid was PCR-verified and a 5 ml LB preculture with the appropriate antibiotics was grown overnight at 30 °C, 200 rpm. 120 ml SOC-medium (supplemented with antibiotics) were inoculated with 100 µl preculture and 100 µl

1 M Arabinose, to induce genes for recombination. Cells were grown at 30 °C, 200 rpm until an OD_{600} of 0.8 was reached, then harvested (4000 rpm, 4 °C, 5 min) and washed twice with 10 ml or 5 ml 10% (v/v) glycerol, respectively. After a last centrifugation at 4000 rpm, 4 °C, 5 min, cell pellet was resuspended in 100 µl 10% (v/v) glycerol. A 50 µl sample of this cell suspension was mixed with 100 ng of the linearized DNA for homologous recombination, transformation by electroporation was conducted and cells were recovered in 900 µl SOC at 30 °C for 1 h. In this step the homologous recombination was performed, subsequently cells were plated on LB containing only the antibiotic for the desired product and grown at 37 °C, to promote the loss of pKD46. Screening for successful recombination was conducted using colony PCR (see 4.2.13) and the plasmid of the respective clone was purified.

4.2.9 Site directed mutagenesis

The introduction of single nucleotide mutations was accomplished using the site directed mutagenesis kit by NEB. Using the QuikChange® primer design software provided by NEB, primers can be designed for amplifying the mutagenized vector by PCR (chapter 4.2.13). The PCR product is processed using the KLD enzyme mix which includes three enzymes, a kinase (K) for phosphate removal, a ligase (L) for ligation, and DpnI (D) for template removal. After gel purification of the PCR product, 0.5 µl of the elution fraction were combined with 2.5 µl of the provided reaction buffer and 0.5 µl enzyme mix in a total volume of 5 µl, the reaction mixture was incubated for 10 min at room temperature and immediately transformed into *E. coli* DH5α chemically competent cells (chapter 4.2.10).

4.2.10 Transformation

A transformation is used to integrate plasmid DNA into *E. coli* cells, either for cloning or expression.

Heat shock transformation

For cloning, the whole reaction mixture of the different ligation protocols was added, for expression, 5 ng DNA were added to chemically competent cells (see 4.2.2). The cells were incubated 30 min on ice, the heat shock was performed for 90 sec at 42 °C and after a further 2 min incubation on ice, 900 µl SOC was added to the cells. Recovery was conducted at 37 °C,

200 rpm for 45 min and subsequently the cells were plated on LB-agar containing the appropriate antibiotic and incubated at 37 °C overnight.

Electroporation transformation

Between 5 to 100 ng DNA were added to electrocompetent cells (see 4.2.2) and incubated for 15 min on ice. The cells were transferred to an electroporation cuvette and pulsed (program EC2) with 2.5 kV. 900 µl SOC were added to the cells which were transferred into a sterile 1.5 ml reaction tube and recovered at 37 °C, 200 rpm, for 45 min. Subsequently, the cells were plated on LB-agar containing the appropriate antibiotic and incubated at 37 °C overnight.

4.2.11 Protoplast transformation

Foreign DNA can be introduced into *Streptomyces* protoplast using a PEG based transformation approach. Before the transformation, R5-agar plates need to be dried at 37 °C. For transformation, protoplast (preparation see chapter 4.2.3) were thawed rapidly, 1 µg DNA was added and gently mixed, and additionally 100 µl transformation buffer (supplemented with PEG) were added and carefully resuspended. The cells were incubated at room temperature for 5 min, plated on the R5-agar, and incubated at 30 °C for 4-6 days.

To proceed with single clones, colonies were picked and resuspend in 50 µl dH_2O (cells needed to be actively squashed, to resuspend mycelia). The 50 µl cell suspension was split in 25 µl that were added to 5 ml CASO (supplemented with Apra and NA) and 25 µl that were plated on MS-agar (supplemented with Apra and NA) and incubate at 28 °C, 200 rpm or 30 °C respectively. Continue cultivation of S*treptomyces* as depicted above (chapter 4.2.1).

4.2.12 *Streptomyces* conjugation

Horizontal conjugation of DNA from *E. coli* to *Streptomyces* is a different approach to integrate foreign DNA to gram positive bacteria. Hereby the ability of *E. coli* ET12567 pUZ8002 to form reproductive pili is used to transfer the DNA into the heterologous *Streptomyces* host.

The expression construct needed to be first transformed into electrocompetent *E. coli* ET12567 pUZ8002 by electroporation (chapter 4.2.10), the cells were selected on LB-agar containing Kan, Cam, and the antibiotic of the respective expression construct. PCR-verified cells (chapter 4.2.13) were grown in 15 ml LB (supplemented with Kan, Cam, Apra) at 37 °C, 200 rpm until an OD_{600} of 0.4 to 0.6 was reached. The cells were washed three times with

15 ml LB without any antibiotic, by harvesting the cells at 4000 rpm, 5 min at 4 °C, discarding the supernatant, and resuspending the cells in LB. At the last washing step, the cells were resuspended in 500 µl LB. Parallel, spores needed to be heat-activated, by resuspending 20 µl spore samples in 500 ml 2xYT medium and heat activating them at 50 °C for 10 min. Spores were cooled on ice before proceeding. Spores were added to washed *E. coli* cells and the mixture was harvested at 4000 rpm, 4 °C for 2 min. The pellet was plated on MS-Agar (supplemented with 10 mM $MgCl_2$ and 60 mM $CaCl_2$) and incubated for 16-20 h at 30 °C. After incubation plates were overlaid with 1 mg NA and 1.25 mg Apra, the plates were sealed with parafilm and incubate for 4-7 d at 30 °C.

To proceed with exconjugants, single clones were picked, resuspend in 50 µl dH_2O (cells need to be actively squashed, to resuspend mycelia). The 50 µl cell suspension was split in 25 µl that were added to 5 ml CASO (supplemented with Apra and NA) and 25 µl that were plated on MS-agar (supplemented with Apra and NA) and incubate at 28 °C, 200 rpm or 30 °C respectively. Continue cultivation of *Streptomyces* as depicted above (chapter 4.2.1).

4.2.13 Polymerase chain reaction (PCR)

Colony PCR

To verify the successful integration of a plasmid into *E. coli* or *Streptomyces* a colony PCR was conducted using the Taq polymerase. Single colonies were resuspended in 50 µl dH_2O and 5 µl of this mixture were used as template for the PCR. Additionally a 25 µl reaction batch consisted of 2.5 µl 10 x reaction buffer (200 mM Tris, 100 mM $(NH_4)_2SO_4$, 100 mM KCl, 20 mM $MgSO_4$, 1% Triton-X, pH=8.8), 1 µl DMSO, 100 nM dNTPs, 200 nM of each primer and 0.125 µl self-made Taq polymerase. The PCR cycling was performed starting with an initial denaturation step at 95 °C for 5 min, followed by 33 cycles of denaturation at 95 °C for 45 sec, annealing at 47-72 °C for 30 sec (according to melting temperature of the primer), and elongation at 72 °C for 1 min per kb, terminal a final elongation is conducted at 72 °C for 5 min and the reaction is cooled at 16 °C. PCR reactions were analyzed using gel agarose electrophoresis (chapter 4.2.15).

Q5 (long amplicon) PCR

The amplification of inserts or whole gene clusters was performed using high fidelity Q5 DNA-polymerase, this enzyme contains a prove-reading function to ensure sequence accuracy for the amplified DNA. A standard PCR was conducted in a 25 µl batch containing 1x reaction

buffer, optional 1x High-GC enhancer, 200 nM dNTPs, 250 nM of each primer, 5 to 100 ng template DNA and 0.125 µl Q5-Polymerase (0.02 U/µl). The PCR cycling was performed with an initial denaturation at 98 °C for 45 sec, followed by 30 cycles of denaturation at 98 °C for 10 sec, annealing at 50-72 °C (according to melting temperature of the primer), and elongation at 72 °C for 20-30 sec per kb, terminal a final elongation is conducted at 72 °C for 5 min and the reaction is stored at 16 °C. For long amplicon PCRs, an elongation velocity of 45 sec per kb was used and the final elongation was adjusted to exceed the normal elongation time. PCR products were analyzed using agarose gel electrophoresis (chapter 4.2.15).

4.2.14 Sanger sequencing

Sanger sequencing was primary conducted using services from GATC Biotech and later switched to Eurofins. For either company, a DNA mixture of 400-500 ng DNA and 5 µl of a 5 pmol/µl concentrated primer solution in a total volume of 10 µl was prepared and sent for sequencing using the LightRun service. Sequencing results were analyzed using the alignment tool of the Geneious software.

4.2.15 Gel electrophoresis

DNA

Analysis of DNA was performed using agarose gel electrophoresis. 1% (w/v) agarose gels were prepared with 1x TAE buffer and supplemented with 4.25 µl SERVA DNA Strain Clear G per 100 ml. DNA samples were mixed with 10x DNA loading dye (Table 28) and loaded to the gel that was run at 120 mV for 30 min. Gels were analyzed in a Bio-Imaging-System Gene Genius using GeneSnap software by Syngene.

Protein

Analysis of proteins was performed using SDS-polyacrylamide gel electrophoresis (SDS-PAGE). SDS-PAGE gels consist of a separating and a stacking gel, in which the protein sample was focused in the stacking gel and separated according to size in the separating gel, which is possible due to the difference in pH of the two gels. In this case 15 ml of the 12% separating gel consisted of 3.75 ml 4x separating gel buffer, 4.5 ml 40% acrylamide, 150 µl 10% (w/v) ammonium peroxydisulfate (APS) and 15 µl tetramethylethylenediamine (TEMED) and 5 ml of the stacking gel consisted of 1.25 ml 4x stacking gel buffer, 0.5 ml 40% acrylamide, 50 ml

10% (w/v) APS and 5 µl TEMED. Protein samples were mixed with 5x SDS loading dye and incubated at 95 °C for 5 min to denature the proteins. For analysis 5 µl of protein sample and 3 µl of the protein ladder (low range unstained protein molecular weight maker, Thermo scientific) were loaded on the gel and the gel was run at 90 V for 45-60 min. The gel was stained using Coomassie staining solution and destained using destaining solution, subsequently the stained gel was analyzed using the Bio-Imaging-System Gene Genius.

4.2.16 Protein expression

The heterologous expression of proteins can be conducted in *E. coli* BL21 cells. The cultivation conditions for protein expression need to be optimized for each protein, however proteins used in this study were all expressed under the same conditions.[126]

A 15 ml LB preculture (supplemented with the respective antibiotic) was grown overnight at 37 °C, 200 rpm. 1 L TB (supplemented with the respective antibiotic) was inoculated with 10 ml preculture (1:100) and incubated at 37 °C, 200 rpm until the OD_{600} 0.6 to 0.8 is reached. Cells were cooled in ice for 20 min, protein expression was induced with 0.5 mM Isopropyl-β-D-thiogalactopyranoside (IPTG) and conducted at 16 °C overnight.

4.2.17 Protein purification

Protein purification is based on affinity chromatography, as the heterologous expressed proteins harbor a 6x histidine tag (His-Tag) for purification. By utilizing a nickel affinity chromatography, the tagged protein can be separated from the other proteins present in the cell.

After protein expression (see 4.2.16), cells were harvested using centrifugation at 4000 rpm, 4 °C for 5 min and resuspended in lysis buffer (see Table 27), 4 mg/ml cell pellet. Cells were lysed using sonication, for 10 min at an intensity of 80%. To remove the cell debris, cells were subsequently centrifuged at 14,000 rpm, 4 °C, 30 min. The supernatant, which contains the protein was incubated with Ni-NTA-resin, 1 ml resin per 1 l starting culture, and incubated on ice for 1 h shaking. At this step His-tagged proteins bind to the resin, the mixture can then be applied to a column, which remains the resin with the bound proteins but lets the unspecific cellular proteins run through. Resin was washed with 15 ml washing buffer 1 (Table 27), 15 ml washing buffer 2 (Table 27) and subsequently eluted with 5 times 500 µl elution buffer (Table 27). To remove the imidazole present in the elution buffer that would interfere in subsequent enzymatic assays, the protein was applied to a PD-10 column and eluted in 3.5 ml of the stated

enzymatic assay buffer. To increase the concentration, proteins were applied to a Viva spin, with a molecular weight cutoff at most half as big as the protein and centrifuged at 8000x g until the volume of the solution was reduced to 200-400 µl. Centrifugation was followed by an elution of the protein and protein concentration determination utilizing spectroscopy using the Nanophotometer P330 by *Implen*.

4.2.18 *In vitro* Assay

The enzymatic conversion of **54** to **39** catalyzed by IkaC was analyzed by an *in vitro* assay. The substrate was dissolved in methanol and added to the assay mixture as described in Table 34. The mixture was incubated at 30 °C and 300 rpm overnight and quenched at the next morning using 1:1 methanol. After centrifugation, the supernatant was measured at the LCQ-fleet (chapter 4.2.20) to evaluate the assay.

Table 34. Components of enzymatic assay.

Amount	Component
2 µM	Substrate (**54**)
10 µg	IkaC
2 mM	NADPH/NADH
200 mM/100 mM	Buffer (NaH_2PO_4/HEPES)
Add 100 µl	dH_2O

4.2.19 Natural product organic solvent extraction

Purifying small molecules from bacterial cultures based on the extraction of supernatant and pellet with different solvents. In a first step, supernatant and cell pellet were separated by centrifugation at 6000 rpm, 10 min at room temperature. The pH of the supernatant was set to 5 with 7% HCl, it was subsequently extracted three times one volume of ethyl acetate (EtOAc) and dried using $MgSO_4$. After filtering, the extract was dried *in vacuo*.

The cell pellet of a 50 ml culture was resuspended in 20 ml of a 1:1 mixture of acetone and methanol and incubated in a sonication bath for 30 min. The extract was centrifuged at 6000 rpm, 10 min at room temperature, the supernatant was transferred into a round bottom flask and dried *in vacuo*. For HPLC analysis, the extracts of a 50 ml cultures were dissolved in 500 µl methanol and filtered through a 0.22 µm syringe filter.

4.2.20 HPLC

Analytical HPLC-MS

For HPLC analysis of cell culture extracts, two different systems were used according to the expected amount of NP present in the extract. The HPLC systems could additionally be coupled to a Mass Spectrometer.

LCQ Fleet MS: HPLC samples possibly containing a low amount of NP were analyzed using the UltiMate 3000 LC System coupled to a LCQ Fleet Ion Trap Mass Spectrometer (Thermo Scientific). Obtained data were evaluated using Thermo Xcalibur Qual Browser 2.2 SP1.48 software. For the chromatographic separation, a XBridge (Waters) C-18 HPLC column (30 x 4.6 mm, 3.5 µm particle size) with solvents H_2O (A) and acetonitrile (B), both supplemented with 0.05% TFA, were utilized. The HPLC gradient is depicted in Table 35.
Jasco HPLC System: HPLC samples harboring enough NP for UV analysis were measured using a computer controlled Jasco HPLC System composed a MD-2010 Plus Multiwavelength Detector, a DG-2080-53 3-Line Degaser, two PU-2086 Plus Intelligent Pumps, a AS-2055 Plus Intelligent Sampler, a MIKA 1000 dynamic mixing chamber, 1000 µl (Portmann Instruments AG Biel-Benken), and a LCNetII/ADC. For LC-MS measurements, this system was coupled to an Expression LCMS-instrument (Advion) containing a single-quadrupole mass analyzer used in combination with a N118LA nitrogen generator (Peak Scientific) and a RV12 high vacuum pump (Edwards). The HPLC system was controlled and obtained data was analyzed using ChromPass Chromatography Data System-Software (Jasco), the mass spectrometer was controlled using a Mass Express software (Advion) and obtained data was analyzed using Data Express software (Advion). For chromatographic separation a Eurosphere C8 column with precolumn (Knauer) (10XE084E2J, 100 x 3 mm, 100-5 C8 A) was utilized. H_2O (A) and acetonitrile (B), both supplemented with 0.05% TFA, with a flowrate of 1 ml/min were used as solvents.

Table 35. HPLC gradient for LCQ-Fleet MS system. LCQ-Fleet.Meth.

Time [min]	% of A	% of B
0	80	20
6	5	95
7.5	5	95
9	80	20

Table 36. HPLC gradient for Jasco System. Eurosphere_IkaFast_CG.Meth.

Time [min]	% of A	% of B
0	60	40
2	60	40
12	0	100
16	0	100
16.5	60	40
20	60	40

Table 37. HPCL gradient for plug-and-play expression extracts. Eurosphere_Ika_lange Methode_AG.Meth.

Time [min]	% of A	% of B
0	100	0
2	100	0
10	55	45
20	55	45
20.5	0	100
24	0	100
24.5	100	0
28	100	0

Semi preparative HPLC

For the purification of NPs present in cell culture extracts a semi preparative HPLC system by Jasco was utilized. The system was composed of an UV-1575 Intelligent UV/VIS detector, two Plus Intelligent Prep. pumps, a MIKA 1000 dynamic mixing chamber, a 5000 µl sample loop (Portman Instruments AG Biel-Benken), a LC-NetII/ADC, and an injection ventil (Rheodyne). The system was controlled, and data was analyzed using Galaxie Chromatography software (Jasco). Two different columns were used, for standard purification a small Eurosphere C8 column (Knauer) (12GE084E2J, 125 x 8 mm, 100-5 C8 A) was utilized whereas for complex molecules a larger Eurosphere C8 column (Knauer) (25GE084E2J, 250 x 8 mm, 100-5 C8 A) was chosen. Both columns were connected to a SecurityGuard™ HPLC column protection (Phenomenex) (C8 4 x 3.0 mm) for the replacement of a precolumn. H_2O (A) and acetonitrile (B), both supplemented with 0.05% TFA, with a flowrate of 2-3 ml/min were used as solvents.

Table 38. HPLC gradient to purify **39** and **54**. Eurosphere_IkaFAST_CG.Meth.

Time [min]	% of A	% of B	Flow rate [ml/min]
0	60	40	2
2	60	40	2
12	0	100	2
16	0	100	2
16.5	60	40	2
20	60	40	2

Table 39. HPLC gradient to purify PoTeM derivatives. Eurosphere_Ika_lange Methode_AG.Meth.

Time [min]	% of A	% of B	Flow rate [ml/min]
0	100	0	2
2	100	0	2
10	55	45	2
20	55	45	2
20.5	0	100	2
24	0	100	2
24.5	100	0	2
28	100	0	2

Table 40. HPLC gradient to purify *spi* NP. Eurosphere_spinosa_3.Meth.

Time [min]	% of A	% of B	Flow rate [ml/min]
0	40	60	2
2	40	60	2
5	35	65	2
9	35	65	2
10	0	100	2
14	0	100	2
14.5	40	60	2
18	40	60	2

Table 41. HPLC gradient to purify PoTeMs. Eurosphere_PTM_lang_3ml-min.Meth.

Time [min]	% of A	% of B	Flow rate [ml/min]
0	95	5	3
2	95	5	3
13	50	50	3
20	50	50	3
20.5	0	100	3
24	0	100	3
24.5	95	5	3
28	95	5	3

NMR

NMR data was recorded using the AVHD500 spectrometer for ^{1}H NMR-spectra and the AV500-cryo spectrometer for ^{13}C NMR-spectra. Both systems were utilized for recording 2D NMR experiments.

Abbreviations

2xYT	2 times yeast, tryptone
6-DEB	6-deoxyerytronolide B
AA	Amino acid
ACN	Acetonitrile
APS	Ammonium peroxydisulfate
AT	Acyltransferase
A	Adenylation
BAC	Bacterial artificial chromosome
BGC	Biosynthetic gene cluster
C	Condensation
CASO	Casein, Soy
CATCH	Cas9-Associated Targeting of Chromosomal segments
CoA	Coenzyme A
CTAB	Cetyltrimethylammonium bromide
db	deg_back, back part of *deg* cluster (*degC*)
deg	PoTeM cluster from *S. degradans*
df	deg_front, front part of *deg* cluster (*degA*)
DH	Dehydratase
DiPaC	Direct Pathway Cloning
DMSO	Dimethyl sulfoxide
DNA	Deoxyribonucleic acid
DOS	Diversity-oriented synthesis
DPG	L-3,4-dihydroxyphenylglycine
DSMZ	German Collection of Microorganisms and Cell Cultures GmbH
E	Epimerization
e.g.	Exempli gratia
eDNA	Environmental DNA
EDTA	ethylenediaminetetraacetic acid
EL	*Eggerthella lenta*
ER	Enoyl reductase
et al.	Et alia
EtOAc	ethyl acetate
EtOH	Ethanol
FAD	Flavin-Adenine-Dinucleotide
FAS	Fatty acid synthase

fb	fer_back, back part of *fer* cluster (*ferC*)
fer	PoTeM cluster from *P. fermentans*
ff	fer_front, front part of *fer* cluster (*ferA*)
FPLC	Fast protein liquid chromatography
gDNA	Genomic DNA
GYM	Glucose, yeast, malt
HEPES	4-(2-hydroxyethyl)-1-piperazineethanesulfonic acid
His-Tag	6x histidine motive
HMG	3-Hydroxy-3-Methylglutaryl
HPG	L-4-hydroxyphenylglycine
HPLC	High performance liquid chromatography
HR	Homologous recombination
HSAF	Heat-stable antifungal factor
HTS	High throughput screening
ika	ikarugamycin BGC
iPKS	Iterative PKS
IPTG	Isopropyl-β-D-thiogalactopyranoside
ISP-4	International *Streptomyces* project 4
KLD	Kinase, ligase, DpnI
KR	Keto reductase
KS	Keto synthase
LB	Luria Bertani
LCHR	Linear plus circular homologous recombination
LC-MS	Liquid-Chromatography-Mass-Spectrometry
LLHR	Linear plus linear homologous recombination
MCS	Multiple cloning site
MeOH	Methanol
mod.	modified
MS-agar	maltose, soy-agar
MS	Mass spectrometry
MT	Methyltransferase
NADH	Nicotinamide adenine dinucleotide
NADPH	Nicotinamide adenine dinucleotide phosphate
NCE	New chemical entities
NEB	New England Biolabs
NMR	Nuclear magnetic resonance
NP	Natural product

NRP	Non-ribosomal peptide
NRPS	Non-ribosomal peptide synthase
OD	Optical density
ORF	Open reading frame
OSMAC	One Strain - Many Compounds
PAC	Phage derived artificial chromosome
PAGE	Polyacrylamide gel electrophorese
PCP	Peptide carrier protein
PCR	Polymerase chain reaction
PEG	Polyethylene glycol
PK	Polyketide
PKS	Polyketide synthase
PoTeM	Polycyclic tetramate macrolactams
RBS	Ribosomal binding site
RNA	Ribonucleic acid
ros	PoTeM cluster from *S. roseosporus*
sb	spi_back, back part of *spi* cluster (*spiDE*)
SDS	Sodium dodecyl sulfate
sf	spi_front, front part of spi cluster (*spiAB*)
SLIC	sequence and ligation independent cloning
SOB	Super optimal broth
SOC	Super optimal broth supplemented with glucose
sp.	Species
spi	PoTeM cluster from *S. spinosa*
T	Thiolation
TAR	Transformation-associated recombination
TB	Terrific broth
TE	Thioesterase
TEMED	Tetramethylethylenediamine
TES	2-[[1,3-dihydroxy-2-(hydroxymethyl)propan-2-yl]amino]ethanesulfonic acid
TFA	Trifluoro acetic acid
ThioR	Thiostreptone resistance
Tris	Tris(hydroxymethyl)aminomethane
Tü6314	PoTeM cluster from *Streptomyces sp.* Tü6314
UV	Ultraviolet
v/v	Volume per volume

w/v	Weight per volume
w/w	Weight per weight
WHO	World health organization
wt	Wildtype
XP	Xylanimonas cellulosilytica
β-OH-Cl-Tyr	β-hydroxychloro tyrosine

Bacterial strains

E. coli	*Escherichia coli*
L. capsici	*Lysobacter capsici*
P. fermentans	*Pelosinus fermentans*
S. albus	*Streptomyces albus*
S. coelicolor	*Streptomyces coelicolor*
S. degradans	*Saccharophagus degradans*
S. lividans	*Streptomyces lividans*
S. spinosa	*Saccharopolispora spinosa*

Units

bp	Base pair
d	Day
g	Gramm
h	Hour
M	Molar
min	Minute
rpm	Rounds per minute
sec	Second
V	Volt
xg	gravitational acceleration (9.80665 m/s^2)
p	pico
n	nano
µ	micro
m	milli
k	kilo

Antibiotics

Amp	Ampicillin
Apra	Apramycin
Cam	Chloramphenicol
Kan	Kanamycin
NA	Nalidixic acid
Thio	Thiostreptone

Amino acids

Alanine	Ala	A
Arginine	Arg	R
Asparagine	Asn	N
Aspartic acid	Asp	D
Cysteine	Cys	C
Glutamine	Gln	Q
Glutamic acid	Glu	E
Glycine	Gly	G
Histidine	His	H
Isoleucine	Ile	I
Leucine	Leu	L
Lysine	Lys	K
Methionine	Met	M
Phenylalanine	Phe	F
Proline	Pro	P
Serine	Ser	S
Threonine	Thr	T
Tryptophan	Trp	W
Tyrosine	Tyr	Y
Valine	Val	V

Bibliography

[1] B. Schäfer, *Naturstoffe der chemischen Industrie*, Springer Spektrum, **2006**.

[2] H.-U. Lammel, in *Kindler Kompakt Klassiker der Naturwissenschaften*, J.B. Metzler, Stuttgart, **2016**, 43-44.

[3] H. Dreser, *Archiv für die gesamte Physiologie des Menschen und der Tiere* **1899**, *76*, 306-318.

[4] W. Sneader, *BMJ (Clinical research ed.)* **2000**, *321*, 1591-1594.

[5] F. W. Serturner, *Trommsdorffs Journal der Pharmazie* **1805**, *14*, 47-98.

[6] C. Sepúlveda, A. Marlin, T. Yoshida, A. Ullrich, *Journal of pain and symptom management* **2002**, *24*, 91-96.

[7] H. King, *Journal of the Chemical Society (Resumed)* **1935**, 1381-1389.

[8] W. C. Bowman, *British Journal of Pharmacology* **2006**, *147*, S277-S286.

[9] A. L. Harvey, *Toxicon* **2014**, *92*, 193-200.

[10] G. M. Cragg, D. J. Newman, *Biochimica et Biophysica Acta (BBA) - General Subjects* **2013**, *1830*, 3670-3695.

[11] D. A. Dias, S. Urban, U. Roessner, *Metabolites* **2012**, *2*, 303-336.

[12] A. Fleming, *British journal of experimental pathology* **1929**, *10*, 226.

[13] A. Fleming, *Reviews of infectious diseases* **1980**, *2*, 129-139.

[14] A. Schatz, E. Bugle, S. A. Waksman, *Proceedings of the Society for Experimental Biology and Medicine* **1944**, *55*, 66-69.

[15] J. F. Murray, D. E. Schraufnagel, P. C. Hopewell, *Annals of the American Thoracic Society* **2015**, *12*, 1749-1759.

[16] L. Katz, R. H. Baltz, *Journal of Industrial Microbiology & Biotechnology* **2016**, *43*, 155-176.

[17] A. L. Demain, *Journal of Industrial Microbiology and Biotechnology* **2006**, *33*, 486-495.

[18] J. Geraci, F. Heilman, D. Nichols, W. Wellman, G. Ross, R. Dorothy, in *Proc. Staff, Meetings Mayo Clinic, Vol. 31*, **1956**, pp. 564-582.

[19] R. Anderson, H. Higgins Jr, C. Pettinga, *Cincinnati Journal of Medicine* **1961**, *42*, 49-60.

[20] M. P. Wilhelm, L. Estes, *Mayo Clinic Proceedings* **1999**, *74*, 928-935.

[21] J. McGuire, W. Boniece, C. Higgens, M. Hoehn, W. Stark, J. Westhead, R. Wolfe, *Antibiotics & Chemotherapy* **1961**, *11*, 320-327.

[22] R. Hamill, M. Haney Jr, M. Stamper, P. Wlley, *Antibiotics & Chemotherapy* **1961**, *11*, 328-334.

[23] M. Debono, M. Barnhart, C. B. Carrell, J. A. Hoffmann, J. L. Occolowitz, B. J. Abbott, D. S. Fukuda, R. L. Hamill, *The Journal of Antobiotics* **1987**, *40*, 761-777.

[24] T. C. Sorrell, D. R. Packham, S. Shanker, M. Foldes, R. Munro, *Annals of Internal Medicine* **1982**, *97*, 344-350.

[25] D. P. Levine, *Clinical infectious diseases : an official publication of the Infectious Diseases Society of America* **2006**, *42 Suppl 1*, S5-12.

[26] L. Barnes, E. Ose, F. Gossett, *Poultry Science* **1960**, *39*, 1376-1381.

[27] D. W. Scott, W. H. Miller Jr, S. M. Cayatte, M. S. Bagladi, *The Canadian Veterinary Journal* **1994**, *35*, 617.

[28] J. Vicca, D. Maes, L. Jonker, A. de Kruif, F. Haesebrouck, *Veterinary Record* **2005**, *156*, 606-610.

[29] K. Lamp, M. Rybak, E. Bailey, G. Kaatz, *Antimicrobial Agents and Chemotherapy* **1992**, *36*, 2709-2714.

[30] R. D. Arbeit, D. Maki, F. P. Tally, E. Campanaro, B. I. Eisenstein, D. 98-01, -. Investigators, *Clinical Infectious Diseases* **2004**, *38*, 1673-1681.

[31] B. I. Eisenstein, F. B. Oleson Jr, R. H. Baltz, *Clinical Infectious Diseases* **2010**, *50*, S10-S15.

[32] A. Endo, *The Journal of antibiotics* **1979**, *32*, 852-854.

[33] A. W. Alberts, J. Chen, G. Kuron, V. Hunt, J. Huff, C. Hoffman, J. Rothrock, M. Lopez, H. Joshua, E. Harris, A. Patchett, R. Monaghan, S. Currie, E. Stapley, G. Albers-Schonberg, O. Hensens, J. Hirshfield, K. Hoogsteen, J. Liesch, J. Springer, *Proceedings of the National Academy of Sciences* **1980**, *77*, 3957-3961.

[34] D. J. Newman, G. M. Cragg, *Journal of Natural Products* **2012**, *75*, 311-335.

[35] D. J. Newman, G. M. Cragg, *Journal of Natural Products* **2016**, *79*, 629-661.

[36] P. F. Chan, D. J. Holmes, D. J. Payne, *Drug Discovery Today: Therapeutic Strategies* **2004**, *1*, 519-527.

[37] T. Henkel, R. M. Brunne, H. Müller, F. Reichel, *Angewandte Chemie International Edition* **1999**, *38*, 643-647.

[38] S. Shang, D. S. Tan, *Current opinion in chemical biology* **2005**, *9*, 248-258.

[39] D. S. Tan, *Nature chemical biology* **2005**, *1*, 74.

[40] I. T. Paulsen, C. M. Press, J. Ravel, D. Y. Kobayashi, G. S. A. Myers, D. V. Mavrodi, R. T. DeBoy, R. Seshadri, Q. Ren, R. Madupu, R. J. Dodson, A. S. Durkin, L. M. Brinkac, S. C. Daugherty, S. A. Sullivan, M. J. Rosovitz, M. L. Gwinn, L. Zhou, D. J. Schneider, S. W. Cartinhour, W. C. Nelson, J. Weidman, K. Watkins, K. Tran, H. Khouri, E. A. Pierson, L. S. Pierson Iii, L. S. Thomashow, J. E. Loper, *Nature Biotechnology* **2005**, *23*, 873.

[41] M. A. Fischbach, C. T. Walsh, *Chemical Reviews* **2006**, *106*, 3468-3496.

[42] T. Kaneda, J. C. Butte, S. B. Taubman, J. W. Corcoran, *Journal of Biological Chemistry* **1962**, *237*, 322-328.

[43] D. E. Cane, H. Hasler, T.-C. Liang, *Journal of the American Chemical Society* **1981**, *103*, 5960-5962.

[44] R. Stanzak, P. Matsushima, R. Baltz, R. Rao, *Bio/technology* **1986**, *4*, 229-232.

[45] S. Omura, *Macrolide antibiotics: chemistry, biology, and practice*, Elsevier, **2002**.

[46] J. Cortes, S. F. Haydock, G. A. Roberts, D. J. Bevitt, P. F. Leadlay, *Nature* **1990**, *348*, 176-178.

[47] S. Donadio, M. J. Staver, J. B. McAlpine, S. J. Swanson, L. Katz, *Science* **1991**, *252*, 675-679.

[48] A. S. Eustáquio, R. P. McGlinchey, Y. Liu, C. Hazzard, L. L. Beer, G. Florova, M. M. Alhamadsheh, A. Lechner, A. J. Kale, Y. Kobayashi, K. A. Reynolds, B. S. Moore, *Proceedings of the National Academy of Sciences* **2009**, *106*, 12295-12300.

[49] R. H. Lambalot, A. M. Gehring, R. S. Flugel, P. Zuber, M. LaCelle, M. A. Marahiel, R. Reid, C. Khosla, C. T. Walsh, *Chemistry & Biology* **1996**, *3*, 923-936.

[50] C. T. Walsh, R. V. O'Brien, C. Khosla, *Angewandte Chemie International Edition* **2013**, *52*, 7098-7124.

[51] S. J. Hammond, M. P. Williamson, D. H. Williams, L. D. Boeck, G. G. Marconi, *Journal of the Chemical Society, Chemical Communications* **1982**, 344-346.

[52] M. Zmijewski, B. Briggs, R. Logan, L. D. Boeck, *Antimicrobial agents and chemotherapy* **1987**, *31*, 1497-1501.

[53] B. K. Hubbard, M. G. Thomas, C. T. Walsh, *Chemistry & biology* **2000**, *7*, 931-942.

[54] H.-T. Chiu, B. K. Hubbard, A. N. Shah, J. Eide, R. A. Fredenburg, C. T. Walsh, C. Khosla, *Proceedings of the National Academy of Sciences* **2001**, *98*, 8548-8553.

[55] B. K. Hubbard, C. T. Walsh, *Angewandte Chemie International Edition* **2003**, *42*, 730-765.

[56] D. H. Williams, M. P. Williamson, D. W. Butcher, S. J. Hammond, *Journal of the American Chemical Society* **1983**, *105*, 1332-1339.

[57] D. P. O'Brien, P. N. Kirkpatrick, S. W. O'Brien, T. Staroske, T. I. Richardson, D. A. Evans, A. Hopkinson, J. B. Spencer, D. H. Williams, *Chemical Communications* **2000**, 103-104.

[58] H. C. Losey, M. W. Peczuh, Z. Chen, U. S. Eggert, S. D. Dong, I. Pelczer, D. Kahne, C. T. Walsh, *Biochemistry* **2001**, *40*, 4745-4755.

[59] D. Boettger, C. Hertweck, *Chembiochem : a European journal of chemical biology* **2013**, *14*, 28-42.

[60] J. F. Leslie, W. Marasas, G. S. Shephard, E. W. Sydenham, S. Stockenström, P. G. Thiel, *Applied and environmental microbiology* **1996**, *62*, 1182-1187.

[61] W. C. Gelderblom, W. F. Marasas, P. S. Steyn, P. G. Thiel, K. J. van der Merwe, P. H. van Rooyen, R. Vleggaar, P. L. Wessels, *Journal of the Chemical Society, Chemical Communications* **1984**, 122-124.

[62] Z. Song, R. J. Cox, C. M. Lazarus, T. J. Simpson, *Chembiochem : a European journal of chemical biology* **2004**, *5*, 1196-1203.

[63] L. Hendrickson, C. R. Davis, C. Roach, T. Aldrich, P. C. McAda, C. D. Reeves, *Chemistry & biology* **1999**, *6*, 429-439.

[64] D. Boettger, H. Bergmann, B. Kuehn, E. Shelest, C. Hertweck, *Chembiochem : a European journal of chemical biology* **2012**, *13*, 2363-2373.

[65] R. Fleischmann, M. Adams, O. White, R. Clayton, E. Kirkness, A. Kerlavage, C. Bult, J. Tomb, B. Dougherty, J. Merrick, e. al., *Science* **1995**, *269*, 496-512.

[66] S. D. Bentley, K. F. Chater, A.-M. Cerdeño-Tárraga, G. L. Challis, N. Thomson, K. D. James, D. E. Harris, M. A. Quail, H. Kieser, D. Harper, *Nature* **2002**, *417*, 141.

[67] H. B. Bode, B. Bethe, R. Hofs, A. Zeeck, *Chembiochem : a European journal of chemical biology* **2002**, *3*, 619-627.

[68] Y.-Q. Shen, J. Heim, N. Solomon, S. Wolfe, A. Demain, *The Journal of antibiotics* **1984**, *37*, 503-511.

[69] A. F. Brana, S. Wolfe, A. L. Demain, *Archives of microbiology* **1986**, *146*, 46-51.

[70] R. Höfs, M. Walker, A. Zeeck, *Angewandte Chemie International Edition* **2000**, *39*, 3258-3261.

[71] C. Puder, P. Krastel, A. Zeeck, *Journal of natural products* **2000**, *63*, 1258-1260.

[72] C. Puder, S. Loya, A. Hizi, A. Zeeck, *Journal of natural products* **2001**, *64*, 42-45.

[73] H. G. Floss, *Journal of biotechnology* **2006**, *124*, 242-257.

[74] S. Omura, H. Ikeda, F. Malpartida, H. Kieser, D. Hopwood, *Antimicrobial agents and chemotherapy* **1986**, *29*, 13-19.

[75] P. J. Solenberg, P. Matsushima, D. R. Stack, S. C. Wilkie, R. C. Thompson, R. H. Baltz, *Chemistry & biology* **1997**, *4*, 195-202.

[76] K. T. Nguyen, D. Ritz, J.-Q. Gu, D. Alexander, M. Chu, V. Miao, P. Brian, R. H. Baltz, *Proceedings of the National Academy of Sciences* **2006**, *103*, 17462-17467.

[77] K. T. Nguyen, X. He, D. C. Alexander, C. Li, J.-Q. Gu, C. Mascio, A. Van Praagh, L. Mortin, M. Chu, J. A. Silverman, *Antimicrobial agents and chemotherapy* **2010**, *54*, 1404-1413.

[78] R. McDaniel, A. Thamchaipenet, C. Gustafsson, H. Fu, M. Betlach, M. Betlach, G. Ashley, *Proceedings of the National Academy of Sciences* **1999**, *96*, 1846-1851.

[79] C. T. Walsh, *Chembiochem : a European journal of chemical biology* **2002**, *3*, 124-134.

[80] F. Y. Chang, S. F. Brady, *Chembiochem : a European journal of chemical biology* **2014**, *15*, 815-821.

[81] A. C. Jones, B. Gust, A. Kulik, L. Heide, M. J. Buttner, M. J. Bibb, *PloS one* **2013**, *8*, e69319.

[82] Y. Zhang, F. Buchholz, J. P. Muyrers, A. F. Stewart, *Nature genetics* **1998**, *20*, 123.

[83] J. Fu, X. Bian, S. Hu, H. Wang, F. Huang, P. M. Seibert, A. Plaza, L. Xia, R. Müller, A. F. Stewart, *Nature biotechnology* **2012**, *30*, 440.

[84] C. Su, X. Zhao, R. Qiu, L. Tang, *Pharmaceutical biology* **2015**, *53*, 269-274.

[85] J. Yin, M. Hoffmann, X. Bian, Q. Tu, F. Yan, L. Xia, X. Ding, A. F. Stewart, R. Müller, J. Fu, *Scientific reports* **2015**, *5*, 15081.

[86] N. Kouprina, V. Larionov, *Nature protocols* **2008**, *3*, 371.

[87] K. Yamanaka, K. A. Reynolds, R. D. Kersten, K. S. Ryan, D. J. Gonzalez, V. Nizet, P. C. Dorrestein, B. S. Moore, *Proceedings of the National Academy of Sciences* **2014**, *111*, 1957-1962.

[88] W. Jiang, X. Zhao, T. Gabrieli, C. Lou, Y. Ebenstein, T. F. Zhu, *Nature communications* **2015**, *6*, 8101.

[89] D. G. Gibson, L. Young, R.-Y. Chuang, J. C. Venter, C. A. Hutchison Iii, H. O. Smith, *Nature Methods* **2009**, *6*, 343.

[90] K. Jomon, Y. Kuroda, M. Ajisaka, H. Sakai, *The Journal of antibiotics* **1972**, *25*, 271-280.

[91] K. Hasumi, C. Shinohara, S. Naganuma, A. Endo, *European journal of biochemistry* **1992**, *205*, 841-846.

[92] T. Luo, B. L. Fredericksen, K. Hasumi, A. Endo, J. V. Garcia, *Journal of virology* **2001**, *75*, 2488-2492.

[93] A. Moscatelli, F. Ciampolini, S. Rodighiero, E. Onelli, M. Cresti, N. Santo, A. Idilli, *Journal of cell science* **2007**, *120*, 3804-3819.

[94] E. Onelli, C. Prescianotto-Baschong, M. Caccianiga, A. Moscatelli, *Journal of experimental botany* **2008**, *59*, 3051-3068.

[95] V. Bandmann, J. D. Muller, T. Kohler, U. Homann, *FEBS letters* **2012**, *586*, 3626-3632.

[96] V. Bandmann, U. Homann, *The Plant journal : for cell and molecular biology* **2012**, *70*, 578-584.

[97] S. P. Gunasekera, M. Gunasekera, P. McCarthy, *The Journal of Organic Chemistry* **1991**, *56*, 4830-4833.

[98] H. Shigemori, M. A. Bae, K. Yazawa, T. Sasaki, J. Kobayashi, *The Journal of Organic Chemistry* **1992**, *57*, 4317-4320.

[99] S. Kanazawa, N. Pusetani, S. Matsunaga, *Tetrahedron Letters* **1993**, *34*, 1065-1068.

[100] M. Jakobi, G. Winkelmann, *The Journal of Antobiotics* **1996**, *49*, 1101-1104.

[101] P. R. Graupner, S. Thornburgh, J. T. Mathieson, E. L. Chapin, G. M. Kemmitt, J. M. Brown, C. E. Snipes, *The Journal of Antobiotics* **1997**, *50*, 1014-1019.

[102] S. Li, L. Du, G. Yuen, S. D. Harris, *Mol Biol Cell* **2006**, *17*, 1218-1227.

[103] F. Yu , K. Zaleta-Rivera, X. Zhu, J. Huffman, J. C. Millet, S. D. Harris, G. Yuen, X.-C. Li, L. Du, *Antimicrobial Agents and Chemotherapy* **2007**, *51*, 64-72.

[104] M.-A. Bae, K. Yamada, Y. Ijuin, T. Tsuji, K. Yazawa, D. Uemura, in *Heterocyclic Communications, Vol. 2*, **1996**, p. 315.

[105] R. J. Capon, C. Skene, E. Lacey, J. H. Gill, D. Wadsworth, T. Friedel, *Journal of Natural Products* **1999**, *62*, 1256-1259.

[106] Y. Hashidoko, T. Nakayama, Y. Homma, S. Tahara, *Tetrahedron Letters* **1999**, *40*, 2957-2960.

[107] M. Bertasso, M. Holzenkampfer, A. Zeeck, E. Stackebrandt, W. Beil, H. P. Fiedler, *The Journal of antibiotics* **2003**, *56*, 364-371.

[108] J. A. Blodgett, D. C. Oh, S. Cao, C. R. Currie, R. Kolter, J. Clardy, *Proc Natl Acad Sci U S A* **2010**, *107*, 11692-11697.

[109] L. Lou, G. Qian, Y. Xie, J. Hang, H. Chen, K. Zaleta-Rivera, Y. Li, Y. Shen, P. H. Dussault, F. Liu, L. Du, *J Am Chem Soc* **2011**, *133*, 643-645.

[110] H. Chen, L. Du, *Applied microbiology and biotechnology* **2016**, *100*, 541-557.

[111] Y. Li, H. Chen, Y. Ding, Y. Xie, H. Wang, R. L. Cerny, Y. Shen, L. Du, *Angew Chem Int Ed Engl* **2014**, *53*, 7524-7530.

[112] J. Antosch, F. Schaefers, T. A. Gulder, *Angew Chem Int Ed Engl* **2014**, *53*, 3011-3014.

[113] G. Zhang, W. Zhang, Q. Zhang, T. Shi, L. Ma, Y. Zhu, S. Li, H. Zhang, Y. L. Zhao, R. Shi, C. Zhang, *Angew Chem Int Ed Engl* **2014**, *53*, 4840-4844.

[114] S. Cao, J. A. Blodgett, J. Clardy, *Organic Letters* **2010**, *12*, 4652-4654.

[115] Y. Luo, H. Huang, J. Liang, M. Wang, L. Lu, Z. Shao, R. E. Cobb, H. Zhao, *Nat Commun* **2013**, *4*, 2894.

[116] C. Olano, I. Garcia, A. Gonzalez, M. Rodriguez, D. Rozas, J. Rubio, M. Sanchez-Hidalgo, A. F. Brana, C. Mendez, J. A. Salas, *Microb Biotechnol* **2014**, *7*, 242-256.

[117] Y. Li, J. Huffman, Y. Li, L. Du, Y. Shen, *MedChemComm* **2012**, *3*.

[118] C. Greunke, J. Antosch, T. A. Gulder, *Chem Commun (Camb)* **2015**, *51*, 5334-5336.

[119] K. Kyeremeh, K. S. Acquah, A. Sazak, W. Houssen, J. Tabudravu, H. Deng, M. Jaspars, *Mar Drugs* **2014**, *12*, 999-1012.

[120] D. P. Labeda, J. R. Doroghazi, K. S. Ju, W. W. Metcalf, *Int J Syst Evol Microbiol* **2014**, *64*, 894-900.

[121] D. A. K. Hopwood, T.; Wright, H. M.; Bibb M. J., *Journal of General Microbiology* **1982**, *129*, 2257-2269.

[122] G. Wang, T. Hosaka, K. Ochi, *Applied and environmental microbiology* **2008**, *74*, 2834-2840.

[123] M. J. Bibb, J. White, J. M. Ward, G. R. Janssen, *Molecular microbiology* **1994**, *14*, 533-545.

[124] L. Kaysser, P. Bernhardt, S.-J. Nam, S. Loesgen, J. G. Ruby, P. Skewes-Cox, P. R. Jensen, W. Fenical, B. S. Moore, *Journal of the American Chemical Society* **2012**, *134*, 11988-11991.

[125] P. Holloway, *Analytical biochemistry* **1973**, *53*, 304-308.

[126] C. Greunke, A. Glockle, J. Antosch, T. A. Gulder, *Angew Chem Int Ed Engl* **2017**, *56*, 4351-4355.

[127] T. Kieser, M. J. Bibb, M. J. Buttner, K. F. Charter, D. A. Hopwood, *John Innes Foundation, Norwich, United Kingdom* **2000**.

[128] B. Bilyk, L. Horbal, A. Luzhetskyy, *Microbial cell factories* **2017**, *16*, 5.

[129] A. C. Jones, S. Ottilie, A. S. Eustáquio, D. J. Edwards, L. Gerwick, B. S. Moore, W. H. Gerwick, *The FEBS journal* **2012**, *279*, 1243-1251.

[130] C. Greunke, E. R. Duell, P. M. D'Agostino, A. Glöckle, K. Lamm, T. A. M. Gulder, *Metabolic Engineering* **2018**, *47*, 334-345.

[131] M. Myronovskyi, A. Luzhetskyy, *Natural product reports* **2016**, *33*, 1006-1019.

[132] Z. Shao, G. Rao, C. Li, Z. Abil, Y. Luo, H. Zhao, *ACS Synthetic Biology* **2013**, *2*, 662-669.

[133] W. Wang, X. Li, J. Wang, S. Xiang, X. Feng, K. Yang, *Applied and environmental microbiology* **2013**, *79*, 4484-4492.

[134] G. Labes, M. Bibb, W. Wohlleben, *Microbiology* **1997**, *143*, 1503-1512.

[135] Y. Luo, L. Zhang, K. W. Barton, H. Zhao, *ACS Synthetic Biology* **2015**, *4*, 1001-1010.

[136] S. Saha, W. Zhang, G. Zhang, Y. Zhu, Y. Chen, W. Liu, C. Yuan, Q. Zhang, H. Zhang, L. Zhang, W. Zhang, C. Zhang, *Chem Sci* **2017**, *8*, 1607-1612.

[137] H. L. Yu, S. H. Jiang, X. L. Bu, J. H. Wang, J. Y. Weng, X. M. Yang, K. Y. He, Z. G. Zhang, P. Ao, J. Xu, M. J. Xu, *Sci Rep* **2017**, *7*, 40689.

[138] Y. Qi, E. Ding, J. A. V. Blodgett, *ACS Synth Biol* **2018**, *7*, 357-362.

[139] P. M. D'Agostino, T. A. M. Gulder, *ACS Synth Biol* **2018**, *7*, 1702-1708.

[140] F. Wright, M. J. Bibb, *Gene* **1992**, *113*, 55-65.

[141] D. J. G. MacNeil, K. M.; Ruby C. L.; Dezeny, G; Gibbons, P. H.; MacNeil, T., *Gene* **1992**, *111*, 61-68.

[142] K. A. Datsenko, B. L. Wanner, *Proceedings of the National Academy of Sciences* **2000**, *97*, 6640-6645.

[143] T. Baba, T. Ara, M. Hasegawa, Y. Takai, Y. Okumura, M. Baba, K. A. Datsenko, M. Tomita, B. L. Wanner, H. Mori, *Molecular Systems Biology* **2006**, *2*, 2006.0008.

[144] M. Z. Li, S. J. Elledge, in *Gene synthesis*, Springer, **2012**, pp. 51-59.

Appendix

Vector Maps

Figure A1. Vector map of expression vector pSET152_ermE.

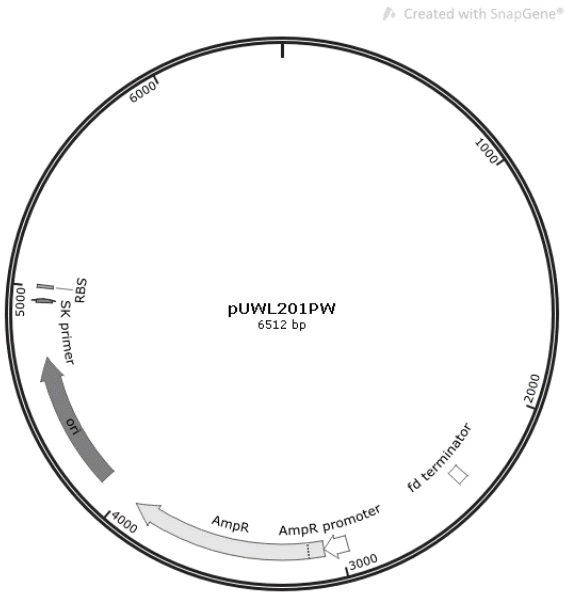


Figure A2. Vector map of expression vector pUWL201PW.

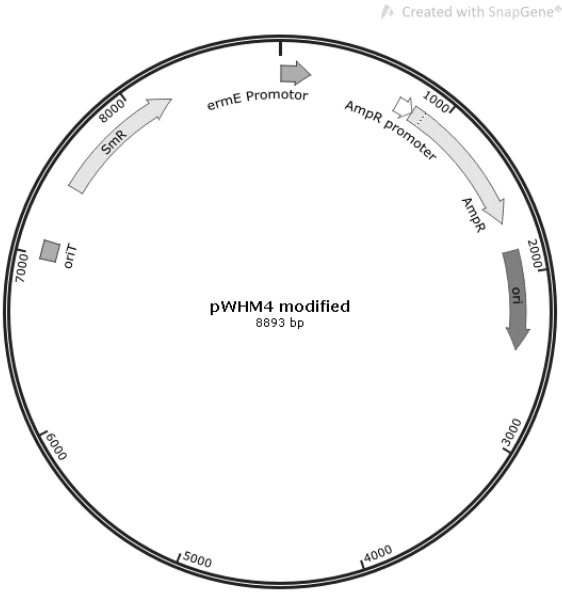


Figure A3. Vector map of expression vector pWHM4*.

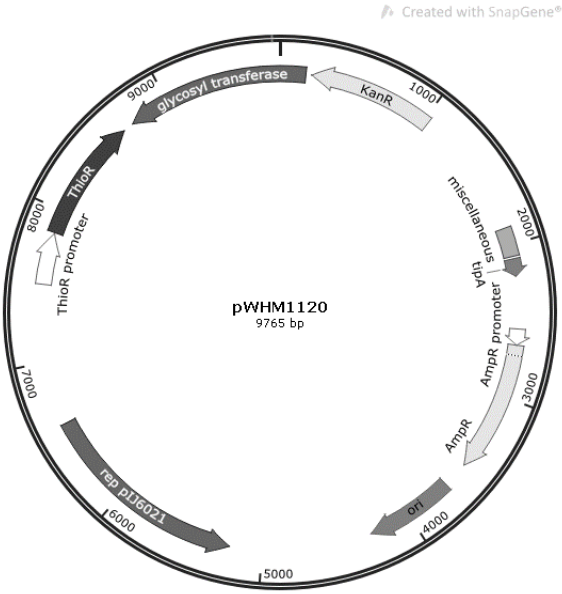


Figure A4. Vector map of expression vector pWHM1120.

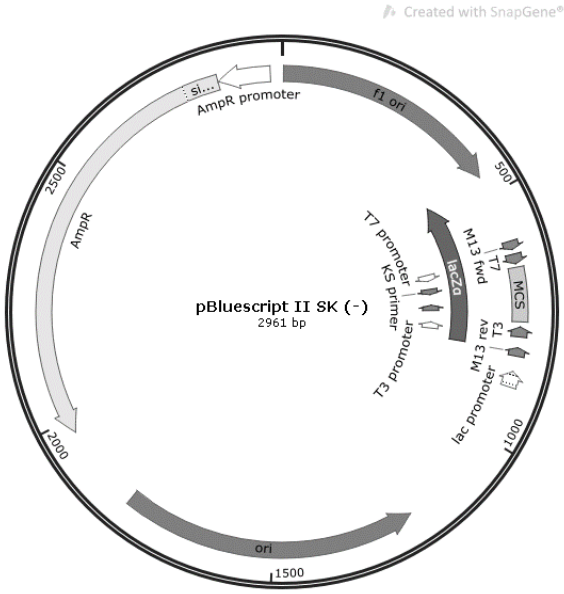


Figure A5. Vector map of cloning vector pBluescript.

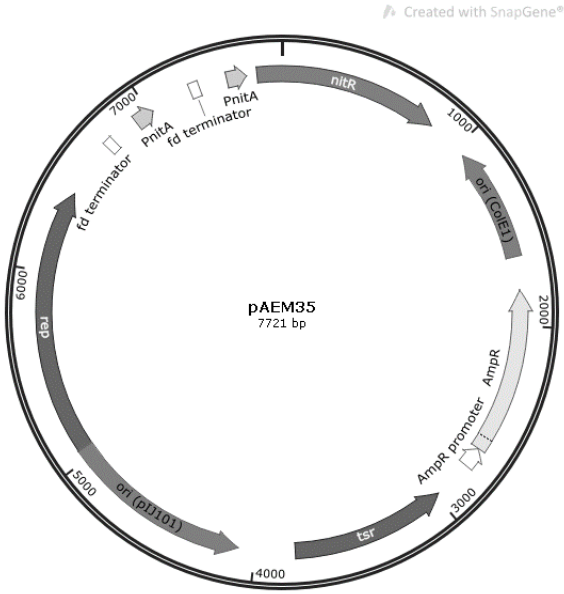


Figure A6. Vector map of expression vector pAEM35.

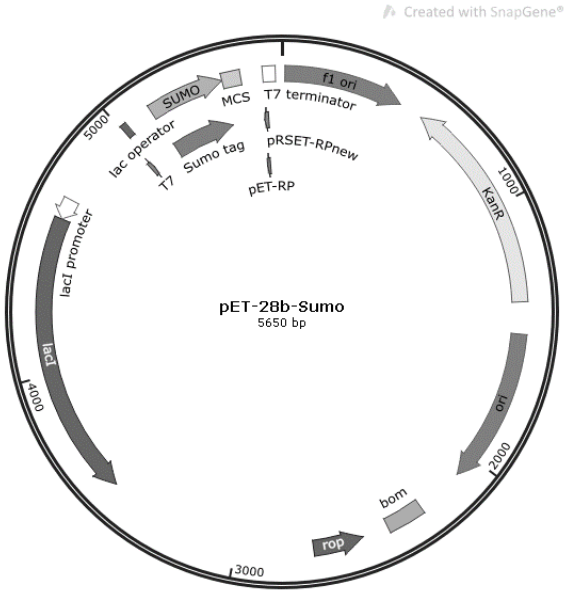


Figure A7. Vector map of cloning vector pET-28b-Sumo.

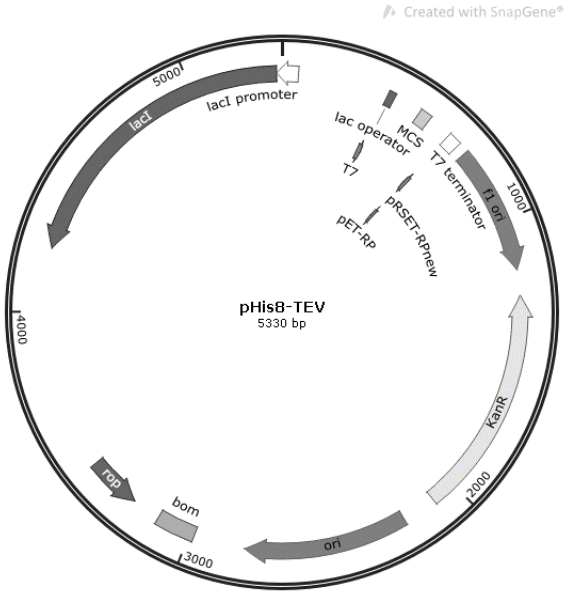


Figure A8. Vector map of cloning vector pHis8-Tev.

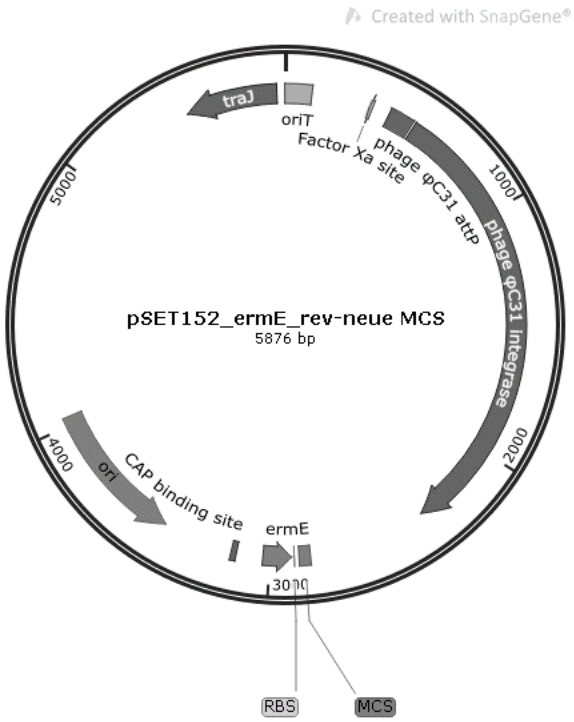


Figure A9. Vector map of expression vector pSET152_ermE_new_MCS.

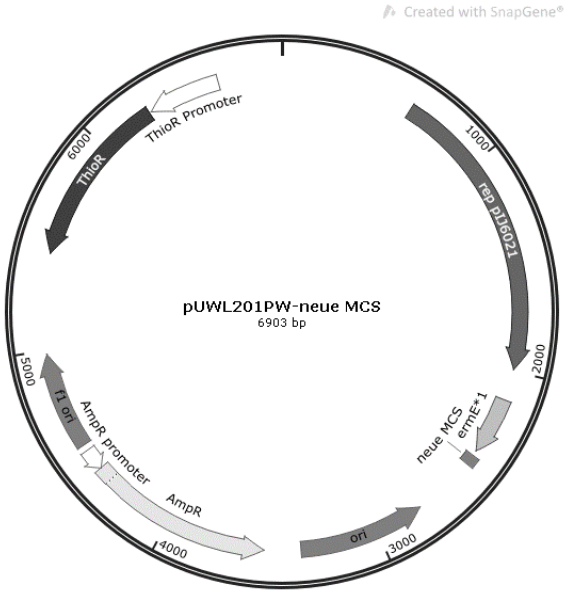


Figure A10. Vector map of expression vector pUWL201PW_new_MCS.

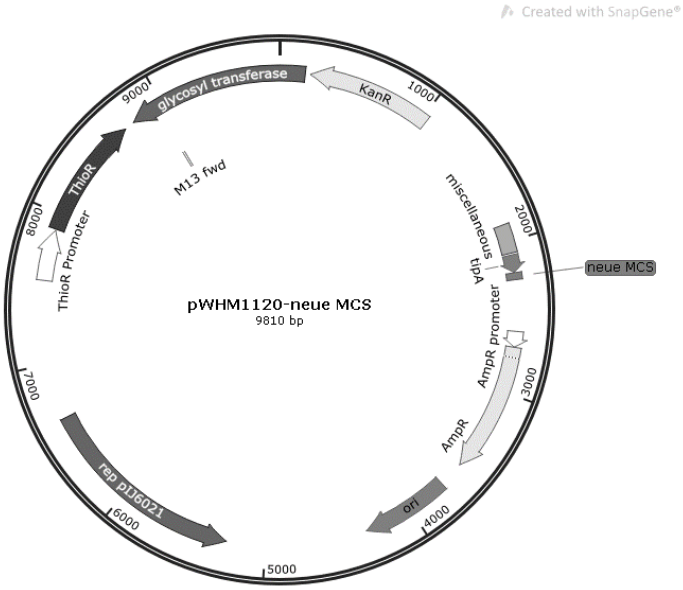


Figure A11. Vector map of expression vector pWHM1120_new_MCS.

Figure A12. Vector map of knockout plasmid pSET152_ermE_tetR_KO.

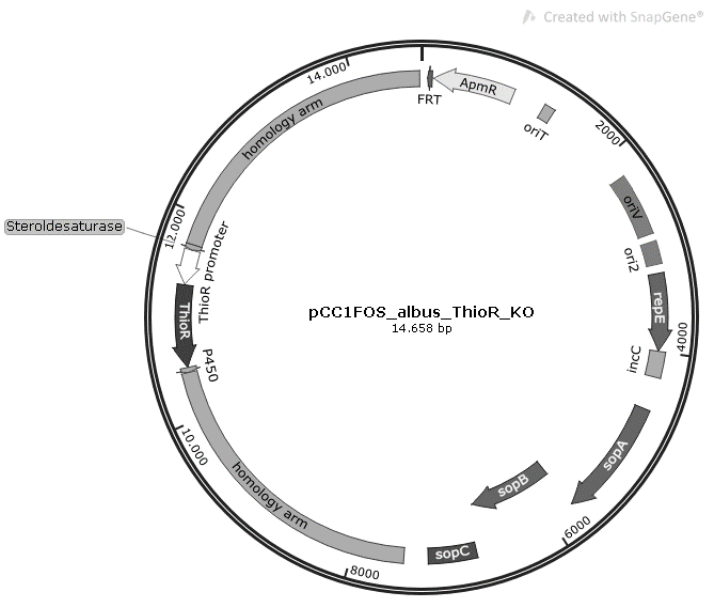


Figure A 13. Vector map of knockout plasmid pCC1FOS_albus_ThioR_KO.

Cloning Results

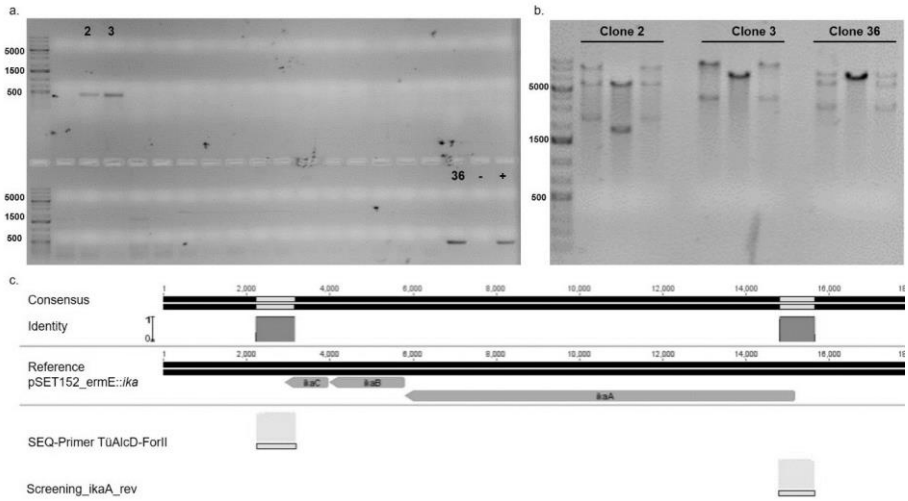


Figure A14. Cloning verification for pSET152_ermE::*ika*. a. Results of the colony PCR with possible positive Clones 2, 3, and 36, negative control conducted with water, positive control 1 µl of Gibson assembly reaction. b. Results of analytical restriction digest with all clones, as only clone 2 showed the predicted restriction pattern this clone was submitted for sequencing. c. Sequencing results of Clone 2 with Primers: SEQ-Primer TüAlcD-ForII and Screening _IKA_rev, verifying a successful cloning of the *ika* BCG into pSET152_ermE.

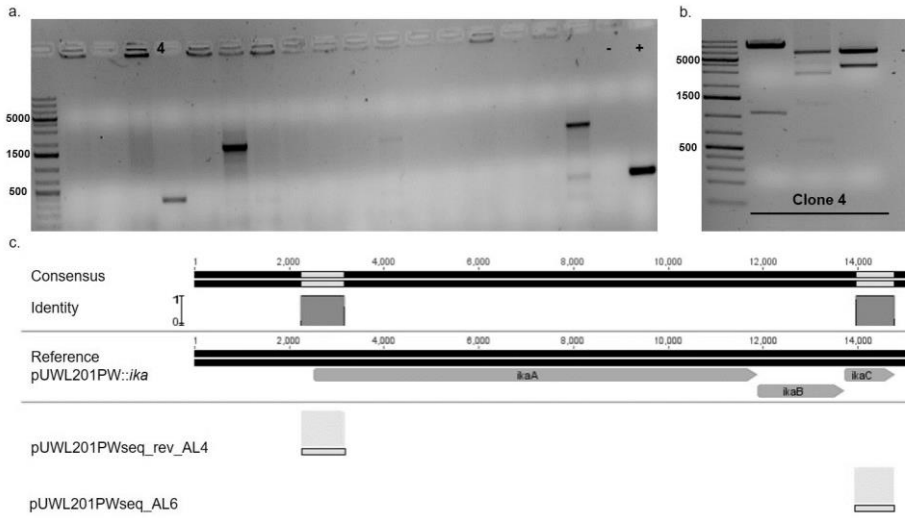


Figure A15. Cloning verification for pUWL201PW::*ika*. a. Results of colony PCR with possible positive clones 4 and 17, negative control conducted with water, positive control 1 µl Gibson assembly reaction mixture. b. Results of analytical restriction digestion for only clone 4, as 17 did not grow, as clone 4 showed the predicted restriction patter this clone was submitted for sequencing. c. Sequencing results of Clone 4 with Primers: pUWL201Pwseq_rev_AL4 and pUWL2019Pwseq_AL6, verifying a successful cloning of the *ika* BGC into pUWL201PW.

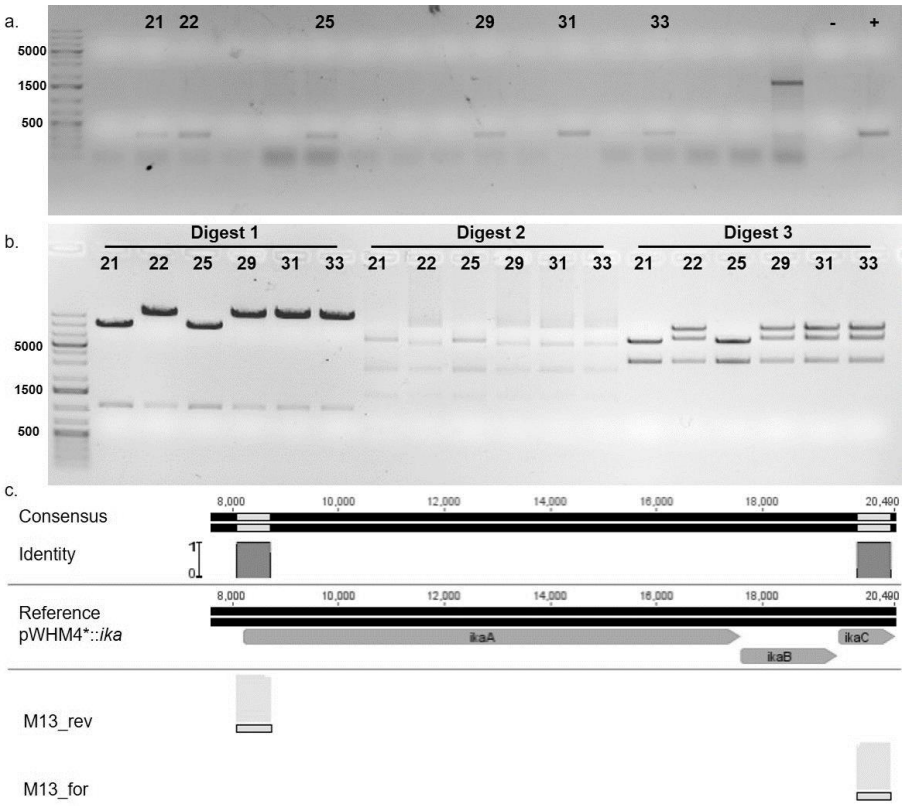


Figure A16. Cloning verification for pWHM4*::*ika*. a. Results of colony PCR with possible positive clones 21, 22, 25, 29, 31, and 33, negative control was conducted with water, positive control 1 µl of Gibson assembly reaction mixture. b. Results of analytical restriction digest of all clones, as clones 22 and 29 showed the predicted restriction pattern, both clones were submitted for sequencing. c. Sequencing result for clone 22 with primers: M13_rev and M13_for, verifying a successful cloning of the *ika* cluster into pWHM4*.

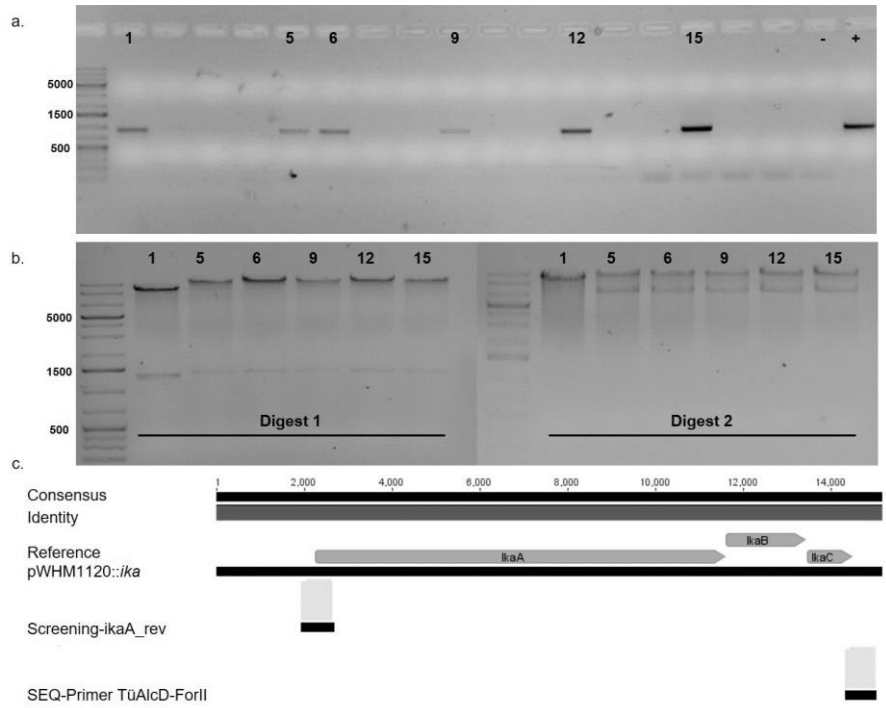


Figure A17. Cloning verification for pWHM1120::*ika*. a. Results of colony PCR with possible positive clone 1, 5, 6, 9, 12, and 15, negative control conducted with water, positive control 1 µl Gibson assembly reaction mixture. b. Results of analytical restriction digest of all clones, as clone 15 showed the predicted restriction pattern this clone was submitted for sequencing. c. Sequencing results for clone 15 with primers: Screening-ikaA_rev and SEQ-Primer TüAlcD-ForII, verifying a successful cloning of the *ika* cluster into pWHM1120.

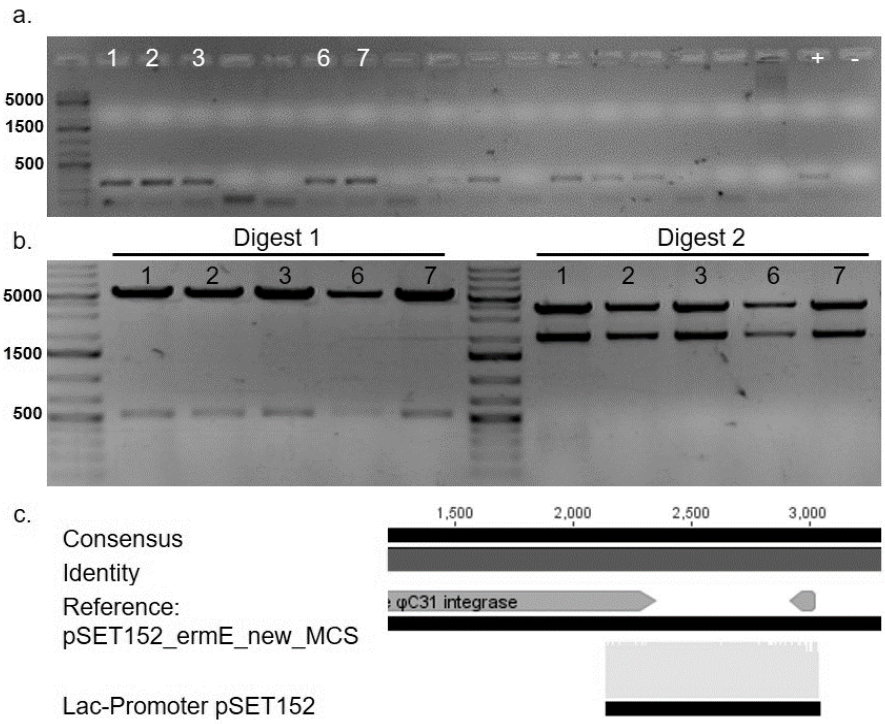


Figure A18. Cloning verification for pSET152_ermE_new_MCS. a. Results of colony PCR with possible positive clone 1, 2, 3, 6, and 7 negative control conducted with water, positive control 1 µl ligation reaction mixture. b. Results of analytical restriction digest of all clones, as clone 1 showed the predicted restriction pattern this clone was submitted for sequencing. c. Sequencing results for clone 1 with primers: Lac-promoter pSET152, verifying a successful cloning of the new MCS into pSET152_ermE.

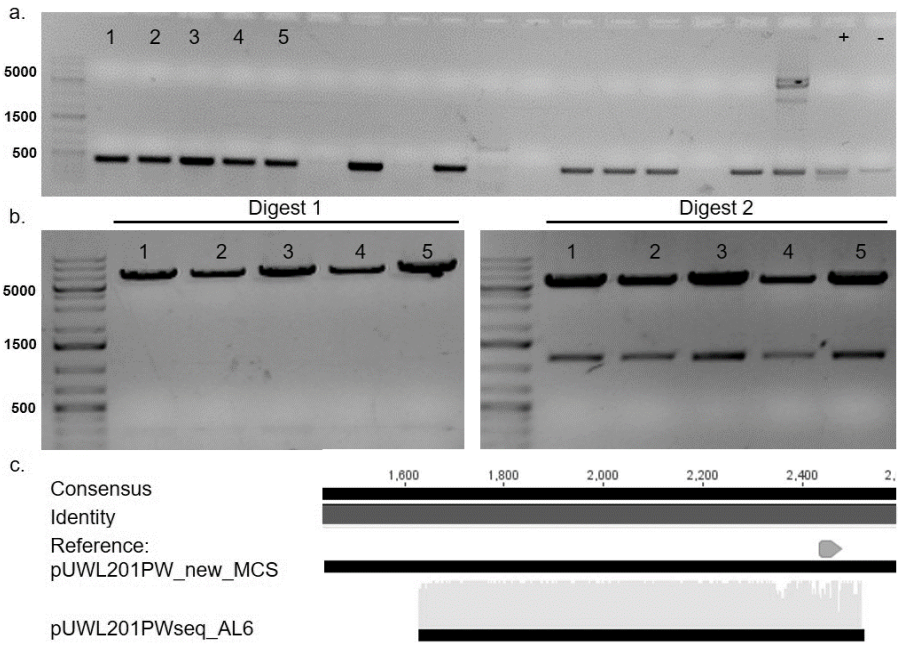


Figure A19. Cloning verification for pUWL201PW_new_MCS. a. Results of colony PCR with possible positive clone 1, 2, 3, 4, and 5 negative control conducted with water, positive control 1 µl ligation reaction mixture. b. Results of analytical restriction digest of all clones, as clone 3 showed the predicted restriction pattern this clone was submitted for sequencing. c. Sequencing results for clone 3 with primers: pUWL201Pwseq_AL6, verifying a successful cloning of the new MCS into pUWL201PW.

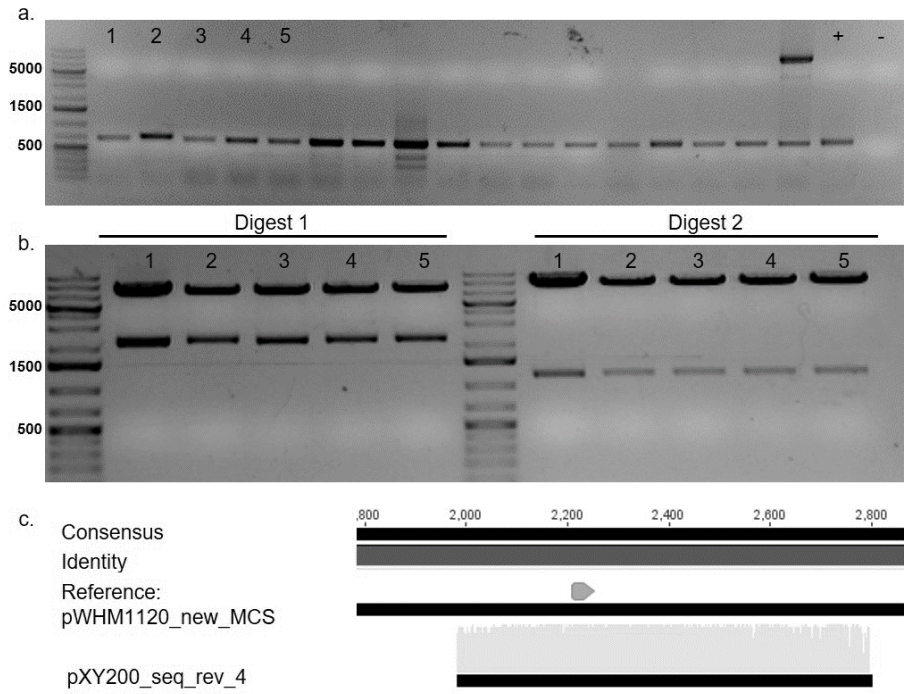


Figure A20. Cloning verification for pWHM1120_new_MCS. a. Results of colony PCR with possible positive clone 1, 2, 3, 4, and 5 negative control conducted with water, positive control 1 µl ligation reaction mixture. b. Results of analytical restriction digest of all clones, as clone 2 showed the predicted restriction pattern this clone was submitted for sequencing. c. Sequencing results for clone with primers: pXY200_seq_rev_4, verifying a successful cloning of the new MCS into pWHM1120.

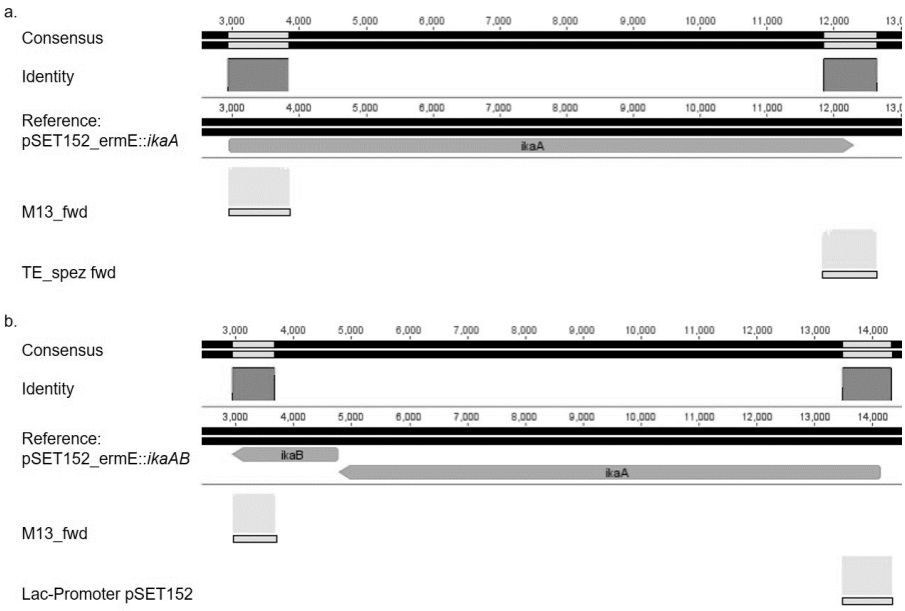


Figure A21. Sequencing results pSET152_ermE_new_MCS::*ikaA and* pSET152_ermE_new_MCS*::ikaAB*. a. Results of sequencing clone 2 containing pSET152_ermE_new_MCS::*ikaA* with primers M13_fwd and TE_spez fwd. b. Results of sequencing clone 4 containing pSET152_ermE_new_MCS::*ikaAB* with Primers M13_fwd and Lac-Promoter pSET152.

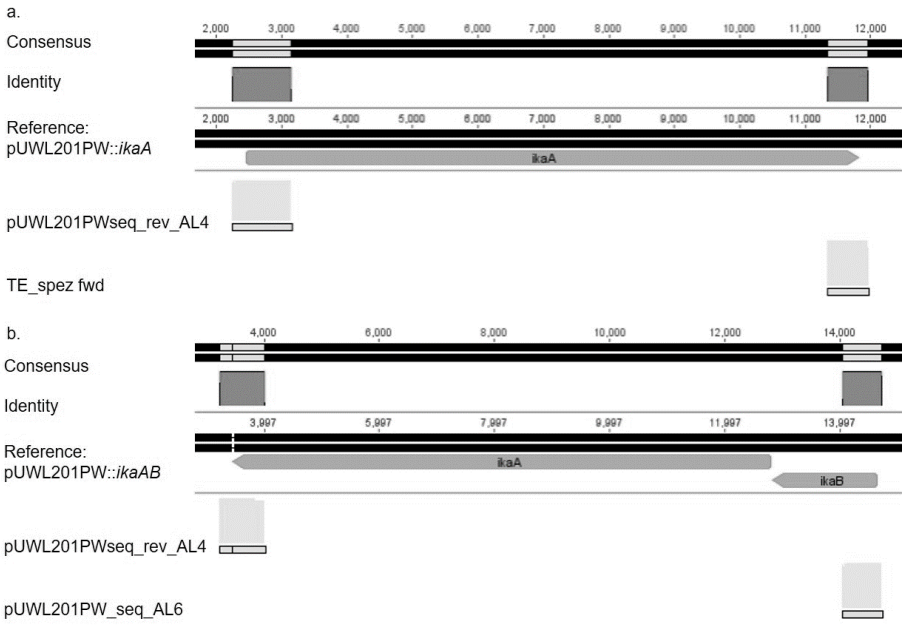


Figure A22. Sequencing results pUWL201PW_new_MCS::*ikaA* and ::*ikaAB*. a. Results of sequencing clone 1 in pUWL201PW_new_MCS::*ikaA* with Primers pUWL201Pwseq_rev_AL4 and TE_spez fwd. b. Results of sequencing clone 20 in pUWL201PW_new_MC::*ikaAB* with Primers pUWL201Pwseq_rev_AL4 and pUWL201PW_seq_AL6. As depicted by the light grey bars, no mutations were detected for pUWL201PW_new_MCS::*ikaA*, however there was an additional ATG in the sequence of pUWL201PW_new_MCS::*ikaAB*.

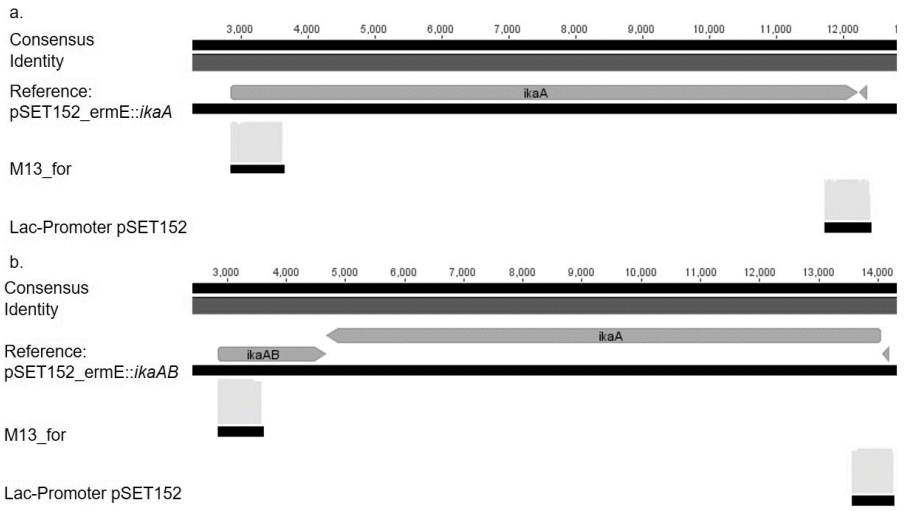


Figure A23. Sequencing results pSET152_ermE::*ikaA* and pSET152_ermE::*ikaAB*. a. Results of sequencing clone 8 containing pSET152_ermE::*ikaA* with Primers M13_fwd and Lac-Promoter pSET152, b. sequencing results of clone 3 containing pSET152_ermE::*ikaAB* with Primers M13_fwd and Lac-Promoter pSET152.

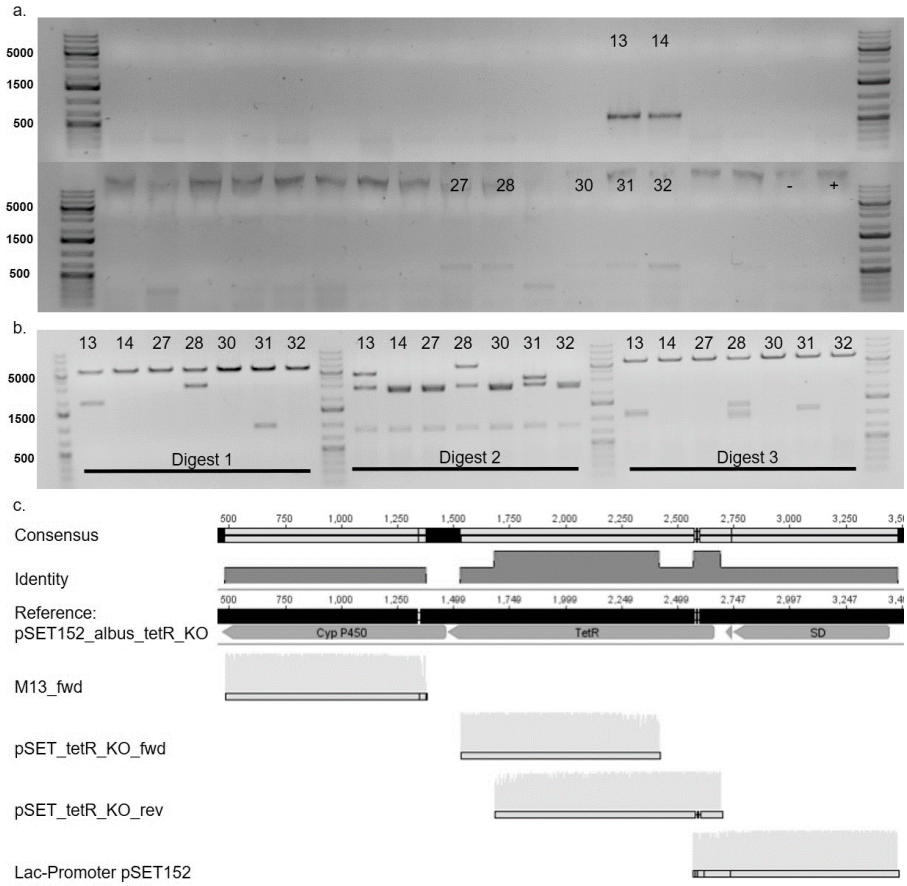


Figure A24. Cloning verification of pSET152_albus_tetR_KO. a. Results of colony PCR with possible positive clone 13, 14, 27, 28, 30, 31, and 32, negative control conducted with water, positive control 1 µl Gibson assembly reaction mixture. b. Results of analytical restriction digest of all clones, as clone 28 showed the predicted restriction pattern this clone was submitted for sequencing. c. Sequencing results for clone 28 with primers: M13_fwd, pSET_tetR_KO_fwd, pSET_tetR_KO_rev, and Lac-Promoter pSET152, verifying integration of the three fragments into pSET152, with several mutations inside the tetR Promoter and a few bases not covered at the beginning of the Cytochrome P450.

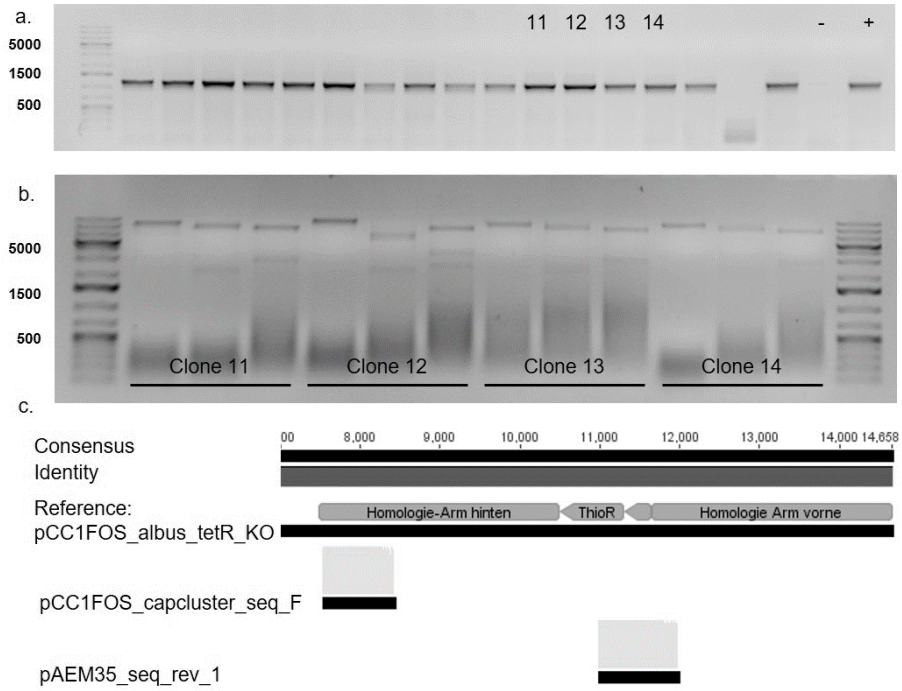


Figure A25. Cloning verification of pCC1FOS_albus_tetR_KO. a. Results of colony PCR with possible positive clone 11, 12, 13, and 14, negative control conducted with water, positive control 1 µl Gibson assembly reaction mixture. b. Results of analytical restriction digest of all clones, as clone 12 showed the predicted restriction pattern this clone was submitted for sequencing. c. Sequencing results for clone 13 with primers: pCC1FOS_capcluster_seq_F and pAEM35_seq_rev_1, verifying assembly of the four fragments.

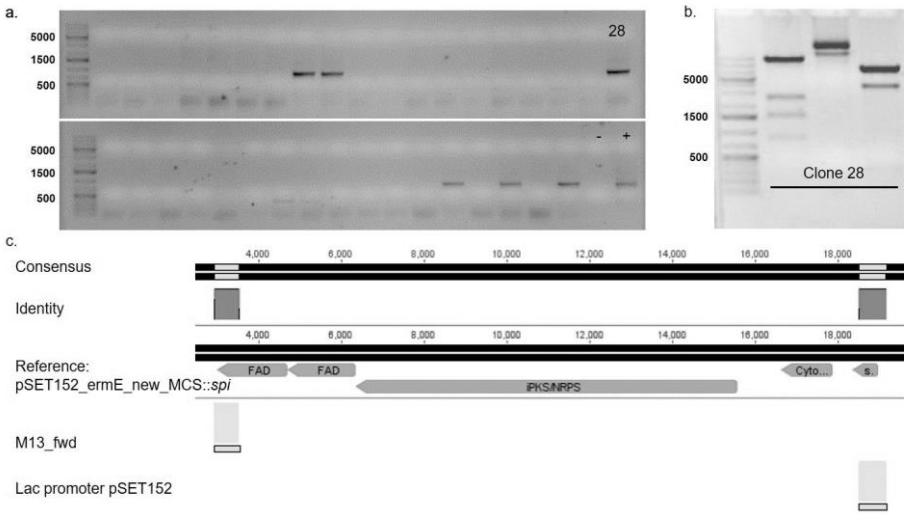


Figure A26. Cloning verification of pSET152_ermE_new_MCS::*spi*. a. Results of colony PCR with possible positive clone 28, negative control conducted with water, positive control 1 µl Gibson assembly reaction mixture. b. Results of analytical restriction digest of clone 28, as clone 28 showed the predicted restriction pattern this clone was submitted for sequencing. c. Sequencing results for clone 28 with primers: M13_fwd, and Lac-Promoter pSET152, verifying integration of the *spi* cluster into pSET152_ermE_new_MCS.

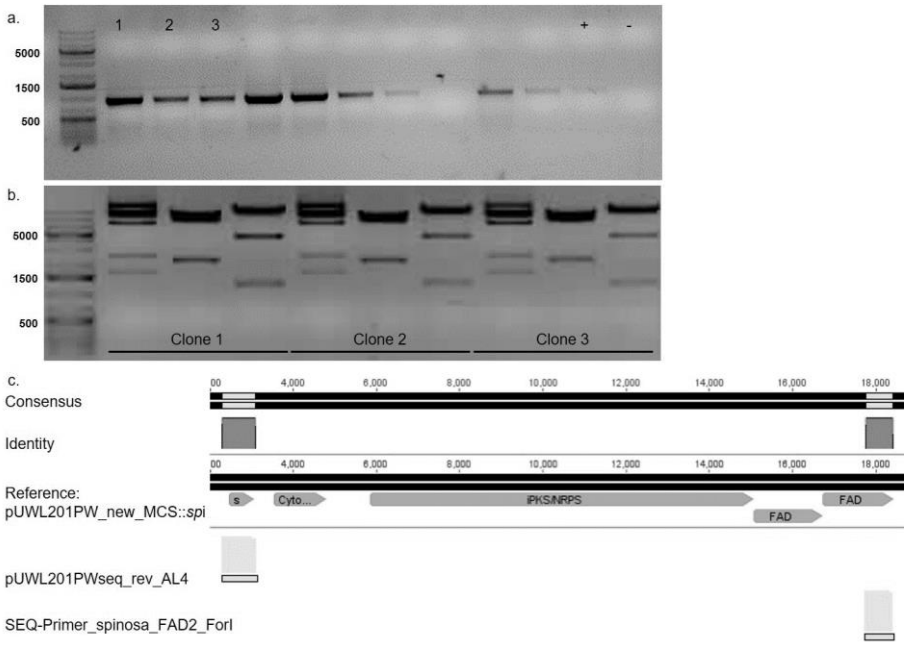


Figure A27. Cloning verification of pUWL201PW_new_MCS::*spi*. a. Results of colony PCR with possible positive clones 1, 2, and 3, negative control conducted with water, positive control 1 µl Gibson assembly reaction mixture. b. Results of analytical restriction digest of all clones, as clone 1 showed the predicted restriction pattern this clone was submitted for sequencing. c. Sequencing results for clone 28 with primers: pUWL201Pwseq_rev_AL4 and SEQ-Primer_spinosa_FAD2_ForI, verifying integration of the *spi* cluster into pUWL201PW_new_MCS.

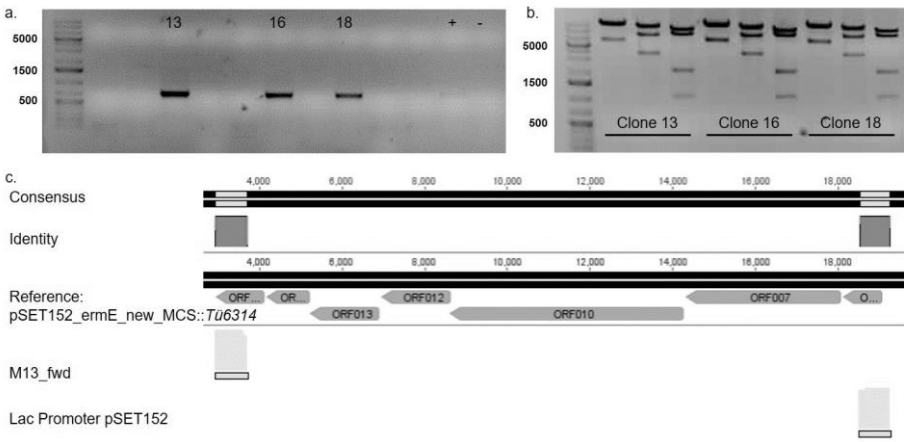


Figure A28. Cloning verification of pSET152_ermE_new_MCS::*Tü6314*. a. Results of colony PCR with possible positive clones 13, 16, and 18, negative control conducted with water, positive control 1 µl Gibson assembly reaction mixture. b. Results of analytical restriction digest of all clones, as clone 13 showed the predicted restriction pattern this clone was submitted for sequencing. c. Sequencing results for clone 13 with primers: Lac Promoter pSET152 and M13_fwd, verifying integration of the *Tü6314* cluster into pSET152_ermE_new_MCS.

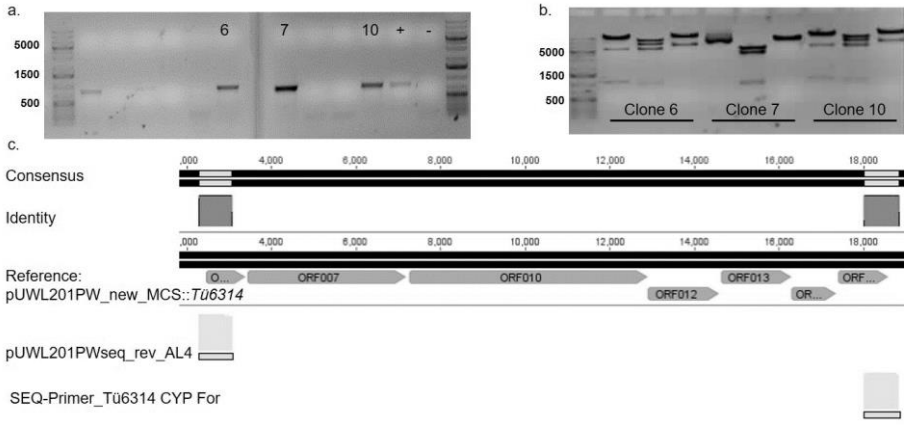


Figure A29. Cloning verification of pUWL201PW_new_MCS::*Tü6314*. a. Results of colony PCR with possible positive clones 6, 7, and 10, negative control conducted with water, positive control 1 µl Gibson assembly reaction mixture. b. Results of analytical restriction digest of all clones, as clone 6 showed the predicted restriction this clone was submitted for sequencing. c. Sequencing results for clone 6 with primers: pUWL201Pwseq_rev_AL4 and SEQ-Primer_Tü6314_CYP_for, verifying integration of the *Tü6314* cluster into pUWL201PW_new_MCS.

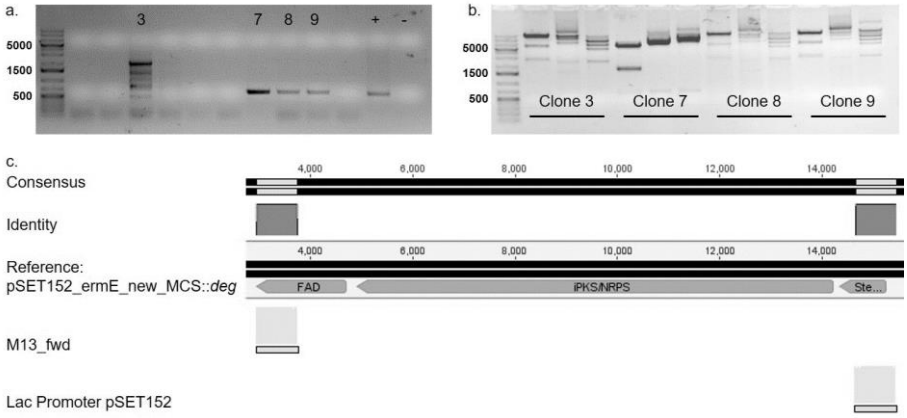


Figure A30. Cloning verification of pSET152_ermE_new_MCS::*deg*. a. Results of colony PCR with possible positive clones 3, 7, 8, and 9, negative control conducted with water, positive control 1 µl Gibson assembly reaction mixture. b. Results of analytical restriction digest of all clones, as clone 7 showed the predicted restriction pattern it was sent for sequencing. c. Sequencing results for clone 7 with primers: Lac Promoter pSET152 and M13_fwd, verifying integration of the *deg* cluster into pSET152_ermE_new_MCS.

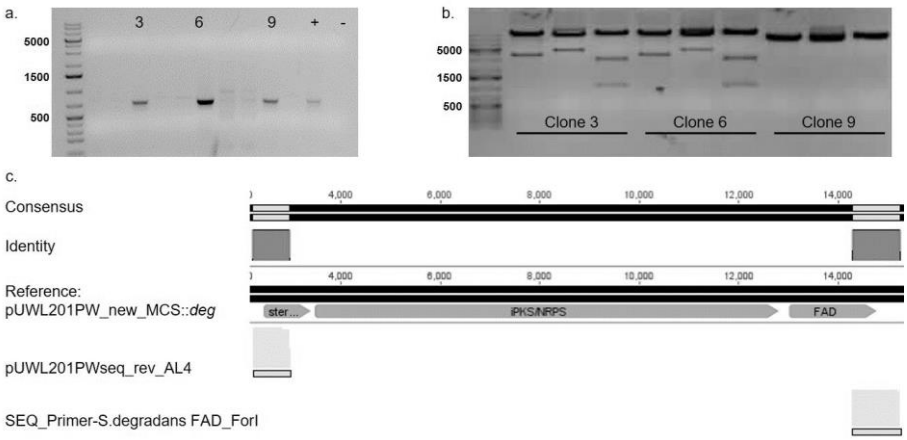


Figure A31. Cloning verification of pUWL201PW_new_MCS::*deg*. a. Results of colony PCR with possible positive clones 3, 6, and 9, negative control conducted with water, positive control 1 µl Gibson assembly reaction mixture. b. Results of analytical restriction digest of all clones, as clone 3 showed the predicted restriction pattern this clone was submitted for sequencing. c. Sequencing results for clone 3 with primers: pUWL201Pwseq_rev_AL4 and SEQ_Primer-S.degradansFAD_ForI, verifying integration of the *deg* cluster into pUWL201PW_new_MCS.

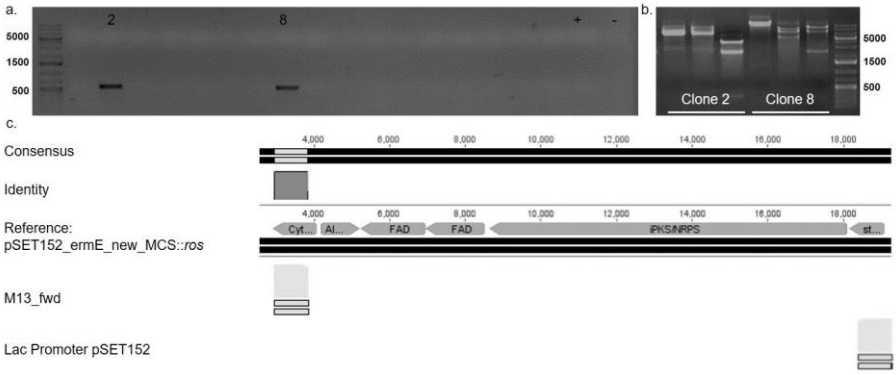


Figure A32. Cloning verification of pSET152_ermE_new_MCS::*ros*. a. Results of colony PCR with possible positive clones 2 and 8, negative control conducted with water, positive control 1 µl Gibson assembly reaction mixture. b. Results of analytical restriction digest of all clones, as clone 8 showed the predicted restriction pattern this clone was submitted for sequencing. c. Sequencing results for clone 8 with primers: Lac Promoter pSET152 and M13_fwd, verifying integration of the *ros* cluster into pSET152_ermE_new_MCS.

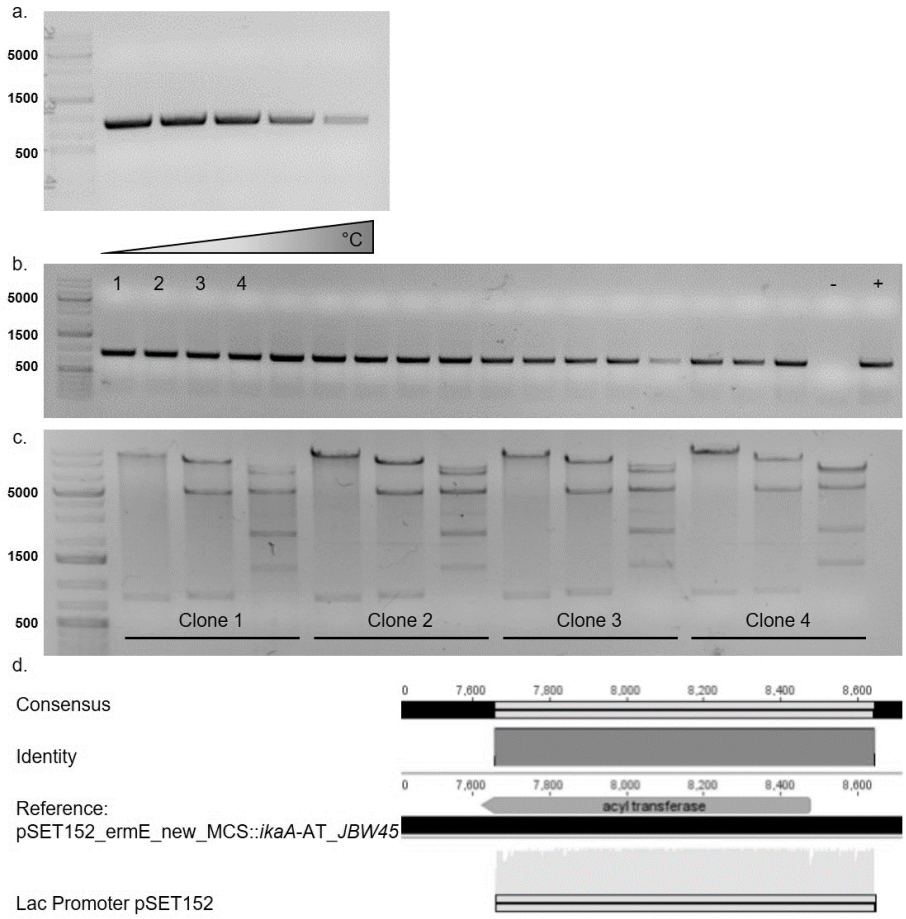


Figure A33. Cloning verification of pSET152_ermE_new_MCS::*ikaA*_AC_*JBW45*. a. Temperature gradient PCR from 50-70 °C to amplify the acyl transferase (895 bp) for Gibson assembly, b. Results of colony PCR with possible positive clones 1, 2, 3, and 4 negative control conducted with water, positive control 1 µl Gibson assembly reaction mixture. b. Results of analytical restriction digest of all clones, as clone 1 showed the predicted restriction this clone was submitted for sequencing. c. Sequencing results for clone 1 with primers: Lac Promoter pSET152, verifying integration of the acyl transferase gene into pSET152_ermE_new_MCS::*ikaA*.

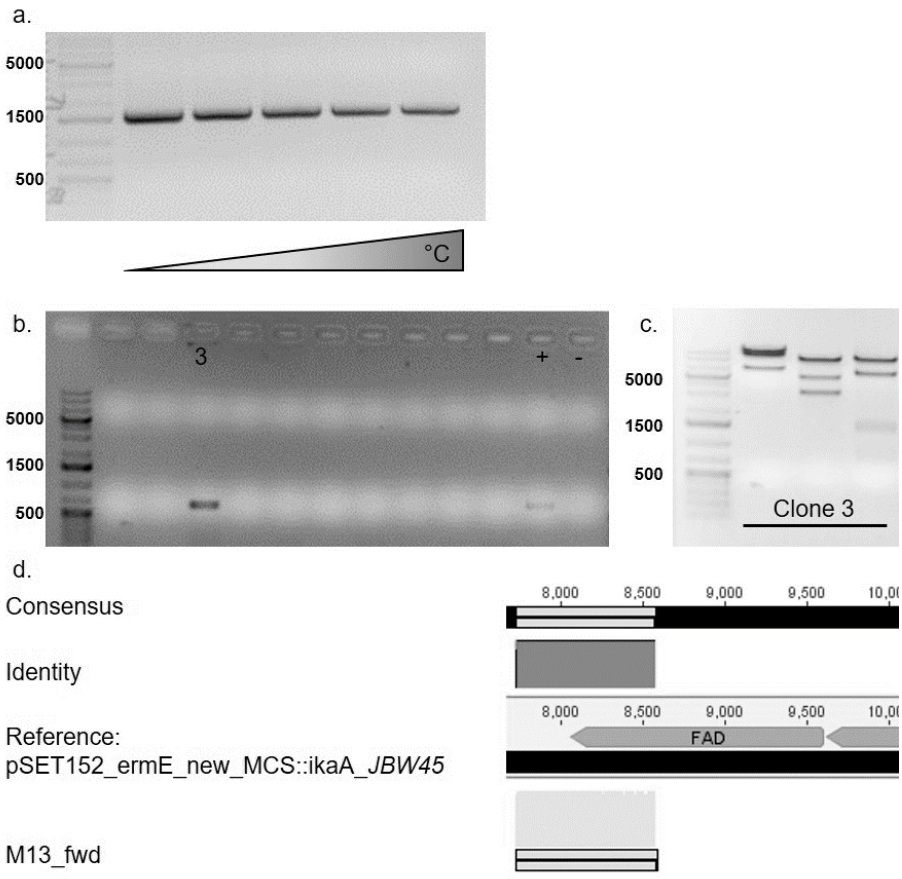


Figure A34. Cloning verification of pSET152_ermE_new_MCS::*ikaA_JBW45*. a. Temperature gradient PCR from 50-70 °C to amplify the FAD dependent oxidoreductase (1591 bp) for Gibson assembly, b. Results of colony PCR with possible positive clone 3, negative control conducted with water, positive control 1 µl Gibson assembly reaction mixture. c. Results of analytical restriction digest of clone 3, as it showed the predicted restriction pattern this clone was submitted for sequencing. d. Sequencing results for clone 3 with primers: M13_fwd verifying integration of the FAD dependent oxidoreductase to accomplish the cloning of pSET152_ermE_new_MCS::*ikaA_JBW45*.

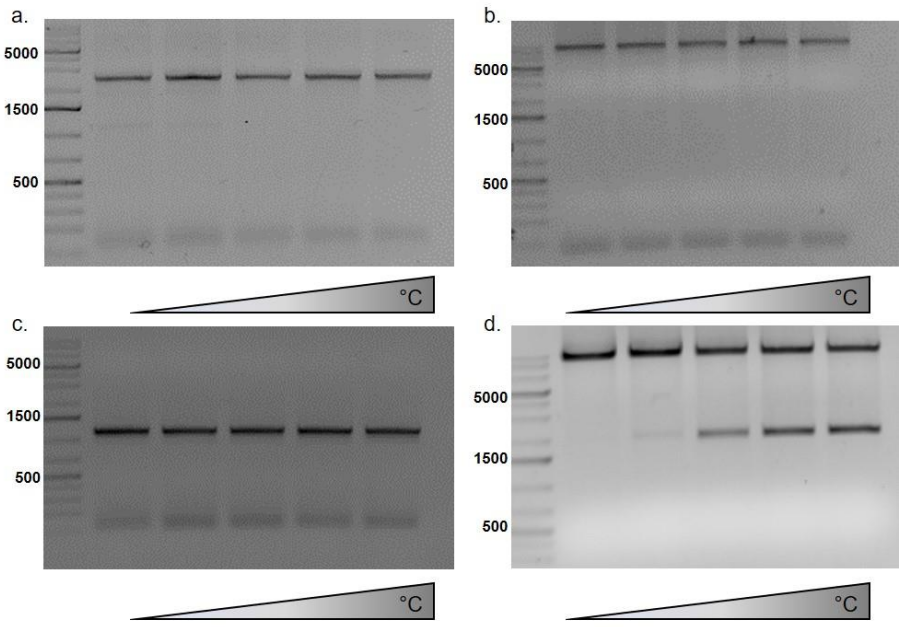


Figure A35. Temperature gradient to amplify cluster parts. Results of temperature gradient PCR 50-70 °C for a. the *spi*_front fragment (2552 bp), b. the *spi*_back fragment (12,683 bp), c. the *deg*_front fragment (1146 bp), and d. the *deg*_back fragment (11,339 bp).

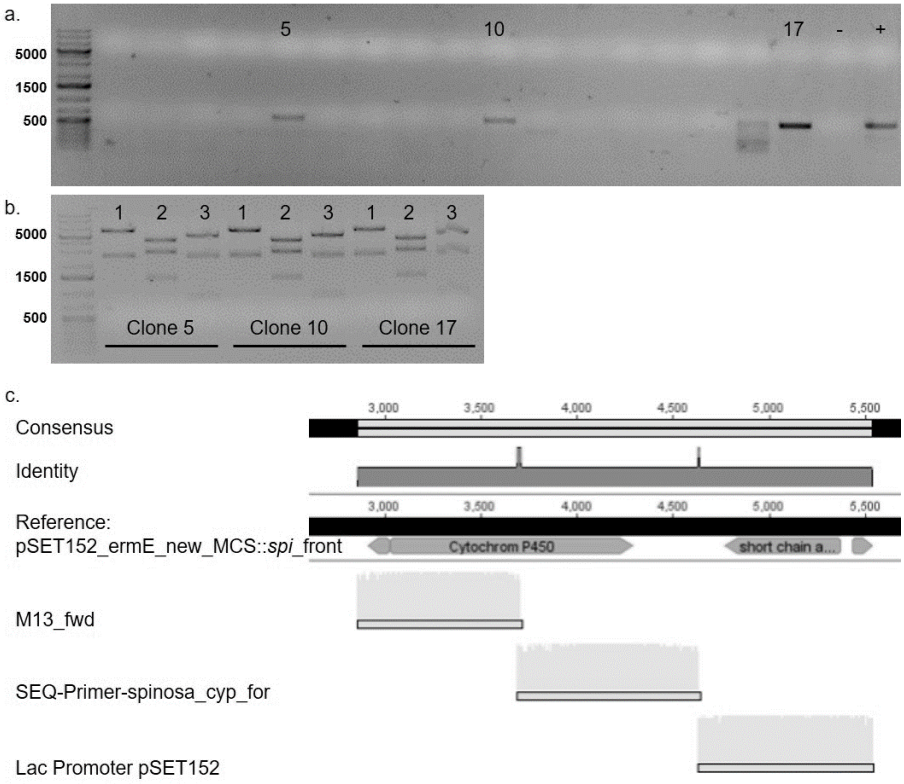


Figure A36. Cloning verification of pSET152_ermE_new_MCS::*spi*_front. a. Results of colony PCR with possible positive clones 5, 10, and 17, negative control conducted with water, positive control 1 µl Gibson assembly reaction mixture. b. Results of analytical restriction digest of all clones, as clone 5 showed the predicted restriction pattern this clone was submitted for sequencing. c. Sequencing results for clone 5 with primers: M13_fwd, SEQ-Primer-spinosa_cyp_for and Lac promoter pSET152 verifying integration of the *spi*_front fragment into pSET152_ermE_new_MCS.

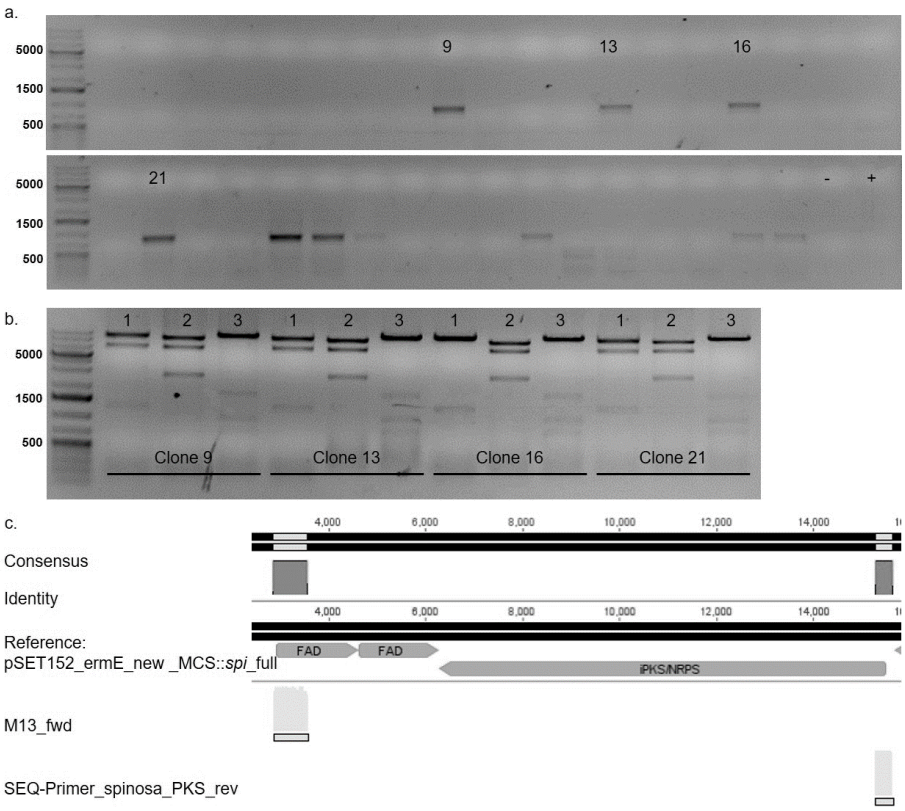


Figure A37. Cloning verification of pSET152_ermE_new_MCS::*spi*_full. a. Results of colony PCR with possible positive clones 9, 13, 16, and 21, negative control conducted with water, positive control 1 µl Gibson assembly reaction mixture. b. Results of analytical restriction digest of all clones, as clone 9 showed the predicted restriction pattern this clone was submitted for sequencing. c. Sequencing results for clone 9 with primers: M13_fwd and SEQ-Primer_spinosa_PKS_rev verifying integration of the *spi*_back fragment into pSET152_ermE_new_MCS::*spi*_front.

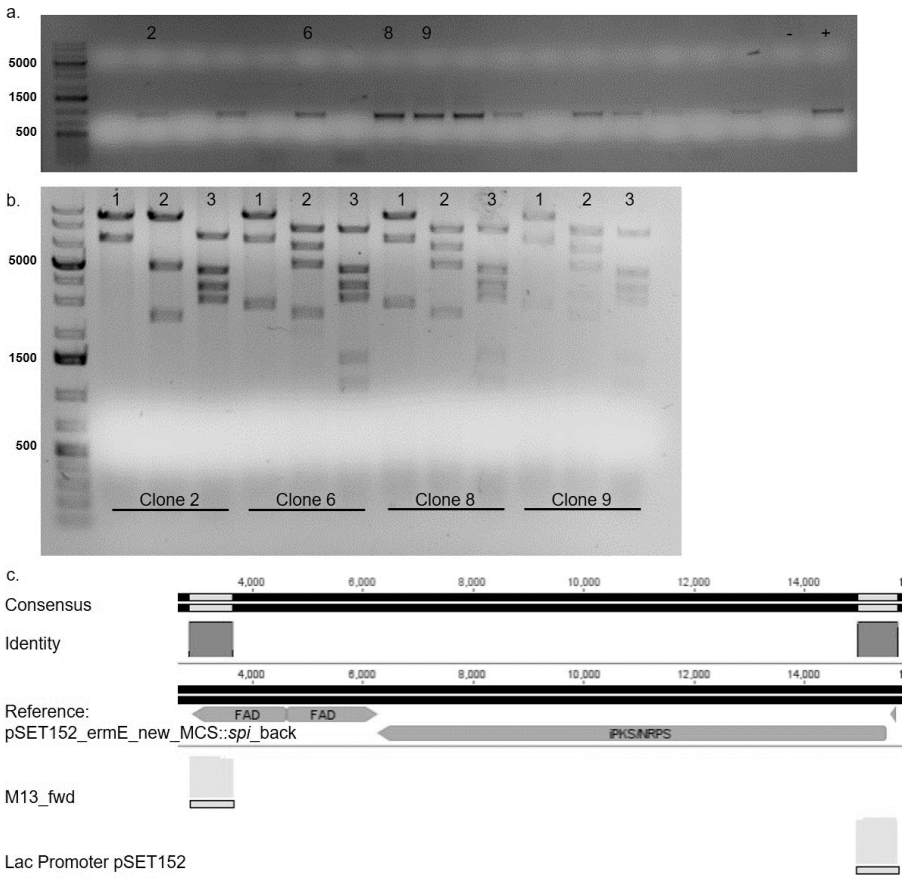


Figure A38. Cloning verification of pSET152_ermE_new_MCS::*spi*_back. a. Results of colony PCR with possible positive clones 2, 6, 8, and 9, negative control conducted with water, positive control 1 µl Gibson assembly reaction mixture. b. Results of analytical restriction digest of all clones, as clone 2 showed the predicted restriction this clone was submitted for sequencing. c. Sequencing results for clone 2 with primers: M13_fwd and Lac Promoter pSET152 verifying integration of the *spi*_back fragment into pSET152_ermE_new_MCS.

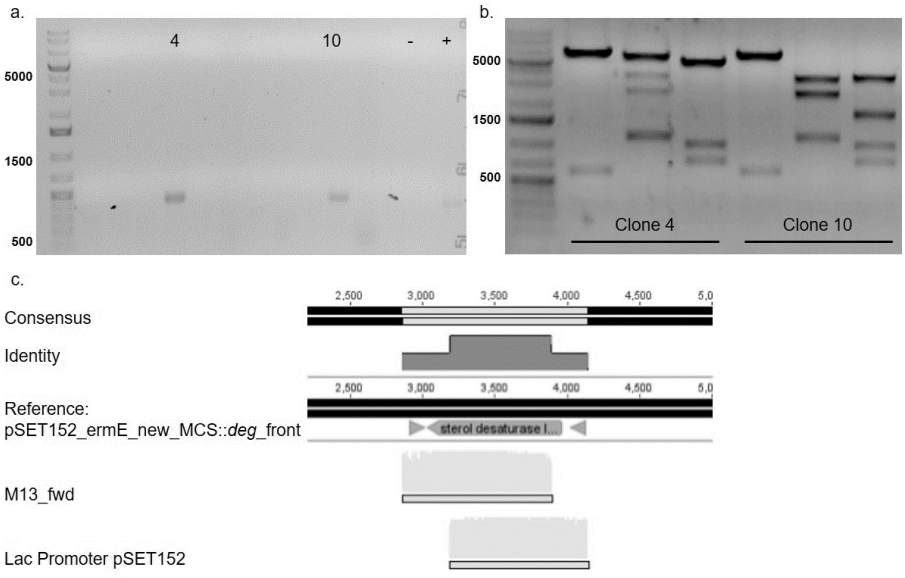


Figure A39. Cloning verification of pSET152_ermE_new_MCS::*deg*_front a. Results of colony PCR with possible positive clones 4 and 10, negative control conducted with water, positive control 1 µl Gibson assembly reaction mixture. b. Results of analytical restriction digest of all clones, as clone 10 showed the predicted restriction pattern this clone was submitted for sequencing. c. Sequencing results for clone 10 with primers: M13_fwd and Lac Promoter pSET152 verifying integration of the *deg*_front fragment into pSET152_ermE_new_MCS.

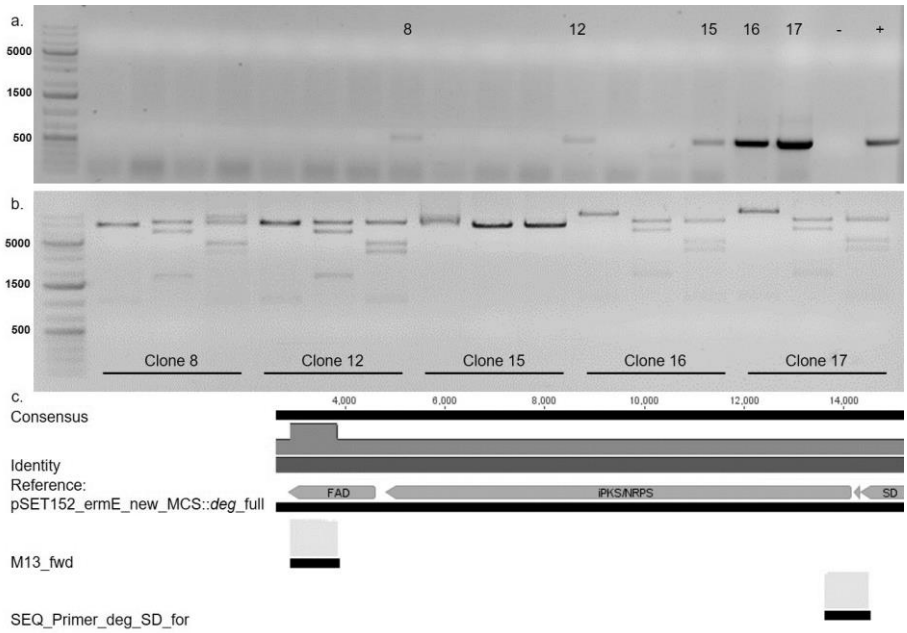


Figure A40. Cloning verification of pSET152_ermE_new_MCS::*deg*_full a. Results of colony PCR with possible positive clones 8, 12, 15, 16 and 17, negative control conducted with water, positive control 1 µl Gibson assembly reaction mixture. b. Results of analytical restriction digest of all clones, as clone 17 showed the predicted restriction pattern this clone was submitted for sequencing. c. Sequencing results for clone 17 with primers: M13_fwd and SEQ_Primer_deg_SD_for verifying integration of the *deg*_back fragment into pSET152_ermE_new_MCS::*deg*_front.

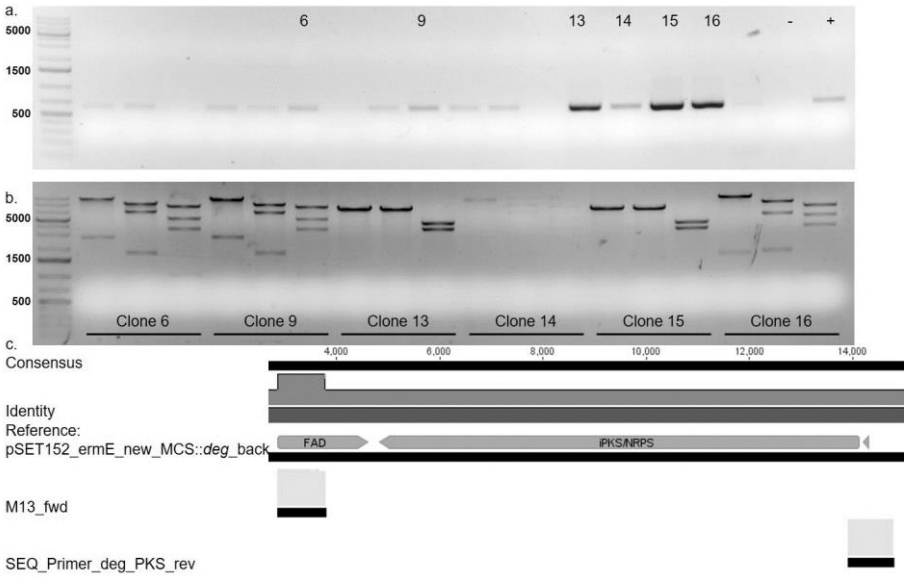


Figure A41. Cloning verification of pSET152_ermE_new_MCS::*deg*_back a. Results of colony PCR with possible positive clones 6, 9, 13, 14, 15 and 16, negative control conducted with water, positive control 1 µl Gibson assembly reaction mixture. b. Results of analytical restriction digest of all clones, as clone 16 showed the predicted restriction pattern this clone was submitted for sequencing. c. Sequencing results for clone 16 with primers: M13_fwd and SEQ_Primer_deg_PKS_rev verifying integration of the *deg*_back fragment into pSET152_ermE_new_MCS.

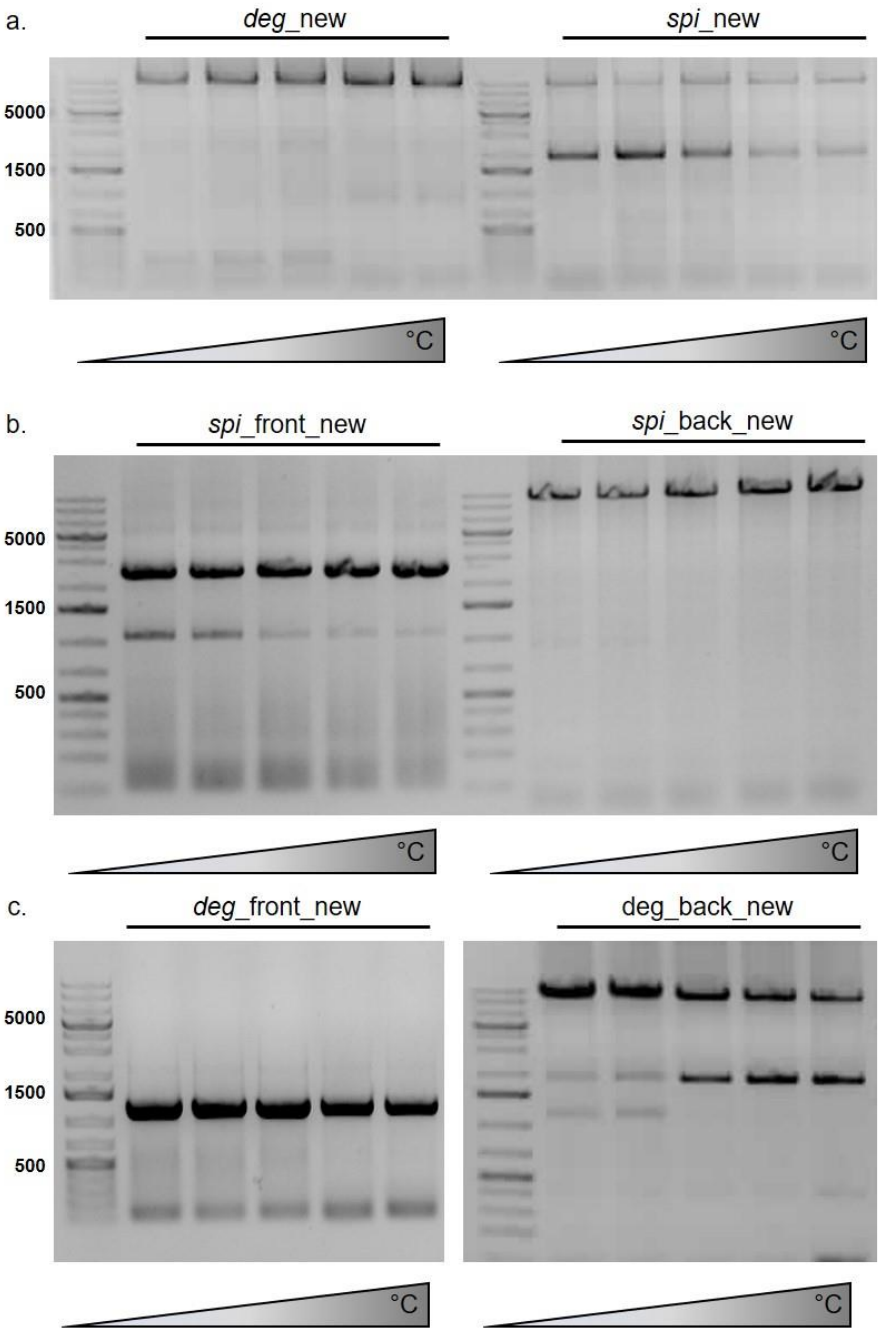


Figure A42. Temperature gradient PCR *spi*_new cluster and *deg*_new cluster. Results of temperature gradient PCR to amplify a. whole *deg* cluster and *spi* cluster, b. partial clusters for PoTeM BGC *spi* and c. partial clusters for PoTeM BGC *deg*.

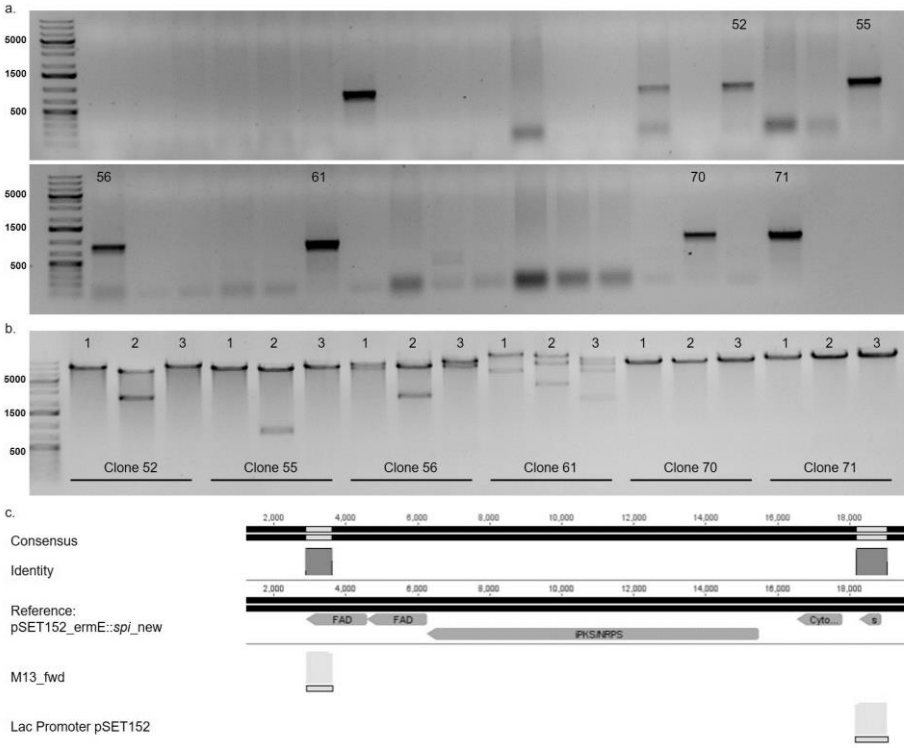


Figure A43. Cloning verification of pSET152_ermE::*spi*_new. a. Results of colony PCR with possible positive clones 52, 55, 56, 61, 70, and 71, negative control conducted with water, positive control 1 µl Gibson assembly reaction mixture. b. Results of analytical restriction digest of all clones, as clone 61 showed the predicted restriction pattern this clone was submitted for sequencing. c. Sequencing results for clone 61 with primers: M13_fwd and Lac Promoter pSET152 verifying integration of *spi*_new into pSET152_ermE.

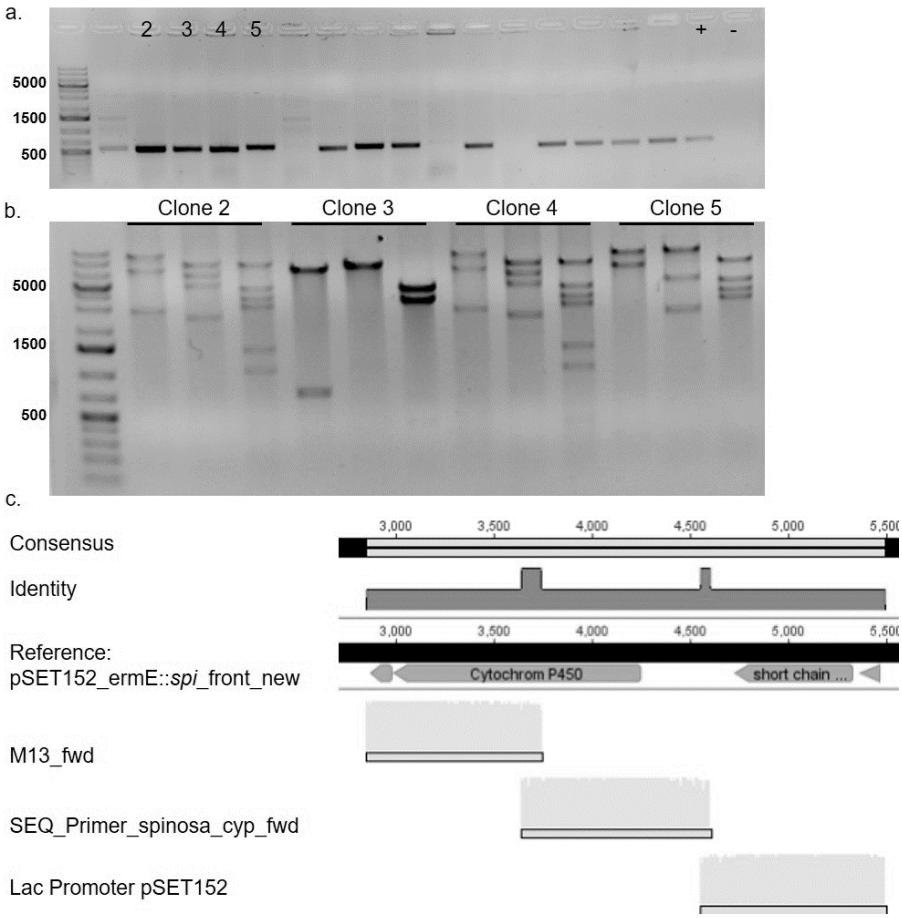


Figure A44. Cloning verification of pSET152_ermE::*spi*_front_new. a. Results of colony PCR with possible positive clones 2, 3, 4, and 5, negative control conducted with water, positive control 1 µl Gibson assembly reaction mixture. b. Results of analytical restriction digest of all clones, as clone 2 showed the predicted restriction pattern this clone was submitted for sequencing. c. Sequencing results for clone 2 with primers: M13_fwd, SEQ_Primer_spinosa_cyp_fwd and Lac Promoter pSET152 verifying integration of *spi*_front_new into pSET152_ermE.

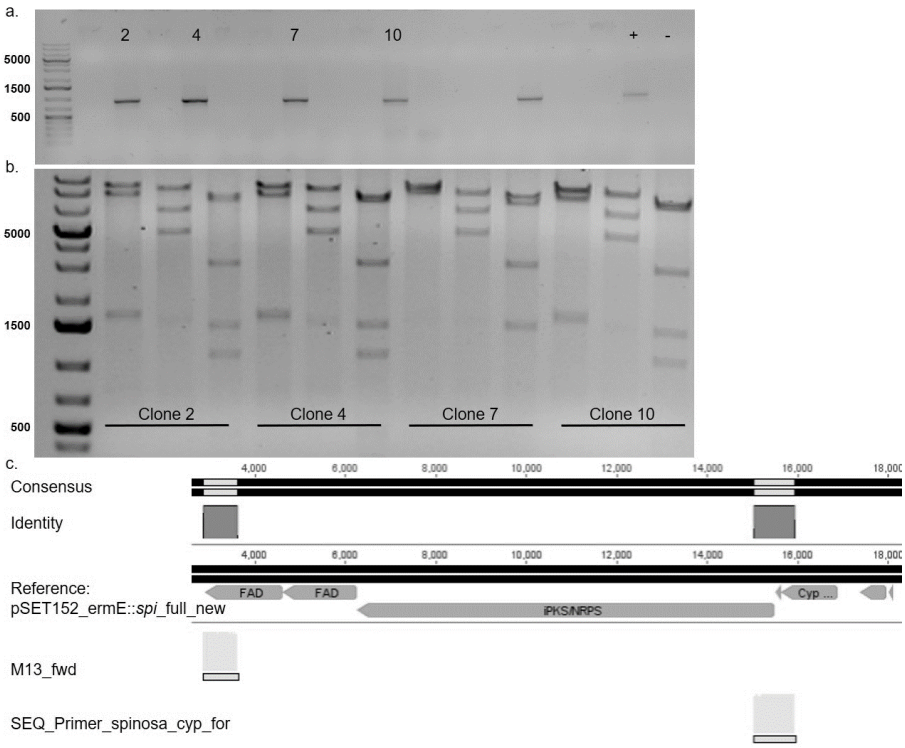


Figure A45. Cloning verification of pSET152_ermE::*spi*_full_new. a. Results of colony PCR with possible positive clones 2, 4, 7, and 10, negative control conducted with water, positive control 1 µl Gibson assembly reaction mixture. b. Results of analytical restriction digest of all clones, as clone 7 showed the predicted restriction pattern this clone was submitted for sequencing. c. Sequencing results for clone 7 with primers: M13_fwd, and SEQ_Primer_spinosa_cyp_for verifying integration of *spi*_back_new into pSET152_ermE::*spi*_front_new.

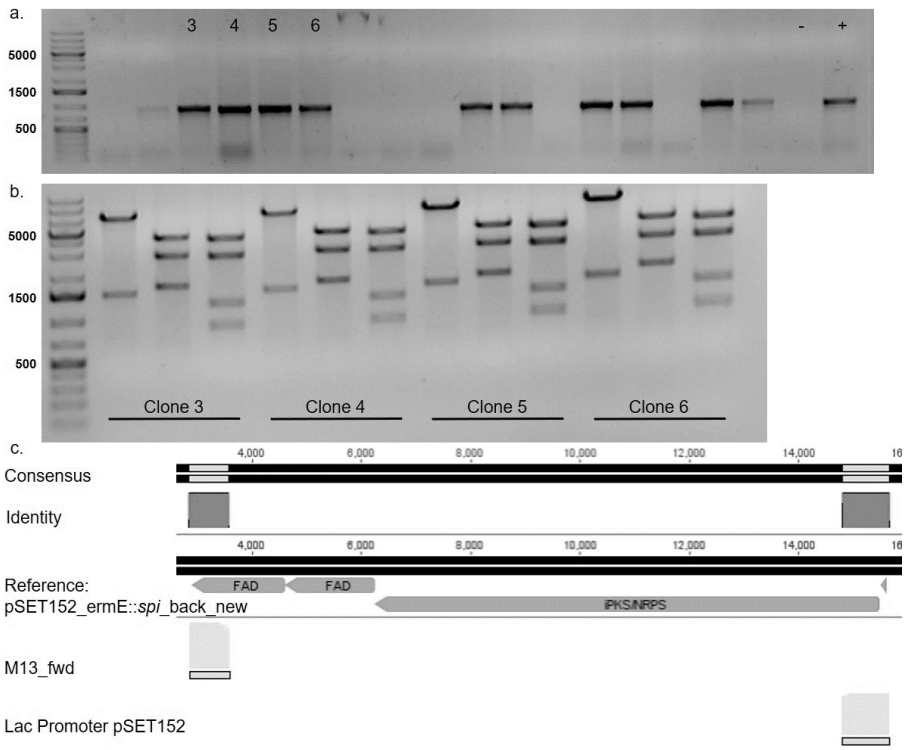


Figure A46. Cloning verification of pSET152_ermE::*spi*_back_new. a. Results of colony PCR with possible positive clones 3, 4, 5, and 6 negative control conducted with water, positive control 1 µl Gibson assembly reaction mixture. b. Results of analytical restriction digest of all clones, as clone 6 showed the predicted restriction pattern this clone was submitted for sequencing. c. Sequencing results for clone 6 with primers: M13_fwd and Lac Promoter pSET152 verifying integration of *spi*_back_new into pSET152_ermE.

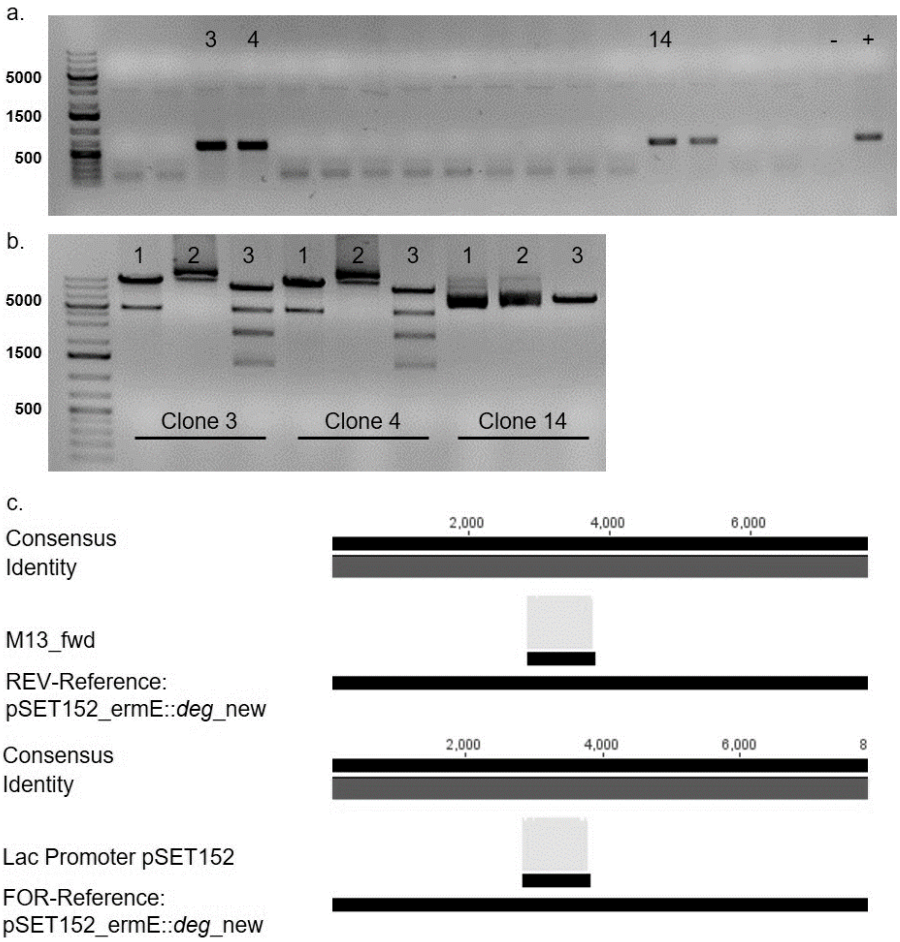


Figure A47. Cloning verification of pSET152_ermE::*deg*_new. a. Results of colony PCR with possible positive clones 3, 4, and 14, negative control conducted with water, positive control 1 µl Gibson assembly reaction mixture. b. Results of analytical restriction digest of all clones, as clone 3 showed the predicted restriction pattern this clone was submitted for sequencing. c. Sequencing results for clone 3 with primers: M13_fwd and Lac Promoter pSET152 verifying integration of *deg*_new into pSET152_ermE.

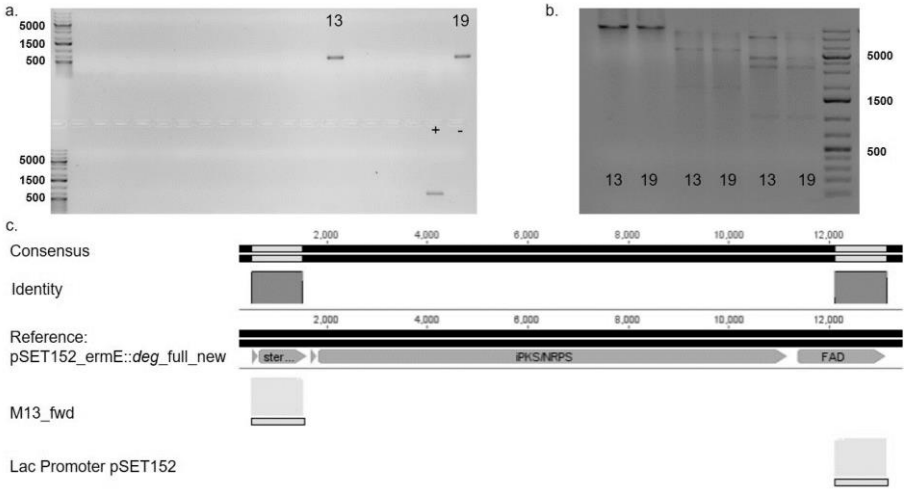


Figure A48. Cloning verification of pSET152_ermE::*deg*_full_new. a. Results of colony PCR with possible positive clones 13 and 19, negative control conducted with water, positive control 1 µl Gibson assembly reaction mixture. b. Results of analytical restriction digest of all clones, as clone 13 showed the predicted restriction pattern this clone was submitted for sequencing. c. Sequencing results for clone 13 with primers: M13_fwd and Lac Promoter pSET152 verifying integration of *deg*_back_new into pSET152_ermE::*deg*_front_new.

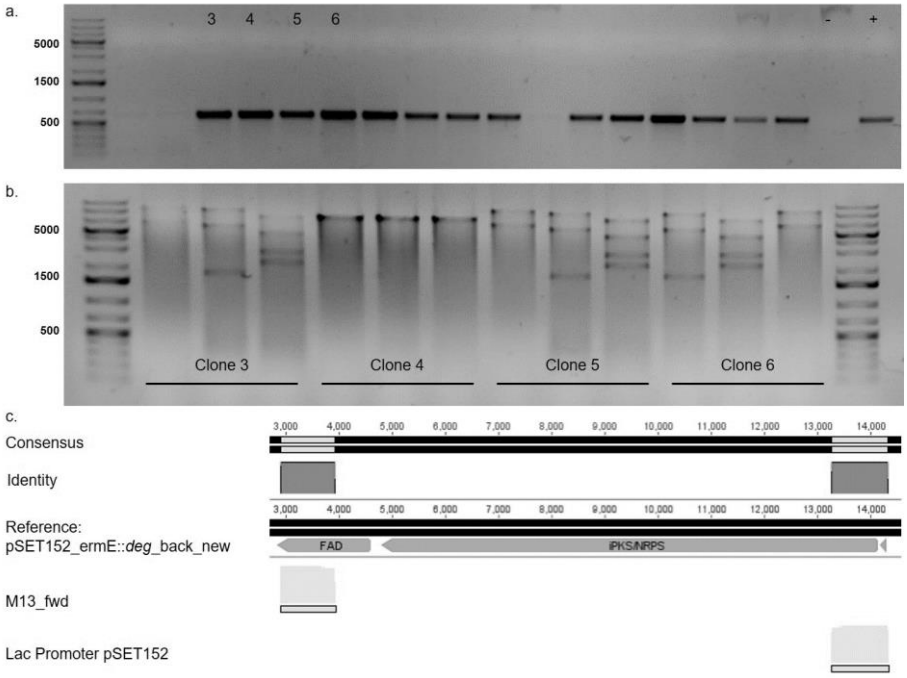


Figure A49. Cloning verification of pSET152_ermE::*deg*_back_new. a. Results of colony PCR with possible positive clones 3, 4, 5, and 6, negative control conducted with water, positive control 1 µl Gibson assembly reaction mixture. b. Results of analytical restriction digest of all clones, as clone 3 showed the predicted restriction pattern this clone was submitted for sequencing. c. Sequencing results for clone 3 with primers: M13_fwd and Lac Promoter pSET152 verifying integration of *deg*_back_new into pSET152_ermE.

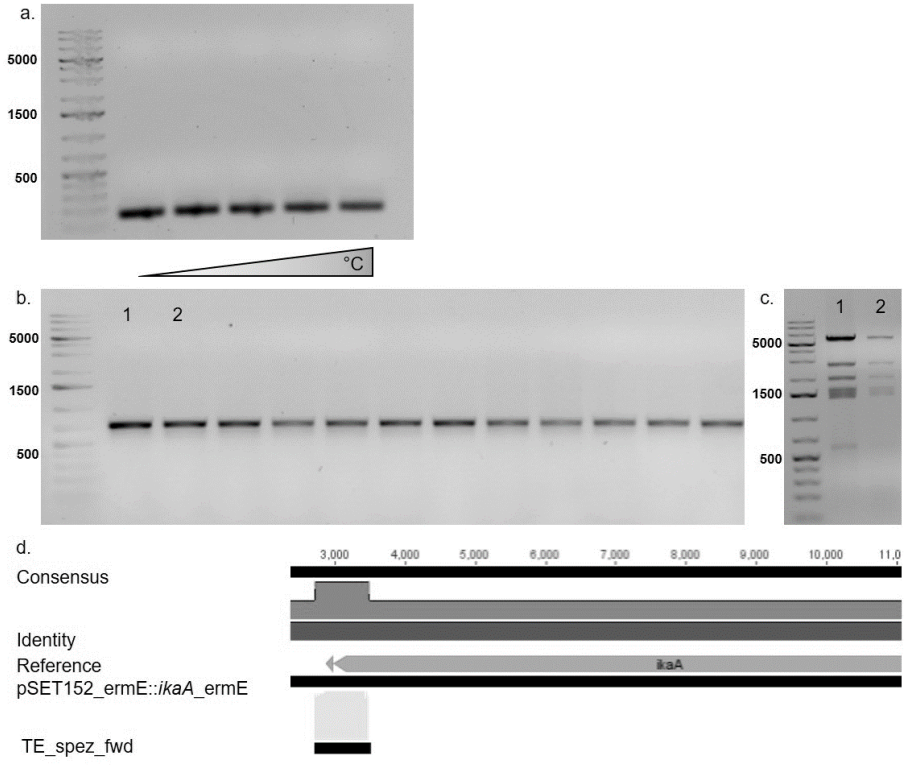


Figure A50. Cloning verification of pSET152_ermE::*ikaA*_ermE. a. Agarose Gel of temperature gradient PCR for ermE amplification. b. Results of colony PCR with possible positive clones 1 and 2. c. Results of analytical restriction digest of all clones, as clone 1 showed the predicted restriction pattern this clone was submitted for sequencing. d. Sequencing results for clone 1 with primers: TE_spez_fwd verifying integration of *ermE* into pSET152_ermE::*ikaA*.

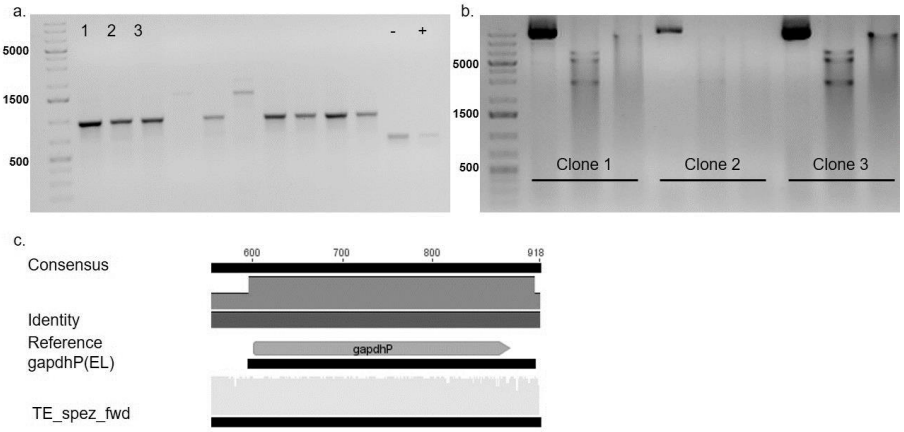


Figure A51. Cloning verification of pSET152_ermE::*ikaA*_gapdhP(EL). a. Results of colony PCR with possible positive clones 1, 2 and 3 negative control conducted with pSET152_ermE_*ikaA*, positive control 1 µl ligation reaction mixture b. Results of analytical restriction digest of all clones, as clone 1 showed the predicted restriction pattern this clone was submitted sequencing. c. Sequencing results for clone 1 with primers: TE_spez_fwd verifying integration of *gapdhP*(EL) into pSET152_ermE::*ikaA*.

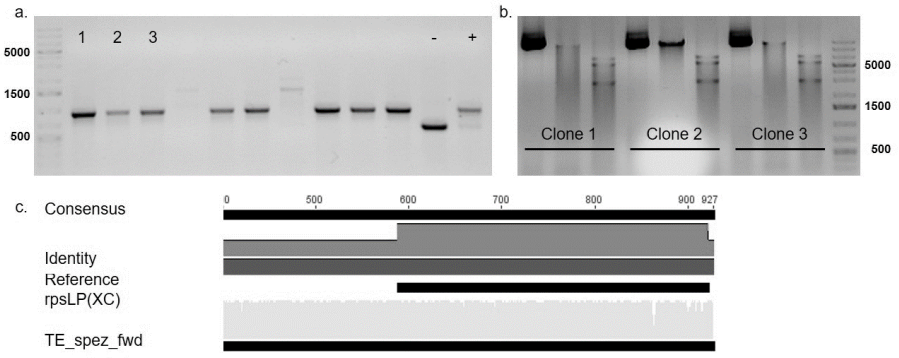


Figure A52. Cloning verification of pSET152_ermE::*ikaA*_rpsLP(XC). a. Results of colony PCR with possible positive clones 1, 2 and 3 negative control conducted with pSET152_ermE_*ikaA*, positive control 1 µl ligation reaction mixture b. Results of analytical restriction digest of all clones, as clone 1 showed the predicted restriction pattern this clone was submitted for sequencing. c. Sequencing results for clone 1 with primers: TE_spez_fwd verifying integration of *rpsLP*(XC) into pSET152_ermE::*ikaA*.

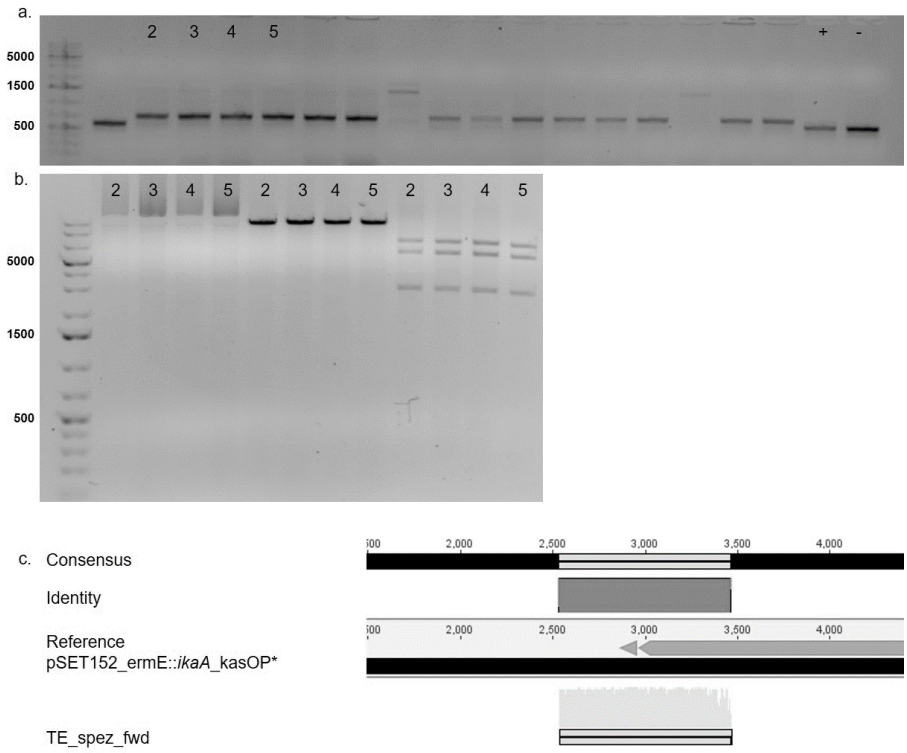


Figure A53. Cloning verification of pSET152_ermE::*ikaA*_kasOP*. a. Results of colony PCR with possible positive clones 2, 3, 4, and 5 negative control conducted with pSET152_ermE_*ikaA*, positive control 1 µl ligation reaction mixture b. Results of analytical restriction digest of all clones, as clone 2 showed the predicted restriction pattern this clone was submitted for sequencing. c. Sequencing results for clone 2 with primers: TE_spez_fwd verifying integration of *kasOP** into pSET152_ermE::*ikaA*.

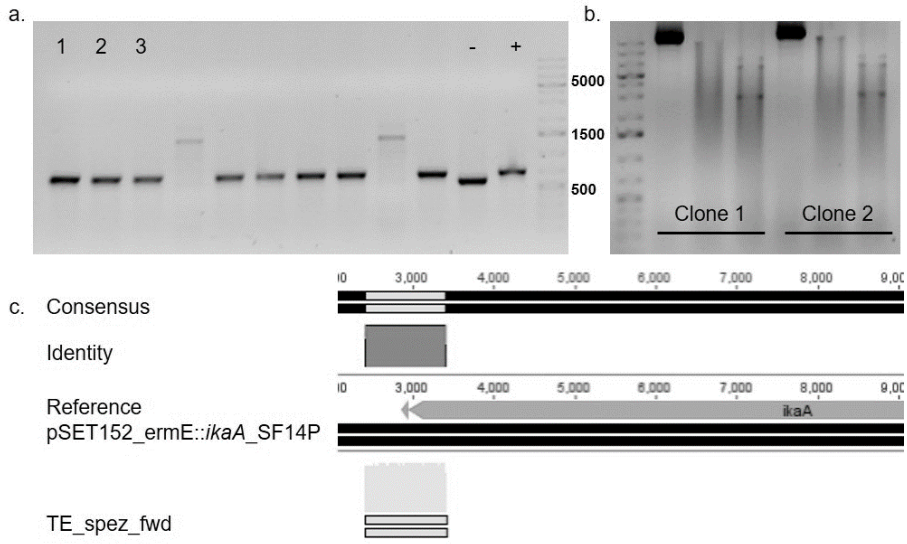


Figure A54. Cloning verification of pSET152_ermE::*ikaA*_SF14P. a. Results of colony PCR with possible positive clones 1, 2, and 3 negative control conducted with pSET152_ermE_*ikaA*, positive control 1 µl ligation reaction mixture b. Results of analytical restriction digest of all clones, as clone 1 showed the predicted restriction pattern this clone was submitted for sequencing. c. Sequencing results for clone 1 with primers: TE_spez_fwd verifying integration of *SF14P* into pSET152_ermE::*ikaA*.

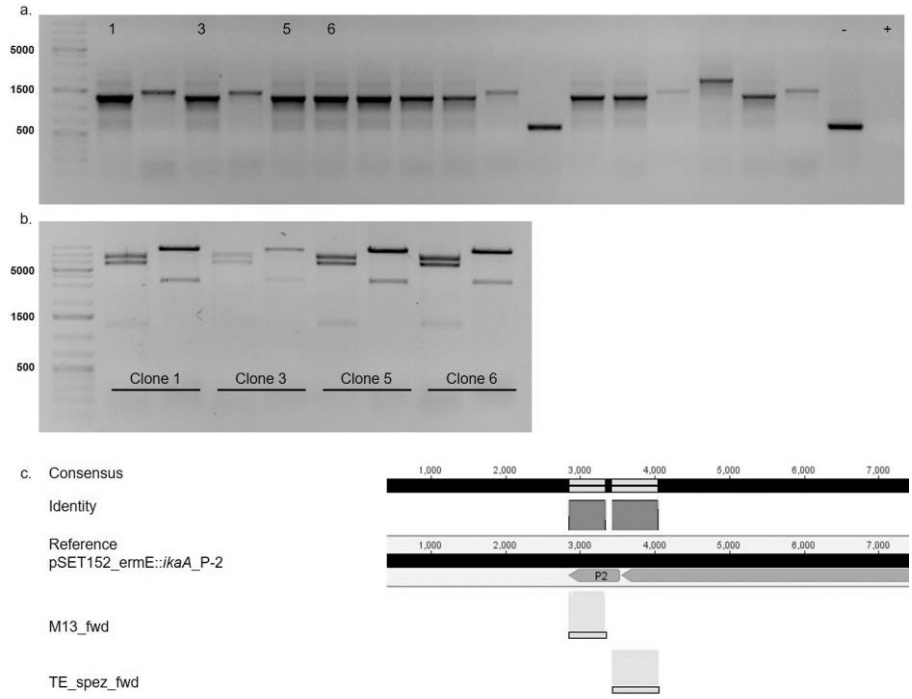


Figure A55. Cloning verification of pSET152_ermE::*ikaA*_P-2. a. Results of colony PCR with possible positive clones 1, 3, 5, and 6 negative control conducted with pSET152_ermE_*ikaA*, positive control 1 µl ligation reaction mixture b. Results of analytical restriction digest of all clones, as clone 1 showed the predicted restriction this clone was submitted for sequencing. c. Sequencing results for clone 1 with primers: TE_spez_fwd and M13_fwd verifying integration of *P-2* into pSET152_ermE::*ikaA*.

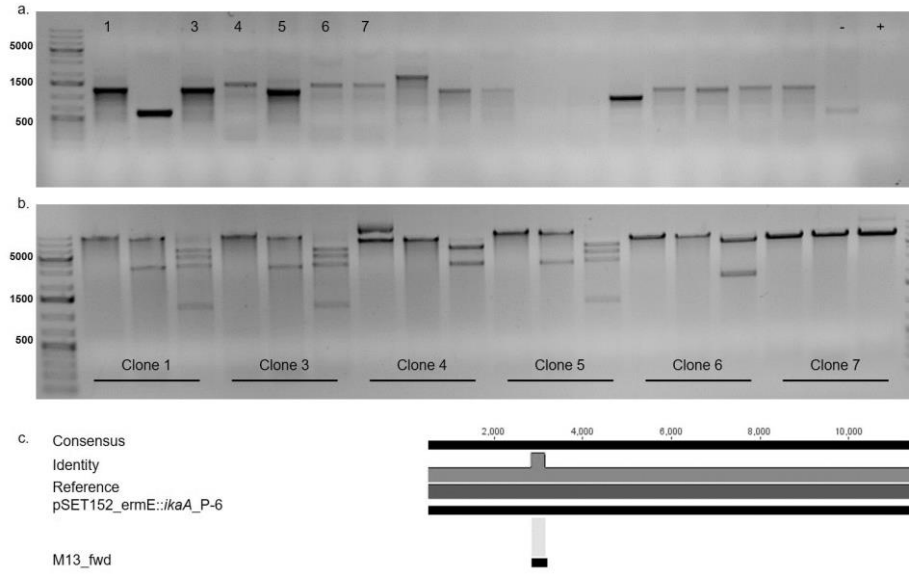


Figure A56. Cloning verification of pSET152_ermE::*ikaA*_P-6. a. Results of colony PCR with possible positive clones 1, 3, 4, 5, 6 and 7 negative control conducted with pSET152_ermE_*ikaA*, positive control 1 µl ligation reaction mixture b. Results of analytical restriction digest of all clones, as clone 1 showed the predicted restriction this clone was submitted for sequencing. c. Sequencing results for clone 1 with primers: M13_fwd verifying integration of *P-6* into pSET152_ermE::*ikaA*.

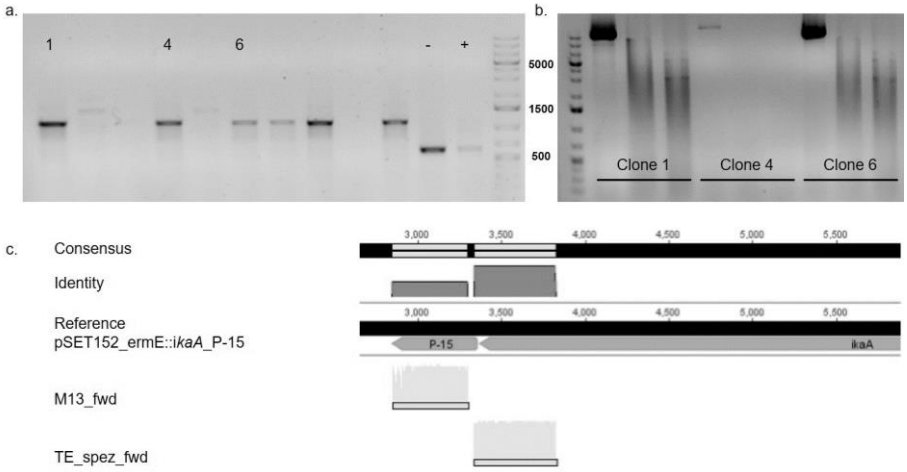


Figure A57. Cloning verification of pSET152_ermE::*ikaA*_P-15. a. Results of colony PCR with possible positive clones 1, 4, and 6 negative control conducted with pSET152_ermE_*ikaA*, positive control 1 µl ligation reaction mixture b. Results of analytical restriction digest of all clones, as clone 1 showed the predicted restriction this clone was submitted for sequencing. c. Sequencing results for clone 1 with primers: M13_fwd and TE_spez_fwd verifying integration of *P-15* into pSET152_ermE::*ikaA*.

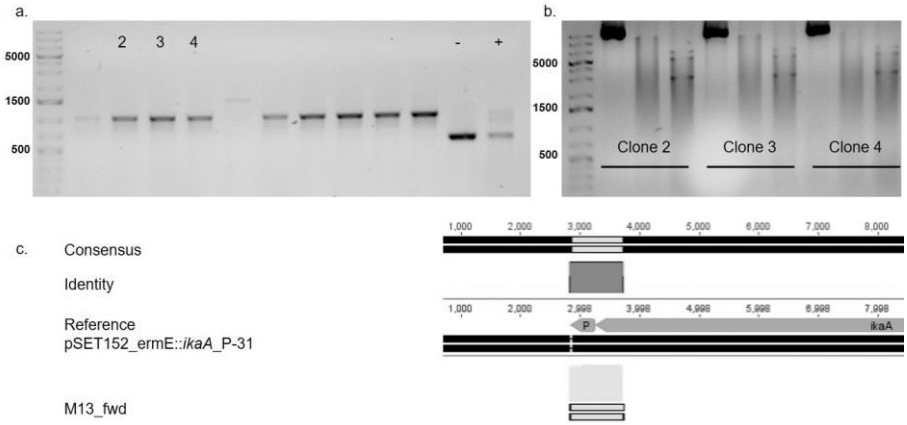


Figure A58. Cloning verification of pSET152_ermE::*ikaA*_P-31. a. Results of colony PCR with possible positive clones 2, 3, and 4 negative control conducted with pSET152_ermE_*ikaA*, positive control 1 µl ligation reaction mixture. b. Results of analytical restriction digest of all clones, as clone 2 showed the predicted restriction this clone was submitted for sequencing. c. Sequencing results for clone 2 with primers: M13_fwd verifying integration of *P-31* into pSET152_ermE::*ikaA*.

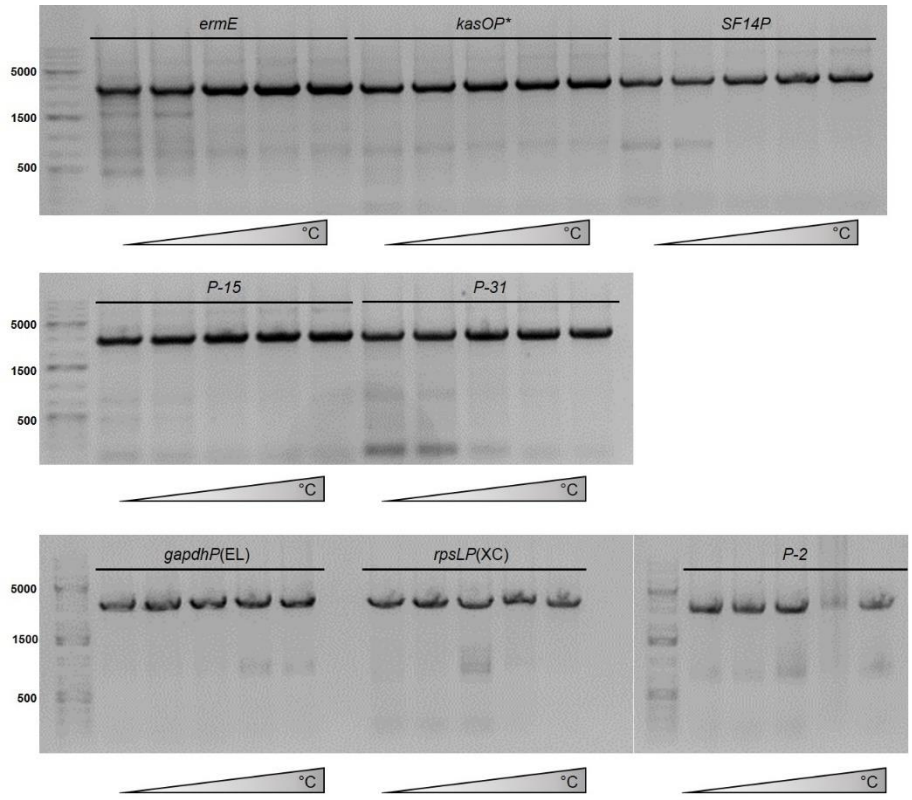


Figure A59. Results of temperature gradient PCR to amplify *ikaBC*.

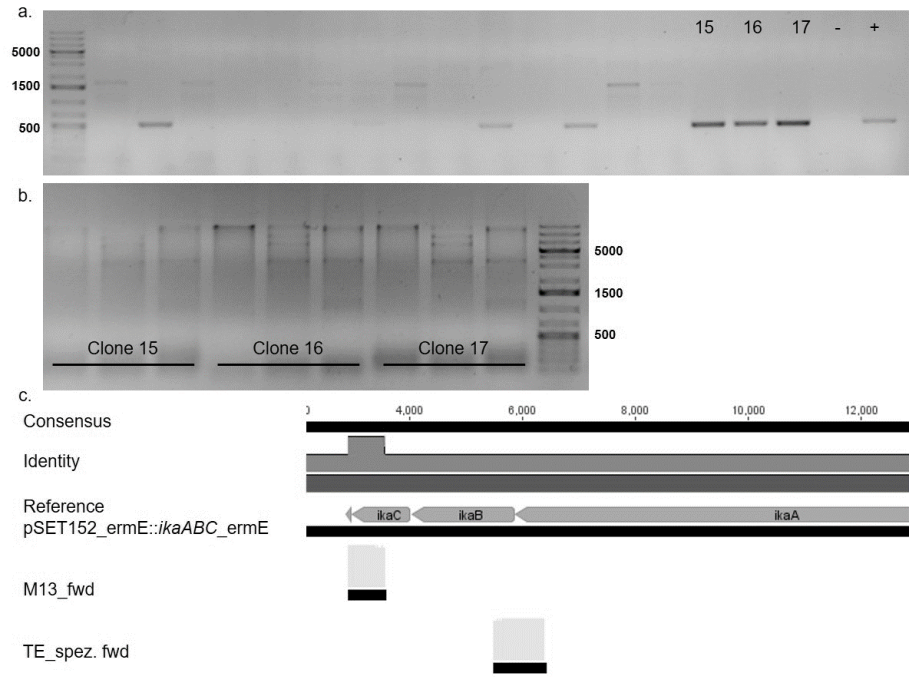


Figure A60. Cloning verification of pSET152_ermE::*ikaABC*_ermE. a. Results of colony PCR with possible positive clones 15, 16, and 17 negative control conducted with water, positive control 1 µl Gibson reaction mixture. b. Results of analytical restriction digest of all clones, as clone 16 showed the predicted restriction pattern this clone was submitted for sequencing. c. Sequencing results for clone 16 with primers: M13_fwd and TE_spez_fwd verifying integration of *ikaBC* into pSET152_ermE::*ikaA*_ermE.

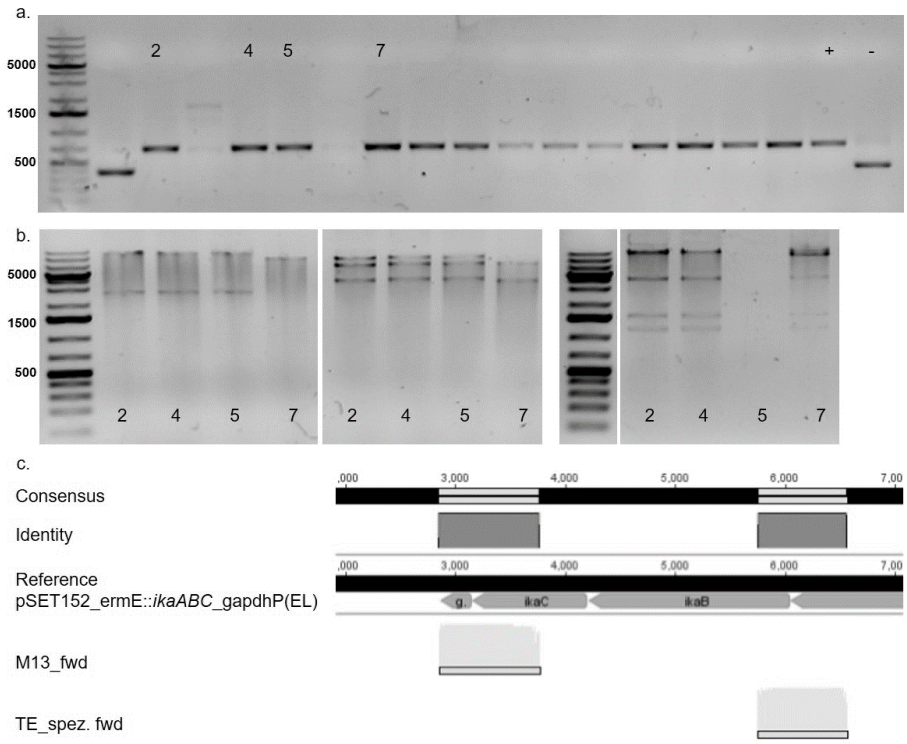


Figure A61. Cloning verification of pSET152_ermE::*ikaABC*_gapdhP(EL). a. Results of colony PCR with possible positive clones 2, 4, 5 and 7 negative control conducted with pSET152_ermE_*ikaA*, positive control 1 µl Gibson reaction mixture. b. Results of analytical restriction digest of all clones, as clone 2 showed the predicted restriction pattern this clone was submitted for sequencing. c. Sequencing results for clone 2 with primers: M13_fwd and TE_spez_fwd verifying integration of *ikaBC* into pSET152_ermE::*ikaA*_gapdhP(EL).

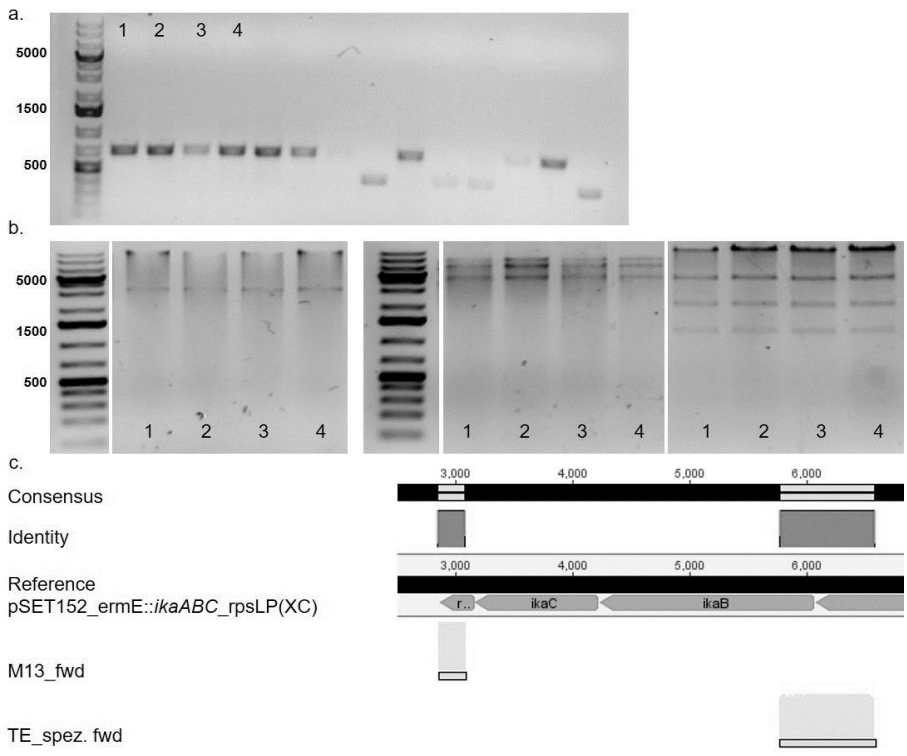


Figure A62. Cloning verification of pSET152_ermE::*ikaABC*_rpsLP(XC). a. Results of colony PCR with possible positive clones 1, 2, 3 and 4. b. Results of analytical restriction digest of all clones, as clone 3 showed the predicted restriction pattern this clone was submitted for sequencing. c. Sequencing results for clone 3 with primers: M13_fwd and TE_spez_fwd verifying integration of *ikaBC* into pSET152_ermE::*ikaA*_rpsLP(XC).

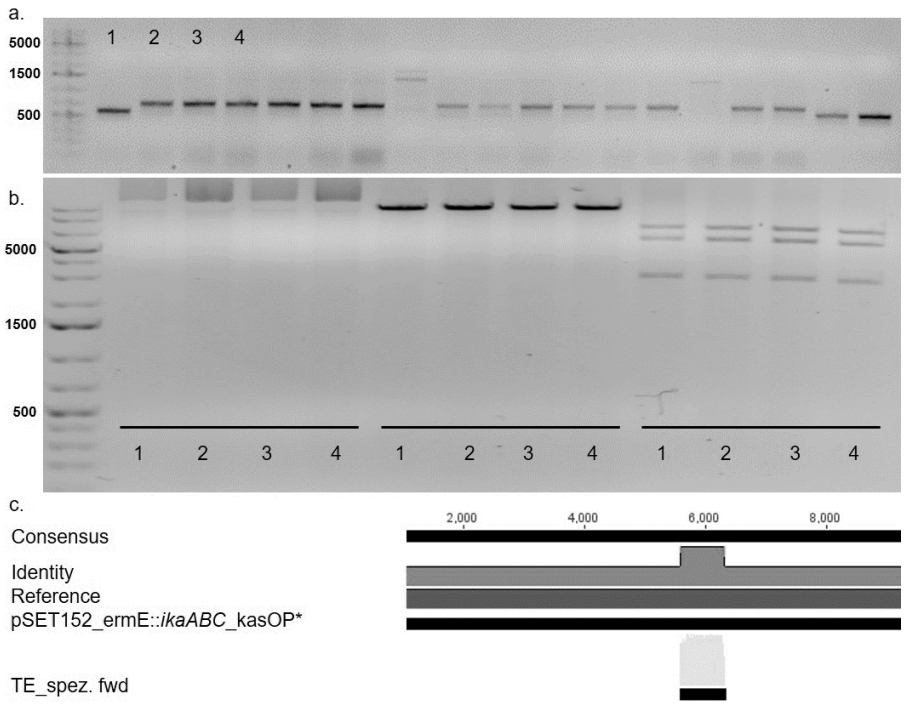


Figure A63. Cloning verification of pSET152_ermE::*ikaABC*_kasOP*. a. Results of colony PCR with possible positive clones 1, 2, 3 and 4. b. Results of analytical restriction digest of all clones, as clone 3 showed the predicted restriction pattern this clone was submitted for sequencing. c. Sequencing results for clone 3 with primers: TE_spez_fwd verifying integration of *ikaBC* into pSET152_ermE::*ikaA*_kasOP*.

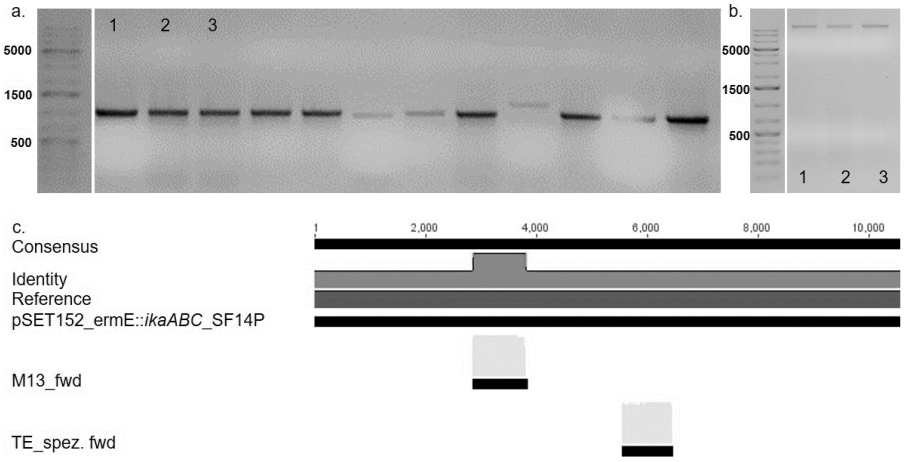


Figure A64. Cloning verification of pSET152_ermE::*ikaABC*_SF14P. a. Results of colony PCR with possible positive clones 1, 2, and 3. b. Results of analytical restriction digest of all clones, as clone 1 showed the predicted restriction pattern this clone was submitted for sequencing. c. Sequencing results for clone 1 with primers: M13_fwd and TE_spez_fwd verifying integration of *ikaBC* into pSET152_ermE::*ikaA*_SF-14.

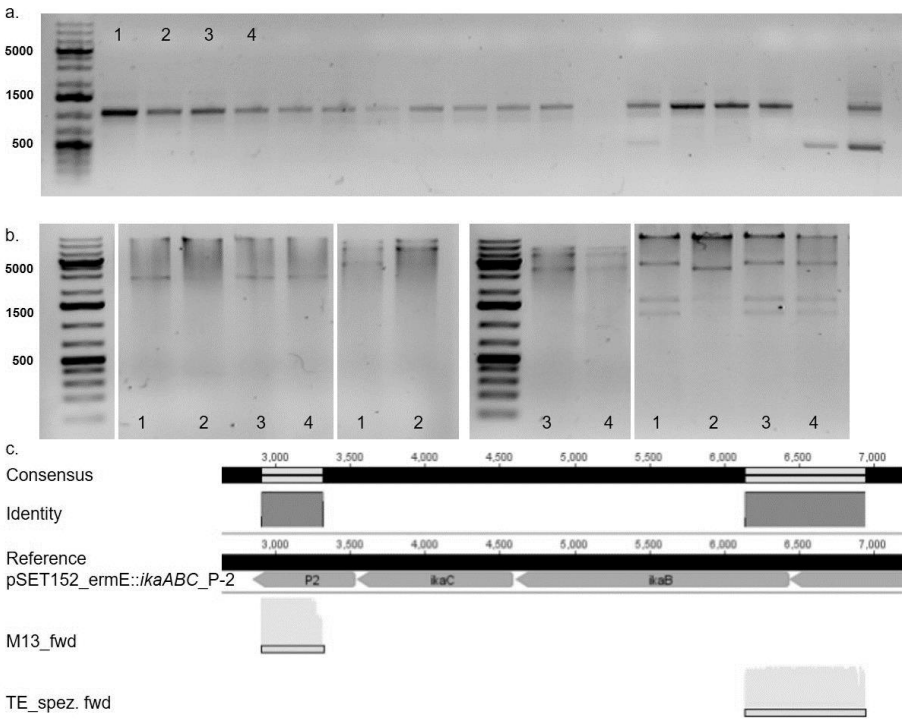


Figure A65. Cloning verification of pSET152_ermE::*ikaABC*_P-2. a. Results of colony PCR with possible positive clones 1, 2, 3, and 4. b. Results of analytical restriction digest of all clones, as clone 1 showed the predicted restriction pattern this clone was submitted for sequencing. c. Sequencing results for clone 1 with primers: M13_fwd and TE_spez_fwd verifying integration of *ikaBC* into pSET152_ermE::*ikaA*_P-2.

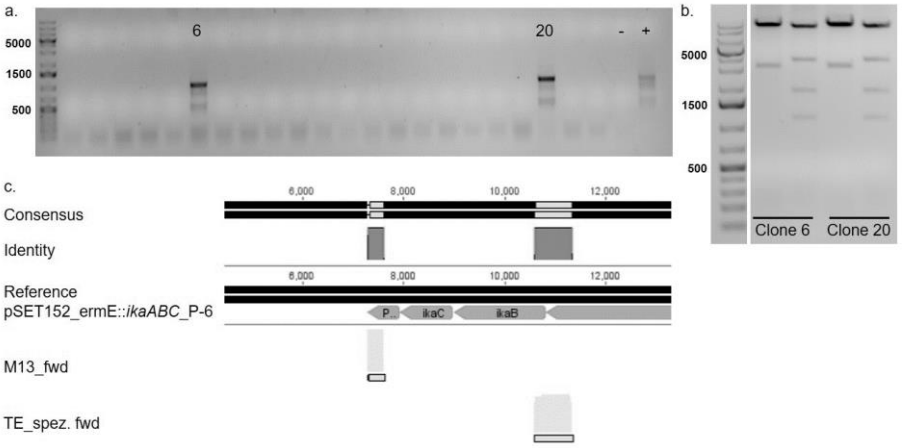


Figure A66. Cloning verification of pSET152_ermE::*ikaABC*_P-6. a. Results of colony PCR with possible positive clones 6 and 20, negative control conducted with water, positive control 1 µl Gibson reaction mixture. b. Results of analytical restriction digest of all clones, as clone 20 showed the predicted restriction pattern this clone was submitted for sequencing. c. Sequencing results for clone 20 with primers: M13_fwd and TE_spez_fwd verifying integration of *ikaBC* into pSET152_ermE::*ikaA*_P-6.

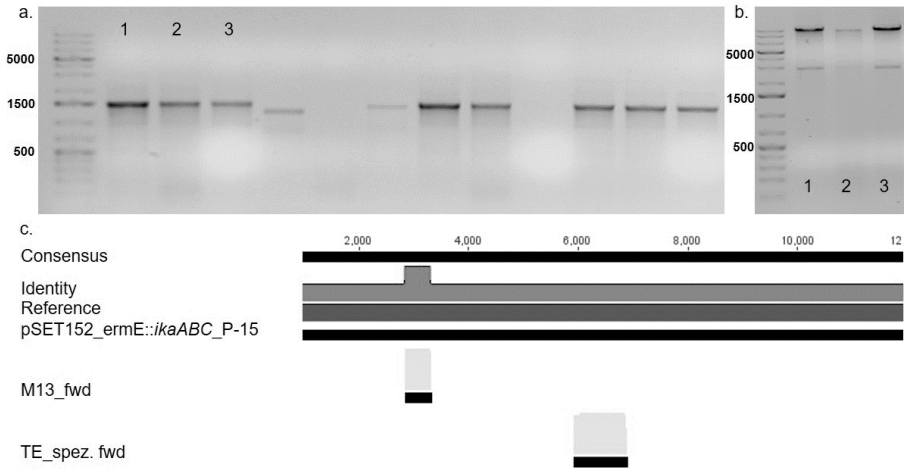


Figure A67. Cloning verification of pSET152_ermE::*ikaABC*_P-15. a. Results of colony PCR with possible positive clones 1, 2, and 3. b. Results of analytical restriction digest of all clones, as clone 1 showed the predicted restriction pattern this clone was submitted for sequencing. c. Sequencing results for clone 1 with primers: M13_fwd and TE_spez_fwd verifying integration of *ikaBC* into pSET152_ermE::*ikaA*_P-15.

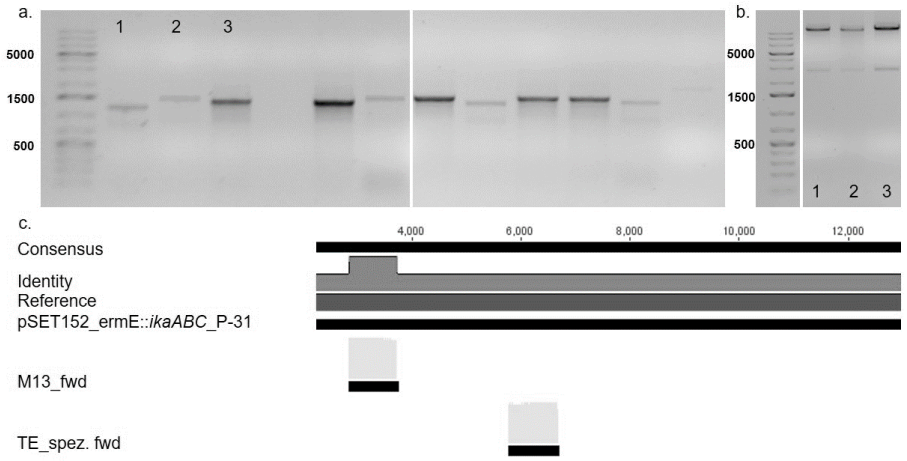


Figure A68. Cloning verification of pSET152_ermE::*ikaABC*_P-31. a. Results of colony PCR with possible positive clones 1, 2, and 3. b. Results of analytical restriction digest of all clones, as clone 1 showed the predicted restriction pattern this clone was submitted for sequencing. c. Sequencing results for clone 1 with primers: M13_fwd and TE_spez_fwd verifying integration of *ikaBC* into pSET152_ermE::*ikaA*_P-31.

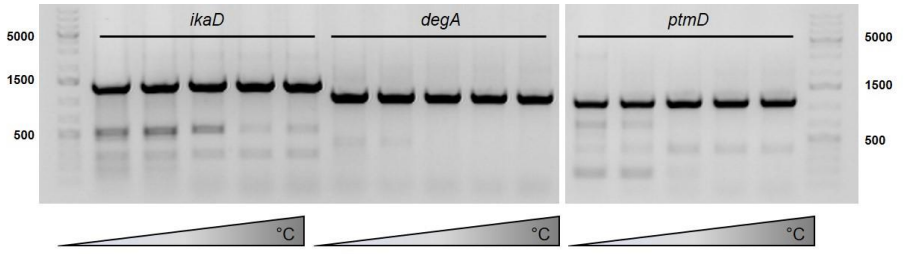


Figure A69. Temperature gradient PCR to amplify *ikaD*, *degA*, and *ptmD*.

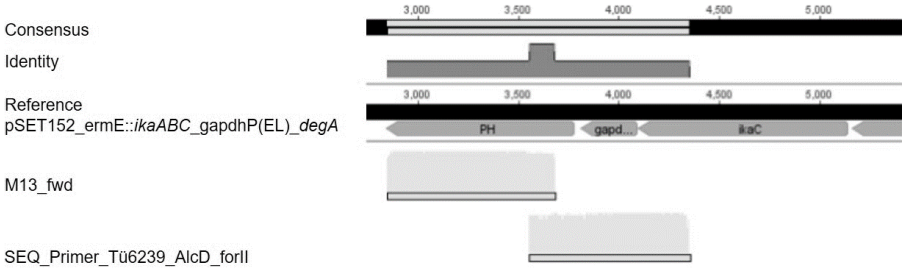


Figure A70. Sequencing result pSET152_ermE::*ikaABC*_gapdhP(EL)_*degA*.

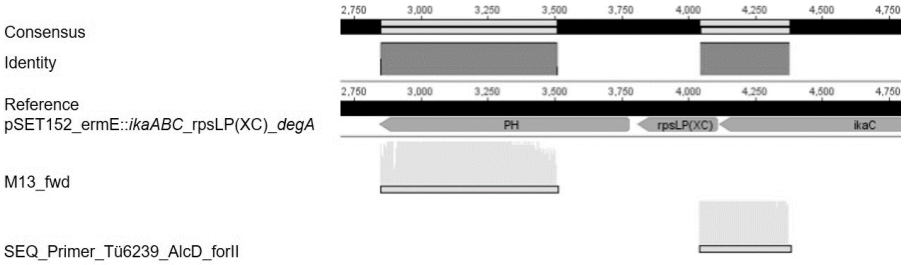


Figure A71. Sequencing result pSET152_ermE::*ikaABC*_rpsLP(XC)_*degA*.

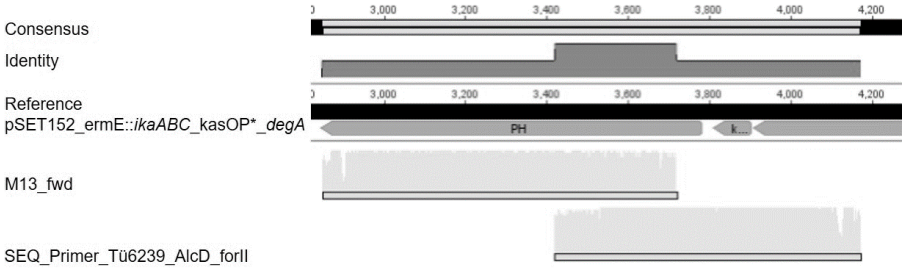


Figure A72. Sequencing result pSET152_ermE::*ikaABC*_kasOP*_*degA*.

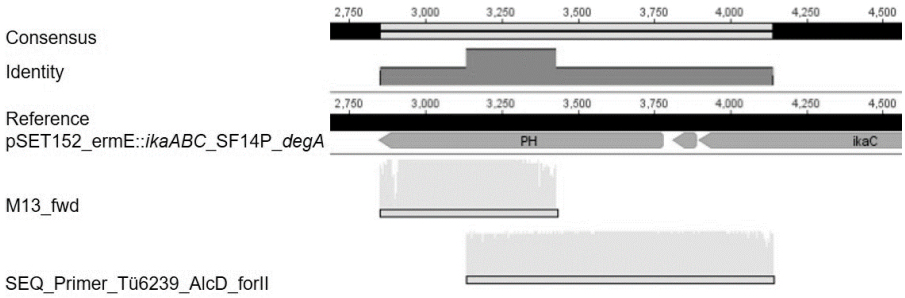


Figure A73. Sequencing result pSET152_ermE::*ikaABC*_SF14P_*degA*.

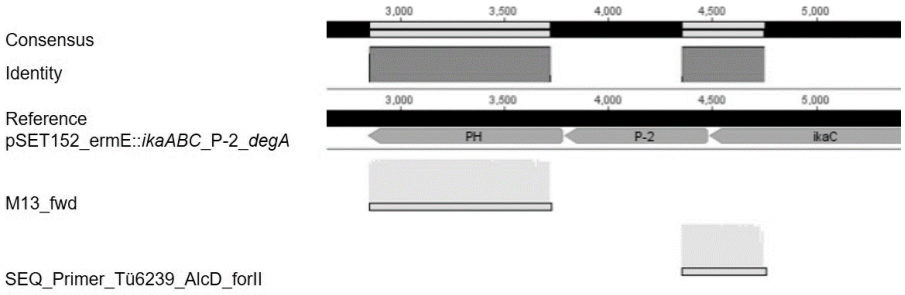


Figure A74. Sequencing result pSET152_ermE::*ikaABC*_P-2_*degA*.

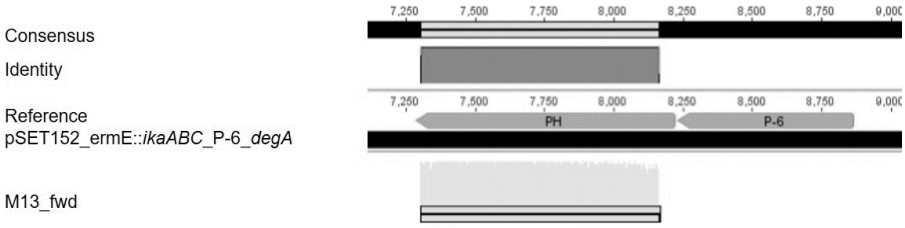


Figure A75. Sequencing result pSET152_ermE::*ikaABC*_P-6_*degA*.

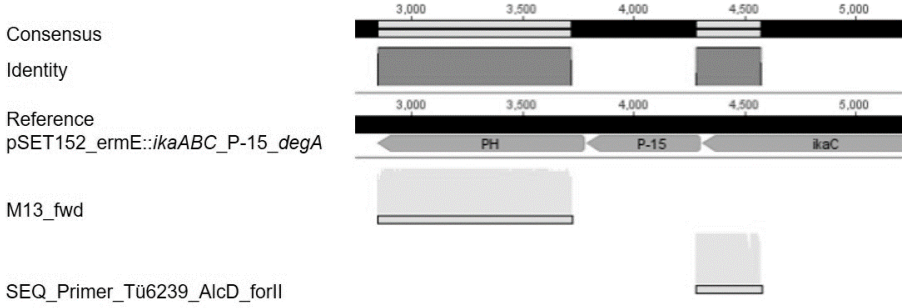


Figure A76. Sequencing result pSET152_ermE::*ikaABC*_P-15_*degA*.

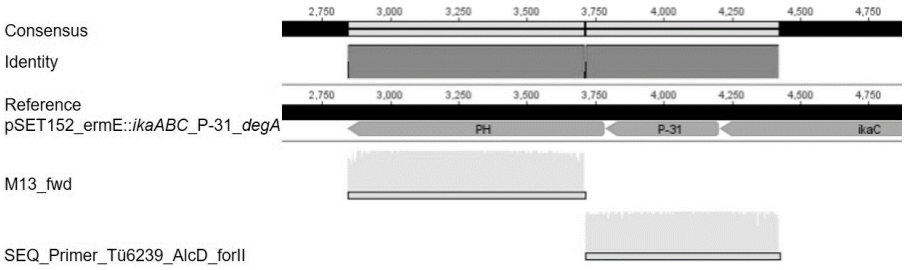


Figure A77. Sequencing result pSET152_ermE::*ikaABC*_P-31_*degA*.

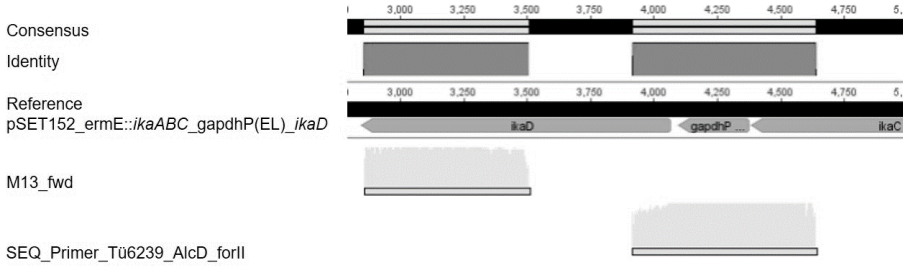


Figure A78. Sequencing result pSET152_ermE::*ikaABC*_gapdhP(EL)_*ikaD*.

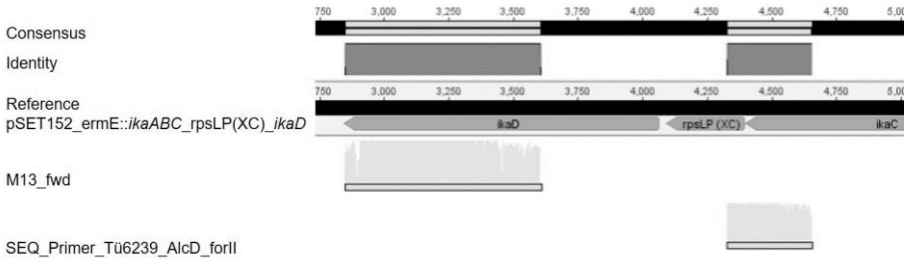


Figure A79. Sequencing result pSET152_ermE::*ikaABC*_rpsLP(XC)_*ikaD*.

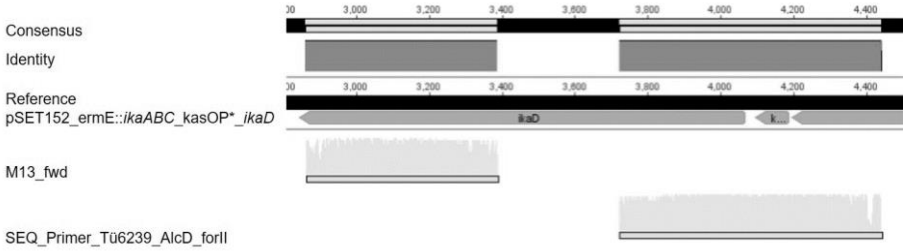


Figure A80. Sequencing result pSET152_ermE::*ikaABC*_kasOP*_*ikaD*.

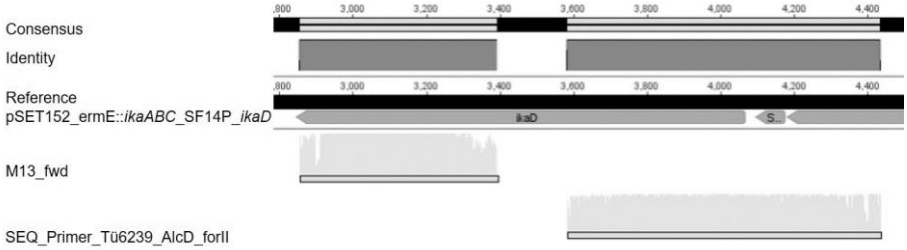


Figure A81. Sequencing result pSET152_ermE::*ikaABC*_SF14P_*ikaD*.

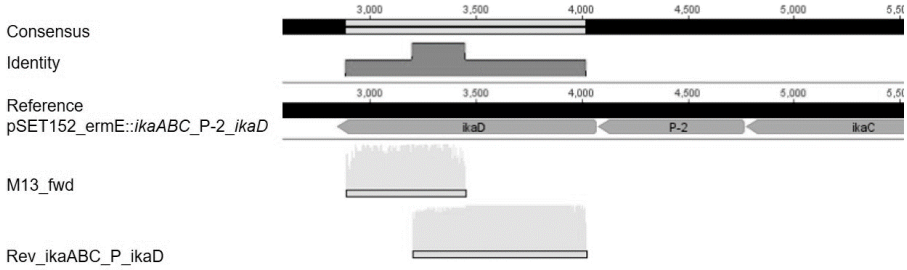


Figure A82. Sequencing result pSET152_ermE::*ikaABC*_P-2_*ikaD*.

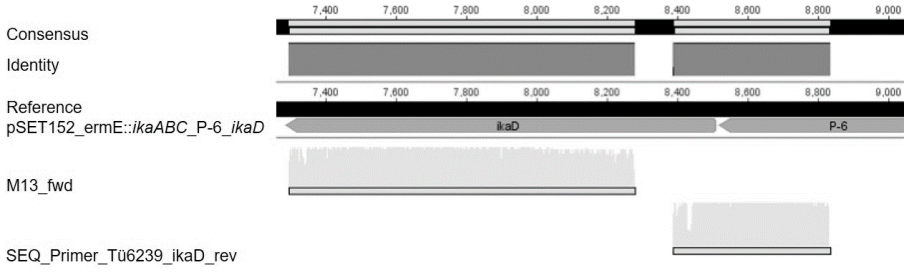


Figure A83. Sequencing result pSET152_ermE::*ikaABC*_P-6_*ikaD*.

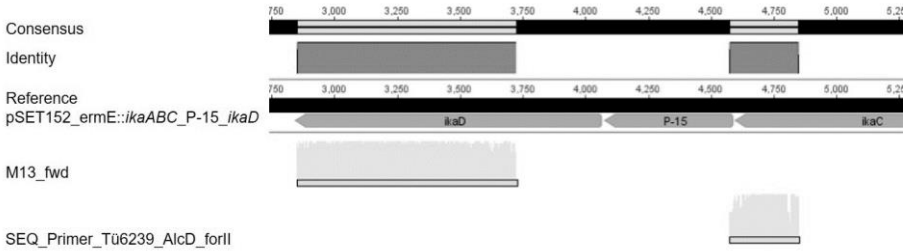


Figure A84. Sequencing result pSET152_ermE::*ikaABC*_P-15_*ikaD*.

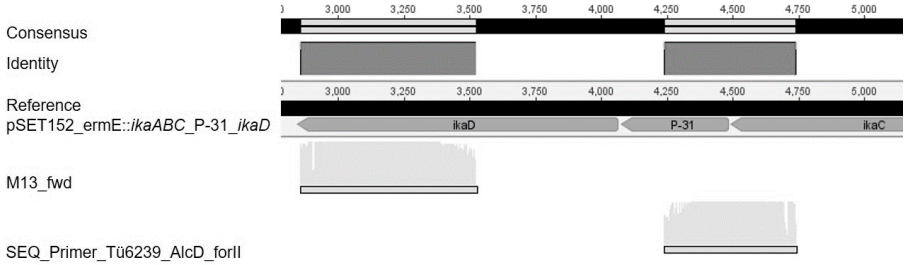


Figure A85. Sequencing result pSET152_ermE::*ikaABC*_P-31_*ikaD*.

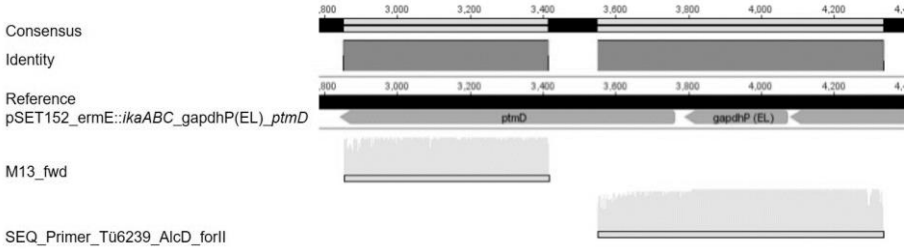


Figure A86. Sequencing result pSET152_ermE::*ikaABC*_gapdhP(EL)_*ptmD*.

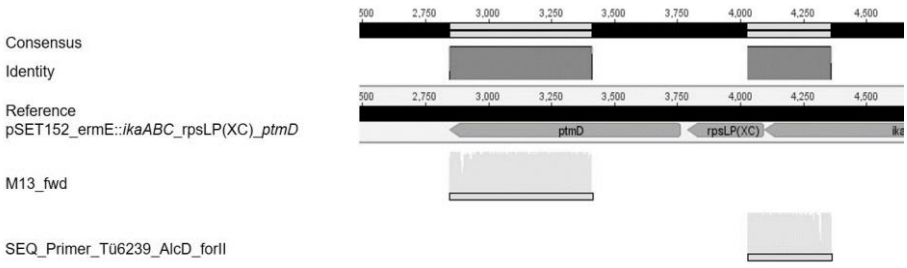


Figure A87. Sequencing result pSET152_ermE::*ikaABC*_rpsLP(XC)_*ptmD*.

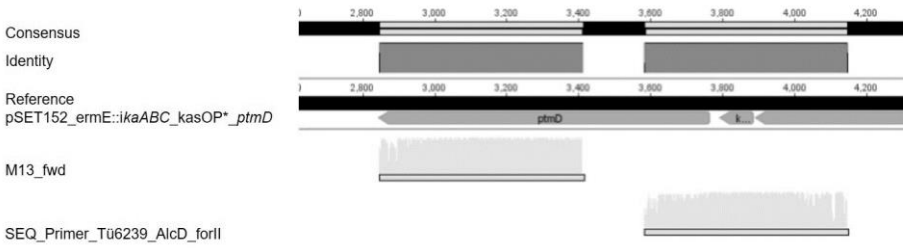


Figure A88. Sequencing result pSET152_ermE::*ikaABC*_kasOP*_*ptmD.*

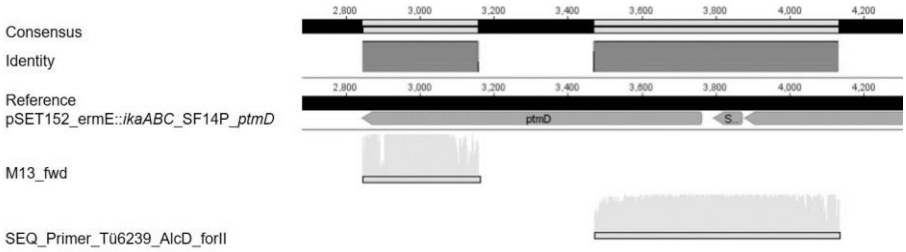


Figure A89. Sequencing result pSET152_ermE::*ikaABC*_SF14P_*ptmD.*

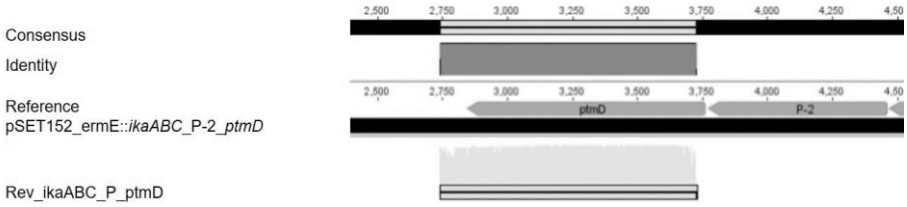


Figure A90. Sequencing result pSET152_ermE::*ikaABC*_P-2_*ptmD.*

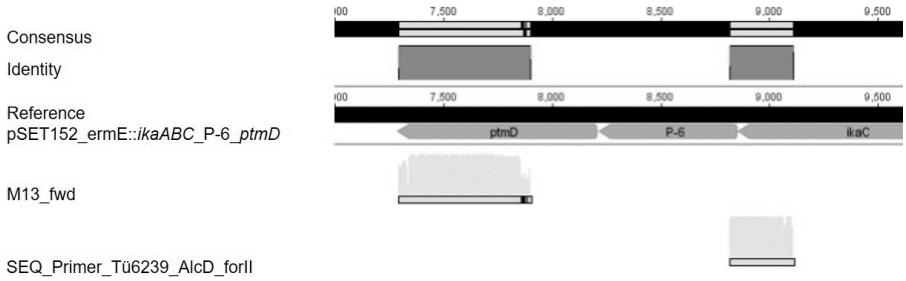


Figure A91. Sequencing result pSET152_ermE::*ikaABC*_P-6_*ptmD.*

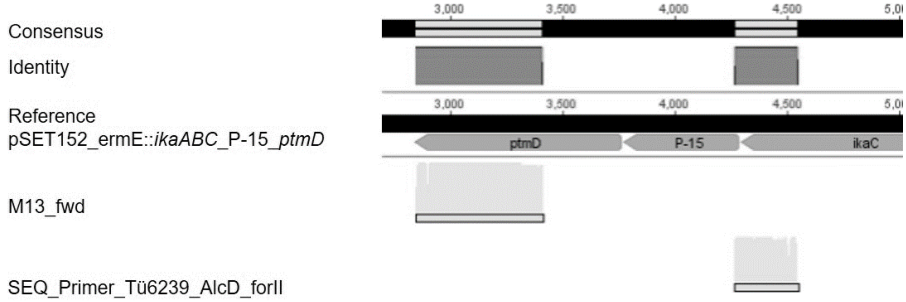


Figure A92. Sequencing result pSET152_ermE::*ikaABC*_P-15_*ptmD.*

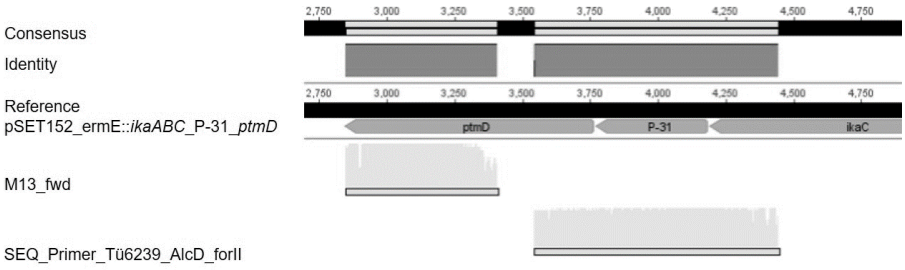


Figure A93. Sequencing result pSET152_ermE::*ikaABC*_P-31_*ptmD.*

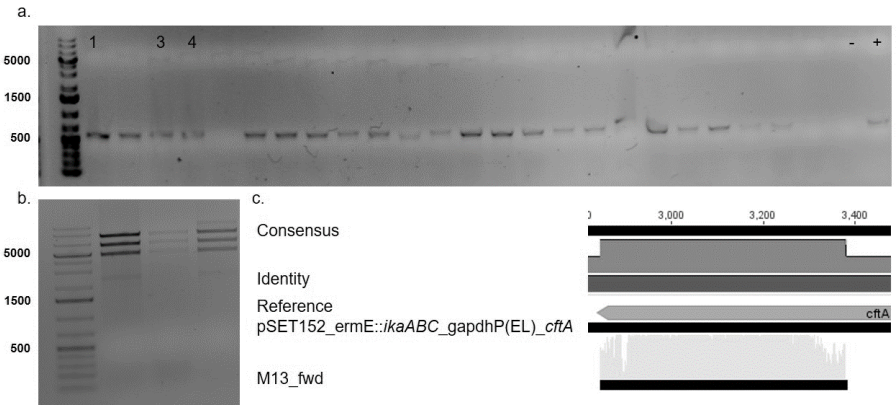


Figure A94. Cloning verification of pSET152_ermE::*ikaABC*_gapdhP(EL)_*cftA*. a. Results of colony PCR with possible positive clones 1, 3, and 4. b. Results of analytical restriction digest of all clones, as clone 1 showed the predicted restriction pattern this clone was submitted for sequencing. c. Sequencing results for clone 1 with primer M13_fwd verifying integration of *cftA* into pSET152_ermE::*ikaABC*_gapdhP(EL).

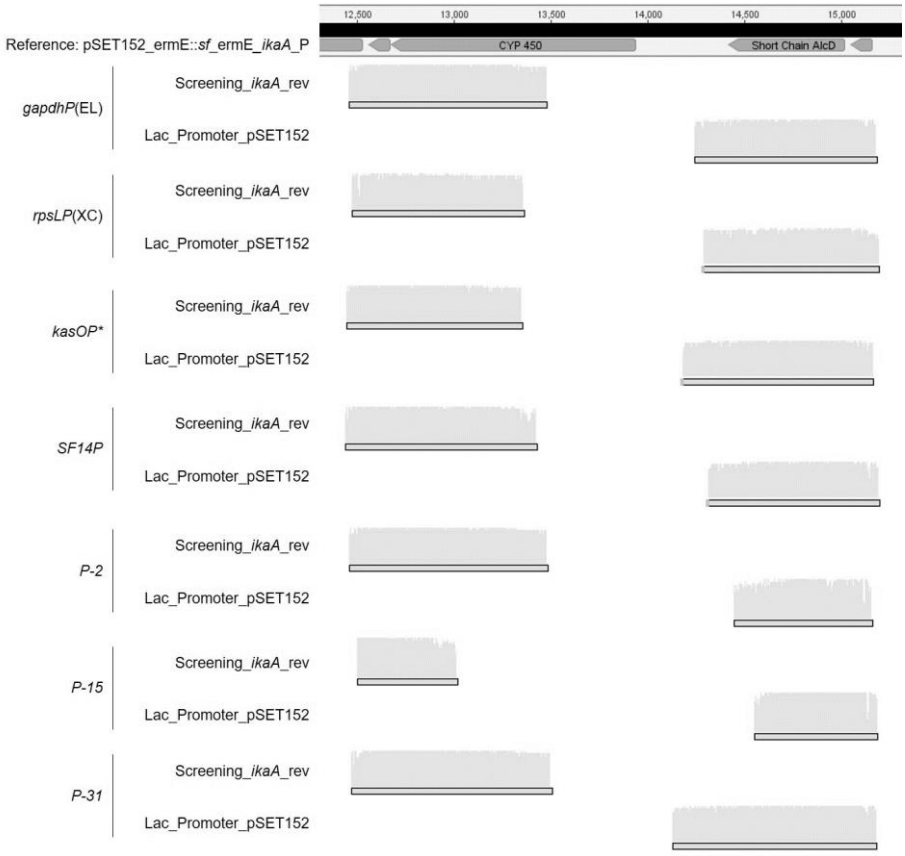


Figure A95. Sequencing results pSET152_ermE::*sf*_ermE_*ikaA*_P constructs.

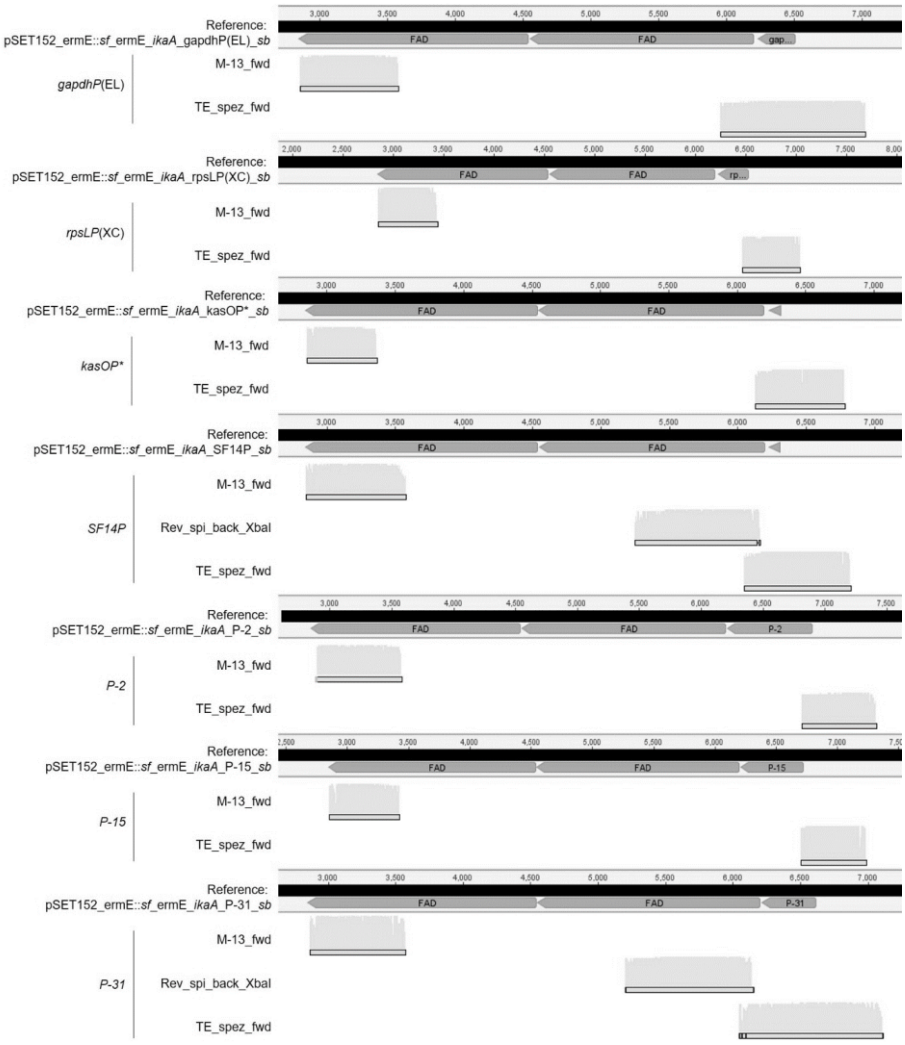


Figure A96. Sequencing results pSET152_ermE::*sf*_ermE_*ikaA*_P_*sb* constructs.

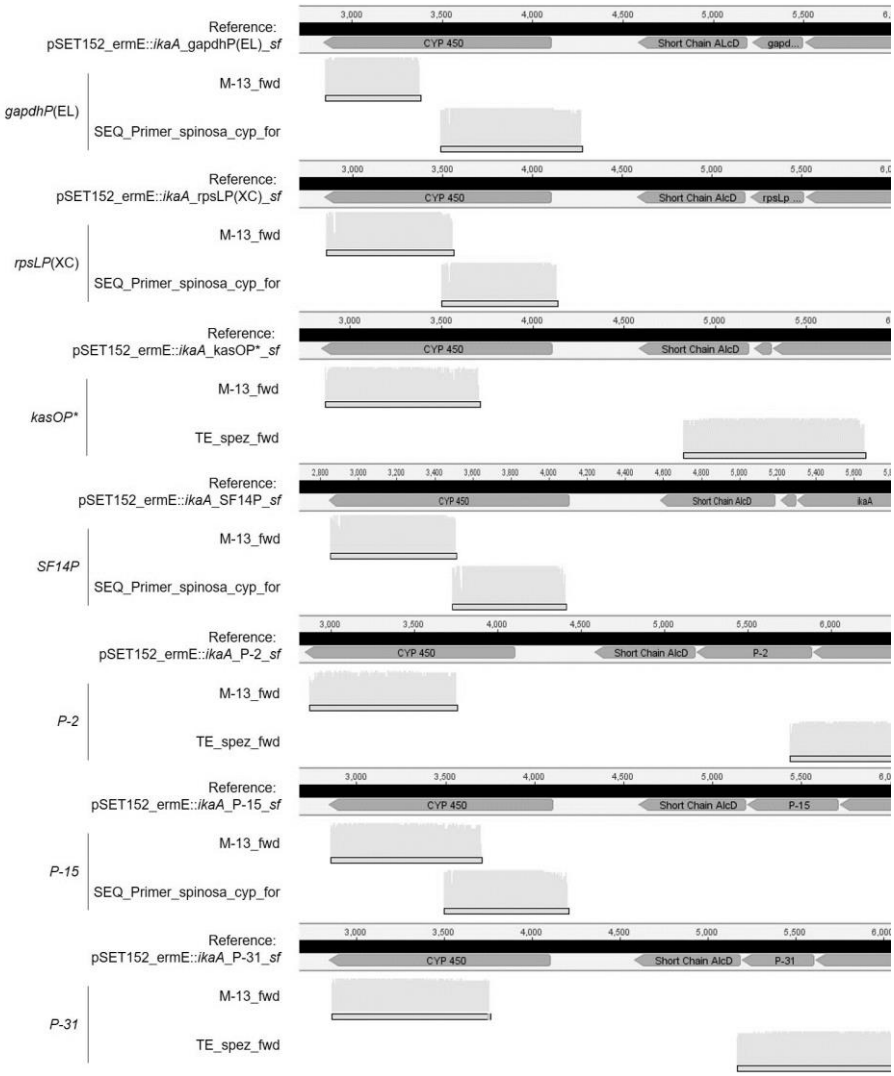


Figure A97. Sequencing results pSET152_ermE::*ikaA*_P_*sf* constructs.

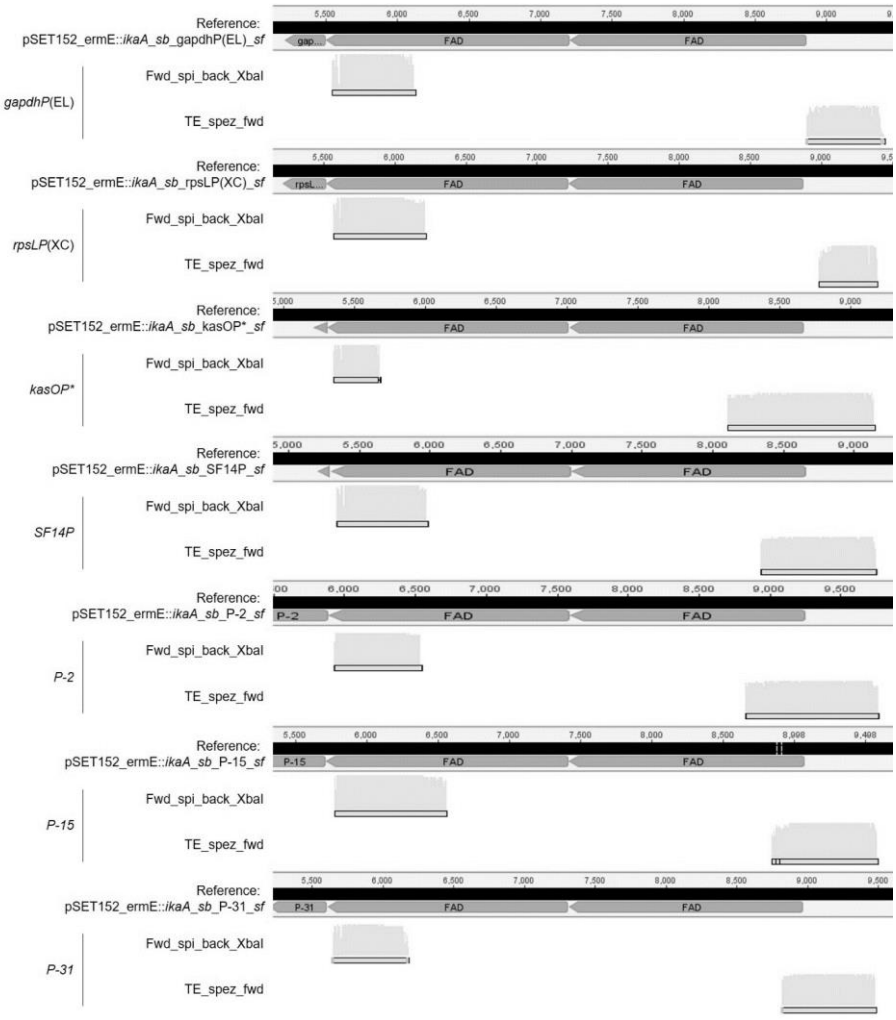


Figure A98. Sequencing results pSET152_ermE::*ikaA_sb*_P_*sf* constructs.

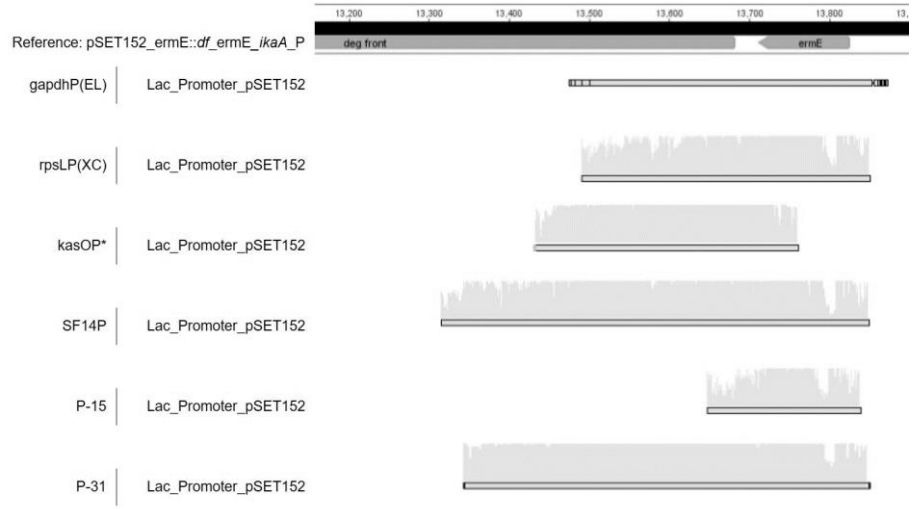


Figure A99. Sequencing results pSET152_ermE::*df*_ermE_*ikaA*_P constructs.

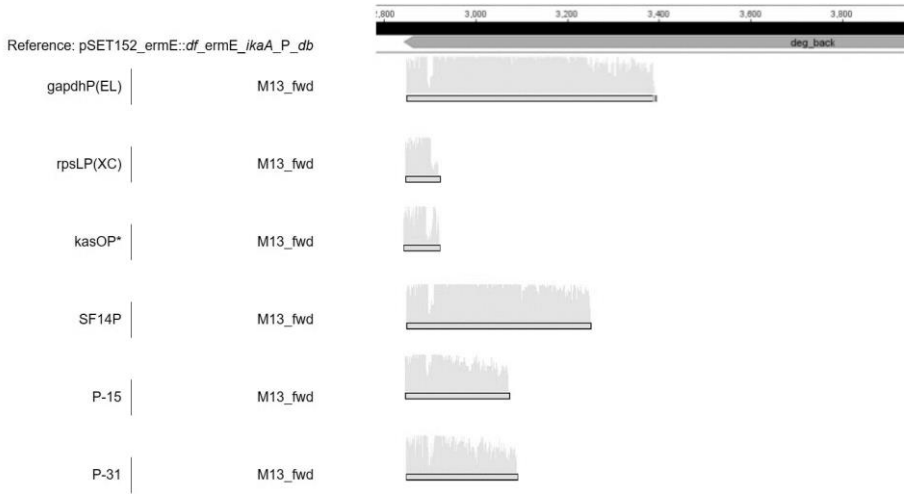


Figure A100. Sequencing results pSET152_ermE::*df*_ermE_*ikaA*_P_*db* constructs.

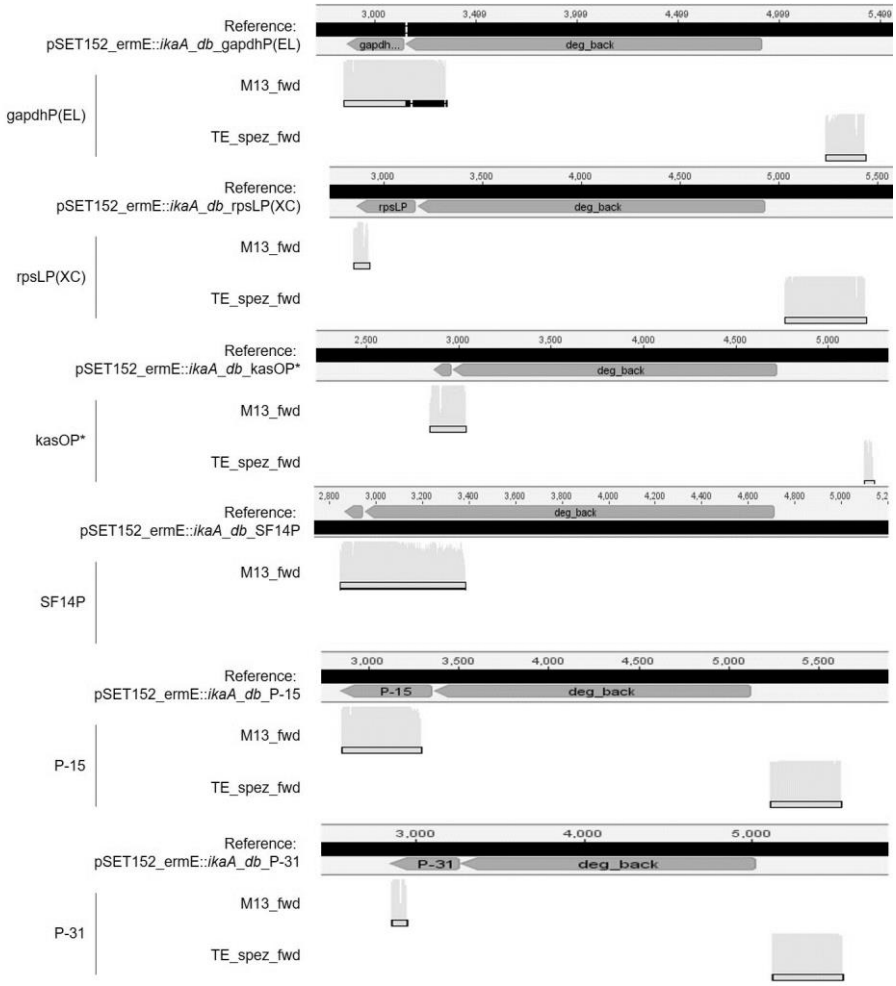


Figure A101. Sequencing results pSET152_ermE::*ikaA_db*_P constructs.

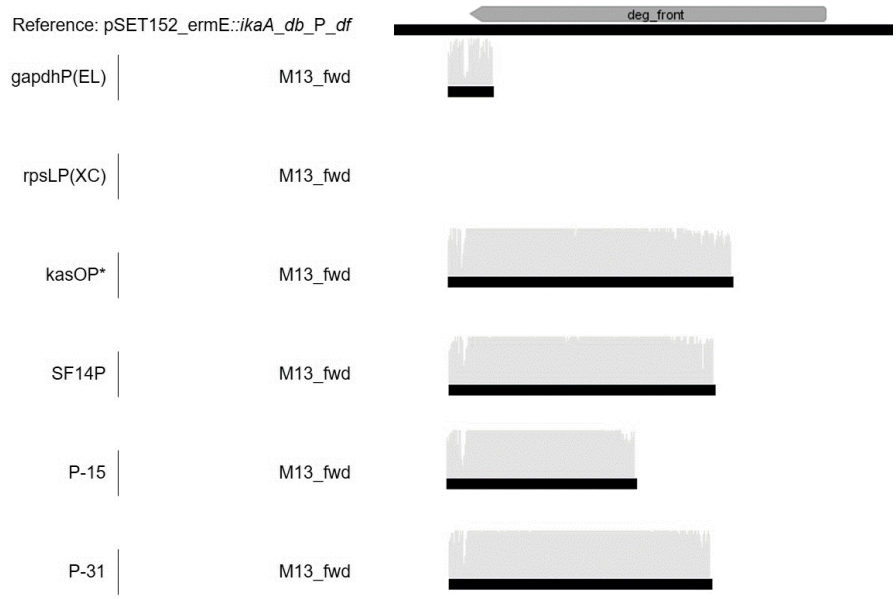


Figure A102. Sequencing results pSET152_ermE::*ikaA_db*_P_*df* constructs.

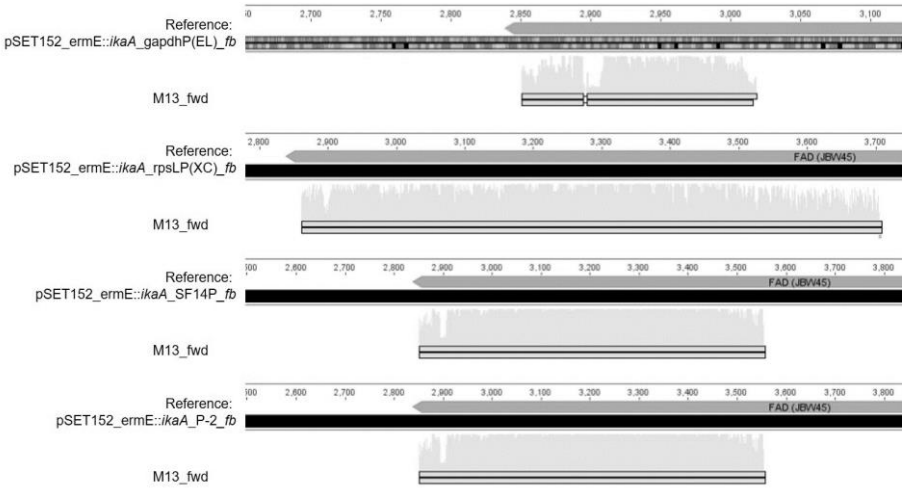


Figure A 103. Sequencing results pSET152_ermE::*ikaA*_P_*fb* constructs.

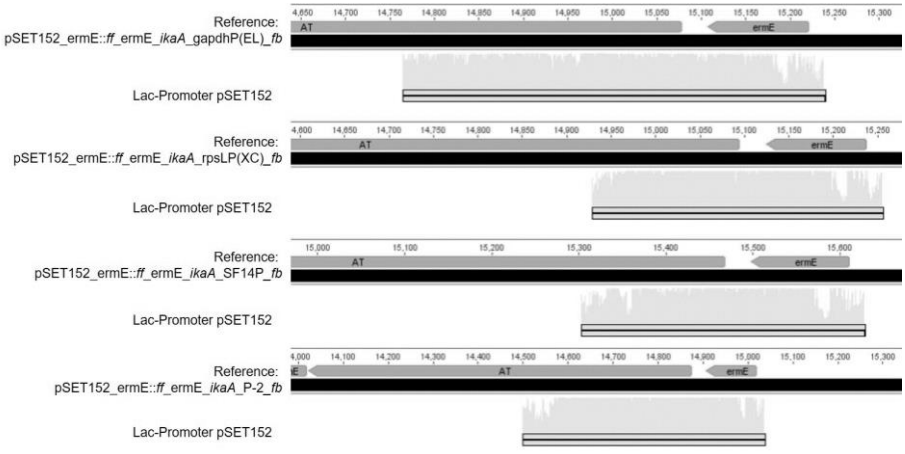


Figure A 104. Sequencing results pSET152_ermE::*ff*_ermE_*ikaA*_P_*fb* constructs.

Additional results

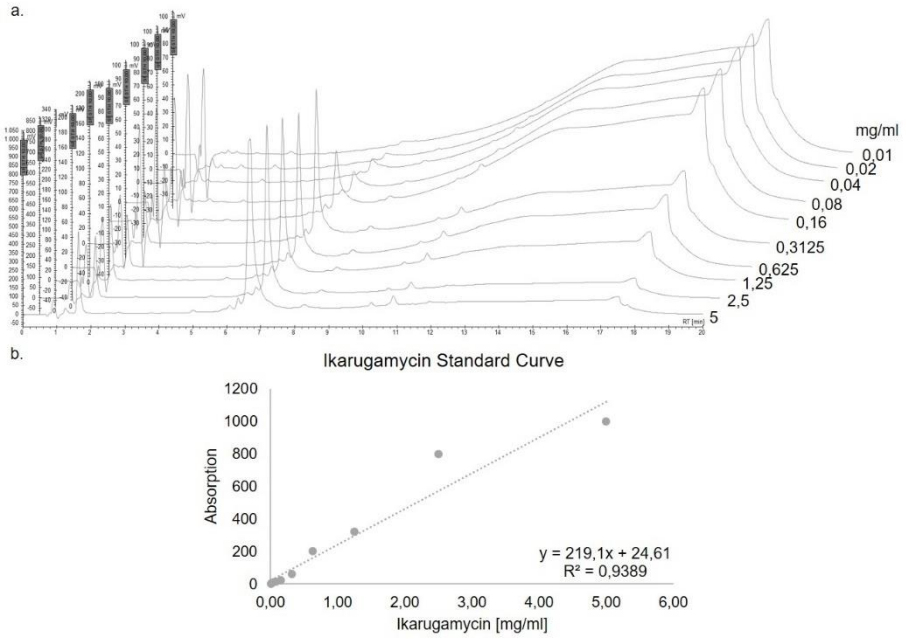


Figure A105. Calculating HPLC standard curve for **39**. a. HPLC chromatograms measuring increasing amounts of **39**. b. calculated standard curve to determine **39** production in heterologous expression system. Analytical method: Eurosphere_IkaFast_CG.Meth.

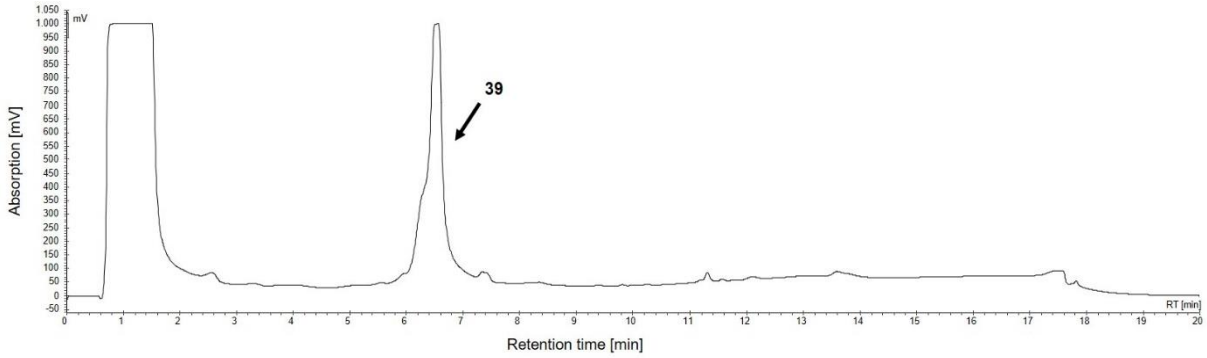


Figure A106. HPLC chromatogram of heterologous expression of **39** using the novel plug-and-play expression vector. *S. albus* cell pellet extracts harboring pSET152_ermE::*ikaABC*_SF14P, as a proof of the heterologous expression of **39** using the novel expression vectors. Analytical method: Eurosphere_IkaFast_CG.Meth.

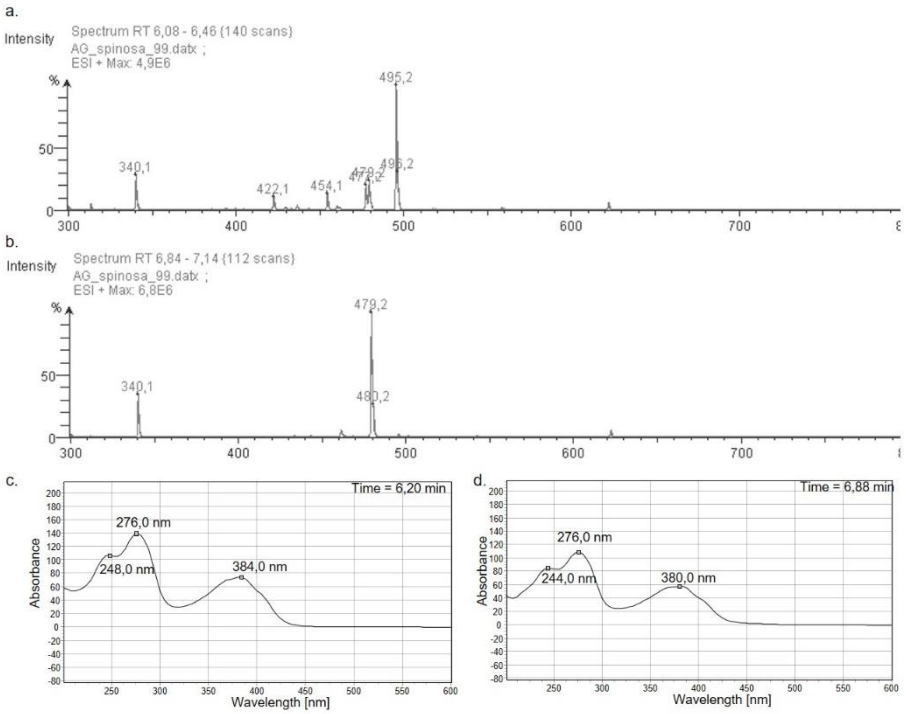


Figure A107. Additional results *S. albus* pSET152_ermE::*spi*_back_new. a. Mass results of first new peak at time 6.08-6.46 min, b. mass results of second new peak at time 6.84-7.14 min. c. UV-spectrum of first new peak at time 6.20 min and d. UV-spectrum of second new peak at time 6.88 min.

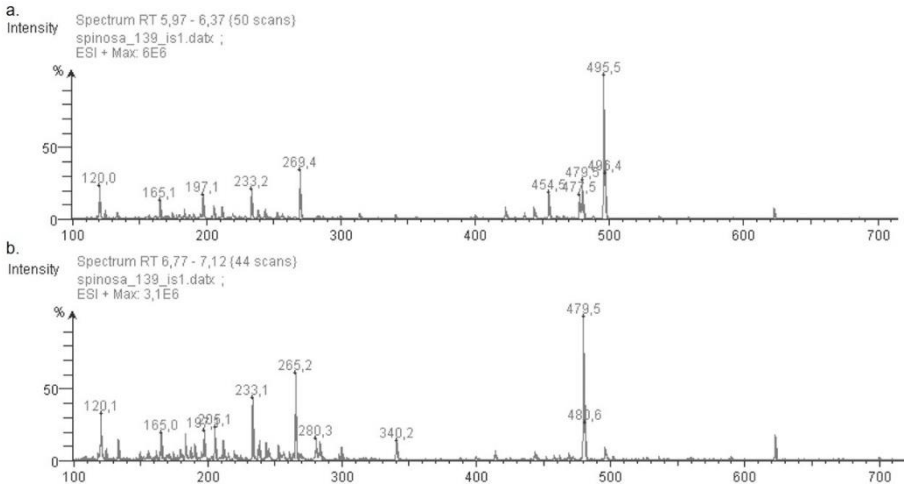


Figure A108. Additional results *S. albus* pSET152_ermE::*spi*_full_new. a. Mass results of first new peak at time 5.97-6.37 min, b. mass results of second new peak at time 6.77-7.12 min.

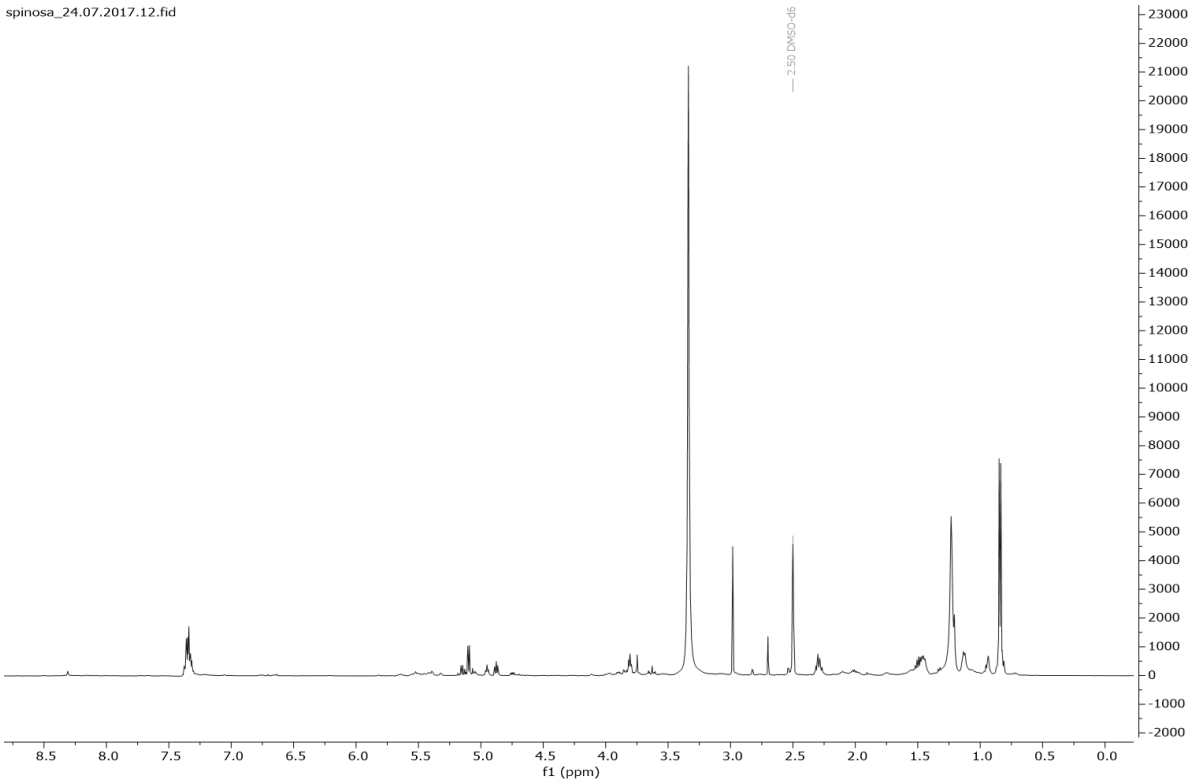


Figure A109. Preliminary NMR data of purified *spi* cluster compound.

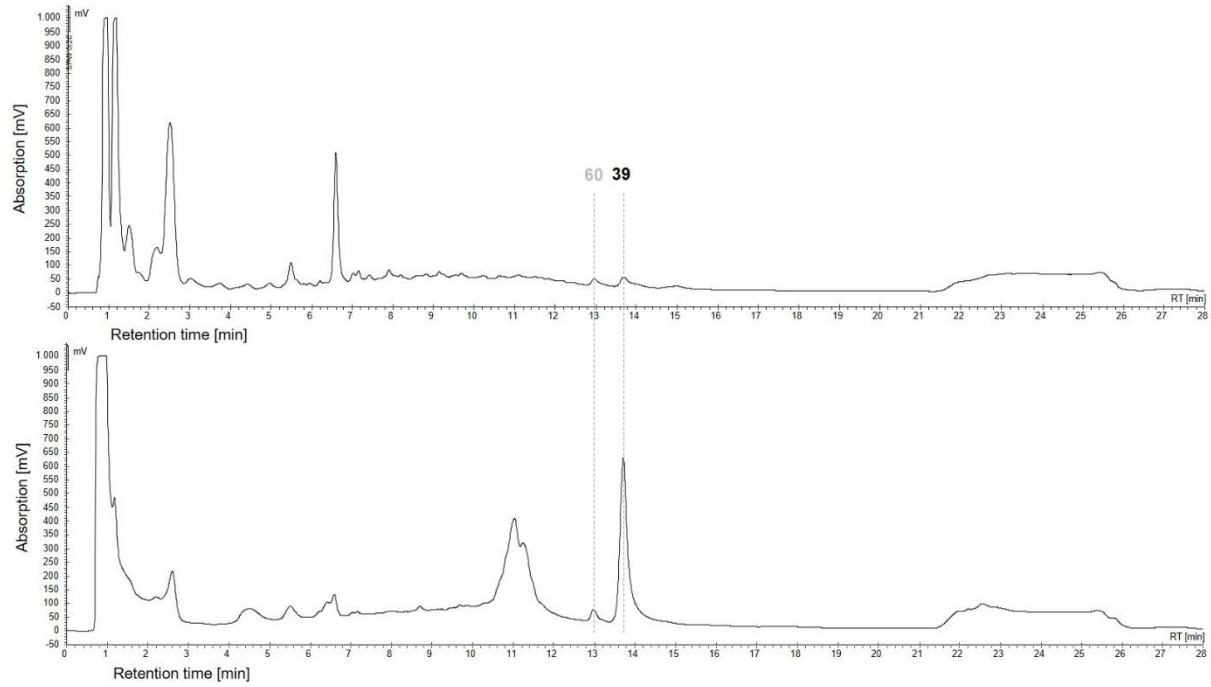


Figure A110. HPLC chromatograms to check for DegA conversion of **39** in a large-scale fermentation. Supernatant (top) and cell pellet (bottom) extracts of 1 L culture of S. albus harboring pSET152_ermE::*ikaABC*_rpsLP(XC)_*degA*. Analytical method: Eurosphere_Ika_lange Methode_AG.Meth.

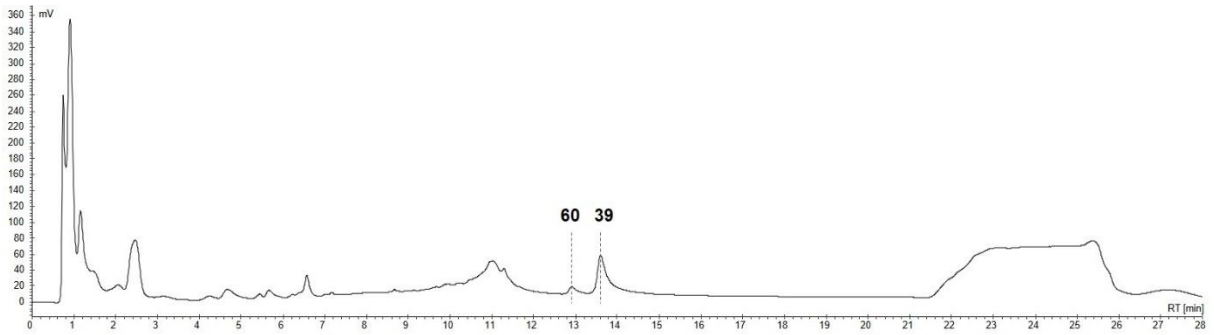


Figure A111. HPLC chromatogram of merged extracts containing DegA and PtmD modified **39**. Extracts of S. albus containing pSET152_ermE::ikaABC_gapdhp(EL)_ptmD and pSET152_erm::ikaABC_rpsLP(XC)_degA were merged in a 1:5 ration (ptmD:degA). Analytical method: Eurosphere_Ika_lange Methode_AG.Meth.

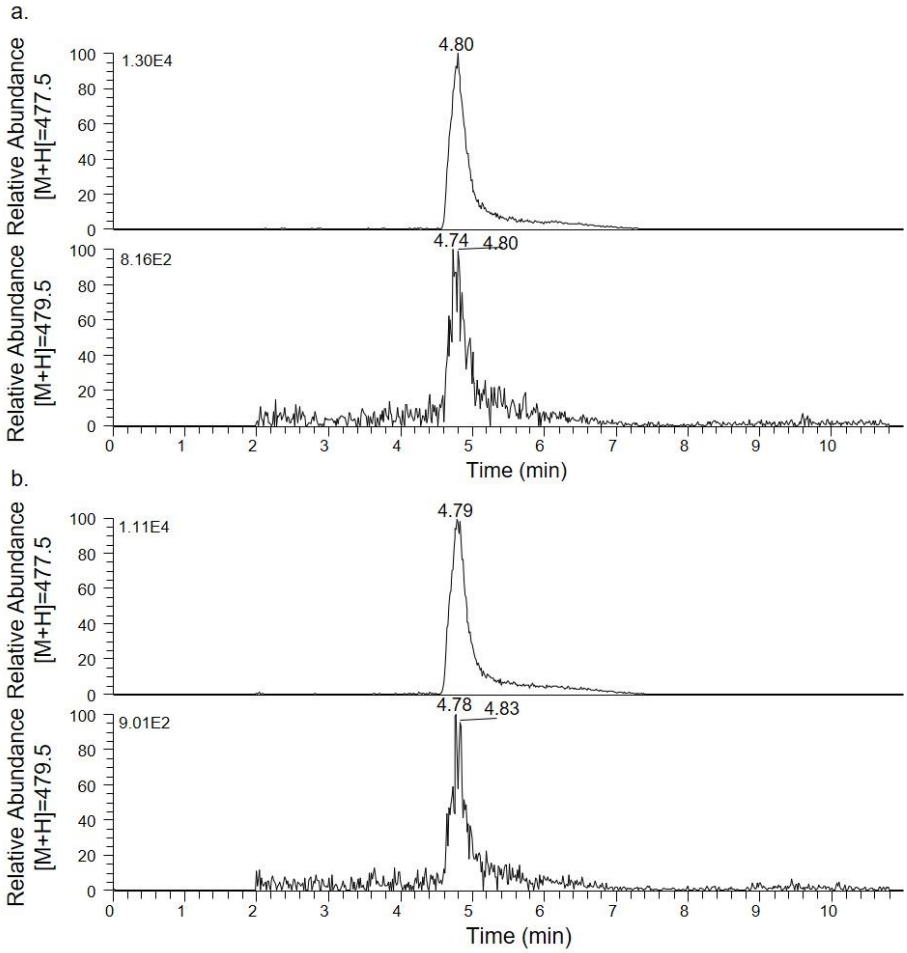


Figure A112. MS results of IkaC assay negative control in NaH_2PO_4 buffer. a. MS results of IkaC assay performed with NADPH, extracted mass for **50** at the top and for **39** at the bottom. b. MS results of IkaC assay performed with NADH, extracted mass for **50** at the top and for **39** at the bottom.

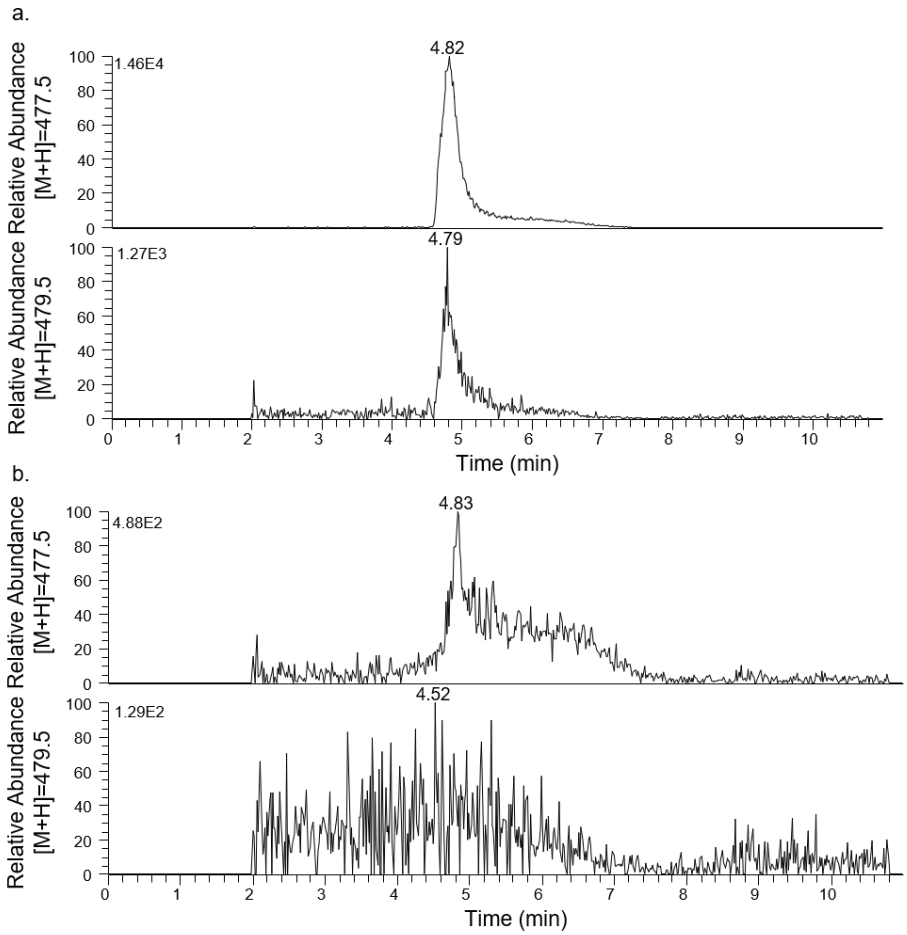


Figure A113. MS results of IkaC assay negative control in HEPES buffer. a. MS results of IkaC assay performed with NADPH, extracted mass for **50** at the top and for **39** at the bottom. b. MS results of IkaC assay performed with NADH, extracted mass for **50** at the top and for **39** at the bottom.

Additional information

Table 42. Promoter sequences used for plug-and-play study. Sequence of promoter, restriction sites are underlined (StuI: AGGCCT and XbaI: TCTAGA), additionally introduced RBS bold.

Promoter	Sequence
ermE	AGGCCTGCGGTCGATCTTGACGGCTGGCGAGAGGTGCGGGGAGGA TCTGACCGACGCGGTCCACACGTGGCACCGCGATGCTGTTGTGGGC ACAATCGTGCCGGTTGGTAGGATCCAGCGGAGCAACGGAGGTACGG ATCTAGA
gapdhP (EL)	AGGCCTGCTGCTCCTTCGGTCGGACGTGCGTCTACGGGCACCTTAC CGCAGCCGTCGGCTGTGCGACACGGACGGATCGGGCGAACTGGCC GATGCTGGGAGAAGCGCGCTGCTGTACGGCGCGCACCGGGTGCGG AGCCCCTCGGCGAGCGGTGTGAAACTTCTGTGAATGGCCTGTTCGGT TGCTTTTTTTATACGGCTGCCAGATAAGGCTTGCAGCATCTGGGCGG CTACCGCTATGATCGGGGCGTTCCTGCAATTCTTAGTGCGAGTATCT GAAAGGGGATACGC**CGAGCAACGGAGGTACGGAC**TCTAGA
rpsLP (XC)	AGGCCTGCCCTGCAGGCGGAAGTCAGGTAGACACGACTTCCGCTAG TCCTTGCAAGGTCTGCTGACGTGAGGCGGGGCGGTCGTTTTTGACC GCCCCGCCTTCGTCATGTAGGCTCGCTCGCTGTGCCTGGCGTGTCTT CAGACGCCCAGGTCCCGGTGCCGTGAGGCCCGGGCCATCGAGCCG GTGGTACGTGGCTGCGGTCCCCTTGTGAGGGCTGCGCGCCGTGTGC TGTCCGGCGCGCACAGCCTTGAATCCACCCGCGGGGGCCGGCCGG TCTCCGTGAGCTCGAGAAGACGACGGAGACGTA**CGAGCAACGGAGG TACGGAC**TCTAGA
*kasOP**	AGGCCTTGTTCACATTCGAACGGTCTCTGCTTTGACAACATGCTGTGC GGTGTTGTAAAGTCGTGGCCAGGAGAATACGACAGCGTGCAGGACT GGGGGAGTT**CGAGCAACGGAGGTACGGAC**TCTAGA
SF14P	AGGCCTCCTATCCAGGAGATATTATGAGTTACGTAGACCTACGCCTT GACCTTGATGAGGCGGCGTGAGCTACAATCAATACTCGATTA**CGAGC AACGGAGGTACGGAC**TCTAGA
P-2	AGGCCTGCCCGGCCATATCCGGCCCGGCCAAATCTCGGCCGGCCAC CTCGGCCTGGCCAGCCTGGCCCGGCCAATCTCGGCCCGACCAACTT CAGCCCGGCCGGCGCTTGAGGCCGATGAGCCGCGGAGCGGCGAGT CTTCCGCCCGGCCGGTCCGGGTGGCCTCAAGCGCCGGCCGGGCTG GTTTTGGTGCGGACACGTCTGACCGTGCCGGTCACCGATGGCCTCA AGCGCCGGCCGGGCTGGGAGTGGTGGCCGAGGCTTCGGGCGTACG TGCCAGCCCGCAAGGGGCTGCGGTGGGGTGGCCTCAAGCGCCGGC CGGGCTGAGGTTGGCTGGCTGGGCCGGGTTCGGCCGGTGGGTCGA GGTGGCCTGGCCGGGCTCGCCAGGGTGAGTTGGCCGACGGGCCGA GGCGGCCCGCCCGGGCTCCCCGGGCCGAGTTGGCGCGGCCAGGC CAGGGCTCAGCAGGGTGGGGGAGTGGGGCAGGCGGCCCGGTAGG GGAGTGCGGGAGGGCAGCGCGCCGCGCGCATTGGCACTCCGCT TGACCGAGTGCTAATCGCGGTCATAGTCTCAGCTCTGGCACTCCCCG CAGGAGAGTGCCAACACAGCGACGGGCAGGTCCGGCACCCGCGAC GACGGATCGACCTGGTCGCCACACTCAGATCAGTTAACCCCGTGATC TCCGAAGGGGGAGGTCGGATCTCTAGA
P-6	AGGCCTGGCGCCGACCGCACCACACTCACGAGGGCCCGCCCCACC AACAGGGGGCGGGCCCTCTGTGCTGGCCTCAGGCGCCGACCGGGC TCGGTGCCCTCAAGCGCCGGCCGGGCTCCAAGGGTGGCCTCAAGC GCCGGCCGGGCTGAGTTGGGCCGGTCTGGGCCCGCACGCGCGCCT CACTGACGGCCTCAAGCGCCGGCCGGGCTATCTATAGCCCGGCCGG CGCTTGAGGCCGTCTTTGGCGCGCGCCTGTGAGCGGACGGCCCGTC AAAGATCAGCCCGGCCGGCGCTTGAGGCCATCTTTCGAGCCCGGCC GGCGTTTGAGGCCACCCCACCCCCGCCCCGGCAGGGGCGGCCTGA CCTCCGCATCCGCCGGCGCGGACAGGGCACCCCCAGTAGACGGGC

	GCGGGGCCGGAGGCCCCTAGCGCCTTGCACTCTCCTACCCCGAGTG CTAATTATTGGCGTTAGCACTCTCCGAGTGAGAGTGACAGAAGGACC GGGTCGGTGAGGCCCGCTGGCCACGCGGGGCAAGGAACCGCGAGG CAGGCAGGCCGTCCGTCGCGGGCGCCAGCACGGTCCGGAGTATCC ACCCTCCCCCAGACAGAGTCCGGGGGGACCCCCAGTCCTGGGAGGA CCACTTCACTCTAGA
P-15	AGGCCTTCCGCGCCGCCGGCCCGACGGTGCCCGGCCCCGTACCCC CCCCGGGTGGTGCGGGGCCGGGCACCGGCCTTTTGGCGCTGCGGA GTTGACGGAAGTTGGCCGAACCGGATGCGCTCGGCGCCCGGGGGC TGAAAGATGCTCACAGCCCCTTTCCACGGCGGTCCGGGAGGGGAGG CCGGGCAACCGGTTTTCGGGGGCGGAGTGTCCGGTATGCGGACGG CCGCGCCCGATAGATGTGTAACGAGTCCGTTTCGCAACCATCTATCT CGGATCGGTTTGTCCGGATTTTGGAAGATGTGAGTGTCAGGTGTGAT CGAACCGAGACCAAAAGGGTGTGGTCGGGCCGAACACCATGGCTAA TAGTTGAGCGCGTAGAGCTCGGGTCAATGGGTCACGCGCTGTGGGG AGCGCCGACTCACGAGCACACTGGGGCACTCGATCTTCGCCGTCAG GGGTGTCGGCGGATCGTCCTGTGCCCTCTCTTGCAGTGAACAAGTG GACTCATGAGGAGGAACCCTCTAGA
P-31	AGGCCTCCGGACCTCTCCTCACGCTCACCCTGCGCGCTTCCGCGCG ACAGGCACAATTACCCGTATATGTCCCGACTCGCCCACAGTCTCCGC CTTCGGCCGGGTCATTCCCCCGACCGACCCGGCCCGGCCCACCCAT TTCCGGCCCGGCCGGCGTTTGAGGCCGACCGGTGACGGACACCCG AAGCCCTCGGAGCGCGCTCGGCATCAGCCCGGACGACGCTTGAGGC CACCTCGACCGCCGCCGGACGGCTTCATCCGAAGTGCCTCTGAACT GGTAAAACGAGCCGTGCTGGCAGCTCTCTGCACAACCAGGCAGAAC AAAACTTGAGCCCGTCCGACTCAACCGCATTGACGCGCCGCGTCCC CTCGTGCATCCTTGAGTGAGTTCCACTCAAGTAGTCAGCTGGAGGAA TTGACTCTAGA

Erklärung

Hiermit versichere ich, dass ich, Anna Glöckle, die vorliegende Arbeit selbstständig verfasst und keine anderen als die angegebenen Quellen und Hilfsmittel benutzt habe, dass alle Stellen der Arbeit, die wörtlich oder sinngemäß aus anderen Quellen übernommen wurden, als solche kenntlich gemacht und dass die Arbeit in gleicher oder ähnlicher Form noch keiner Prüfungsbehörde vorgelegt wurde.

Ort, Datum Unterschrift

Acknowledgements

An erster Stelle möchte ich mich bei Prof. Dr. TAM Gulder bedanken, für die Möglichkeit meine Doktorarbeit in seiner Arbeitsgruppe durchzuführen. Ich hatte das Glück an ein vielversprechendes Thema anschließen zu dürfen. Besonders möchte ich die ermutigende und aufmunternde Art im Umgang mit Doktoranten hervorheben, die es einem ermöglicht immer an den Erfolg des eigenen Projektes zu glauben.

Zudem möchte ich mich bei Prof. Dr. K. Lang bedanken, für das zweite Gutachten meiner Arbeit und Prof. Dr. M. Groll für die Übernahme des Vorsitzes der Prüfungskommission.

Ein besonderer Dank geht an Hülya, Katha und Paul. Danke für die Mühe euch diesen *** (Scheißdreck) wirklich durchzulesen, ich habe es geschrieben, ich kann mir gut vorstellen, wie anstrengend es war ihn zu korrigieren! Hülya danke ich besonders dafür, dass sie während meiner Zeit als Doktoranten zu meiner Adoptiv-Mutter geworden ist, auf mich aufgepasst hat, mich in die richtigen Bahnen gelenkt hat und mir immer mit Rat und Tat beiseite stand. Zudem war und ist es mir immer eine große Freude Freizeit mit dir zu verbringen! Katha, als mein S-Bahn Kumpel kennst du mich wohl besser als die meisten. Ehrlich gesagt gehen mir unsere morgendlichen Gespräche in der Bahn schon ein bisschen ab, bei wem soll ich mich denn jetzt auskotzen, um dann entspannt in den Tag starten zu können. Nachdem ich es jetzt wohl doch noch geschafft habe zu promovieren, weiß ich, dass auch du das ganz bald schaffen wirst! Paul, I really enjoyed working with you, your ‚no-worries' attitude is extremely supportive, especially when the third year of PhD kicks in. Bei den ‚alten' Gulder Mitglieder, allen voran Fritzi, Anna und Jana möchte ich mich für die vielen guten Zeiten bedanken, ihr standet mir immer mit einem offenen Ohr, fachlicher Expertise und zur Not auch Schnaps zur Seite. Bei allen aktuellen Gulders möchte ich mich für eine gute Zusammenarbeit bedanken, vor allem Mert und den Küken: Manu, TobiM und Julia, die mich herzlich in ihre Kaffeepause aufgenommen haben.

Vor allem in der finalen Phase hat mich Marie sehr unterstützt, vielen Dank, dass du auch noch die dümmsten Fragen zum Einreichen mit mir geklärt hast!

Ein besonderer Dank geht an meinen Traummann, Pascal, und meine Familie. Von euch allen habe ich während meiner Promotion und auch in der langen Zeit des Schreibens viel Kraft bekommen. Ihr habt mich nicht auf das Thema angesprochen, wenn ich es nicht wollte und habt mich unterstützt, als ich es brauchte.