Engineering of *Streptomyces albus* J1074 and *Streptomyces lividans* TK24 for Natural Products Production

Dissertation

zur Erlangung des Grades

des Doktors der Naturwissenschaften

der Naturwissenschaftlich-Technischen Fakultät

der Universität des Saarlandes

Von

Yousra Ahmed

Saarbrücken

2019

Tag des Kolloquiums:	09. August 2019
Dekan:	Prof. Dr. rer. nat. Guido Kickelbick
Berichterstatter:	Prof. Dr. Andriy Luzhetskyy
	Prof. Dr. Rolf Müller
Vorsitz:	Prof. Dr. Volkhard Helms
Akad. Mitarbeiter:	Dr. Stefan Boettcher

Diese Arbeit entstand unter der Anleitung von Prof. Dr. Andriy Luzhetskyy in der Fachrichtung Pharmazeutische Biotechnologie der Naturwissenschaftlich-Technischen Fakultät der Universität des Saarlandes von Mai 2013 bis August 2018.

Acknowledgments

Firstly, I would like to express my deep sense of gratitude and respectful regards to my supervisor Prof. Dr. Andriy Luzhetskyy for giving me the opportunity to work in his group on these exciting projects. I highly appreciate his invaluable guidance and continuous support throughout my PhD studies. I am also thankful to him for reading the manuscript of this thesis and for his recommendations to improve it.

I would like to thank Prof. Dr. Rolf Müller for his insightful comments and suggestions during the committee meetings and for being the second reviewer of this thesis.

I am forever grateful to Dr. Yuriy Rebets, my lab supervisor, for his patience, motivation, and for sparing his valuable time whenever I needed. His useful experimental advice, constructive suggestions and scientific inspiration made it possible for me to achieve my goal. I am indebted to him for his help, suggestions and correction of the entire thesis. It has been an absolute pleasure to work with and learn from him.

I am also very grateful to all my colleagues in AMEG group for their wonderful collaboration. They supported me impressively and were always willing to help me. Thanks a lot for the nice time we have spent together outside of work. I will cherish these memories forever.

I would like to express my deepest gratitude to my parents, Samaa and MhD. Saleh, my sister Souhaila and my brothers: Ossama, Ahmed, Mouwafak and Abderrahman, who always believe in me and encourage me to follow my dreams. Words fail to express my appreciation and respect to them for their moral and emotional support in all aspects of my life. I would like to thank my nieces and nephews that made me laugh from my heart in the hardest moments. To Mouwafak and Abderrahman I am so grateful to have you here beside me and for helping in whatever way you could during all challenging times. Thanks a lot for the very delicious meals and for making sure that I did not have to cook when I was very busy. I am also obliged to my other family members and friends who have supported me along the way.

I sincerely express my gratitude from my heart to my beloved husband, Ahmad, for his warm love, continued patience, and endless support during the happy and hard times. He has been a true and great supporter for the last two and half year of my studies, living every single minute of it. I thank God for enlightening my life with his presence. Without him this journey would not have been possible.

Yousra Ahmed Saarbrücken, April 01, 2019

Publications

Amin, R., Franz-Wachtel, M., Tiffert, Y., Heberer, M., Meky, M., **Ahmed, Y.**, Matthews, A., Krysenko, S., Jakobi, M., Hinder, M., et al. (2016) Post-translational Serine/Threonine Phosphorylation and Lysine Acetylation: A Novel Regulatory Aspect of the Global Nitrogen Response Regulator GlnR in *S. coelicolor* M145. Front. Mol. Biosci., 3, 38

Ahmed, Y., Rebets, Y., Tokovenko, B., Brötz, E., Luzhetskyy, A. 2017. Identification of butenolide regulatory system controlling secondary metabolism in *Streptomyces albus* J1074. Sci Rep 7:9784.

Busche, T., Tsolis, K.C., Koepff, J., Rebets Y., Ruckert, C., Hamed, M.B., Bleidt, A., Wiechert, W., Lopatniuk, M., **Ahmed, Y.** *et al*. Multi-Omics and Targeted Approaches to Determine the Role of Cellular Proteases in *Streptomyces* Protein Secretion. *Front Microbiol* 2018, 9:1174.

Rebets, Y., Tsolis, K.C., Guðmundsdóttir, E.E., Koepff, J., Wawiernia, B., Busche, T., Bleidt, A., Horbal, L., Myronovskyi, M., **Ahmed, Y.**, Wiechert, W., Rückert, C., Hamed, M.B., Bilyk, B., Anné, J., Friðjónsson, Ó., Kalinowski, J., Oldiges, M., Economou, A., Luzhetskyy, A. Characterization of Sigma Factor Genes in *Streptomyces lividans* TK24 Using a Genomic Library-Based Approach for Multiple Gene Deletions. *Frontiers in Microbiology* **9**, doi:10.3389/fmicb.2018.03033 (2018).

Ahmed, Y., Rebets, Y., Rodríguez Estévez, M., Luzhetskyy, A. Engineering of Streptomyces lividans for heterologous expression of secondary metabolites gene clusters (manuscript in preparation)

Conference Contributions

Amin, R., Tiffert, Y., Heberer, M., **Ahmed, Y.**, Krysenko, S., Jakobi, M., Hinder, M., et al. Post-translational Serine/Threonine Phosphorylation and Lysine Acetylation: A Novel Regulatory Aspect of the Global Nitrogen Response Regulator GlnR in *S. coelicolor* M145, the Annual Meeting of the part "Association for General and Applied Microbiology“ Tuebingen, Germany (VAAM). (Poster)

Ahmed, Y., Rebets, Y., Brötz, E. and Luzhetskyy, A. new strategy in awaking silent secondary metabolites biosynthesis gene clusters in *Streptomyces albus* J1074, VAAM workshop, Freiburg, Germany 2014. (Poster)

Ahmed, Y., Rebets, Y., Brötz, E. and Luzhetskyy, A. Regulatory elements silencing the secondary metabolism in Streptomyces albus J1074. 5th HIPS Symposium on Pharmaceutical Sciences Devoted to Infection Research, Saarbruecken, 2015. (Poster)

Abstract

Actinobacteria have remarkable chemical potential that is not explored due to low level of corresponding genes expression. In order to uncover this reservoir of natural products we developed a reporter-guided screening strategy combined with transposon mutagenesis. It was used to activate silent polycyclic tetramate macrolactam biosynthesis gene cluster in *Streptomyces albus* J1074. As result, the mutant with awaken secondary metabolism was obtained. Analysis of this strain led to identification of new regulatory system consisting of transcriptional regulator XNR_3174 and bacterial hormone-like compound butenolide. *XNR_3174* and butenolide biosynthesis genes orthologues are present in the genomes of different *Streptomyces*. The identified regulatory system comprises a new condition-depended cascade controlling secondary metabolism in Actinobacteria. We also developed new host strains for heterologous production of natural products by deleting 11 endogenous secondary metabolite gene clusters from chromosome of *S. lividans* and introducing up to 2 sites for integration of foreign DNA. When expressing three heterologous gene clusters the generated hosts have shown better performance than the parental strain. *S. lividans* TK24 was also improved as a host for heterologous protein production by deleting a set of proteases encoding genes. The developed strains represent a step forward to a better panel of organisms for bioprospecting and genome mining of novel natural products.

Zusammenfassung

Actinobakterien besitzen ungeahntes chemisches Potenzial, das aufgrund niedriger Exprimierung entsprechender Gene nicht erforscht werden kann. Um Dieses aufzudecken, wurde eine Kombination aus reportergeführter Screening-Strategie und Transposon Mutagene entwickelt. Die Verwendung dieser Strategie führte zur Aktivierung des polyzyklischen Tetramat-Makrolaktam Gen-Clusters in *Streptomyces albus* J1074. Der erhaltene Stamm weißt eine aktivierte Produktion von Sekundärmetaboliten auf und Analysen führten zur Identifizierung eines Regulationssystems, bestehend aus dem transkriptionellen Regulator XNR_3174 und dem hormonähnlichen Naturstoff Butenolid. Orthologe von *XNR_3174* und der Butenolid Gene findet man in Genomen verschiedener Streptomyceten. Das identifizierte System umfasst die neue „condition-dependent"-Kaskade, die den Sekundärstoffwechsel in Actinobakterien steuert. Zusätzlich entwickelten wir, durch das Entfernen von 11 endogenen Gen-Clustern aus dem Genom von *S. lividans* und dem Einfügen von zwei DNA Integrationsstellen, neue Wirtsstämme für die heterologe Produktion von Naturstoffen. Bei der heterologen Exprimierung von drei Gen-Clustern zeigten die optimierten Wirte bessere Resultate als der ursprüngliche Stamm. Durch das Entfernen der Proteasegene wurde *S. lividans* TK24 als Wirt für die Proteinproduktion verbessert. Die entwickelten Stämme vereinfachen die Bioprospektion und die Entdeckung neuer Naturstoffe unter Verwendung des „Genome-Mining" Ansatzes.

Table of Contents

Table of Contents

1. Introduction

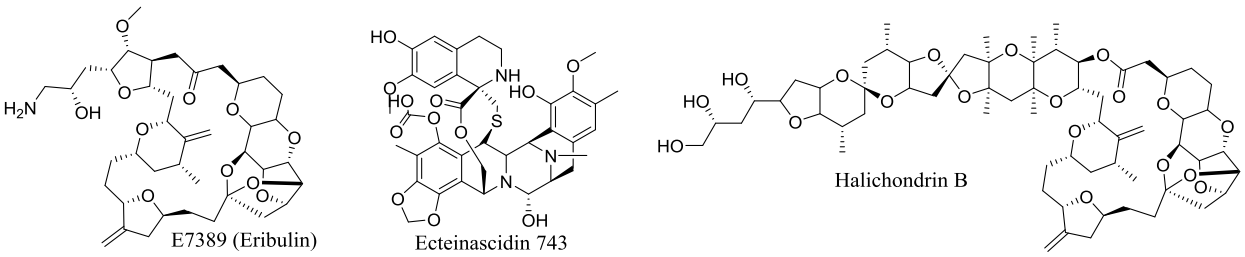
1.1. Importance of Natural Products in Medicines Development

Natural products (NPs) represent a large group of structurally diverse organic molecules produced by plants, bacteria, fungi and marine animals (Fig. 1). These compounds often exhibit biological activities (such as antimicrobials, anticancer, antiviral, immunosuppressive, etc.) and therefore have been used in different applications such as human medicine, veterinary medicine and agriculture [1-4]. As of the end of 2013, all together 1453 new molecular entities have been approved for use as therapeutics by the US Food and Drug Administration (FDA) [5]. About 40 % of them are NPs, their derivatives or synthetic mimetics related to NPs that emphasize the essential role these compounds play in modern medicine [1,5,6]. Since 2000 more than three quarter (77%) of approved antibiotics are developed from the NPs, 100% of which are derived from microorganisms [6,7].

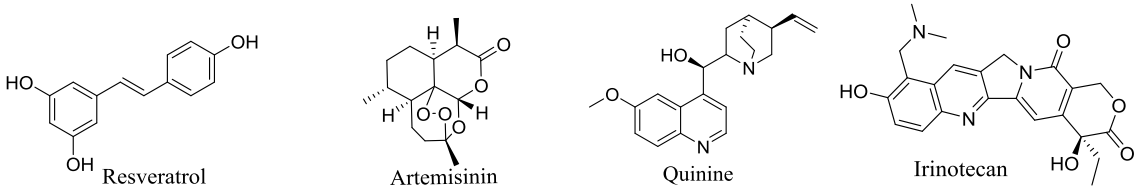
The discovery of penicillin by Alexander Fleming in 1928 [8] and its development into a clinical use by Florey and Chain in the 1940s [9] brought microbial NPs into spotlight. As a result, the so called "golden era" of antibiotic discovery (1940s-1970s) started. During this period an intensive screening of microorganism resulted in discovery of a wealth of bioactive compounds including commonly approved antibacterial medicines such as streptomycin, erythromycin, tetracycline, vancomycin and chemotherapeutic drugs such as doxorubicin [10,11]. These compounds were the first effective treatment for life-threatening microbial diseases that were responsible for the major infections in human history. In spite of the high success rate of these compounds, with the time, the efficiency of the natural products discovery program decreased due to the rediscovery of already known compounds, difficulties in isolating new producers and therefore new antibiotics and low production level [12]. Instead alternative strategies such as high-throughput screening of synthetic compound libraries and fragment-based design were gained popularity to counter the slowdown in NPs discovery [13,14]. Unfortunately, these strategies had limited success so far due to the fact that not all successful synthetic compounds are able to penetrate the bacterial cell wall [15]. As an example, in the period between 1970s and 1980s, only one group of fully synthetic antibiotics, fluoroquinolones, was approved for clinical use [16]. In fact, microbial products or synthetic analogues are outperforming synthetic compounds in effectiveness. It is worth mentioning that since the turn of the century several new microbial products or synthetic analogues have been approved as medicines to treat serious diseases such as daptomycin (against

infections caused by methicillin-resistant *Staphylococcus aureus* (MRSA)) and fidaxomicin (against gastrointestinal infection by *Clostridium difficile*) [17,18].

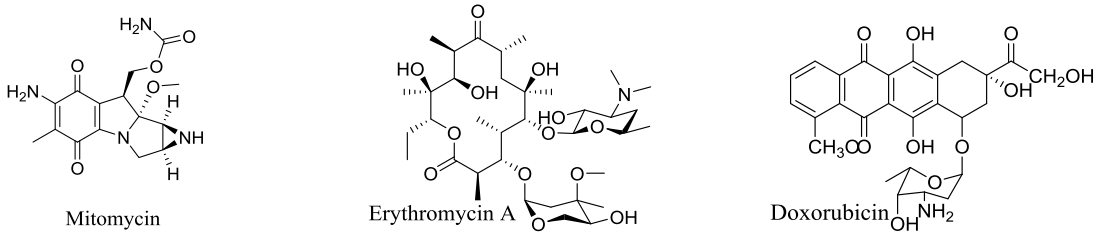
A



B



C



D

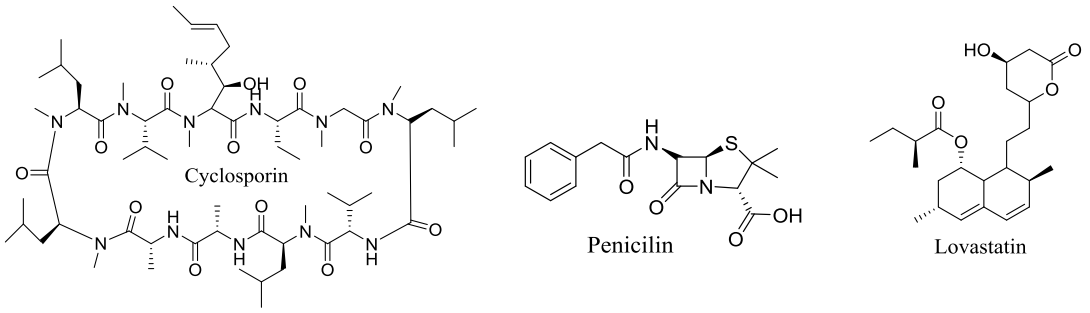


Figure 1. NPs-derived from different sources: marine organisms-derived NPs **(A)**, plant-derived NPs **(B)**, bacterial-derived NPs **(C)** and fungal-derived NPs **(D)**.

The lack of new antibiotics coupled with the rise of multidrug resistant microorganisms caused infections has created a serious global public health threat [19-21]. In addition, the evolution and spread of the antibiotic resistance genes has further exacerbated the situation [22,23]. The United States Centers for Disease Control and Prevention reported that more than 2 million people get antibiotic resistant infections every year in the U.S. alone, resulting in close to 23,000 deaths [24]. According to a recent study, if current situation continues, by the year 2050 10 million individuals will die per year from antimicrobial resistant infections [25]. In light of these reports, it is clear that there is an urgent need to develop new antibiotics with novel mode of action. Microbial NPs have thousands of years of evolution behind their structure optimization in order to become efficient killers, which make them a preferred choice for further exploitation as medicines. Between 2008 and 2013, 25 NP and NP-derived drugs were approved including 5 being classified as NPs, 10 as semi-synthetic NPs (SS NPs) and 10 as NP-derived drugs [26]. 14 of them have being isolated from fungus and bacteria, among them 7 were derived from actinomycetes [26] (Fig. 2).

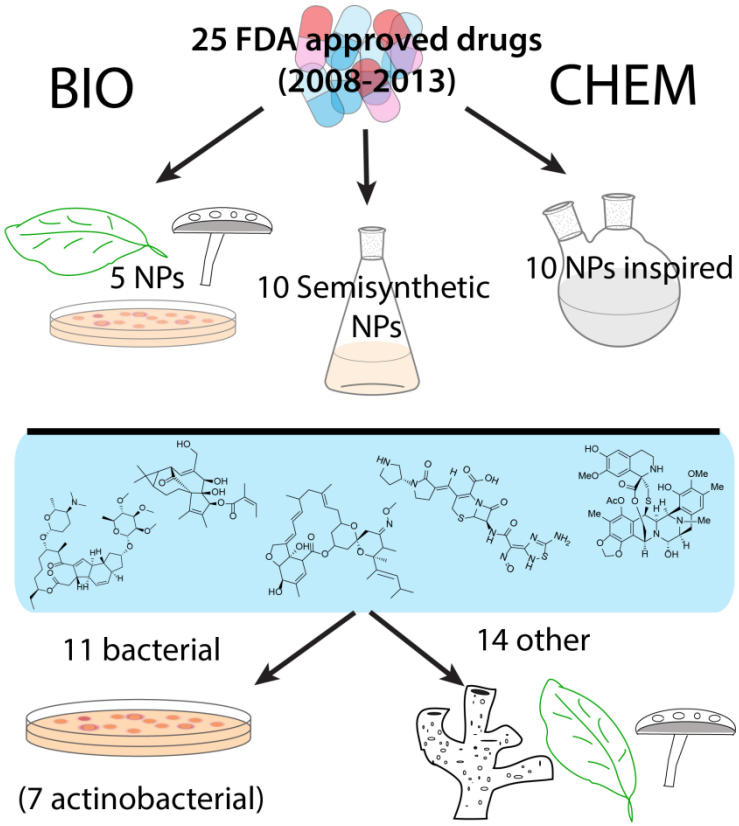


Figure 2. Classification of 25 NP and NP-derived drugs approved between 2008 and 2013. Five are classified as NPs, 10 as semi-synthetic NP (SS NPs) and 10 as NP-derived drugs [26]. Eleven of them have being isolated from bacteria, among which 7 were derived from actinomycetes [26].

1.2. Natural Products of Actinobacteria

The phylum Actinobacteria includes Gram-positive mycelial bacteria with high genomic G+C content. They are typically found in terrestrial and aquatic ecosystems. These bacteria, mainly of the *Streptomyces* genus, are one of the major sources of microbial natural products [27]. They produce about three quarters of all known marketed antibiotics of microbial origin [28]. Moreover, there is a significant structural diversity among NPs produced by *Streptomyces*, which is reflected in antibacterial, antifungal, antiviral, antihypertensive, immune suppressive and antitumor activities [1,29]. The very first active compounds isolated from *Streptomyces* were actinomycin and streptomycin in 1940 and 1943, respectively [30,31]. Actinomycin is the first anticancer drug of NPs origin and is still in clinical usage. The aminoglycoside antibiotic streptomycin has remarkable antibacterial activity and have been used to treat *Mycobacterium tuberculosis* infections for a long time. Subsequently, a great number of bioactive compounds were discovered including antibiotics (erythromycin, vancomycin, and daptomycin), antifungal (nystatin), antitumor (bleomycin) and immunosuppressive (FK506) to mention but a few [10,32,33]. Thus, based on existing knowledge regarding actionobacterial NPs variety one can predict high chances of finding new biological active compounds produced by these fascinating bacteria. Recently, several new NPs have been isolated from *Streptomyces* such as metatricycloene [34] and streptovaricin which has potent activity against methicillin-resistant *Staphylococcus aureus* (MRSA) [35].

Despite the decades of actinobacterial NPs exploitation only in 2002 with *S. coelicolor* genome sequence published it became obvious that these bacteria are hiding enormous chemical potential encrypted within their chromosomes [36]. This understanding was further enhanced by genome sequencing of several other actinomycetes strains such as *Saccharopolyspora erythraea* [37], *S. avermitilis* [38], *S. griseus* [39], and *S. lividans* [40]. The genes responsible for secondary metabolite biosynthesis are organized into clusters. The typical clusters include structural genes, coding for biosynthetic enzymes, as well as regulatory and resistance genes. The genomes of Actinobacteria contain more than 20 biosynthetic gene clusters [41]. The expression of only small fraction of these clusters results in detectable quantities of respective compound production under standard laboratory conditions. Therefore, *Streptomyces* comprises a marvelous treasure of NPs, making these species the perfect objects for NPs discovery.

With the development of new generation sequencing technologies, the genomes information of Actinobacteria became accessible resulting in accumulation of enormous amount of genomic data bearing the potential NP biosynthesis gene clusters. This raised the need for bioinformatics

tools for identification and classification of NPs biosynthetic genes, and structure prediction of corresponding compounds [42,43]. As a result, several software have been developed to help identifying putative biosynthetic gene clusters such as antibiotics and Secondary Metabolites Analysis Shell (antiSMASH) [44], Cluster Scanner (ClustScan) [45], structure based sequence analysis of polyketide synthases (SBSPKS) [46] and NP searcher [47]. The genome sequencing and bioinformatics tools have actually made an important step forwards in exploitation of this untapped reservoir of metabolites. However, in order to turn this genomics-based understanding of actinobacterial NPs potential into biologically active compounds the development of methodologies and tools to manipulate the silent or poorly expressed gene clusters became of high demand.

1.3. Activation of Silent Biosynthetic Gene Clusters

The genome sequencing data highlighted the difference between the number of secondary metabolite gene clusters identified by bioinformatic tools and the actual number of chemically characterized compounds produced by a microorganism. The so called "silent" gene clusters are usually expressed at very low level or not at all under laboratory growth conditions. Activation of these clusters could unlock the corresponding biosynthetic pathways they encode, which will lead to discovery of novel compounds with medical and biotechnological interest. Therefore, several strategies have been intensively developed and applied to activate the silent gene clusters [48,49]. They can be grouped into pleiotropic and pathway-specific approaches (Fig. 3; Fig. 7). Pleiotropic approaches, such as OSMAC (One Strain Many Compounds), engineering of transcriptional and translational machinery, manipulation of global regulators, etc., alter the expression of different genes simultaneously and thus introduce global changes in the regulation of secondary metabolism pathways. The main advantage of these strategies is that the knowledge of a gene cluster type and how they are regulated are not mandatory. Thus, pleiotropic approaches allow greater throughput than pathway-specific one. However, the output is more complex that makes identification and isolation of new produced compounds more difficult. Conversely, pathway-specific approaches, such as manipulation of pathway-specific regulators, refactoring the gene cluster often combined with heterologous expression, are based on modifying transcription and translation regulatory elements of the cluster of interest. These methods require sufficient information about overall genetic organization and regulatory components of the cluster and availability of efficient tools for genetic manipulations. As a standalone approach the reporter-guided screening in combination with either pleiotropic or

pathway-specific methods was developed in order to detect the changes in transcription of NPs gene clusters (Fig. 3; Fig. 10).

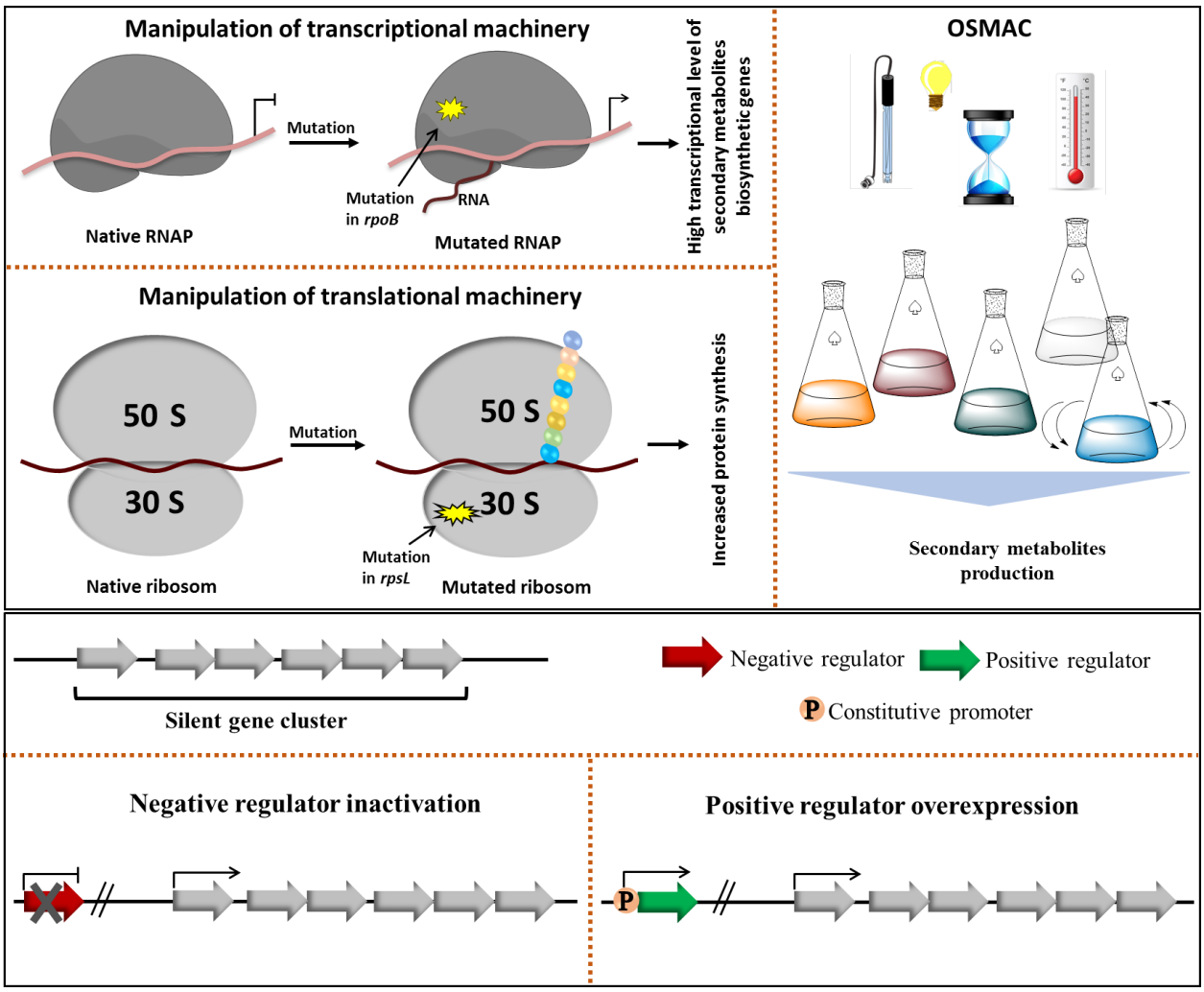


Figure 3. Different pleiotropic approaches used for the activation of silent gene clusters in Actinobacteria.

1.2.1. OSMAC: One Strain Many Compounds

The regulation of secondary metabolites biosynthesis occurs in response to internal and external signals which are almost always correlated with changes in the growth conditions [50]. The approach named OSMAC introduced by Bode and co-workers was used to trigger the potential of one single microbe to produce many bioactive compounds by manipulating the growth conditions [51]. Changing the cultivation parameters (such as media composition, aeration, pH and addition of enzyme inhibitors) will stimulate the expression of silent gene clusters and thus production of so far unknown metabolites. There are many successful examples of application of

the OSMAC approach showing that simple alteration in cultivation condition leads to different metabolites profiles of one strain. Rateb and co-workers have studied the effect of different carbon sources in culture media on the natural products biosynthesis in *Streptomyces* strain C34 [52]. Chaxamycins A and B (Fig. 4) were obtained when the strain was grown in ISP2 medium containing glucose as a carbon source. However, substitution of glucose with glycerol caused accumulation of two new derivatives chaxamycins C and D (Fig. 4). Chaxamycin D showed antibacterial activity against *Staphylococcus aureus* strains, including a panel of methicillin resistant clinical isolates with MIC value lower than 1µg/ml.

In recent study the marine-derived *Streptomyces* sp. YB104 was grown in seven different media and screened for bioactive natural products [53]. Analysis of metabolic profiles from different conditions leads to detection of new compounds in the extract of the strain when cultured in ACM medium but not others. One of the corresponding compounds inthomycin B was purified and its structure was determined. Inthomycin B possesses antimicrobial and herbicidal activities. The ACM is quite simple, inexpensive and chemically defined medium and mainly consists of soluble starch and mineral element. Interestingly, even under non-optimized conditions *Streptomyces* sp. YB104 was shown to produce inthomycin B at high yield (25 mg/l). This study indicates that production of natural products by *Streptomyces* strains is highly dependent on the cultivation media.

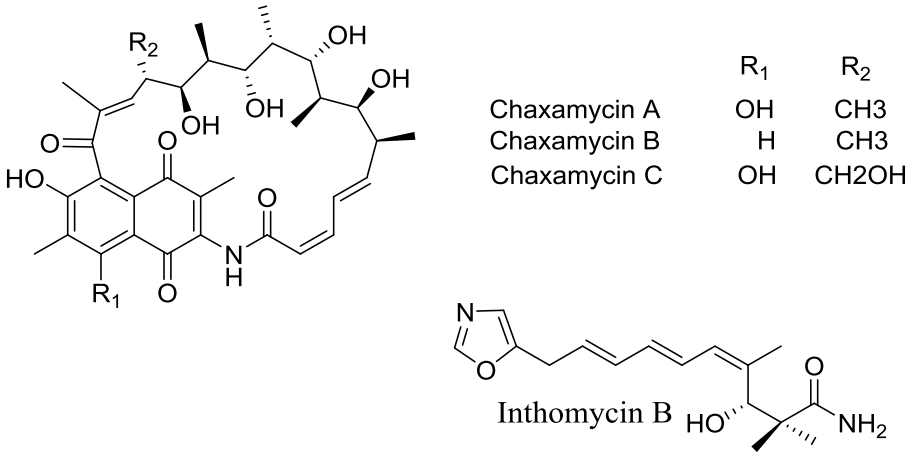


Figure 4. Examples of chemical structures of compounds identified using OSMAC approach.

The influence of the light on secondary metabolism production was observed in *S. coelicolor* mutant, lacking actinorhodin and undecylprodigiosin production [54]. The strain was found to accumulate yellow pigments when it was exposed to the light compared to the one grown in the dark condition. The yellow fraction was containing several carotenoid compounds including

isorenieratene and β-carotene. This observation indicates that the carotenoid production is controlled by regulatory system for light signal transduction. Evidence showed that the carotenoid biosynthesis gene cluster (*crt*) is flanked with *lit* (light-induced transcription) genes containing *litS* and *litR* genes coding for ECF sigma factor (LitS) and MerR-type transcriptional regulator (LitR) respectively. LitS is directing the transcription initiation of *crt* gene cluster whose activity in turn controlled by LitR transcriptional regulator with a photosensory function.

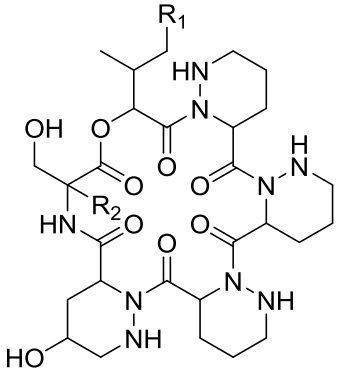
The OSMAC approach is relatively simple, versatile and low-cost tool to elicit NPs potential of the strain by changing the composition of medium and growth parameters. Moreover, this approach can be easily applied to microorganisms that are not amenable to genetic manipulations. Nevertheless, it is difficult to predict which conditions are appropriate to grow the tested microorganism to achieve the production of new NPs. This is why multiple growth parameters should be probed.

1.2.2. Manipulation of Transcriptional and Translational Machineries

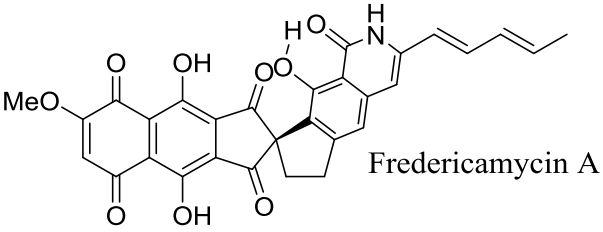
Another example of pleiotropic approach includes modulation of the transcriptional and translational machineries by generating mutations in a *rpoB* gene (encoding β-subunit of RNA polymerase (RNAP)), a *rpsL* gene (30S ribosomal protein S12) and other ribosomal proteins genes, which confer resistance to antibiotics [55]. These point mutations are selected by growing the bacteria in presence of different antibiotics such as streptomycin, rifampicin and gentamycin. Screenings for antibiotic resistant mutants is a simple method to activate silent gene clusters. The role of a translational machinery in controlling the antibiotics production was for the first time demonstrated in a case of streptomycin resistance *S. lividans* strain [56]. The mutant was found to produce actinorhodin in contrast to the wild type in which the corresponding gene cluster is silent. Genetic analysis of this mutant showed that a reason for activation of actinorhodin biosynthesis is a point mutation in the *rpsL* gene. This phenomenon was intensively studied by Ochi and co-workers and widely used for new NPs discovery [57]. In this way eight novel compounds (named piperidamycins) produced by *S. mauvecolor* mutants were identified and characterized (Fig. 5). Production of these compounds occurred after a selection step using streptomycin and rifampicin antibiotics for generating spontaneous *rpsL* or *rpoB* point mutations, respectively.

Using the same strategy, the cryptic fredericamycin A biosynthetic gene cluster was activated when marine *S. somaliensisn* SCS10 ZHH was subjected to high concentration of rifampicin. Upon inspection R444H mutation in the β-subunit of RNA polymerase was found [58]. Unlike the

wild type, the mutant strain produced a dark brown pigment which was extracted, purified and identified as an anticancer drug fredericamycin A (Fig. 5).



	R_1	R_2
Piperidamycin A	H	H
Piperidamycin D	H	CH_3
Piperidamycin F	CH_3	CH_3



Fredericamycin A

Figure 5. Examples of chemical structures of compounds identified by manipulation of the transcriptional and translational machineries.

Generally, mutations in RNAP have an effect on its promoters binding affinity, thereby may enhance the transcription of silent gene clusters involved in the biosynthesis of small molecules. This can be correlated to the bacterial alarmone guanosine tetraphosphate (ppGpp). The ppGpp bind to RNAP and modify its ability to initiate the transcription at different promoter regions [59,60]. It was shown that ppGpp has positive effect on antibiotic production in *S. coelicolor* A3(2) [61]. Thus, the mutated RNAP may mimic the ppGpp-bound conformation resulting in transcriptional activation of silent gene clusters [61]. In the case of the *rpsL* mutations, studies have shown that the protein synthesis in mutants increases during the stationary phase when the most biosynthetic gene clusters are expressed, thus causing the enhanced production of corresponding compound [62].

1.2.3. Manipulation of Global and Pathway-Specific Regulators

The transcription of secondary metabolite gene clusters in Actinobacteria is usually under the tight control of a transcriptional regulatory system that alter the expression level of these clusters in response to the extracellular environmental and intracellular physiological changes [50,63,64]. These regulators either enhance (activators) or block (repressors) the expression of biosynthetic gene clusters. The regulators which are located outside gene clusters and have global effect on different metabolic pathways and often morphological processes are called global regulators,

whereas those located within the gene clusters and directly control their transcription are known as pathway-specific regulators. Therefore, identification and subsequent manipulation with the components of this intracellular regulatory network is a basis of approach, which can be exploited to activate or enhance the production level of secondary metabolites.

a) Manipulation of Global Regulators

A set of global transcriptional regulators have been shown to alter the expression of secondary metabolite gene clusters and their manipulation enabled the identification of novel secondary metabolites [65,66]. There are several highly conserved global regulators found in many species of *Streptomyces*, such as AdpA, WblA and *bldA* controlling more than one pathway [67-69].

WblA (WhiB-like protein) is a global regulator that has dual function in antibiotic production control, as a repressor or an activator, depending on a species [70,71]. The disruption of the *wblA* gene affects morphological differentiation and antibiotic production in *S. ansochromogenes* [69]. In the WblA deficient mutant (ΔwblA) the production of nikkomycin was completely eliminated whereas production of two novel tylosin analogues was activated (Fig. 6). Both compounds showed much higher activity (at MIC around 7 µg/ml) against *Streptococcus pneumoniae* than tylosin (MIC >100 µg/ml). Identified compounds are promising new derivatives of tylosin.

AdpA (A-factor-dependent protein A) is another important pleiotropic regulator, which can also act as an activator or a repressor depending on the regulated genes [72,73]. AdpA controls the expression of a large set of genes including those involved in morphological differentiation and secondary metabolites production [74]. Therefore, the genetic manipulation of *adpA* was used for activation of silent oviedomycin biosynthesis gene cluster in *S. ansochromogenes* [75]. Pan, *et al.*, showed that AdpA positively affects nikkomycin production by activating the expression of the pathway-specific regulator gene *sanG* [76]. The disruption of the *adpA* gene completely abolished nikkomycin production but led to the discovery of oviedomycin which belongs to the angucycline family [75] (Fig. 6). AdpA blocks the biosynthesis of oviedomycin by repressing the positive regulator genes (*ovmZ* and *ovmW*) within the corresponding *ovm*-gene cluster.

The presented above examples are only covering few of many other global regulatory genes found in Streptomycetes. The highly complex lifestyle of these bacteria caused development of a complex regulatory network. Often the secondary metabolism and morphological development are tightly related and controlled by the same regulators [50]. Vice versa, some global regulatory factors are strictly dedicated either to morphological processes or NPs biosynthesis [77].

Manipulating such regulatory cascades is a straightforward and simple way to induce global changes in the NPs metabolic profile of producing strain without knowledge about actual

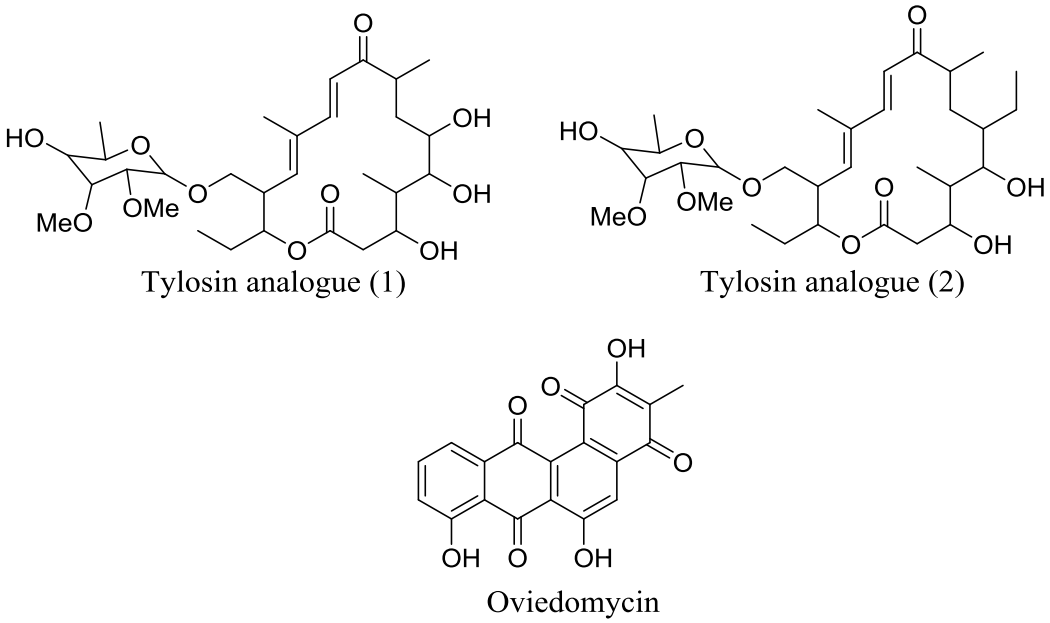


Figure 6. Examples of chemical structures of compounds identified by manipulation of global regulators.

environmental triggers. In this the described approach is somewhat similar to OSMAC. On other hand, the clear understanding of the regulatory network components and their interactions is prerequisite of success of this approach, since often the morphology, growth, sporulation and biomass accumulation might be affected as well, that is often undesired side effects.

b) Manipulation of Pathway-Specific Regulators

Many NPs biosynthetic gene clusters, but not all, contain dedicated transcriptional factor genes controlling the expression of structural genes [50]. The overexpression of genes coding for pathway-specific activators is often performed by replacing the native promoter with constitutive or inducible promoters or by cloning the regulatory genes in high copy number cloning vectors (Fig. 7).

The constitutive expression of the gene coding for pathway specific activator ChaI within the silent type II PKS gene cluster (*cha*) resulted in the production of new compounds by *S. chattanoogensis* L10 [78]. Based on *in silico* analysis, the *cha* gene cluster was found to contain three genes *chaK, chaK1* and *chaI* that are putatively encoding pathway specific regulators. The ChaI protein shows a high identity to JadR1 in *S. venezeulae* and LanI in *S. cyanogenus*, both of which are known as positive regulators of angucyclines biosynthesis [79,80]. The corresponding

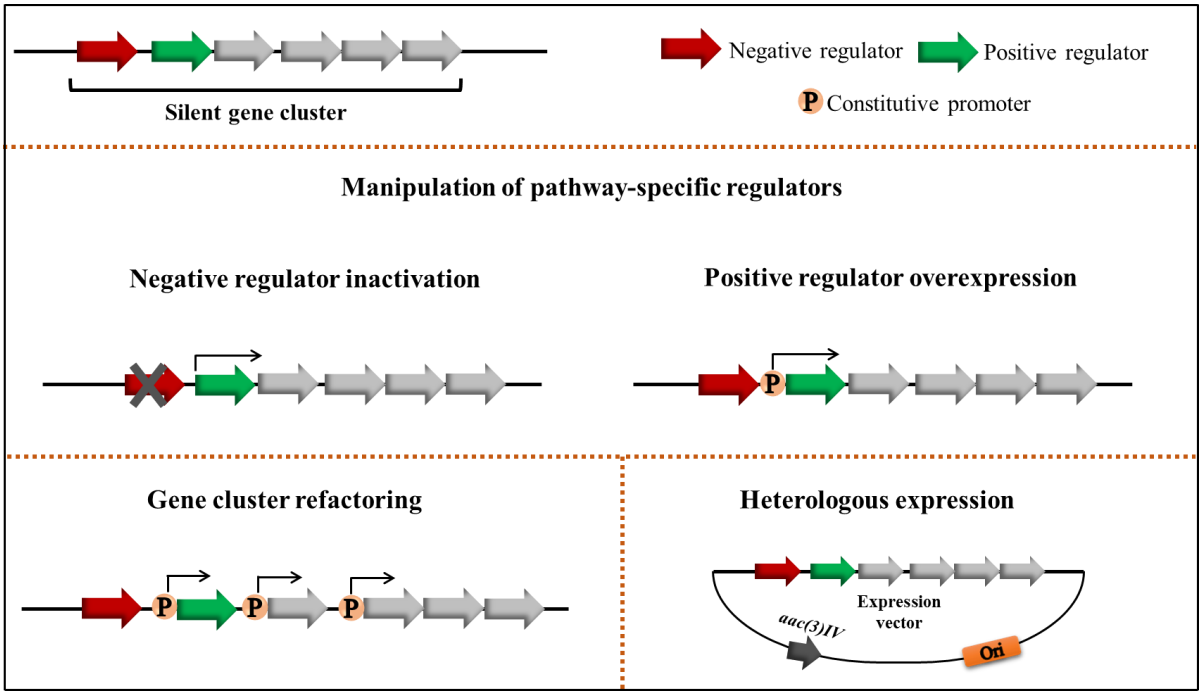


Figure 7. Pathway-specific approaches used for activation of silent gene clusters in Actinobacteria.

gene was cloned under the control of constitutive *ermE** promoter. As a result, two novel natural products named chattamycins A and B were identified (Fig. 8). The new members of angucycline family showed good antitumor and weak antibacterial activity.

Wang and co-workers used similar approach to activate the silent ansamycin biosynthesis gene cluster (*ast*) in *Streptomyces* sp. XZQH13 [81]. Genetic analysis revealed the presence of *astGI* gene encoding a putative positive regulator of LAL family. The overexpression of the LAL family transcription factors has been previously shown to activate the polyketide biosynthesis in Streptomycetes [82,83]. Therefore, the *astGI* gene was cloned into an integrative vector under the control of the *ermE** promoter and introduced into *Streptomyces* sp. XZQH13. As a result, two new peaks were obtained which were identified as hydroxymycotrienin and thiazinotrienomycin (Fig. 8).

Deletion of pathway-specific repressor genes has also been shown to influence the expression of natural product gene clusters [84,85]. For instance, the pathway-specific regulatory gene (*scbR2*) in *S. coelicolor* is located in the silent type I polyketide synthase gene cluster (*cpk*). ScbR2 is a homologous to ScbR which is the γ-butyrolactone receptor protein and negatively control the expression of pathway specific activator *cpkO* within the *cpk* gene cluster [86,87]. To study the role of ScbR2 in secondary metabolite production, Gottelt *et al.* generated a *scbR2* deletion mutant, which resulted in the production of two novel compounds: antibiotic cryptic polyketide (abCPK)

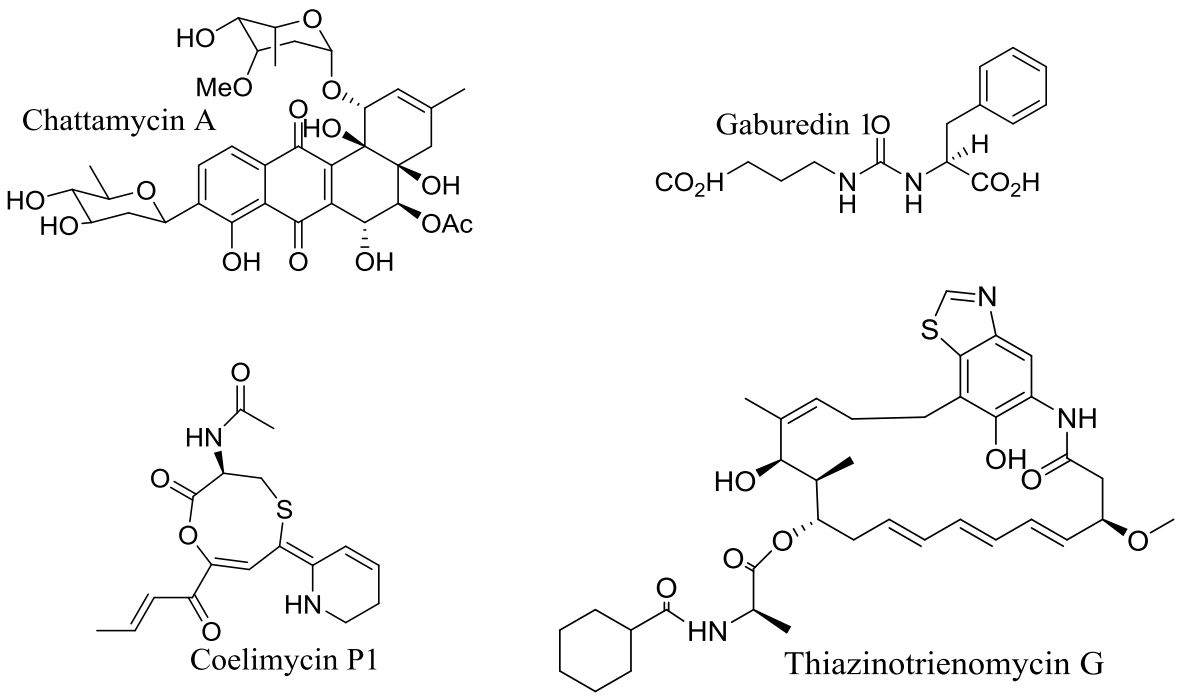


Figure 8. Examples of compounds obtained by manipulation of specific- pathways regulators.

and yellow cryptic polyketide (yCPK) [88]. The later compound was structurally elucidated by Challis and coworkers and named coelimycin P1 (Fig. 8) [89]. Similarly, the inactivation of the *arpA*-like putative transcriptional repressor coding gene *(gbnR* gene) in *Streptomyces venezuelae* ATCC 10712 lead to activation of the *gbn* gene cluster [90]. As a result, a new family of γ-aminobutyrate (GABA)-derived urea, the gaburedin, was discovered (Fig. 8).

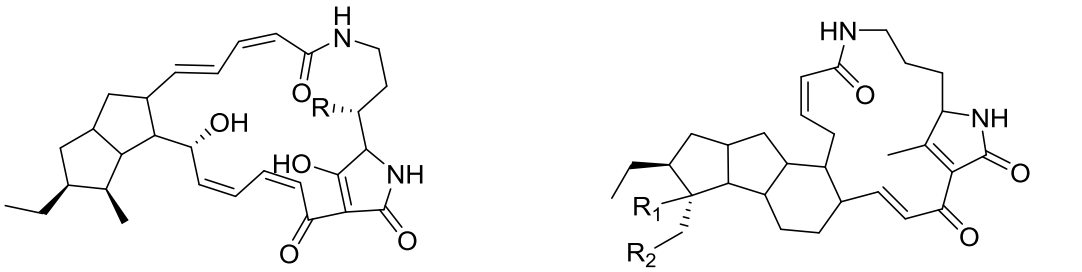
In conclusion, the manipulations of pathway-specific regulatory genes result in activation of the corresponding gene clusters, which could potentially result in production of new NP(s). On the other hand, A clear understanding of a function of the cluster-based transcriptional factor is required for successful application of this method, since often such regulators are controlling only some of cluster functions, like transport of final product, and are not affecting the structural genes activity. Moreover, the genetic manipulations of the corresponding biosynthetic pathways should be also amenable.

1.2.4. Gene Cluster Refactoring

With the recent development of synthetic biology tools, the activation of NP biosynthetic pathways could be achieved by complete refactoring of corresponding gene cluster. Refactoring the cluster of interest usually involves the exchange of the native promoters with constitutive

promoters to decouple the expression of biosynthetic genes from a complex native regulatory network [91] (Fig. 7). This approach is often combined with the expression of refactored gene cluster in a heterologous host when the native producer is not genetically tractable.

This strategy was used to activate several polycyclic tetramate macrolactam (PTM) gene clusters in *Streptomyces* species. The PTM biosynthetic gene clusters are conserved and widely distributed among phylogenetically diverse bacteria; however, most of their products are still



6-epi-alteramide A R= OH
6-epi-alteramide B R=H

Pactamide A R1= H, R2= H
Pactamide F R1= OH, R2= Cl

Figure 9. Examples of chemical structures of compounds that produced after refactoring their corresponding gene clusters.

uncharacterized [92,93]. PTMs originated from a conserved PKS-NRPS which resembles iterative enzymes typically found in fungi [92]. Utilizing promoter engineering strategy, Olano and co-workers [94] succeeded to activate the PTM gene cluster in *S. albus* J1074 by introducing *ermE** promoter upstream of two genes coding for PKS-NRPS pathway and hydroxylase enzyme. Two new compounds were identified as 6-*epi*-alteramide A and 6-*epi*-alteramide B (Fig. 9). Following similar strategy, the putative PTM gene cluster from marine-derived *S. pactum* SCSIO 02999 was successfully activated and six new PTMs, pactamides A-F were obtained (Fig. 9) [95]. These compounds displayed potent cytotoxic activity against four cancer cell lines (SF-268, MCF-7, NCI-H460 and Hep-G2). The silent PTM gene cluster (SGR810-815) from *S. griseus* was reconstructed and activated by using plug-and-play synthetic biology strategy [96]. The six-gene operon which encodes putative PKS-NRPS and processing enzymes was reconstructed by decoupling and placing each gene under the control of a constitutive promoter. Expression of this synthetic construct in *S. lividans* resulted in production of three novel PTMs.

Although PTM gene clusters described above share the same genetic organization and have high degree of sequence similarity, their activation resulted in production of structurally diverse PTMs. This observation suggests that manipulating other silent PTM gene clusters in phylogenetically distinct bacteria provides an opportunity to discover more bioactive PTMs [92,93].

Biosynthetic pathway refactoring offers a great potential and possibilities for flexible and targeted control on metabolites production and structural diversification. With the rapid advance in synthetic biology of Actinobacteria, various new synthetic BioBricks with defined performance such as promoters, RBSs, terminators, and DNA recombination and assembly tools can be employed to diverse genetic designs.

1.2.5. Reporter-Guided Mutant Selection (RGMS)

The aforementioned approaches have been developed and used to activate silent gene clusters in Actinobacteria. Their success often relies on analytical instrumental methods in order to identify clones that are producing the desired compound(s). However, the use of HPLC, GC-MS and LC-MS are time consuming and costly. In order to simplify the screening procedure several strategies were developed such as utilization of biosensor and reporter-guided mutant selection **(**RGMS) [97,98]. When combined with the mutagenesis, either UV and chemical or transposon based, both these strategies represent standalone approach for silent gene cluster activation.

The principle of RGMS is based on direct monitoring of target gene promoter activity with the assumption that the production level of compound of interest directly depends on the level of expression of biosynthetic genes. To achieve this promoter from the biosynthetic gene cluster of interest is fused to the reporter gene allowing simple identification of clones with activated or enhanced transcription (Fig. 10).

For the first time RGMS in actinomycetes was applied by Xiang and co-workers to increase clavulanic acid (CA) production in *S. clavuligerus* [99]. The authors used single reporter strategy based on kanamycin resistance reporter gene (*neo*) placed under control of promoter of CA biosynthesis transcriptional activator gene *ccaR*. The recombinant *S. clavuligerus* carrying respective plasmid was subjected to chemical mutagenesis and plated on medium supplemented with high concentration of kanamycin. In such way colonies that survived were assumed to have high level of *ccaR* and thus structural *cca* genes transcription that supposed to be reflected in high yield of clavulanic acid. However, only 51 % of mutants showed improvement in CA titer compared to the parental strain.

The appearance of false positive colonies most probably is caused by development of kanamycin resistance mediated by other than *neo* mechanism. To reduce the false positive rate the system was modified by including the second reporter *xylE*. The kanamycin resistance phenotype was used for initial selection of the mutants while *xylE* allowed distinguishing between spontaneous kanamycin resistant and *ccaR* overexpressing colonies. The double reporter approach resulted in

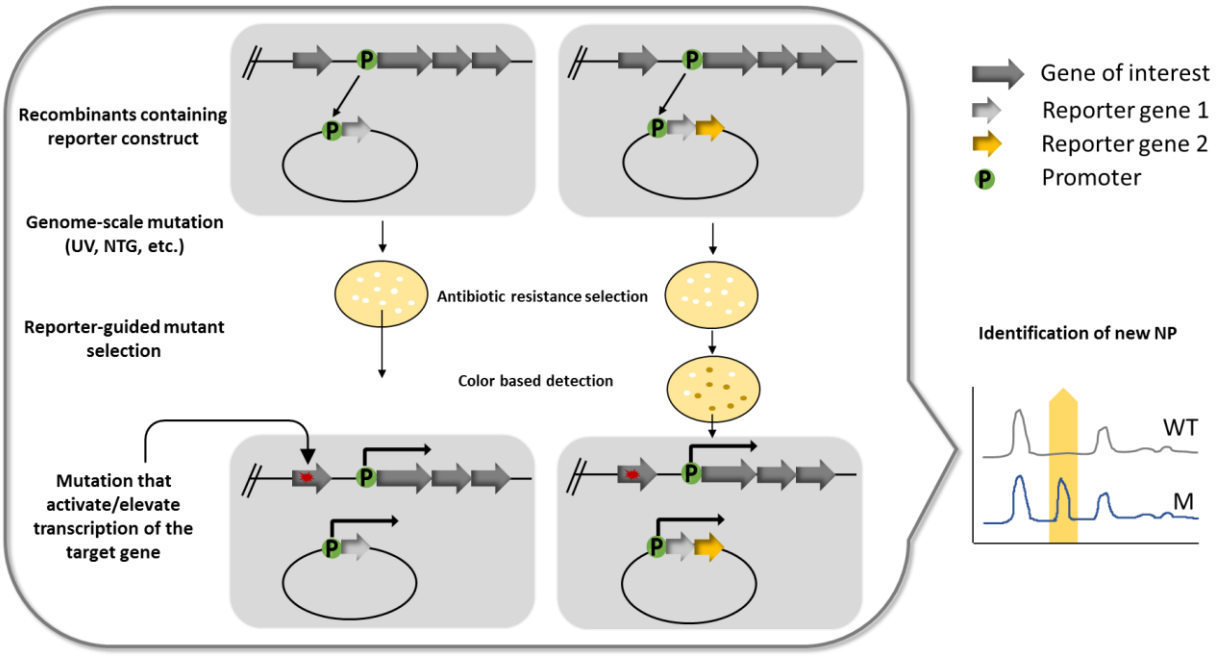


Figure 10. Schematic representation of the principle of reporter-guided mutant selection method (RGMS). Modified from [99].

90 % success rate. It was found to be more efficient than the single reporter system since it resulted in much fewer false-positive hits. The same double reporter cassette was successfully applied to activate the silent gaudimycin (*pga*) gene cluster in *Streptomyces* sp. PGA64 and the poorly expressed jadomycin (*jad*) gene cluster in *S. venezuelae* ISP5230 (Fig. 11) [98].

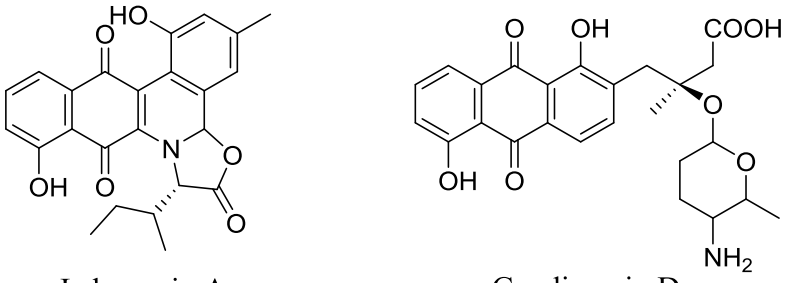


Figure 11. Representative compounds that are produced using RGMS approach.

In conclusion, this approach is very efficient since the expression level of a biosynthetic gene cluster is coupled to the production level of the corresponding metabolite, thus excluding the non-producers from further analysis. On other hand, the random mutagenesis used in these examples introduces mutations that are difficult to identify. To overcome this obstacle transposon mutagenesis can be utilized.

1.3. Heterologous Expression of Secondary Metabolite Gene Clusters

The majority of the sequenced gene clusters are in a dormant state under laboratory conditions. One of the main ways to study these silent biosynthetic pathways is heterologous expression of corresponding genes in well-characterized host strain as metagenomic or genomic libraries. This approach is especially useful, since it allows for fast and simple manipulations of a gene cluster of interest, which otherwise can be difficult to handle in a natural producer. Utilization of currently available cloning methods such as transformation-associated recombination (TAR) in *Saccharomyces cerevisiae* [100], Gibson assembly [101], Red/ET-mediated recombination [102], genomic library construction [43,103] has made it possible to clone the gene cluster of interest into an appropriate heterologous host.

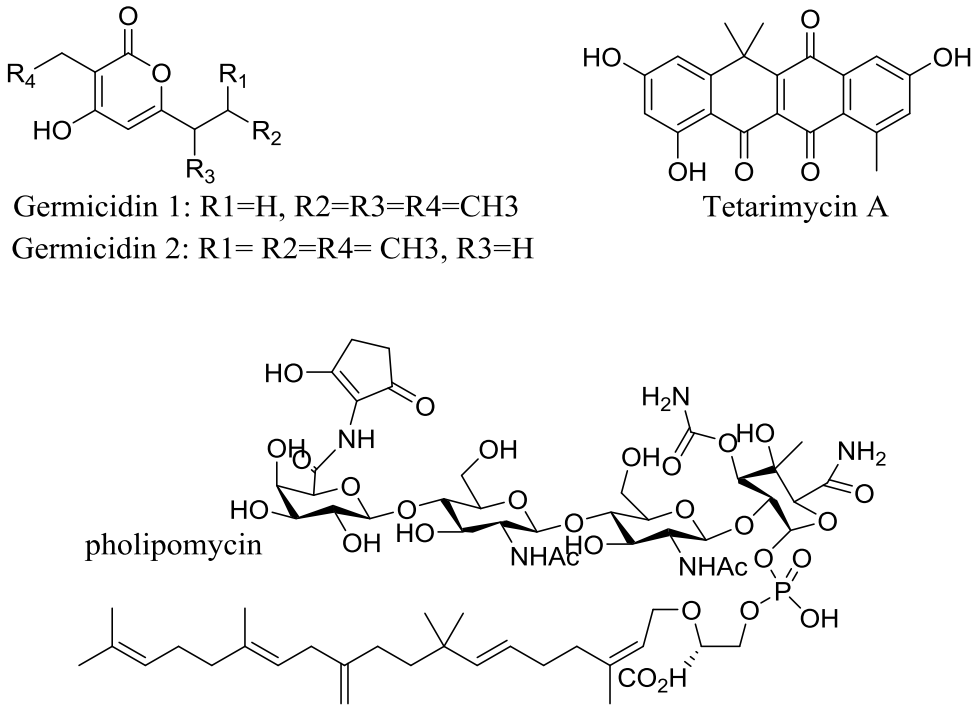


Figure 12. Examples of compounds identified by heterologous expression in different hosts.

Heterologous expression is often used to activate silent gene clusters [100,104], generate unnatural metabolites by combinatorial biosynthesis or mutasynthesis [105,106] and to obtain better yield of a compound of interest [107]. To successfully apply this approach a panel of suitable expression hosts has to be available. Typically, such host strain is expected to be well-studied, genetically amenable, provide high pool of precursors and energy as well as have available set of tools and BioBricks for reconstruction of gene cluster of interest. Since *Streptomyces* have been shown to produce a wide variety of NPs, they are preferable candidates to be used as heterologous host for bacterial natural product biosynthesis gene expression (Table 1). At the same time only several

Streptomyces species such as *S. coelicolor* M145, *S. avermitilis*, *S. albus* J1074 and *S. lividans* TK24 were intensively used. However, the production level of compound of interest in these heterologous hosts is usually low and very often the biosynthetic gene clusters cannot be efficiently expressed [108,109]. This creates the difficulties in identification of produced compounds that is often related to the accumulation of endogenous NPs of the host strain. To overcome these drawbacks several strategies are applied: optimizing production media [110], overexpressing regulatory or transport genes [111] [109], engineering precursor supply [112], refactoring the transcriptional control of the cluster [113,114], deleting negative regulators within the gene cluster [100], increasing the copy number of the gene cluster [115] and using a endogenous secondary metabolites clean hosts [116,117].

1.4.1. *Streptomyces coelicolor*

Among all Actinobacteria *S. coelicolor* M145 is the most studied species and a lot of tools have been developed for genetic engineering of this model strain [118,119]. The genome of *S. coelicolor* was sequenced and annotated in 2002 [36]. *S. coelicolor* CH999 is the first constructed host strain, which lack the production of Act and Red and was used to study the functional peculiarities of type II PKS system [120]. Since the level of production of aberrant type II PKS products is very low they were not detectable in parental *S. coelicolor* M145 due to the presence of Act and Red in the extract. The use of CH999 simplified the detection of these compounds and strongly benefited to our understanding of functional features of polyketide synthase. Following this strategy Bibb and coworkers deleted four active gene clusters from *S. coelicolor* M145 resulting in M1146 strain [117]. In addition, this strain was further improved by introducing point mutation in *rpoB* gene to generate M1152 which was further modified by introducing second point mutation in *rpsL* gene resulting in M1154 [117]. These modifications were introduced in order to increase the production yield of heterologous compounds [57]. All three modified strains have been successfully used for the expression of different biosynthetic gene clusters even from the species phylogenetically distant from *Streptomyces* [105].

The host strain can be specifically designed for the heterologous expression of a particular type of natural product. For example, *S. coelicolor* M1317 was generated for expression of actinobacterial type III PKS clusters [121]. M1317 is derivative from M1152 by deleting all type III PKS genes. The germicidin biosynthesis genes from *S. coelicolor* and *rppA* homologue from *S. venezeulae* were successfully expressed in the engineered host (Fig. 12). These results indicate that the absence of metabolic pathways competing for the same precursor molecules leads to the observed over-producing phenotypes.

1.4.2. *Streptomyces avermitilis*

Streptomyces avermitilis is the producer of avermectin and oligomycin [116]. This strain is considered to be genetically stable compared to other *Streptomyces* since it has short terminal inverted repeats (TIRs) at the ends of the linear chromosome [114]. *S. avermitilis* has been optimized to provide a high pool of primary metabolic precursors and energy for the biosynthesis of secondary metabolites [122]. Therefore, it was of a great importance to generate a host based on this strain. Komatsu and Ikeda constructed the genome-minimized derivatives of *S. avermitilis*: SUKA5, SUKA17 and SUKA22 using general homologous recombination and site-specific recombination (Cre-*loxP*) techniques [116]. *S. avermitilis* SUKA5 strain has deletion of the oligomycin biosynthetic gene cluster, as well the left subtelomeric region covering avermectin and filipin biosynthetic gene clusters. SUKA17 and SUKA22 are isogenic strains that have additional deletions of neopentalenolacton, carotenoid, and geosmin gene clusters. The biosynthetic gene clusters encoding different classes of secondary metabolites were successfully expressed in the engineered *S. avermitilis* strains [114,116]. For example, the cryptic pholipomycin biosynthetic gene cluster from *S. clavuligerus* ATCC 27065 was activated in SUKA22 with yield of 20 mg/L [114] (Fig. 12).

1.4.3. *Streptomyces albus*

S. albus J1074, the derivative of the *S. albus* strain G, is defective in *Sal*I restriction endonuclease gene [123,124]. It has one of the smallest genome within *Streptomyces* [125]. *S. albus* J1074 was first explored as a host by the Salas group [126]. This strain is now one of the most widely used hosts for the heterologous expression of secondary metabolite gene clusters such as fredericamycin, isomigrastatin, napyradiomycin, and meonomycin from various *Streptomyces* species [127,128]. *S. albus* J1074 has also been used for the heterologous expression of gene clusters from non- *Streptomyces* strains. The thiocoraline biosynthesis gene from marine *Micromonospora* strain was successfully expressed in *S. albus* when the *tioA* positive regulatory gene was co-expressed from the *ermEp** promoter [129]. The use of *S. albus* as a host was expanded by expressing the metagenomic DNA isolated from soil that resulted in production of novel natural products such as tetarimycin A [130] (Fig. 12), clarepoxcins A–F and landepoxcins A and B [131]. Recently, Myronovskyi, *et al.*, generated new *S. albus* Del14 chassis by deleting 15 gene clusters from the genome of *S. albus* J1074 using IMES system [132]. The engineered strain was further modified by introducing additional phage *phi*C31 *attB* sites (*S. albus* B2P1 and B4) for improved heterologous production of secondary metabolites. Heterologous expression of several characterized gene clusters revealed that the generated host strains have higher

production yields compared to the *S. albus* J1074 and *S. coelicolor* hosts. *S. albus* Del14 was successfully used to activate cryptic type I PKS gene cluster from metagenomic library resulting in production of pyridinopyrone. Furthermore, two other cryptic clusters from *Frankia alni* ACN14a and *Frankia* sp. CcI3 were successfully expressed in *S. albus* Del14 leading to production of salicylic acid, fralnimycin (from *Frankia alni* ACN14a), bhimamycin A and aloesaponarin II (from *Frankia* sp. CcI3).

1.4.4. *Streptomyces lividans*

S. lividans is closely related to *S. coelicolor* and is one of the well characterized Actinobacteria species. *S. lividans* accepts methylated DNA and has low protease activity. Therefore, it has been used as a host for heterologous expression of many secondary metabolite gene clusters and secreted production of recombinant proteins [133-137]. Daptomycin gene cluster was heterologously expressed in *S. lividans* TK64 and *S. lividans* TK23 strains [138]. However, the daptomycin yield in heterologous strains was lower than in the native strain. The yield of daptomycin was improved by deleting Act biosynthesis genes and adjustment of phosphate level in the medium [110]. Ziermann and co-authors constructed new host *S. lividans* K4-114 and K4-155 by removing the entire *act* gene cluster from the chromosome of *S. lividans* TK24 [139]. The expression of genes coding for biosynthesis of erythromycin precursor 6-deoxyerythronolide *B (*6-dEB) in *S. lividans* K4-114, K4-155 and *S. coelicolor* CH999 resulted in approximately the same level of compound accumulation. Similarly, *S. lividans* based host with deletion of the *act* and *red* gene clusters was successfully used to produce granaticin [140].

Heterologous expression of silent biosynthetic genes in more amenable and well-characterized host strains facilitates the discovery and engineering of novel natural products. Despite that described above *Streptomyces* species are able to produce variety of heterologous NPs there is strong need in improvement their characteristics such as: 1) increasing the success rate of cluster expression; 2) yield improvement; 3) lowering metabolic background for downstream processing, etc.

Table 1. Examples of gene clusters produced in different heterologous hosts. n.r: not reported. Where yields from the native producer were not described in the expression paper, the estimated levels were determined from the literature.

Native strain	Product	Expression host	Yield (mg/l)	Reference
S. sp.2014/011-12	Alpiniamides	*S. lividans* TK24 *S. albus* Sel14	n.r	[141]
Streptomyces sp. T676	Anthracimycin	*S. coelicolor* 1152 *S. coelicolor* 1154	9.7 8.6	[142]
S. sp. BC16019	Bottromycin A2	*S. albus* *S. lividans* TK24	0 5-50 fold increase	[143]
S. scopuliridis SCSIO ZJ46	Desotamide G	*S. coelicolor* 1146 *S. coelicolor* 1152	13.8 9.7	[144]
Salinispora pacifica CNT-150	Enterocin	*S. coelicolor* M1146 / *S. lividans* TK23	n.r	[145]
S. griseus	Fredericamycin	*S. albus* J1074 *S. albus* J1074 *S. lividans* K4-114	120 132 0	[146] [147] [147]
S. platensis	Iso-migrastatin	*S. albus* J1074 *S. lividans* K4-114 *S. avermitilis* SUKA5	46 25 4.2	[148] [148] [148]
S. collinus Tü 365	Streptocollin	*S. coelicolor* 1146 *S. coelicolor* 1152	110 5.4	[149]

1.5. Outline of The Dissertation

The overall goal of the work described in this thesis is engineering of *Streptomyces albus* J1074 and *Streptomyces lividans* TK24 for efficient endogenous and heterologous NPs production. Two strategies were developed and applied in this work: the reporter-guided screening combined with transposon mutagenesis in order to awake the *S. albus* silent gene clusters; and generation of streamlined *S. lividans* chassis lacking 11 endogenous gene clusters for heterologous production of NPs.

In **Chapter 2** we describe the development and application of reporter-guided screening strategy for the activation of silent polycyclic tetramate macrolactam cluster in *S. albus* J1074. The promoter of the PKS-NRPS gene from the corresponding gene cluster was cloned and fused with the in-house developed β-glucuronidase encoding reporter gene and subsequently introduced into *S. albus* J1074 resulting in *S. albus* ATGSal2P2. The *himar1* transposon system was used to create a mutants library of *S. albus* ATGSal2P2. The most promising strain *S. albus* ATGSal2P2::TN14 was producing several polycyclic tetramate macrolactams due to the

activation of the target gene cluster. In addition, the strain showed increased accumulation of two other groups of NPs candicidins and antimycins. Transposon insertion in *S. albus* ATGSal2P2::Tn14 was mapped to the *XNR_3174* gene encoding uncharacterized transcriptional regulator. Deletion of the *XNR_3174* resulted in similar to *S. albus* ATGSal2P2::Tn14 phenotype. This suggests that *XNR_3174* is a negative regulator of secondary metabolism in *S. albus*. The mutant strain was found to produce avenolide-like compound butenolide 4. The deletion of *XNR_2339* gene involved in the biosynthesis of butenolide 4 in *S. albus* Δ3174 mutant caused not only loss of butenolide production but also significant decrease in PTMs, candicidins, and antimycins accumulation suggesting the signaling role of butenolides in regulation of secondary metabolism. On other hand, the promoter activity of *XNR_2339* was higher in *S. albus* Δ3174 than in *S. albus* J1074 leading to conclusion that production of butenolides is in turn controlled by *XNR_3174* encoded transcriptional factor. The homologues of *XNR_3174* and the butenolide biosynthesis genes were found in genomes of multiple *Streptomyces* species. The influence of *XNR_3174* gene deletion on production of heterologous secondary metabolites in *S. albus* was also examined.

In **Chapter 3** we tested the influence of proteases on the secreted protein stability in *S. lividans* TK24. For this, several interesting proteases candidates were selected based on transcriptomics data and deleted in *S. lividans* TK24. Subsequently, the protease deletion strains were validated for their potential for heterologous protein secretion by expressing the gene encoding SP^{vsi}-mRFP.

Lastly, **Chapter 4**, of the thesis deals with construction of host strain based on the *S. lividans* TK24 for heterologous gene cluster expression. The success with the engineered *Streptomyces* strains encouraged us to develop a new *S. lividans* based chassis with clean metabolic profile which would allow easy detection and purification of new compounds. To fulfil this aim, 11 native gene clusters were deleted from chromosome of *S. lividans* TK24 while keeping track of the strains' fitness. We further engineered our streamlined strains to include up to 3 sites for site-directed integration of foreign DNA. The engineered strains efficiency in production of secondary metabolites was tested by expressing four heterologous gene clusters encoding biosynthesis of different classes of NPs.

Chapter 2

The Reporter-Guided Activation of Cryptic Gene Clusters: Identification of New Hidden Triggers of Secondary Metabolism in *Streptomyces albus* J1074

Yousra Ahmed[1], Yuriy Rebets[1,2], Bogdan Tokovenko[1,3], Elke Brötz[1.4], Andriy Luzhetskyy[1,2] *

Scientific Reports **volume 7**, Article number: 9784 (2017)
Published online 29 August 2017

Supplementary information of this paper is available at:
doi:10.1038/s41598-017-10316-y

2. Activation of Silent Gene Cluster

2.1. Abstract

A large majority of genome-encrypted chemical diversity in Actinobacteria remains to be discovered, which is related to the low level of secondary metabolism genes expression. Here, we report the application of a reporter-guided screening strategy to activate cryptic polycyclic tetramate macrolactam gene clusters in *Streptomyces albus* J1074. The analysis of the *S. albus* transcriptome revealed an overall low level of secondary metabolism genes transcription. Combined with transposon mutagenesis, reporter-guided screening resulted in the selection of two *S. albus* strains with altered secondary metabolites production. Transposon insertion in the most prominent strain, *S. albus* ATGSal2P2::TN14, was mapped to the *XNR_3174* gene encoding an unclassified transcriptional regulator. The mutant strain was found to produce the avenolide-like compound butenolide 4. The deletion of the gene encoding a putative acyl-CoA oxidase, an orthologue of the *Streptomyces avermitilis* avenolide biosynthesis enzyme, in the *S. albus XNR_3174* mutant caused silencing of secondary metabolism. The homologues of *XNR_3174* and the butenolide biosynthesis genes were found in the genomes of multiple *Streptomyces* species. This result leads us to believe that the discovered regulatory elements comprise a new condition-dependent system that controls secondary metabolism in Actinobacteria and can be manipulated to activate cryptic biosynthetic pathways.

2.2. Introduction

The genomics era in Actinobacteria research has led to rapid changes in our understanding of secondary metabolite diversity [1,2]. A great surprise was that only a few compounds are produced under laboratory conditions by the Actinobacteria strains harboring 30-40 secondary metabolite gene clusters in their genomes [3-5]. The majority of these genes are considered silent or expressed at a level insufficient for the detection of corresponding compounds. The strains (such as *S. coelicolor, S. lividans,* and *S. avermitilis*) that have been used for decades as model actinomycetes with well-studied metabolic profiles were found to produce new, unknown secondary metabolites [6,7]. The extensive genomics-based re-thinking of the *S. coelicolor* secondary metabolome increased the number of compounds isolated from this strain from 4 to 17 [8]. A simple calculation makes it obvious that a large majority of the chemical potential of Actinobacteria remains unexplored, opening opportunities for the discovery of natural products with new chemical scaffolds and biological activities [2]. The development of tools and strategies

for efficient genome mining for secondary metabolism genes and their activation is now the primary goal in Actinobacteria genetics [9].

Several approaches have been used to activate cryptic gene clusters in Actinobacteria (for reviews see [9,10]). Both "unselective", based on introducing global changes into intracellular processes, and "selective", manipulating a specific cluster of interest, strategies have been intensively developed and applied. "Selective" approaches include engineering the transcriptional and translational processes within a cluster of interest [11-13], overexpression of pathway-specific regulatory genes [14], refactoring of the pathway in a plug-and-play manner [15,16] and heterologous expression [17]. These methods require the development of tools for manipulating large DNA fragments, efficient heterologous hosts and well-characterized libraries of gene expression elements [7,18]. Changes in media composition [19,20], ribosomal and metabolic engineering [21,22] and manipulating pleiotropic regulatory genes [23] often induces global changes in the secondary metabolites profile of a producing strain. A necessary prerequisite for success when using "unselective" approaches is the availability of an efficient screening platform to recognize and select colonies with a desired phenotype.

Streptomyces albus J1074, the derivative of the *S. albus* strain G [24], is one of the most widely used hosts for the heterologous expression of antibiotic biosynthesis genes [25]. The great advantage of this strain is an absence of the production of endogenous secondary metabolites. Sequencing of the *S. albus* J1074 genome led to the identification via various annotations of 22 to 26 gene clusters [11,26]. Recently [11], a successful application of "selective" approaches to activate several secondary metabolism pathways in *S. albus* resulted in the accumulation of indigoidine, alteramides, antimycins and candicidins by recombinant strains.

Despite the obvious success of both strategies in activating cryptic antibiotic biosynthesis pathways in Actinobacteria, we still lack an answer to one of the major questions raised by the discovery of the cryptic secondary metabolome: why are these genes silent? The current approaches do not provide insight into the regulatory processes silencing secondary metabolism. A reporter-guided selection of mutants with activated metabolic pathways of interest might be a tool to answer this question. This approach is based on a simple assumption that the yield of the final metabolite directly correlates with the expression level of the corresponding biosynthetic gene cluster. This strategy was successfully applied in the selection of *Aspergillus terreus* strains with increased titers of lovastatin [27] and was later adapted to screen for *S. clavuligerus* mutants with improved clavulanic acid production [28]. Recently, this approach was used to activate jadomycin (*jad*) and gaudimycin (*pga*) biosynthesis pathways in *S. venezuelae* and *Streptomyces*

sp. PGA64, respectively [29]. Here, we report the further development and exploitation of this technique by combining it with transposon mutagenesis in *S. albus* J1074. This method enables fast and simple identification and characterization of the DNA locus responsible for the activation and production of the desired metabolite, providing insights into the regulatory processes silencing the secondary metabolism in Actinobacteria.

2.3. Results

Secondary metabolism gene clusters in *S. albus* are poorly transcribed under laboratory conditions

A manual correction of the AntiSMASH 2.0 outcome of *S. albus* J1074 genome analysis resulted in a list of 26 secondary metabolite biosynthetic gene clusters (BGCs) by removing two regions (*XNR_0391-0405* and *XNR_2192*) and adding the recently identified paulomycin BGC (*XNR_0573-0610*) [11]. To evaluate the transcriptional activity of the identified BGCs, we aggregated RNAseq datasets from several experiments in which *S. albus* J1074 was grown in various liquid media and MS agar medium [13,26,30] (Fig. 1, Table 3S). The transcriptome data from these experiments were normalized by rating the level of individual gene expression in percentile from the highest expressed gene. This conversion was applied to compare data from various experiments and sequencing providers. Although this process loses most gene expression information (absolute values and fold-changes within samples), it maintains the relative order of ranked gene expression, which is sufficient for basic qualitative comparisons. The highest expressed gene for all conditions was *XNR_3521*, which encodes the putative excisionase/DNA-binding protein with RPKM values ranging from 8,641 to 15,730 (Fig. 1, Table 3S). These values were taken as 100%.

We found that *S. albus* J1074 secondary metabolism BGCs are transcriptionally active over a wide range of percentile values (Fig. 1, Table 3S). However, in the majority of BGC-medium combinations, the percentile expression was below 60-70%, with the 50 percentile ranking corresponding to values as low as 6-28 RPKM in various experiments. For comparison, the observed percentile expression across all media for *rpoB* was 93.4-98.2% (448-1,504 RPKM) and fo r *rps12* was 98.5-99.7% (2,683-9,949 RPKM). The only highly expressed BGC (number 17) under tested conditions was predicted to be responsible for the biosynthesis of ectoine. Its rank-expression was 76.5-92.3% across all media (70-310 RPKM). Surprisingly, a high level of transcription was detected for the putative type I lantipeptide gene cluster (BGC 11) when the strain was grown in GYM medium (ranked at 98%, 2218 RPKM) (Table 3S).The above-

described BGCs are exceptions; the majority of genes involved in secondary metabolism in the *S. albus* J1074 genome are poorly transcribed under the examined conditions at levels that appear insufficient for producing detectable amounts of metabolites.

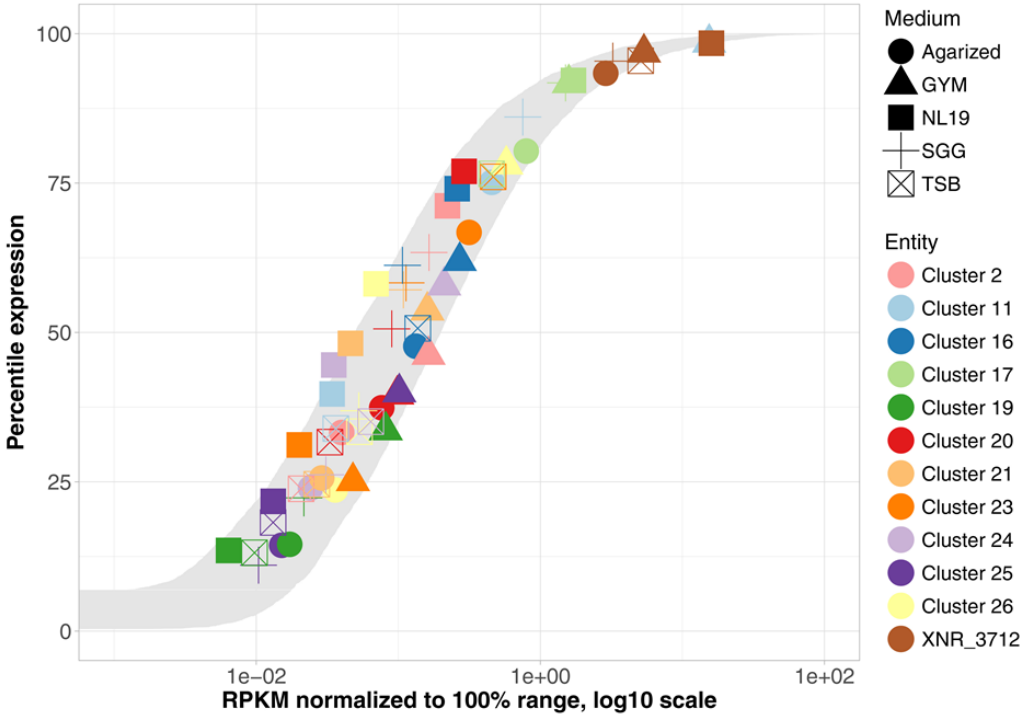


Figure 1. The gray ribbon shows the range of percentile-logRPKM curves for all genes within the *S. albus* J1074 genome, grown in all 7 media (for any given percentile, the ribbon covers all observed logRPKM values, including minima and maxima). Subsets of 5 media (shape-coded) and 11 gene clusters (color-coded) are also shown, with additional horizontal and vertical jitter to alleviate data-point overlaps. The rpoB (*XNR_3712*; brown) gene expression is shown for reference.

Reporter-guided selection of mutants with activation of polycyclic tetramate macrolactam (PTM) biosynthesis gene cluster

The gene cluster BGC 2 encodes a hybrid type I PKS-NRPS. Its activation, by the insertion of a strong promoter, led to accumulation of the 6-*epi*-alteramides A and B [11]. Cluster 2 is among the most poorly transcribed in the *S. albus* genome, ranked at the 32-70 percentile with average read counts of 6 to 23 RPKM in the various media (Fig. 1, Table 3S). The transcriptome data clearly showed that BGC 2 is organized into two transcriptional units. The ORFs *XNR_0204*, *0203*, *0202* and *0201* encode a hybrid type I PKS-NRPS, two oxidoreductases and a zinc-dependent alcohol dehydrogenase (cyclase) to form an operon, with the predicted transcription start site at 29 bp upstream from the start codon of *XNR_0204* (Fig. 2S). From the available RNAseq data,

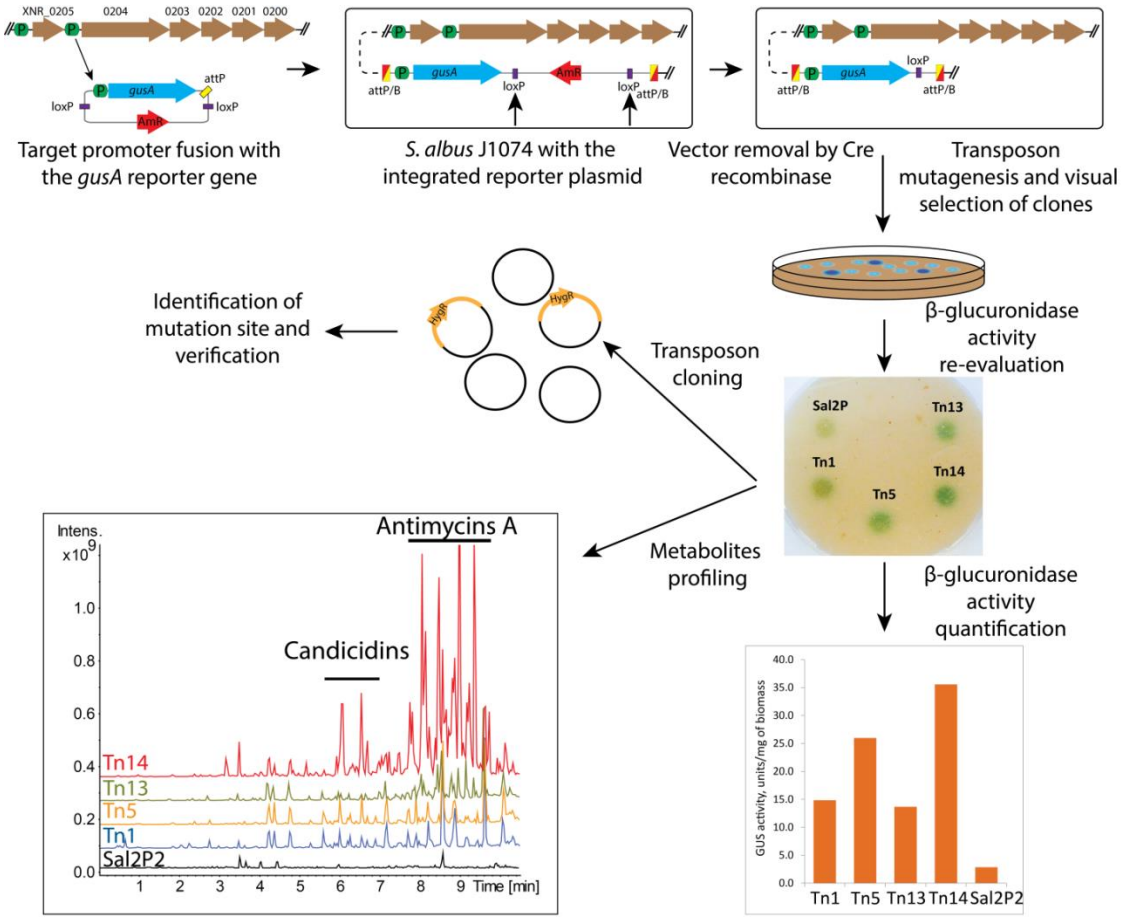


Figure 2. Schematic illustration of the workflow.

we cannot tell whether *XNR_0200*, which encodes the putative cytochrome P450, is also transcribed as part of this operon. *XNR_0205*, coding for putative hydroxylase/desaturase, is transcribed as monocistronic mRNA from the promoter with the (+1) position mapped 31 bp upstream of the start codon.

We cloned a 184-bp DNA fragment, including a promoter, RBS and the first 18 bp of the coding region of the *XNR_0204* gene into the plasmid carrying the ATG allele of the *gusA* reporter gene (Fig. 2, 1S, 2S) [31]. The transcriptional fusion construct facilitated rapid and simple evaluation of *XNR_0204* promoter activity, enabling visual selection of clones with increased transcription. The fusion construct was introduced into the *S. albus* J1074 strain, and the β-glucuronidase activity was tested. The recombinant strain *S. albus* ATGSal2P2 displayed a very low level of β-glucuronidase activity *in vivo* (Fig. 2). The *himar1* transposon system [32] was applied to generate a mutant library of *S. albus* ATGSal2P2, which was further visually screened for clones with increased β-glucuronidase activity. As a result, nine clones were pre-selected from approximately 20,000 colonies. The selected strains were cultivated in NL19 medium, a liquid analog of the MS medium used for screening, and were analyzed for β-glucuronidase activity in

cell-free lysate. Only four of nine mutants exhibited increased *gusA* expression (Fig. 2). Among these, *S. albus* ATGSal2P2::Tn14 showed the highest level of reporter activity. The strains were cultivated in several media and the produced metabolites were analyzed by LC-MS. *S. albus* ATGSal2P2::Tn13 and ATGSal2P2::Tn14 were found to produce candicidins and antimycins

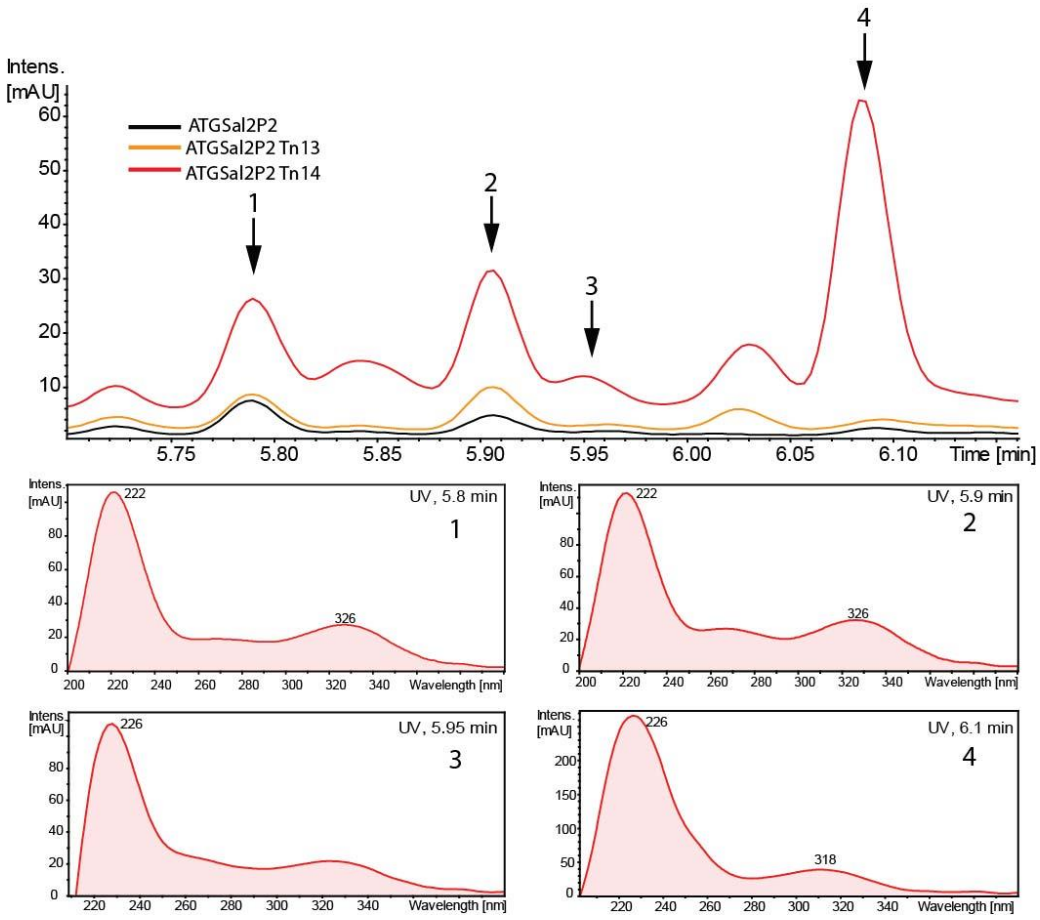


Figure 3. LC-MS-based identification of PTMs produced by *S. albus* ATGSal2P2::Tn13 and *S. albus* ATGSal2P2::Tn14 strains. Samples were separated with a 10 min gradient.

when grown in R5A and NL19 media, with the latter strain ATGSal2P2::Tn14 showing a significant increase in the accumulation of these compounds (Fig. 2, 3S, 4S). The metabolic profiles of two other mutants were not affected when compared to the parental strain. The rate of false positive clones was approximately 50%, as expected for a single reported gene selection procedure [29]. Two compounds with absorption spectra typical of PTMs [11] can be found in the extracts of *S. albus* ATGSal2P2::Tn13 and *S. albus* ATGSal2P2::Tn14, as well as in the parental strain (compounds **1** and **2**; retention time of 5.8 and 5.9 minutes, respectively) (Fig. 3). *S. albus* ATGSal2P2::Tn13 showed a minor increase in the production of compound **2**. The *S. albus* ATGSal2P2::Tn14 strain not only accumulates more of compounds **1** and **2** but also produces several other compounds with similar spectral characteristics (Fig. 3, 5S). Due to the low yield of

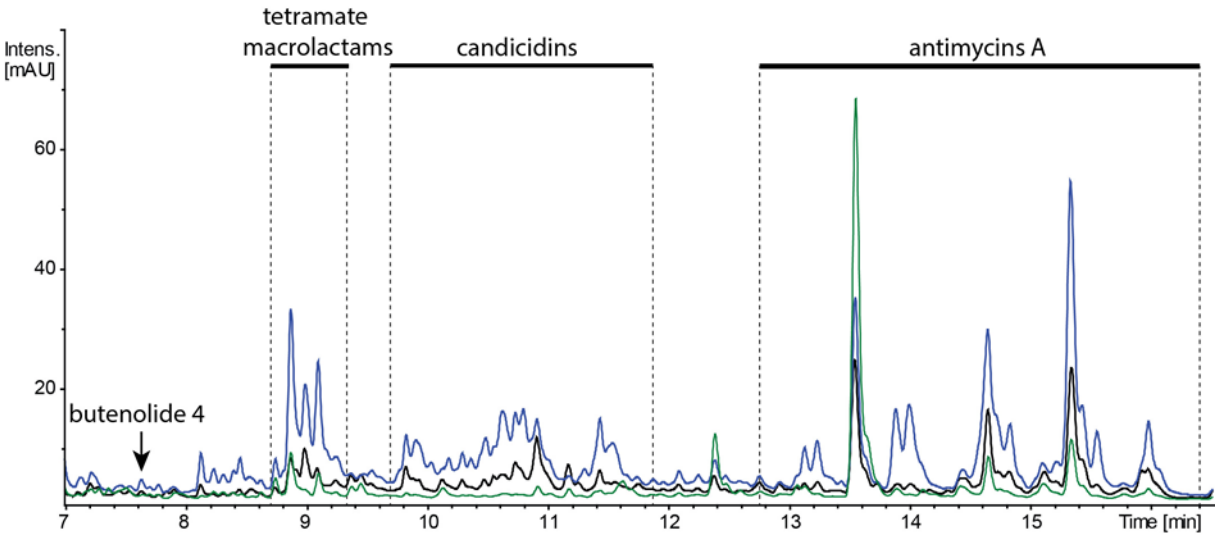


Figure 4. LC-MS chromatogram (at 320 nm) of extracts of *S. albus* J1074 (black), *S. albus* Δ3174 (blue) and *S. albus* Δ3174Δ2339 (green). Groups of identified compounds are highlighted. Metabolites were separated with a 20 min gradient.

these metabolites and their partial overlap with candicidins during LC separation, we grew *S. albus* ATGSal2P2::Tn14 in 5 liters of NL19 medium and fractionated the extract by size-exclusion chromatography prior to high resolution LC-MS and MS/MS analysis. Fraction 11 was found to contain five different compounds, two of which (**1** and **2**) were originally detected in the extracts of *S. albus* ATGSal2P2 and both mutants, and three (**3**, **4**, **5**) were present only in the extract of the *S. albus* ATGSal2P2::Tn14 strain (Fig. 3, 5S). While we were not able to purify the

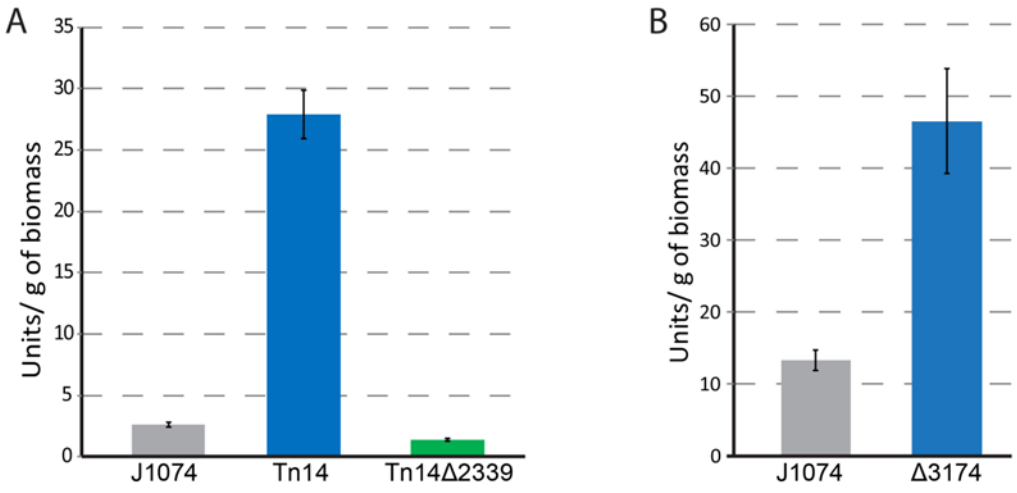


Figure 5. (A) β-glucuronidase activity of cell-free extracts of *S. albus* J1074 (grey), *S. albus* ATGSal2p2::Tn14 (blue) and *S. albus* ATGSal2p2::Tn14Δ2339 (green) resulting from the *gusA* reporter expressed with the promoter of *XNR_0204* gene. (**B)** β-glucuronidase activity of cell-free extracts of *S. albus* J1074 (grey) and *S. albus* Δ3174 (blue) resulted from the *gusA* reporter expressed with the promoter of *XNR_2339* gene.

detected compounds for detailed structure elucidation, all of them have MS/MS fragmentation pattern similar to pure ikarugamycin sample with three characteristic ions, that seems to be common for entire PTMs family (Fig. 6S). Based in exact mass and fragmentation pattern we think that compound **2** might be alteramide A (m/z of 511.2678 $[M+H]^+$; calculated m/z 511.2763 $[M+H]^+$) (Fig. 5S, 6S), when ccompounds **3** with m/z 509.2562 $[M+H]^+$ and **5** with m/z of 525.2552 $[M+H]^+$ most probably are frontalamides B and A, respectively (calculated m/z B 509.2645 $[M+H]^+$ and A 525.2594 $[M+H]^+$) (Fig. 7S). Nevertheless, despite the lack of exact structures of detected compounds, we can clearly tell that they belong to PTMs family and originated from the activation of BGC2.

XNR_3174 represses biosynthesis of secondary metabolites in *S. albus* J1074

The transposon insertion loci were retrieved from the chromosome of mutant strains and sequenced (Table 4S). The transposon insertion in *S. albus* ATGSal2P2::Tn13 occurred in the *XNR_3521* gene, which encodes a putative MerR family transcription factor. This protein was found to be highly conserved among various Actinobacteria, with the degree of amino acid identity ranging from 86% to 100%. In the case of *S. albus* ATGSal2P2::Tn14, the transposon was mapped to the 3' end of the coding region of the *XNR_3174* gene. The *XNR_3174* product is annotated as a putative transcriptional regulator with unknown function [26]. This gene is transcribed as monocystronic mRNA and showed high transcriptional activity in the cultures of *S. albus* J1074 grown in NL19 or solid MS media, but not in SGG or TSB (Table 3S). Transposon insertion disrupts the *XNR_3174* gene by cutting the last 13 codons. The domain BLAST analysis showed that this region of the protein is predicted to form a helix-turn-helix DNA binding motif.

Because *S. albus* ATGSal2P2::Tn14 transposon insertion showed the most prominent phenotypic expression, the *XNR_3174* gene was deleted in the chromosome of *S. albus* J1074 to verify the mutant phenotype. The obtained strain *S. albus* Δ3174 demonstrated a similar metabolic profile to *S. albus* ATGSal2P2::Tn14, accumulating PTMs, candicidins and antimycins when grown in NL19 medium (Fig. 4). The reintroduction of *XNR_3174* on a plasmid into *S. albus* Δ3174 resulted in cessation of PTM and candicidins production and a significant decrease in antimycins accumulation.

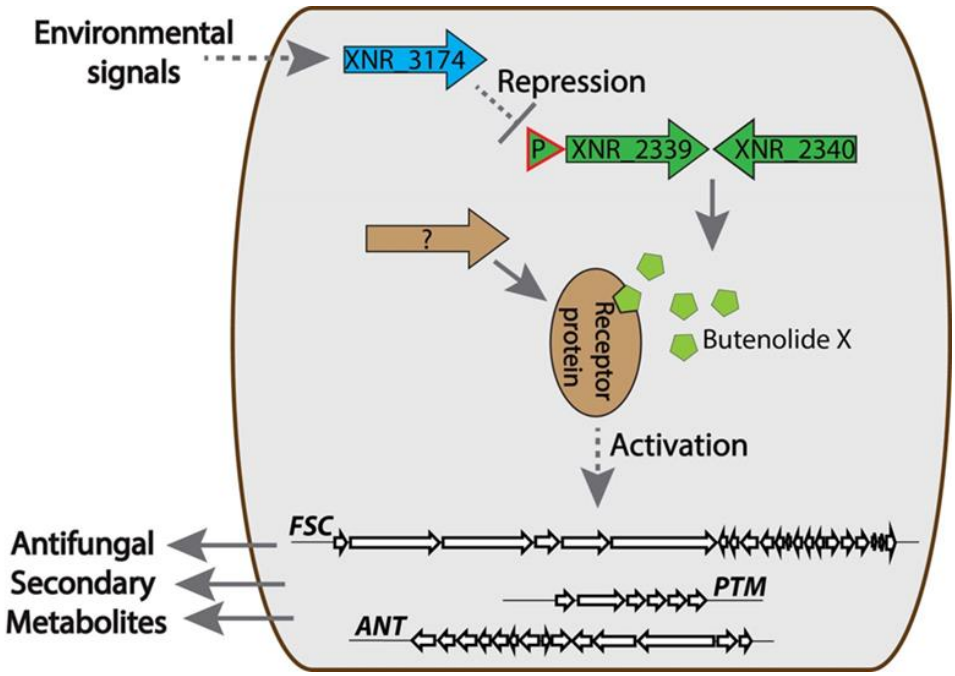


Figure 6. Proposed scheme of regulation of antifungal secondary metabolites (PTMs, candicidins and antimycins) biosynthesis in *S. albus* J1074.

Putative butenolide biosynthesis genes are involved in the regulation of secondary metabolism in *S. albus* J1074

Analysis of extracts from the *S. albus* ATGSal2P2::Tn14 and Δ3174 strains revealed one compound with *m/z* 225.1373 $[M+H]^+$ that was not produced by the parental strain under the tested conditions (Fig. 4, 8S). This compound was purified, and its structure was established by NMR experiments to be 4-hydroxy-10-methyl-11-oxo-dodec-2-en-1,4-olide (butenolide 4) (Table 5S, Fig. 9S) This compound, together with several other derivatives, was previously isolated from extracts of marine *Streptomyces* sp. SM8 and *S.* sp. B3497 [33,34], and is structurally related to avenolide, the chemical trigger of secondary metabolism in *S. avermitilis* [35]. However, we did not detect avenolide itself in the extracts of the *S. albus* ATGSal2P2::Tn14 or Δ3174 strains.

The biosynthesis of avenolide in *S. avermitilis* is proposed to be governed by two enzymes: acyl-CoA oxidase and cytochrome P450 hydroxylase [35]. Genes encoding these enzymes (*aco* and *cyp17*, respectively) are clustered with the gene for the avenolide receptor protein AvaR1. We searched the genome of *S. albus* J1074 for the *aco* and *cyp17* orthologues, excluding *avaR1* from the search. Two genes, *XNR_2339* and *XNR_2340*, encoding a putative acyl-CoA oxidase and P450 hydroxylase, were found to be the closest homologues of *aco* and *cyp17* in the *S. albus* J1074 genome (Table 6S). These genes are poorly transcribed in *S. albus* J1074 according to

RNAseq data (Table 3S). Additional genome mining of the *avaR1* homologue resulted in detection of the *XNR_4681* gene as a possible hit (31% identity and 43% similarity of protein sequences). The *XNR_2339* gene was deleted from the chromosome of *S. albus* ATGSal2P2::Tn14 and *S. albus* Δ3174. The obtained mutants lost the ability to accumulate butenolide 4 and had significantly decreased production of PTMs, candicidins and antimycins (Fig. 4, 8S). At the same time, transcription from the *XNR_0204* promoter driving the expression of the PTM synthase gene was severely decreased in the double mutant compared to the *S. albus* ATGSal2P2::Tn14 strain (Fig. 5A). The genetic complementation of *S. albus* ATGSal2P2::Tn14Δ2339 with native *XNR_2339* led to restoration of secondary metabolite production. Furthermore, the overexpression of the *XNR_2339* gene from the *ermE* promoter induced secondary metabolism in *S. albus* J1074 (Fig. 10S).

We generated the transcriptional fusion of the *XNR_2339* gene promoter (a 273-bp DNA fragment including the promoter and the 98-bp coding region of *XNR_2339*) with the *gusA* reporter. The resulting construct was introduced into *S. albus* J1074 and *S. albus* Δ3174, and the β-glucuronidase activity was measured in the cell-free extract (Fig. 5B). The *XNR_2339* promoter was found to be 6- to 10-fold more active in *S. albus* Δ3174 than in the wild-type strain, suggesting that the *XNR_3174* transcriptional factor represses the expression of the *XNR_2339* and *XNR_2340* genes responsible for production of the avenolide-like chemical trigger, rather than directly influencing secondary metabolites biosynthesis genes expression in *S. albus*. The deletion of *XNR_4681*, which encodes a putative butenolide receptor protein, did not influence production of secondary metabolites by *S. albus* J1074.

XNR_3174 influence the production of heterologous secondary metabolites in *S. albus*

Four gene clusters encoding biosynthesis pathways for pamamycin, aranciamycin, griseorhodin and tunicamycin were introduced into *S. albus* J1074 and Δ3174. Production of corresponding compounds was examined in three different media (NL19, SM17 and SG for aranciamycin, griseorhodin and tunicamycin or SGG for pamamycin). Metabolites of polyketide origin (pamamycin, aranciamycin and griseorhodin) were overproduced in the mutant strain when cultivated in NL19 and SG (or SGG for pamamycin) media (Fig. 11S). In average, up to 5 fold increase in griseorhodin, 14 fold in pamamycin and 5 fold in aranciamycin accumulation was observed. The observed phenomenon was clearly media dependent, since there were no changes in production of these metabolites by SM17 grown cultures of *S. albus* Δ3174. In the same time, biosynthesis of tunicamycin and moenomycin (data not shown) was not affected by the deletion of *XNR_3174* at any tested conditions. Therefore, *XNR_3174* influence the expression of

heterologous polyketide biosynthesis gene clusters in *S. albus*. This correlates with the observed increase in production of intrinsic PKS and hybrid PKS-NRPS products in *S. albus* Δ3174.

Butenolides are an alternative hormone system in Actinobacteria dedicated to secondary metabolism

To estimate the distribution of the butenolide regulatory system among different Actinobacteria, we performed a BLAST search for homologs of the Aco and Cyp17 proteins from *S. avermitilis*, XNR_2339, XNR_2340 and XNR_3174 from *S. albus* J1074 and the A-factor synthase AfsA from *S. griseus* within the Actinobacteria genomes available from public databases (Table 6S). As many genomes are only available as draft sequences, the proximity of the *aco* and *cyp17* orthologues was excluded from the search criteria. The putative butenolide biosynthesis genes were only found in the genus *Streptomyces* (Table 6S). Among the 415 Streptomycetes genomes searched, putative butenolide biosynthesis enzymes were present in 74 strains (aa identity threshold set to 50%). Among these, 19 strains can be categorized into a provisional *S. albus* sub-group (Table 6S, Fig. 12S). These strains not only showed a high degree of conservation of XNR_2339, XNR_2340 and XNR_3174 proteins but also carried the putative antimycin, candicidin and PTM biosynthesis gene clusters within their genomes. Several of these strains also encode putative AfsA homologues, whereas others did not. One of these strains, *S. albidoflavus NRRL B-1271,* was grown in NL19 media, and the extracted metabolites were analyzed by LC-MS. The strain was found to produce a set of candicidins and butenolide 4 (Fig. 13S). At least two other strains from the *S. albus* sub-group, *Streptomyces sp.* S4 and *Streptomyces sp.* SM8, are known to produce antimycins and candicidins [34,36].

2.4. Discussion

The production of secondary metabolites in Actinobacteria is controlled by environmental factors through diverse regulatory genes and networks [37]. In all cases, regulation occurs in response to internal or external signals that are often related to changes in cultivation conditions and aging of the culture. The laboratory conditions are obviously different from those that bacteria face in natural environments. These differences are believed to be reflected in the secondary metabolite repertoire of Actinobacteria. Manipulation of cultivation conditions is often used to induce the production of new chemicals by seemingly well-characterized strains [19,38]. However, the utility of this approach emphasizes the gaps in our understanding of environmental factors and corresponding intracellular regulatory processes that cause silencing of some secondary metabolism pathways while maintaining others.

The majority of secondary metabolism gene clusters discovered within Actinobacteria genomes are considered silent under laboratory conditions. It is thought that silencing occurs at the transcriptional level, suggesting that activation approaches can be developed on the basis of modifying transcriptional control elements of biosynthetic genes [18]. The transcriptomic data analyzed in this work suggest that secondary metabolism genes in *S. albus* J1074 are transcriptionally active (Fig. 1, Table 3S). However, most BGCs are transcribed at a very low level that appears to be insufficient for producing detectable amounts of corresponding metabolites. In some cases, such as BGC 11, the transcription of the structural genes is sufficiently high for producing the compound, and the limitations lie in the area of metabolite detection and identification. This was clearly shown in the case of paulomycins [11]. However, efficient expression of some secondary metabolite gene clusters, such as the one for desferrioxamine, appears to be indispensable [39].

Many approaches and strategies have been developed and successfully applied for the activation of silent secondary metabolite biosynthesis gene clusters (for a review see [9]). Some of these strategies (mainly "unselective" ones) involve intensive screening procedures. However, the ability to genetically manipulate Actinobacteria strains outperforms the throughput of existing screening techniques. This is when the reporter-guided selection of clones with increased transcription of genes for a metabolic pathway of interest can be truly beneficial. This approach, combined with mutagenesis, was first applied to select lovastatin-overproducing *Aspergillus terreus* variants [27] and was recently adapted to facilitate the activation of secondary metabolism in Streptomycetes [29]. In both cases, conventional mutagenesis was used to generate mutant libraries, making localization of mutations difficult. Here, we combined the reporter-guided screening approach with transposon mutagenesis to awake the poorly expressed *S. albus* J1074 PTM gene cluster. The choice of *gusA* as a reporter was prompted by the simplicity of β-glucuronidase activity detection and quantification both *in vivo* and *in vitro* compared to other reporter systems for Actinobacteria [7,31]. However, the use of a single reporter resulted in a high rate of false positive clones selection (approximately 50%). The utilization of a second reporter gene, as proposed by Guo and co-authors, will decrease this number [29]. Moreover, the combination of the described approach with transposon mutagenesis enables simple and rapid identification of mutated loci, eliminating the need for genome re-sequencing [32]. The obvious bottleneck of the transposon technique, the ability to generate only inactivation mutants, can be easily overcome with a transposon carrying strong constitutive promoters at the ends of the transposable construct [40].

Using the proposed approach, we selected four strains with increased promoter activity for the *XNR_0204* gene, encoding a putative NRPS-PKS enzyme from the PTM biosynthesis gene cluster. This cluster shares high similarity with frontalamide biosynthesis genes from *Streptomyces sp.* SPB78 (61-83% amino acid identity and 73-90% amino acid similarity among individual proteins, including the putative oxidoreductases/cyclases XNR_0203, 0202 and 0201 and hydroxylase XNR_0205) [41]. However, the attempt to activate BGC 2 resulted in the production of *6-epi*-alteramides rather than frontalamides [11]. We detected the compound with similar to alteramide A features in the extract of *S. albus* J1074 and two obtained mutants *S. albus* ATGSal2P2::Tn13 and ATGSal2P2::Tn14 (Fig. 3, 5S). At the same time, the ATGSal2P2::Tn14 strain was found to produce two other PTMs with exact mass, adsorption spectra and MS/MS fragmentation pattern similar to frontalamides A and B or highly related compounds. The alteramides were previously proposed to be the shunt products of the PTMs assembly line [42]. Inactivation of the cyclase gene *ikaC* in ikaguramycin biosynthesis or OX2 or OX4 from the HSAF gene cluster (heat-stable antifungal factor, with the same cyclization pattern as frontalamides) was shown to lead to accumulation of alteramides [43,44]. Furthermore, the ectopic expression of several PTM gene clusters resulted in accumulation of alteramides as a consequence of spontaneous cyclization of the common biosynthetic precursor [15,45]. This result suggests that the final products of *S. albus* BGC 2 most probably are frontalamides or related PTM compounds rather than alteramides, as previously thought [11]. Notably, the properly balanced expression of the secondary metabolite biosynthesis gene clusters appears to be crucial for formation of the appropriate product. Similar observations were recently made during an attempt to express the landomycin A biosynthesis gene cluster from a strong constitutive promoter [13]. In addition to the PTMs, *S. albus* ATGSal2P2::Tn13 and ATGSal2P2::Tn14 were found to produce antimycins and candicidins.

The mutation in *S. albus* ATGSal2P2::Tn13 was mapped to the gene encoding the MerR family transcriptional regulator, whereas in *S. albus* ATGSal2P2::Tn14, transposon insertion occurred in the *XNR_3174* gene, which encodes a putative uncharacterized transcriptional factor with a typical helix-turn-helix DNA binding motif (Table 4S). Deletion of *XNR_3174* in the *S. albus* J1074 did not affect the growth rate or morphological development of the strain and only influenced secondary metabolism. Furthermore, the effect of mutation in *XNR_3174* was more evident when the strain was cultivated on MS or in NL19 media, which was used in the original screening procedure. This result suggests that the identified regulatory system is dependent on growth conditions and responds to factors present in the particular medium.

Both *S. albus* ATGSal2P2::Tn14 and of *S. albus* Δ3174 were found to produce butenolide 4. This compound structurally resembles avenolide, the chemical trigger of avermectin production in *S. avermitilis* [35]. Butenolide 4 itself was not able to induce antibiotic production in *S. albus* J1074. However, the overexpression of the *XNR_2339* gene, encoding the orthologue of the avenolide biosynthesis acyl-CoA oxidase from the *ermE* promoter in *S. albus* J1074, led to activation of PTMs, antimycins and candicidins production. This result suggests that the activation of secondary metabolism in *S. albus* ATGSal2P2::Tn14 and *S. albus* Δ3174 strains is triggered by an as yet unidentified avenolide-like chemical messenger rather than is the direct outcome of *XNR_3174* inactivation (Fig. 6). In turn, the *XNR_3174* gene products appear to act as a repressor of this chemical messenger production because the expression of *XNR_2339* is negligible in *S. albus* J1074 and severely increased in *S. albus* ATGSal2P2::Tn14 and Δ3174 (Fig. 5). Additionally, expression of the *XNR_0204* gene strongly depends on the presence of *XNR_2339* in the chromosome of the strain. Interestingly, the discovered regulatory cascade not only controls the production of intrinsic secondary metabolites, but also influences the expression of heterologous gene clusters. This knowledge can be used to generate a better host strains for production of secondary metabolites.

Homologues of avenolide biosynthesis genes and the *XNR_3174* transcriptional regulator were found within the genomes of 74 *Streptomyces* strains (Table 6S). Among them, 19 strains can clearly be distinguished based on the presence of XNR_3174, XNR_2339 and XNR_2340 orthologues. Furthermore, almost all 19 strains carry putative gene clusters for antimycin, candicidin and PTM biosynthesis within their genomes. One strain, *Streptomyces* sp. SM8, is known to produce both antimycins and butenolides 1-3 and 4 [34]. We were also able to find candicidins and butenolide 4 in the extract of *S. albidoflavus NRRL B-1271* (Fig. 13S). The high degree of conservation of the XNR_3174, XNR_2339 and XNR_2340 proteins among various Streptomycetes and a clear link between these regulatory elements and several secondary metabolite biosynthesis gene clusters indicate the universality of the detected regulatory system. Furthermore, the identified regulatory mechanism controls the production of secondary metabolites with antifungal activities in response to environmental triggers. This finding may indicate the physiological and ecological significance of such a regulatory cascade for the producing strains.

2.5. Materials and Methods

Strains, plasmids and cultures conditions

All strains and plasmids used in this work are listed in Table 1S. *E. coli* strains were grown in LB medium [46]. *S. albus* strains were grown on mannitol soy flour agar (MS agar) [47] and in liquid TSB medium (Sigma-Aldrich, USA). NL19 (MS medium without agar), SG [48], NL5 [49], NL111 [50], SGG [51], SM17 (glucose 2 g/L, glycerol 40 g/L, soluble starch 2 g/L, soya flour 5 g/L, peptone (Oxoid L37) 5 g/L, yeast extract 5 g/L, NaCl 5 g/L, $CaCO_3$ 2 g/L, tap water) and R5A [47] media were used for secondary metabolites production. The following antibiotics were used when required: apramycin (50 µg/mL), kanamycin (30 µg/mL), hygromycin (50 µg/mL), thiostrepton (50 µg/mL) and phosphomycin (100 µg/mL) (Carl Roth, Germany, Sigma-Aldrich, USA).

Isolation and manipulation of DNA

DNA extraction and manipulation, *E. coli* transformation and *E. coli*/*Streptomyces* intergeneric conjugation were performed according to standard protocols [46,47,52]. Dream Taq polymerase (Thermo Fisher Scientific, USA) was used to amplify DNA fragments for cloning, for PCR verification constructs and strains. DNA fragments were purified from agarose gels using the QIAquick Gel Extraction Kit (Qiagen, Germany). Plasmid and chromosomal DNA were purified with QIAprep Spin miniprep kit and DNeasy Blood and Tissue Kit (Qiagen, Germany). Restriction endonucleases and ligase were used accordingly to manufacturer recommendations (New England Biolabs, USA). Oligonucleotides used in this study are listed in Table 2S (Eurofins Genomics, Germany).

Construction of *S. albus* ATGSal2P2 and transposon mutagenesis

Promoter of *XNR_0204* gene was amplified with the primers Sal2P2F and Sal2P2R and cloned using the CloneJET PCR Cloning Kit (Thermo Fisher Scientific, USA). The insert was sequenced and cloned by *Kpn*I-*Eco*RV into pGUS-MF-PVV vector carrying *gusA* reporter gene with ATG start codon (Fig. 1S) [31]. The resulting plasmid was introduced into *S. albus* J1074 and vector backbone was removed by expressing the Cre recombinase [53]. Obtained strain *S. albus* ATGSal2P2 was tested for β-glucuronidase activity when grown on MS and in NL19 media as described [31]. 1 mL of pre-culture was inoculated into 50 mL of NL19 media and grown for 36 h or 48h. 2 mL of culture were taken for β-glucuronidase activity measurements in cell free extract. Data was normalized to biomass weight. Measurements were triplicated.

Transposon mutagenesis was performed as described [32] with small modifications. Briefly, the transposon was introduced on pHTM vector, carrying *himar1* transposase gene and apramycin resistance gene within the transposable cassette and hygromycin resistance in the delivery vector. Culture was grown for 2 h in TSB medium at 30°C and transposition was induced with 5μg/mL of thiostrepton. After incubation at 30°C for 4 h the temperature was raised to 39°C for 12 or 24 h. Cells were plated on MS plates containing apramycin, incubated for 4 days at 30°C, spores were harvested, diluted and plated out on MS plates supplemented with apramycin. Obtained colonies were transferred to TSB plates supplemented with apramycin or hygromycin to test the frequency of vector lost. The transposon library was taken for further screening if the frequency of hygromycin resistant colonies was lower than 5%. The serial dilutions of selected transposon library were plated on MS plates supplemented with 50 μg/mL of X-Gluc (X-CLUC Direct, USA), incubated for 3 days and colonies were selected by intensity and size of the blue hallo.

Cloning and sequencing of transposon insertion loci

The genomic DNA of transposon mutants was isolated, digested with *Sac*II, precipitated, self-ligated with T4-DNA ligase and transformed into *E. coli* TransforMax™ EC100D™ pir-116 (Epicentre, Madison, WI, USA). The plasmids were isolated, and chromosome-transposon junctions were sequenced using the primer p3-pALG-ch Sequences were mapped to the *S. albus* J1074 genome using Geneious software, version 8.1.7 (Biomatters Ltd, New Zealand).

Recombinant BAC construction and gene deletions

Gene deletion was performed using a Red/ET recombination approach using BAC clones 1G13 (for *XNR_3174*), 2C15 (for *XNR_2339*) and 2L15 (for *XNR_4681*) from an *S. albus* ordered genomic library and IMES antibiotic resistance cassettes [39]. The cassette was excised from a carrier plasmid with *Pvu*II and amplified using 3174_F/3174_R, 4681_F/4681_R and 2339_F/2339_R primers Red/ET was performed as previously described [54]. Deletions were confirmed by PCR using the primer pairs 3174C_F1/3174C_R1, 4681C_F/4681C_R and 2339C_F/2339C_R. The recombinant BACs were introduced into the *S. albus* strains. The double-crossover mutants were screened on MS medium supplemented with apramycin or hygromycin and 50 μg/mL of X-Gluc. The resistance marker was removed from the chromosome of the generated mutants by expression of øC31 actinophage integrase [39], and the resulting strain genotypes were confirmed by PCR with the above-described primers.

Construction of complementation plasmids

XNR_3174 was amplified from the genomic DNA of *S. albus* J1074 using the primer pair 3174E_BamHIF/3174E_HindIIIR. The PCR product was digested with *Bam*HI/*Hind*III and cloned into pUWLHFLP (Dr. Fedoryshyn, personal communication), replacing the *flp* gene and resulting in pUWLH3174. *XNR_2339* was amplified with the primer pair 2339E_F/2339E_R. The amplified fragment was digested with *Bam*HI/*Hind*III and ligated into pUWLoriT [55], resulting in pUWLT2339.

Transcriptional fusion of *XNR_2339* promoter to the *gusA* gene

The promoter region of *XNR_2339* was amplified with the primer pair GUS-XbalF/GUS-KpnIR. The amplified fragment was digested with *Kpn*I/*Xba*I and cloned into a *Kpn*I/*Xba*I-digested pGUS-MF-PVV plasmid. The resulting construct was assigned as pGUS*P*aco. The plasmid pGUS*P*aco was introduced into *S. albus* J1074 and *S. albus* Δ3174. The β-glucuronidase activity was tested as described above [31].

Metabolite extraction and analysis

S. albus strains were grown in 10 mL of TSB for 1 day, and 1 mL of each culture was used to inoculate the 50 mL of production media. Cultures were grown for 5 days at 30 °C. Metabolites were extracted with ethyl acetate from the supernatant and the acetone:methanol (1:1) mixture from biomass. Extracts were evaporated and dissolved in methanol. 1μL of each sample was separated using a Dionex Ultimate 3000 UPLC (Thermo Fisher Scientific, USA) and a 10-cm ACQUITY UPLC® BEH C18 column, 1.7 μm (Waters, USA) and a linear gradient of acetonitrile against 0.1% formic acid solution in water from 5% to 95% in 10 or 18 minutes at a flow rate of 0.6 mL/min. Samples were analyzed using an amaZon speed mass spectrometer or, when needed, maXis high-resolution LC-QTOF system (Bruker, USA). Pure ikarugamycin was used as standard in MS/MS experiments (Sigma-Aldrich, USA).

Isolation and structure elucidation of compounds

S. albus ATGSal2P2::Tn14 was grown at 30°C for 3 days in 6 x 500-mL flasks containing 50 mL of TSB, and pre-culture was used to inoculate 100 x 500-mL flasks containing 50 mL of NL19 media. Cultures were incubated at 30 °C for 5 days. Metabolites were extracted as described above. The extracts from biomass and the supernatant were combined and fractionated by size-exclusion chromatography on an LH 20 Sephadex column (Sigma-Aldrich, USA) using methanol as the solvent. The fractions were collected every 15 minutes. , evaporated and

dissolved in 0.5 mL of MeOH. Samples were further separated by preparative HPLC (Dionex UltiMate 3000, Thermo Fisher Scientific, USA) using a NUCLEODUR® C18 HTec column (250 x 10 mm, 5 µm) (Macherey-Nagel, Germany) with a linear gradient of solvent B (acetonitrile with 0.1% of formic acid) against solvent A (water with 0.1% of formic acid) at a flow rate of 4.5 mL/min at 45 °C. Compounds were separated using a gradient starting from 30% and increasing to 70% of B over 30 min. UV spectra were recorded with a DAD detector at 280 nm. Individual peaks were collected and analyzed by LC-MS as described above.

NMR spectra were acquired on a Bruker Ascend 700 MHz NMR spectrometer equipped with a 5 mm TXI cryoprobe (Bruker, USA). Deuterated $CDCL_3$ was used as a solvent and HSQC, HMBC and ^{1}H-^{1}H COSY spectra were recorded using standard pulse programs (Table 5S).

Expression of heterologous secondary metabolism gene clusters

Constructs carrying gene clusters for biosynthesis of pamamycin, griseorhodin, aranciamycin and tunicamycin (Table 1S) were transferred into *S. albus* J1074 and *S. albus* Δ3174 by mean of intergeneric conjugation. Three independent clones of each recombinant strain were grown in 10 mL of TSB for 2 days and 200 mg of biomass of each pre-culture was used to inoculated 50 mL of production media (NL19, SM17 and SG or SGG). Strains were cultivated for 5 days at 28 °C and metabolites were extracted and analyzed as described above. Metabolite accumulation was averaged and normalized by biomass. The production by *S. albus* J1074 was taken as 100% in each particular experiment.

Secondary metabolism gene clusters transcription analysis

Multiple *S. albus* J1074 RNAseq experimental datasets generated in our laboratory over the last 5 years were used to compare biosynthetic gene cluster expression across multiple media/conditions. First, gene RPKM expression values were ranked from highest to lowest and expressed as percentiles within each sample. The resulting values from all samples were combined into a single summary table for comparison. Conversion to percentiles was used as the most reliable normalization technique for comparing data from different experiments and sequencing providers. This process loses most of the information about gene expression (absolute values and fold-changes within samples) but maintains the relative order of ranked gene expression, which is sufficient for basic qualitative comparisons. To verify that the percentile yield is sufficiently robust for cross-experimental comparison, the β-subunit of the RNA polymerase gene (*XNR_3712*) was included. To accurately depict the meaning of

percentile expression, RPKM values within each sample were normalized to 100% and plotted (after log10 transformation of the RPKM % axis) together with percentile values.

2.6. References

1 Doroghazi, J. R. & Metcalf, W. W. Comparative genomics of actinomycetes with a focus on natural product biosynthetic genes. *BMC Genomics* **14**, 611, doi:10.1186/1471-2164-14-611 (2013).

2 Doroghazi, J. R. *et al.* A roadmap for natural product discovery based on large-scale genomics and metabolomics. *Nat. Chem. Biol.* **10**, 963-968, doi:10.1038/nchembio.1659 (2014).

3 Bentley, S. D. *et al.* Complete genome sequence of the model actinomycete *Streptomyces coelicolor* A3(2). *Nature* **417**, 141-147, doi:Doi 10.1038/417141a (2002).

4 Omura, S. *et al.* Genome sequence of an industrial microorganism *Streptomyces avermitilis*: deducing the ability of producing secondary metabolites. *Proc. Natl. Acad. Sci. USA.* **98**, 12215-12220, doi:10.1073/pnas.211433198 (2001).

5 Oliynyk, M. *et al.* Complete genome sequence of the erythromycin-producing bacterium *Saccharopolyspora erythraea* NRRL23338. *Nat. Biotechnol.* **25**, 447-453, doi:10.1038/nbt1297 (2007).

6 Nett, M., Ikeda, H. & Moore, B. S. Genomic basis for natural product biosynthetic diversity in the actinomycetes. *Nat. Prod. Rep.* **26**, 1362-1384, doi:10.1039/b817069j (2009).

7 Rebets, Y., Brotz, E., Tokovenko, B. & Luzhetskyy, A. Actinomycetes biosynthetic potential: how to bridge *in silico* and *in vivo*? *J. Ind. Microbiol. Biotechnol.* **41**, 387-402, doi:10.1007/s10295-013-1352-9 (2014).

8 Challis, G. L. Exploitation of the *Streptomyces coelicolor* A3(2) genome sequence for discovery of new natural products and biosynthetic pathways. *J. Ind. Microbiol. Biotechnol.* **41**, 219-232, doi:10.1007/s10295-013-1383-2 (2014).

9 Rutledge, P. J. & Challis, G. L. Discovery of microbial natural products by activation of silent biosynthetic gene clusters. *Nat. Rev. Microbiol.* **13**, 509-523, doi:10.1038/nrmicro3496 (2015).

10 Zerikly, M. & Challis, G. L. Strategies for the discovery of new natural products by genome mining. *ChemBioChem.* **10**, 625-633, doi:10.1002/cbic.200800389 (2009).

11 Olano, C. *et al.* Activation and identification of five clusters for secondary metabolites in *Streptomyces albus* J1074. *Microb. Biotechnol.* **7**, 242-256, doi:Doi 10.1111/1751-7915.12116 (2014).

12 Saleh, O. *et al.* Activation of a silent phenazine biosynthetic gene cluster reveals a novel natural product and a new resistance mechanism against phenazines. *MedChemComm.* **3**, 1009-1019, doi:10.1039/c2md20045g (2012).

13 Myronovskyi, M. *et al.* Generation of new compounds through unbalanced transcription of landomycin A cluster. *Appl. Microbiol. Biotechnol.*, doi:10.1007/s00253-016-7721-3 (2016).

14 Laureti, L. *et al.* Identification of a bioactive 51-membered macrolide complex by activation of a silent polyketide synthase in *Streptomyces ambofaciens*. *Proc. Natl. Acad. Sci. USA.* **108**, 6258-6263, doi:10.1073/pnas.1019077108 (2011).

15 Luo, Y. *et al.* Activation and characterization of a cryptic polycyclic tetramate macrolactam biosynthetic gene cluster. *Nat. Commun.* **4**, 2894, doi:10.1038/ncomms3894 (2013).

16 Bai, C. *et al.* Exploiting a precise design of universal synthetic modular regulatory elements to unlock the microbial natural products in *Streptomyces*. *Proc. Natl. Acad. Sci. USA.* **112**, 12181-12186, doi:10.1073/pnas.1511027112 (2015).

17 Komatsu, M. *et al.* Engineered *Streptomyces avermitilis* host for heterologous expression of biosynthetic gene cluster for secondary metabolites. *ACS Synth. Biol.* **2**, 384-396, doi:10.1021/sb3001003 (2013).

18 Myronovskyi, M. & Luzhetskyy, A. Native and engineered promoters in natural product discovery. *Nat. Prod. Rep.* **33**, 1006-1019, doi:10.1039/c6np00002a (2016).

19 Bode, H. B., Bethe, B., Hofs, R. & Zeeck, A. Big effects from small changes: Possible ways to explore nature's chemical diversity. *ChemBioChem* **3**, 619-627, doi:Doi 10.1002/1439-7633(20020703)3:7<619::Aid-Cbic619>3.0.Co;2-9 (2002).

20 Kawai, K., Wang, G., Okamoto, S. & Ochi, K. The rare earth, scandium, causes antibiotic overproduction in *Streptomyces* spp. *FEMS. Microbiol. Lett.* **274**, 311-315, doi:10.1111/j.1574-6968.2007.00846.x (2007).

21 Shima, J., Hesketh, A., Okamoto, S., Kawamoto, S. & Ochi, K. Induction of actinorhodin production by *rpsL* (encoding ribosomal protein S12) mutations that confer streptomycin resistance in *Streptomyces lividans* and *Streptomyces coelicolor* A3(2). *J. Bacteriol.* **178**, 7276-7284 (1996).

22 Hosaka, T. *et al.* Antibacterial discovery in actinomycetes strains with mutations in RNA polymerase or ribosomal protein S12. *Nat. Biotechnol.* **27**, 462-464, doi:10.1038/nbt.1538 (2009).

23 McKenzie, N. L. *et al.* Induction of antimicrobial activities in heterologous Streptomycetes using alleles of the *Streptomyces coelicolor* gene *absA1*. *J. Antibiot.* **63**, 177-182, doi:10.1038/ja.2010.13 (2010).

24 Chater, K. F. & Wilde, L. C. Restriction of a bacteriophage of *Streptomyces albus* G involving endonuclease *Sal*I. *J. Bacteriol.* **128**, 644-650 (1976).

25 Baltz, R. H. *Streptomyces* and *Saccharopolyspora* hosts for heterologous expression of secondary metabolite gene clusters. *J. Ind. Microbiol. Biotechnol.* **37**, 759-772, doi:DOI 10.1007/s10295-010-0730-9 (2010).

26 Zaburannyi, N., Rabyk, M., Ostash, B., Fedorenko, V. & Luzhetskyy, A. Insights into naturally minimised *Streptomyces albus* J1074 genome. *BMC Genomics* **15**, doi:10.1186/1471-2164-15-97 (2014).

27 Askenazi, M. *et al.* Integrating transcriptional and metabolite profiles to direct the engineering of lovastatin-producing fungal strains. *Nat. Biotechnol.* **21**, 150-156, doi:10.1038/nbt781 (2003).

28 Xiang, S. H. *et al.* Application of a double-reporter-guided mutant selection method to improve clavulanic acid production in *Streptomyces clavuligerus*. *Metab. Eng.* **11**, 310-318, doi:10.1016/j.ymben.2009.06.003 (2009).

29 Guo, F. *et al.* Targeted activation of silent natural product biosynthesis pathways by reporter-guided mutant selection. *Metab. Eng.* **28**, 134-142, doi:10.1016/j.ymben.2014.12.006 (2015).

30 Myronovskyi, M. *et al.* Genome rearrangements of *Streptomyces albus* J1074 lead to the carotenoid gene cluster activation. *Appl. Microbiol. Biotechnol.* **98**, 795-806, doi:10.1007/s00253-013-5440-6 (2014).

31 Myronovskyi, M., Welle, E., Fedorenko, V. & Luzhetskyy, A. beta-Glucuronidase as a sensitive and versatile reporter in Actinomycetes. *Appl. Environ. Microbiol.* **77**, 5370-5383, doi:10.1128/Aem.00434-11 (2011).

32 Bilyk, B. *et al.* In vivo random mutagenesis of Streptomycetes using mariner-based transposon *Himar1*. *Appl. Microbiol. Biotechnol.* **97**, 351-359, doi:10.1007/s00253-012-4550-x (2013).

33 Mukku, V. J., Speitling, M., Laatsch, H. & Helmke, E. New butenolides from two marine Streptomycetes. *J. Nat. Prod.* **63**, 1570-1572 (2000).

34 Viegelmann, C. *et al.* Metabolomic profiling and genomic study of a marine sponge-associated *Streptomyces* sp. *Marine drugs* **12**, 3323-3351, doi:10.3390/md12063323 (2014).

35 Kitani, S. *et al.* Avenolide, a *Streptomyces* hormone controlling antibiotic production in *Streptomyces avermitilis*. *Proc. Natl. Acad. Sci. USA.* **108**, 16410-16415, doi:10.1073/pnas.1113908108 (2011).

36 Seipke, R. F. *et al.* A single *Streptomyces* symbiont makes multiple antifungals to support the fungus farming ant *Acromyrmex octospinosus*. *PlOS ONE* **6**, e22028, doi:10.1371/journal.pone.0022028 (2011).

37 Liu, G., Chater, K. F., Chandra, G., Niu, G. & Tan, H. Molecular regulation of antibiotic biosynthesis in *Streptomyces*. *Microbiol. Mol. Biol. Rev.* **77**, 112-143, doi:10.1128/MMBR.00054-12 (2013).

38 Hewage, R. T., Aree, T., Mahidol, C., Ruchirawat, S. & Kittakoop, P. One strain-many compounds (OSMAC) method for production of polyketides, azaphilones, and an isochromanone using the endophytic fungus *Dothideomycete* sp. *Phytochemistry* **108**, 87-94, doi:10.1016/j.phytochem.2014.09.013 (2014).

39 Myronovskyi, M., Rosenkranzer, B. & Luzhetskyy, A. Iterative marker excision system. *Appl. Microbiol. Biotechnol.* **98**, 4557-4570, doi:10.1007/s00253-014-5523-z (2014).

40 Horbal, L., Fedorenko, V., Bechthold, A. & Luzhetskyy, A. A transposon-based strategy to identify the regulatory gene network responsible for landomycin E biosynthesis. *FEMS. Microbiol. Lett.* **342**, 138-146, doi:10.1111/1574-6968.12117 (2013).

41 Blodgett, J. A. *et al.* Common biosynthetic origins for polycyclic tetramate macrolactams from phylogenetically diverse bacteria. *Proc. Natl. Acad. Sci. USA.* **107**, 11692-11697, doi:10.1073/pnas.1001513107 (2010).

42 Xu, L. X., Wu, P., Wright, S. J., Du, L. C. & Wei, X. Y. Bioactive polycyclic tetramate macrolactams from *Lysobacter enzymogenes* and their absolute configurations by theoretical ECD calculations. *J. Nat. Prod.* **78**, 1841-1847, doi:10.1021/acs.jnatprod.5b00099 (2015).

43 Zhang, G. T. *et al.* Mechanistic insights into polycycle formation by reductive cyclization in ikarugamycin biosynthesis. *Angew. Chem. Int. Ed.* **53**, 4840-4844, doi:10.1002/anie.201402078 (2014).

44 Xie, Y. X., Wright, S., Shen, Y. M. & Du, L. C. Bioactive natural products from *Lysobacter*. *Nat. Prod. Rep.* **29**, 1277-1287, doi:10.1039/c2np20064c (2012).

45 Li, Y. *et al.* Iterative assembly of two separate polyketide chains by the same single-module bacterial polyketide synthase in the biosynthesis of HSAF. *Angew. Chem. Int. Ed.* **53**, 7524-7530, doi:10.1002/anie.201403500 (2014).

46 Green, M. R. & Sambrook, J. *Molecular cloning : a laboratory manual*. 4th edn, (Cold Spring Harbor Laboratory Press, 2012).

47 Kieser T., B. M. J., Buttner M.J., Charter K.F., Hopwood D. *Practical Streptomyces Genetics.* (John Innes Foundation, 2000).

48 Rebets, Y. *et al.* Production of landomycins in *Streptomyces globisporus* 1912 and *S. cyanogenus* S136 is regulated by genes encoding putative transcriptional activators. *FEMS. Microbiol. Lett.* **222**, 149-153, doi:Doi 10.1016/S0378-1097(03)00258-1 (2003).

49 Rebets, Y. *et al.* Complete genome sequence of producer of the glycopeptide antibiotic Aculeximycin *Kutzneria albida* DSM 43870(T), a representative of minor genus of *Pseudonocardiaceae*. *Bm.c Genomics* **15**, doi:10.1186/1471-2164-15-885 (2014).

50 Durr, C. *et al.* Biosynthesis of the terpene phenalinolactone in *Streptomyces* sp. Tu6071: analysis of the gene cluster and generation of derivatives. *Chemistry & Biology* **13**, 365-377, doi:10.1016/j.chembiol.2006.01.011 (2006).

51 Rebets, Y. *et al.* Insights into the pamamycin biosynthesis. *Angew. Chem. Int. Ed.* **54**, 2280-2284, doi:10.1002/anie.201408901 (2015).

52 Rebets, Y., Kormanec, J., Lutzhetskyy, A., Bernaerts, K. & Anne, J. Cloning and Expression of Metagenomic DNA in *Streptomyces lividans* and Subsequent Fermentation for Optimized Production. *Methods Mol. Biol.* **1539**, 99-144, doi:10.1007/978-1-4939-6691-2_8 (2017).

53 Fedoryshyn, M., Welle, E., Bechthold, A. & Luzhetskyy, A. Functional expression of the Cre recombinase in actinomycetes. *Appl. Microbiol. Biot.* **78**, 1065-1070, doi:10.1007/s00253-008-1382-9 (2008).

54 Gust, B. *et al.* lambda red-mediated genetic manipulation of antibiotic-producing *Streptomyces*. *Adv. Appl. Microbiol.* **54**, 107-128, doi:Doi 10.1016/S0065-2164(04)54004-2 (2004).

55 Luzhetskyy, A. *et al.* Generation of novel landomycins M and O through targeted gene disruption. *ChemBioChem* **6**, 675-678, doi:10.1002/cbic.200400316 (2005).

2.7. Supplementary Information

Table 1S. Strains and plasmids used in this work.

Strains	**Features**	**Reference or source**
Streptomyces strains		
S. albus J1074	Wild-type strain	[1]
S. albus ATGSal2P2	*S. albus* J1074 harboring pGUSATGSal2P2 construct	This work
S. albus ATGSal2P2::TN1	*S. albus* ATGSal2P2 strain with *Himar1* minitransposon insertion	This work
S. albus ATGSal2P2::TN5	*S. albus* ATGSal2P2 strain with *Himar1* minitransposon insertion	This work
S. albus ATGSal2P2::TN13	*S. albus* ATGSal2P2 strain with *Himar1* minitransposon insertion	This work
S. albus ATGSal2P2::TN14	*S. albus* ATGSal2P2 strain with *Himar1* minitransposon insertion	This work
S. albus Δ3174	*S. albus* J1074 strain with deletion of *XNR_3174* gene	This work
S. albus Δ3174 pUWLH3174	*S. albus* Δ3174 strain harboring pUWLH3174 construct	This work
S. albus Tn14Δ2339	*S. albus* ATGSal2P2Tn14 strain with deletion of *XNR_2339* gene	This work
S. albus Tn14Δ2339 pUWLT2339	*S. albus* ATGSal2P2Tn14 Δ2339 strain with pUWLT2339 construct	This work
S. albus Δ3174 Δ2339	*S. albus* Δ3174 strain with deletion of *XNR_2339* gene	This work
S. albus J1074 pGUS*P*aco	*S. albus* J1074 strain with pGUS*P*aco construct	This work
S. albus Δ3174 pGUS*P*aco	*S. albus* Δ3174 with pGUS*P*aco construct gene	This work
S. albus Δ4681	*S. albus* J1074 strain with deletion of *XNR_4681* gene	This work
E. coli strains		
DH10β	General cloning host	[2]
ET12567 pUB307	Donor strain for intergeneric conjugation	[3]
TransforMaxTMEC100DTM pir-116	*E.coli* with *pir* gene for replication of plasmids containing the *R6Kγ* origin of replication	Epicenter
Plasmids		
pHTM	*Himar1* minitransposon delivery plasmid.	Bilyk et al. (2013)
pUWL-Dre	Plasmid containing synthetic *dre(a)* gene under the *tipA* promoter, pSG5rep and *oriT*	Fedoryshyn (2008b)
patt-saac-oriT	Resistance cassette plasmid containing a synthetic fragment with *aac(3)IV*, *oriT*, *B-CC*, *P-GG* and *loxP* sites	Myronovskyi et al. (2014)
patt-shyg-oriT	Resistance cassette plasmid containing a synthetic fragment with *hyg*, *oriT*, *B-CC*, *P-GG* and *loxP* sites	Myronovskyi et al. (2014)
pUWLHFLP	Replicative vector with pIJ101 replicon, *oriT*, *tsr*, *bla* and *hyg* resistance genes with the *flp(a)* gene under an *ermE* promoter	Dr.Fedoryshyn personal communication
pUWLH3174	pUWLHFLP derivative with *XNR_3174* gene cloned under control of *ermE* promoter and replacing *flp* gene	This work
pUWLoriT	Replicative vector with pIJ101 replicon, *oriT*, *tsr*, *bla* resistance genes.	[4]

pUWLT2339	pUWLoriT derivative with *XNR_2339* gene cloned under control of *ermE* promoter	This work
pGUS-MF-PVV	Plasmid containing a synthetic *gusA* gene flanked with transcriptional terminators, *aac(3)IV, aadA, oriT, intøC31.*	Dr.Manderscheid, personal communication
pGUSATGSal2P2	pGUS-MF-PVV with promoter of *XNR_0204* gene fused with *gusA* reporter	This work
pGUS*P*aco	pGUS-MF-PVV with promoter of *XNR_2339* gene fused with *gusA* reporter	This work
BACs		
pSMART	Cm^R; BAC vector	Lucigen
pSMARTgus	Derivative of pSMART with the *gusA* gene	Dr. Myronovskyi, personal communication
1G13	pSMARTgus derivative containing a fragment of the *S. albus* chromosome	[5]
2C15	pSMARTgus derivative containing a fragment of the *S. albus* chromosome	[5]
2L15	pSMARTgus derivative containing a fragment of the *S. albus* chromosome	[5]
1G13am	Derivative of 1G13 with *XNR_3174* gene substituted with a cassette from patt-saac-oriT	This work
2C15hyg	Derivative of 2C15 with *XNR_2339* gene substituted with a cassette from patt-shyg-oriT	This work
2L15am	Derivative of 2L15 with *XNR_4681* gene substituted with a cassette from patt-saac-oriT	This work
p41-2C-06	pOJ436 derivative, containing aranciamycin biosynthetic gene cluster	[6]
PMP31	pOJ436 derivative, containing the griseorhodin biosynthetic cluster	[7]
pIJ12003a	12.9 Kbp *tun*-gene cluster cloned into the pRT802	[8]
pOJR2	pOJ436 with cloned *pam*-gene cluster	[9]

Table 2S. Primers used in this work.

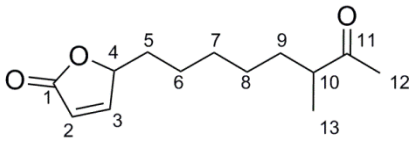
Primers	Sequence	Purpose
Sal2P2F	ATTGGTACCGTTGATCGCGTCAGCCAAGT	*XNR_0204* promoter cloning
Sal2P2R	ATTGATATCTCCACGACCCCACGACGTT	
3174-F	TCCGGCCGGATTCCGCCGGGCCCCGGGCTCAGTCCGTCTATTCCGGGGATCCGTCGACC	Deletion of *XNR_3174*
3174-R	GTTCCTGTAATGCATGACCCTGAGGGGGGTTCTTATGGGTGTAGGCTGGAGCTGCTTC	
2339-F	CATGACTCCCTGGCGACCGTTACTTTACAAAGTGTGATTACTACGCCCCCAACTGAGAG	Deletion of *XNR_2339*
2339-R	CCGTCCCGCGTCCCGCCGTCGTGCGCCTCAGCCCGTCATTCGACCCGGTACCGGAGTA	
3174C_ F1	AGCGGATACAGGCATCATGT	Verification of *XNR_3174* deletion
3174C_R1	GGAAGCAGTACGTACTGCTT	
2339C_F	ACTGGCTTGAACTGGTCACT	Verification of *XNR_2339* deletion
2339C_R	GGAGGACGAGAGACGCGA	
3174E_HindIIIR	AAAAGCTTGAGGGGGGTTCTTATGGGG	Expression of *XNR_3174*
3174E_BamHIF	AAGGATCCCCCGGGCTCAGTCCGTCT	
2339E-F	TTGGATCCCCGTTACTTTACAAAGTGTG	Expression of *XNR_2339*
2339E-R	TTAAGCTTTGCGCCTCAGCCCGTCAT	
GUS-Xbal F	AATCTAGATTCCGCTTTTCGCCCTGG	*XNR_2339* promoter cloning
GUS-KpnI R	AAGGTACCTGTTCTTCGCCGCGTGTCG	
p3-pALG-ch	GTGAGCCGCCGCGTGCCGTCG	*Himar1* transposon sequencing primer
4681-F	TCGTCACGGAAACGTTGAACGAAGGTGGACAGCACGTTGCTACGCCCCCAACTGAGAG	Deletion of *XNR_4681*
4681-R	GCGATGTCACCGGAGGGAGCCCGGTCGCGTCGTCACTCC TCGACCCGGTACCGGAGTA	
4681C-F	TCGAATGGCCAGCAACCG	Verification of *XNR_4681* deletion
4681C-R	GTCTGCCGACTCCGCTCT	

Table 4S. Localization of transposon insertion in the chromosome of selected *S. albus* ATGSal2p2 mutants

Strain name.	Gene ID.	Position of insertion.	Putative gene product.
ATGSal2p2::Tn1	*XNR_2800*	3,180,172	Putative type IV secretory pathway
	XNR_3855	4,381,847	Putative glutaminase
ATGSal2p2::Tn5	*XNR_4186*	4,745,403	Putative hypothetical protein upstream
ATGSal2p2::Tn13	*XNR_3521*	4,038, 469	Putative MerR family transcriptional regulator
ATGSal2p2::Tn14	*XNR_3174*	3,624,446	Putative LuxR family transcriptional regulator

Table 5S. NMR spectroscopic data for butenolide 4 in $CDCl_3$ (700 MHz, 25 °C)

Pos.	**δ_C**	**δ_H, mult, (*J* in Hz)**	**COSY**	**HMBC**
1	173.4	-	-	2-H, 3-H
2	121.9	6.10, dd (5.7, 1.7)	4-H, 3-H	^1J, 4-H, 3-H
3	156.5	7.43, ddd (5.5, 5.7, 1.4)	2-H, 4-H	^1J, 5-H_2, 2-H, 4-H
4	83.6	5.02, ddd (5.5, 1.7, 7.0)	2-H, 3-H, 5-H_2	5-H_2, 2-H, 3-H
5	33.2	1.76, m 1.64, m	5-H_b, 4-H, 6-H_2 5-H_a, 4-H, 6-H_2	4-H, 6-H_2, 7-H_2
6	25.0	1.44, m	(5-H_a)	^1J, 4-H, 5-H_2, 7-H_2,
7	29.5	1.33, m	8-H, 9-H_a	5-H_b, 9-H_a
8	27.1	1.25, m	7-H_2, 9-H_b	7-H_2, 9-H_a, 10-H
9	32.8	1.64, m 1.31, m	9-H_b, 8-H, 10-H 9-H_a, 10-H	8-H_2, 10-H, 13-H_3
10	47.3	2.49, m	9-H_2, 13-H_3	9-H_2, 12-H_3, 13-H_3
11	212.6	-	-	9-H_2, 10-H, 12-H_3, 13-H_3
12	28.2	2.13, s	-	^1J
13	16.5	1.08, d (7.2)	10-H	^1J, 9-H_2, 10-H



Compound Y3 (0.4 mg) was obtained as oily substance and showed molecular ion *m*/*z* = 225.1483 $[M+H]^+$ corresponding to a molecular formula of $C_{13}H_{21}O_3$ for the protonated ion, UV (AcN) λmax. 208 nm. NMR data and literature survey determined the compound as 5-(6-Methyl-7-oxooctyl)-2(*5H*)-furanone. The chemical shift δ_C = 173,4 ppm is pointing on carbonyl group at C-1. C-1 C-1 is monitoring C-1/2-H and C1/3-H HMBC cross-peaks showing the connection of C-1 to the methines which couple to each other in ^{1}H spectrum ^3J = 5,7 Hz and in COSY spectrum (2-H/3-H; 3-H/2-H). Further 2-H shows COSY cross-peak and ^4J = 1,7 Hz to 4-H. 3-H shows COSY cross-peaks 3-H/4-H and ^3J = 5,5 Hz to 4-H. Methine C-4 (δ_C = 83,6 ppm) supports the furanone structure (HMBC signals 4-H/3-H and 4-H/2-H). The side-chain at C-4 is defined by COSY (4-H/5-H_2) and HMBC (C-4/5-H_2) signals. Methylenes 5-H_2–9-H_2 in the chain shows multiplet signals due to coupling to their neighbors. COSY and HMBC signals lead through the chain sequence. The first methylene group 5-H_2 (δ_H = 1,76 ppm and 1,64 ppm) shows COSY cross-peaks 5-H_2/6-H_2, 5-H_2/4-H and HMBC-signals C-5/6-H_2 and C-5/7-H_2. 7-H_2 monitors HMBC signal C-7/5-H_2 supporting the connection to the left chain sequence and HMBC signal C-7/9-H_2 in combination with COSY cross-peaks to 8-H_2 and 9-H_2 supporting the connection to the right chain sequence. Methylene group 9-H_2 (δ_H = 1,64 ppm and 1,31 ppm) shows COSY coupling 9-H_2/10-H and the HMBC signal C-9/13-H_3. 13-H_3 methyl group (δ_C = 16,5 ppm) shows a doublet with ^3J = 7,2 Hz and a COSY cross-peak 13-H3/10-H determining the 10-H as a neighbor. The HMBC signal C-13/9-H_2 clearly assigns the connection to the chain. The positon of the carbonyl (δ_C = 212,6 ppm) is determined based on several HMBC signals C-11/9-H_2, C-11/10-H and C-11/13-H_3. Finally, the single methyl group C-12 is showing no coupling but the position is supported by HMBC signals C-10/12-H_3 and C-11/12-H_3.

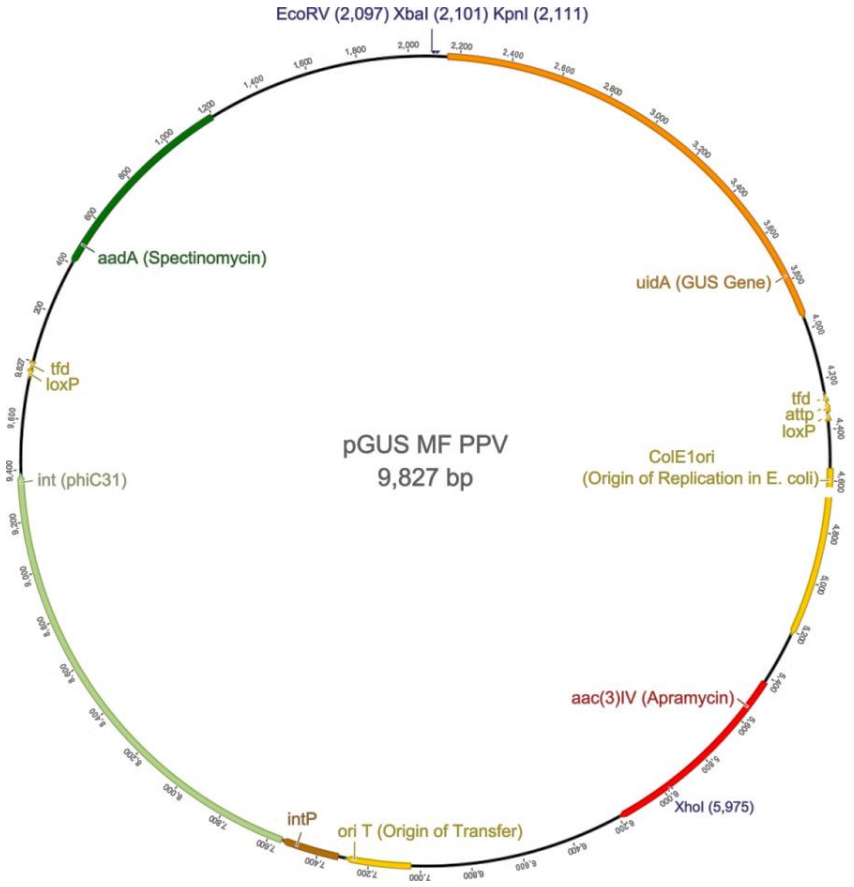


Figure 1S. Map of the plasmid pGUS MF PPV used in this work (Dr. Niko Manderscheid, personal communication). The plasmid carries the reporter gene *gusA* coding for β-glucuronidase with the ATG start codon. Components of the vector are shown with arrows: *uidA* – synthetic *gusA* gene with codon usage adapted for Actinobacteria [10]; *aac(3)IV* – apramycin resistance gene; *aadA* – spectinomycin resistance gene; *oriT* – origin of transfer; *attP* – phage øC31 attachment site; *int* – phage øC31 integrase gene; *loxP* – Cre recombinase recognition sites. The reporter construct is flaked with the t_{fd} transcriptional terminators.

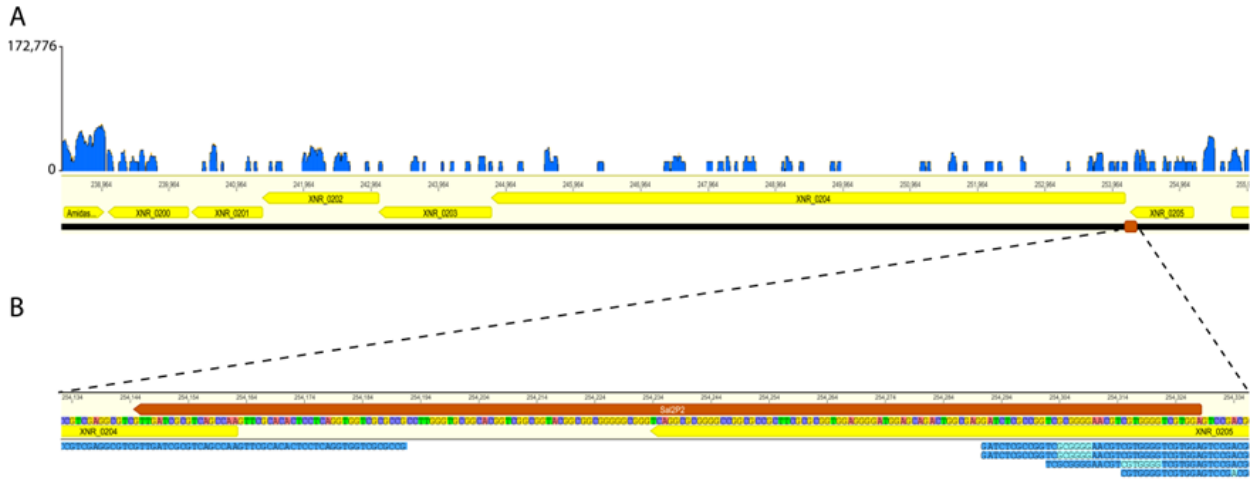


Figure 2S. Transcriptional profile of the *S. albus* J1074 secondary metabolites biosynthesis gene cluster 2 based on RNAseq data. The gene cluster is poorly transcribed in all tested conditions resulting in a very low number of reads mapped. The data used in this experiment was obtained from *S. albus* J1074 culture grown in NL19 medium. **A.** Overall organization of BGC 2. Yellow arrows indicate the size and location of individual genes. The coordinates of the region within the *S. albus* genome are shown above the arrows. The blue plot on top shows the coverage (the number of non-end gap characters at each position). The scale bar indicates the mean coverage level (log scale). The highest coverage is shown. **B.** Zoomed-in representation of *XNR_0204* gene promoter region. Fragment, cloned into the promoter-probing vector is indicated with brown arrow. Sequences highlighted in blue show the RNAseq reads mapped to the corresponding region of the *S. albus* genome. The data analysis, reads mapping and visualization was performed using Geneious software, version 8.1.7 (Biomatters Ltd, New Zealand).

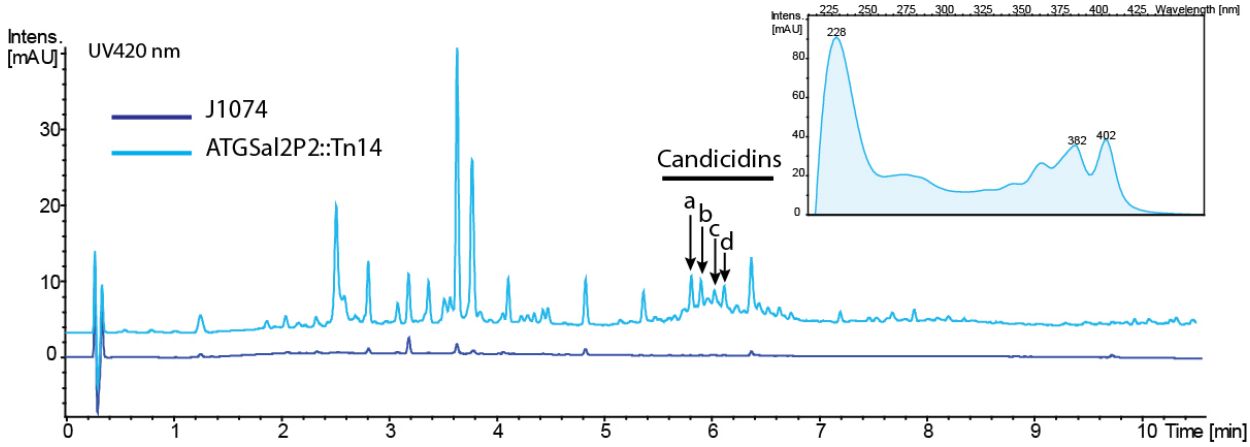


Figure 4S. LC-MS chromatogram (at 420 nm) of secondary metabolites extracts of *S. albus* J1074 and *S. albus* ATGSal2p2::Tn14 strains grown in NL19 medium. Samples were separated with the 10 min gradient protocol (see Materials and Methods). Compounds identified as candicidins are indicated by arrows and marked as: **a** – candicidin I (detected m/z 1111.59 $[M+H]^+$, calculated m/z 1111.5952 $[M+H]^+$), **b** and **c** – candicidins II and III (detected m/z 1109.58 $[M+H]^+$, calculated m/z 1109.5792 $[M+H]^+$), and **d** – candicidin IV (detected m/z 1093.58 $[M+H]^+$, calculated m/z 1093.5842 $[M+H]^+$). Detected m/z values corresponds to the described one for candicidins.

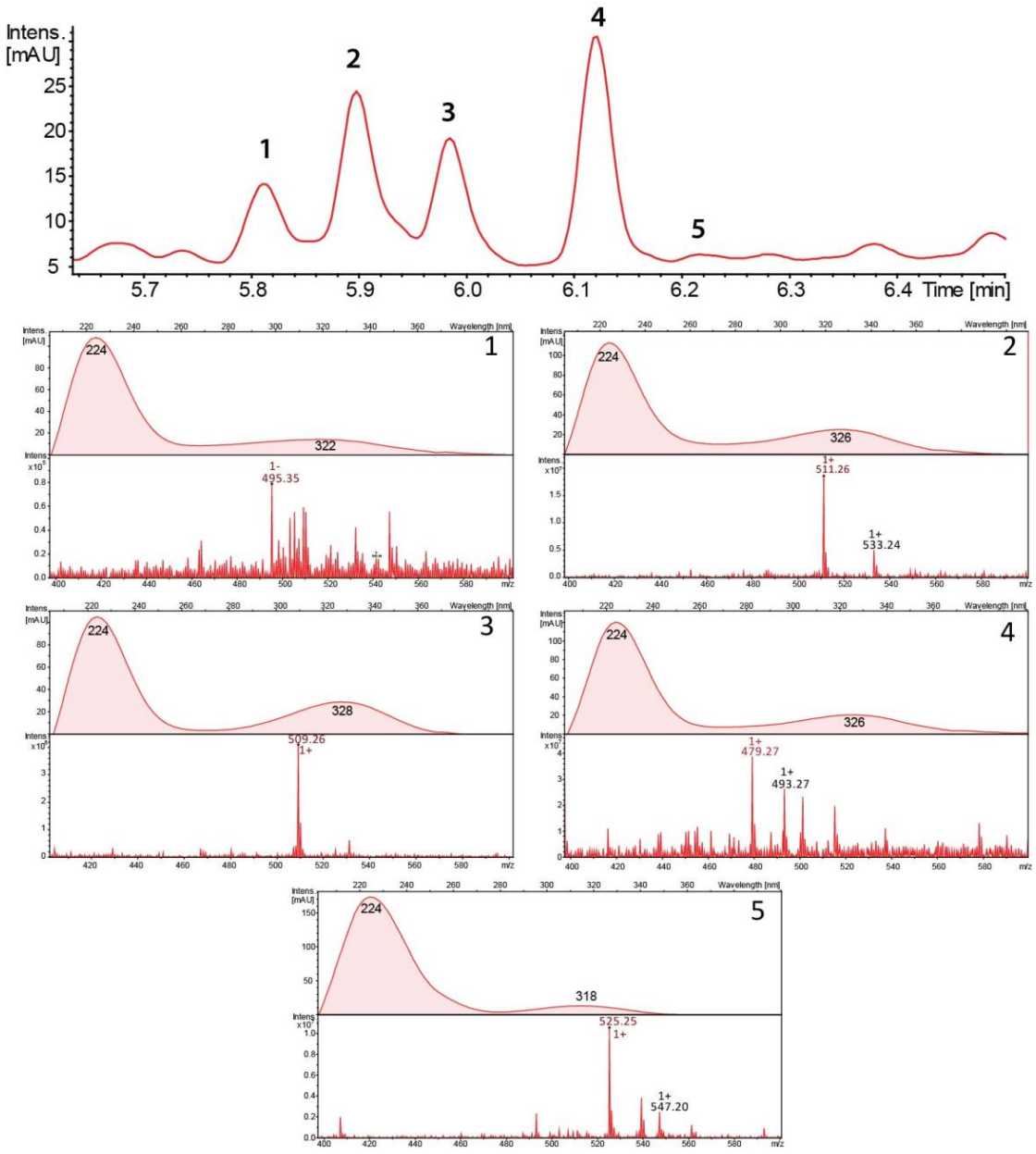


Figure 5S. LC-MS based identification of polycyclic tetramate mactolactams produced by the *S. albus* ATGSal2P2::Tn14 strain. Initial extract from 5L of *S. albus* ATGSal2P2::Tn14 culture (concentrated x 250 folds to final volume of 20 ml in methanol) grown in NL19 medium was fractionated by size exclusion chromatography and collected fractions were analyzed by LC-MS with the 10 minutes gradient protocol. Fraction 11 (shown; chromatogram at 320 nm) contains several compounds with spectral characteristics typical for polycyclic tetramate mactolactams. The m/z values are shown for each peak.

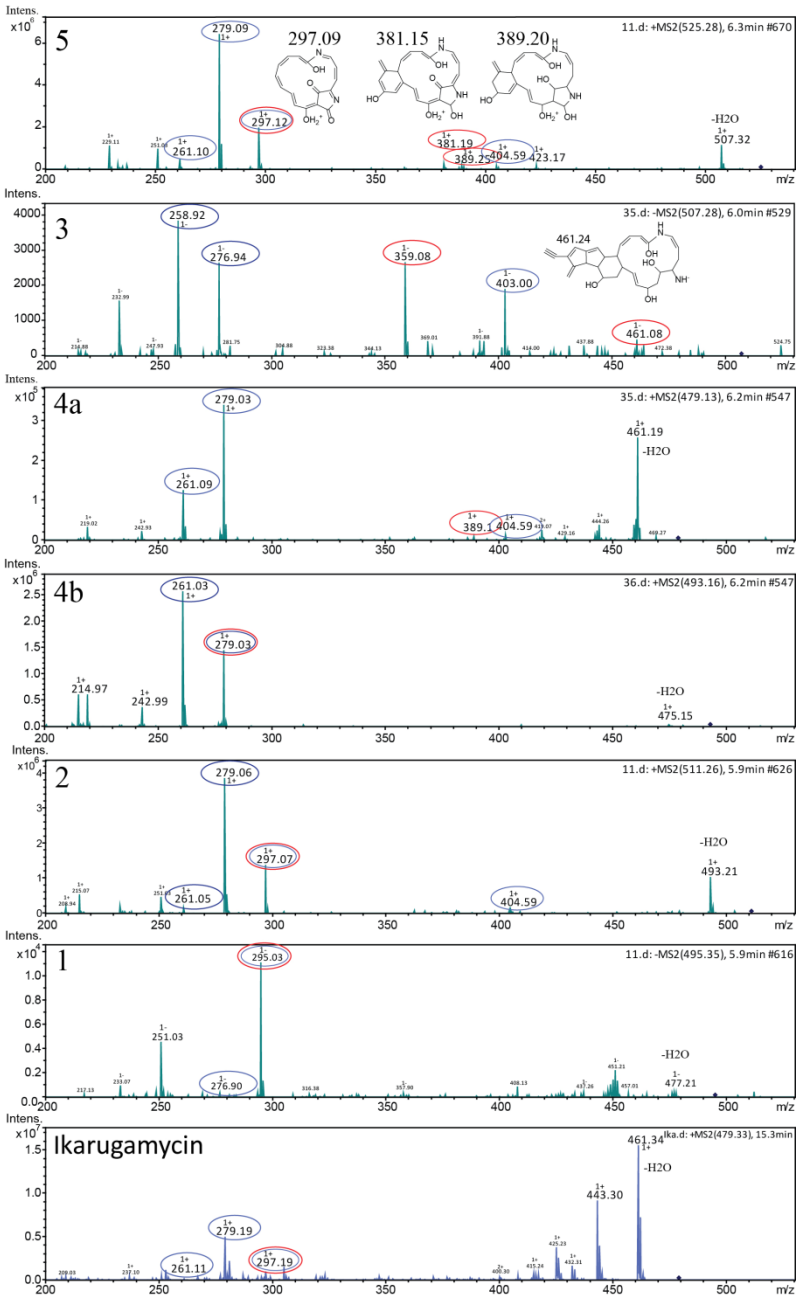


Figure 6S. MS/MS fragmentation pattern of compounds indetified in the extract of *S. albus* ATGSal2P2::Tn14 as PTMs and pure sample of ikarugamycin (Sigma-Aldrich, USA). Fragments marked with blue are common for all studied metabolites, including ikarugamycin. Fragments marked with red match the predicted fragmentation by CFM-ID software for particular compound [11]. Compound **5** fragmentation is similar to predicted for positive ion of frontalamide A. Compound **3** fragmentation is similar to predicted for negative ion of frontalamide B. Possible structures of fragments are shown (structures were generated by CFM-ID) Fragments with m/z of 261 $[M+H]^+$, 279 $[M+H]^+$ and 297 $[M+H]^+$ can be found in patterns of alteramide A and B deposited in GNPS Library (alteramide A ID: CCMSLIB00000077249; B: CCMSLIB00000077250) [12]. These fragments seem to be common for PTMs family. Mass-spec data was collected on LC-MS amaZon speed system (Bruker Daltonics, Germany).

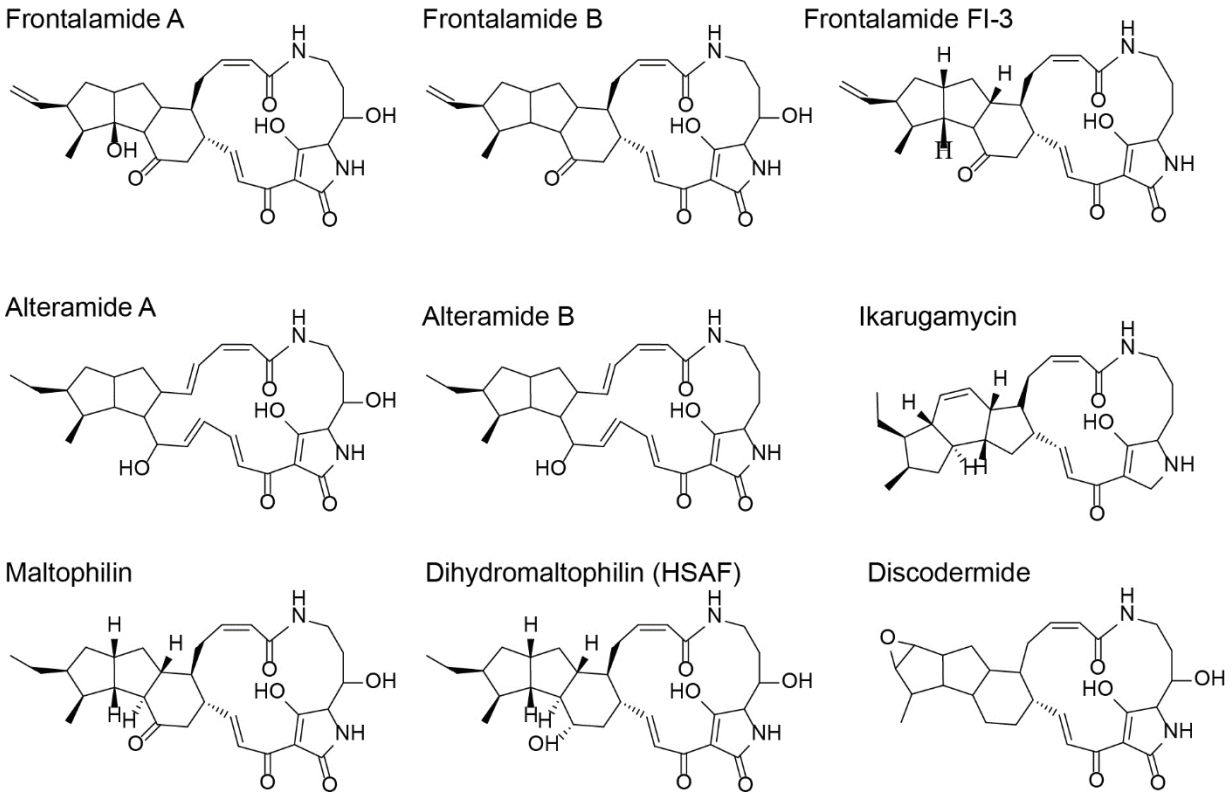


Figure 7S. Structures of several representatives of polycyclic tetramate mactolactams.

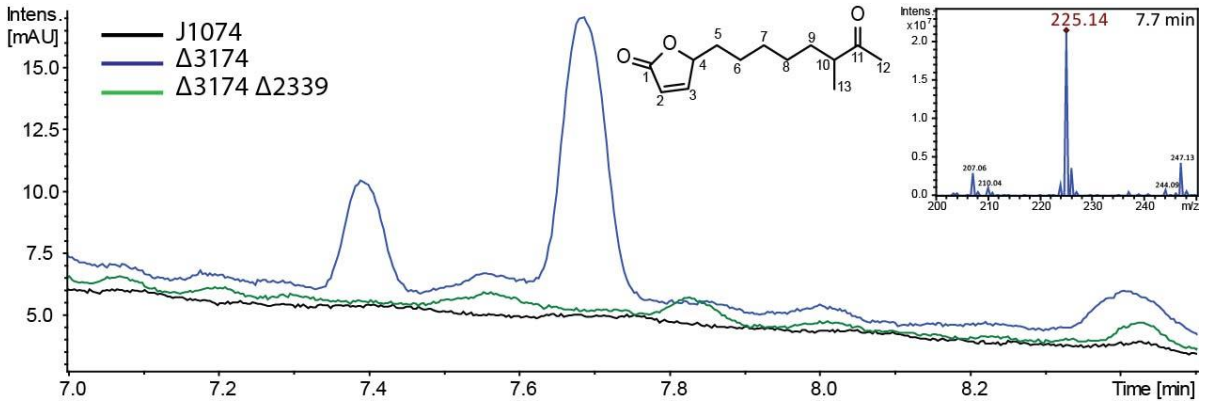


Figure 8S. Production of butenolide 4 by strains *S. albus* J1074 (black trace), *S. albus* Δ3174 (blue) and *S. albus* Δ3174Δ2339 (green). Chromatogram of 20 min gradient protocol (see Materials and Methods) at 320 nm is shown. Compound with retention time of 7.7 min was purified and its structure was proved with the NMR.

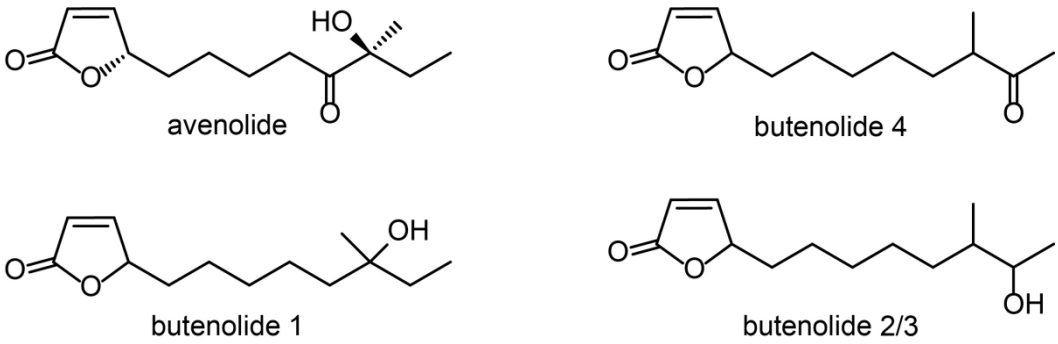


Figure 9S. Structures of avenolide from *S. avermitilis*, butenolide 4 from *S. albus*, and butenolides 1-4 from marine *Streptomyces* sp. SM8 and *S.* sp. B3497 [13,14].

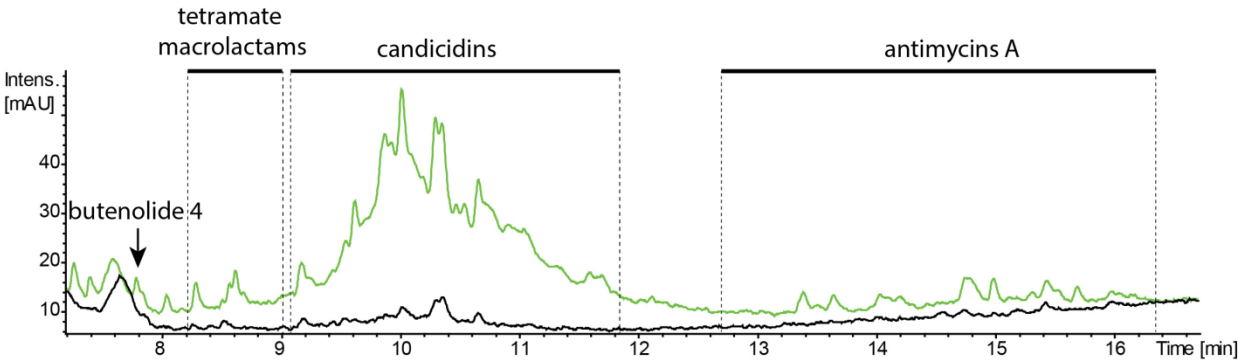


Figure 10S. LC-MS chromatogram (at 320 nm) of secondary metabolites extracts of *S. albus* J1074 (black) and *S. albus* J1074/pUWLH2339 (green) strains grown in NL19 medium. Samples were separated with the 20 min gradient protocol. Identified compounds are highlighted.

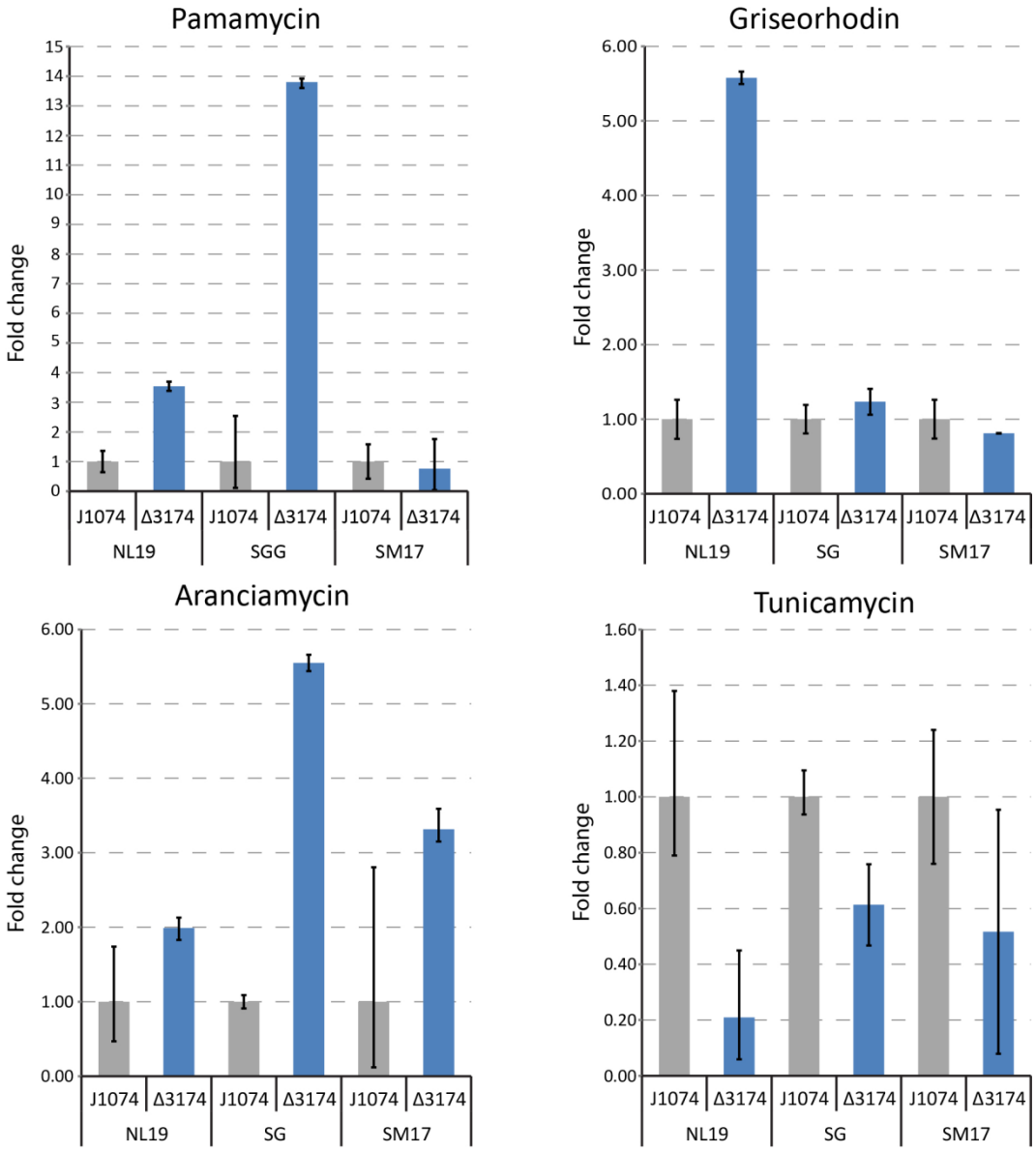


Figure 11S. Heterologous production of different secondary metabolites in *S. albus* J1074 and Δ3174.

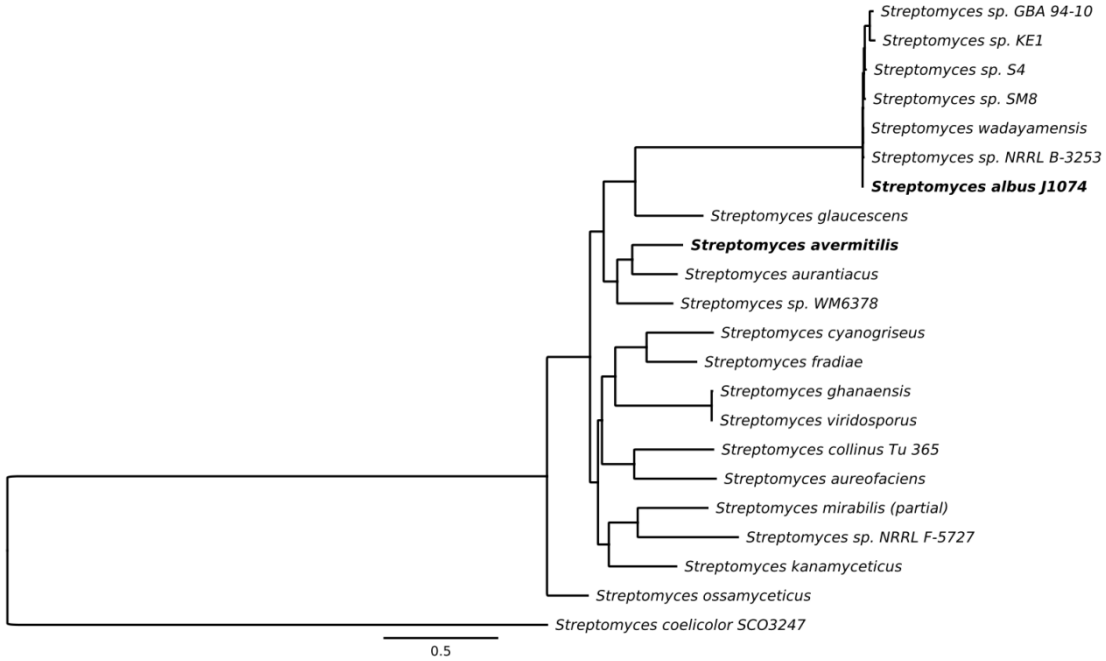


Figure 12S. Phylogeny of avenolide/butenolide biosynthesis acyl-CoA oxidase protein from different Actinobacteria. Closest homologue from S. *coelicolor* was used as outgroup. The analysis and visualization was performed using Geneious software, version 8.1.7 (Biomatters Ltd, New Zealand).

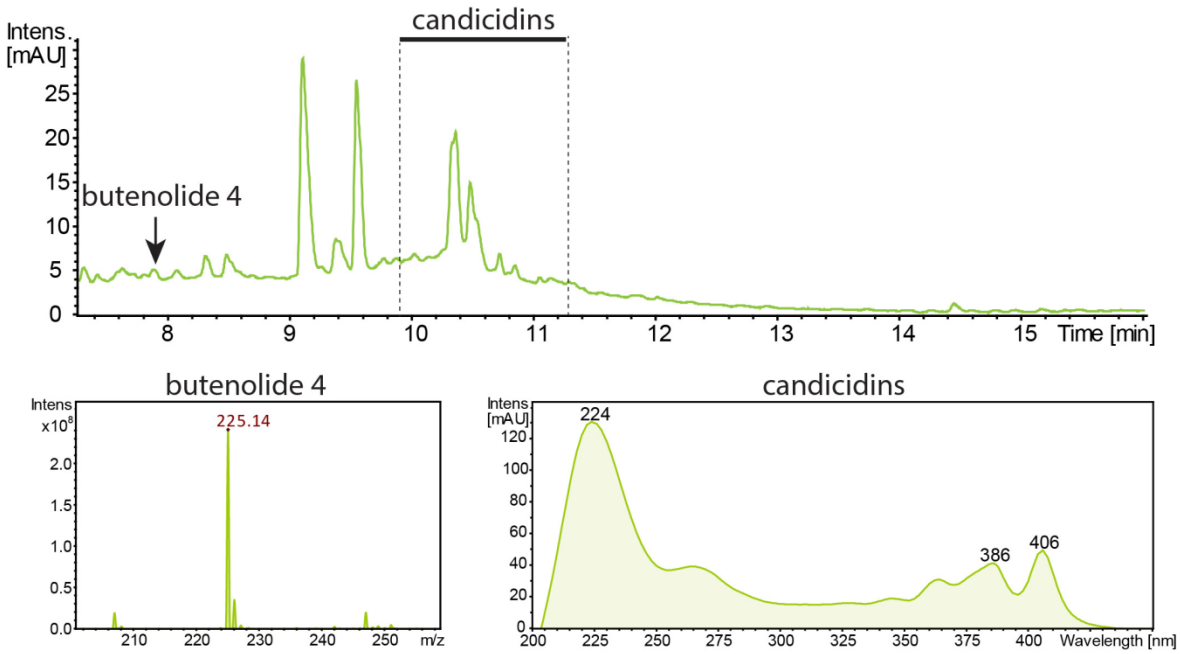


Figure 13S. LC-MS chromatogram (at 320 nm) of secondary metabolites extracts of *Streptomyces albidoflavus* NRRL B-1271 strain grown in NL19 medium. Sample was separated with the 20 min gradient protocol. Identified compounds are highlighted.

2.8. References for the Supplementary Information

1 Chater, K. F. & Wilde, L. C. Restriction of a bacteriophage of *Streptomyces albus* G involving endonuclease *Sal*I. *J. Bacteriol.* **128**, 644-650 (1976).

2 Grant, S. G., Jessee, J., Bloom, F. R. & Hanahan, D. Differential plasmid rescue from transgenic mouse DNAs into *Escherichia coli* methylation-restriction mutants. *Proc. Natl. Acad. Sci. USA* **87**, 4645-4649 (1990).

3 Flett, F., Mersinias, V. & Smith, C. P. High efficiency intergeneric conjugal transfer of plasmid DNA from *Escherichia coli* to methyl DNA-restricting Streptomycetes. *FEMS Microbiol. Lett.* **155**, 223-229 (1997).

4 Luzhetskyy, A. *et al.* Generation of novel landomycins M and O through targeted gene disruption. *ChemBioChem* **6**, 675-678, doi:10.1002/cbic.200400316 (2005).

5 Zaburannyi, N., Rabyk, M., Ostash, B., Fedorenko, V. & Luzhetskyy, A. Insights into naturally minimised *Streptomyces albus* J1074 genome. *BMC Genomics* **15**, doi:10.1186/1471-2164-15-97 (2014).

6 Luzhetskyy, A. *et al.* Cloning and heterologous expression of the aranciamycin biosynthetic gene cluster revealed a new flexible glycosyltransferase. *ChemBioChem* **8**, 599-602, doi:10.1002/cbic.200600529 (2007).

7 Li, A. & Piel, J. A gene cluster from a marine *Streptomyces* encoding the biosynthesis of the aromatic spiroketal polyketide griseorhodin A. *Chemistry & Biology* **9**, 1017-1026 (2002).

8 Wyszynski, F. J. *et al.* Biosynthesis of the tunicamycin antibiotics proceeds via unique exo-glycal intermediates. *Nat. Chem.* **4**, 539-546, doi:10.1038/nchem.1351 (2012).

9 Rebets, Y. *et al.* Insights into the pamamycin biosynthesis. *Angew. Chem. Int. Ed.* **54**, 2280-2284, doi:10.1002/anie.201408901 (2015).

10 Myronovskyi, M., Welle, E., Fedorenko, V. & Luzhetskyy, A. beta-Glucuronidase as a sensitive and versatile reporter in Actinomycetes. *Appl. Environ. Microbiol.* **77**, 5370-5383, doi:10.1128/Aem.00434-11 (2011).

11 Allen, F., Pon, A., Greiner, R. & Wishart, D. Computational Prediction of Electron Ionization Mass Spectra to Assist in GC/MS Compound Identification. *Anal. Chem.* **88**, 7689-7697, doi:10.1021/acs.analchem.6b01622 (2016).

12 Wang, M. *et al.* Sharing and community curation of mass spectrometry data with Global Natural Products Social Molecular Networking. *Nat. Biotechnol.* **34**, 828-837, doi:10.1038/nbt.3597 (2016).

13 Mukku, V. J., Speitling, M., Laatsch, H. & Helmke, E. New butenolides from two marine Streptomycetes. *J. Nat. Prod.* **63**, 1570-1572 (2000).

14 Viegelmann, C. *et al.* Metabolomic profiling and genomic study of a marine sponge-associated *Streptomyces* sp. *Marine Drugs* **12**, 3323-3351, doi:10.3390/md12063323 (2014).

Chapter 3

Multi-Omics and Targeted Approaches to Determine the Role of Cellular Proteases in *Streptomyces* Protein Secretion

Busche, Tobias[1,6*], Tsolis, C. Konstantinos[2*], Koepff, Joachim[3], Rebets, Yuriy[4,5], Rückert, Christian[1], Hamed, B. Mohamed[2,7], Bleidt, Arne[3], Wiechert, Wolfgang[3], Lopatniuk, Mariia[5], Ahmed, Yousra[5], Jozef Anné[2], Karamanou, Spyridoula[2], Oldiges, Marco[3,8], Kalinowski, Jörn [1], Luzhetskyy, Andriy[4] and Economou, Anastassios[2,9].

Journal of Fronteries Microbiology. 2018 Jun 4; 9:1174.

doi: 10.3389/fmicb.2018.01174. eCollection 2018

Published online 04 June 2018

Supplementary Material for this article can be found online at:
https://www.frontiersin.org/articles/10.3389/fmicb.2018.01174/full#supplementary-material

3. Proteases in *Streptomyces* Secretion

3.1. Abstract

Gram positive *Streptomyces* bacteria are profuse secretors of polypeptides using complex, yet unknown mechanisms. Many of their secretory proteins are proteases that play important roles in the acquisition of amino acids from the environment. Other proteases regulate cellular proteostasis. To begin dissecting the possible role of proteases in *Streptomyces* secretion, we applied a multi-omics approach. We probed the role of the 190 proteases of *Streptomyces lividans* strain TK24 in protein secretion in defined media at different stages of growth. Transcriptomics analysis revealed transcripts for 93% of these proteases and identified that 41 of them showed high abundance. Proteomics analysis identified 57 membrane-embedded or secreted proteases with variations in their abundance. We focused on 17 of these proteases and putative inhibitors and generated strains deleted of their genes. These were characterized in terms of their fitness, transcriptome and secretome changes. In addition, we performed a targeted analysis in deletion strains that also carried a secretion competent mRFP. One strain, carrying a deletion of the gene for the regulatory protease FtsH, showed significant global changes in overall transcription and enhanced secretome and secreted mRFP levels. These data provide a first multi-omics effort to characterize the complex regulatory mechanisms of protein secretion in *Streptomyces lividans* and lay the foundations for future rational manipulation of this process.

3.2. Introduction

Microorganisms of the family *Streptomycetaceae* are filamentously growing key players in soil habitats all around the planet, where they vitally contribute as decomposers to recycle organic material (Barka et al., 2016; Ranjani et al., 2016; Hopwood, 2007). To fulfill this duty, *Streptomyces* produce and secrete a large arsenal of extracellular enzymes including proteases, enabling them to exploit complex proteinogenic resources (Chater, 2016).

Extracellular peptide-cleaving enzymes have a dual function in heterologous protein production. While on one hand being necessary for protein folding and essential steps in the secretion machinery (Gilbert et al., 1995; Neef et al., 2017), proteases are also a challenge to industrial protein production (van den Hombergh et al., 1997), due to their undesired capacity to degrade the recombinant product itself. Therefore, reducing the extracellular protease activity by gene deletions is a typical path towards higher protein production titers and yields in other industrial relevant organisms, such as *Aspergillus* (van den Hombergh et al., 1997; Xu et al., 2000),

Bacillus (Pohl et al., 2013) and in higher eukaryotic systems like insect cell cultures (Gotoh et al., 2001).

S. lividans, a well-investigated member of the *Streptomyces* family, already innately displays a rather low extracellular protease activity in comparison to other related species (Butler et al., 1993; Gilbert et al., 1995; Liu et al., 2013), while at the same time maintaining a highly active secretion machinery (Anne et al., 2017) and being easily manipulated genetically. Therefore, this strain has become an important model organism for heterologous protein production (Gilbert et al., 1995; Anne et al., 2012; Chater, 2016).

In addition to the extracellular environment, proteases play essential roles within the membrane border of the cell. Amongst these functions are the recycling of misfolded proteins, degradation of unused enzymes as well as general housekeeping tasks (Krishnappa et al., 2013). Regulated proteolysis is a post-translational mechanism with a direct influence on the amount of certain proteins (Langklotz et al., 2012). Regulatory proteolysis is accomplished in Gram-negative bacteria by five ATP-dependent proteases: ClpAP, ClpXP, Lon, HslUV and FtsH, and three other proteases: ClpCP, ClpEP and the proteasome (Gur et al., 2011). Deletion of these regulatory proteases may affect the proteome of the cell. For example, deletion of FtsH strongly increases the abundance of ten cytoplasmic and membrane proteins in *Corynebacterium glutamicum* without effect on its growth (Ludke et al., 2007).

Here, we undertook a broad analysis of TK24 cellular proteases using a transcriptomics and secretomics approach to define proteases of potential interest in the regulation of endogenous and heterologous protein secretion. Using expression levels, growth phase-specific synthesis or apparent biochemical properties we narrowed down our focus to 17 protease genes that were subsequently deleted. Eight of the derivative strains were analyzed with respect to their effect on whole secretome (or "exoproteome") export and in a more targeted approach, the secretion of a mRFP derivative carrying a Sec pathway signal peptide was studied in 14 of the deletion strains. These experiments revealed that the most significant effect was seen when the core regulatory protease FtsH, which is embedded in the plasma membrane (Walker et al., 2007), was removed. FtsH removal led to significant improvement of secretion at the whole secretome level and, also, at the targeted level of a secreted mRFP. A complex network of both transcription and protein level effects might explain this novel role of FtsH in protein secretion. This study lays the foundation for application of multi-omics tools to the study of several aspects of protein secretion in TK24 and paves the way towards better understanding and rationally redesigning heterologous protein secretion in these bacteria.

3.3. Results

Analysis and identification of highly transcribed proteases

S.lividans TK24 contains 190 protease-encoding genes (Table S3.1; SToPSdb (Tsolis et al., 2018)[1]. Of these, 50% are secreted and another 10% are membrane-embedded. In addition, TK24 secretes 3 proteins that act as protease inhibitors. To determine whether proteases play a role in the secretory processes of TK24, we first classified potential proteases based on several parameters, as follows: (I) the presence of a Sec or Tat secretion signal using multiple bioinformatics tools (Tsolis et al., 2018), (II) whether the protein in question belongs to the set of core genes of *S. lividans* TK24, based on the comparison to the core genome of 13 *Streptomyces* species, and (III) the maximal transcript level, determined by RNAseq. For the latter, TK24 was grown under two media regimes (minimal medium supplemented with glucose and minimal medium with glucose supplemented with casamino acids) and cells were harvested at three different growth phases (early and late logarithmic, and stationary).

Transcription signals for 93.8 % of the annotated encoded proteases could be determined and were used to rank them (Fig. 1A, Table S1). Forty one of the protease-encoding genes were transcribed at high levels, in a growth-phase- and medium-dependent manner. They code for secreted (13), cytoplasmic (18) and membrane-embedded (19) proteases. Some of the membrane-embedded proteases, such as FtsH, are well known proteostatic regulators in other organisms like *E.coli* (Gottesman, 1996). In addition to the proteases, transcription of a gene encoding for a secreted protease inhibitor was detected.

To determine if these transcription signals give rise to detectable secreted protein products, we undertook a proteomics analysis of the secretome of TK24 under the same growth conditions and sampling times along the growth curve of the cell (Fig. 3). In total, 82 of the proteases could be detected at the protein level, 57 of them were secreted or membrane-integrated (Table S2). A medium correlation of their average transcript and protein abundance was observed between transcriptomics and proteomics experiments (Fig. S1). In minimal media, the levels of about half of them peaked at early-log and late-log phase and while the others were most abundant in the stationary phase. In contrast, when growing in casamino acid-supplemented minimal media, abundance peaks were for most of the proteases at early-log or stationary phase. This late phase expression is reminiscent of the secretion from TK24 of some heterologous proteins in various growth media, in which high levels of secretion were linked to the cells entering the stationary phase (Pozidis et al., 2001; Hamed et al., 2017).

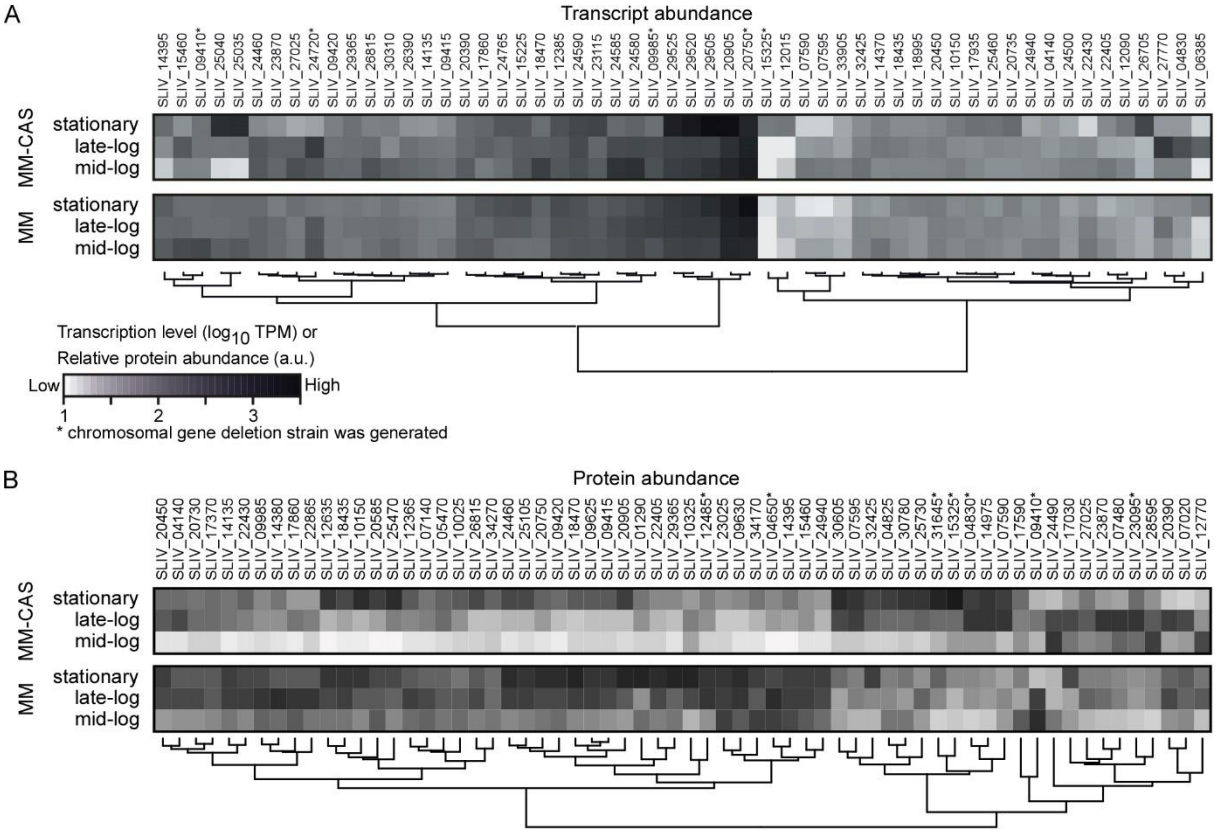


Figure 1. Transcriptomics and secretomics analysis of proteases in TK24. **(A)** Transcript abundance profile for protease encoding genes transcribed highly (TPM > 100) in MM or MM-CAS media for early-log, late-log and stationary growth phase cultures in at least one condition. Transcript abundance (TPM values) are log-transformed and grouped by hierarchical clustering using the Wards method and the Euclidean distance between values. Proteases whose genes were deleted (see Table 1) are shown with (*). **(B)** Abundance profile for secreted proteases identified by proteomics in MM or MM-CAS media for earl-log, late-log and stationary growth phase cultures. Protein abundance (iBAQ values) are log-transformed, centered, scaled and grouped by hierarchical clustering using the Euclidean distance between values. Proteases whose genes were deleted (see Table 1) are shown with (*). *n= 6-8.*

In addition to the proteases, secretomics revealed the presence of 3 protease inhibitors including that of subtilisin inhibitor (Uniprot Accession: D6EYB7), commonly seen as the most abundant exported protein in TK24 (Hamed et al., 2018).

Selection of protease genes of interest and creation of a protease deletion library

Given the transcriptomics and the secretomics results, we decided to focus on a small number of proteases and test their role in protein secretion. For this we chose 15 proteases, a protease inhibitor gene and membrane-embedded regulator (Table 1). The former included the likely proteostatic regulator FtsH (Langklotz et al., 2012) and its second homologue in *S. lividans* FtsH3 (both membrane-embedded), and Lon, a major proteostatic regulator in *E.coli* (Lee and Suzuki, 2008). We also included 9 secreted proteases. We chose the protease targets based on at

least one of the following criteria: high level transcription, possible role in overall regulation, possible direct effect on secreted polypeptides, or secretion patterns that were consistent with early or late appearance in the secretome. Following the gene selection process, we generated strains with specific single deletions of all the genes of interest, employing an established protocol used extensively for gene deletion experiments (Myronovskyi et al., 2014).

Fitness- and production testing of strains with protease deletions

The 14 TK24 derivative strains with deleted protease and protease inhibitor genes, were next characterized for the effect of protease gene deletions on growth and protein secretion. For this, the deletion mutant library was phenotyped towards growth-related parameters using a microbioreactor-based pipeline (Koepff et al., 2017). Most strains did not exhibit very strong phenotypes concerning the evaluated criteria (Fig. 2A). Three of the four parameters (maximum specific growth rate (μ_{max}), cell dry weight concentration (CDW) and cultivation time demand until stationary phase (t_{batch}) directly quantified growth performance. These values vary in the deletion strains mostly within ~20% to those of the WT. No deletion derivative strain showed consistently higher growth performance than that of TK24. TK24Δ*ftsH* exhibited by far the strongest phenotype in all three parameters (Fig. 2A). This highly transcribed membrane-embedded protease FtsH is an essential, hexameric, membrane-anchored metalloprotease in *E. coli* with a wide substrate diversity (Bieniossek et al., 2006; Walker et al., 2007). It targets multiple cellular processes including lipopolysaccharide biosynthesis, heat-shock sigma factor degradation, protein secretion, periplasmic chaperone functions and stress adaptation (Bittner et al., 2017). In comparison to TK24, TK24Δ*ftsH* revealed a 45% reduced μ_{max}, coupled with 37% reduced CDW and a 53% prolonged t_{batch}. FtsH3, a homologue of FtsH present in TK24 but absent from *E. coli*, was also deleted. However, this gene deletion has significantly less pronounced effects on growth than those of Δ*ftsH*. FtsH3 may have auxiliary roles in the cell that are less critical than those of FtsH.

Secretome analysis of strains with protease deletions

To test specifically the effect of the protease gene deletions on protein secretion we examined the secretomes of the derivative strains. Equal amounts of total secretome polypeptides were first analyzed by SDS-PAGE and silver staining (Fig. 3A). The patterns of the various derivatives seemed similar overall at this level of analysis except for TK24Δ*ftsH* that gave rise to several different polypeptides and had lost others (Fig. 3A, stars). Moreover, upon loading of secretome material derived from the same volume of culture, TK24Δ*ftsH* was also found to be a profuse

secretor of polypeptides. As seen in other studies (Hamed et al., 2017; Tsolis et al., 2018), there appears to be a good correlation between suppressed growth and improved secretion as seen by comparison of the total secretome expressed per unit cell biomass (Fig. 3B).

Table 1. List of protease and protease inhibitor genes selected for deletion analysis.

Name of derivative strain	Deleted gene name	Function	Topology (secretion system)	Structural/ Functional domains	TPM	SC	mRFP
TK24ΔSLIV_09985	SLIV_09985	Peptidase S8, subtilisin-related protein	Integral membrane protein (Sec)	IPR015500	639.67		+
TK24ΔSLIV_20750	SLIV_20750	ATP-dependent zinc metalloprotease FtsH (EC 3.4.24.-)	Integral membrane protein (Sec)	IPR005936	2528.62	+	+
TK24ΔSLIV_10535	SLIV_10535	ATP-dependent zinc metalloprotease FtsH3	Integral membrane protein (Sec)	IPR005936	6.46	+	+
TK24ΔSLIV_10025	SLIV_10025	T7SS peptidase S8A, mycosin-1, component of T7S export system	Integral membrane protein (Sec)	IPR015500	48.02		+
TK24ΔSLIV_11935	SLIV_11935	ATP-dependent serine protease Lon	Cytoplasmic	IPR027065	87.45	+	+
TK24ΔSLIV_11275	SLIV_11275	Neutral zinc metalloprotease	Secreted protein (Sec)	IPR023612	18.57	+	+
TK24ΔSLIV_11270	SLIV_11270	Neutral zinc metalloprotease	Secreted protein (TAT)	IPR023612	2.22	+	+
TK24ΔSLIV_15325	SLIV_15325	Peptidase, Leupeptin-inactivating enzyme 1	Secreted protein (Sec)	Peptidase_M 28 domain IPR007484	121.08		+
TK24ΔSLIV_09410	SLIV_09410	Peptidase	Secreted protein (Sec)	Peptidase_M 23 domain IPR016047	498.44		+
TK24ΔSLIV_24720	SLIV_24720	Protein containing Tachylectin 2 domain	Secreted protein (Sec)	IPR023294	733.23		+
TK24ΔSLIV_02150	SLIV_02150	Extracellular small neutral protease (EC 3.4.24.77)	Secreted protein (Sec)	IPR000013	1.09		+
TK24ΔSLIV_17030	SLIV_17030	Peptidase M1, alanine aminopeptidase/leukotriene A4 hydrolase	Secreted protein (Sec)	IPR001930	43.28		+
TK24ΔSLIV_31645	SLIV_31645	Tripeptidyl aminopeptidase (EC 3.4.14.-)	Secreted protein (Sec)	IPR000073	69.42	+	
TK24ΔSLIV_12485	SLIV_12485	Peptidase S33 tripeptidyl aminopeptidase-like protein	Secreted Lipoprotein (Sec)	IPR029058	59.60	+	
TK24ΔSLIV_04650	SLIV_04650	Peptidase containing LysM_dom and Peptidase M23 domains	Secreted protein (Sec)	LysM_dom (IPR018392) Peptidase M23 (IPR016047) domains	37.16	+	
TK24ΔSLIV_34120	SLIV_34120	Probable subtilase-type protease inhibitor	Secreted protein (Sec)	IPR000691	8162.43		+
TK24ΔSLIV_28740	SLIV_28740	Stomatin family	Integral membrane (Sec)	IPR001972	539.22		+

Maximal transcription level (TPM);, Secretomics (SC); mRFP secretion (mRFP).

Subcellular topology classification was obtained from STopSdb http://www.stopsdb.eu (Tsolis et al., 2018) using the updated topological annotation of the *Streptomyces lividans* proteome. IPR: Proteins signature based on Integrative proteins signature data base (InterPro database). (+) Represents strains that are included in secretomics analysis or mRFP secretion or both.

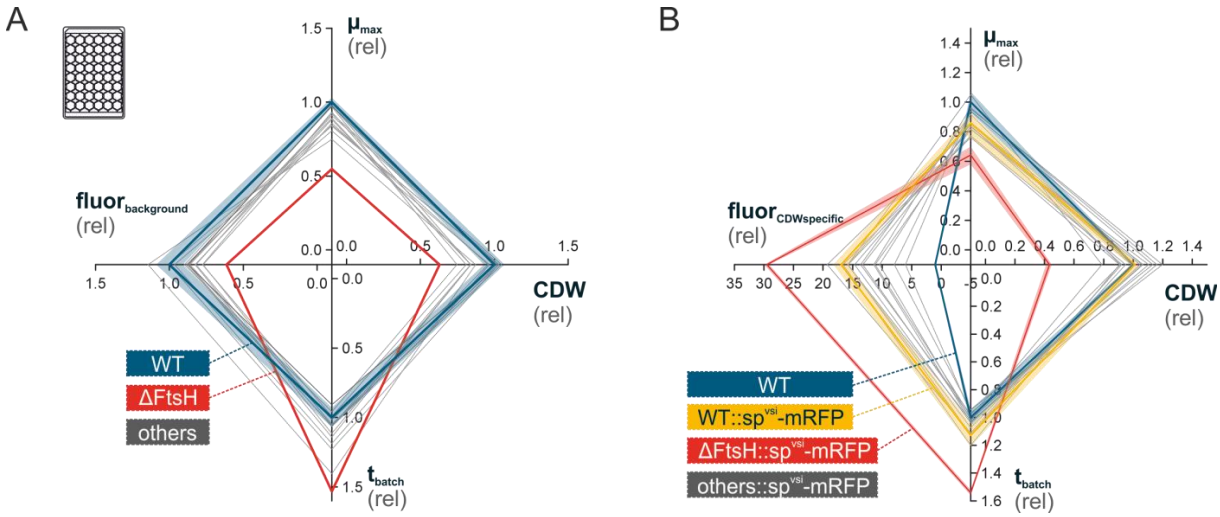


Figure 2. Phenotyping of *S. lividans* protease gene deletion strains, using microtiter-plate cultivation (Koepff et al. 2017).(**A)** Characterization of the protease deletion strain library for growth phenotypes, in advance to recombinant mRFP integration. In total, four parameters maximum specific growth rate (μ_{max}, top), cell-dry-weight (CDW, right), cultivation time until stationary phase (t_{batch}, bottom) and background RFP fluorescence ($fluor_{background}$, left, ex: 550 nm, em: 589 nm) were evaluated. All data was normalized on the wild-type *S. lividans* TK24 (WT, solid blue line), for which the obtained standard deviation (semi-transparent background) is additionally provided. The mutant, exhibiting the strongest phenotype, ΔFtsH, is outlined by a solid red line. **(B)** Phenotyping of the strain library with recombinant mRFP production. Axe labeling is identical to **(A)** except that $fluor_{background}$ was replaced by CDW-specific mRFP fluorescence ($fluor_{CDWspecific}$). Standard deviation is provided for the TK24 (blue) and TK24::*sp*^{vsi}*-mRFP* (yellow), as well as for TK24Δ*ftsH*::*sp*^{vsi}*-mRFP* (red).

The same samples were also analyzed by label-free nanoLC-MSMS and the identity of the polypeptides determined, and their amounts quantified (Table S3). The abundance of proteins in the secretome of the deletion derivative strains was compared to that of the wild type (Fig. 3C). In all cases, several polypeptides, representing 10-20% of the secretome, were identified at different abundances suggesting that the secretome is very sensitive to removal of proteases and yet, given the minor effect on fitness (Fig. 2A), the cellular system remains robust (Table S4). Functional characterization of these differentially abundant proteins, suggests that multiple hydrolases, proteins of housekeeping function and transport related proteins are oversecreted in the mutants showing the most severe phenotype (TK241ftsH, TK241lon) (Fig. S2 and Table S3).

Fitness- and production testing of protease deletion strains secreting a heterologous protein

In view of the wide-ranging effects of the deletion of protease genes on the secretome, we sought to corroborate these results with a more targeted approach that would also allow us to evaluate the potential of these strains for heterologous protein secretion. For this, protease deletion strains of interest were selected for a study focusing on the secretion of a single model protein SP^{vsi}-mRFP (Hamed et al., 2018). For this, the gene encoding SP^{vsi}-mRFP that was previously stably

integrated into the genome of TK24 (Hamed et al., 2018) was now integrated in the same position of the genomes of all the individual TK24 derivative strains carrying the different protease and protease inhibitor gene deletions using the phage VWB attachment site as described (Hamed et al., 2018).

Streptomyces species secrete various secondary metabolites, some of which fluoresce (Tenconi et al., 2013). It was therefore important prior to the SP^{vsi}-mRFP secretion analysis to ensure that the background fluorescence of the various deletion strains was compatible with experimental detection and quantification of secreted SP^{vsi}-mRFP. To this end a fourth parameter in the fitness/growth testing was the background fluorescence ($fluor_{background}$) (Fig. 2A). This was evaluated using an excitation/emission filter set (ex. 550 nm / em. 589 nm) appropriate for mRFP fluorescence detection and quantification. This analysis can identify possible fluorescence superposition effects, provoked by the various gene deletions that induced the production of endogenous fluorescent compounds by *Streptomyces* cells. However, none of the strains tested showed strongly enhanced background fluorescence with properties that could optically interfere with the fluorescence of SP^{vsi}-mRFP. Therefore, TK24 derivatives with specific protease gene deletions were used to determine their effect on SP^{vsi}-mRFP secretion.

The 14 strains with the integrated gene encoding SP^{vsi}-mRFP were characterized, using the microbioreactor (Fig. 2B, Table 1). All strains showed significantly increased biomass-specific mRFP fluorescence ($fluor_{CDWspecific}$) in comparison to TK24. However, this parameter showed large variation (~6 to ~29-fold). TK24::sp^{vsi}*-mRFP* showed a ~17-fold increase in comparison to the other strains that carried not only the SP^{vsi}-mRFP reporter but also protease and protease inhibitor gene deletions.

To directly determine the extent to which the measured fluorescence represented secreted mRFP, we separated cells from spent growth media by centrifugation and measured the individual fluorescence in the two fractions. These data revealed that >89% of the mRFP fluorescence derived from secreted protein, in agreement with previous observations (Hamed et al., 2018). Additionally, we have tested the effect of deletion of aforementioned genes on diamide tolerance in presence and absence of sp^{vsi}*-mRFP* construct. Diamide is a thiol oxidant, causing generation of nonnative disulfide bonds, resulting in damage of cytoplasmic proteins (Hochgräfe et. al., 2005). Most of the strains showed no or insignificant changes in diamide sensitivity. In contrast, the Δ*ftsH* and Δ*ftsH3* mutants were found to be more susceptible to the thiol oxidative stress when expressing the sp^{vsi}*-mRFP* then the parental strain TK24 (Fig. S3, Table S6). At the same time, small increase in sensitivity was observed also in the case of *S. lividans* deficient in

SLIV_09985 encoding putative integral membrane peptidase S8. This finding might indicate the involvement of these genes in the protein quality control or protein processing, at least during growth on solid media.

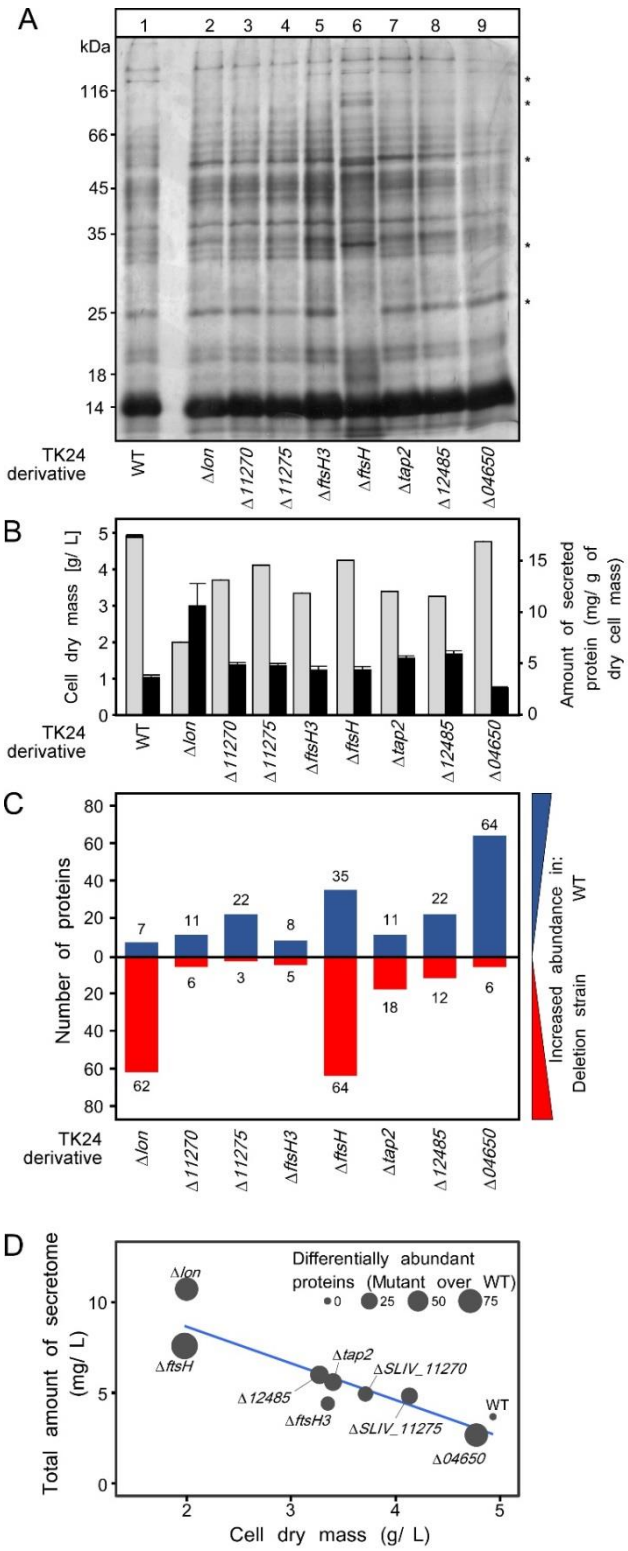


Figure 3. Secretome analysis of protease deletion strains. (**A)** Secretome profile of WT TK24 and derivatives resolved by SDS-PAGE and silver stained, loading equal amount of secretome polypeptides (3 µg). Representative samples from each TK24 derivative was loaded, for the strains that were used in the proteomic analysis. (**B)** Cell dry weight (g/L) (light grey) and amount of total proteins secreted (mg/L) (black) for the strains used in the proteomic analysis. (**C)** Number of differentially abundant secreted proteins between each TK24 derivative against the WT. Proteins with increased abundance in WT are shown in blue and proteins more abundant in deletion strains are coloured red. Samples were loaded for proteomic analysis normalized to the same amount of cell biomass. (**D)** Correlation between dry cell weight, amount of total proteins secreted and number of differentially abundant proteins over the WT. (*n = 4-7*)

In-depth multi-omics analysis of the strain deleted for *ftsH*

Given the higher biomass-specific mRFP fluorescence performance of TK24Δ*ftsH*::*SP*^{vsi}*-mRFP* and the significant effects on the secretome overall seen with TK24Δ*ftsH*, we further characterized the effects, caused by the deletion of this highly expressed, non-secreted core-protease.

TK24Δ*ftsH* and TK24 were cultivated in the microbioreactor, to generate sufficient biological replicates for subsequent omics analysis. Samples were taken, during early and late growth phase, as well as during stationary phase, membrane filtered and snap-frozen. Given the role of *ftsH* in *E. coli* in regulating the turn-over of sigma-factors (Bittner et al., 2017) and the severity of the changes seen at the secretome level when it is deleted (Fig. 3), it was of interest to determine whether some effect of Δ*ftsH* is already exerted at the transcriptome level. To test this, we compared the transcriptome profiles of TK24 with those of TK24Δ*ftsH* (Fig. 4). This analysis revealed that loss of FtsH has a significant effect on the transcriptome, regardless of the growth phase sampled. Overall, the transcript abundance of 2,240 genes was reduced at least twofold in all three growth phases sampled, while only 166 genes showed a consistent increase in transcript abundance (Fig. 4A). This reduction in transcript levels was even more pronounced in the early and late-log phase (an additional 1,547 genes with reduced pools compared to an additional 157 genes with increased pools), while the stationary phase alone shows a significant increase in some transcript pools (816 genes). This major disturbance of the transcriptome seems to correlate with the reduced growth of TK24Δ*ftsH*, but makes it nearly impossible to pinpoint any specific molecular cause. This becomes apparent when the average change in transcript levels over all three time points analyzed is examined (Fig. 4B): With more than 1,567 genes with significantly changed transcript pools (1,466 with decreased pools and 101 with increased pools), the effects of the *ftsH* deletion can only be described as global.

Finally, we analyzed the secretome of TK24Δ*ftsH* and compared it to that of the wild type. The abundance of 99 secreted proteins is statistically different in the two strains, reflecting proteins that are seen at both lower and higher levels in the mutant strain (Fig. 5A). Proteins affected include: a quinoprotein amine dehydrogenase, a solute-binding lipoprotein and a D-alanyl-D-alanine carboxypeptidase that are synthesized/secreted 3-7 times less than in TK24 and a Phospholipase-A2 domain-containing protein, a spore-associated protein A and a branched chain amino acid binding protein that are secreted >4 times more. Overall, the affected proteins fall in 6 main functional classes (Fig. 5B). This is suggestive of an extensive regulatory role of *ftsH* in *S. lividans.*

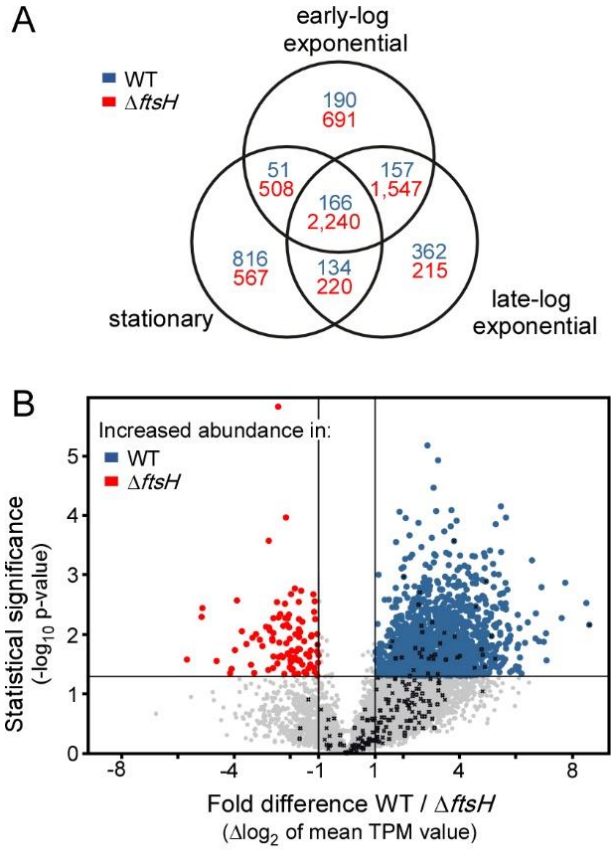


Figure 4. Transcriptomic activity of TK24Δ*ftsH* against that of TK24. **(A)** Venn diagram showing the number of genes with a change of transcript levels of at least two-fold in the three growth stages sampled. Numbers in blue refer to genes with increased transcript abundance in the strain TK24Δ*ftsH* while numbers in red refer to genes with increased transcript abundance in the WT TK24. **(B)** Volcano-plot showing the differences in transcript abundance of all genes between strains TK24Δ*ftsH* and WT TK24. Each dot represents one gene. On the x axis is plotted the fold difference (in $\log_2$ scale) of the transcript abundance in the Δ*ftsH* strain compared to that in the WT, averaged over the three time points sampled. On the y axis, the p-value derived from a t-test between the two strains ($-\log_{10}$) is given. Transcripts with higher abundance in TK24Δ*ftsH* are given in blue, those more abundant in the WT are given in red, while the abundance of transcripts given in grey not significantly changed ($-1 < m < 1$ and or $p > 0.05$). Genes encoding proteases are listed in **Table 1**.

3.4. Discussion

Our analysis aimed to determine whether proteases play a role in protein secretion and whether disruption of the secreted protease equilibria affects the secretome of *S. lividans* overall. Our data suggested that deletion of multiple proteases is possible in *S. lividans*. While, significant qualitative and quantitative effects can be seen on the secretome, in most cases the strains maintain metabolic robustness.

The effect of the various deletions on the secretome would indicate that secretome proteases are under strict metabolic control, although the molecular mechanism underlying this regulation remains unknown. Nitrogen provided by the casamino acids may down-regulate the synthesis of some of the proteases until this resource is depleted. The apparent clustering pattern of several of the proteases (Fig. 1B) also raises the possibility that they may be under the same transcriptional regulatory control. The multiplicity of sigma and anti-sigma factors in TK24 (Rebets et al., 2018) and the currently incomplete understanding of promoter usage and promiscuity, renders buildup of regulon networks challenging. Another parameter that would regulate these equilibria would be protease inhibitors, particularly given that many of them will inhibit multiple proteases (e.g.

there are 6 secreted subtlilisin-like proteases that may be inhibited by the same molecules). Better understanding of these networks is expected to derive from RNAseq analyses combined with strains deleted for specific sigma factors. Despite the obvious qualitative and quantitative changes observed, overall most of the deletions had no obvious effect on fitness under the conditions tested. This is suggestive of a high degree of build-in redundancy and/or the peripheral contribution of the removed secreted proteases in house-keeping processes of the cellular network. This was also true for removal of the cytoplasmic and membrane proteases of known regulatory importance in cell proteostasis such as Lon and FtsH3 (Fig. 3).

Some secretory proteases would be expected to have regulatory roles. For example, serine proteases were reported to coordinately regulate the cellular protein turnover associated with secondary metabolism and morphogenesis (Taguchi et al., 1995), also, their physiological roles in terms of mycelial growth, autolysis of mycelia after stationary phase in submerged cultures (Kim and Lee, 1995). Furthermore, leupeptin-inactivating enzyme also seems to play a critical role in mycelium differentiation of *S. exfoliates* SMF13 by controlling the amount of leupeptin that regulates trypsin-like proteases activity (Kim et al., 1998). In these cases, their removal would lead to the stable accumulation of their potential polypeptide-targets that now become more stable and can be detected. In some cases, secreted polypeptides of enhanced abundance were indeed observed. However, in addition, certain secreted polypeptides also became less abundant when specific proteases were removed (Fig. 4C, Table S3). This would suggest that some secreted proteases are co-regulated together with other secreted polypeptides in a way that is independent of the proteolytic activity. The molecular basis of these apparently complex balanced networks remains unknown.

FtsH, was the only protease with a significant effect on fitness (Fig. 3, lane 6). While, the biological function of FtsH and the need for a second copy in *Streptomyces* is not well understood, we assume that the severity of the effect, reflects important regulatory roles in TK24 as seen previously in *E. coli* (Langklotz et al., 2012). This is further corroborated by the major disturbance of the transcriptome and secretome of TK24Δ*ftsH*, and hence, FtsH is expected to be a major global regulator affecting 6 main functional classes of secretome proteins (Fig. 5B). The regulatory role of *ftsH* in *S. lividans* may be exerted through a combination of mechanisms that involve both transcriptional (Fig. 4) and other means of regulation, e.g. protein degradation. Little is known about the molecular basis of these effects and how they might be inter-connected. The extent of the transcriptome alterations in TK24Δ*ftsH* precludes any guidance as to a specific molecular pathway. At least two hypotheses can be entertained: (a) as FtsH is membrane-

embedded and secretion is a process of transmembrane-crossing, the actual translocation of some proteins or their clearing from translocase sites may require FtsH, as seen in *E. coli* (Langklotz et al., 2012). (b) removal of FtsH may bring the cell to a stress state that could be analogous to the

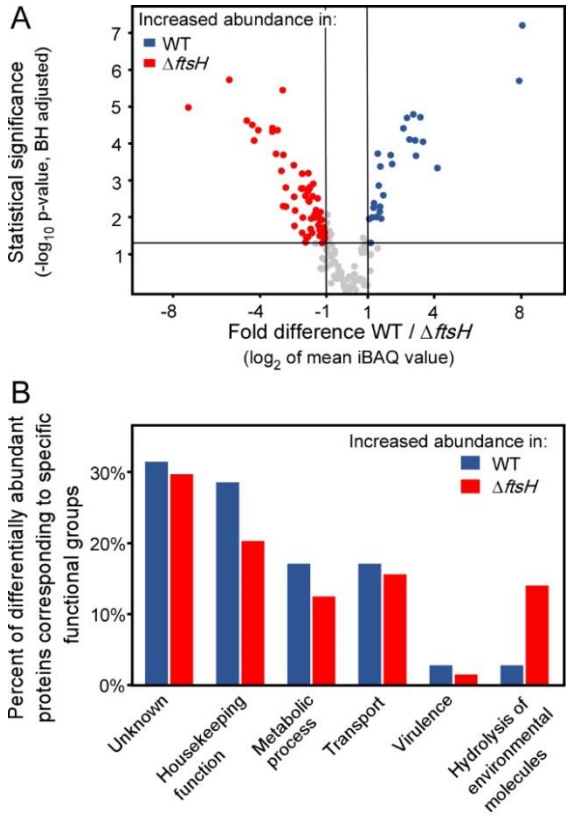


Figure 5. Comparative secretomics analysis of TK24Δ*ftsH* against TK24. **(A)** Volcano-plot showing the differentially abundant proteins between TK24 and TK24Δ*ftsH*. Each dot represents one protein. On the x axis is plotted the fold difference (in $\log_2$ scale) of the mean protein abundance in the TK24 (marked: WT) over that in TK24Δ*ftsH*, and on the y axis the p-value derived from a t-test between the two strains ($-\log_{10}$, adjusted by the Benjamini–Hochberg method). With blue are colored proteins more abundant in the WT and in red, the significantly more abundant proteins in TK24Δ*ftsH*. **(B)** Functional classification of differentially abundant proteins based on their biological function. The ratio of the abundant proteins corresponding to a specific functional category over the number of over-represented proteins in the specific strain is plotted.

role of FtsH in dealing with membrane protein stress in *E.coli* (Bittner et al., 2017). This may correlate with transcripts of known stress-related genes such as SLIV_24510 (encoding superoxide dismutase), SLIV_12445 (encoding a glutaredoxin-like protein), being 2-45-fold elevated in TK24Δ*ftsH*. Whichever the mechanism, FtsH appears to be a core proteostatic component since no other proteases can replace it and maintain the same transcript level trends as those of the other genes (Fig. 4B, marked with black Xs).

The stable activity of secreted mRFP underlines the suitability of *S. lividans* for heterologous protein production and is reminiscent of what has been observed with several other heterologous proteins (Pozidis et al., 2001; Sianidis et al., 2006; Hamed et al., 2017). Perhaps the deletion of multiple secreted proteases could be more beneficial for heterologous protein production, but this can be seen now under new light. Previously, multiple deletions in Gram positive bacteria were used to reduce proteolysis of the heterologous product [e.g. in *B. subtilis* (Westers et al., 2004)]. Furthermore, deletion of 8 secreted proteases in *B. subtilis* affects the heterologous production not only by reducing its degradation but by increasing the extra-cytoplasmic chaperons and

quality control factors PrsA, HtrA and HtrB as well (Krishnappa et al., 2013; Krishnappa et al., 2014). However, we now see that removal of proteases can also act in a regulatory role. This is more prominently seen with deletion of genes for the regulatory protease FtsH, which has a role in quality control of membrane proteins (Dalbey et al., 2012). Therefore, a finer regulation of proteases such as FtsH might have implication for improving the quality and secretion yield.

3.5. Materials and Methods

Generation of protease gene deletions

To delete selected genes BAC clones from ordered *S. lividans* BAC library were selected and mutagenized using the Red/ET technique combined with the apramycin resistance IMES cassette from patt-saac-oriT (Myronovskyi et al., 2014). Primers used to amplify the cassette and to verify the mutation are listed in Table S7. Red/ET recombineering of the BACs using amplified apramycin resistance IMES cassette fragment was performed as described previously (Gust et al., 2004). The resulting recombinant BACs were introduced in the *S. lividans* TK24 via conjugation (Rebets et al., 2017). Screening for double-crossover mutants was performed on MS medium (per liter: 20 g agar, 20 g mannitol, 20 g soya flour) supplemented with 50 μg/ml of apramycin and 70 μg/ml of X-gluc. Gene deletions were confirmed via PCR using appropriate check primers.

The fragment containing the *SP^{vsi}-mRFP* gene, i.e. the *S. venezuelae* subtilisin inhibitor signal peptide *SP^{vsi}* fused to *mRFP* behind the strong *vsi* promotor (P_{vsi}) was excised from the plasmid pIJ486**V**-SP^{vsi}-mRFP (Hamed et al., 2018) using *Xba*I and *Hind*III and ligated into the respective sites of pTOS (Herrmann et al., 2012) yielding pTOS+mRFP that contained the SP^{vsi}-mRFP-encoding gene and *attB* of VWB-phage flanked by two *rox*-sites. This plasmid was introduced into the genome of *S. lividans* TK24 and protease gene deletion strains by intergeneric conjugation with *E.coli* ET1326::pUZ8002 (Kieser et al., 2000) as described (Herrmann et al., 2012). For each conjugated strain, genomic DNAs of four randomly chosen exconjugants were isolated and verified by PCR for proper integration of the pTOS+mRFP plasmid. Then, a plasmid containing the gene of the *Dre* recombinase (pUWLDre), was introduced into the respective mutant strains and the pTOS+mRFP-backbone was excised as described (Herrmann et al., 2012).

Rapid *S. lividans* strain phenotyping

Time-efficient characterization of the strain library was realized using a previously published

workflow, which utilizes parallelized microbioreactor cultivation in 48-well microplate with a working volume of 1000 µL in each well (Koepff et al., 2017). All strains were, at minimum, cultivated in biological triplicates. The WT was incorporated in each separate microtiter-plate run. The corresponding WT results were averaged and used to normalize the data, obtained from all other deletion mutants. By this methodology, batch-related differences could be compensated. Detailed cultivation and data processing pipeline is described in detail in (Rebets et al., 2018)

Monomeric red fluorescence protein florescence was relatively quantified during cultivation by using a excitation/emission: 550 nm/ 589 nm filter set, incorporated in the automated microbioreactor (m2p-labs, Baesweiler, Germany). To efficiently compare the strain performance, mRFP fluorescence intensity values, obtained at the transition to stationary phases were identified and evaluated. Error propagation was applied to calculate CDW-specific mRFP production.

Lab-scale bioreactor cultivation

Lab-scale cultivations of *S. lividans* were carried out in parallelized glass bioreactors (DasGip, Jülich, Germany) with a working volume of 1000 mL. In principle, the same media composition as in the microbioreactor was applied, with the exception, that no MES buffer was used, but a constant pH of 6.8 was maintained by titrating 4 M NaOH or 4 M HCl solutions if required.

Transcriptomics, identification of highly transcribed proteases

Samples were taken during the early and late log growth phase as well as after entry into the stationary phase. Harvesting and RNA isolation was performed as described previously (Busche et al., 2012). Samples from different biological replicates were isolated separately and pooled after quality control. The RNA quality was checked via Agilent 2100 Bioanalyzer (Agilent Technologies, Böblingen, Germany) and Trinean Xposesystem (Gent, Belgium) prior and after rRNA depletion using the Ribo Zero rRNA Removal Kit for Bacteria (Epicentre, Madison, WI, United States). The TruSeq Stranded mRNA Library Prep Kit from Illumina was used to prepare the cDNA libraries, which were then sequenced in paired-end mode on an Illumina HiSeq 1500 system with 28 respectively 70 bases read length. Transcripts per kilobase million (TPM) (Wagner et al., 2012) were calculated using READXPLORER v.2.2 (Hilker et al., 2016). For differential RNA-Seq analyses the signal intensity value (A-value) was calculated by average $\log_2$(TPM) of each gene and the signal intensity ratio (M-value) by the difference of $\log_2$(TPM). In cases where the TPM for a gene was 0, a TPM of 0.1 was used instead to avoid $\log_2(0)$. To identify protaeses that were strongly transcribed and differentially expressed under at least one

condition, the RNA-Seq data were filtered using a TPM cut-off of 100 and an M-value cut-off of >1.0 under at least one condition. The raw sequence data sets are available at the NCBI SRA under study ID SRP144344, SRA accessions SRR7093716-SRR7093727.

For analysis of the transcriptome comparisons of the Δ*ftsH* and WT strains, the average of the *M*-values over all three time points sampled as well as a *P*- value based on the $\log_2$(TPM) values using a Student's t-test (two tailed, heteroscedastic) were calculated. Genes with an average *M*-value above/below 1/-1 and a *P*-value < 0.05 were considered to be differentially transcribed.

Secretomics sample preparation and measurement

Cells were removed by centrifugation (4,500 x *g*; 5 min; 4 ^{o}C) and subsequent filtration (syringe filter, 0.2 µm, cellulose acetate). Proteins contained in culture supernatants were precipitated via 25% v/v TCA precipitation (4 ^{o}C; 20 min). Precipitated proteins were pelleted via centrifugation (20,000 x *g*; 20 min; 4 ^{o}C), on a bench-top centrifuge. The pellet was washed twice with ice-cold acetone and re-pelleted via centrifugation (20,000 x *g*; 20 min; 4 ^{o}C). The protein pellet was then solubilized in 8M Urea in 1M ammonium bicarbonate solution (ABS). Polypeptide concentrations were measured using the Bradford reagent. Polypeptides (3 µg) were separated by 12% SDS-PAGE and visualized by silver staining (Shevchenko et al. 1996).

Analysis of secretomes by nanoLC-MS/MS

A volume corresponding to the secreted polypeptides derived from 3 x 10^6 cells (usually a volume equivalent to 20-40 µL of the initial cell culture) was used for in-solution tryptic digestion. The protein solution was initially diluted into urea (2 M final concentration in 50 mM Ammonium bicarbonate solution (ABS), followed by reduction of cysteines with 1 mM DTT (45 min; 56 ^{o}C), alkylation using 10 mM Iodoacetamide (IAA) (45 min; 22 ^{o}C; dark) and digestion using 0.015 µg Trypsin for 1.5 µg protein (Trypsin Gold, Promega, Fitchburg, Wisconsin; ratio trypsin/protein 1/100; overnight; 37^{o}C). Digested peptide solutions, were acidified with trifluoroacetic acid (TFA) to pH<2, desalted using STAGE tips (Rappsilber et al., 2007; Tsolis and Economou, 2017), and stored lyophilized at -20^{o}C, until the MS analysis.

Lyophilized peptide samples were re-suspended in an aqueous solution containing 0.1% v/v formic acid (FA) and 5% v/v Acetonitrile (ACN) and analyzed using nano-Reverse Phase LC coupled to a QExactive Hybrid Quadrupole - Orbitrap mass spectrometer (Thermo Scientific, Bremen, Germany) through a nanoelectrospray ion source (Thermo Scientific, Bremen, Germany). Peptides were initially separated using a Dionex UltiMate 3000 UHPLC system on an

EasySpray C18 column (Thermo Scientific, OD 360 µm, ID 50 µm, 15 cm length, C18 resin, 2 µm bead size) at a flow rate of 300 nL min^{-1}. The LC mobile phase consisted of two different buffer solutions, an aqueous solution containing 0.1% v/v FA (Buffer A) and an aqueous solution containing 0.08% v/v FA and 80% v/v ACN (Buffer B). A 60 min multi-step gradient was used from Buffer A to Buffer B as follows [0–3 min constant (96:4), 3–15 min (90:10); 15–35 min (65:35); 35–40 min (35:65); 40-41 min (5:95); 41-50 min (5:95); 50-51 min (95:5); 51-60 min (95:5)].

The separated peptides were analyzed in the Orbitrap QE operated in positive ion mode (nanospray voltage 1.5 kV, source temperature 250 °C). The instrument was operated in data-dependent acquisition (DDA) mode with a survey MS scan at a resolution of 70,000 FWHM for the mass range of m/z 400-1600 for precursor ions, followed by MS/MS scans of the top 10 most intense peaks with +2, +3 and +4 charged ions above a threshold ion count of 16,000 at 35,000 resolution. MS/MS was performed using normalized collision energy of 25% with an isolation window of 3.0 m/z, an apex trigger 5-15 sec and a dynamic exclusion of 10 s. Data were acquired with Xcalibur 2.2 software (Thermo Scientific).

Raw MS files were analyzed by the MaxQuant v1.5.3.3 proteomics software package (Cox, Mann 2008). MS/MS spectra were searched by the Andromeda search engine against the Uniprot *S. lividans* TK24 proteome (taxonomy: 457428, last modified May, 2016, 7320 protein entries; (Ruckert et al., 2015) and common contaminants (e.g. trypsin, keratins). Enzyme specificity was set to trypsin and a maximum of two missed cleavages were allowed. Dynamic (methionine oxidation and N-terminal acetylation) and fixed (S-carbamidomethylation of cysteinyl residues) modifications were selected. Precursor and MS/MS mass tolerance was set to 20 ppm for the first search (for the identification of the maximum number of peptides for mass and retention time calibration) and 4.5 ppm for the main search (for the refinement of the identifications). Protein and peptide false discovery rate (FDR) were set to 1%. FDR was calculated based on the number of spectra matched to peptides of a random proteome database (reversed sequence database) in relation to the number of spectra matching to the reference proteome. Peptide features were aligned between different runs and masses were matched (“match between runs” feature), with a match time window of 3 min and a mass alignment window of 20 min. Protein quantification was performed using the iBAQ algorithm (Schwanhausser et al., 2011) through MaxQuant software. Differentially abundant proteins were selected using the t-test and by comparing the fold difference of average protein intensities between the samples. P-values were further corrected for multiple hypothesis testing error using the Benjamini-Hochberg method

(Benjamini and Hochberg, 1995).Thresholds for the analysis were set to adjusted p-value < 0.05 and fold difference > 2. Functional characterization of proteomics results was performed after filtering the dataset only to secreted proteins, excluding cytoplasmic contamination, using proteome annotation as described in the SToPSdb (Tsolis et al., 2018) (*www.stopsdb.eu*).The percentage of differentially abundant proteins that match a specific term over the total differentially abundant proteins for each condition was plotted. Keywords were derived after manual curation of the proteome.

Analysis of diamide sensitivity

Sensitivity of the *S. lividans* mutants to diamide was tested by plating spores (apr. 10^9) of each strain on fresh MS or TSB agar plates. Immediately after plating, paper discs loaded with 5 μM or 10 μM of diamide were placed, and plates were analyzed after 48 h incubation at 30°C. The tests were repeated three times for each strain. The inhibition zones were measured.

Miscellaneous

Chemicals were from Sigma. DNA enzymes were from New England Biolabs and oligonucleotides from Eurogentec. The mass spectrometry proteomics data have been deposited to the ProteomeXchange Consortium via the PRIDE partner repository with the dataset identifiers PXD006819 (Vizcaino et al., 2016).

3.7. References

Anne, J., Economou, A., and Bernaerts, K. (2017). Protein Secretion in Gram-Positive Bacteria: From Multiple Pathways to Biotechnology. *Curr Top Microbiol Immunol* 404**,** 267-308. doi: 10.1007/82_2016_49.

Anne, J., Maldonado, B., Van Impe, J., Van Mellaert, L., and Bernaerts, K. (2012). Recombinant protein production and Streptomycetes. *J Biotechnol* 158(4)**,** 159-167. doi: 10.1016/j.jbiotec.2011.06.028.

Barka, E.A., Vatsa, P., Sanchez, L., Gaveau-Vaillant, N., Jacquard, C., Meier-Kolthoff, J.P., et al. (2016). Taxonomy, Physiology, and Natural Products of Actinobacteria. *Microbiol Mol Biol Rev* 80(1)**,** 1-43. doi: 10.1128/MMBR.00019-15.

Benjamini, Y., and Hochberg, Y. (1995). Controlling the False Discovery Rate - a Practical and Powerful Approach to Multiple Testing. *Journal of the Royal Statistical Society Series B-Methodological* 57(1)**,** 289-300. doi: citeulike-article-id:1042553doi: 10.2307/2346101.

Bieniossek, C., Schalch, T., Bumann, M., Meister, M., Meier, R., and Baumann, U. (2006). The molecular architecture of the metalloprotease FtsH. *Proc Natl Acad Sci U S A* 103(9)**,** 3066-3071. doi: 10.1073/pnas.0600031103.

Bittner, L.M., Arends, J., and Narberhaus, F. (2017). When, how and why? Regulated proteolysis by the essential FtsH protease in Escherichia coli. *Biol Chem* 398(5-6)**,** 625-635. doi: 10.1515/hsz-2016-0302.

Busche, T., Šilar, R., Pičmanová, M., Pátek, M., and Kalinowski, J. (2012). Transcriptional regulation of the operon encoding stress-responsive ECF sigma factor SigH and its anti-sigma factor RshA, and control of its regulatory network in Corynebacterium glutamicum. *BMC Genomics* 13**,** 445. doi: 10.1186/1471-2164-13-445.

Butler, M.J., Bergeron, A., Soostmeyer, G., Zimny, T., and Malek, L.T. (1993). Cloning and characterisation of an aminopeptidase P-encoding gene from Streptomyces lividans. *Gene* 123(1)**,** 115-119.

Chater, K.F. (2016). Recent advances in understanding Streptomyces. *F1000Res* 5**,** 2795. doi: 10.12688/f1000research.9534.1.

Dalbey, R.E., Wang, P., and van Dijl, J.M. (2012). Membrane proteases in the bacterial protein secretion and quality control pathway. *Microbiol Mol Biol Rev* 76(2)**,** 311-330. doi: 10.1128/MMBR.05019-11.

Gilbert, M., Morosoli, R., Shareck, F., and Kluepfel, D. (1995). Production and secretion of proteins by streptomycetes. *Crit Rev Biotechnol* 15(1)**,** 13-39. doi: 10.3109/07388559509150530.

Gotoh, T., Miyazaki, Y., Sato, W., Kikuchi, K., and Bentley, W.E. (2001). Proteolytic activity and recombinant protein production in virus-infected Sf-9 insect cell cultures supplemented with carboxyl and cysteine protease inhibitors. *J Biosci Bioeng* 92(3)**,** 248-255.

Gottesman, S. (1996). Proteases and their targets in Escherichia coli. *Annu Rev Genet* 30**,** 465-506. doi: 10.1146/annurev.genet.30.1.465.

Gur, E., Biran, D., and Ron, E.Z. (2011). Regulated proteolysis in Gram-negative bacteria--how and when? *Nat Rev Microbiol* 9(12)**,** 839-848. doi: 10.1038/nrmicro2669.

Gust, B., Chandra, G., Jakimowicz, D., Yuqing, T., Bruton, CJ., Chater, KF. (2004). Lambda red-mediated genetic manipulation of antibiotic-producing *Streptomyces*. *Adv Appl Microbiol* 54:107–128. doi: 10.1016/s0065-2164(04) 54004-54002

Hamed, M.B., Karamanou, S., Olafsdottir, S., Basilio, J.S.M., Simoens, K., Tsolis, K.C., et al. (2017). Large-scale production of a thermostable Rhodothermus marinus cellulase by heterologous secretion from Streptomyces lividans. *Microb Cell Fact* 16(1)**,** 232. doi: 10.1186/s12934-017-0847-x.

Hamed, M.B., Vranken, K., Bilyk, B., Koepff, J., Novakova, R., K., T.C., et al. (2018). Monitoring protein secretion in *Streptomyces* using fluorescent proteins. Front. Microbiol. 9:3019. doi: 10.3389/fmicb.2018.03019

Herrmann, S., Siegl, T., Luzhetska, M., Petzke, L., Jilg, C., Welle, E., et al. (2012). Site-specific recombination strategies for engineering actinomycete genomes. *Appl Environ Microbiol* 78(6)**,** 1804-1812. doi: 10.1128/AEM.06054-11.

Hilker, R., Stadermann, K.B., Schwengers, O., Anisiforov, E., Jaenicke, S., Weisshaar, B., et al. (2016). ReadXplorer 2-detailed read mapping analysis and visualization from one single source. *Bioinformatics* 32(24)**,** 3702-3708. doi: 10.1093/bioinformatics/btw541.

Hochgrafe, F., Mostertz, J., Albrecht, D., and Hecker, M. (2005). Fluorescence thiol modification assay: oxidatively modified proteins in Bacillus subtilis. *Mol. Microbiol.* 58, 409–425. doi: 10.1111/j.1365-2958.2005.04845.x

Hopwood, D.A. (2007). Streptomyces in nature and medicine. The antibiotic makers. Oxford, New York: Oxford University Press.

Kieser, T., Bibb, M.J., Buttner, M.J., Chater, K.F., and Hopwood, D.A. (2000). Practical streptomyces genetics. Norwich, England: John Innes Foundation.

Kim, I.S., Kim, Y.B., and Lee, K.J. (1998). Characterization of the leupeptin-inactivating enzyme from Streptomyces exfoliatus SMF13 which produces leupeptin. *Biochemical Journal* 331(Pt 2)**,** 539-545.

Kim, I.S., and Lee, K.J. (1995). Physiological roles of leupeptin and extracellular proteases in mycelium development of Streptomyces exfoliatus SMF13. *Microbiology* 141 (Pt 4)**,** 1017-1025. doi: 10.1099/13500872-141-4-1017.

Koepff, J., Keller, M., Tsolis, K.C., Busche, T., Ruckert, C., Hamed, M.B., et al. (2017). Fast and reliable strain characterization of Streptomyces lividans through micro-scale cultivation. *Biotechnol Bioeng* 114(9)**,** 2011-2022. doi: 10.1002/bit.26321.

Krishnappa, L., Dreisbach, A., Otto, A., Goosens, V. J., Cranenburgh, R. M.,Harwood, C. R., et al. (2013). Extracytoplasmic proteases determining the cleavage and release of secreted proteins, lipoproteins, and membrane proteins in *Bacillus subtilis*. J. Proteome Res. 12, 4101–4110. doi: 10.1021/pr400433h.

Krishnappa, L., Monteferrante, C.G., Neef, J., Dreisbach, A., and van Dijl, J.M. (2014). Degradation of extracytoplasmic catalysts for protein folding in Bacillus subtilis. *Appl Environ Microbiol* 80(4)**,** 1463-1468. doi: 10.1128/aem.02799-13.

Langklotz, S., Baumann, U., and Narberhaus, F. (2012). Structure and function of the bacterial AAA protease FtsH. *Biochim Biophys Acta* 1823(1)**,** 40-48. doi: 10.1016/j.bbamcr.2011.08.015.

Lee, I., and Suzuki, C.K. (2008). Functional mechanics of the ATP-dependent Lon protease- lessons from endogenous protein and synthetic peptide substrates. *Biochim Biophys Acta* 1784(5)**,** 727-735. doi: 10.1016/j.bbapap.2008.02.010.

Liu, L., Yang, H., Shin, H.D., Li, J., Du, G., and Chen, J. (2013). Recent advances in recombinant protein expression by *Corynebacterium*, *Brevibacterium*, and *Streptomyces*: from transcription and translation regulation to secretion pathway selection. *Appl Microbiol Biotechnol* 97(22)**,** 9597-9608. doi: 10.1007/s00253-013-5250-x.

Ludke, A., Kramer, R., Burkovski, A., Schluesener, D., and Poetsch, A. (2007). A proteomic study of *Corynebacterium glutamicum* AAA+ protease FtsH. *BMC Microbiol* 7(1)**,** 6. doi: 10.1186/1471-2180-7-6.

Myronovskyi, M., Rosenkranzer, B., Luzhetskyy, A. (2014). Iterative marker excision system. Appl Microbiol Biotechnol, 98(10):4557-4570. doi: 10.1007/s00253-014-5523-z.

Neef, J., Bongiorni, C., Goosens, V.J., Schmidt, B., and van Dijl, J.M. (2017). Intramembrane protease RasP boosts protein production in Bacillus. *Microb Cell Fact* 16(1)**,** 57. doi: 10.1186/s12934-017-0673-1.

Pohl, S., Bhavsar, G., Hulme, J., Bloor, A.E., Misirli, G., Leckenby, M.W., et al. (2013). Proteomic analysis of *Bacillus subtilis* strains engineered for improved production of heterologous proteins. *Proteomics* 13(22)**,** 3298-3308. doi: 10.1002/pmic.201300183.

Pozidis, C., Lammertyn, E., Politou, A.S., Anne, J., Tsiftsoglou, A.S., Sianidis, G., et al. (2001). Protein secretion biotechnology using *Streptomyces lividans*: large-scale production of functional trimeric tumor necrosis factor alpha. *Biotechnol Bioeng* 72(6)**,** 611-619.

Ranjani, A., Dhanasekaran, D., and Gopinath, P.M. (2016). An Introduction to Actinobacteria. In Dharumadurai Dhanasekaran, Yi Jiang (Eds.): Actinobacteria - Basics and Biotechnological Applications: InTech.

Rappsilber, J., Mann, M., and Ishihama, Y. (2007). Protocol for micro-purification, enrichment, pre-fractionation and storage of peptides for proteomics using StageTips. *Nat Protoc* 2(8)**,** 1896-1906. doi: 10.1038/nprot.2007.261.

Rebets, Y., Kormanec, J., Lutzhetskyy, A., Bernaerts, K., Anné, J. (2017). Cloning and Expression of Metagenomic DNA in *Streptomyces lividans* and Subsequent Fermentation for Optimized Production. *Methods Mol Biol*.;1539:99-144. doi: 10.1007/978-1-4939-6691-2_8

Rebets, Y., Tsolis, K.C., Guðmundsdóttir, E.E., Koepff, J., Wawiernia, B., Busche, T., et al. (2018). Characterization of Sigma Factor Genes in *Streptomyces lividans* TK24 Using a Genomic

Library-Based Approach for Multiple Gene Deletions. *Front. Microbiol.* 9:3033**.**, doi:10.3389/fmicb.2018.03033

Ruckert, C., Albersmeier, A., Busche, T., Jaenicke, S., Winkler, A., Friethjonsson, O.H., et al. (2015). Complete genome sequence of *Streptomyces lividans* TK24. *J Biotechnol* 199**,** 21-22. doi: 10.1016/j.jbiotec.2015.02.004.

Schwanhausser, B., Busse, D., Li, N., Dittmar, G., Schuchhardt, J., Wolf, J., et al. (2011). Global quantification of mammalian gene expression control. *Nature* 473(7347)**,** 337-342. doi: 10.1038/nature10098.

Sianidis, G., Pozidis, C., Becker, F., Vrancken, K., Sjoeholm, C., Karamanou, S., et al. (2006). Functional large-scale production of a novel Jonesia sp. xyloglucanase by heterologous secretion from *Streptomyces lividans*. *J Biotechnol* 121(4)**,** 498-507. doi: 10.1016/j.jbiotec.2005.08.002.

Tenconi, E., Guichard, P., Motte, P., Matagne, A., and Rigali, S. (2013). Use of red autofluorescence for monitoring prodiginine biosynthesis. *J Microbiol Methods* 93(2)**,** 138-143. doi: 10.1016/j.mimet.2013.02.012.

Tsolis, K.C., and Economou, A. (2017). Quantitative Proteomics of the E. coli Membranome. *Methods Enzymol* 586**,** 15-36. doi: 10.1016/bs.mie.2016.09.026.

Tsolis, K.C., Tsare, E.-P., Orfanoudaki, G., Busche, T., Kanaki, K., Ramakrishnan, R., et al. (2018). Comprehensive subcellular topologies of polypeptides in *Streptomyces* (SToPSdb). *Microbial cell Factory* 17(1):43. doi: 10.1186/s12934-018-0892-0

van den Hombergh, J.P., van de Vondervoort, P.J., Fraissinet-Tachet, L., and Visser, J. (1997). Aspergillus as a host for heterologous protein production: the problem of proteases. *Trends Biotechnol* 15(7)**,** 256-263.

Vizcaino, J.A., Csordas, A., del-Toro, N., Dianes, J.A., Griss, J., Lavidas, I., et al. (2016). 2016 update of the PRIDE database and its related tools. *Nucleic Acids Res* 44(D1)**,** D447-456. doi: 10.1093/nar/gkv1145.

Wagner, G.P., Kin, K., and Lynch, V.J. (2012). Measurement of mRNA abundance using RNA-seq data: RPKM measure is inconsistent among samples. *Theory Biosci* 131(4)**,** 281-285. doi: 10.1007/s12064-012-0162-3.

Walker, D., Mosbahi, K., Vankemmelbeke, M., James, R., and Kleanthous, C. (2007). The role of electrostatics in colicin nuclease domain translocation into bacterial cells. *J Biol Chem* 282(43)**,** 31389-31397. doi: 10.1074/jbc.M705883200.

Westers, L., Westers, H., and Quax, W.J. (2004). *Bacillus subtilis* as cell factory for pharmaceutical proteins: a biotechnological approach to optimize the host organism. *Biochim Biophys Acta* 1694(1-3)**,** 299-310. doi: 10.1016/j.bbamcr.2004.02.011.

Xu, J., Wang, L., Ridgway, D., Gu, T., and Moo-Young, M. (2000). Increased heterologous protein production in Aspergillus niger fermentation through extracellular proteases inhibition by pelleted growth. *Biotechnol Prog* 16(2)**,** 222-227. doi: 10.1021/bp000006s.

3.8. Supplementary Information

Supplementary Material for this article can be found online at: https://www.frontiersin.org/articles/10.3389/fmicb.2018.01174/full#supplementary-material

Table S4. Strain phenotyping data of protease deletion strain versions before mRFP integration. Maximum specific growth rate (μ_{max}), cell-dry-weight concentration at the transition to stationary phase (CDW) as well as cultivation time duration until this point (t_{batch}) and background fluorescence at excitation 550 nm and emission 589 nm ($fluor_{background}$) are provided as values and standard deviation for biological replicates of each strain and relative to the wild-type (WT) of the corresponding experiment. See method section for detailed methodology.

ID Jülich	**Deleted gene locus**	**Deleted gene product**	**mRFP integration**	**μ_{max}**		**CDW**		**t_{batch}**		**$fluor_{background}$**	
				value	range	value	std	value	std	value	std
SL000	-	-	-	1,00	0,03	1,00	0,07	1,00	0,07	1,00	0,09
SL318	SLIV_09985	Integral membrane Peptidase S8, subtilisin-related protein	-	0,86	0,02	0,96	0,00	1,06	0,06	0,77	0,00
SL317	SLIV_20750	Integral membrane ATP-dependent zinc metalloprotease FtsH (EC 3.4.24.-)	-	0,55	0,05	0,63	0,02	1,53	0,15	0,61	0,06
SL316	SLIV_10535	Integral membrane ATP-dependent zinc metalloprotease FtsH3	-	0,84	0,06	1,02	0,02	1,06	0,13	0,88	0,01
SL315	SLIV_11275	Secreted Neutral zinc metalloprotease	-	0,92	0,03	1,05	0,01	1,05	0,04	0,82	0,03
SL314	SLIV_11270	Secreted Neutral zinc metalloprotease	-	0,93	0,03	0,81	0,00	0,99	0,01	0,78	0,02
SL329	SLIV_15325	Secreted Peptidase, Leupeptin-inactivating enzyme 1	-	0,89	0,06	0,84	0,04	1,14	0,00	0,87	0,02
SL328	SLIV_09410	secreted peptidase	-	0,84	0,05	1,03	0,02	1,23	0,03	1,00	0,02
SL327	SLIV_24720	secreted Protein containing Tachylectin 2 domain	-	0,81	0,03	0,97	0,01	1,19	0,05	0,96	0,03
SL326	SLIV_34120	Secreted Probable subtilase-type protease inhibitor	-	0,75	0,01	0,87	0,07	1,40	0,04	1,14	0,02
SL320	SLIV_28740	Integral membrane Stomatin family	-	0,91	0,04	1,03	0,03	1,01	0,11	0,86	0,05
SL307	SLIV_02150	Secreted extracellular small neutral protease	-	0,97	0,01	0,75	0,04	0,96	0,02	0,90	0,06
SL321	SLIV_17030	Secreted Peptidase M1, alanine aminopeptidase/leukotriene A4 hydrolase	-	0,98	0,02	0,84	0,02	1,07	0,06	0,87	0,02
SL319	SLIV_10025	Integral membrane T7SS peptidase S8A, mycosin-1, component of T7S export system	-	0,88	0,04	0,84	0,05	1,11	0,04	0,83	0,01
SL309	SLIV_11935	Cytoplasmic ATP-dependent serine protease Lon	-	0,97	0,03	1,01	0,01	0,91	0,01	0,87	0,02

Table S5. Strain phenotyping data of protease deletion strain versions after mRFP integration. Annotation is identical to Table S1, in exception to mRFP fluorescence is provided as CDW specific values and standard deviation ($fluor_{CDWspecific}$). Again all values are normalized by the WT strain without mRFP integration.

ID Jülich	Deleted gene locus	Deleted gene product	mRFP integration	μ_{max}		CDW		t_{batch}		$fluor_{CDWspecific}$	
				value	range	value	std	value	std	value	std
SL000	-	-	-	1,00	0,07	1,00	0,01	1,00	0,06	1,00	0,13
SL348	-	-	x	0,86	0,06	1,01	0,03	1,13	0,08	16,95	1,32
SL350	SLIV_09985	Integral membrane Peptidase S8, subtilisin-related protein	x	0,82	0,02	1,05	0,02	1,04	0,05	14,64	0,22
SL351	SLIV_20750	Integral membrane ATP-dependent zinc metalloprotease FtsH (EC 3.4.24.-)	x	0,64	0,06	0,43	0,02	1,54	0,02	29,53	0,44
SL352	SLIV_10535	Integral membrane ATP-dependent zinc metalloprotease FtsH3	x	0,84	0,05	0,92	0,01	1,11	0,07	19,31	1,43
SL353	SLIV_11275	Secreted Neutral zinc metalloprotease	x	0,96	0,04	1,16	0,03	1,03	0,02	11,42	0,44
SL354	SLIV_11270	Secreted Neutral zinc metalloprotease	x	0,82	0,03	1,10	0,01	0,99	0,02	11,20	0,32
SL355	SLIV_15325	Secreted Peptidase, Leupeptin-inactivating enzyme 1	x	0,94	0,05	1,13	0,01	0,95	0,03	7,92	0,93
SL356	SLIV_09410	secreted peptidase	x	0,77	0,07	0,95	0,05	1,12	0,09	15,75	0,54
SL357	SLIV_24720	secreted Protein containing Tachylectin 2 domain	x	0,83	0,02	1,07	0,01	1,00	0,02	13,28	0,23
SL358	SLIV_34120	Secreted Probable subtilase-type protease inhibitor	x	0,85	0,07	0,99	0,02	1,07	0,08	15,82	0,53
SL359	SLIV_28740	Integral membrane Stomatin family	x	1,05	0,07	0,79	0,03	1,02	0,05	15,81	3,87
SL360	SLIV_02150	Secreted extracellular small neutral protease	x	0,86	0,06	1,06	0,07	1,01	0,06	10,42	0,92
SL361	SLIV_17030	Secreted Peptidase M1, alanine aminopeptidase/leukotriene A4 hydrolase	x	0,92	0,05	0,91	0,00	0,98	0,03	13,66	2,37
SL362	SLIV_10025	Integral membrane T7SS peptidase S8A, mycosin-1, component of T7S export system	x	0,76	0,06	0,91	0,06	1,20	0,04	16,63	1,01
SL363	SLIV_11935	Cytoplasmic ATP-dependent serine protease Lon	x	0,96	0,01	1,20	0,00	0,97	0,04	6,07	0,33

3.9. References for the Supplementary Information

Tsolis, K.C., Tsare, E.-P., Orfanoudaki, G., Busche, T., Kanaki, K., Ramakrishnan, R., et al. (2018). Sub- cellular topologies of polypeptides in *Streptomyces* (SToPSdb). *Microbial cell Fact.* 17(1)**,** 43. doi: 10.1186/s12934-018-0892-0.

Chapter 4

Engineering of *Streptomyces lividans* for Heterologous Expression of Secondary Metabolites Gene Clusters

Yousra Ahmed[1], Yuriy Rebets[1], Marta Rodríguez Estévez[1], Josef Zapp[2], Andriy Luzhetskyy[1, 3,*]

4. *Streptomyces lividans* Chassis

4.1. Abstract

Background: Heterologous expression of secondary metabolism gene clusters is used to achieve better production of desired compounds, activate cryptic gene clusters, manipulate clusters from genetically unamenable strains, obtain natural products from uncultivable species, create new unnatural pathways, etc. Several *Streptomyces* species are genetically engineered to be used as hosts for heterologous expression of gene clusters. *S. lividans* TK24 is one of the most studied and genetically tractable Actinobacteria, which remain untapped. It was therefore important to generate *S. lividans* chassis strains with clean metabolic background.

Results: In this study we generated a set of *S. lividans* chassis strains by deleting endogenous gene clusters and introducing additional φC31 *attB* loci for site-specific integration of foreign DNA. Beside the simplified metabolic background, the engineered *S. lividans* have better growth characteristics in liquid production medium. The utility of developed strains was validated by expressing four secondary metabolism gene clusters responsible for production of different classes of natural products. Engineered strains were found to be superior to the parental strain in producing heterologous natural product. At the same time, *S. lividans* based strains were better producers of amino acid based natural products than other tested common hosts. Expression of *Streptomyces albus* subsp. *chlorinus* NRRL B-24108 genomic library in modified *S. lividans* ΔYA9 and *S. albus* Del14 strains resulted in production of 7 potentially new compounds, only one of which was produced in both strains.

Conclusion: The constructed *S. lividans* based strains are a great complement to the panel of heterologous hosts for the actinobacterial secondary metabolites gene expression. The expansion of a number of such engineered strains will contribute to the higher success rate in isolating new natural products originated from expression of genomic and metagenomic libraries thus will raise chances to obtain novel biologically active compounds.

4.2. Background

Streptomyces is considered as one of the most explored genera of Actinobacteria. These bacteria produce large number of pharmaceutically important compounds as a part of their secondary metabolism [1, 2]. The secondary metabolites biosynthesis genes are typically grouped into clusters that include structural genes, coding for biosynthetic enzymes, as well as regulatory and resistance/transport genes. Such organization simplifies the identification of biosynthetic genes and facilitates their cloning and expression in heterologous hosts. This heterologous expression

is especially useful, since it allows fast and simple manipulations with gene cluster of interest, which otherwise can be difficult to handle in natural producer since many Actinobacteria are poorly genetically tractable. Heterologous expression is often used to activate silent gene clusters [3, 4], generate unnatural metabolites by combinatorial biosynthesis or mutasynthesis [5, 6], and to obtain better yield of compound of interest [7]. To get the best outcome from this procedure, appropriate expression host, that is well studied, genetically amenable as well as provides a high pool of precursors and energy, have to be used. Therefore, several *Streptomyces* species such as *S. coelicolor*, *S. lividans*, *S. avermitilis* and *S. albus* are used as hosts for expression of gene clusters cloned from actinomycetes. Originally, the native strains were employed in such experiments. However, often this caused difficulties in identification of produced metabolite or interaction between expressed and endogenous pathways resulting in aberrant products formation [8, 9]. To overcome these complications the first modified host strain *S. coelicolor* CH999 deficient in production of internal natural products was constructed [10]. This strain was generated by deleting the actinorhodin and inactivating undecylprodigiosin gene clusters and was primary used to study the functional peculiarities of type II PKS system. With the entrance to genomic area in Actinobacteria research the idea of "clean" chassis strain became feasible. As result, several improved hosts were constructed with simplified metabolic background, enhanced supply of precursors and increased productivity of desired exogenous secondary metabolites. Komatsu and Ikeda generated the genome-minimized derivatives of *S. avermitilis*: SUKA5, SUKA17 and SUKA22 [11]. *S. avermitilis* SUKA5 strain has deletion of the oligomycin biosynthetic gene cluster, as well the left subtelomeric region covering avermectin and filipin biosynthetic gene clusters. SUKA 17 and SUKA22 are isogenic strains that contain additional deletions of neopentalenolacton, geosmin and a carotenoid encoding gene clusters. Different heterologous secondary metabolites were successfully produced in engineered *S. avermitilis* strains including aminoglycoside streptomycin, β-lactam cephamycin C and the macrocyclic compound pladienolide [11]. Gomez-Escribano and Bibb developed *S. coelicolor* M1152 and M1154 strains with deletion of four internal gene clusters (*act*, *red*, *cpk*, coelimycin; and *cda*, calcium-dependent antibiotic). In addition to that, M1152 carries *rpoB* (rifampicin resistance) mutation and M1154 has *rpoB* and *rpsL* (streptomycin resistance) mutations [12]. These mutations were reported to enhance the production of secondary metabolites in Actinobacteria due to increased RNA-polymerase promoter affinity (*rpoB*) and induction of protein synthesis in the stationary growth phase (*rpsL*) [13, 14]. *S. coelicolor* strains were widely used to express different types of secondary metabolism biosynthetic gene clusters (reviewed in [5]). *S. albus* J1074 derivative *S. albus* Del14 with deletion of 15 endogenous gene clusters is another

engineered host strain with the "clean" genetic and metabolic background for secondary metabolite gene clusters expression [15]. *S. albus* Del14 was successfully used to activate cryptic type I PKS gene cluster from metagenomic library resulting in production of pyridinopyrone. Furthermore, two other cryptic clusters from *Frankia alni* ACN14a and *Frankia* sp. CcI3 were successfully expressed in *S. albus* Del14 leading to production of salicylic acid, fralnimycin (from *Frankia alni* ACN14a), bhimamycin A and aloesaponarin II (from *Frankia* sp. CcI3) [15].

S. lividans is closely related to *S. coelicolor* and both are well characterized Actinobacteria species. Since *S. lividans* is accepting methylated DNA and has low endogenous protease activity, it is, therefore, often a preferable host for heterologous expression of secondary metabolite gene clusters and secreted production of recombinant proteins [16-20]. Ziermann and co-authors have generated two hosts *S. lividans* K4-114 and K4-155 by removing the entire actinorhodin gene cluster from chromosome of *S. lividans* TK24 [21]. The expression of genes for biosynthesis of erythromycin precursor 6-deoxyerythronolide *B (*6-dEB) in *S. lividans* K4-114, K4-155 and *S. coelicolor* CH999 resulted in approximately same level of compound accumulation. Similar, *S. lividans*-based host with deletion of the *act* and *red* gene clusters was successfully used to produce granaticin [22]. Recently, the heterologous production of mithramycin in *S. lividans* was reported to be significantly improved by deleting the *act, red* and *cda* gene clusters [23]. However, no deep engineering like in the case of *S. albus* Del14 or *S. avermitilis* SUKA strains was performed with *S. lividans* TK24 in order to generate the completely "clean" strain.

The success with the engineered *Streptomyces* strains prompted us to generate a new *S. lividans* based host with clean metabolic profile and thus simplified detection and purification of new compounds. Herein, we report construction of optimized *S. lividans* host strains by removing 11 endogenous secondary metabolite gene clusters, accounting for 228.5 kb of the chromosome. Further modification of the strains to improve the production level was performed by introducing additional integration sites (*attB* sites) for φC31-based vectors.

4.3. Results

Transcriptome based identification of actively expressed secondary metabolite gene clusters

The genome of *S. lividans* TK24 was sequenced and twenty-five gene clusters, potentially involved in biosynthesis of secondary metabolites were identified (Table 1S) [24]. For long time

the only secondary metabolites produced by the strain and only under certain laboratory conditions were considered to be actinorhodin and undecylprodigiosin and the remaining gene

Table 1. Secondary metabolite gene clusters deleted within the chromosome of *S. lividans* TK24 and resulting strains.

Cluster	Product	Coordinates of deleted genes	Size of the gene cluster (kb)	Size of deleted region (Kb)	Strain ID
Cluster 10	Undecylprodigiosin	09100 - 09185	≈47.65	≈ 26.5	**ΔYA1**
Cluster 14	Actinorhodin	12885 - 12965	≈41.55	≈ 15.3	**ΔYA2**
Cluster 17	Melanin	24135 - 24185	≈9.17	≈ 8	**ΔYA3**
Cluster 13	Terpene	12220 - 12265	≈21.67	≈ 11.7	**ΔYA4**
Cluster 19	Germicidin	31668 - 31860	≈42.26	≈ 40	**ΔYA5**
Cluster 15	CDA	21535 - 21540	≈80.87	≈ 27	**ΔYA6**
Cluster 6	Coelimycin	06755 - 06775	≈79.56	≈ 20	**ΔYA7**
Cluster 2	Coelibactin	00885 - 00925	≈73.28	≈ 17	**ΔYA8**
Cluster 24	eicosapentaenoic acid	37285 - 37305	≈53.55	≈ 13	**ΔYA9**
Cluster 21	Coelichelin	35395 - 35490	≈51.22	≈ 32::attB2	**ΔYA10**
Cluster 5	NRPS	06215 - 06270	≈48.64	≈18::attB3	**ΔYA11**

clusters were thought to be silent. In order to estimate the transcription of secondary metabolism genes we performed analysis of RNAseq data from previous described dataset obtained for *S. lividans* TK24 cultivated in minimal medium in mini-bioreactor [25]. The transcription of secondary metabolism genes was calculated as average RPKM (reads per kilo base per million mapped reads) of all genes within the particular cluster identified by antiSMASH with ClusterFinder algorithm for clusters border prediction (Table 1S). We found that majority of secondary metabolism gene clusters are transcriptionally active during stationary phase of growth. The most expressed gene clusters are responsible for biosynthesis of putative rhizobactin-like siderophore (number 11, average RPKM 918.1) [26, 27] and desferrioxamine (number 16, average RPKM 844.3) [28]. The uncharacterized terpene cluster (number 13) comes next with average 803.1 RPKM; then cluster number 21 coding for biosynthesis of coelichelin with average RPKM of 707.8. The later compound also acts as hydroxamate siderophore [29]. In order to detect the production of endogenous secondary metabolites by *S. lividans* TK24 the strain was grown in SG medium and the metabolites were extracted and analyzed by LC-MS (Fig. 1). As result, we could identify undecylprodigiosin (peak with RT of 14.4 min) and coelibactin (peak with RT of 12.9 min) in the extract of *S. lividans* TK24 (Fig. 1, Fig. 1S). Additionally, several compounds can be found in the extract but we were not able to assign them to particular *S. lividans* metabolites. Actinorhodin, CDA, coelimycin P1 and other typical

secondary metabolites of the strain were not found in the extract despite the observed transcriptional activity of corresponding gene clusters.

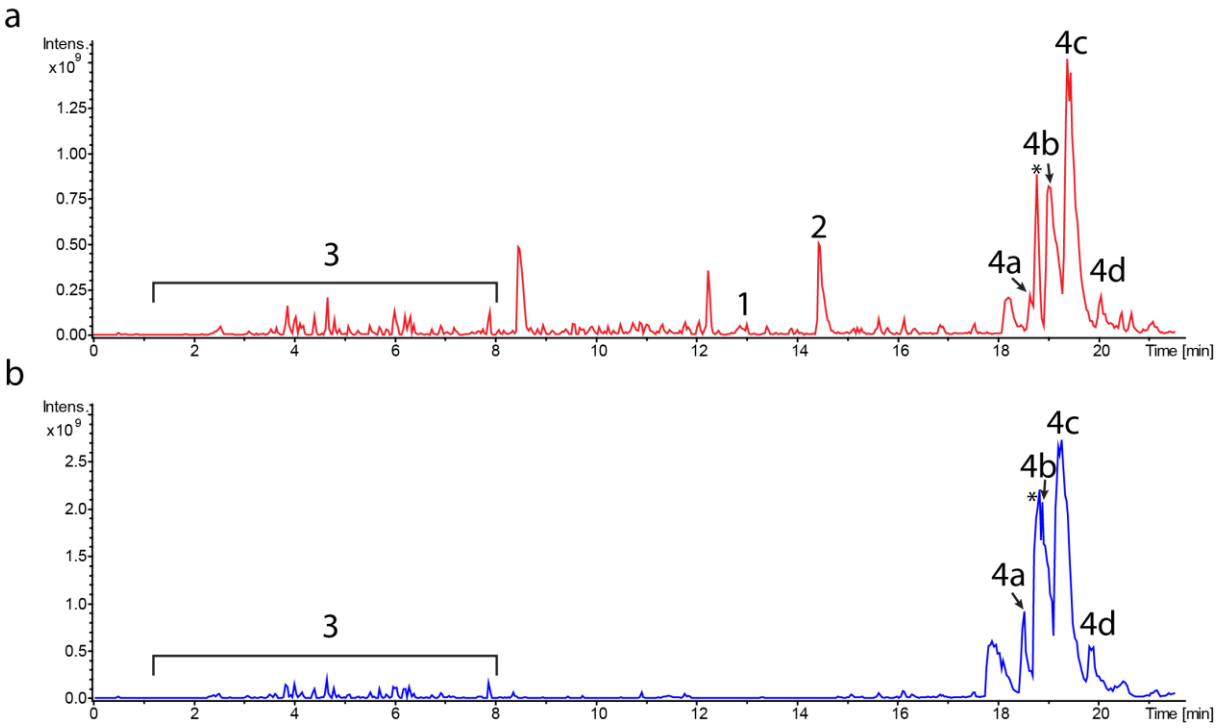


Figure 1. BPC chromatogram of ethyl-acetate extracts of *S. lividans* TK24 (red) and *S. lividans* YA9 (blue) grown in SG medium. The identified compounds are indicated as: 1 – coelibactin; 2 – undecylprodigiosin; 3 – media components; 4 –monoglyceride lipids found in many other Streptomycetes extracts. * - unidentified contamination from the method (column bleed).

Generation of *S. lividans* ΔYA9 chassis strain

To simplify the metabolic profile of the strain, nine secondary metabolite biosynthesis gene clusters were deleted within the chromosome of *S. lividans* TK24, including *act*, *red*, and *cpk* clusters (Table. 1). The markerless deletions were achieved by using the *S. lividans* TK24 genomic BAC library [30] and Red/ET recombination technique [31] combined with iterative marker excision system (IMES) [32]. The recombinant BACs were constructed by replacing most of the targeted gene cluster or the core genes (in case the single BAC was not covering the entire cluster) with the IMES apramycin resistance cassette. The BAC clone was transferred into *S. lividans* and double crossover was selected with the help of β-glucuroidase mediated blue-white selection [30]. To perform the sequential removal of the nine gene clusters in a single strain, deletions were intertwined with the excision of the resistance marker by expressing φC31 phage integrase gene, leaving a 48 bp inactive scar in place of a targeted gene cluster [32]. The final engineered strain, lacking 9 gene clusters with a deletion of 178.5 kbp of chromosome, was designated as *S. lividans* ΔYA9. The comparison of metabolic profiles of TK24 and ΔYA9

clearly shows for ΔYA9 an almost complete loss of compounds detected in the extract of parental strain, including coelibactin (**1**) and undecylprodigiosin (**2)** (Fig. 1). Only one group of closely related hydrophobic metabolites (**4a-d**) (Fig. 1), also found in many other Streptomycetes but unidentified so far, seems to be unaffected, if not even overproduced after deletion of the gene clusters. We have isolated compounds **4a-c** and partially elucidated their structures by NMR analysis. As result, compounds **4a-d** seems to be mono-substituted glyceride lipids with the highly unusual unsaturated β-hydroxy fatty acids attached (Fig. 2S). However, due to the highly hydrophobic nature of these compounds we were not able to purify enough of the material to solve the structure of **4a-d** in full. Although, the origin of these compounds is not clear they do not interfere the identification of compounds resulted from heterologous gene cluster expression.

Introduction of additional *attB* sites into chromosome of *S. lividans* ΔYA9 strain

Most of the gene cluster expression constructs are based on actinobacterial vectors integrating into host chromosome. Among them the most popular are BAC or cosmids vectors based on φC31 and VWB actinophages integrations systems [33-35]. *S. lividans* TK24 has one major and three *pseudo attB* sites for integration of the φC31-based vectors [36]. To increase the production level in our engineered strain we decided to introduce two additional *attB* sites into *S. lividans* ΔYA9. Recombinant BACs were constructed in the way to integrate *attB* sites in place of gene clusters 5 (uncharacterized NRPS) and 21 (coelichelin) (see methods section). The final constructs were sequentially introduced into *S. lividans* ΔYA9 strain by conjugation and secondary cross-over strains were selected resulting in markerless replacement of targeted gene cluster by *attB* site.

As results, we generated three different *S. lividans* strains, ΔYA9 with deleted 9 gene clusters and one native *attB* site, *S. lividans* ΔYA10 with 10 gene clusters deleted and two *attB* sites and *S. lividans* ΔYA11 with lacking 11 gene clusters and containing three *attB* sites (Table 1). The final strain *S. lividans* ΔYA11 lacks all together 228.5 kb of chromosome (2.7%).

Growth rate and sporulation of *S. lividans* strains

Actinobacterial secondary metabolites are generally accepted to have no influence on strain growth and viability, except if they act as siderophores [37]. To examine the growth rate of the obtained strains, 100 mg of biomass of *S. lividans* TK24, *S. lividans* ΔYA9 and *S. lividans* ΔYA11 pre-cultures were used to inoculate 70 ml of DNPM medium. One ml samples were harvested every 24 hours for 6 days and the wet biomass was weighted. In general, ΔYA9 and

ΔYA11 are accumulating slightly more biomass than TK24 (Fig. 2a). However, the growth rate was significantly affected in the case of engineered strains. After 72 hours of cultivation *S. lividans* TK24 reached the end of exponential phase and entered into the stationary phase of growth. In the case of ΔYA9 and ΔYA11 strains the stationary phase of growth starts at between 24 and 48 hours of cultivation. The engineered strains did not show any differences in growth and morphological features on solid MS medium (Fig. 2b).

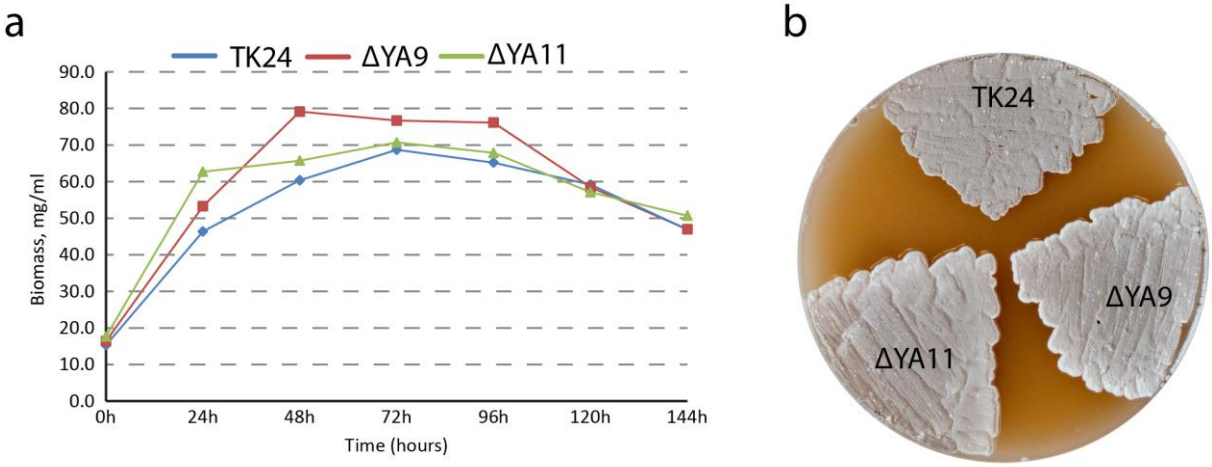


Figure 2. Effect of clusters deletion on fitness and growth parameters. **(a)** Growth curve analysis of the wild type *S. lividans* TK24 (blue) and the mutants *S. lividans* ΔYA9 (red) and *S. lividans* ΔYA11 (green) grown for 6 days in DNPM media. Samples were taken every 24 hours and the wet biomass was measured. **(b)** Phenotype of the *S. lividans* TK24, *S. lividans* ΔYA9, and *S. lividans* ΔYA11 after 6 days growth on MS media.

Validation of engineered *S. lividans* strains performance

To examine the effect of gene cluster deletion and multiple-copy of *attB* sites on the production level of exogenous secondary metabolites, three gene clusters, representing different classes of biosynthetic pathways were chosen. Constructs carrying gene cluster of tunicamycin [38], griseorhodin [39], and deoxycinnamycin (derivative of cinnamycin with deletion of OH group) [40] (Fig. 3) were introduced into parental and engineered *S. lividans* strains and two other engineered hosts: *S. albus* Del14 [15] and *S. coelicolor* M1154 [12]. We observed that conjugation frequency in *S. lividans* TK24, *S. lividans* ΔYA9 and *S. albus* Del14 was similar and higher than in *S. lividans* ΔYA10, *S. lividans* ΔYA11 and *S. coelicolor* M1154. The latter strain had the lowest ex-conjugants count. Obtained recombinant strains were inoculated in two different production media: modified SG and DNPM (see materials and methods). Strains were cultivated for 6 days and metabolites were extracted and analyzed by LC-MS.

The production level of the nucleoside antibiotic tunicamycin in engineered *S. lividans* strains was slightly higher than in the wild type TK24 in both media (Fig. 4a). At the same time, all engineered *S. lividans* strains produced at approximately same level, indicating that the copy

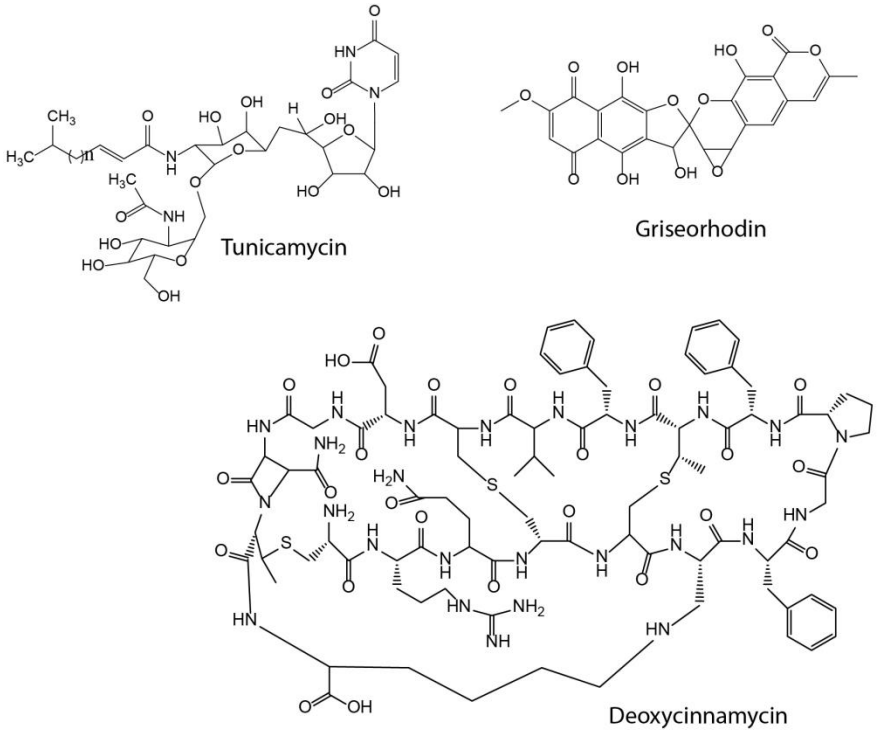


Figure 3. Structure of compounds described in the study.

number of the gene cluster has almost no influence on accumulation of tunicamycins. No production was observed in the case of *S. coelicolor* M1154. *S. albus* Del14 was superior to all other tested strains, superseding the best performing *S. lividans* ΔYA11 strain for around 20%.

Engineered *S. lividans* showed higher production level of lantipeptide deoxycinnamycin than other tested strains (Fig. 4b). The production level in *S. lividans* ΔYA9 was increased more than 3.5 and 3-fold (in SG and DNPM, respectively) compared to the parental TK24. *S. lividans* ΔYA10 and ΔYA11, containing one and two additional *attB* sites, respectively, showed further increase in accumulation of deoxycinnamycin. Among all tested strains the highest productivity was observed in *S. lividans* ΔYA10 with 4.5-fold (SG) and 3.5-fold (DNPM) increase compared to the wild type. In the case of *S. albus* Del14 accumulation of deoxycinnamycin was slightly lower or on par with *S. lividans* ΔYA9 when *S. coelicolor* M1154 produced less than other tested strains.

In the case of aromatic polyketide griseorhodin surprisingly no production was observed by *S. lividans* TK24 and *S. lividans* ΔYA11 when cultivated in SG media, originally developed for

polyketides production (Fig. 4c). Two other *S. lividans* strains ΔYA9 and ΔYA10 were able to produce this antibiotic in relative low amounts. *S. albus* Del14 showed highest production level of griseorhodin in SG medium. A different behavior was observed when strains were cultivated

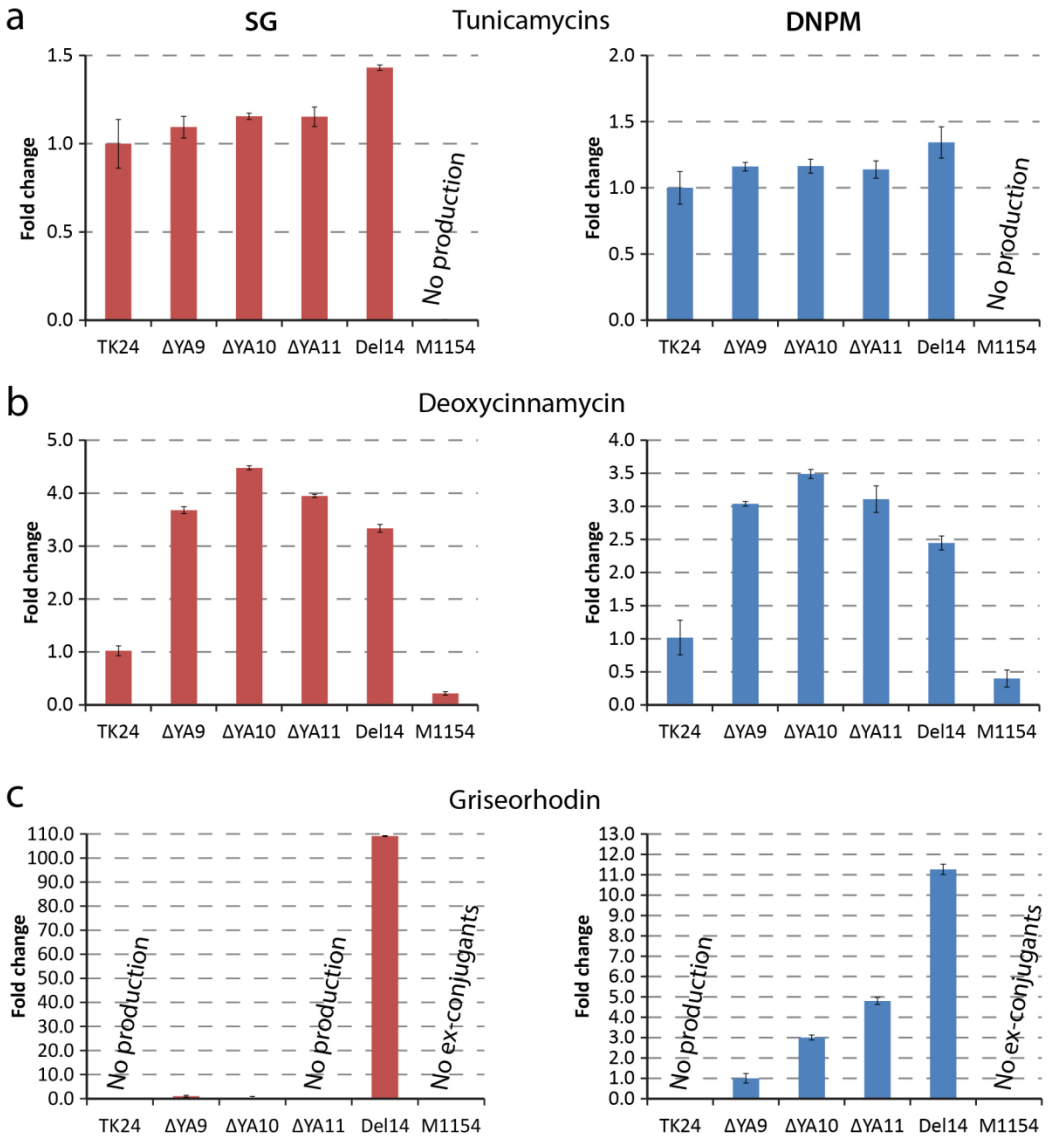


Figure 4. Production level of tunicamycin, deoxycinnamycin and griseorhodin in *S. lividans* TK24, *S. lividans* ΔYA9, *S. lividans* ΔYA10, *S. lividans* ΔYA11, *S. albus* Del14, *S. coelicolor* M1154. Three independent colonies from each strain were grown for 6 days in SG media (left) and DNPM (right). The production by *S. lividans* TK24 was taken as 100% in each particular experiment, except for griseorhodin, when production ΔYA9 was taken as 100%.

in DNPM medium. When *S. lividans* TK24 still lack the production of griseorhodin, strain YA9 deficient in large part of its own secondary metabolism was able to produce griseorhodin. The production level was further increased with the use of strains carrying additional *attB* sites (2-fold in ΔYA10 and almost 4-fold in ΔYA11 compared to *S. lividans* ΔYA9). However, like in the case of tunicamycin, the highest accumulation of griseorhodin was observed in *S. albus*

Del14 in both media. For *S. coelicolor* no ex-conjugants were obtained with the griseorhodin gene cluster construct after 20 attempts.

Table 2. Expression of *S. albus* subsp. *chlorinus* NRRL B-24108 BAC library clones in engineered host strains.

BAC	Type of cluster	*S. lividans* ΔYA9	*S. albus* Del14
4M20	Type I PKS	-	-
2O18	NRPS/siderophore	-	+
6O24	NRPS/siderophore	-	-
5A14	Siderophore	-	-
2P5	Type II PKS	+	+
1P15	Lantipeptide	-	-
4E8	NRPS	-	+
3D8	Siderophore	-	-
6E10	Type I PKS	-	-
3N1	Lantipeptide	-	-
1K1	NRPS	-	-
5K5	Lantipeptide	-	-
5F24	Phenazine	-	+
5H22	NRPS	-	-
2I4	NRPS	+	-

Expression of BAC library of *S. albus* subsp. *chlorinus* NRRL B-24108

The big advantage of engineered *Streptomyces* host strains is that simplified metabolic background allows fast identification of compounds produced by expressing unknown biosynthetic gene cluster(s). To test the feasibility of the developed strains we have used *S. albus* Del14 and *S. lividans* ΔYA9 to screen the BAC library of *S. albus* subsp. *chlorinus* NRRL B-24108 [41] for production of new secondary metabolites. For this purpose, 17 clones carrying secondary metabolite gene clusters predicted by antiSMASH were systematically expressed in *S. lividans* ΔYA9 and *S. albus* Del14 and extracts were screened for new compounds by LC-MS. As a result, six clones were successfully expressed either in *S. lividans* (two clones) or *S. albus*

Del14 (five clones) resulting in appearance of new peaks on LC-MS chromatogram of corresponding extracts (Table 2). One cluster, encoding a type II PKS with unknown products, was successfully expressed in both strains. The other five were active either in *S. lividans* or in *S. albus* host. Among the compounds produced by *S. albus* Del14 was the promising antimycobacterial compound nybomycin [42]. On the other hand, BAC clone 2I4 was expressed only in *S. lividans* ΔYA9 leading to accumulation of three new compounds with RT of 8.0 min (**1**), 8.2 min (**2**), 9.3 min (**3**) and *m/z* 259.1425 $[M+H]^+$, 243.1473 $[M+H]^+$ and 273.1587 $[M+H]^+$, respectively (Fig. 5). The BAC 2I4 clone harbors an uncharacterized gene cluster with the predicted NRPS core enzyme showing high degree of homology to several members of the

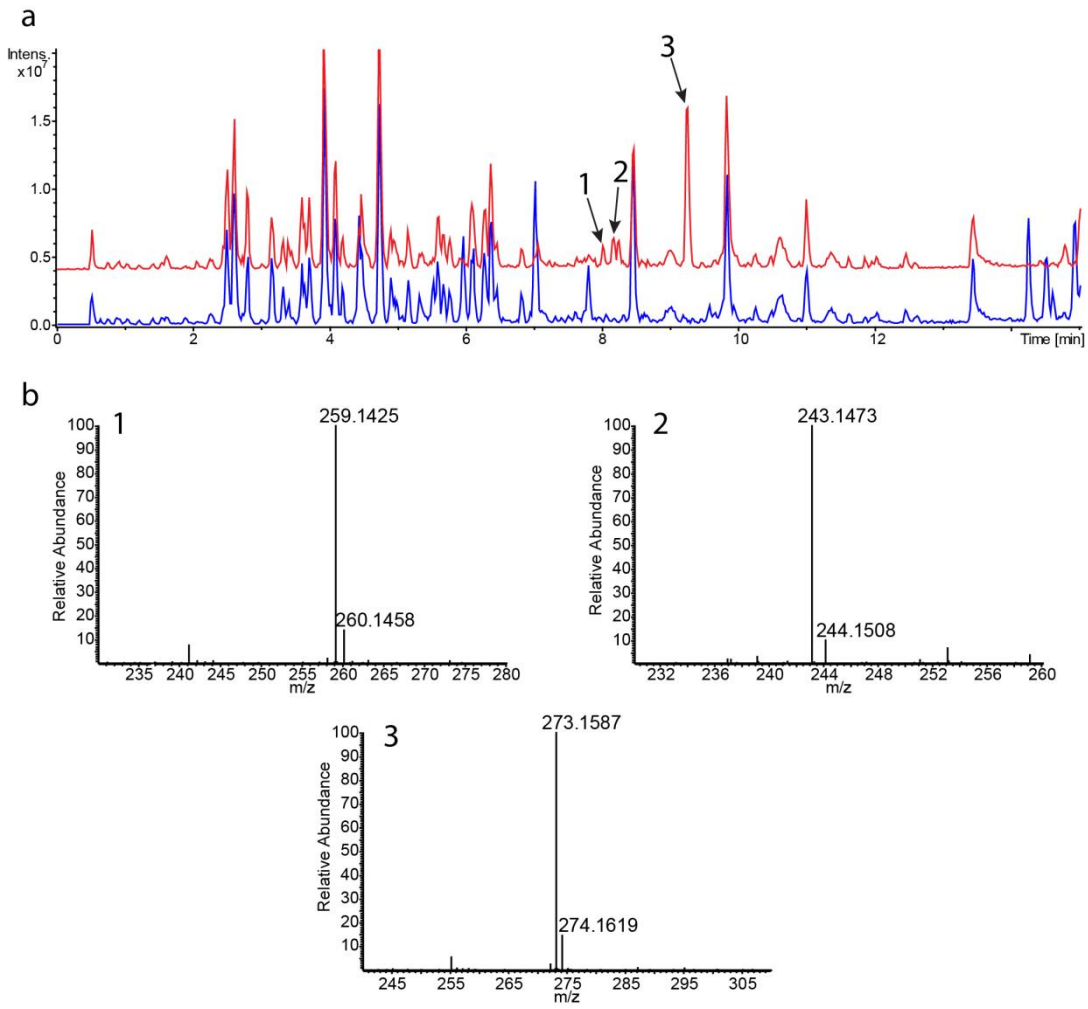


Figure 5. LC-MS analysis of *S. lividans* ΔYA9 with and without BAC2I4. Base peak chromatogram of ethyl acetate extracts of *S. lividans* ΔYA9 (blue) and *S. lividans* ΔYA9/BAC2I4 (red) grown in DNPM medium for 6 days. The new PBDs are marked as **1** (RT 8.0 min, *m/z* 259.14 [*M*+H]+), **2** (8.2 min, *m/z* 243.14 [*M*+H]+) and **3** (9.3 min, *m/z* 273.15 [*M*+H]+). Samples were separated with the 20 min gradient protocol.

pyrrolobenzodiazepines (PBDs) biosynthesis NRPSs, like those involved in biosynthesis of anthramycin and porothramycin [43, 44] (Fig. 3S, 4S, Table 2S). Based on the exact mass the closest DNP hits for the identified compounds are usabamycin A (calculated *m/z* 273.15583 $[M+H]^+$), B (calculated *m/z* 259.14018 $[M+H]^+$) and C (calculated *m/z* 243.14527 $[M+H]^+$) (Fig. 5, Fig. 4S). However, MS/MS patterns of produced compounds significantly differ from predicted for usabamycins, since all three compounds are losing water during fragmentation (Fig. 5S). Despite the fact that the exact nature of these metabolites still has to be established by more precise structural analysis, we can clearly say that they are new representatives of the PBD family of natural products.

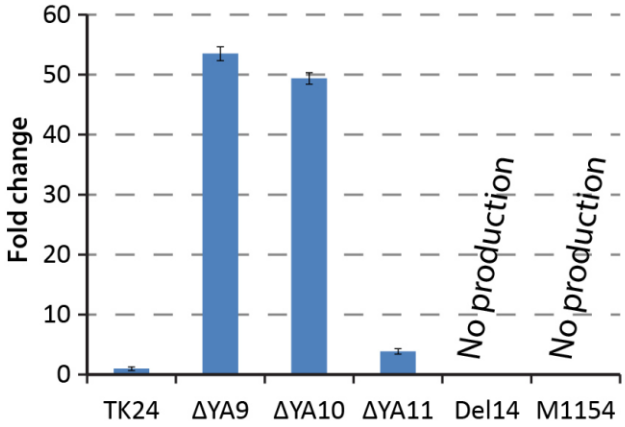


Figure 6. Production levels of PDB-like compound 3 in different hosts. The recombinant colonies were grown for 6 days in DNPM. The production by *S. lividans* TK24 was taken as 100%.

In order to investigate the influence of copy number on the production of identified PBDs, we introduced 2I4 BAC clone into *S. lividans* TK24, ΔYA9, ΔYA10 and ΔYA11. Strains were grown in DNPM and metabolites were extracted and analyzed with LC-MS. As result, the two engineered *S. lividans* ΔYA9 and *S. lividans* ΔYA10 showed 40-fold and 60-fold increase in accumulation of PBD-compounds, respectively (Fig. 6). In fact, the parental strain was producing trace amounts of these compounds. The strain with three *attB* sites ΔYA11 demonstrated 2-fold elevation in production level when compared to the parental TK24 that might be caused by the frequent loss of the gene cluster due to recombination events.

4.4. Discussion

Despite accessibility of a large number of actinobacterial strains, only several of them *S. coelicolor*, *S. lividans* and *S. albus* are well studied in terms of general physiology, biochemistry and genetics, and thus are used as heterologous hosts for antibiotics production. In rare cases

some industrial strains or their derivatives are also used as expression hosts. In all cases, the industrial and general strains have their advantages and disadvantages, among which poor genetic tractability (for instance *S. coelicolor*), high level of production of endogenous metabolites (industrial hosts), limited precursors supply (general hosts), cross-regulation and biochemical interactions between endogenous and heterologous gene clusters, etc. [9]. As a solution to some of these problems, several attempts to modify *Streptomyces* strains for better production of exogenous secondary metabolites were reported [11, 12, 15]. In all cases the endogenous secondary metabolism was targeted in order to minimize the genome of the host to improve the fitness and growth rate as well as redirect the precursors and regulatory factors to the heterologous biosynthetic pathway. As such, the engineered strains of *S. albus, S. coelicolor, S. avermitilis,* and *S. chattanoogensis* were generated [45]. The engineering of such "clean" metabolic background strains is if a high importance, first of all since they allow to shift from chemical to bioinformatics dereplication of bacterial natural products. The dereplication step is must have in current natural products discovery project due to the often re-discovery of same compounds. It is used to eliminate tedious and laborious process of compound isolation and structure elucidation by applying the high resolution MS techniques. On the other hand the chemical dereplication often excludes the true hits. With the advances in genome sequencing the identification of the gene cluster can be used as another criterion in the dereplication process [46]. On the other hand, when it goes about heterologous expression of genomic or metagenomic libraries the dereplication can be entirely shifted from chemical to bioinformatics by analysis the sequence and architecture of gene clusters in order to preselect the one with unique features. Expressing such pre-selected gene clusters in the host with well-defined or preferably "clean" secondary metabolites background will provide high chances for discovery of new natural products.

To expand the range of such host strains, herein, we report a development of new *S. lividans* based chassis for heterologous expression of secondary metabolism gene clusters. On top of the balanced and amenable *S. lividans* TK24 strain, we worked our way towards a simplified metabolic background by deleting nine endogenous gene clusters, while keeping track on the strains' fitness. We further engineered our streamlined strain to include two additional sites for integration of foreign DNA in place of two endogenous secondary metabolism gene clusters. The engineered strains efficiency in production of secondary metabolites was validated by expressing four heterologous gene clusters of different nature of produced secondary metabolites.

The general cloning strains that for last decades were (and still are) a preferred choice for heterologous gene cluster expression are *S. lividans* TK24, *S. coelicolor* M415, and *S. albus* J1074. Later two became a basis for generation of engineered chassis strains by inactivating a large portion of endogenous secondary metabolism. Also, there were several attempts to generate such hosts based on *S. lividans* TK24 by deleting *act* and *red* gene clusters [21, 22], still no systematic engineering of the strain to create a "clean" background host was reported till far. This is important first of all due to the fact that undecylprodigiosin and especially actinorhodin are produced by *S. lividans* only at certain and very limited number of conditions when other secondary metabolites obviously have more impact on heterologous natural products biosynthesis [47]. With its fast growth and easy to handle in general *S. lividans* is a preferable host strain for heterologous production of proteins and secondary metabolites.

To fulfill the gap in the range of engineered hosts we have chosen to delete eleven gene clusters from the chromosome of *S. lividans* TK24. Typically two different approaches are considered in chassis strain construction. The first one is involving a deletion of either empirically chosen large region of chromosome that also include several gene clusters for secondary metabolites and avoiding the core genome region. This method is often accompanied by bioinformatic analysis of the essentiality of the genes in selected region in order to avoid fitness costs [11, 45]. In such way *S. avermitilis* and *S. chattanoogensis* chassis strains were generated. Different approach was used to create *S. coelicolor* M1154 and *S. albus* Del14 strains [12, 15]. In this case a precise selection and sequential deletion of a set of gene clusters for endogenous secondary metabolites biosynthesis was performed. This approach allows avoiding any problems with the loss of genes important for growth, regulation and morphological development and produces the chassis strain with clean metabolic profile.

To generate *S. lividans* chassis strains the gene clusters for deletion were selected based on high transcriptional level and positive detection of associated compounds in production media. As a result, all PKS and NRPS encoding clusters were removed from the chromosome of the strain except for type II PKS *whiE* gene cluster (number 12), which is responsible for the spores' pigment biosynthesis [48]. The *whiE* cluster is ectopically expressed in aerial mycelium right before sporulation, and since *S. lividans* is not sporulating in liquid culture, is not in competition for CoA substrates. The ectoine and desferoxamine biosynthesis gene clusters were also kept in the genome of engineered strains because of the essential role of these compounds in physiological responses to the growth under osmotic stress and low iron conditions, respectively [26, 49]. Deletion of the desferoxamine biosynthesis gene cluster had a deleterious effect in *S.*

albus, showing its essentiality even under normal growth condition [32]. Similar effect was observed when we attempted to delete gene cluster annotated as such involved in biosynthesis of terpene (number 3). Despite being able to obtain the double cross-over clones, they were found to grow poorly after re-plating on fresh medium. This gene cluster showed 100% similarity to the hopene biosynthesis genes in *S. coelicolor* [50]. The hopanoids are produced by *S. coelicolor* during transition from substrate to aerial mycelia growth on solid medium but are not accumulated in liquid culture [51]. These compounds are components of the cytoplasmic membrane and their function is to reduce the stress in aerial hyphae by decreasing the water permeability across the cell membrane. Hopanoids have been reported to be not essential for growth of *S. scabies* [52], although we have observed their strong influence on growth of *S. lividans* lacking the respective gene cluster. Thus the cluster 3 remained within the genomes of engineered strains. Out of remaining 5 terpene gene clusters we deleted cluster 13 coding for albaflavenone biosynthesis [53]. Two gene clusters (number 23 and 25, coding for biosynthesis of isorenieratene and isoprenoid, respectively) are not covered by the BAC clones in the used *S. lividans* genomic library [30]. Two remaining terpene clusters are coding for geosmin (number 8) and uncharacterized terpene (number 1, absent in *S. coelicolor* genome) biosynthetic enzymes and were left in engineered strains. One of the major contaminant in the extract of *S. lividans* TK24 is melanin that gives dark color to the media and interferes with downstream processing of produced compound of interest, especially after long cultivation. This reasoned us to include the melanin gene cluster (number 17) as one of the major candidates for deletion, despite the low level of its transcription under tested conditions.

For Actinobacteria genome engineering mostly the Cre/*loxP* recombination system was used [54-57]. However, after recombination event it leaves the active scar that can be recognized by the recombinase thus significantly complicating the next engineering step by large chromosomal rearrangements. To generate *S. lividans* strains the IMES system was used [32]. This system allows recycling the resistance marker in the same genomic background without any interference with the remaining scar in the chromosome. As a result, we constructed a set of strains *S. lividans* ΔYA 9, *S. lividans* ΔYA10 and *S. lividans* ΔYA11 carrying deletions of 9, 10, and 11 endogenous gene clusters, respectively, with the last two strains harboring additional *attB* sites for heterologous cluster integration. These strains beside the clean metabolic background have slightly better growth characteristics in liquid production medium possibly due to smaller size of chromosome and lack of materials and energy draining into secondary metabolism.

To assess the usefulness of the engineered *S. lividans* strains as a heterologous host in comparison with the well-established chassis like *S. coelicolor* M1154 and *S albus* Del14 we introduce several characterized gene clusters into these strains. As of a general observation, the conjugation frequency in all tested strains was different. In the case of *S. albus* Del14, *S. lividans* TK24 and *S. lividans* ΔYA9 the conjugation efficiency was comparable. However, the introduction of additional *attB* sites impaired the gene transfer in the case of *S. lividans* ΔYA10 and *S. lividans* ΔYA11. Similar behavior was observed in previous study in our group, when *S. albus* SAM2 strain with only one *attB* site in the chromosome demonstrated highest conjugation efficiency when compared to other *S. albus* strains with multiple integration sites [58]. At least in *S. albus* case this phenomenon was postulated to be not caused by recombination between different *attB* sites within the same genome and is rather related to the efficiency of integrase enzyme to simultaneously integrate several copies of the vector DNA. However, the scenario when the recombination between the integrated sequences occurs cannot be excluded. In the original study pSET152 vector was used, that is only 5.5 kbp in size. The gene clusters provide a much larger template for the recombination that might result in deletion of a part of the chromosome, leading to lethality of some portion of transconjugants.

Although, the engineered *S. lividans* strains performed better than the parental TK24, they also demonstrated a larger degree of deviation in production level between clones. The degree of instability increased with the increase of number of *attB* sites in the strains' chromosome. This makes us to think that these two factors are related and the deviation between different clones first of all could be caused by the recombination events, eliminating part of the integrated gene clusters. Another reason for observed deviation might be the difference in efficiency of integration into *attB* and pseudo *attB* sites resulting in different number of gene clusters being integrated into chromosome of individual clones. It is a well-established fact, that integration of φC31-based vectors in *S. lividans* TK24 into different *attB* sites occurs with different efficiency [59]. Independently from the reasons, the *S. lividans* ΔYA10 and especially ΔYA11 transconjugants population obviously have a high degree of heterogeneity resulting in dramatic variation in the production level from clone to clone and in long-term cultivation. Thus the care should be taken when selecting the proper strain and individual clone(s) for scaled up production of desired compound in case of ΔYA10 and ΔYA11 strains.

The engineered *S. lividans* strains were found to perform better with the gene clusters encoding biosynthesis of amino acid derived compounds, like deoxycinnamycin and identified PBD family natural product. The later metabolite was not produced by *S. albus* Del14 when the

respective construct was introduced. It is also well-known fact that *S. lividans* is a long-term preferred host for ribosomally synthesized and post-translationally modified peptides (RiPPs) gene cluster expression [60, 61]. At the same time the polyketide griseorhodin was not produced by *S. lividans* TK24 at all; and *S. albus* Del14 was superior to all other strains tested in this work. Similar, pamamycin gene cluster was successfully expressed solely in *S. albus*, but not any other strain tested [62]. Deletion of endogenous secondary metabolism gene clusters in *S. lividans* resulted in activation of griseorhodin biosynthesis in *S. lividans* ΔYA9, ΔYA10 and ΔYA11 strains with the obvious copy-number effect observed. The reasons for this are not very clear, perhaps the sink of precursors into internal polyketide pathways for actinorhodin, undecylprodigiosine, coelimycin, herboxidiene resulted in high competition and thus lack of griseorhodin production. Also, the interplay between regulatory systems controlling expression of secondary metabolism in *S. lividans* cannot be excluded. In the case of *S. albus* the deletion of endogenous secondary metabolism gene clusters resulted in elevated transcription level of heterologous biosynthetic genes, clearly indicating a possibility of such interactions [15]. Furthermore, recently we have demonstrated that *S. albus* butenolides regulatory system is influencing production of heterologous secondary metabolites [63]. It is difficult to say if this situation is true also for *S. lividans*, but at least in the case of griseorhodin the phenotype is very prominent and pointing on possibility of such regulatory interplay. Similar phenomenon was observed in the case of deoxycinnamycin and PBD-like compound production, when the accumulation of both compounds was significantly elevated in the ΔYA9 strain when compared with TK24. In fact, most probably we would not be able to detect the PBD compounds if the library was screened in the parental strain.

The clean metabolic background of engineered strains provides another benefit in their utilization as hosts for screening genomic and metagenomic libraries for new secondary metabolites. The absence of endogenously produced compounds significantly improves detection and identification of products. On the other hand, it is quite obvious, that success of such projects strongly depends on efficiency of utilized chassis strain(s). As such, the griseorhodin was produced in *S. albus* Del14 but not in *S. lividans* TK24. Vice versa, expression of gene cluster for potent ribosome inhibitor bottromycin was successful in *S. lividans* but not *S. albus* [64]. In fact there are many examples when the gene cluster was expressed in one but not another host strain. Unfortunately, there is no universal chassis strain for such experiments. Thus, the success rate of genomic or metagenomics library screening for new secondary metabolites will strongly benefit from the use of more than one expression host. The engineered *S. lividans* complement the panel of metabolically clean heterologous gene clusters expression chassis strains.

4.5. Conclusion

As conclusion, developed strains represent a step forward towards a better panel of organisms for bioprospecting and genome mining of novel natural products. In fact, it is becoming obvious that different chassis stain do have a preference for certain type of natural products. Such as, *S. albus* is successful host for expression of polyketide secondary metabolites, when *S. lividans* performs better with the gene clusters coding for production of amino-acid based natural products, like RiPPs and nonribosomal peptides. However, despite the obvious progress in this direction, the current panel of expression hosts is obviously far from being able to satisfy the needs of screening programs and still would have to be expanded first of all with non-*Streptomyces* Actinobacteria strains in order to achieve the higher success rate in new natural products identification and isolation.

4.6. Methods

Strains, plasmids and growth conditions

All bacterial strains, plasmids and BACs used in this study are listed in Additional file 8: Table 3S. *E. coli* strains were grown in LB (lysogeny broth) medium [65]. *Streptomyces* strains were grown on mannitol soy flour agar (MS agar) [66] for spores formation and in liquid tryptic soy broth (TSB) medium for pre-culture (Sigma-Aldrich, USA). Modified SG (glucose 20 g, yeast extract 5 g, soytone 10 g, calcium carbonate 2 g, dH_2O 1000 ml) and DNPM [67] (dextrin 40 g, soytone 7.5 g, fresh yeasts 5 g, MOPS 21 g, dH_2O 1000 ml, pH 6.8) were used for secondary metabolites production. The following antibiotics were used when necessary: apramycin (50µg/mL), kanamycin (30µg/mL), hygromycin (50µg/mL), thiostrepton (50µg/mL) phosphomycin (100 µg/mL) and chloramphenicol (50 µg/mL) (Carl Roth, Germany, and Sigma-Aldrich, USA).

Isolation and manipulation of DNA

DNA isolation and manipulation, *E. coli* transformation and *E. coli*/*Streptomyces* intergeneric conjugation were performed according to standard protocols [65, 66, 68]. Dream Taq polymerase (Thermo Fisher Scientific, USA) was used to amplify DNA fragments for cloning, for PCR verification of constructs and strains. DNA fragments were purified from agarose gels using the QIAquick Gel Extraction Kit (Qiagen, Germany). Restriction enzymes and ligase were used accordingly to manufacturer recommendations (New England Biolabs, USA). All primers used in this study are listed in Additional file 9: Table. 4S (Eurofins Genomics, Germany).

Construction of *S. lividans* TK24 mutants

To delete the desired gene clusters BAC clones (Table 3S) from ordered *S. lividans* genomic library were selected [30] and modified using Red/ET recombination approach combined with IMES antibiotic resistance cassettes [31, 69]. The IMES cassette containing apramycin resistance marker and *oriT* was excised from a carrier plasmid (patt-saac-*oriT*) with *Pvu*II and amplified using primers listed in (Table 4S). Red/ET was performed as previously described [70]. Deletions were confirmed by PCR using check primers listed in Additional file 9: Table 4S. The recombinant BACs were introduced stepwise into the *S. lividans* strains. The double-crossover mutants were screened on MS medium supplemented with apramycin and 50 μg/mL of X-Gluc (X-CLUC Direct, USA). After each deletion the resistance marker was removed from the chromosome of the generated strains by expression of φC31 actinophage integrase [69], and the resulting strains genotype was confirmed by PCR and PCR product sequencing.

Introduction of additional *attB*-sites into the chromosome of *S. lividans* ΔYA9

To introduce *attB* sites into *S. lividans* strains the synthetic cassette consisting of hygromycin resistance marker flanked by *Mss*I restriction sites and *attB* sequence was used [15]. This cassette was PCR amplified using primers for Red/ET (Table 4S). Large fragments of clusters 21 and 5 were replaced with the above mentioned cassettes in BACs 1468 and 1092 respectively using Red/ET recombination. Hygromycin resistance gene was removed by digesting with *Mss*I restriction enzyme followed by self-ligation. The resulting BACs were named 1468::attB and 1092::attB, respectively. Recombinant BACs were further modified by substituting the chloramphenicol gene (*cat*) with cassette containing apramycin resistance gene (*aac(3)IV*) for selection in *Streptomyces* and *oriT* sequence to conjugation transfer of construct from *E. coli* to *Streptomyces*. The final BACs (1468::attBamoriT and 1092::attBamoriT, respectively) were introduced sequentially into *S. lividans* ΔYA9 strain by conjugation. The secondary cross-over colonies were selected by lack of β-glucuronidase activity and apramycin sensitivity phenotype caused by loss of the vector backbone. The deletion of the gene clusters and introducing additional *attB* sites were confirmed by PCR using primers listed in Additional file 9: Table 4S. Additionally, the PCR products were sequenced to confirm presence and correct sequence of *attB* sites.

Growth and sporulation study

For biomass measurement the wild type *S. lividans* TK24 and the obtained strains, *S. lividans* ΔYA9 and *S. lividans* ΔYA11 were grown in 30 mL of TSB for 2 days. 100 mg of biomass of

each pre-culture was used to inoculate 70 mL of DNPM production medium. 1 ml of each sample was taken every 24 hours for 6 days. The biomass was harvested by centrifugation for 10 min at 14000 rpm. The supernatant was discarded, and the weight of wet biomass was measured.

Heterologous expression of secondary metabolite gene clusters

Constructs carrying gene clusters for biosynthesis of the nucleoside antibiotic tunicamycin, the aromatic polyketide griseorhodin, the lantipeptide deoxycinnamycin (Table 3S) were introduced into *S. lividans* TK24, *S. lividans* ΔYA9, *S. lividans* ΔYA10, *S. lividans* ΔYA11, *S. albus* Del14 and *S. coelicolor* M1154 by mean of intergeneric conjugation [66].

The genomic library of *S. albus* subsp. *chlorinus* NRRL B-24108 was constructed on pSMART vector by Intact Genomics, USA. First, 17 clones were introduced into *S. lividans* ΔYA9 and *S. albus* Del14. Resulting strains were cultivated and the metabolites were extracted and analysed as described below. Out of 17 clones the BAC2I4 was chosen for further study and was introduced into above-mentioned tested hosts by conjugation.

Cultivation conditions and extraction of secondary metabolites

Three independent clones of each recombinant strain were grown in 30 mL of TSB for 2 days (except *S. albus* Del14 were grown for 24 hours) and 100 mg of biomass of the pre-culture was used to inoculate 50 mL of production media. Strains were cultivated for 6 days at 28 °C and metabolites were extracted as following: 1) the deoxycinnamycin was extracted with butanol; 2) tunicamycins were extracted with butanol and acetic acid; 3) the PBD-like compounds were extracted with ethyl acetate; 4) griseorhodin was extracted with ethyl acetate and acetic acid. Extracts were evaporated and dissolved in 300 µl of DMSO/MeOH (1:1). 3µl were analyzed by LC-MS as described below. Metabolites accumulation was averaged and the production by *S. lividans* TK24 was taken as 100% in each particular experiment (except for griseorhodin the *S. lividans* ΔYA9 was taken as 100%).

Analysis of secondary metabolites production

3µL of each sample was separated using a Dionex Ultimate 3000 UPLC (Thermo Fisher Scientific, USA) and a 10-cm ACQUITY UPLC® BEH C18 column, 1.7 µm (Waters, USA) and a linear gradient of acetonitrile against 0.1% formic acid solution in water from 5% to 95% in 10 or 18 minutes at a flow rate of 0.6 mL/min. Samples were analyzed using an amaZon speed mass spectrometer (Bruker Daltonics, Germany) using ESI source. Mass spectra were acquired in centroid mode ranging from 200 to 2000 *m/z* at a 2 Hz scan rate. Data was collected and

analyzed with the Bruker Compass Data Analysis software, version 4.2 (Bruker, Billerica, USA). For determination of accurate mass Thermo LTQ Orbitrap XL coupled to UPLC Thermo Dionex Ultimate 3000 RS was used. Data were acquired with Xcalibur 2.2 software (Thermo Scientific). The separation conditions were identical to those used for quantification study. For fragmentation pattern the mass spectra were ranging from 50 to 2000 *m/z*.

Purification of compounds 4a-d

S. lividans ΔYA11 was grown at 30 °C for 3 days in 6 × 500-mL flasks containing 50 mL of TSB, and pre-culture was used to inoculate 100 × 500-mL flasks containing 50 mL of SG media. Cultures were incubated at 30 °C for 6 days. Metabolites were extracted as described above. The extracts from biomass and the supernatant were combined and fractionated by size-exclusion chromatography on an LH 20 Sephadex column (Sigma-Aldrich, USA) using methanol as the solvent. The fractions were collected every 15 minutes, evaporated and dissolved in 0.5 mL of MeOH. The obtained fractions with mixture of compounds 4 were further purified by Agilent 1260 Series from Agilent Technologies (semipreparative). Semipreparative HPLC was performed using a Jubiter LC-Column (250×10 mm, 4μm; Phenomenex) with a multistep gradient from 5 to 80% B (B: acetonitrile with 0.1% formic acid; A: water with 0.1% formic acid) over 5 min and then increased to 95% B over 12 min and ended with a 7 min step of 95% B before re-equilibration under the initial conditions at a flow rate of 4 ml/min and 45◦C.

Nuclear magnetic resonance (NMR) spectroscopy and structure elucidation of 4a-d

^{1}H-NMR and 2D HHCOSY, HSQC and HMBC were recorded on a Bruker Avance III 700 spectrometer (Bruker, BioSpin GmbH, Rheinstetten, Germany) at 298 K equipped with a 5 mm TCI cryo probe using deuterated methanol (Deutero, Kastellaun, Germany) as solvent. The chemical shifts were reported in parts per million (ppm) relative to the solvent peaks (δ_H 3.30 and δ_C 49.00 respectively). All spectra were performed using standard pulse programs from the Bruker pulse program library.

The mass spectra taken from the BPC chromatogram of *S. lividans* TK24 extract (Fig.1S) showed a set of at least four members of a homologous series with $[M+H]^{+)}$ = 597.58, 611.58, 625,62, 639,64 amu (**4a-d**). Purification of the extract led to enriched fractions of **4b** and **4c**, which could be subjected to NMR studies. Both ^{1}H NMR (Fig 6S, (6+1)S, (6+2)S and (6+3)S) were very close to each other and differed only in the integration of a few resonances. Intensive research and careful comparison of the data with those from literature [71] revealed a 1-monoglyceride structure (δ_H 4.14 dd, J=11.5 and 4.5 Hz, 4.05 dd, J=11.5 and 6.5 Hz: ROC**H**$_2$-

CHOH-CH_2OH; δ_H 3.81 m: $ROCH_2$-CHO**H**-CH_2OH; δ_H 3.55 dd and 3.53, both J=11.5 and 5.5 Hz: $ROCH_2$-CHOH-C**H_2**OH). In contrast to ordinary fats the fatty acids in **4b** and **4c** showed resonances for a α-hydroxy group. Their chemical shift (δ_H 5.21) proved an esterification at this position. The acid substructure of this ester remained unidentified, mainly due to signal overlapping. The small amount of substance did not allow a needful chemical degradation, purification and further NMR investigations

4.7. References

1. Demain AL, Sanchez S: Microbial drug discovery: 80 years of progress. *J Antibiot (Tokyo)* 2009, 62(1):5-16.
2. Berdy J: Bioactive microbial metabolites. *J Antibiot (Tokyo)* 2005, 58(1):1-26.
3. Yamanaka K, Reynolds KA, Kersten RD, Ryan KS, Gonzalez DJ, Nizet V, Dorrestein PC, Moore BS: Direct cloning and refactoring of a silent lipopeptide biosynthetic gene cluster yields the antibiotic taromycin A. *Proc Natl Acad Sci U S A* 2014, 111(5):1957-1962.
4. Pait IGU, Kitani S, Roslan FW, Ulanova D, Arai M, Ikeda H, Nihira T: Discovery of a new diol-containing polyketide by heterologous expression of a silent biosynthetic gene cluster from *Streptomyces lavendulae* FRI-5. *J Ind Microbiol Biotechnol* 2018, 45(2):77-87.
5. Gomez-Escribano JP, Bibb MJ: Heterologous expression of natural product biosynthetic gene clusters in *Streptomyces coelicolor*: from genome mining to manipulation of biosynthetic pathways. *J Ind Microbiol Biotechnol* 2014, 41(2):425-431.
6. Young TS, Dorrestein PC, Walsh CT: Codon randomization for rapid exploration of chemical space in thiopeptide antibiotic variants. *Chem Biol* 2012, 19(12):1600-1610.
7. Du D, Zhu Y, Wei JH, Tian YQ, Niu G, Tan HR: Improvement of gougerotin and nikkomycin production by engineering their biosynthetic gene clusters. *Appl Microbiol Biot* 2013, 97(14):6383-6396.
8. Penn J, Li X, Whiting A, Latif M, Gibson T, Silva CJ, Brian P, Davies J, Miao V, Wrigley SK *et al*: Heterologous production of daptomycin in *Streptomyces lividans*. *J Ind Microbiol Biotechnol* 2006, 33(2):121-128.
9. Zhao ZL, Shi T, Xu M, Brock NL, Zhao YL, Wang YM, Deng ZX, Pang XH, Tao MF: Hybrubins: Bipyrrole Tetramic Acids Obtained by Crosstalk between a Truncated Undecylprodigiosin Pathway and Heterologous Tetramic Acid Biosynthetic Genes. *Org Lett* 2016, 18(3):572-575.
10. McDaniel R, Ebertkhosla S, Hopwood DA, Khosla C: Engineered Biosynthesis of Novel Polyketides. *Science* 1993, 262(5139):1546-1550.
11. Komatsu M, Uchiyama T, Omura S, Cane DE, Ikeda H: Genome-minimized *Streptomyces* host for the heterologous expression of secondary metabolism. *Proc Natl Acad Sci U S A* 2010, 107(6):2646-2651.
12. Gomez-Escribano JP, Bibb MJ: Engineering *Streptomyces coelicolor* for heterologous expression of secondary metabolite gene clusters. *Microb Biotechnol* 2011, 4(2):207-215.
13. Shima J, Hesketh A, Okamoto S, Kawamoto S, Ochi K: Induction of actinorhodin production by *rpsL* (encoding ribosomal protein S12) mutations that confer streptomycin resistance in *Streptomyces lividans* and *Streptomyces coelicolor* A3(2). *J Bacteriol* 1996, 178(24):7276-7284.
14. Hu H, Zhang Q, Ochi K: Activation of antibiotic biosynthesis by specified mutations in the *rpoB* gene (encoding the RNA polymerase beta subunit) of *Streptomyces lividans*. *J Bacteriol* 2002, 184(14):3984-3991.

15. Myronovskyi M, Rosenkranzer B, Nadmid S, Pujic P, Normand P, Luzhetskyy A: Generation of a cluster-free *Streptomyces albus* chassis strains for improved heterologous expression of secondary metabolite clusters. *Metab Eng* 2018.
16. Barkei JJ, Kevany BM, Felnagle EA, Thomas MG: Investigations into Viomycin Biosynthesis by Using Heterologous Production in *Streptomyces lividans*. *Chembiochem* 2009, 10(2):366-376.
17. Shao L, Li J, Liu AJ, Chang Q, Lin HM, Chen DJ: Efficient Bioconversion of Echinocandin B to Its Nucleus by Overexpression of Deacylase Genes in Different Host Strains. *Appl Environ Microb* 2013, 79(4):1126-1133.
18. Anne J, Vrancken K, Van Mellaert L, Van Impe J, Bernaerts K: Protein secretion biotechnology in Gram-positive bacteria with special emphasis on *Streptomyces lividans*. *Bba-Mol Cell Res* 2014, 1843(8):1750-1761.
19. Eustaquio AS, Gust B, Galm U, Li SM, Chater KF, Heide L: Heterologous expression of novobiocin and clorobiocin biosynthetic gene clusters. *Appl Environ Microb* 2005, 71(5):2452-2459.
20. Voller GH, Krawczyk JM, Pesic A, Krawczyk B, Nachtigall J, Sussmuth RD: Characterization of new class III lantibiotics--erythreapeptin, avermipeptin and griseopeptin from *Saccharopolyspora erythraea*, *Streptomyces avermitilis* and *Streptomyces griseus* demonstrates stepwise N-terminal leader processing. *Chembiochem* 2012, 13(8):1174-1183.
21. Ziermann R, Betlach MC: Recombinant polyketide synthesis in *Streptomyces*: Engineering of improved host strains. *Biotechniques* 1999, 26(1):106-+.
22. Martinez A, Kolvek SJ, Yip CL, Hopke J, Brown KA, MacNeil IA, Osburne MS: Genetically modified bacterial strains and novel bacterial artificial chromosome shuttle vectors for constructing environmental libraries and detecting heterologous natural products in multiple expression hosts. *Appl Environ Microbiol* 2004, 70(4):2452-2463.
23. Novakova R, Nunez LE, Homerova D, Knirschova R, Feckova L, Rezuchova B, Sevcikova B, Menendez N, Moris F, Cortes J *et al*: Increased heterologous production of the antitumoral polyketide mithramycin A by engineered *Streptomyces lividans* TK24 strains. *Appl Microbiol Biotechnol* 2018, 102(2):857-869.
24. Ruckert C, Albersmeier A, Busche T, Jaenicke S, Winkler A, Friethjonsson OH, Hreggviethsson GO, Lambert C, Badcock D, Bernaerts K *et al*: Complete genome sequence of *Streptomyces lividans* TK24. *J Biotechnol* 2015, 199:21-22.
25. Busche T, Tsolis KC, Koepff J, Rebets Y, Ruckert C, Hamed MB, Bleidt A, Wiechert W, Lopatniuk M, Yousra A *et al*: Multi-Omics and Targeted Approaches to Determine the Role of Cellular Proteases in *Streptomyces* Protein Secretion. *Front Microbiol* 2018, 9:1174.
26. Lambert S, Traxler MF, Craig M, Maciejewska M, Ongena M, Van Wezel GP, Kolter R, Rigali S: Altered desferrioxamine-mediated iron utilization is a common trait of bald mutants of *Streptomyces coelicolor*. *Metallomics* 2014, 6(8):1390-1399.
27. Challis GL, Hopwood DA: Synergy and contingency as driving forces for the evolution of multiple secondary metabolite production by *Streptomyces* species. *P Natl Acad Sci USA* 2003, 100:14555-14561.
28. Barona-Gomez F, Wong U, Giannakopulos AE, Derrick PJ, Challis GL: Identification of a cluster of genes that directs desferrioxamine biosynthesis in *Streptomyces coelicolor* M145. *J Am Chem Soc* 2004, 126(50):16282-16283.
29. Challis GL, Ravel J: Coelichelin, a new peptide siderophore encoded by the *Streptomyces coelicolor* genome: structure prediction from the sequence of its non-ribosomal peptide synthetase. *FEMS Microbiol Lett* 2000, 187(2):111-114.
30. Rebets Y, Tsolis KC, Guðmundsdóttir EE, Koepff J, Wawiernia B, Busche T, Bleidt A, Horbal L, Myronovskyi M, Ahmed Y *et al*: Characterization of Sigma Factor Genes in *Streptomyces*

lividans TK24 Using a Genomic Library-Based Approach for Multiple Gene Deletions. *Frontiers in Microbiology* 2018, 9(3033).

31. Zhang YM, Buchholz F, Muyrers JPP, Stewart AF: A new logic for DNA engineering using recombination in *Escherichia coli*. *Nat Genet* 1998, 20(2):123-128.
32. Myronovskyi M, Rosenkranzer B, Luzhetskyy A: Iterative marker excision system. *Appl Microbiol Biotechnol* 2014, 98(10):4557-4570.
33. Lomovskaya ND, Mkrtumian NM, Gostimskaya NL, Danilenko VN: Characterization of temperate actinophage phi C31 isolated from *Streptomyces coelicolor* A3(2). *J Virol* 1972, 9(2):258-262.
34. Lomovskaya ND, Chater KF, Mkrtumian NM: Genetics and molecular biology of *Streptomyces* bacteriophages. *Microbiol Rev* 1980, 44(2):206-229.
35. Anne J, Wohlleben W, Burkardt HJ, Springer R, Puhler A: Morphological and Molecular Characterization of Several Actinophages Isolated from Soil Which Lyse *Streptomyces cattleya* or *Streptomyces venezuelae*. *J Gen Microbiol* 1984, 130(Oct):2639-2649.
36. Groth AC, Olivares EC, Thyagarajan B, Calos MP: A phage integrase directs efficient site-specific integration in human cells. *Proc Natl Acad Sci U S A* 2000, 97(11):5995-6000.
37. Barona-Gomez F, Lautru S, Francou FX, Leblond P, Pernodet JL, Challis GL: Multiple biosynthetic and uptake systems mediate siderophore-dependent iron acquisition in *Streptomyces coelicolor* A3(2) and *Streptomyces ambofaciens* ATCC 23877. *Microbiology* 2006, 152(Pt 11):3355-3366.
38. Wyszynski FJ, Lee SS, Yabe T, Wang H, Gomez-Escribano JP, Bibb MJ, Lee SJ, Davies GJ, Davis BG: Biosynthesis of the tunicamycin antibiotics proceeds via unique exo-glycal intermediates. *Nat Chem* 2012, 4(7):539-546.
39. Li AY, Piel J: A gene cluster from a marine *Streptomyces* encoding the biosynthesis of the aromatic spiroketal polyketide griseorhodin A. *Chemistry & Biology* 2002, 9(9):1017-1026.
40. Lopatniuk M, Myronovskyi M, Luzhetskyy A: *Streptomyces albus*: A New Cell Factory for Non-Canonical Amino Acids Incorporation into Ribosomally Synthesized Natural Products. *Acs Chem Biol* 2017, 12(9):2362-2370.
41. Hahn DR, Graupner PR, Chapin E, Gray J, Heim D, Gilbert JR, Gerwick BC: Albucidin: a novel bleaching herbicide from *Streptomyces albus* subsp. *chlorinus* NRRL B-24108. *J Antibiot (Tokyo)* 2009, 62(4):191-194.
42. Rodriguez Estevez M, Myronovskyi M, Gummerlich N, Nadmid S, Luzhetskyy A: Heterologous Expression of the Nybomycin Gene Cluster from the Marine Strain *Streptomyces albus* subsp. *chlorinus* NRRL B-24108. *Mar Drugs* 2018, 16(11).
43. Hu Y, Phelan V, Ntai I, Farnet CM, Zazopoulos E, Bachmann BO: Benzodiazepine biosynthesis in *Streptomyces* refuineus. *Chem Biol* 2007, 14(6):691-701.
44. Najmanova L, Ulanova D, Jelinkova M, Kamenik Z, Kettnerova E, Koberska M, Gazak R, Radojevic B, Janata J: Sequence analysis of porothramycin biosynthetic gene cluster. *Folia microbiologica* 2014, 59(6):543-552.
45. Bu QT, Yu P, Wang J, Li ZY, Chen XA, Mao XM, Li YQ: Rational construction of genome-reduced and high-efficient industrial *Streptomyces* chassis based on multiple comparative genomic approaches. *Microb Cell Fact* 2019, 18(1):16.
46. Paulus C, Rebets Y, Tokovenko B, Nadmid S, Terekhova LP, Myronovskyi M, Zotchev SB, Ruckert C, Braig S, Zahler S *et al*: New natural products identified by combined genomics-metabolomics profiling of marine *Streptomyces* sp. MP131-18. *Sci Rep* 2017, 7:42382.
47. Shima J, Hesketh A, Okamoto S, Kawamoto S, Ochi K: Induction of actinorhodin production by *rpsL* (encoding ribosomal protein S12) mutations that confer streptomycin resistance in *Streptomyces lividans* and *Streptomyces coelicolor* A3(2). *J Bacteriol* 1996, 178(24):7276-7284.

48. Yu TW, Hopwood DA: Ectopic Expression of the *Streptomyces coelicolor Whie* Genes for Polyketide Spore Pigment Synthesis and Their Interaction with the Act Genes for Actinorhodin Biosynthesis. *Microbiol-Uk* 1995, 141:2779-2791.
49. Lippert K, Galinski EA: Enzyme Stabilization by Ectoine-Type Compatible Solutes - Protection against Heating, Freezing and Drying. *Appl Microbiol Biot* 1992, 37(1):61-65.
50. Bentley SD, Chater KF, Cerdeno-Tarraga AM, Challis GL, Thomson NR, James KD, Harris DE, Quail MA, Kieser H, Harper D *et al*: Complete genome sequence of the model actinomycete *Streptomyces coelicolor* A3(2). *Nature* 2002, 417(6885):141-147.
51. Poralla K, Muth G, Hartner T: Hopanoids are formed during transition from substrate to aerial hyphae in *Streptomyces coelicolor* A3(2). *FEMS Microbiol Lett* 2000, 189(1):93-95.
52. Seipke R, Loria R: Hopanoids are not essential for growth of *Streptomyces scabies* 87-22. *J Bacteriol* 2009, 191(16):5216-5223.
53. Blin K, Wolf T, Chevrette MG, Lu X, Schwalen CJ, Kautsar SA, Suarez Duran HG, de Los Santos ELC, Kim HU, Nave M *et al*: antiSMASH 4.0-improvements in chemistry prediction and gene cluster boundary identification. *Nucleic Acids Res* 2017.
54. Fedoryshyn M, Welle E, Bechthold A, Luzhetskyy A: Functional expression of the Cre recombinase in actinomycetes. *Appl Microbiol Biot* 2008, 78(6):1065-1070.
55. Herrmann S, Siegl T, Luzhetska M, Petzke L, Jilg C, Welle E, Erb A, Leadlay PF, Bechthold A, Luzhetskyy A: Site-Specific Recombination Strategies for Engineering Actinomycete Genomes. *Appl Environ Microb* 2012, 78(6):1804-1812.
56. Raynal A, Karray F, Tuphile K, Darbon-Rongere E, Pernodet JL: Excisable cassettes: New tools for functional analysis of *Streptomyces* genomes. *Appl Environ Microb* 2006, 72(7):4839-4844.
57. Fedoryshyn M, Petzke L, Welle E, Bechthold A, Luzhetskyy A: Marker removal from actinomycetes genome using Flp recombinase. *Gene* 2008, 419(1-2):43-47.
58. Bilyk B, Luzhetskyy A: Unusual site-specific DNA integration into the highly active pseudo-*attB* of the *Streptomyces albus* J1074 genome. *Appl Microbiol Biotechnol* 2014, 98(11):5095-5104.
59. Combes P, Till R, Bee S, Smith MC: The *streptomyces* genome contains multiple pseudo-attB sites for the (phi)C31-encoded site-specific recombination system. *J Bacteriol* 2002, 184(20):5746-5752.
60. Krawczyk JM, Voller GH, Krawczyk B, Kretz J, Bronstrup M, Sussmuth RD: Heterologous expression and engineering studies of labyrinthopeptins, class III lantibiotics from *Actinomadura namibiensis*. *Chem Biol* 2013, 20(1):111-122.
61. Zhang Y, Chen MY, Bruner SD, Ding YS: Heterologous Production of Microbial Ribosomally Synthesized and Post-translationally Modified Peptides. *Frontiers in Microbiology* 2018, 9.
62. Rebets Y, Brotz E, Manderscheid N, Tokovenko B, Myronovskyi M, Metz P, Petzke L, Luzhetskyy A: Insights into the pamamycin biosynthesis. *Angew Chem Int Ed Engl* 2015, 54(7):2280-2284.
63. Ahmed Y, Rebets Y, Tokovenko B, Brotz E, Luzhetskyy A: Identification of butenolide regulatory system controlling secondary metabolism in *Streptomyces albus* J1074. *Sci Rep* 2017, 7(1):9784.
64. Horbal L, Marques F, Nadmid S, Mendes MV, Luzhetskyy A: Secondary metabolites overproduction through transcriptional gene cluster refactoring. *Metabolic Engineering* 2018, 49:299-315.
65. Green MR, Sambrook J: Molecular cloning : a laboratory manual, 4th edn. Cold Spring Harbor, N.Y.: Cold Spring Harbor Laboratory Press; 2012.
66. Kieser T. BMJ, Buttner M.J., Charter K.F., Hopwood D.: Practical *Streptomyces* Genetics. . Norwich, United Kingdom: John Innes Foundation; 2000.

67. Myronovskyi M, Brotz E, Rosenkranzer B, Manderscheid N, Tokovenko B, Rebets Y, Luzhetskyy A: Generation of new compounds through unbalanced transcription of landomycin A cluster. *Appl Microbiol Biotechnol* 2016, 100(21):9175-9186.
68. Rebets Y, Kormanec J, Lutzhetskyy A, Bernaerts K, Anne J: Cloning and Expression of Metagenomic DNA in *Streptomyces lividans* and Subsequent Fermentation for Optimized Production. *Methods Mol Biol* 2017, 1539:99-144.
69. Myronovskyi M, Rosenkranzer B, Luzhetskyy A: Iterative marker excision system. *Appl Microbiol Biotechnol* 2014, 98(10):4557-4570.
70. Fu J, Bian X, Hu S, Wang H, Huang F, Seibert PM, Plaza A, Xia L, Muller R, Stewart AF *et al*: Full-length RecE enhances linear-linear homologous recombination and facilitates direct cloning for bioprospecting. *Nat Biotechnol* 2012, 30(5):440-446.
71. Nieva-Echevarria B, Goicoechea E, Manzanos MJ, Guillen MD: A method based on H-1 NMR spectral data useful to evaluate the hydrolysis level in complex lipid mixtures. *Food Res Int* 2014, 66:379-387.

4.7. Supplementary Information

Table 1S. Secondary metabolites gene clusters in *S. lividans* TK24 and their transcriptional level Average in in minimal medium in mini-bioreactor [1].

Cluster	Name	Coordinates, bp	ORFs	RPKM
Cluster 1	terpene	16255 - 39174	SLIV_00085 - SLIV_00185	17.5
Cluster 2	coelibactin	145838 - 218926	SLIV_00745 - SLIV_01060	198.4
Cluster 3	terpene	1013598 - 1039716	SLIV_04650 - SLIV_04765	205.6
Cluster 4	lantipeptide	1116997 - 1135734	SLIV_05075 - SLIV_05165	157.8
cluster 5	NRPS	1380581 - 1429153	SLIV_06175 - SLIV_06340	62.1
Cluster 6	coelimycin	1504477 - 1583755	SLIV_06690 - SLIV_06900	33.4
Cluster 7	siderophore	1601312 - 1613393	SLIV_06985 - SLIV_07035	36.1
Cluster 8	terpene	1777266 - 1799442	SLIV_07735 - SLIV_07820	109.3
Cluster 9	bacteriocin	1811603 - 1821470	SLIV_07885 - SLIV_07930	294.4
Cluster 10	undecylprodigiosin	2068946 - 2116414	SLIV_09055 - SLIV_09240	91.1
Cluster 11	siderophore	2198827 - 2209792	SLIV_09600 - SLIV_09630	918.1
Cluster 12	type II PKS	2688586 - 2730052	SLIV_11695 - SLIV_11900	325.5
Cluster 13	terpene	2804344 - 2825931	SLIV_12210 - SLIV_12310	803.1
Cluster 14	actinorhodin	2944091 - 2985628	SLIV_12835 - SLIV_13045	351.4
Cluster 15	CDA	4737997 - 4818586	SLIV_21425 - SLIV_21620	53.9
Cluster 16	siderophore	5289644 - 5301384	SLIV_23725 - SLIV_23770	844.3
Cluster 17	melanin	5385376 - 5394552	SLIV_24140 - SLIV_24195	94.1
Cluster 18	ectoine	6336179 - 6347102	SLIV_28360 - SLIV_28410	360.3
Cluster 19	germicidin	7040159 - 7082308	SLIV_31665 - SLIV_31865	226.3
Cluster 20	bacteriocin	7572312 - 7580633	SLIV_34155 - SLIV_34180	35.9
Cluster 21	coelichelin	7827171 - 7878351	SLIV_35400 - SLIV_35580	707.8
Cluster 22	lantipeptide	8101385 - 8137004	SLIV_36525 - SLIV_36675	323.1
Cluster 23	terpene	8180179 - 8205399	SLIV_36890 - SLIV_37000	398
Cluster 24	eicosapentaenoic acid	8242078 - 8295559	SLIV_37190 - SLIV_37380	38.1
Cluster 25	terpene	8308420 - 8328145	SLIV_37430 - SLIV_37535	6.9
RNA polymerase β- subunit		3408742 - 3404843	SLIV_25060	958.5
S. lividans TK24 genome average RPKM				426.4

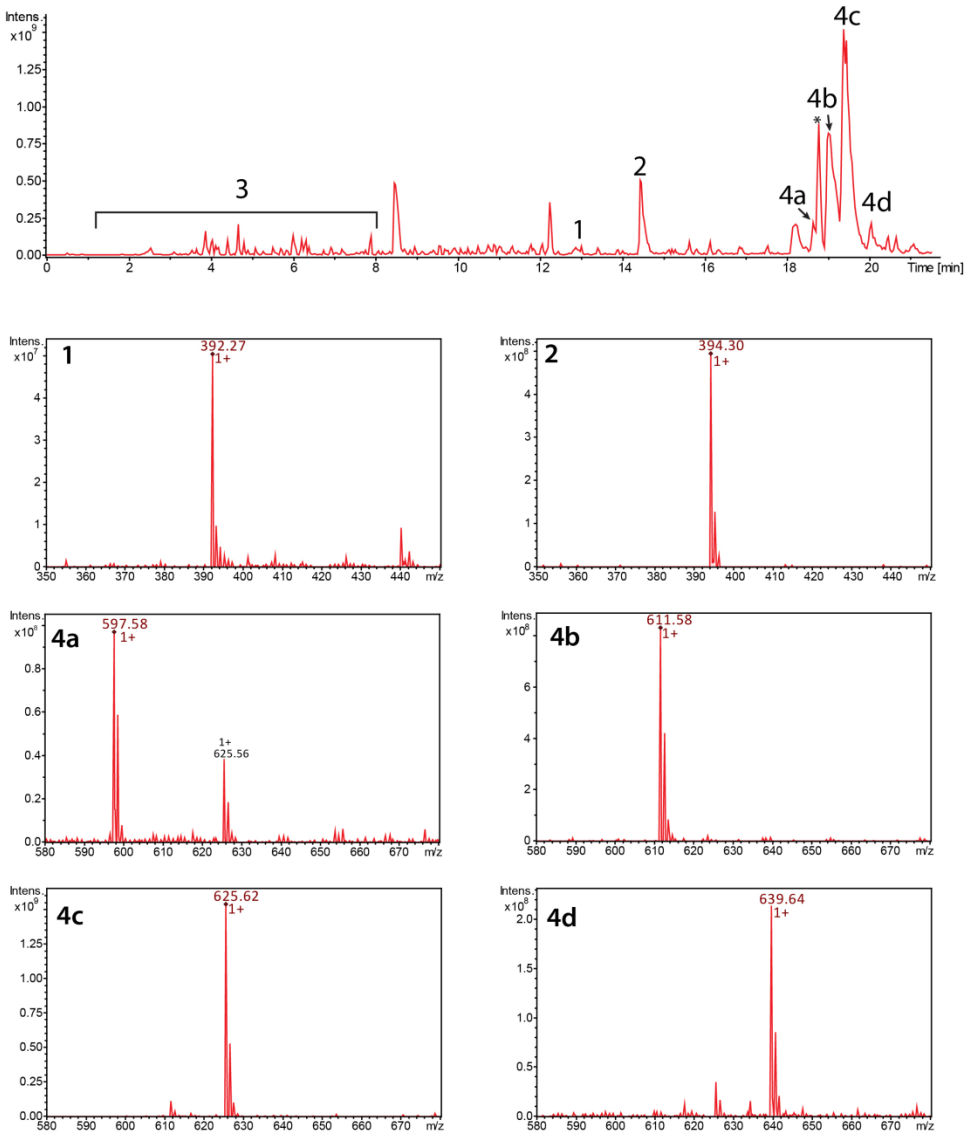
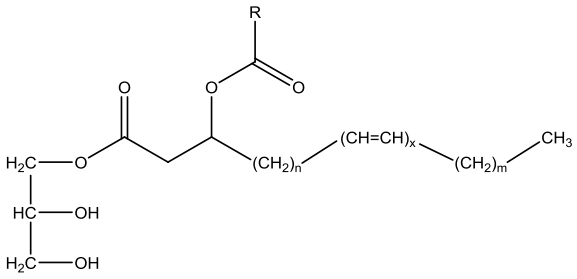


Figure 1S. BPC chromatogram extract of *S. lividans* TK24. The strain was grown in SG medium for 6 days at 28°C. The metabolites were extracted with ethyl acetate and measured on LC-MS amaZon speed system (Bruker Daltonics, Germany) with 18 min gradient protocol. Mass spectra is shown for compounds: Coelibactin (**1**), *m/z* 392.27 $[M+H]^+$ (calculated 392.08 $[M+H]^+$); Undecylprodigiosin (**2**), *m/z* 394.30 $[M+H]^+$ (calculated 394.28 $[M+H]$, the homologous series of 1-MG derivatives (**4a-d**), *m/z* 597.58, 611.58, 625,62, 639,64 $[M+H]^+$, respectively.



X: 0,1

Figure 2S. Proposed structures for 4a-d.

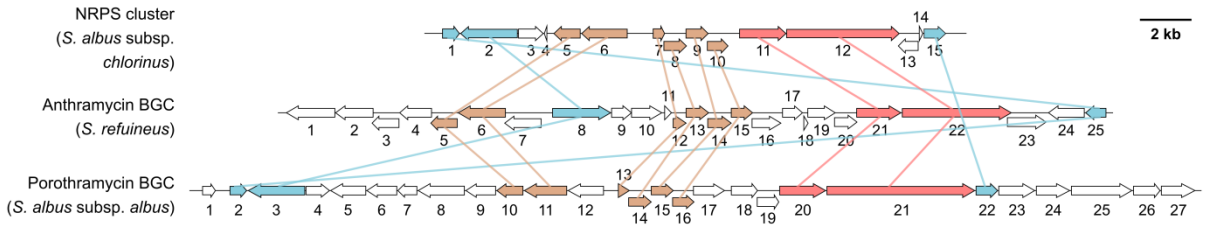


Figure 3S. Comparison of *S. albus* subsp. *chlorinus* NRPS gene cluster with described PBD biosynthetic gene clusters: anthramycin BGC from *S. refuineus* and porothramycin BGC from *S. albus* subsp. *albus*. Numbers refer to the respective genes from each cluster: in *S. albus* subsp. *chlorinus* NRPS cluster, *chl1*; in anthramycin gene cluster, *orf1*; in porothramycin gene cluster, *por1*. Genes coding for regulatory or transport proteins are coloured in blue; genes coding for APD precursor (4-alkyl-L-proline derivatives) biosynthetic enzymes are coloured in brown and genes coding for NRPSs are coloured in red. Homologous genes are connected by lines.

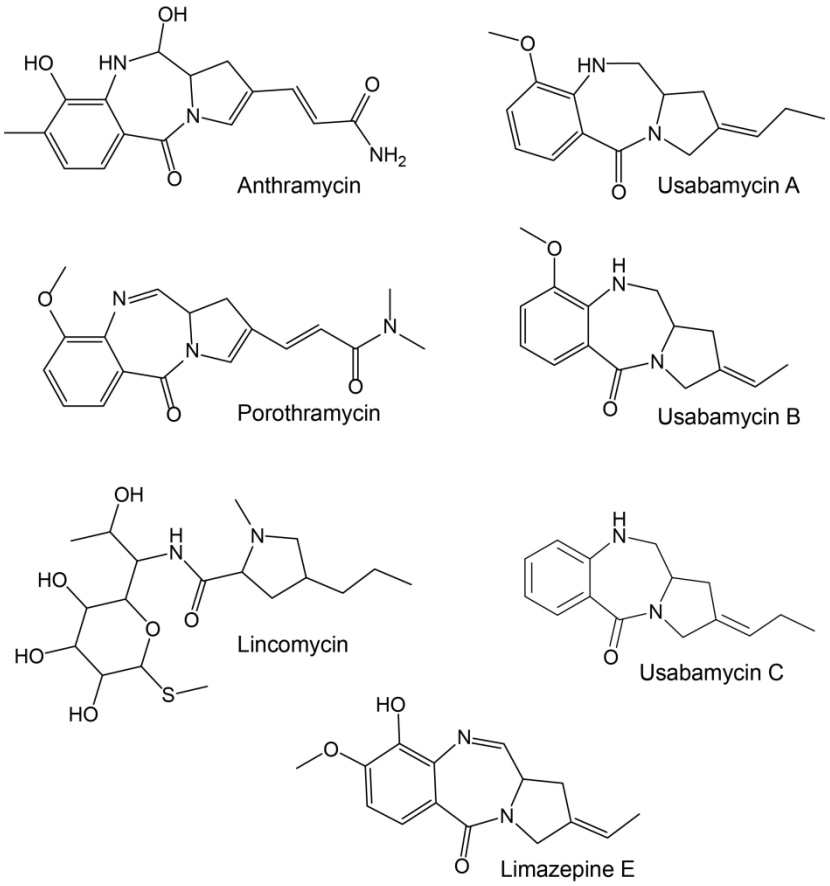


Figure 4S. Structures of several representatives of of pyrrolobenzodiazepines (PBDs) family.

Table 2S. Homologous genes in described PBD biosynthetic gene clusters.

Streptomyces albus subsp. *chlorinus* NRPS cluster gene	Predicted function	Homologue in anthramycin BGC	Identity (%)	Homologue in porothramycin BGC	Identity (%)
chl1	Transcriptional regulatory protein	*orf25*	67	*por2*	91
chl2	UvrABC system protein	*orf8*	49	*por3*	93
chl3	HTH-domain protein	-	-	*por4*	92
chl4	YafY family transcriptional regulator	-	-	-	-
chl5	Putative methyltransferase	*orf5*	76	*por10*	90
chl6	Gamma-glutamyltransferase	*orf6*	80	*por11*	92
chl7	Putative L-DOPA 2,3-dioxygenase	*orf12*	74	*por13*	89
chl8	Putative tyrosine hydroxylase	*orf13*	58	*por14*	83
chl9	Putative F-420 dependent reductase	*orf14*	82	*por15*	91
chl10	Putative isomerase	*orf15*	76	*por16*	86
chl11	Putative NRPS	*orf21*	72	*por20*	87
chl12	Putative NRPS + kinurenine monooxygenase	*orf22*	71	*por21*	89
chl13	Hypothetical protein	-	-	-	-
chl14	Hypothetical protein	-	-	-	-
chl15	NmrA family transcriptional regulator	-	-	por22	90

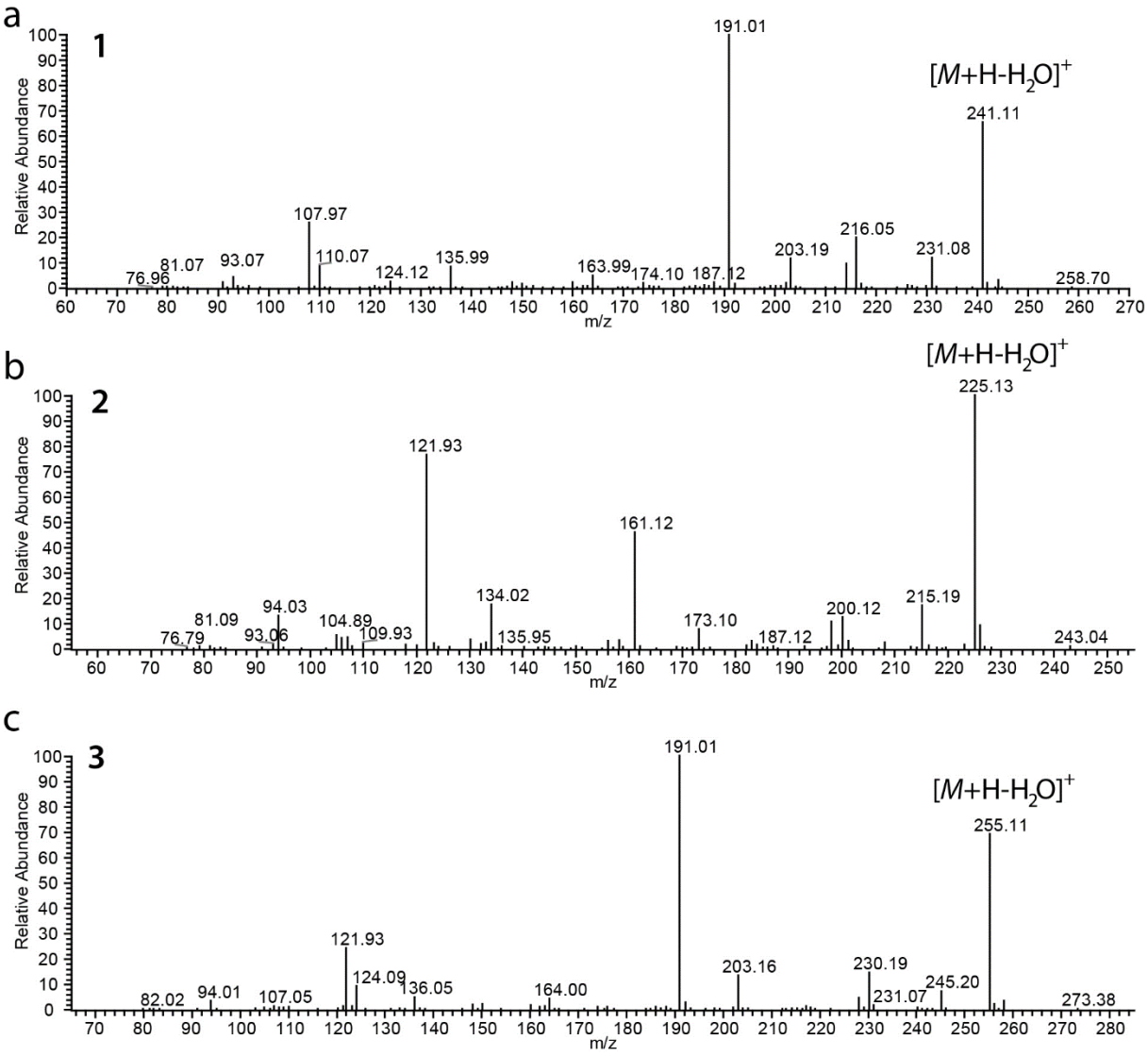


Figure 5S. Fragmentation pattern of compounds indetified in the extract of *S. lividans* ΔYA9 with BAC2I4. Mass chromatogram of metabolites of *S. lividans* ΔYA9 with BAC2I4 grown in DNPM medium for 6 days at 28°C. The metabolites were extracted with ethyl acetate and the accurate mass was determined using Thermo LTQ Orbitrap XL coupled to UPLC Thermo Dionex Ultimate 3000 RS with 20 min gradient protocol. **(a)** Fragmentation pattern of compound **1** with *m/z* 259.14377 $[M+H]^+$, **(b)** fragmentation pattern of compound **2** with *m/z* 243.14883 $[M+H]^+$, and **(c)** fragmentation pattern of compound **3** with *m/z* 273.15933 $[M+H]^+$.

Table 3S. Strains, Plasmids and BACs used in this study.

Strains	Features	Reference or source
Streptomyces strains		
Streptomyces lividans TK24	*S. lividans* wild type	[2]
Streptomyces albus Del14	Derivative of S. albus J1074 with deletion of 15 gene clusters	[3]
Streptomyces coelicolor M1154	Derivative of *S. coelicolor* M1152 with *rpsL* mutation	[4]
Streptomyces lividans ΔYA9	Derivative of *S. lividans* TK24 with 9 inactivated gene clusters	This work
*Streptomyces lividans Δ*YA10	Derivative of *S. lividans* ΔYA9 with 10 inactivated gene clusters and one additional *attB* site	This work
Streptomyces lividans ΔYA11	Derivative of *S. lividans* ΔYA10 with 11 inactivated gene clusters and two additional *attB* site	This work
E. coli strains		
E. coli ET12567 pUB307	Donor strain for intergeneric conjugation	[5]
E. coli WM6026	Donor strain for intergeneric conjugation	[6]
E. coli GB05-red	Strain used for Red/ET	[7]
Plasmids		
patt-saac-oriT	Resistance cassette plasmid containing a synthetic fragment with *aac(3)IV*, *oriT*, *B-CC*, *P-GG* and *loxP* sites	[8]
pUWLint31	pUWLCREdeltaKpnI with the *XbaI/BamHI* fragment of pKHint31 containing *int*	[8]
phygattB	Resistance cassette plasmid containing hygromycin resistance marker flanked with *MssI* restriction sites and *attB* sequence	[3]
BACs		
pSMARTgus	Derivative of pSMART with the *gusA* gene	[8]
1629	pSMARTgus derivative containing a fragment of the *S. lividans* chromosome with Undecylprodigiosin cluster №10	[9]
1704	pSMARTgus derivative containing a fragment of the *S. lividans* chromosome with Actinorhodin №14	[9]
1443	pSMARTgus derivative containing a fragment of the *S. lividans* chromosome with melanin cluster №17	[9]
1320	pSMARTgus derivative containing a fragment of the *S. lividans* chromosome with melanin cluster №13	[9]
1715	pSMARTgus derivative containing a fragment of the *S. lividans* chromosome with germicidin cluster №19	[9]
1008	pSMARTgus derivative containing a fragment of the *S. lividans* chromosome with CDAcluster №15	[9]
1161	pSMARTgus derivative containing a fragment of the *S. lividans* chromosome with coelimycin cluster №6	[9]
1250	pSMARTgus derivative containing a fragment of the *S. lividans* chromosome with coelibactin cluster №2	[9]
490	pSMARTgus derivative containing a fragment of the *S. lividans* chromosome with t1PKS cluster №24	[9]
1468	pSMARTgus derivative containing a fragment of the *S. lividans* chromosome with coelichelin cluster №21	[9]
1092	pSMARTgus derivative containing a fragment of the *S. lividans* chromosome with abCPK cluster №5	[9]
1629-am	Derivative of 1629 with №10 gene cluster fragment substituted with a cassette from patt-saac-oriT	This work
1704-am	Derivative of 1704 with №14 gene cluster fragment substituted with a cassette from patt-saac-oriT	This work
1443-am	Derivative of 1443 with №17 gene cluster fragment substituted with a cassette from patt-saac-oriT	This work
1320-am	Derivative of 1320 with №13 gene cluster fragment substituted with a cassette from patt-saac-oriT	This work
1715-am	Derivative of 1715 with №19 gene cluster fragment substituted with a cassette from patt-saac-oriT	This work
1008-am	Derivative of 1008 with №15 gene cluster fragment substituted with a cassette from patt-saac-oriT	This work
1161-am	Derivative of 1161 with№6 gene cluster fragment substituted with a cassette from patt-saac-oriT	This work
1250-am	Derivative of 1250 with№2 gene cluster fragment substituted with a cassette from patt-saac-oriT	This work

490-am	Derivative of 490 with №24 gene cluster fragment substituted with a cassette from patt-saac-oriT	This work
1468::hygattB	Derivative of 1468 with №24 gene cluster fragment substituted with a cassette from phygattB	This work
1092::hygattB	Derivative of 1092 with №5 gene cluster fragment substituted with a cassette from phygattB	This work
1468::attB	Derivative of 1468::hygattB after cutting out the *hyg* gene with *MssI*	This work
1092::attB	Derivative of 1092::hygattB after cutting out the *hyg* gene with *MssI*	This work
1468::attBamoriT	Derivative of 1468::attB with substitution of *cat* gene with *aac(3)IV*-OriT fragment	This work
1092::attBamoriT	Derivative of 1092::attB with substitution of *cat* gene with *aac(3)IV*-OriT fragment	This work
2I4	pSMART derivative containing a fragment from *S. albus* subsp. *chlorinus* NRRL B-24108 with NRPS gene cluster	Intact Genomics, USA
PMP31	pOJ436 derivative, containing the griseorhodin biosynthetic cluster	[10]
pIJ12003a	12.9 Kbp *tun*-gene cluster cloned into the pRT802	[11]
pCindelX	Plasmid containing cinnamycin with deletion of OH group gene cluster and *aac(3)IV* resistance marker	[12]

Table 4S. Primers used in this study.

Primers	Sequences	Description
Cl10_F	CTACGTGGACGAGGAGGAGTCCTTCGTACGGTCACGGTCCTGGTGGTGGTTCGACCCGGTA CCGGAGTA	Deletion of cluster Nr° 10
Cl10_R	CAACGCGCGACATGGCGGGGCAGCAGGAACGGATACGGGCGGCGCAGGATACTACGCCCC CAACTGAGAG	
Cl13_F	GTGAACGACGGCGTGAACGGCGTCGCACTCGCCGTCTTCATCTTCTTCTTTCGACCCGGTAC CGGAGTA	Deletion of cluster Nr° 13
Cl13_R	GTGAGGATCTTCGGCAAGGCACGGCACCGGCCCTCCGCCTCTTGGCGGCAACTACGCCCCC AACTGAGAG	
Cl14_F	GAGCCCCCGAACACTCCCGGCCACGGCCCTGCTTGAGGGGCTGCCCGACGTCGACCCGGTA CCGGAGTA	Deletion of cluster Nr° 14
Cl14_R	TGGTCGCGGATCGCCCGGCTGGTGTCCGGCCGTCCGCGTGCCGTCTGGGTACTACGCCCCC AACTGAGAG	
Cl17_F	TCCACCGGGTCGCCGACTCGTGCTCGTCGGCCGGGCGGTGGGTGCGGAACTCGACCCGGTA CCGGAGTA	Deletion of cluster Nr° 17
Cl17_R	GAGGCGCTGCTCGCCGCGCTGCACGAGGGCCGCATCGCGGGTGCGGGCGTACTACGCCCCC AACTGAGAG	
Cl19_F	TGAGGTCAACTCCGCCGGGCGGCTGCCGTACGGGCAGCTTTCCGGGGATCCGTCGACCC	Deletion of cluster Nr° 19
Cl19_R	ACTGGTACCTGACGAGCGTTCAGGCGGGGTTCGAGCCGTTGTAGGCTGGAGCTGCTTCG	
Cl15_F	TAGTGCGGGTCGATCGGCAGGTAGGCGGCCCCGGACTTGTTCCGGGGATCCGTCGACCC	Deletion of cluster Nr° 15
Cl15_R	GTGTCTGAGAACTCTTCGGTTCGGCACGGTCTGACGAGCTGTAGGCTGGAGCTGCTTCG	
Cl6_F	TCCCTTCAGGCGCTGGTGACGGGCTCTTCGCGGGCCGCGTTCCGGGGATCCGTCGACCC	Deletion of cluster Nr° 6
Cl6_R	ATGGTGTGGTGGAGGGGTTGGTCGAGGTGGGGGTCGAGTTGTAGGCTGGAGCTGCTTCG	
Cl2_F	AGGAACTCGATGGTCCCGTCGGGCAGGTAGCGGCCGAGGTTCCGGGGATCCGTCGACCC	Deletion of cluster Nr° 2
Cl2_R	TGGTACGGCGAGAAGGAAGTCACCGCGGTCGCCCAGGCGTGTAGGCTGGAGCTGCTTCG	
Cl24_F1	AACGCATGCGGGTCTCACTTCACGGCGGGTTCGGACAGGTTCCGGGGATCCGTCGACCC	Deletion of cluster Nr° 24
Cl24_R	AAACGGCAGCCTGGAGAGCGCCCCCTGCGGTGTCAACCTTGTAGGCTGGAGCTGCTTCG	
Cl21attB_F	CGTGGTCGAAGATCATGAACCGGGTGCTGGTGGTTCCCTTGTAGGCTGGAGCTGCTTCG	Deletion of cluster Nr° 21
Cl21attB_R	AACAGGCCGGTGCGACCGTGCGGGTCTCCAGGTCGGCCTTTCCGGGGATCCGTCGACCC	
Cl5attB_F	GGCGTGTTCAGGGCCGGCGGCCGTCCGGGATCAGGACGGTGTAGGCTGGAGCTGCTTCG	Deletion of cluster Nr° 5
Cl5attB_R	ACGCATGTCGTTCCTGCATGACCTGCTCACGGCGCAGGCTTCCGGGGATCCGTCGACCC	
Cm_F	TTCAGGAGCTAAGGAAGCTAAAATGGAGAAAAAAATCACTTCCGGGGATCCGTCGACCC	Substitution of *cat* gene with *aac (3)IV* gene and *oriT* sequence
Cm_R	AGGCTTTTGACTTCTGTCACCTAGGTTACGCCCCGCCCTTGTAGGCTGGAGCTGCTTCG	
Cl10C_F	GCGAGACCGGCTCCGAGA	Verification cluster Nr° 10 deletion
Cl10C_R	GTGCGTGATCCACAGCGGA	
Cl13C_F	TCGGCATCCCGTTCTTCTA	Verification cluster Nr° 13 deletion
Cl13C_R	TCCCGAGGCAGCCTTACCGA	
Cl14C_F	CTGGGGGAACGACCGGATT	Verification cluster Nr° 14 deletion
Cl14C_R	TGGTCGGCCTGCTGTCGAT	
Cl17C_F	TCGGGACGCGGGCCGAC	Verification cluster Nr° 17 deletion
Cl17C_R	CTCCGTGCACCTGGCGCT	
Cl19C_F	AGTTGCTGCACGGTCTCAC	Verification cluster Nr° 19 deletion
Cl19C_R	AACGGCTCGAACGACGTG	
Cl15C_F	AGGTGTAGATCGTGTACGC	Verification cluster Nr° 15deletion
Cl15C_R	AATAAGCATATTGACCAGCT	
Cl6C_F	AAGTCGAGTACGGCGATCT	Verification cluster Nr° 6 deletion
Cl6C_R	TGAGTTCGCCGATGGAGTGT	
Cl2C_F	CGAGTCGTCCACGACGA	Verification cluster Nr° 2deletion
Cl2C_R	GGACCTCTTCGCCGCGT	
Cl24C_F1	CCGCCGTCAGACTCGTA	Verification cluster Nr° 24 deletion
Cl24C_R	GCGTGCTGTCGGTAGGAA	
Cl21C_F	TTGCGCAGCGCGTTCTGCAT	Verification cluster Nr° 21 deletion
Cl21C_R	TGTCGCGTGTCGCCTGTCA	
Cl5C_F	CACGCCGTCCGTACGAGT	Verification cluster Nr° 5 deletion
Cl5C_R	AACCGGCGCTCACCGGC	
CmC_F	AATCTGATCGGCACGTAAGA	Verification Substitution of *cat* gene
CmC_R	GTCCGTGGAATGAACAATG	

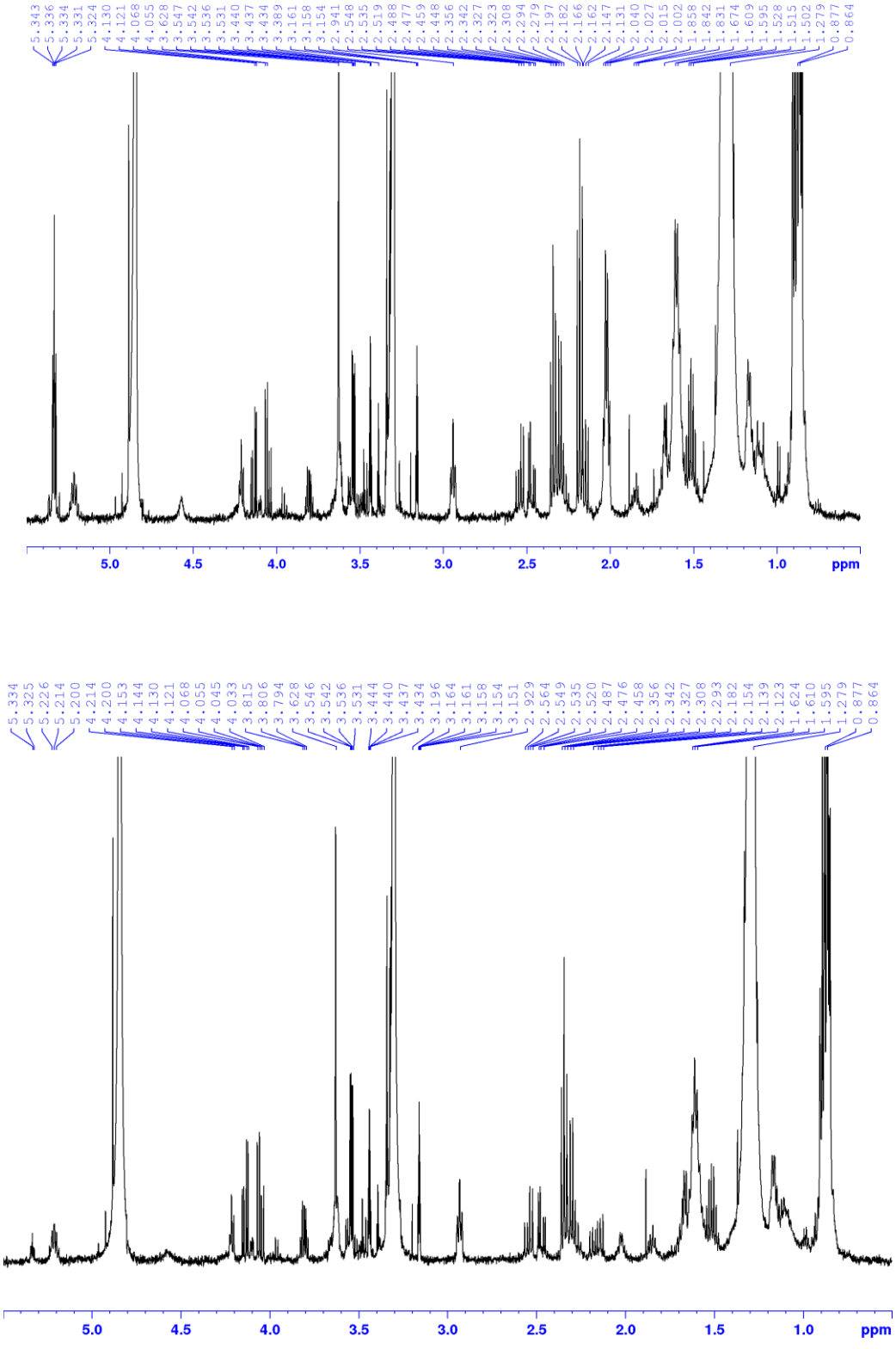


Figure 6S. [1]H NMR overview spectrum of 4b (above) and 4c (below).

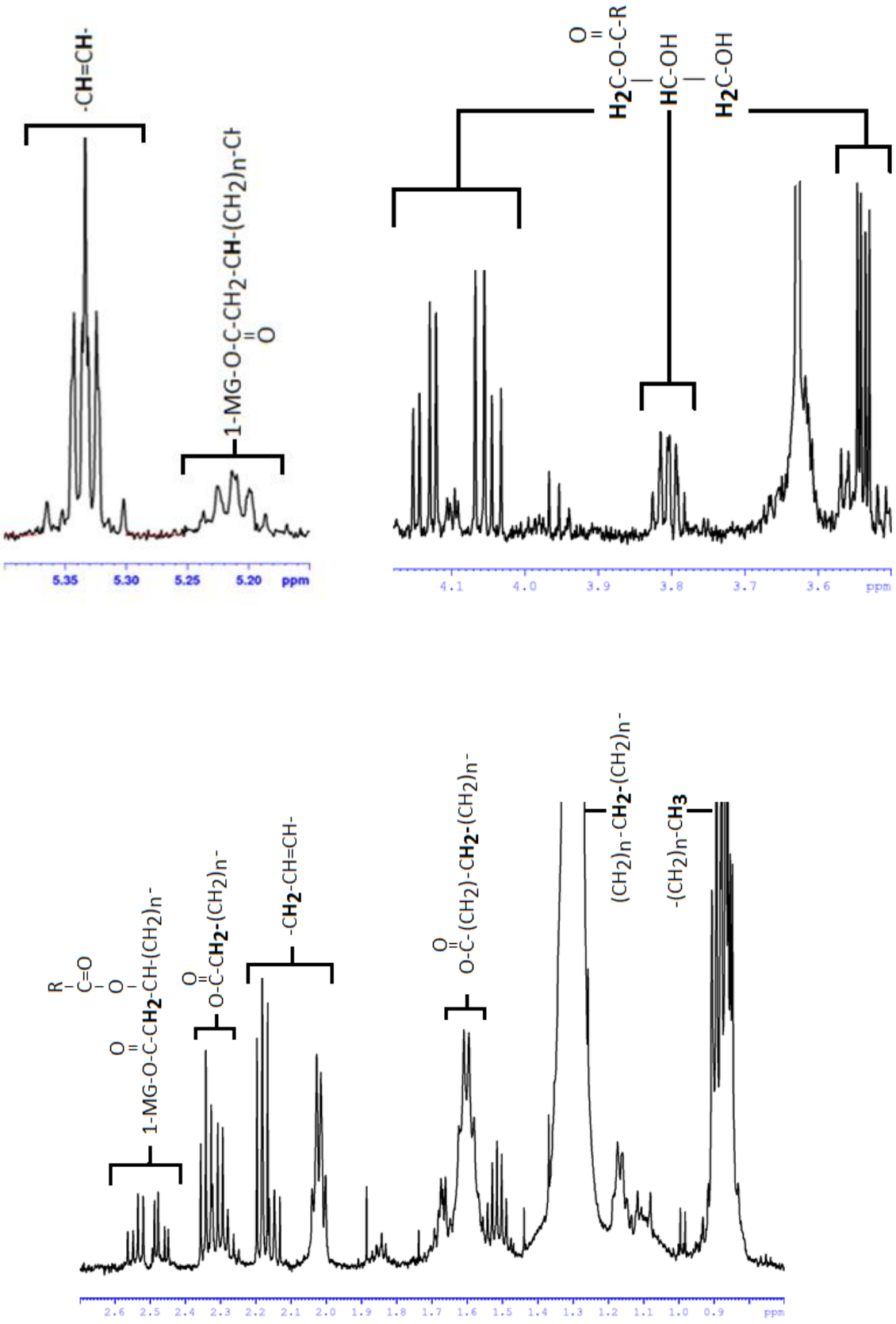


Figure (6+1)S. Three enlarged views of the ^{1}H NMR of 4b. Relevant signals for the 1-MG are assigned.

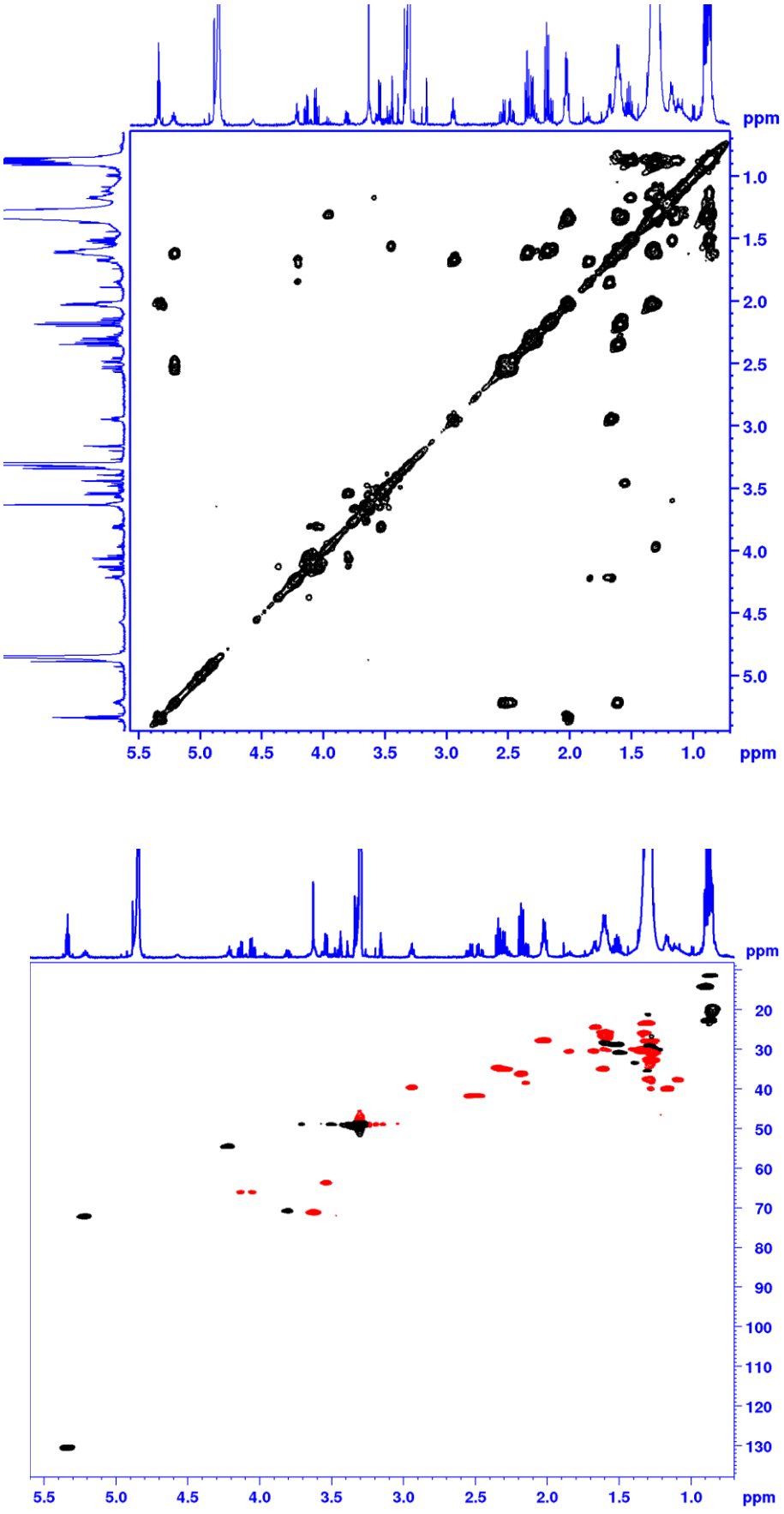


Figure (6+2)S. 2D NMR HHCOSY (above) and HSQC (below) of 4b.

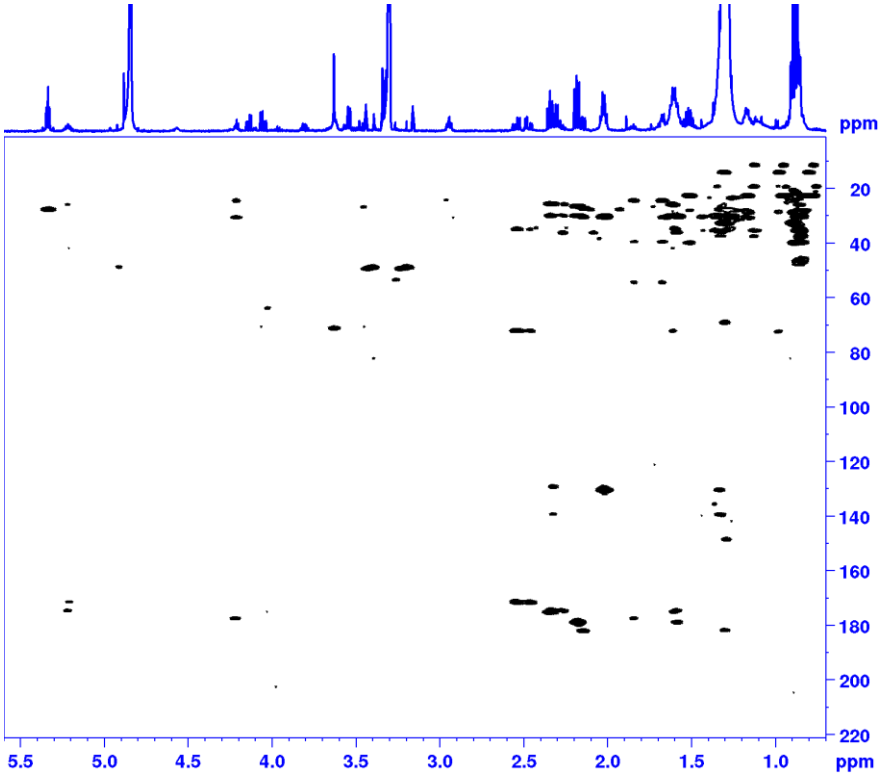


Figure (6+3)S. 2D NMR HMBC of 4b

4.8. References for the Supplementary Information

1. Busche T, Tsolis KC, Koepff J, Rebets Y, Ruckert C, Hamed MB, Bleidt A, Wiechert W, Lopatniuk M, Yousra A *et al*: Multi-Omics and Targeted Approaches to Determine the Role of Cellular Proteases in *Streptomyces* Protein Secretion. *Front Microbiol* 2018, 9:1174.
2. Ruckert C, Albersmeier A, Busche T, Jaenicke S, Winkler A, Friethjonsson OH, Hreggviethsson GO, Lambert C, Badcock D, Bernaerts K *et al*: Complete genome sequence of *Streptomyces lividans* TK24. *J Biotechnol* 2015, 199:21-22.
3. Myronovskyi M, Rosenkranzer B, Nadmid S, Pujic P, Normand P, Luzhetskyy A: Generation of a cluster-free *Streptomyces albus* chassis strains for improved heterologous expression of secondary metabolite clusters. *Metab Eng* 2018.
4. Gomez-Escribano JP, Bibb MJ: Engineering *Streptomyces coelicolor* for heterologous expression of secondary metabolite gene clusters. *Microb Biotechnol* 2011, 4(2):207-215.
5. Flett F, Mersinias V, Smith CP: High efficiency intergeneric conjugal transfer of plasmid DNA from Escherichia coli to methyl DNA-restricting Streptomycetes. *Fems Microbiol Lett* 1997, 155(2):223-229.
6. Blodgett JA, Thomas PM, Li G, Velasquez JE, van der Donk WA, Kelleher NL, Metcalf WW: Unusual transformations in the biosynthesis of the antibiotic phosphinothricin tripeptide. *Nat Chem Biol* 2007, 3(8):480-485.
7. Fu J, Teucher M, Anastassiadis K, Skarnes W, Stewart AF: A Recombineering Pipeline to Make Conditional Targeting Constructs. *Method Enzymol* 2010, 477:125-144.

8. Myronovskyi M, Rosenkranzer B, Luzhetskyy A: Iterative marker excision system. *Appl Microbiol Biotechnol* 2014, 98(10):4557-4570.
9. Rebets Y, Tsolis KC, Guðmundsdóttir EE, Koepff J, Wawiernia B, Busche T, Bleidt A, Horbal L, Myronovskyi M, Ahmed Y *et al*: Characterization of Sigma Factor Genes in *Streptomyces lividans* TK24 Using a Genomic Library-Based Approach for Multiple Gene Deletions. *Frontiers in Microbiology* 2018, 9(3033).
10. Li A, Piel J: A gene cluster from a marine Streptomyces encoding the biosynthesis of the aromatic spiroketal polyketide griseorhodin A. *Chem Biol* 2002, 9(9):1017-1026.
11. Wyszynski FJ, Lee SS, Yabe T, Wang H, Gomez-Escribano JP, Bibb MJ, Lee SJ, Davies GJ, Davis BG: Biosynthesis of the tunicamycin antibiotics proceeds via unique exo-glycal intermediates. *Nat Chem* 2012, 4(7):539-546.
12. Lopatniuk M, Myronovskyi M, Luzhetskyy A: *Streptomyces albus*: A New Cell Factory for Non-Canonical Amino Acids Incorporation into Ribosomally Synthesized Natural Products. *Acs Chem Biol* 2017, 12(9):2362-2370.

5. Discussion

Actinomycetes, especially Streptomycetes, produce a variety of NPs that have major importance in the medical and industrial use [32]. These compounds come from diverse secondary metabolism comprising a set of specialized biochemical pathways leading to formation of variety of chemical structures. Most of these compounds, with some exceptions like siderophores, are not required for strain survival. But at some growth conditions they provide advantages for the producing bacteria. On the other hand, these compounds are used as antibiotics, anticancer, antifungal and veterinary drugs [1,29]. The biosynthesis of NPs by Actinomycetes is controlled by complex regulatory cascades [150]. These include both pleiotropic regulatory elements influencing the global cellular processes and specific regulators dedicated to particular metabolic pathway [50]. Such complex regulatory network was developed by Actinobacteria in order to respond efficiently to different external stimuli by changing both primary and especially secondary metabolisms. Since the production of these compounds is beneficial to the strain only in certain conditions it is not surprising that most of gene clusters responsible for production of NPs are silent in laboratory conditions [151]. Only a small proportion of this incredible chemical potential is accessible in laboratory or production facilities. The major obstacle preventing the exploitation of silent biosynthetic pathways is the lack of understanding of their regulation and stimuli triggering their expression. During last decade numerous methods and strategies have been developed in order to gain access to secondary metabolism potential of Actinobacteria [48,49]. They were successfully applied resulting in production of new NPs some of which are very promising pharmaceutical leads like teixobactin [152]. The complexity and multi-level organization of regulatory network controlling secondary metabolism in Actinobacteria allows development of different approaches for activation of silent gene clusters covering from single to multiple pathways [48,49]. Both pleiotropic and pathway-specific strategies have their advantages and limitations. When specific approaches based on entire cluster refactoring or a manipulation of pathway associated regulatory genes is often more effective in obtaining the desired compound, they do not provide insights why the pathway is silent. On the other hand, although the pleiotropic approaches shed light on elements controlling secondary metabolism, they often are not selective and applied procedures and are tied with the laborious screening processes.

The reporter-guided selection of clones with the increased transcription level of genes for particular secondary metabolism pathway is a great solution to the existing screening problem. The simple idea that the level of metabolite production correlates with the level of respective biosynthetic genes expression was originally turned into a method for selection of lovastatin

overproducing *Aspergillus terreus* strains [153]. To monitor the transcription level of target gene its promoter is fused to the promoterless reporter that allows simple detection by antibiotic resistance, fluorescence and enzymatic activity [154-156]. In such way colonies with activated or increased expression of desired pathway can be easily distinguished. This idea was further adapted to improve clavulanic acid production in *S. clavuligerus* [99]. However, the use of single antibiotic resistance reporter resulted in high rate of false positive clones (around 50 %). Same approach with two reporter genes was used to activate two silent gene clusters: *jad* in *S. venezuelae* and *pga* in *Streptomyces* sp. PGA64 [157]. The use of two reporter genes strongly decreased the rate of the false positive clones [99]. However, the traditional mutagenesis used in these studies to create mutant libraries did not allow simple identification of mutation loci. Thus, the described procedure cannot be used in full to discover new genes controlling NPs biosynthetic pathways.

In this thesis, we modified a reporter-guided screening strategy first of all by introducing gene encoding β-glucuronidase (*gusA*) as a reporter. This reporter is easy to use, inexpensive and allows simple visual selection of mutants and quantitative analysis of gene expression both *in vivo* and *in vitro* [156]. The availability of different alleles of *gusA* with different efficiency of translation initiation allows tuning the basal level of output signal according the screening needs. On the other hand, the use of antibiotic resistance reporter gene when combined with UV or chemical mutagenesis often result in arising of resistant clones caused by secondary mutations in antibiotic target or overexpression of a multidrug efflux pump rather than actual increase in fused promoter activity [99]. By applying β-glucuronidase reporter we avoided the antibiotic pressure and thus potentially expected the lower rate of false-positive clones. However, with the switch of reporter the procedure was also changed from selection to screening, which involved visual assessment of reporter activity of thousand colonies. Due to this some number of true-positive clones were abandoned as well as some colonies were assessed incorrectly as those with activation of gene cluster of interest. Indeed, the use of two reporter genes one of which provide the possibility of direct selection and second is used for post-selection screening and validation will provide the best results. In most cases, for both single and double reporter approaches, the observed drawbacks are related to the fact that the compound production is monitored indirectly by estimating the corresponding biosynthetic gene promoter activity rather than by direct sensing. The latter could be achieved by the use of whole cell biosensors, applying NP specific transcriptional factor as a sensing part and reporter gene fusion with the controlled by this transcriptional factor promoter [97,158].

The choice of promoter for reporter-guided screening is determinative for success of procedure; therefore, the regulatory genes and the genes for the earlier steps in the biosynthesis should be of the primary interest. This will guarantee the proper expression of entire gene cluster and consequently production of mature compound rather than shunt products accumulation. As the example of later case could be the attempt to activate the PTM gene cluster in *S. albus* by inserting the constitutive promoter in front of PKS-NRPS encoding gene and refactoring of similar PTM gene cluster from *S. griseus* [94,96]. In both cases the strains were accumulating alteramides that are known shunt products in the assembly line of PTMs [159,160]. We have selected promoters of *XNR_0204*, *XNR_5634*, *XNR_5808* and *XNR_5915* key genes from PKS-NRPS (N° 2), NRPS (N° 16), PKSIII (N° 18) and PKS-NRPS (N° 21) gene clusters putatively responsible for secondary metabolites production of *S. albus* and fused them to two alleles of *gusA* gene with ATG and CTG start codons. This allowed selecting truly silent gene cluster by assessing the corresponding promoter activity and thus to set the baseline for further screening procedure. As result, promoter of PKS-NRPS gene *XNR_0204* from PTM gene cluster in combination with ATG variant of *gusA* was selected for further studies.

The second important modification that was applied by us was the use of transposon mutagenesis. First of all, this prevents mutations within the reporter construct that will affect the screening results. Of course, the appearance of spontaneous mutations cannot be excluded, but their frequency is significantly lower than in the case of UV or chemical mutagenesis. Moreover, the transposon insertion loci, responsible for the secondary metabolism alteration can be easily identified thus providing new insights into functionality of regulatory network controlling NPs production in Actinobacteria. A prominent example of such transposon utilization is a work of Xu *et al*. describing development and use of Tn5 transposon system to obtain mutants with altered production level of undecylprodigiosin in *S. coelicolor* [161]. As result several new genes influencing undecylprodigiosin biosynthesis were identified, including 17 genes inside undecylprodigiosin biosynthetic gene cluster. Similar, Horbal and co-authors used Tn5 derivative to obtain mutants of *S. globisporus* with affected production of landomycins thereby leading to identification of regulatory elements controlling biosynthesis of this antibiotic [162]. Beside the two mentioned cases, numerous other experiments have proven the efficiency of transposon approach for identifying the regulatory genes controlling both morphological and metabolic differentiation in Actinobacteria and thus increasing our understanding of complexity of regulatory cascades in these bacteria [163,164]. Most of such experiments were performed using the Tn5 based transposons, however recently a new transposition element based on *Himar1* transposon was developed for use in actinomycetes [165]. The *Himar1* transposon was used to

generate random mutants in *S. coelicolor* M145, resulting in the identification of four novel regulators controlling actinorhodin production. This transposon system is independent from host-specific factors [166] and it requires TA dinucleotides for successful insertion in contrast to Tn5 which needs longer consensus motifs [167]. The transposon efficiency of *Himar1* is higher than of the Tn5 [168]. These advantages of *Himar1* based system made it a preferred way for generating the genetic diversity in *S. albus* to be used in combination with reporter-guided screening procedure.

PTMs are antifungal compounds that share a macrocyclic lactam ring core structure combined with tetramic acid along with different sets of carbocyclic rings. They originate from a conserved PKS-NRPS assembly line that acts iteratively [92]. Different tailoring modifications lead to further structural diversity of PTMs. These gene clusters are widely distributed among phylogenetically diverse bacteria and most of them are silent including the one encrypted in *S. albus* genome [92,93]. As in many other cases the reasons for this silencing are not very clear and is hidden within regulatory network of the strain, many elements of which remains to be discovered. For now, we have only a shallow understanding of mechanisms and triggering factors controlling the secondary metabolism of Actinobacteria and this knowledge is solely coming from studies of few strains like *S. coelicolor, S. lividans* and *S. griseus* [50].

Four transposon mutants with increased promoter activity of *XNR_0204* gene were selected as a result of designed screening. However, only two of them had affected metabolic profiles. *S. albus* ATGSal2P2::Tn14 strain have shown the most significant changes in secondary metabolites production. Beside the accumulation of altermides and frontalamides resulting from activation of targeted cluster the strain was found to produce the variety of other compounds including antimycins and candicidins. Similar but less prominent phenotype was observed in the case of *S. albus* ATGSal2P2::Tn13 mutant. This means that mutations in both strains occurred in loci that is pleiotropically controlling several secondary metabolism pathways in *S. albus*. Cross-regulation among different biosynthetic pathways is a well-known fact described in *Streptomyces* [169,170]. Similar fact was observed by Salas group when the activation of candicidins production in *S. albus* J1074 also simultaneously triggered accumulation of antimycins [94]. The presence of candicidins biosynthetic gene clusters is relatively common in the genome of actinomycetes [171,172]. Moreover, it is also quite frequent that antimycin gene clusters is in the vicinity of the candicidins clusters [173,174]. It is interesting, that all three compounds, antimycins, candicidins and PTMs, have an antifungal activity. Besides that, *S. albus* is known to carry the silent gene cluster that when activated resulted in production of another antifungal substances surugamides [175].

Altogether, at least four NPs with activity against fungi are produced by the strain. This fact together with our finding of co-regulation of at least three of these compounds production makes us to think that the strain has acquired the corresponding clusters as a mechanism of defence against fungal invasion. So, we believe that the coordinated control over the production of at least three different antifungal NPs ensure the efficient adaptation of the strain to the competitive environment. The parental *S. albus* strain G was isolated from air in the laboratory of Welsch, Belgium and was originally reported to be highly active against at least 10 species of fungi [176].

We have identified that the reason for activation of antifungal NPs production in the *S. albus* ATGSal2P2::Tn14 is insertion of transposon into the *XNR_3174* gene cutting off the region coding for DNA binding helix-turn-helix domain of putative transcriptional regulator. The deletion of *XNR_3174* had same phenotype as the ATGSal2P2::Tn14 mutant. This gene is located in the core genome of the strain and could be found in the genome of large group of Streptomycetes strains. The percentage of amino acid identity of different orthologues is ranging from 60 to 100%. The high degree of conservation of *XNR_3174* orthologues sequence implies the importance of its function. Interestingly, the mutant phenotype is condition specific since it is the most drastic on MS (and its liquid analogue NL19) media, which were used in the initial screening procedure.

Besides the antifungal compounds the *XNR_3174* was found to negatively control production of another NP butenolide 4. This compound was originally isolated from *Streptomycete* strain B 3497 [177]. Three other butenolide derivatives (**1**-**3**) were isolated from *Streptomycete* strain B 5632. Compound **4** had no biological activity against *Escherichia coli, Staphylococcus aureus*, *Candida albicans*, *Mucor miehei,* or *S. viridochromogenes* when tested. Besides production of butenolides, both strains were also accumulating antimycins. Structurally related compound avenolide was found to be involved in control of avermectin production in *S. avermitilis* [178]. Similar to A-factor it was proposed to act as hormone-like compound. The biosynthesis of avenolide in *S. avermitilis* requires two enzymes: acyl-CoA oxidase (Aco) and a putative cytochrome P450 hydroxylase (Cyp17) [178]. The orthologue of Aco encoding gene *XNR_2339* was found within the genome of *S. albus*. Its deletion in the *S. albus* Δ3174 caused complete cessation of butenolide 4 production as well as accumulation of antimycins, candicidins and PTMs. Also, *gusA* based assessment of *XNR_0204* promoter activity assay showed that the expression of PTMs biosynthetic gene cluster depends on the presence of butanolide 4 biosynthesis gene *XNR_2339*. This make us to believe that the actual trigger of expression of the antifungal part of *S. albus* clustume is butenolide 4, which production is in turn controlled by the

XNR_3174 transcriptional regulator. Surprisingly, the deletion of putative butenolide receptor protein encoding gene, did not effect the production of secondary metabolites in *S. albus* J1074. Interestingly, genes coding for Aco were found within the genomes of all strains carrying the orthologues of *XNR_3174*, pointing on the universality of proposed regulatory mechanism.

The universal nature of butenolides regulatory function was recently studied by Nihira and co-authors [179]. They have used the *S. avermitilis* Δaco mutant, lacking production of avenolide and thus deficient in avermectin accumulation. The dual-culture assays of *S. avermitilis* mutant with *S. albus* led to the production of avermectin. This suggests that the butenolides produced by *S. albus* can substitute for avenolide function in *S. avermitilis*. Finally, co-cultivation of *S. avermitilis* Δaco with 51 actinomycetes strains (40 *Streptomyces* and 11 endophytic actinomycetes) resulted in restoring the avermectin production in 24% of cases [180]. This clearly proves that butenolides, similar to γ-butyrolactone-type signaling molecules, are also common in Actinobacteria and seems to play an important role in controlling secondary metabolism and as interspecies communication. We found that at least 19 strains are carrying genes encoding butenolides biosynthesis, *XNR_3174* orthologues and antimycin, candicidin and PTMs gene clusters within their genomes. These strains were grouped into "*S. albus*" sub-group. Such clear correlation of gene clusters and butenolide regulatory system indicates a common evolutional trend for this group of species.

The universality of butenolide regulatory cascade made us to assume that deletion of *XNR_3174* can have an effect on expression of heterologous gene clusters. As result, the strain *S. albus* Δ3174 indeed was found to be superior for production of griseorhodin, aranciamycin and pamamycin. This result indicates that modulating the butenolides regulatory system functionality can be used to create better hosts for the heterologous production of secondary metabolites based on strains natively carrying the system.

The discovery of the regulatory elements silencing secondary metabolism in *S. albus* has proven the efficiency of developed mutagenesis and screening approach. The use of a double reporter system will further increase the output of the proposed procedure by decreasing the number of falls-positive hits [99]. The use of transposon with the constitutive promoters will broaden the number of regulatory genes that could be identified by including positive regulators. With these improvements the described approach can be used to activate the silent gene clusters in different Streptomycetes as well as to discover new regulatory elements, controlling secondary metabolism. This will consequently deepen our understanding of the complex life processes of these fascinating bacteria.

Heterologous expression plays an important role in drug discovery and studying of NPs biosynthesis. With the advances in sequencing and bioinformatics tools and cloning technology development, it has become possible to access an enormous untapped resource for the discovery of new NPs through heterologous expression. This includes not only the genomic but also a metagenomics DNA that often is even more interesting, since allows potentially reaching for the natural chemistry of uncultivable bacteria. *Escherichia coli*, *Saccharomyces cerevisiae*, *Pseudomonas putida, Myxococcus xanthus* and *Streptomyces* spp. are frequently used as heterologous hosts for the production of NPs depending on their origin. *P. putida* and *M. xanthus* are used to express gene clusters originated from myxobacteria [181]. *S. cerevisiae* is a typical species to be used for production of eukaryotic NPs like those from plants and fungi [182]. *E. coli*, despite being the most accessible and studied bacterium, typically has a limited usage for NPs production that is caused by lack of some precursors, like ethylmalonyl-CoA, and differences in genetic code usage [183]. Naturally, *Streptomyces* strains are the most used for NPs biosynthesis genes expression. Among them, *S. lividans* TK24, *S. coelicolor* A3(2), *S. albus* J1074 have been widely implicated as heterologous expression hosts. This is first of all caused by historical reasons, since these strains were studied in detail for a long time and are genetically amendable. Beside the long-term successful application of these Actinobacteria species they are still far from being ideal hosts. This is caused by the presence of endogenous secondary metabolism gene clusters that are often competing for precursors and parts of regulatory network controlling the gene expression [184]. This also in many cases complicates the identification of heterologous products due to the typical low yield of the later one. The understanding of the drawbacks of native species as heterologous hosts led to generation of strains deficient in endogenous secondary metabolites production [127]. At certain point it became obvious that partial or absolute deletion of endogenous gene clusters is an effective tool to streamline the precursors flow to the biosynthesis of compound of interest. As an example, the inactivation of avermectin pathway resulted in overproduction of oligomycin in *S. avermitilis* [185]. Understanding of this phenomenon promoted the construction of *S. albus, S. coelicolor* and *S. avermitilis* strains lacking the internal secondary metabolism gene clusters. These engineered strains were successfully used to express both genomic and metagenomic DNA [186].

S. lividans is closely related to *S. coelicolor* species and both are used for a long time in the genetic, biochemical, physiological and morphological studies. This makes these two species the most studied Actinobacteria, thus providing deep and clear understanding of regulatory network controlling secondary metabolism and morphological development processes. Such knowledge makes these strains perfect candidates for heterologous host generation. Indeed several *S.*

coelicolor host strains were reported and used to express secondary metabolites gene cluster. However, the strain is difficult to hand due to the low transformation and conjugation efficiency. At the same time, the absence of two DNA methylases and two secreted serine proteases explain the ability of *S. lividans* to uptake methylated DNA and reduced extracellular protease activity, respectively. Additionally, *S. lividans* is easier to cultivate. These characteristics often make this strain preferred heterologous host for the production of proteins and specialized metabolites [137,187].

Several groups attempted to generate better host based on *S. lividans* TK24 for production of secondary metabolites by removing *act* and *red* gene clusters [140]. However, there is no systematic strain engineering for creating simplified background host reported so far. To fulfill the gap in the range of engineered hosts we have chosen to delete eleven gene clusters from the chromosome of *S. lividans* TK24. The deleted gene clusters were chosen according to their high transcriptional level and the positive detection of the associated compounds in production media. As a result, all PKS and NRPS encoding gene clusters were removed from the chromosome of *S. lividans* except for type II PKS *whiE* gene cluster (Cluster N° 12), which is responsible for the spores' pigment biosynthesis [188]. The *whiE* cluster is ectopically expressed in aerial mycelium right before sporulation, and since *S. lividans* is not sporulating in liquid culture, it is not in competition for CoA substrates [188]. We attempted to delete gene cluster annotated to be responsible for biosynthesis of terpene (Cluster N° 3). Despite being able to obtain the double cross-over clones, they were found to grow poorly after re-plating on fresh medium. This gene cluster showed 100% similarity to the hopene biosynthesis genes in *S. coelicolor* [36]. The hopanoids are produced by *S. coelicolor* during transition from substrate to aerial mycelia growth on solid medium but are not accumulated in liquid culture [189]. These compounds are components of the cytoplasmic membrane and their function is to reduce the stress in aerial hyphae by decreasing the water permeability across the cell membrane [189]. Hopanoids have been reported to be not essential for growth of *S. scabies* [190], although we have observed their strong influence on growth of *S. lividans* lacking the respective gene cluster. Thus, the cluster N° 3 remained untouched in the engineered strains.

For sequential markerless deletion of nine gene clusters the IMES system was used [191]. This system allows recycling the resistance marker in the same genomic background without any interference with the remaining scar in the chromosome. The IMES was developed to generate cluster-free *S. albus* chassis strain [132]. As a result, we constructed strain *S. lividans* ΔYA9 carrying deletions of 9 endogenous gene clusters. This strain was growing slightly faster than

parental *S. lividans* TK24 and its morphological features were not affected. It already had a significantly cleaner secondary metabolites profile. With the idea to increase the production level of the host we have constructed two additional strains *S. lividans* ΔYA10 and *S. lividans* ΔYA11 harboring additional *attB* sites for integration of expressed cluster in place of remaining coelichelin and unknown NRPS gene clusters. These two *attB* sites were placed on opposite arms of *S. lividans* chromosome and distantly from the native *attB* locus. Such topology of *S. lividans* ΔYA11 genome should prevent the undesired recombination events between the integrated heterologous clusters. The deletion of gene clusters led to slightly better growth characteristics of obtained strains in liquid medium. These strains are supposed to have the biosynthetic precursors and energy supply at least partially redirected from secondary metabolism to primary metabolism and growth functions. Surprisingly, the secondary metabolite profile of engineered strains was not as expected. All three strains under certain growth conditions in some media commonly used for NPs production accumulate a large amount of hydrophobic compounds that were identified as mono-substituted glyceride lipids with the highly unusual unsaturated β-hydroxy fatty acids attached. The compounds are accumulated only during growth in some of the media but not others. Their functions and biosynthetic origin are not clear but the high level of accumulation may interfere with heterologous NPs production, especially those of polyketide origin.

In general, the engineered *S. lividans* performed better as a host for expression of heterologous gene clusters than a parental strain TK24 and absolutely outperform the engineered *S. coelicolor* M1154. Despite a long-standing history of *S. coelicolor* applications as a host for heterologous expression, similar observation was shown before when the production level in *S. coelicolor* hosts were less than in *S. albus* hosts [192]. On the other hand, in our experience, *S. lividans* and *S. albus* strains have preferences to particular biosynthesis pathways. When *S. albus* was more efficient with the polyketides biosynthesis, *S. lividans* and especially the engineered strains performed better in producing amino acid based compounds, like deoxycinnamycin. This is not surprising, since *S. lividans* was widely used for production of ribosomally synthesized and post-translationally modified peptides (RiPPs) by many groups for decades [193,194]. For instance, bottromycin biosynthesis gene cluster was successfully expressed in *S. lividans* but not *S. albus* [143]. The strain has yet unidentified proteinase(s) that are capable to efficiently process the peptide precursors of these antibiotics. At the same time the griseorhodin was not produced by *S. lividans* TK24 at all; and *S. albus* Del14 was superior to all other strains tested in this work for polyketides production. Similar, pamamycin gene cluster was successfully expressed solely in *S. albus* [195]. Interestingly, deletion of endogenous secondary metabolite gene clusters in *S. lividans*

resulted in activation of griseorhodin biosynthesis in *S. lividans* ΔYA9. This indicates that the biosynthetic precursors supply is a limiting factor preventing the efficient production of some polyketides in *S. lividans*. By reducing the competition for malonyl-CoA and redirecting it to the griseorhodin biosynthesis the production of latter compound was achieved in engineered but not parental strain. However, the possibility that endogenous and heterologous gene clusters are competing for the same regulatory elements cannot be excluded. Such situation was observed in the case of *S. albus* J1074 and its derivative Del14 [132]. The deletion of internal secondary metabolism gene clusters in *S. albus* resulted in increased transcription level of heterologous genes [132].

The largest problem faced during heterologous production of natural products is a low yield. Despite the universality of this approach, typically the production level is in few mg or even μg per liter range that is often not sufficient for identification and purification of produced compounds. In a native producing strains a vast variety of techniques and tools were developed to increase the yield of the final product. Among them optimizing production media [110], overexpressing regulator or transporter genes [111] [109], adding precursors or introducing constitutive or inducible promoters [112] [113] and also introducing additional *attB* sites to increase the copy number of the gene clusters inside the genome [115]. The often observed result of the selection for overproducers is the amplification of secondary metabolite biosynthetic gene cluster of interest. As such, kanamycin over-producing *S. kanamyceticus* 12-6 was obtained after several round of mutagenesis accompanied by selection for kanamycin resistance and screening for enhanced kanamycin production [196]. The strain was found to be a heterogenic mixture of cells each carrying different number of *kan* gene cluster in their chromosomes. Similar, during the attempt to circularize the linear chromosome of *S. albus* Del1 by deleting the right and left arms the amplification of carotenoid gene clusters at the edge of the chromosome occurred leading to the overproduction of the corresponding NP [197]. The gene dosage effect was also studied for heterologous system [115,192]. The additional *attB* sites for phiC31 based vectors integration were introduced into chromosome of *S. albus* J1074. In this case a clear correlation between gene clusters number and level of corresponding NP was observed. This observation promoted us to further engineer *S. lividans* ΔYA9 by increasing the number of *attB* sites to 2 and 3, resulting in *S. lividans* ΔYA10 and *S. lividans* ΔYA11 respectively. The influence of gene cluster amplification on production level was clearly observed in the case of griseorhodin gene cluster expression. When *S. lividans* TK24 was not able to produce this polyketide and ΔYA9 was accumulating a small quantity, two strains ΔYA10 and ΔYA11 showed gradual increased in griseorhodin accumulation. Also in the case of deoxycinnamycin *S. lividans* ΔYA10 showed the

highest production level among all tested strains. However, in some cases *S. lividans* ΔYA11 performed worse than ΔYA9 and ΔYA10. This can be explained by toxicity of the produced compound or overflow of resources into secondary metabolites production resulting in decrease in fitness of the strain.

Minimizing the host metabolites background improves detection of products from expression of heterologous gene clusters. For instance, *S. albus* Del14 was used to express two genomic libraries of *Streptomyces* and *Frankia* species resulting in isolation of pyridinopyrone (from *Streptomyces* sp.), salicylic acid, fralnimycin, bhimamycin A and aloesaponarin II (from *Frankia* species) [192]. Similar, screening for new NPs while expressing clones from genomic library of *S. albus* subsp. *chlorinus* in *S. lividans* ΔYA9 leads to identification of new members of pyrrolobenzodiazepines (PBD) group of NPs. The latter compounds are potent anticancer agents. In overall, it should be clear that there is no ideal and universal heterologous host. Such strain is basically not possible due to the particularities in metabolic pathways, physiology, genetics and overall behavior of different *Streptomyces* species. Furthermore, the expression of non-*Streptomyces* Actinobacteria genomic libraries or metagenomic DNA of unidentified origin will face additional problems due to the differences in transcriptional and translational regulatory elements. So, to our great regret, the unified universal host strain for natural products production remains an elusive and mystique unicorn. With such, it should be clear that the only solution to this situation is development of panel of strains with different preferences to different natural products production together with an efficient set of synthetic well-controllable BioBricks (delivery vectors, promoters, RBSs, terminators) for refactoring the gene cluster of interest. This will increase the success rate of screening programs for new NPs from heterologous gene clusters expression. For instance, the combined application of *S. albus* Del14 and *S. lividans* ΔYA9 for expression of *S. albus* subsp. *chlorinus* genomic library resulted in overall 35% success rate with 6 new compounds identified, when each strain separately produced only 2 (*S. lividans* ΔYA9) and 5 (*S. albus* Del14) metabolites (1 was produced by both strains).

Streptomyces produce and secrete a wide range of proteins and thus are used for protein production in industrial scale [198]. Among *Streptomyces* species, *S. lividans* has reduced extracellular protease activity and highly active secretion machinery [199]. This makes *S. lividans* strain a preferred heterologous host for the production of proteins. The strain was used for a long time as a producer of extracellular recombinant proteins [135,198]. In order to determine the role of proteases in the secretion of native and heterologous proteins we generated *S. lividans* TK24 mutants lacking a set of proteases encoding genes mutants. The protease deletion strains were

validated for heterologous protein secretion by expressing the gene encoding SP^{vsi}-mRFP. Among all tested proteases the deletion of *ftsH* gene showed significant improvement in SP^{vsi}-mRFP secretion level and also demonstrated changes in overall protein secretome. FtsH is ATP-dependent protease for both cytoplasmic and membrane proteins [65]. This membrane-embedded protease plays a role in the quality control of integral membrane proteins [200,201]. Thus, it is not surprising that FtsH is playing global regulatory role in protein secretion and manipulations with this gene can lead to improving the quality and secretion yield of target protein. There are multiple evidences that deletion of proteases encoding genes in Gram-positive bacteria effected the heterologous production of proteins [202,203]. Thus further improvement of *S. lividans* as a host for heterologous protein production can be achieved by deep investigation of proteolytic activity of the strain and factors influencing it as well as gaining an understanding of protein secretion machineries and their functional properties.

Summary

We have developed a strategy for direct selection of clones of *Streptomyces* spp. with activated expression of silent secondary metabolism gene cluster that allows also quick and simple identification of new regulatory loci controlling NPs production in these bacteria. This technique can be applied to increase the production yield of compound of interest also in heterologous conditions. On the other hand, we have generated *S. lividans* "clean" host strain for heterologous production of NPs of actinobacterial origin. This strain is completing the panel of engineered hosts for NPs discovery, potentially increasing the success rate of screening programs and contributing to diversification of obtained new compounds. Lastly, we have tested the utility of proteolytic activity modulation in *S. lividans* on its ability to produce secreted proteins. When combined, the developed approach and host will not only increase the chances for new NPs discovery but provide an opportunity to gain a new knowledge about complex regulatory events controlling secondary metabolism both in native and heterologous systems.

6. References

1 Demain, A. L. Importance of microbial natural products and the need to revitalize their discovery. *J Ind Microbiol Biotechnol* **41**, 185-201, doi:10.1007/s10295-013-1325-z (2014).

2 Giddings, L. A. & Newman, D. J. Microbial natural products: molecular blueprints for antitumor drugs. *J Ind Microbiol Biotechnol* **40**, 1181-1210, doi:10.1007/s10295-013-1331-1 (2013).

3 Newman, D. J., Cragg, G. M. & Snader, K. M. The influence of natural products upon drug discovery. *Nat Prod Rep* **17**, 215-234, doi:DOI 10.1039/a902202c (2000).

4 Newman, D. J. & Cragg, G. M. Natural Products As Sources of New Drugs over the 30 Years from 1981 to 2010. *J Nat Prod* **75**, 311-335, doi:10.1021/np200906s (2012).

5 Kinch, M. S., Haynesworth, A., Kinch, S. L. & Hoyer, D. An overview of FDA-approved new molecular entities: 1827-2013. *Drug Discov Today* **19**, 1033-1039, doi:10.1016/j.drudis.2014.03.018 (2014).

6 Katz, L. & Baltz, R. H. Natural product discovery: past, present, and future. *J Ind Microbiol Biotechnol* **43**, 155-176, doi:10.1007/s10295-015-1723-5 (2016).

7 Patridge, E., Gareiss, P., Kinch, M. S. & Hoyer, D. An analysis of FDA-approved drugs: natural products and their derivatives. *Drug Discov Today* **21**, 204-207, doi:10.1016/j.drudis.2015.01.009 (2016).

8 Fleming, A. On the antibacterial action of cultures of a penicillium, with special reference to their use in the isolation of B. influenzae. 1929. *Bull World Health Organ* **79**, 780-790 (2001).

9 Chain, E. *et al.* The Classic - Penicillin as a Chemotherapeutic Agent. *Clin Orthop Relat R*, 3-7 (1993).

10 Berdy, J. Bioactive microbial metabolites. *J Antibiot (Tokyo)* **58**, 1-26, doi:10.1038/ja.2005.1 (2005).

11 Berdy, J. Thoughts and facts about antibiotics: Where we are now and where we are heading. **65**, 385-395 (2012).

12 Li, J. W. & Vederas, J. C. Drug discovery and natural products: end of an era or an endless frontier? *Science* **325**, 161-165, doi:10.1126/science.1168243 (2009).

13 Payne, D. J., Gwynn, M. N., Holmes, D. J. & Pompliano, D. L. Drugs for bad bugs: confronting the challenges of antibacterial discovery. *Nat Rev Drug Discov* **6**, 29-40, doi:10.1038/nrd2201 (2007).

14 Hajduk, P. J. & Greer, J. A decade of fragment-based drug design: strategic advances and lessons learned. *Nat Rev Drug Discov* **6**, 211-219, doi:10.1038/nrd2220 (2007).

15 von Bubnoff, A. Seeking new antibiotics in nature's backyard. *Cell* **127**, 867-869, doi:10.1016/j.cell.2006.11.021 (2006).

16 Lewis, K. Antibiotics: Recover the lost art of drug discovery. *Nature* **485**, 439-440, doi:10.1038/485439a (2012).

17 Fowler, V. G., Jr. *et al.* Daptomycin versus standard therapy for bacteremia and endocarditis caused by Staphylococcus aureus. *N Engl J Med* **355**, 653-665, doi:10.1056/NEJMoa053783 (2006).

18 Cornely, O. A. *et al.* Fidaxomicin versus vancomycin for infection with Clostridium difficile in Europe, Canada, and the USA: a double-blind, non-inferiority, randomised controlled trial. *Lancet Infect Dis* **12**, 281-289, doi:10.1016/S1473-3099(11)70374-7 (2012).

19 Boucher, H. W. *et al.* Bad bugs, no drugs: no ESKAPE! An update from the Infectious Diseases Society of America. *Clin Infect Dis* **48**, 1-12, doi:10.1086/595011 (2009).

20 Antibiotic Resistance Threats in the United States. Antibiotic/Antimicrobial Resistance | CDC. Available at: https://www.cdc.gov/drugresistance/threat-report-2013/index.html. (2013).

21 WHO. | Antimicrobial resistance: global report on surveillance 2014. (2016).

22 Liu, Y. Y. *et al.* Emergence of plasmid-mediated colistin resistance mechanism MCR-1 in animals and human beings in China: a microbiological and molecular biological study. *Lancet Infect Dis* **16**, 161-168, doi:10.1016/S1473-3099(15)00424-7 (2016).

23 Walsh, T. R., Weeks, J., Livermore, D. M. & Toleman, M. A. Dissemination of NDM-1 positive bacteria in the New Delhi environment and its implications for human health: an environmental point prevalence study. *Lancet Infect Dis* **11**, 355-362, doi:10.1016/S1473-3099(11)70059-7 (2011).

24 Control, C. f. D. & Prevention. Threat Report 2013: Antibiotic/Antimicrobial Resistance. *Centers for Disease Control and Prevention* (2013).

25 Review on Antimicrobial Resistance. Antimicrobial resistance: tackling acrisisfor the health and wealth of nations. (2014).

26 Butler, M. S., Robertson, A. A. & Cooper, M. A. Natural product and natural product derived drugs in clinical trials. *Nat Prod Rep* **31**, 1612-1661, doi:10.1039/c4np00064a (2014).

27 Chandra, G. & Chater, K. F. Developmental biology of *Streptomyces* from the perspective of 100 actinobacterial genome sequences. *Fems Microbiol Rev* **38**, 345-379, doi:10.1111/1574-6976.12047 (2014).

28 Genilloud, O. Actinomycetes: still a source of novel antibiotics. *Nat Prod Rep* **34**, 1203-1232, doi:10.1039/c7np00026j (2017).

29 Barka, E. A. *et al.* Taxonomy, Physiology, and Natural Products of Actinobacteria. *Microbiol Mol Biol Rev* **80**, 1-43, doi:10.1128/MMBR.00019-15 (2016).

30 Schatz, A., Bugie, E. & Waksman, S. A. Streptomycin, a substance exhibiting antibiotic activity against Gram-positive and Gram-negative bacteria (vol 437, pg 3, 2005). *Clin Orthop Relat R*, 379-379 (2005).

31 Waksman, S. A., Geiger, W. B. & Reynolds, D. M. Strain Specificity and Production of Antibiotic Substances: VII. Production of Actinomycin by Different Actinomycetes. *Proc Natl Acad Sci U S A* **32**, 117-120 (1946).

32 Demain, A. L. & Sanchez, S. Microbial drug discovery: 80 years of progress. *J Antibiot (Tokyo)* **62**, 5-16, doi:10.1038/ja.2008.16 (2009).

33 Procopio, R. E., Silva, I. R., Martins, M. K., Azevedo, J. L. & Araujo, J. M. Antibiotics produced by *Streptomyces*. *Braz J Infect Dis* **16**, 466-471, doi:10.1016/j.bjid.2012.08.014 (2012).

34 Iqbal, H. A., Low-Beinart, L., Obiajulu, J. U. & Brady, S. F. Natural Product Discovery through Improved Functional Metagenomics in *Streptomyces*. *J Am Chem Soc* **138**, 9341-9344, doi:10.1021/jacs.6b02921 (2016).

35 Liu, Y. *et al.* Functional Analysis of Cytochrome P450s Involved in Streptovaricin Biosynthesis and Generation of Anti-MRSA Analogues. *Acs Chem Biol* **12**, 2589-2597, doi:10.1021/acschembio.7b00467 (2017).

36 Bentley, S. D. *et al.* Complete genome sequence of the model actinomycete *Streptomyces coelicolor* A3(2). *Nature* **417**, 141-147, doi:10.1038/417141a (2002).

37 Oliynyk, M. *et al.* Complete genome sequence of the erythromycin-producing bacterium *Saccharopolyspora erythraea* NRRL23338. *Nat Biotechnol* **25**, 447-453, doi:10.1038/nbt1297 (2007).

38 Ikeda, H. *et al.* Complete genome sequence and comparative analysis of the industrial microorganism *Streptomyces avermitilis*. *Nat Biotechnol* **21**, 526-531, doi:10.1038/nbt820 (2003).

39 Ohnishi, Y. *et al.* Genome sequence of the streptomycin producing microorganism *Streptomyces griseus* IFO 13350. *J Bacteriol* **190**, 4050-4060, doi:10.1128/JB.00204-08 (2008).

40 Ruckert, C. *et al.* Complete genome sequence of *Streptomyces lividans* TK24. *J Biotechnol* **199**, 21-22, doi:10.1016/j.jbiotec.2015.02.004 (2015).

41 Challis, G. L. & Hopwood, D. A. Synergy and contingency as driving forces for the evolution of multiple secondary metabolite production by *Streptomyces* species. *P Natl Acad Sci USA* **100**, 14555-14561, doi:10.1073/pnas.1934677100 (2003).

42 Zerikly, M. & Challis, G. L. Strategies for the discovery of new natural products by genome mining. *Chembiochem* **10**, 625-633, doi:10.1002/cbic.200800389 (2009).

43 Weber, T. *et al.* Metabolic engineering of antibiotic factories: new tools for antibiotic production in actinomycetes. *Trends Biotechnol* **33**, 15-26, doi:10.1016/j.tibtech.2014.10.009 (2015).

44 Medema, M. H. *et al.* antiSMASH: rapid identification, annotation and analysis of secondary metabolite biosynthesis gene clusters in bacterial and fungal genome sequences. *Nucleic Acids Res* **39**, W339-346, doi:10.1093/nar/gkr466 (2011).

45 Starcevic, A. *et al.* ClustScan: an integrated program package for the semi-automatic annotation of modular biosynthetic gene clusters and in silico prediction of novel chemical structures. *Nucleic Acids Res* **36**, 6882-6892, doi:10.1093/nar/gkn685 (2008).

46 Anand, S. *et al.* SBSPKS: structure based sequence analysis of polyketide synthases. *Nucleic Acids Res* **38**, W487-496, doi:10.1093/nar/gkq340 (2010).

47 Li, M. H., Ung, P. M., Zajkowski, J., Garneau-Tsodikova, S. & Sherman, D. H. Automated genome mining for natural products. *Bmc Bioinformatics* **10**, 185, doi:10.1186/1471-2105-10-185 (2009).

48 Zarins-Tutt, J. S. *et al.* Prospecting for new bacterial metabolites: a glossary of approaches for inducing, activating and upregulating the biosynthesis of bacterial cryptic or silent natural products. *Nat Prod Rep* **33**, 54-72, doi:10.1039/c5np00111k (2016).

49 Rutledge, P. J. & Challis, G. L. Discovery of microbial natural products by activation of silent biosynthetic gene clusters. *Nat Rev Microbiol* **13**, 509-523, doi:10.1038/nrmicro3496 (2015).

50 Liu, G., Chater, K. F., Chandra, G., Niu, G. & Tan, H. Molecular regulation of antibiotic biosynthesis in *streptomyces*. *Microbiol Mol Biol Rev* **77**, 112-143, doi:10.1128/MMBR.00054-12 (2013).

51 Bode, H. B., Bethe, B., Hofs, R. & Zeeck, A. Big effects from small changes: possible ways to explore nature's chemical diversity. *Chembiochem* **3**, 619-627, doi:10.1002/1439-7633(20020703)3:7<619::AID-CBIC619>3.0.CO;2-9 (2002).

52 Rateb, M. E. *et al.* Chaxamycins A-D, bioactive ansamycins from a hyper-arid desert Streptomyces sp. *J Nat Prod* **74**, 1491-1499, doi:10.1021/np200320u (2011).

53 Wu, Q. *et al.* Production and Identification of Inthomycin B Produced by a Deep-Sea Sediment-Derived Streptomyces sp. YB104 Based on Cultivation-Dependent Approach. *Curr Microbiol* **75**, 942-951, doi:10.1007/s00284-018-1469-1 (2018).

54 Takano, H., Obitsu, S., Beppu, T. & Ueda, K. Light-induced carotenogenesis in *Streptomyces coelicolor* A3(2): identification of an extracytoplasmic function sigma factor that directs photodependent transcription of the carotenoid biosynthesis gene cluster. *J Bacteriol* **187**, 1825-1832, doi:10.1128/JB.187.5.1825-1832.2005 (2005).

55 Ochi, K. & Hosaka, T. New strategies for drug discovery: activation of silent or weakly expressed microbial gene clusters. *Appl Microbiol Biotechnol* **97**, 87-98, doi:10.1007/s00253-012-4551-9 (2013).

56 Shima, J., Hesketh, A., Okamoto, S., Kawamoto, S. & Ochi, K. Induction of actinorhodin production by *rpsL* (encoding ribosomal protein S12) mutations that confer streptomycin resistance in *Streptomyces lividans* and *Streptomyces coelicolor* A3(2). *J Bacteriol* **178**, 7276-7284 (1996).

57 Hosaka, T. *et al.* Antibacterial discovery in actinomycetes strains with mutations in RNA polymerase or ribosomal protein S12. *Nat Biotechnol* **27**, 462-464, doi:10.1038/nbt.1538 (2009).

58 Zhang, Y. H. *et al.* Activation and enhancement of Fredericamycin A production in deepsea-derived Streptomyces somaliensis SCSIO ZH66 by using ribosome engineering and response surface methodology. *Microb Cell Fact* **14**, doi:ARTN 64 10.1186/s12934-015-0244-2 (2015).

59 Artsimovitch, I. *et al.* Structural basis for transcription regulation by alarmone ppGpp. *Cell* **117**, 299-310 (2004).

60 Srivatsan, A. & Wang, J. D. Control of bacterial transcription, translation and replication by (p)ppGpp. *Curr Opin Microbiol* **11**, 100-105, doi:10.1016/j.mib.2008.02.001 (2008).

61 Xu, J., Tozawa, Y., Lai, C., Hayashi, H. & Ochi, K. A rifampicin resistance mutation in the rpoB gene confers ppGpp-independent antibiotic production in *Streptomyces coelicolor* A3(2). *Mol Genet Genomics* **268**, 179-189, doi:10.1007/s00438-002-0730-1 (2002).

62 Hosaka, T., Xu, J. & Ochi, K. Increased expression of ribosome recycling factor is responsible for the enhanced protein synthesis during the late growth phase in an antibiotic-overproducing *Streptomyces coelicolor* ribosomal *rpsL* mutant. *Mol Microbiol* **61**, 883-897, doi:10.1111/j.1365-2958.2006.05285.x (2006).

63 Martin, J. F. Phosphate control of the biosynthesis of antibiotics and other secondary metabolites is mediated by the PhoR-PhoP system: an unfinished story. *J Bacteriol* **186**, 5197-5201, doi:10.1128/JB.186.16.5197-5201.2004 (2004).

64 Rigali, S. *et al.* Feast or famine: the global regulator DasR links nutrient stress to antibiotic production by *Streptomyces*. *EMBO Rep* **9**, 670-675, doi:10.1038/embor.2008.83 (2008).

65 McKenzie, N. L. *et al.* Induction of antimicrobial activities in heterologous Streptomycetes using alleles of the *Streptomyces coelicolor* gene absA1. *J Antibiot* **63**, 177-182, doi:10.1038/ja.2010.13 (2010).

66 Yin, W. B. & Keller, N. P. Transcriptional Regulatory Elements in Fungal Secondary Metabolism. *J Microbiol* **49**, 329-339, doi:10.1007/s12275-011-1009-1 (2011).

67 Guyet, A. *et al.* Identified members of the *Streptomyces lividans* AdpA regulon involved in differentiation and secondary metabolism. *BMC Microbiol* **14**, 81, doi:10.1186/1471-2180-14-81 (2014).

68 Kalan, L. *et al.* A cryptic polyene biosynthetic gene cluster in *Streptomyces calvus* is expressed upon complementation with a functional *bldA* gene. *Chem Biol* **20**, 1214-1224, doi:10.1016/j.chembiol.2013.09.006 (2013).

69 Lu, C., Liao, G., Zhang, J. & Tan, H. Identification of novel tylosin analogues generated by a wblA disruption mutant of *Streptomyces ansochromogenes*. *Microb Cell Fact* **14**, 173, doi:10.1186/s12934-015-0359-5 (2015).

70 Crack, J. C. *et al.* Mechanistic Insight into the Nitrosylation of the [4Fe-4S] Cluster of WhiB-like Proteins. *J Am Chem Soc* **133**, 1112-1121, doi:10.1021/ja109581t (2011).

71 Larsson, C. *et al.* Gene Expression of *Mycobacterium tuberculosis* Putative Transcription Factors whiB1-7 in Redox Environments. *Plos One* **7**, doi:ARTN e3751610.1371/journal.pone.0037516 (2012).

72 Tobes, R. & Ramos, J. L. AraC-XylS database: a family of positive transcriptional regulators in bacteria. *Nucleic Acids Res* **30**, 318-321 (2002).

73 Akanuma, G., Hara, H., Ohnishi, Y. & Horinouchi, S. Dynamic changes in the extracellular proteome caused by absence of a pleiotropic regulator AdpA in *Streptomyces griseus*. *Mol Microbiol* **73**, 898-912, doi:10.1111/j.1365-2958.2009.06814.x (2009).

74 Ohnishi, Y., Yamazaki, H., Kato, J. Y., Tomono, A. & Horinouchi, S. AdpA, a central transcriptional regulator in the A-factor regulatory cascade that leads to morphological development and secondary metabolism in *Streptomyces griseus*. *Biosci Biotech Bioch* **69**, 431-439 (2005).

75 Xu, J. *et al.* Activation and mechanism of a cryptic oviedomycin gene cluster via the disruption of a global regulatory gene, *adpA*, in *Streptomyces ansochromogenes*. *J Biol Chem* **292**, 19708-19720, doi:10.1074/jbc.M117.809145 (2017).

76 Pan, Y., Liu, G., Yang, H., Tian, Y. & Tan, H. The pleiotropic regulator AdpA-L directly controls the pathway-specific activator of nikkomycin biosynthesis in *Streptomyces ansochromogenes*. *Mol Microbiol* **72**, 710-723, doi:10.1111/j.1365-2958.2009.06681.x (2009).

77 Choi, S. U., Lee, C. K., Hwang, Y. I., Kinoshita, H. & Nihira, T. Cloning and functional analysis by gene disruption of a gene encoding a gamma-butyrolactone autoregulator receptor from *Kitasatospora setae*. *J Bacteriol* **186**, 3423-3430, doi:10.1128/JB.186.11.3423-3430.2004 (2004).

78 Zhou, Z. *et al.* Genome mining-directed activation of a silent angucycline biosynthetic gene cluster in *Streptomyces chattanoogensis*. *Chembiochem* **16**, 496-502, doi:10.1002/cbic.201402577 (2015).
79 Yang, K., Han, L., He, J., Wang, L. & Vining, L. C. A repressor-response regulator gene pair controlling jadomycin B production in *Streptomyces venezuelae* ISP5230. *Gene* **279**, 165-173 (2001).
80 Rebets, Y. *et al.* Function of lanI in regulation of landomycin A biosynthesis in *Streptomyces cyanogenus* S136 and cross-complementation studies with *Streptomyces* antibiotic regulatory proteins encoding genes. *Arch Microbiol* **189**, 111-120, doi:10.1007/s00203-007-0299-5 (2008).
81 Xie, C., Deng, J. J. & Wang, H. X. Identification of AstG1, A LAL family regulator that positively controls ansatrienins production in *Streptomyces* sp. XZQH13. *Curr Microbiol* **70**, 859-864, doi:10.1007/s00284-015-0798-6 (2015).
82 Laureti, L. *et al.* Identification of a bioactive 51-membered macrolide complex by activation of a silent polyketide synthase in *Streptomyces ambofaciens*. *Proc Natl Acad Sci U S A* **108**, 6258-6263, doi:10.1073/pnas.1019077108 (2011).
83 Li, S. *et al.* Activating a Cryptic Ansamycin Biosynthetic Gene Cluster To Produce Three New Naphthalenic Octaketide Ansamycins with n-Pentyl and n-Butyl Side Chains. *Org Lett* **17**, 3706-3709, doi:10.1021/acs.orglett.5b01686 (2015).
84 O'Rourke, S. *et al.* Extracellular signalling, translational control, two repressors and an activator all contribute to the regulation of methylenomycin production in *Streptomyces coelicolor*. *Mol Microbiol* **71**, 763-778, doi:10.1111/j.1365-2958.2008.06560.x (2009).
85 Onaka, H. *et al.* Cloning and characterization of the A-factor receptor gene from *Streptomyces griseus*. *J Bacteriol* **177**, 6083-6092 (1995).
86 Takano, E. Gamma-butyrolactones: *Streptomyces* signalling molecules regulating antibiotic production and differentiation. *Curr Opin Microbiol* **9**, 287-294, doi:10.1016/j.mib.2006.04.003 (2006).
87 Takano, E., Chakraburtty, R., Nihira, T., Yamada, Y. & Bibb, M. J. A complex role for the gamma-butyrolactone SCB1 in regulating antibiotic production in *Streptomyces coelicolor* A3(2). *Mol Microbiol* **41**, 1015-1028 (2001).
88 Gottelt, M., Kol, S., Gomez-Escribano, J. P., Bibb, M. & Takano, E. Deletion of a regulatory gene within the cpk gene cluster reveals novel antibacterial activity in *Streptomyces coelicolor* A3(2). *Microbiol-Sgm* **156**, 2343-2353, doi:10.1099/mic.0.038281-0 (2010).
89 Gomez-Escribano, J. P. *et al.* Structure and biosynthesis of the unusual polyketide alkaloid coelimycin P1, a metabolic product of the cpk gene cluster of *Streptomyces coelicolor* M145. *Chem Sci* **3**, 2716-2720, doi:10.1039/c2sc20410j (2012).
90 Sidda, J. D. *et al.* Discovery of a family of gamma-aminobutyrate ureas via rational derepression of a silent bacterial gene cluster. *Chem Sci* **5**, 86-89, doi:10.1039/c3sc52536h (2014).
91 Shao, Z. *et al.* Refactoring the silent spectinabilin gene cluster using a plug-and-play scaffold. *ACS Synth Biol* **2**, 662-669, doi:10.1021/sb400058n (2013).
92 Blodgett, J. A. *et al.* Common biosynthetic origins for polycyclic tetramate macrolactams from phylogenetically diverse bacteria. *Proc Natl Acad Sci U S A* **107**, 11692-11697, doi:10.1073/pnas.1001513107 (2010).
93 Zhang, G., Zhang, W., Saha, S. & Zhang, C. Recent Advances in Discovery, Biosynthesis and Genome Mining of Medicinally Relevant Polycyclic Tetramate Macrolactams. *Curr Top Med Chem* **16**, 1727-1739 (2016).
94 Olano, C. *et al.* Activation and identification of five clusters for secondary metabolites in *Streptomyces albus* J1074. *Microb Biotechnol* **7**, 242-256, doi:10.1111/1751-7915.12116 (2014).
95 Saha, S. *et al.* Activation and characterization of a cryptic gene cluster reveals a cyclization cascade for polycyclic tetramate macrolactams. *Chem Sci* **8**, 1607-1612, doi:10.1039/c6sc03875a (2017).

96 Luo, Y. *et al.* Activation and characterization of a cryptic polycyclic tetramate macrolactam biosynthetic gene cluster. *Nat Commun* **4**, 2894, doi:10.1038/ncomms3894 (2013).
97 Sun, Y. Q. *et al.* Development of a Biosensor Concept to Detect the Production of Cluster-Specific Secondary Metabolites. *ACS Synth Biol* **6**, 1026-1033, doi:10.1021/acssynbio.6b00353 (2017).
98 Guo, F. *et al.* Targeted activation of silent natural product biosynthesis pathways by reporter-guided mutant selection. *Metab Eng* **28**, 134-142, doi:10.1016/j.ymben.2014.12.006 (2015).
99 Xiang, S. H. *et al.* Application of a double-reporter-guided mutant selection method to improve clavulanic acid production in *Streptomyces clavuligerus*. *Metab Eng* **11**, 310-318, doi:10.1016/j.ymben.2009.06.003 (2009).
100 Yamanaka, K. *et al.* Direct cloning and refactoring of a silent lipopeptide biosynthetic gene cluster yields the antibiotic taromycin A. *Proc Natl Acad Sci U S A* **111**, 1957-1962, doi:10.1073/pnas.1319584111 (2014).
101 Gibson, D. G. *et al.* Enzymatic assembly of DNA molecules up to several hundred kilobases. *Nat Methods* **6**, 343-345, doi:10.1038/nmeth.1318 (2009).
102 Gust, B. Chapter 7. Cloning and analysis of natural product pathways. *Methods Enzymol* **458**, 159-180, doi:10.1016/S0076-6879(09)04807-1 (2009).
103 Sosio, M. *et al.* Artificial chromosomes for antibiotic-producing actinomycetes. *Nat Biotechnol* **18**, 343-345, doi:10.1038/73810 (2000).
104 Pait, I. G. U. *et al.* Discovery of a new diol-containing polyketide by heterologous expression of a silent biosynthetic gene cluster from *Streptomyces lavendulae* FRI-5. *J Ind Microbiol Biotechnol* **45**, 77-87, doi:10.1007/s10295-017-1997-x (2018).
105 Gomez-Escribano, J. P. & Bibb, M. J. Heterologous expression of natural product biosynthetic gene clusters in *Streptomyces coelicolor*: from genome mining to manipulation of biosynthetic pathways. *J Ind Microbiol Biotechnol* **41**, 425-431, doi:10.1007/s10295-013-1348-5 (2014).
106 Young, T. S., Dorrestein, P. C. & Walsh, C. T. Codon randomization for rapid exploration of chemical space in thiopeptide antibiotic variants. *Chem Biol* **19**, 1600-1610, doi:10.1016/j.chembiol.2012.10.013 (2012).
107 Du, D. *et al.* Improvement of gougerotin and nikkomycin production by engineering their biosynthetic gene clusters. *Appl Microbiol Biot* **97**, 6383-6396, doi:10.1007/s00253-013-4836-7 (2013).
108 Binz, T. M., Wenzel, S. C., Schnell, H. J., Bechthold, A. & Muller, R. Heterologous expression and genetic engineering of the phenalinolactone biosynthetic gene cluster by using red/ET recombineering. *Chembiochem* **9**, 447-454, doi:10.1002/cbic.200700549 (2008).
109 Huo, L., Rachid, S., Stadler, M., Wenzel, S. C. & Muller, R. Synthetic biotechnology to study and engineer ribosomal bottromycin biosynthesis. *Chem Biol* **19**, 1278-1287, doi:10.1016/j.chembiol.2012.08.013 (2012).
110 Julia, P. *et al.* Heterologous production of daptomycin in Streptomyces lividans. *J Ind Microbiol Biot* **33**, 121-128, doi:10.1007/s10295-005-0033-8 (2006).
111 Dangel, V., Eustaquio, A. S., Gust, B. & Heide, L. novE and novG act as positive regulators of novobiocin biosynthesis. *Arch Microbiol* **190**, 509-519, doi:10.1007/s00203-008-0396-0 (2008).
112 Rebets, Y. *et al.* Insights into the pamamycin biosynthesis. *Angew. Chem. Int. Ed.* **54**, 2280-2284, doi:10.1002/anie.201408901 (2015).
113 Du, D. *et al.* Improvement of gougerotin and nikkomycin production by engineering their biosynthetic gene clusters. *Appl Microbiol Biotechnol* **97**, 6383-6396, doi:10.1007/s00253-013-4836-7 (2013).
114 Komatsu, M. *et al.* Engineered *Streptomyces avermitilis* host for heterologous expression of biosynthetic gene cluster for secondary metabolites. *ACS Synth Biol* **2**, 384-396, doi:10.1021/sb3001003 (2013).

115 Manderscheid, N. *et al.* An influence of the copy number of biosynthetic gene clusters on the production level of antibiotics in a heterologous host. *J Biotechnol* **232**, 110-117, doi:10.1016/j.jbiotec.2016.05.038 (2016).
116 Komatsu, M., Uchiyama, T., Omura, S., Cane, D. E. & Ikeda, H. Genome-minimized *Streptomyces* host for the heterologous expression of secondary metabolism. *Proc Natl Acad Sci U S A* **107**, 2646-2651, doi:10.1073/pnas.0914833107 (2010).
117 Gomez-Escribano, J. P. & Bibb, M. J. Engineering *Streptomyces coelicolor* for heterologous expression of secondary metabolite gene clusters. *Microb Biotechnol* **4**, 207-215, doi:10.1111/j.1751-7915.2010.00219.x (2011).
118 Gust, B. *et al.* Lambda red-mediated genetic manipulation of antibiotic producing *Streptomyces*. *Adv Appl Microbiol* **54**, 107-128, doi:10.1016/S0065-2164(04)54004-2 (2004).
119 T. Kieser, M. J. B., M. J. Buttner, K. F. Chater, D. A. Hopwood. . Practical *Streptomyces* Genetics. *International Microbiology* **3**, 260-261 (2000).
120 McDaniel, R., Ebertkhosla, S., Hopwood, D. A. & Khosla, C. Engineered Biosynthesis of Novel Polyketides. *Science* **262**, 1546-1550, doi:DOI 10.1126/science.8248802 (1993).
121 Thanapipatsiri, A., Claesen, J., Gomez-Escribano, J. P., Bibb, M. & Thamchaipenet, A. A *Streptomyces coelicolor* host for the heterologous expression of Type III polyketide synthase genes. *Microb Cell Fact* **14**, 145, doi:10.1186/s12934-015-0335-0 (2015).
122 Ikeda, H., Kazuo, S. Y. & Omura, S. Genome mining of the *Streptomyces avermitilis* genome and development of genome-minimized hosts for heterologous expression of biosynthetic gene clusters. *J Ind Microbiol Biotechnol* **41**, 233-250, doi:10.1007/s10295-013-1327-x (2014).
123 Chater, K. F. & Wilde, L. C. *Streptomyces albus* G Mutants Defective in the *Sal*GI Restriction-Modification System. *J Gen Microbiol* **116**, 323-334 (1980).
124 Chater, K. F. & Wilde, L. C. Restriction of a bacteriophage of *Streptomyces albus* G involving endonuclease *Sal*I. *J Bacteriol* **128**, 644-650 (1976).
125 Zaburannyi, N., Rabyk, M., Ostash, B., Fedorenko, V. & Luzhetskyy, A. Insights into naturally minimised *Streptomyces albus* J1074 genome. *Bmc Genomics* **15**, doi:Artn 9710.1186/1471-2164-15-97 (2014).
126 Fernandez, E., Lombo, F., Mendez, C. & Salas, J. A. An ABC transporter is essential for resistance to the antitumor agent mithramycin in the producer *Streptomyces argillaceus*. *Mol Gen Genet* **251**, 692-698, doi:DOI 10.1007/s004380050218 (1996).
127 Baltz, R. H. *Streptomyces* and *Saccharopolyspora* hosts for heterologous expression of secondary metabolite gene clusters. *J Ind Microbiol Biotechnol* **37**, 759-772, doi:10.1007/s10295-010-0730-9 (2010).
128 Makitrynskyy, R. *et al.* Genetic factors that influence moenomycin production in Streptomycetes. *J Ind Microbiol Biot* **37**, 559-566, doi:10.1007/s10295-010-0701-1 (2010).
129 Lombo, F. *et al.* Deciphering the biosynthesis pathway of the antitumor thiocoraline from a marine *actinomycete* and its expression in two *streptomyces* species. *Chembiochem* **7**, 366-376, doi:10.1002/cbic.200500325 (2006).
130 Kallifidas, D., Kang, H. S. & Brady, S. F. Tetarimycin A, an MRSA-active antibiotic identified through induced expression of environmental DNA gene clusters. *J Am Chem Soc* **134**, 19552-19555, doi:10.1021/ja3093828 (2012).
131 Owen, J. G. *et al.* Multiplexed metagenome mining using short DNA sequence tags facilitates targeted discovery of epoxyketone proteasome inhibitors. *Proc Natl Acad Sci U S A* **112**, 4221-4226, doi:10.1073/pnas.1501124112 (2015).
132 Myronovskyi, M. *et al.* Generation of a cluster-free *Streptomyces albus* chassis strains for improved heterologous expression of secondary metabolite clusters. *Metab Eng*, doi:10.1016/j.ymben.2018.09.004 (2018).

133 Barkei, J. J., Kevany, B. M., Felnagle, E. A. & Thomas, M. G. Investigations into Viomycin Biosynthesis by Using Heterologous Production in *Streptomyces lividans*. *Chembiochem* **10**, 366-376, doi:10.1002/cbic.200800646 (2009).

134 Shao, L. *et al.* Efficient Bioconversion of Echinocandin B to Its Nucleus by Overexpression of Deacylase Genes in Different Host Strains. *Appl Environ Microb* **79**, 1126-1133, doi:10.1128/Aem.02792-12 (2013).

135 Anne, J., Vrancken, K., Van Mellaert, L., Van Impe, J. & Bernaerts, K. Protein secretion biotechnology in Gram-positive bacteria with special emphasis on *Streptomyces lividans*. *Bba-Mol Cell Res* **1843**, 1750-1761, doi:10.1016/j.bbamcr.2013.12.023 (2014).

136 Eustaquio, A. S. *et al.* Heterologous expression of novobiocin and clorobiocin biosynthetic gene clusters. *Appl Environ Microb* **71**, 2452-2459, doi:10.1128/Aem.71.5.2452-2459.2005 (2005).

137 Voller, G. H. *et al.* Characterization of new class III lantibiotics--erythreapeptin, avermipeptin and griseopeptin from *Saccharopolyspora erythraea*, *Streptomyces avermitilis* and *Streptomyces griseus* demonstrates stepwise N-terminal leader processing. *Chembiochem* **13**, 1174-1183, doi:10.1002/cbic.201200118 (2012).

138 Penn, J. *et al.* Heterologous production of daptomycin in *Streptomyces lividans*. *J Ind Microbiol Biotechnol* **33**, 121-128, doi:10.1007/s10295-005-0033-8 (2006).

139 Ziermann, R. & Betlach, M. C. Recombinant polyketide synthesis in *Streptomyces*: Engineering of improved host strains. *Biotechniques* **26**, 106-+ (1999).

140 Martinez, A. *et al.* Genetically modified bacterial strains and novel bacterial artificial chromosome shuttle vectors for constructing environmental libraries and detecting heterologous natural products in multiple expression hosts. *Appl Environ Microbiol* **70**, 2452-2463 (2004).

141 Paulus, C. *et al.* New Alpiniamides From *Streptomyces* sp. IB2014/011-12 Assembled by an Unusual Hybrid Non-ribosomal Peptide Synthetase Trans-AT Polyketide Synthase Enzyme. *Front Microbiol* **9**, 1959, doi:10.3389/fmicb.2018.01959 (2018).

142 Alt, S. & Wilkinson, B. Biosynthesis of the Novel Macrolide Antibiotic Anthracimycin. *ACS Chem Biol* **10**, 2468-2479, doi:10.1021/acschembio.5b00525 (2015).

143 Horbal, L., Marques, F., Nadmid, S., Mendes, M. V. & Luzhetskyy, A. Secondary metabolites overproduction through transcriptional gene cluster refactoring. *Metabolic Engineering* **49**, 299-315, doi:10.1016/j.ymben.2018.09.010 (2018).

144 Li, Q. *et al.* Identification of the Biosynthetic Gene Cluster for the Anti-infective Desotamides and Production of a New Analogue in a Heterologous Host. *J Nat Prod* **78**, 944-948, doi:10.1021/acs.jnatprod.5b00009 (2015).

145 Bonet, B., Teufel, R., Crusemann, M., Ziemert, N. & Moore, B. S. Direct capture and heterologous expression of Salinispora natural product genes for the biosynthesis of enterocin. *J Nat Prod* **78**, 539-542, doi:10.1021/np500664q (2015).

146 Wendt-Pienkowski, E. *et al.* Cloning, sequencing, analysis, and heterologous expression of the fredericamycin biosynthetic gene cluster from Streptomyces griseus. *J Am Chem Soc* **127**, 16442-16452, doi:10.1021/ja054376u (2005).

147 Chen, Y., Wendt-Pienkowski, E. & Shen, B. Identification and utility of FdmR1 as a Streptomyces antibiotic regulatory protein activator for fredericamycin production in Streptomyces griseus ATCC 49344 and heterologous hosts. *Journal of Bacteriology* **190**, 5587-5596, doi:10.1128/Jb.00592-08 (2008).

148 Feng, Z. *et al.* Engineered production of iso-migrastatin in heterologous *Streptomyces* hosts. *Bioorg Med Chem* **17**, 2147-2153, doi:10.1016/j.bmc.2008.10.074 (2009).

149 Iftime, D. *et al.* Streptocollin, a Type IV Lanthipeptide Produced by *Streptomyces collinus* Tu 365. *Chembiochem* **16**, 2615-2623, doi:10.1002/cbic.201500377 (2015).

150 Bibb, M. J. Regulation of secondary metabolism in Streptomycetes. *Curr Opin Microbiol* **8**, 208-215, doi:10.1016/j.mib.2005.02.016 (2005).
151 Challis, G. L. Genome mining for novel natural product discovery. *J Med Chem* **51**, 2618-2628, doi:10.1021/jm700948z (2008).
152 Ling, L. L. *et al.* A new antibiotic kills pathogens without detectable resistance. *Nature* **517**, 455-+, doi:10.1038/nature14098 (2015).
153 Askenazi, M. *et al.* Integrating transcriptional and metabolite profiles to direct the engineering of lovastatin-producing fungal strains. *Nat Biotechnol* **21**, 150-156, doi:10.1038/nbt781 (2003).
154 Cadenas, R. F., Martin, J. F. & Gil, J. A. Construction and Characterization of Promoter-Probe Vectors for *Corynebacteria* Using the Kanamycin-Resistance Reporter Gene. *Gene* **98**, 117-121, doi:Doi 10.1016/0378-1119(91)90113-P (1991).
155 Han, L., Khetan, A., Hu, W. S. & Sherman, D. H. Time-lapsed confocal microscopy reveals temporal and spatial expression of the lysine epsilon-aminotransferase gene in *Streptomyces clavuligerus*. *Mol Microbiol* **34**, 878-886, doi:DOI 10.1046/j.1365-2958.1999.01638.x (1999).
156 Myronovskyi, M., Welle, E., Fedorenko, V. & Luzhetskyy, A. Beta-glucuronidase as a sensitive and versatile reporter in actinomycetes. *Appl Environ Microbiol* **77**, 5370-5383, doi:10.1128/AEM.00434-11 (2011).
157 Guo, F. *et al.* Targeted activation of silent natural product biosynthesis pathways by reporter-guided mutant selection. *Metabolic Engineering* **28**, 134-142, doi:10.1016/j.ymben.2014.12.006 (2015).
158 Rebets, Y. *et al.* Design, development and application of whole-cell based antibiotic-specific biosensor. *Metab Eng* **47**, 263-270, doi:10.1016/j.ymben.2018.03.019 (2018).
159 Xu, L., Wu, P., Wright, S. J., Du, L. & Wei, X. Bioactive Polycyclic Tetramate Macrolactams from *Lysobacter enzymogenes* and Their Absolute Configurations by Theoretical ECD Calculations. *J Nat Prod* **78**, 1841-1847, doi:10.1021/acs.jnatprod.5b00099 (2015).
160 Li, Y. *et al.* Biosynthesis of the Polycyclic System in the Antifungal HSAF and Analogues from *Lysobacter enzymogenes*. *Angew Chem Int Ed Engl* **57**, 6221-6225, doi:10.1002/anie.201802488 (2018).
161 Xu, Z. *et al.* Large-Scale Transposition Mutagenesis of *Streptomyces coelicolor* Identifies Hundreds of Genes Influencing Antibiotic Biosynthesis. *Appl Environ Microbiol* **83**, doi:10.1128/AEM.02889-16 (2017).
162 Horbal, L., Fedorenko, V., Bechthold, A. & Luzhetskyy, A. A transposon-based strategy to identify the regulatory gene network responsible for landomycin E biosynthesis. *FEMS Microbiol Lett* **342**, 138-146, doi:10.1111/1574-6968.12117 (2013).
163 Petzke, L. & Luzhetskyy, A. In vivo Tn5-based transposon mutagenesis of Streptomycetes. *Appl Microbiol Biotechnol* **83**, 979-986, doi:10.1007/s00253-009-2047-z (2009).
164 Sprusansky, O., Zhou, L., Jordan, S., White, J. & Westpheling, J. Identification of three new genes involved in morphogenesis and antibiotic production in *Streptomyces coelicolor*. *J Bacteriol* **185**, 6147-6157 (2003).
165 Bilyk, B. *et al.* In vivo random mutagenesis of Streptomycetes using mariner-based transposon *Himar1*. *Appl Microbiol Biot* **97**, 351-359, doi:10.1007/s00253-012-4550-x (2013).
166 Lampe, D. J., Grant, T. E. & Robertson, H. M. Factors affecting transposition of the Himar1 mariner transposon in vitro. *Genetics* **149**, 179-187 (1998).
167 Goryshin, I. Y., Miller, J. A., Kil, Y. V., Lanzov, V. A. & Reznikoff, W. S. Tn5/IS50 target recognition. *Proc Natl Acad Sci U S A* **95**, 10716-10721 (1998).
168 Ding, Q. & Tan, K. S. *Himar1* Transposon for Efficient Random Mutagenesis in *Aggregatibacter actinomycetemcomitans*. *Front Microbiol* **8**, 1842, doi:10.3389/fmicb.2017.01842 (2017).
169 Huang, J. *et al.* Cross-regulation among disparate antibiotic biosynthetic pathways of *Streptomyces coelicolor*. *Mol Microbiol* **58**, 1276-1287, doi:10.1111/j.1365-2958.2005.04879.x (2005).

170 de la Fuente, A., Lorenzana, L. M., Martin, J. F. & Liras, P. Mutants of *Streptomyces clavuligerus* with disruptions in different genes for clavulanic acid biosynthesis produce large amounts of holomycin: possible cross-regulation of two unrelated secondary metabolic pathways. *J Bacteriol* **184**, 6559-6565 (2002).
171 Haeder, S., Wirth, R., Herz, H. & Spiteller, D. Candicidin-producing Streptomyces support leaf-cutting ants to protect their fungus garden against the pathogenic fungus Escovopsis. *Proc Natl Acad Sci U S A* **106**, 4742-4746, doi:10.1073/pnas.0812082106 (2009).
172 Jorgensen, H. *et al.* Candicidin biosynthesis gene cluster is widely distributed among Streptomyces spp. isolated from the sediments and the neuston layer of the Trondheim fjord, Norway. *Appl Environ Microbiol* **75**, 3296-3303, doi:10.1128/AEM.02730-08 (2009).
173 Seipke, R. F. *et al.* A single Streptomyces symbiont makes multiple antifungals to support the fungus farming ant Acromyrmex octospinosus. *PLoS One* **6**, e22028, doi:10.1371/journal.pone.0022028 (2011).
174 Yan, Y. *et al.* Biosynthetic pathway for high structural diversity of a common dilactone core in antimycin production. *Org Lett* **14**, 4142-4145, doi:10.1021/ol301785x (2012).
175 Xu, F., Nazari, B., Moon, K., Bushin, L. B. & Seyedsayamdost, M. R. Discovery of a Cryptic Antifungal Compound from *Streptomyces albus* J1074 Using High-Throughput Elicitor Screens. *J Am Chem Soc* **139**, 9203-9212, doi:10.1021/jacs.7b02716 (2017).
176 Waksman, S. A. H., Elizabeth S.; Welsch, Maurice; Woodruff, H. Boyd. Distribution of Antagonistic Actinomycetes in Nature. *Soil Science* **54**, 281-296 (1942).
177 Mukku, V. J., Speitling, M., Laatsch, H. & Helmke, E. New butenolides from two marine Streptomycetes. *J Nat Prod* **63**, 1570-1572 (2000).
178 Kitani, S. *et al.* Avenolide, a *Streptomyces* hormone controlling antibiotic production in *Streptomyces avermitilis*. *Proc Natl Acad Sci U S A* **108**, 16410-16415, doi:10.1073/pnas.1113908108 (2011).
179 Nguyen, T. B., Kitani, S., Shimma, S. & Nihira, T. Butenolides from *Streptomyces albus* J1074 Act as External Signals To Stimulate Avermectin Production in *Streptomyces avermitilis*. *Appl Environ Microbiol* **84**, doi:10.1128/AEM.02791-17 (2018).
180 Thao, N. B., Kitani, S., Nitta, H., Tomioka, T. & Nihira, T. Discovering potential *Streptomyces* hormone producers by using disruptants of essential biosynthetic genes as indicator strains. *J Antibiot (Tokyo)* **70**, 1004-1008, doi:10.1038/ja.2017.85 (2017).
181 Wenzel, S. C. & Muller, R. The impact of genomics on the exploitation of the myxobacterial secondary metabolome. *Nat Prod Rep* **26**, 1385-1407, doi:10.1039/b817073h (2009).
182 Zhang, M. M., Wang, Y., Ang, E. L. & Zhao, H. Engineering microbial hosts for production of bacterial natural products. *Nat Prod Rep* **33**, 963-987, doi:10.1039/c6np00017g (2016).
183 Kim, E. J., Yang, I. & Yoon, Y. J. Developing *Streptomyces venezuelae* as a cell factory for the production of small molecules used in drug discovery. *Arch Pharm Res* **38**, 1606-1616, doi:10.1007/s12272-015-0638-z (2015).
184 Zhao, Z. L. *et al.* Hybrubins: Bipyrrole Tetramic Acids Obtained by Crosstalk between a Truncated Undecylprodigiosin Pathway and Heterologous Tetramic Acid Biosynthetic Genes. *Org Lett* **18**, 572-575, doi:10.1021/acs.orglett.5b03609 (2016).
185 Xiong, W., Liang, Y. X. & Zheng, Y. H. Enhancement and selective production of oligomycin through inactivation of avermectin's starter unit in *Streptomyces avermitilis*. *Biotechnol Lett* **28**, 911-916, doi:10.1007/s10529-006-9012-z (2006).
186 Huo, L. *et al.* Heterologous expression of bacterial natural product biosynthetic pathways. *Nat Prod Rep*, doi:10.1039/c8np00091c (2019).

187 Anne, J., Vrancken, K., Van Mellaert, L., Van Impe, J. & Bernaerts, K. Protein secretion biotechnology in Gram-positive bacteria with special emphasis on *Streptomyces lividans*. *Biochim Biophys Acta* **1843**, 1750-1761, doi:10.1016/j.bbamcr.2013.12.023 (2014).
188 Yu, T. W. & Hopwood, D. A. Ectopic Expression of the *Streptomyces coelicolor Whie* Genes for Polyketide Spore Pigment Synthesis and Their Interaction with the Act Genes for Actinorhodin Biosynthesis. *Microbiol-Uk* **141**, 2779-2791, doi:Doi 10.1099/13500872-141-11-2779 (1995).
189 Poralla, K., Muth, G. & Hartner, T. Hopanoids are formed during transition from substrate to aerial hyphae in *Streptomyces coelicolor* A3(2). *FEMS Microbiol Lett* **189**, 93-95 (2000).
190 Seipke, R. & Loria, R. Hopanoids are not essential for growth of *Streptomyces scabies* 87-22. *J Bacteriol* **191**, 5216-5223 (2009).
191 Myronovskyi, M., Rosenkranzer, B. & Luzhetskyy, A. Iterative marker excision system. *Appl Microbiol Biotechnol* **98**, 4557-4570, doi:10.1007/s00253-014-5523-z (2014).
192 Myronovskyi, M. *et al.* Generation of a cluster-free *Streptomyces albus* chassis strains for improved heterologous expression of secondary metabolite clusters. *Metab Eng* **49**, 316-324, doi:10.1016/j.ymben.2018.09.004 (2018).
193 Krawczyk, J. M. *et al.* Heterologous expression and engineering studies of labyrinthopeptins, class III lantibiotics from *Actinomadura namibiensis*. *Chem Biol* **20**, 111-122, doi:10.1016/j.chembiol.2012.10.023 (2013).
194 Zhang, Y., Chen, M. Y., Bruner, S. D. & Ding, Y. S. Heterologous Production of Microbial Ribosomally Synthesized and Post-translationally Modified Peptides. *Frontiers in Microbiology* **9**, doi:ARTN 180110.3389/fmicb.2018.01801 (2018).
195 Rebets, Y. *et al.* Insights into the pamamycin biosynthesis. *Angew Chem Int Ed Engl* **54**, 2280-2284, doi:10.1002/anie.201408901 (2015).
196 Yanai, K., Murakami, T. & Bibb, M. Amplification of the entire kanamycin biosynthetic gene cluster during empirical strain improvement of *Streptomyces kanamyceticus*. *Proc Natl Acad Sci U S A* **103**, 9661-9666, doi:10.1073/pnas.0603251103 (2006).
197 Myronovskyi, M. *et al.* Genome rearrangements of *Streptomyces albus* J1074 lead to the carotenoid gene cluster activation. *Appl Microbiol Biotechnol* **98**, 795-806, doi:10.1007/s00253-013-5440-6 (2014).
198 Anne, J., Maldonado, B., Van Impe, J., Van Mellaert, L. & Bernaerts, K. Recombinant protein production and Streptomycetes. *Journal of Biotechnology* **158**, 159-167, doi:10.1016/j.jbiotec.2011.06.028 (2012).
199 Anne, J., Economou, A. & Bernaerts, K. Protein Secretion in Gram-Positive Bacteria: From Multiple Pathways to Biotechnology. *Curr Top Microbiol Immunol* **404**, 267-308, doi:10.1007/82_2016_49 (2017).
200 Ludke, A., Kramer, R., Burkovski, A., Schluesener, D. & Poetsch, A. A proteomic study of *Corynebacterium glutamicum* AAA+ protease FtsH. *BMC Microbiol* **7**, 6, doi:10.1186/1471-2180-7-6 (2007).
201 Gottesman, S., Wickner, S. & Maurizi, M. R. Protein quality control: triage by chaperones and proteases. *Genes Dev* **11**, 815-823 (1997).
202 Westers, L., Westers, H. & Quax, W. J. *Bacillus subtilis* as cell factory for pharmaceutical proteins: a biotechnological approach to optimize the host organism. *Bba-Mol Cell Res* **1694**, 299-310, doi:10.1016/j.bbammcr.2004.02.011 (2004).
203 Krishnappa, L. *et al.* Extracytoplasmic Proteases Determining the Cleavage and Release of Secreted Proteins, Lipoproteins, and Membrane Proteins in *Bacillus subtilis*. *Journal of Proteome Research* **12**,4101-4110,doi:10.1021/pr400433h(2013).

Author's Contribution in the Work Presented in this Thesis

Chapter 2

The author carried out all genetic manipulation experiments. The author performed the transposon mutagenesis and screening for clones with activated gene cluster. Further the author generated Transcriptional fusion of *XNR_2339* and *XNR_3174* promoter to the *gusA* gene. The author performed the metabolic extraction and analysis as well as the purification of the compounds. The author performed the introduction of heterologous gene clusters into the wild type and mutant, cultivation in production medium and analysis of the expression of heterologous secondary metabolism gene clusters.

Chapter 3

The author deleted the markers from protease mutants and introduced the mRFP construct into their chromosome.

Chapter 4

The author carried out the generation of *S. lividans* chassis strains. The author performed the growth study of the engineered strains and the wild type. Further the author performed the validation of engineered *S. lividans* strains in producing heterologous compounds. The author performed the metabolic extraction and analysis as well as the purification of the compounds 4.