DEVELOPMENT AND APPLICATION OF GENETIC ENGINEERING METHODS FOR *ACTINOPLANES* SP. SE50/110

A Dissertation

Submitted to the Faculty of Biology

of

Bielefeld University

by

Tetiana Gren

In Partial Fulfillment of the

Requirements for the Degree

of

Doctor of Natural Sciences (Dr. rer. nat)

May 2017

Bielefeld University

Bielefeld, Germany

Primary Supervisor:

Prof. Dr. Jörn Kalinowski

Center for Biotechnology (CeBiTec)

Research group "Microbial Genomics and Biotechnology"

Bielefeld University

Abstract

The α-glucosidase inhibitor acarbose is used for treatment of diabetes mellitus type 2, and is manufactured industrially with overproducing derivatives of *Actinoplanes* sp. SE50/110. This strain was reportedly optimized through step-by-step conventional mutagenesis procedures in the past, however this strategy seems to reach its limits by now. Despite of high industrial significance, only limited information exists regarding acarbose metabolism, function and regulation of these processes, due to the absence of proper genetic engineering methods and tools developed for this strain. In this work, a full toolkit and set of methods for genetic engineering of *Actinoplanes* sp. SE50/110 were developed. A standardized protocol for a DNA transfer through *E. coli* - *Actinoplanes* conjugation was adjusted and applied for the transfer of φC31, φBT1 and VWB actinophage-based integrative vectors and pSG5-based replicative vector. Integration sites, occurring once per genome for all integrative vectors, were sequenced and characterized for the first time in *Actinoplanes* sp. SE50/110. Notably, the studied plasmids were proven to be stable and neutral with respect to strain morphology and acarbose production, enabling future use for genetic manipulations of *Actinoplanes* sp. SE50/110. To further broaden the spectrum of available tools, a GUS reporter system, was established in *Actinoplanes* sp. SE50/110. The set of different methods for gene knockouts was tested, which included integrative and replicative vector based knockouts, ReDirect system based knockouts and CRISPR-Cas9 genetic engineering. ReDirect system was further used to create a library of *Actinoplanes* single knockout strains. Two of the strains, *Actinoplanes* $\Delta acbD$ and *Actinoplanes* $\Delta cadC$ were further characterized in detail regarding their phenotype.

The developed gene cloning system offers multiple possibilities to solve fundamental questions regarding acarbose production, in particular, formulation and verification of the complete acarbose metabolism model, as well as the rational design of acarbose overproducing strains.

TABLE OF CONTENTS

Page

Page

LIST OF TABLES

LIST OF FIGURES

Figure Page

ABBREVIATIONS

NP	Natural Product
NMR	Nuclear magnetic resonance
MS	Mass spectrometry
DM	Diabetes mellitus
T1DM	Type 1 diabetes mellitus
T2DM	Type 2 diabetes mellitus
CM	Complete medium
MM	Minimal medium
SFM	Soya flour medium
OM	Oatmeal medium
LB	Luria-Bertani medium
Mal-MM	Maltose minimal medium
Glu-MM	Glucose minimal medium
MT-MM	Maltotriose minimal medium
Cpur-MM	Cpur minimal medium
Mal-MM-TE	Maltose minimal medium with trace elements
Glu-CM	Glucose complex medium
CeBiTec	Center of Biotechnology, Bielefeld University
GUS	reporter system, based on the activity of β-glucuronidase enzyme
RPM	Rounds per minute
CDS	Protein-coding sequences
DNA	Deoxyribonucleic acid
RNA	Ribonucleic acid
BP	Base pair

cDNA	Copy-DNA
ncRNA	Non-coding RNA
SEM	Scanning electron microscope
HPLC	High performance liquid chromatography
ENA	European Nucleotide Archive
IFNUL	Ivan Franko National University of Lviv
CDA	Calcium dependent antibiotic
ORF	Open reading frame
LC-MS	Liquid chromatography-mass spectrometry
PCR	Polymerase chain reaction
qRT-PCR	Quantitative reverse transcription PCR
HEWL	Hen egg-white lysozyme
SSR	Site-specific recombinase
PM	Phenotype microarray
HMM	Hidden Markov Model
CRISPR	Clustered regularly interspaced short palindromic repeats

1. INTRODUCTION

1.1 The history and challenges of natural product discovery

Natural products (NPs) can be defined as a large group of diverse chemical compounds with a wide variety of biological activities that have found various, multiple uses, notably in human and veterinary medicine and in agriculture [65]. NPs have a variety of original sources, such as: bacterial, fungal, plant, marine and animal origins. The bacterial natural product, which is the main target of this work is a product of "secondary metabolism" of bacteria. Secondary metabolites are the molecules that are not required for survival of the host organism under laboratory conditions, but which have a potential to provide an advantage to the host organism in its native environment with defensive or adaptive capabilities [65, 118].

Since the discovery of penicillin in 1928, estimated more than 23,000 NPs were discovered and characterized, most of them being produced by bacteria and fungi [65, 57]. Since than, NPs discovery has changed in its speed and methods. The first thirty years (1940s-1970s) of NPs discovery are frequently being characterized as "the Golden Age" and usually implied utilization of the simplest techniques of phenotypic screening of soil isolates against "test cultures". In general, a traditional phenotypic screen can be described by employing a bacteria, yeast, eukaryotic cells or tissues and a method to determine a response to an applied compound. It is usually conducted without knowledge of the mechanism of action of tested compound. The easiest output is the inhibition of the cell growth. The paradigm is as follows: 1) phenotypic screening; 2) compound isolation and structural characterization; 3) studies regarding mode of action in several cases; 4) pre-clinical development; 5) clinical development and commercialization. In the "Golden Age" era this pipeline was used to identify more than 1,000 NPs that had antibacterial or antifungal activity.

The systematic screening of soil isolates during this era led to the discovery of such important antibiotics as actinomycin, streptothricin and streptomycin.

The second thirty years period of NPs discovery (1970s-2000s) has seen the miniaturization of phenotypic screening processes: microtiter plates replaced the need for shake flask cultivation [91]. Also, major improvements in nuclear magnetic resonance (NMR) spectroscopy and mass spectrometry (MS) analyses have made structure determination of complex compounds straightforward and less time consuming. Screening of soil samples continued in the 1970s and 1980s, but was expanded beyond the search for antibacterial or antifungal drugs in soil samples. In addition, new sources were used for NPs discovery, this included marine environments, with a heavy focus on the discovery of novel anti-cancer agents. Samples, which were screened, included bacteria, algae, and marine invertebrates, such as tunicates, corals, bryozoans, sea slugs and sponges [65]. Advancements in recombinant DNA techniques in this era and other related technologies enabled researchers to rapidly determine mechanism of action of active compounds, their semi-synthetic analogues and synthetic compounds, and led to the development of biochemical or whole cell assays, collectively called as "the target-based screening" [65, 57].

It is difficult to estimate the exact amount of new NPs being discovered in this period, but the number of patents filed worldwide is reported to exceed 1,300 [71]. However, only less than two dozen novel NPs, their semi-synthetic or completely synthetic derivatives, that were discovered and studied during this period, were later developed and applied in medicine. It is claimed, that the only novel anti-bacterial NP, that went through clinical development, was daptomycin.

During the 1990s, combinatorial chemistry was developed, that gave a possibility to obtain thousands of unique compounds from hundreds of available chemical scaffolds. Most of the pharmaceutical companies developed their own libraries that numbered in the 5×10^5 to 4×10^6 range of compounds. Advances in robotics enabled the possibility to run these libraries through multiple high-throughput screens simultaneously. Numerous 2^{nd}, 3^{rd}, and in the case of the cephalosporins, 4^{th} generation

semi- or completely synthetic derivatives of NPs, representing many chemical classes, were successfully developed during this period [65].

The era of genomics-based methods (2000s and till now) has seen the rise in the amount of sequenced bacterial genomes and the sharp fall in sequencing costs. In the early 2000s, the first two *Streptomyces* genomes, one of which is of the model strain of Actinobacteria genetics, *S.coelicolor*, revealed the striking observation, that more secondary metabolite gene clusters are encoded in the large genomes of Actinobacteria than was predicted from their expressed secondary metabolomes. It was estimated in different reviews that less than 10% of secondary metabolite clusters are expressed in sufficient quantities to be detected under typical screening conditions. The others were shown to require specific conditions, manipulation of media or genetic engineering of their bacterial hosts to be produced in sufficient titers. [65, 57, 118]. In order to activate and harness the power of such "hidden" or so-called cryptic gene clusters, the vast number of genetic and fermentation methods was developed. Genetic methods include, but are not limited to: 1) chemical and transposon mutagenesis; 2) gene cluster duplication/amplification; 3) overexpression/knockout of regulator genes; 4) modulation of the transcription apparatus; 5) modulation of the translation; 6) modulation of post translation of carrier proteins; 7) heterologous expression of genes or biosyntheic clusters; 8) synthetic biology methods, e.g. refactoring of promoters and ribosome binding sites; and 9) combinations of 1) to 8) [65].

Scientific community worldwide acknowledges an urgent need in new medically potent NPs, resulting from the increasing number of multidrug-resistant pathogens and diseases to combat. It is believed that advances in next-generation genome sequencing, bioinformatics, analytical chemistry are going to be combined to overcome barriers in exploration of NPs. Coupled with new strategies in compound discovery, including inhibition of resistance, novel drug combinations, and new targets, NPs are predicted to enter the new "Golden Age" of discovery [65, 139].

1.2 Actinobacteria – the biggest source of industrially relevant natural products

Actinobacteria are Gram-positive filamentous soil bacteria that have a complex life cycle, embodying several stages of morphological differentiation. They produce many bioactive compounds as secondary metabolites, such as antibiotics, antitumor agents, immunomodulators, anthelmintic and insect control agents [68]. Actinobacteria, sometimes described with a more general term "actinomycetes", are frequently claimed to be a source of two thirds of all known secondary metabolites [14] or by some more modest estimations, 35% of all marketed antibiotic formulations contain an active ingredient derived from Actinobacteria [118]. In more specific terms, just one *Streptomyces* genus alone, is claimed to be utilized to produce more than 40% of all known microbial bioactive NPs [118]. Actinomycetes are also frequently described as a high G+C branch of Gram-positive bacteria [68].

One actinomycete is often capable of producing several to many active secondary metabolite compounds. A gentamicin-producing strain of *Micromonospora* forms around fifty secondary metabolites, that could be isolated to date [23]. Various compounds obtained from the same producer strain can be achieved by varying nutritional conditions, physical parameters or adding inhibitors [23, 57]. Biosynthetic genes of actinomycetes are frequently present in clusters coding for large multidomain and multimodular enzymes, e.g. polyketide synthases, prenyltransferases, non-ribosomal peptide synthases and terpene cyclases. Multiple gene clusters encoding secondary metabolites are common in species of *Streptomyces*, *Actinoplanes* and Mycobacteria. Genes adjacent to the biosynthetic gene clusters frequently encode regulatory proteins, oxidases, hydroxylases and transporters [23]. However, there are examples of biosynthetic gene clusters, that lack regulators or which are split in two parts which are located far from each other in the genome, e.g. *S. ghanaensis* [101].

As regarding their ecology, actinomycetes have a great ability to colonize multiple environments, water, animal and plant tissues and namely the soil. This ability is

further facilitated by growth as a vegetative hyphal mass which can differentiate into spores that assist in spread, survival and persistence of bacteria. The spores are a semi-dormant stage in the life cycle that can survive in soil for long periods: viable *Streptomyces* cultures were recovered from 70 year old soil samples. The spores of the representatives of Actinobacteria typically show resistance to low amount of nutrients and water, while the mycelial stage is highly sensitive to drought [68, 57].

1.3 Genus *Actinoplanes*, its members and their relevance for natural product discovery

Sporeforming actinomycetes related to the genus *Actinoplanes* are harboured taxonomically in the family Micromonosporaceae, order Micromonosporales, which belongs to the broad and diverse class Actinobacteria [126, 15]. The genus *Actinoplanes* was first described in 1950 [21] and characterized by bacteria-like, flagellated and capable of swimming spores, formed in sporangia. Later on, as a part of the description of formerly distinguished family Actinoplanaceae, it was concluded, that vegetative mycelia is formed in water on a variety of plant or animal substrates. Interestingly, an aerial mycelium is usually absent. If present, it can be composed of short rudimentary sterile hyphae [102]. Reproduction is carried out via spores in sporangia, which can be spherical, sub-spherical, cylindrical and irregular. Spores may vary in shape from subglobose, rod-shaped, bent rodes to spiral; motile or non-motile. Flagella of motile spores are mostly of the bacterial type. Colour of the mycelium is usually the shade of orange, though also other colours occur, e.g. red, blue, violet, and green pigmented strains are described [126].

Ecology of the genus *Actinoplanes* is extremely diverse. Its representatives can be found in soil, freshwater and marine environments. It is believed, that life cycles of the sporangiate genera are based on an alternation between terrestrial and aquatic habitats. The growth of vegetative mycelia, which is usually composed of thin, twisted, and branched hyphae, on plant or animal residues culminates in the

differentiation into sporangia. The sporangia can easily lose their connection to the degenerating mycelium and can be disseminated as spores by the wind or by soil fauna. The sporangia can survive for decades by withstanding prolonged desiccation and other unfavouring conditions [126, 102]. The sporangial envelope is usually water resistant and repellent. But if sporangia become re-hydrated by sufficient moisture, e.g., during periods of fog or rain, the spores inside the sporangia begin to swell, the sporangial envelope bursts, and the flagellated spores are released. [22].

A broad library of soil isolates was generated and characterized [22] in these studies, however struggle to define proper taxonomic description of all strains continued. As one of such efforts to determine taxonomic positions of members of former Actinoplanaceae family was carried out in [117]. Hydrolytic residues of cell walls of 48 different strains of this family, which were previously assigned to four different genera were examined by paper and column chromatography. Based on results and morphological studies, a single genus concept was suggested [117]. Frequently, cell wall composition is used as a characteristic feature. The cell wall composition was determined as the type II according to Lechevalier, since it contains meso-2,6-diaminopimelic acid, LL-2,6-diaminopimelic acid, and/or hydroxy-diaminopimelic acid and glycine as major components [75]. Further studies, using chemical and numerical taxonomic procedures, were carried out to investigate the taxonomy of the genus *Actinoplanes* and related genera [41].

Extensive studies on the members of the genus *Actinoplanes* and their potential to produce secondary metabolites were carried in [102]. It is noted in this study, that *Actinoplanes* members are able to synthesize polyene-type macrolides, glycolipids or aromatic polycyclic antibiotics. Also, they show an ability to polymerize amino acids to generate polypepteids and depsipeptides. As a noticeable feature, an absence of the amino cyclitol-containing polysaccharide antibiotics, e.g. the aminoglycosides and streptothricins, is highlighted. The authors conclude, that genus *Actinoplanes* is a versatile group of secondary metabolite producers, worthy of further investigations.

1.4 The object of the study: acarbose producer strain *Actinoplanes* sp. SE50/110

In December 1969, strain *Actinoplanes* sp. SE50 was isolated from a soil sample taken at the coffee plantation near to the city Ruiru, Kenya, Africa [34]. Among all other isolates being studied in parallel, it was tested in the course of valuable substance with inhibitory effects on glycoside hydrolases by the company Bayer AG. It was shown, that the strain *Actinoplanes* sp. SE50/110 is able to secrete a complex oligosaccharides mixture with a remarkable inhibitory effect on mammalian intestinal amylases, maltases, and sucrases [34, 108]. Strain *Actinoplanes* sp. SE50/110, which is the main object of this study, is a spontaneous mutant of *Actinoplanes* SE50. It was selected and characterized by a production of acarbose in rates close to 1 g/L [33]. At the moment derivatives of the strain *Actinoplanes* sp. SE50/110 are used for the worldwide production of anti-diabetic drug acarbose, one of the prominent secondary metabolites of the strain. Acarbose is mainly used for treatment of diabetes mellitus type II, but also can be regarded as a prevention drug. It helps patients with starch- and sucrose-containing diets delaying the digestion of ingested carbohydrates in the human intestine, thereby resulting in a steady, slowed rise in blood glucose concentration [106, 110].

1.5 Acarbose, antidiabetic agent and secondary metabolite of *Actinoplanes* sp. SE50/110

Metabolic disorders and illnesses like obesity, diabetes and atherosclerosis, necessitate carbohydrate-controlled diets for the patients. Restricting level of carbohydrates, which are being consumed daily, is considered to be a very complicated task. Therefore, use of the agents, which lower appetite, slow down the digestion of complex carbohydrates, and intensify carbohydrate and fat metabolism have been under intensive research in last decades [90, 130]. This work is dedicated to acarbose, one of

such prominent glucosidase inhibitors, wildly used in combination with other drugs, in attempt to prevent and cure diabetes and other metabolic diseases.

This section of introduction is designed to give an overview on the structure, modes of action and biochemical properties of acarbose and its homologues. Also taking into account its pharmaceutical and economic relevance, the scale of the problem of life threatening metabolic disorders, possible reasons of their spreading and development.

1.5.1 Pharmaceutical and economic relevance of acarbose, as an anti-diabetic drug

The term diabetes mellitus (DM) describes a chronic, progressive metabolic disorder of multiple etiology, characterized by hyperglycemia with disturbances of carbohydrate, fat and protein metabolism, resulting from defects in insulin secretion, insulin action, or both. Its effects include long-term damage, dysfunction and failure of various organs. It may present with characteristic symptoms such as thirst, polyuria, blurring vision, and weight loss. In its most severe forms, ketoacidosis or a non-ketotic hyperosmolar may develop and lead to stupor, coma and, in absence of effective treatment, death [2]. Diabetes is traditionally classified into type 1 diabetes mellitus (T1DM) and diabetes mellitus type 2 (T2DM) [7].

Complications and consequences of DM of all types to human health and economy are tremendous. Over time, high blood glucose damages nerves and blood vessels, leading to complications such as heart disease, stroke, kidney disease, blindness, dental disease, and amputations. Other complications of diabetes may include increased susceptibility to other diseases, loss of mobility with aging, depression and pregnancy complications [7, 2, 120].

Therefore, DM is recognized as an increasingly important public health problem. It was stated to be one of four priority non communicable diseases targeted for an immediate action by the world community. The number of cases of all types of DM has been steadily increasing over the past few decades. To put this problem in numbers,

an estimated 422 million adults were living with diabetes in 2014, compared to 108 million in 1980. The global age-standardized prevalence of DM has nearly doubled since 1980, rising from 4.7% to 8.5% in the adult population [138]. According to the World Health Organization report, which was published in 2016, DM caused 1.5 million deaths in 2012. Additional 2.2 million deaths were caused by "higher than optimal" blood glucose levels. Such condition can lead to the elevated risks of cardiovascular and other diseases. It is important to mention, that 43% of these 3.7 million deaths occur before reaching the age of 70 years [138].

Other studies confirm this data, e.g., the Global Burden of Disease Study 2013, designed to provide the latest and most comprehensive assessment of causes of death, detecting 240 specific causes of death in 188 countries, identified DM as one of the major causes of reduced life expectancy globally [95]. Further on, a general population-based survey of 98,658 adults conducted in China in 2010, estimated the prevalence of diabetes to be 11.6% and that of prediabetes to be 50.1%. This evidence is making China the country with the highest prevalence of diabetes in Asia and the largest absolute disease burden for diabetes in the world. Furthermore, poor awareness, treatment and disease control among patients with diabetes was revealed, which highlights an urgent need for a novel approaches in diabetes prevention and care [95, 140].

Current therapeutic approaches to treat T2DM include oral anti-diabetic drugs such as sulfonylureas, thiazolidinediones, metformin, α-glucosidase inhibitors and glycosurics. Inhibiting or blocking the digestion and absorption of carbohydrates has high therapeutic implications for controlling postprandial hyperglycemia in T2DM. In this regard, inhibition of digestive α-glucosidases is one therapeutic approach that slows down carbohydrate digestion and glucose absorption, therefore, stabilizing blood glucose level and preventing hyperglycemia in diabetic patients [36].

Acarbose has recently been recommended in guidelines for treating T2DM, receiving preferred status when compared to other oral glucose lowering drugs, due to its proven ability to reduce cardiovascular events [3]. Additionally, acarbose has minimal risk for hypoglycaemia and, when titrated slowly, is generally well tolerated.

Acarbose has also been found to improve vascular health to have antiplatelet effects [25] and to magnify incretin secretion and positively change gut microbiota [60].

1.5.2 Molecular structure of acarbose and its homologues

The α-glucosidase inhibitor acarbose, produced by strains of the genera *Actinoplanes* and *Streptomyces* [104], is a member of an unusual group of bacterial, mainly actinomycete, secondary metabolites, called amylostatins. This group includes oligostatins, trestatins and validamycins as well, all of which inhibit various glucosidases [129]. Structurally acarbose (O 4,6-dideoxy-4 [1S- (1,4,6/5) -4,5,6- trihydroxy -3- hydroxymethyl -2- cyclohexen-1-yl]-amino-α-D-glucopyranosyl-(1-4)-O-α-D-glucopyranosyl-(1-4)-D-glucopyranose) can be characterized as a pseudotetrasaccharide (Figure 1). Acarviosine part comprises a pseudodisaccharide, based on the unsaturated C7-cyclitol (valienamine or valienol) bound via an imino-bridge to 4-amino-4,6-dideoxyglucose. Acarviosyl moiety is essential for the inhibitory effect of the whole compound, as long as imino-bridge can not be hydrolyzed by α-glucosidic enzymes [129].

Apart from acarbose *Actinoplanes* sp. SE50/110 is able to produce a wide variety of acarviosyl-based pseudooligosaccharides. First studies on the structure of acarviosyl-containing metabolites were conducted very early [108]. It was established, that *Actinoplanes* can produce a series of compounds of different structure with different substrate specificity. It was strikingly dependent on the fermentation conditions, especially the carbon source in the medium [108]. These compounds can differ from acarbose in the number of glucose units connected to each other by α-1,4-glycosidic bonds which are attached to the acaviosyl core either at the reducing or non-reducing ends as shown in Figure 2. The inhibitory specificity against different α-glucosidases depends on the number of glucose units. Furthermore compounds with differences in the type of the terminal glycosidic bond or in the nature of the terminal sugar moiety can be found [142, 51]. The outlook of the general characteristics of acarviosyl-containing compounds is given in Figure 2.

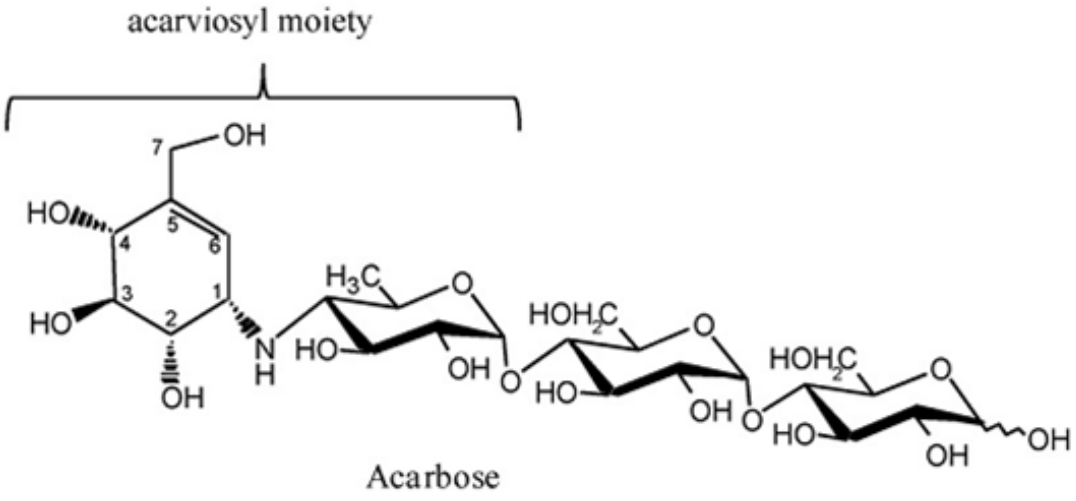

Fig. 1. Schematic structure of the acarbose. Acarviosyl moiety (valienaminyl-4-amino-4,6-dideoxyglucose) is α-1,4-bound to a maltose residue. Modified from figure in [129]. Copyright 2003 by the Taylor & Francis Ltd. Adapted and reprinted with permission.

Name	m	n
Component 2	0	1
Component 3 (acarbose)	0	2
Component 4-isomer*	1	2
Component 5-isomer*	2	2
Component 6-isomer*	3	2

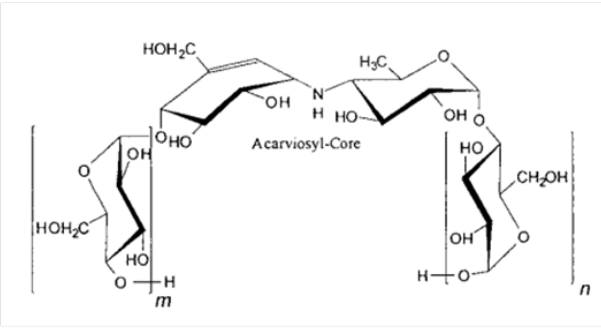

Fig. 2. Chemical structures of the acarbose and amylostatin family of α-glucosidase inhibitors synthesized by *Actinoplanes* sp. SE50/110. For components marked with an asterisk, the main ingredient of the isomer mixture with (m + n) is 3 (4 or 5) [51]. Copyright 2001 by the American Society of Microbiology. Reprinted with permission.

It was determined, that the different acarbose homologues are formed dependent on the sugar source supplementation of the media. If glucose or maltose are supplied as the sole carbon source inhibitors with a small number of glucose units are produced, preferentially acarbose, while addition of starch leads to compounds with a higher number of glucose units. This process can be easily understood, taking into account, that acarbose mostly act as a strong inhibitor of disaccharidase and long-chain acarbose derivatives mostly are active against amylases [108]. Now it is known that acarbose in contact with α-amylases and cyclodextrin glucanotransferases in the presence or absence of maltooligodextrins becomes converted to longer-chain derivatives containing at least two acarviosyl residues, as shown in crystallized enzyme-inhibitor complexes. Therefore, acarbose can be regarded as a prodrug which forms more active inhibitors by the catalytic activity of its target site [51, 129, 110].

1.5.3 Modes of acarbose action as an inhibitor of various glucosidases

Carbohydrates are being consumed as a part of human diet either as simple sugars or as more complex molecules, called polysaccharides. The body absorbs carbohydrates from small intestine in the form of simple sugar units, called monosaccharides, e.g. glucose, fructose. Polysaccharide molecules routinely require enzymatic digestion, which starts in the mouth by the action of an enzyme, called α-amylase. Enzyme causes endohydrolysis of (1-4)-α-D-glucosidic linkages in the starch chain and produces maltose as a resulting unit. Apart from non branched molecules of amylose (15-25%), usually linked by α 1-4 connection, starch contains also amylopectin (75-82%), which has highly branched chains with α 1-6 connections [90, 130]. Latest are almost not affected by α-amylases in saliva. Therefore, as a result of primary digestion, maltose and so called "limit dextrins" are formed. The process of digestion continues further in the stomach and intestines by the action of gastric and pancreatic enzymes, which contain also α-amylase, similar to the one in saliva and other enzymes, such as glucoamylase, oligo- and disaccharidases. Dextrins, formed earlier

are attacked by enzyme, capable of disrupting inter-chain branching links. Further on, specific enzyme maltase splits maltose into two glucose units, which are later absorbed into blood stream [34].

Reports on natural inhibitors of α-glucosidases date back to thirties of the previous century. The discoveries were made from the range of natural sources: buckwheat, rye and wheat flour, beans, corn and peanuts [130, 90]. Some of the preparations, that were isolated, found to be active inhibitors of saliva and pancreas α-amylases. New concept of implantation of α-glucosidase inhibitors to the treatment of metabolic disorders was developed by Puls [106]. It gave a further push for screening for new active substances, especially of actinomycetes origin. [130]. It can be explained by fact, that, not all the inhibitors acting on the digesting ability of α-glucosides with respect to starch, acted on the hydrolysis of saccharose. Therefore, application of such inhibitors to animals and clinical patients, did not noticeably influence the blood glucose content and was not reflected in the resulting body weight. Therefore, many inhibitors of microbial origin, which have shown a broader spectrum of action with respect to α-glucosidases, were paid more attention accordingly to these findings [90]. Many of the isolated inhibitors, belong to various classes of chemical compounds and have different substrate specificity. Among the isolated inhibitors are amino sugars, oligosaccharides, proteins, and glycopolypeptides. Acarbose, An α-glucosidase inhibitor, which is a subject of this work is a pseudo oligosaccharide [90].

Very early after the first identification and isolation of α-glucosidase inhibitors from the cultural broth of *Actinoplanes*, it was defined that inhibitors of various structure and substrate specificity are being synthesized in regard of the medium composition. [108]. Acarbose, as one of those compounds, has an inhibiting action on γ-amylase, saccharase, and certain other α-glucosidases from the mucous membrane in small intestine: maltase, dextrinase, glucominase, although it does not inhibit α-amylase and lactase. Studies on the inhibition kinetics and mechanism showed that acarbose is a competing inhibitor of α-glucosidases [90, 18]. Crystals of enzyme-inhibitor complexes were used to study the mode of action of the inhibitory

activity of acarbose on several α-glucosidases and cyclodextrin glucanotransferases. The acarviosil- moiety binds with high affinity to the active centers of the enzymes [129]. This binding is particularly strong in case of saccharases. It was found, that the affinity of acarbose to saccharases is about 15,000 fold higher than the affinity of the enzymes for saccharose, demonstrating that acarbose forms uniquely stable enzyme/inhibitor complexes [18, 129].

1.6 Structure of *acb* gene cluster in *Actinoplanes* sp. SE50/110 and putative functions of encoded proteins

Gene cluster for the acarbose biosynthesis, *acb*-cluster, was cloned and studied prior to the sequencing of the *Actinoplanes* sp. SE50/110 genome [121]. As long as structurally acarbose resembles aminoglycosides, genes *strE* and *strD* from the *S. griseus* streptomycin gene cluster, were chosen as a probes for the hybridization with *Actinoplanes* sp. SE50/110 genomic DNA [103]. In this way, at first genes *acbABC*, coding for dTDP-D-glucose synthase, dTDP-D-glucose 4,6-dehydratase, dehydroquinate synthase respectively, were cloned [116]. Later on, utilizing this approach a 17 kb fragment that includes genes *acbQKMLNOCBAEDGF* was identified, isolated and studied [116, 115]. The full length *acb* gene cluster, covering 35 kbp of genomic DNA and consisting of 25 ORFs, was identified for the first time in work of [121], through the screening of the *Actinoplanes* sp. SE50/110 cosmid library [116, 115]. It was suggested to be organized in at least eight transcription units; three of which are operons: *acbZ*, *acbWXY*, *acbVUSRPIJQKMLNOC*, *acbB*, *acbA*, *acbE*, *acbD* and *acbHGF* [143].

In further studies composition of the *acb*-cluster was corrected. According to the crystallographic analysis of the extracellular binding protein AcbH the suggestion was made for the putative acarbose importer AcbHGF to bind galactose, rather than acarbose [76]. These data was supported by comparative RNA-sequencing of *Actinoplanes* sp. SE50/110 transcriptome, being conducted in minimal and complete

Fig. 3. Schematic organization of acarbose biosynthesis gene cluster from *Actinoplanes* sp. SE50/110. Modified from figure in [110], the Open Access article distributed under the terms of the Creative Commons Attribution License (http://creativecommons.org/licenses/by/2.0).

growth media. It was shown, that *acbGFH* genes are not expressed in coordination with the whole gene cluster [111]. Therefore, at the moment *acb*-cluster is considered to contain twenty two genes, seven independent transcription units, among which there are two operons *acbWXY* and *acbVUSRPIJQKMLNOC*, Figure 3.

Evidences for the functionality of *acb* gene cluster to represent the full set of genes, needed for acarbose production stem from: 1) the enzymology of individual Acb-proteins, studied mostly through heterologous expression; 2) the expression of the cosmid pHTWCos6 which contains the full-length *acb* cluster of *Actinoplanes* sp. SE50/110 in *S. lividans*, resulting in heterologous production of small amounts of acarbose or a compound, which closely resembles acarbose [121].

Several *acb* genes have been characterized experimentally already. Among them, the *acbC* gene which is coding for putative dehydroquinate synthase. It was cloned and expressed heterologously in *S. lividans*. Isolated protein, product of this gene was shown to be a C7-cyclitol synthase using sedo-heptulose-7-phosphate as a substrate for the production of 2-epi-5-epi-valiolone [116].

The gene *acbE* was shown to encode a resistant to acarbose α-amylase, which is able to use regular starch and amylose as substrates. It was also shown to be induced by the addition of maltotriose to media [115]. The neighboring gene *acbD*,

which product acarviosyl transferase (ATase) was shown to catalyze the transfer of the acarviosyl moiety of acarbose to maltooligosaccharides [51].

The most of *acb* genes so far have not been experimentally investigated. Their functions are postulated only on the basis of the bioinformatic comparisons of the deduced protein sequences with known proteins. Interestingly, among the *acb* genes for which a function is not yet evident, no regulator protein-encoding gene has been identified [129].

1.7 Complex "–omics" technologies to characterize metabolism of acarbose in *Actinoplanes* sp. SE50/110

The extensive body of data on *Actinoplanes* sp. SE50/110 genomics, proteomics and transcriptomics was obtained in a course of research being conducted in Bielefeld University, Center of Biotechnology (CeBiTec) in recent years. In order to better characterize and review this data, it is represented here in subchapters, in regards to the methods utilized while obtaining it. Several provisional models of acarbose biosynthesis, transport and regulation of these processes, were built in the course of these and other studies. Therefore, it is of great importance to learn the result of the application of complex "-omics" technologies to study acarbose metabolism in *Actinoplanes* sp. SE50/110 in order to proceed with further investigations.

1.7.1 Sequencing of *Actinoplanes* sp. SE50/110 genome and its annotation

As a prerequisite for performing any genetic modifications or "-omics" studies, complete genome sequence of the strain has to be known. Therefore, *Actinoplanes* sp. SE50/110 strain was selected for whole genome sequencing, as it possesses elevated acarbose production of up to 1 g/L in comparison to the wild type. At the moment of sequencing, the *acb* gene cluster was already identified and sequenced in this strain [121].

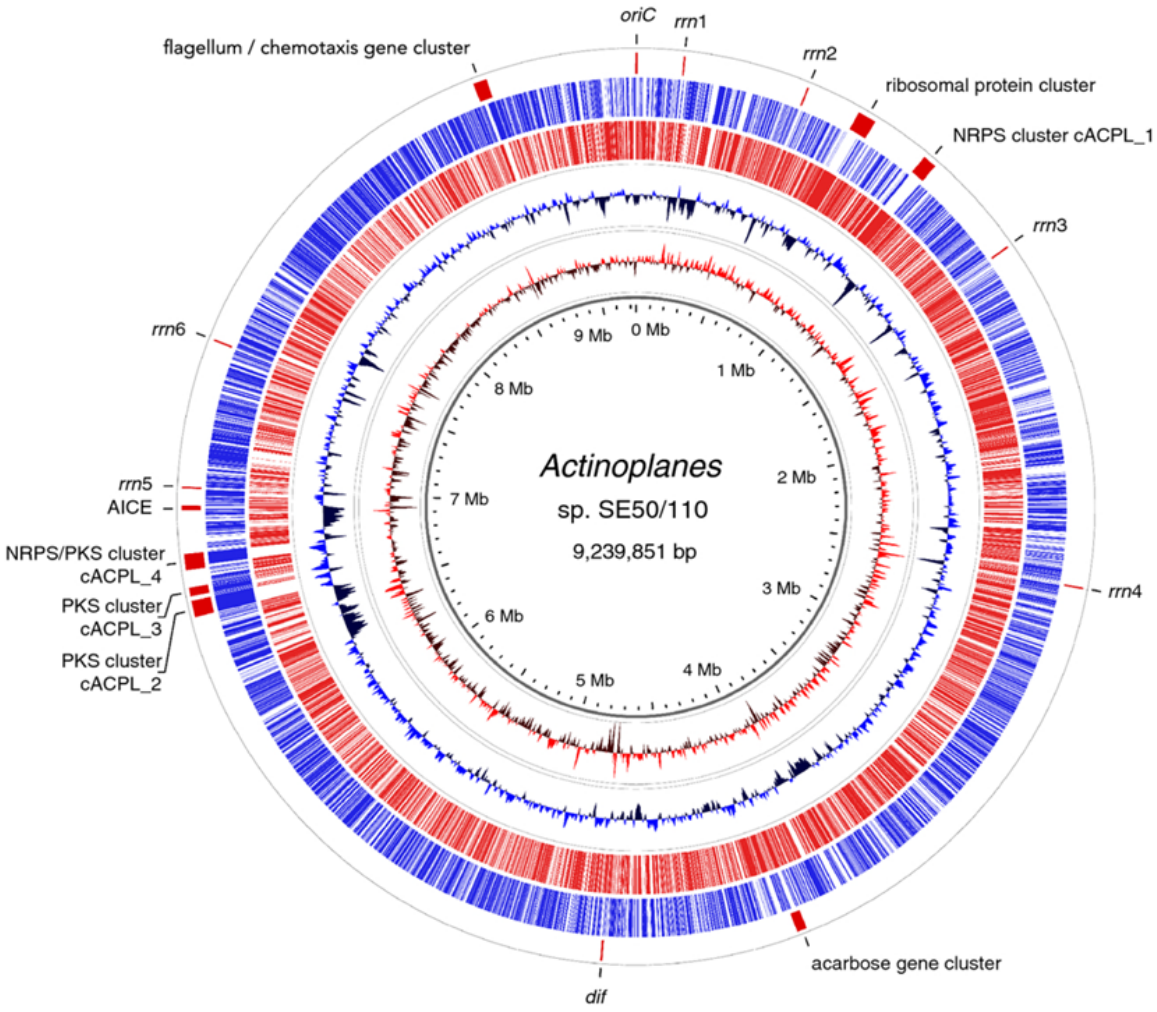

Fig. 4. Plot of the complete *Actinoplanes* sp. SE50/110 genome. The circles represent from the inside: 1) scale in million base pairs; 2) GC skew; 3) G + C content (blue above- and black below genome average); 4) genes in backward direction; 5) genes in forward direction; 6) gene clusters and other sites of special interest. Abbreviations were used as follows: *oriC*, origin of replication; *dif*, chromosomal terminus region; *rrn*, ribosomal operon; NRPS, non-ribosomal peptidesynthetase; PKS, polyketide synthase; AICE, actinomycete integrative and conjugative element. Figure originates from [110], the Open Access article distributed under the terms of the Creative Commons Attribution License (http://creativecommons.org/licenses/by/2.0).

The complete genome determination of *Actinoplanes* sp. SE50/110 was accomplished by combining the sequencing data, which was generated by paired end and whole-genome shot-gun pyrosequencing strategies. The sequencing resulted in the final assembly of a complete single circular chromosome of 9,239,851bp with an average G + C content of 71.36% (Figure 4). The origin of replication *oriC* was identified as a 1266 nt intergenic region between the two genes *dnaA* and *dnaN*. As reported in [110], a total of 8,270 protein-coding sequences (CDS) were determined, which include: 4,999 genes (60.5%) with an associated functional COG category, 2,202 genes (26.6%) with a fully qualified EC number and 973 orphan genes (11.8%) with no similar sequences in public databases. Classification of 4,999 CDS with annotated COG-category with the use of bioinformatic tools revealed a strong emphasis(47%) on enzymes related to metabolism: amino acid (10%) and carbohydrate metabolism (11%). Furthermore, 16% of the COG-classified CDS were found to code for proteins involved in transcriptional processes, which suggests a multi level of gene regulation cascades. Consistently, 4% of all identified CDSs were found to be involved in secondary metabolite biosynthesis, in similarity to other Actinobacteria, e.g. *S. coelicolor*, for which it reaches 5% [11].

It is interesting to note, that *acb* cluster was found to occur only once in the genome of *Actinoplanes* sp. SE50/110, however many of the genes from this cluster have a close homologs in the genome. By utilizing BLAST analysis, single genes and gene sets with equal functional annotation and amino acids sequence similarity were found scattered throughout the genome [110]. The annotation of genome of *Actinoplanes* sp. SE50/110 was further improved in following studies, e.g. with the help of RNA sequencing technology[111, 109]. Results of these studies are reviewed in the next chapter.

1.7.2 Transcriptomic studies of *Actinoplanes* sp. SE50/110 metabolism

In order to gain insights into the differential expression of genes involved in the acarbose metabolism under different growth conditions, high-throughput sequencing of copy-DNA (cDNA) libraries, commonly referred to as RNA-sequencing or RNA-seq, was performed [109]. Shortly described, RNA-seq involves: 1) isolation of RNA from cells; 2) the removal of ribosomal RNA; 3) the reverse transcription of messenger RNAs in order to create a cDNA library; 4) sequencing of cDNA library and analysis [111]. RNA-seq provides precise characterization of transcriptome with a single-base resolution, therefore it is very well suited for the improvement and correction of the annotation of bacterial genomes.

In this study, the assignment of transcription start sites was done by sequencing and mapping enriched 5′ ends of primary transcripts to the previously sequenced genome of *Actinoplanes* sp. SE50/110. Further on, the location and number of read starts of transcription start sites were used to determine the percentage of leaderless transcripts; conserved promoter motifs of genes with and without 5′-untranslated regions were identified in similar fashion. Determination of the precise transcription start sites is important not only in order to improve a current annotation of *Actinoplanes* genome, but also to detect novel protein coding genes and previously unknown non-coding RNA sequences. In this work, *Actinoplanes* was cultivated in three different liquid growth media: minimal media with maltose, the same media supplemented with trace elements and glucose supplemented complex medium, in order to reach a necessary variety of transcripts.

As a result of this study, altogether 1,427 putative transcription start sites were detected. With the help of the annotated genome sequence, 661 transcription start sites were found to belong to the leader region of protein-coding genes. Interestingly, only approximately 20% of these genes rank among the class of leaderless transcripts. In case of the canonical transcripts,a 5′-UTR length was ranging from 3 to 100 bases. In addition to this, 9 novel protein-coding genes and 122 previously undetermined

non-coding RNA (ncRNA) genes of *Actinoplanes* sp. SE50/110 were identified. This study demonstrates well, that sequencing of 5′ enriched ends of primary transcripts and identification of transcription starts are both valuable tools for advanced genome annotation of bacteria [109].

Further on, a whole transcriptome study using RNA-sequencing technology was conducted in order to compare *Actinoplanes* sp. SE50/110 gene expression while grown in maltose minimal medium (Mal-MM), maltose minimal medium with trace elements (Mal-MM-TE) and glucose complex medium (Glu-CM) [111]. This study was conducted and designed in order to answer some basic questions, namely: which gene clusters and genes are expressed highly and the most differently in these three media and the effect of added trace elements on gene transcription. As a result of these studies no acarbose was detected in the supernatant of the glucose complex medium and the supplied trace elements were shown to promote cell growth, which putatively caused an increase in acarbose production. Among top twenty most expressed genes over all involved conditions include ones coding for: three conserved membrane proteins of the unknown function, putative transcriptional regulator CarD, ribosomal proteins and putative starch binding enzyme Cgt. As to speak specifically about addition of trace elements to maltose minimal media, special software detected around 70 significantly differentially expressed genes, including two adjacent operons *acpl3030/copY* and *cydAB(CD)R* and *modHABCR* gene cluster. As expected, all of these genes in various manner are connected to homeostasis of metal ions. While comparing cells grown in Glu-CM and Mal-MM, it was found a sharp reduction of *acb* gene cluster expression in the first one in comparison to the second. This finding was claimed by authors to be in line with previous observations [111].

1.7.3 Proteomic and metabolomic studies of acarbose production

Substantial proteomic and metabolomic studies on acarbose producing *Actinoplanes* sp. SE50/110 were conducted in recent years in CeBitec, Bielefeld University.

In the first study, intra- and extracellular proteoms were analyzed. The number of Acb-proteins were identified, namely nine of them intracellularly and three extracellularly. The secretome of *Actinoplanes* was shown to be dominated mainly by seven proteins, which are all involved in carbohydrate metabolism. Taking this into account, MalEFG system was suggested as a missing acarbose-import system [131]. On the base of this study, corrections to the model of acarbose biosynthesis, proposed in works of [142] and [128], were suggested (see below).

Further proteomic and metabolomic studies were carried out in [132]. Cells of *Actinoplanes* were grown in maltose, glucose, galactose or mixed minimal media and a combination of HPLC and LCMS measurements of metabolite production was deployed. Further on, a qualitative analysis of extra- and intracellular proteome for the presence of the *acb* cluster proteins, have shown that they were synthesized in all media studied. Based on the results, a new model for the intracellular biosynthesis of major and minor acaviose components was developed (discussed below).

A year later, novel proteomics approach, which combines subcellular fractionation, shotgun proteomics and spectral counting, was used to study the abundance of *Actinoplanes* proteins within different fractions [133]. Four different fractions were studied in the course of this study: cytosolic, enriched membrane, membrane shaving and extracellular. Surprisingly, intracellular Acb proteins were detected in both cytosolic, but also enriched membrane fraction. This finding indicates, that biosynthesis of acarbose might take place at the inner membrane itself. In a related recent study, similar proteomics approach was used to study the differences in the proteome of *Actinoplanes* wild type strain cultures grown in Mal-MM and Glu-MM. [134]. When comparing protein quantities, detected in both conditions, differences were observed for saccharide transport and metabolism proteins. At the same time, differences for acarbose biosynthesis gene cluster proteins were almost absent.

1.8 Putative models of biosynthesis of acarviosyl-containing metabolites in *Actinoplanes* sp. SE50/110

In all of the studies conducted on the *Actinoplanes* acarbose production, several putative models of acarbose metabolism were formulated. The first model was based on the previously conducted feeding experiments, the heterologous expression of *acbKLMNOC* genes and study of the biochemical and enzymatic activities of the products of the respective genes. Also, information obtained from sequencing of the *acb* gene cluster was taken into account [142]. However, no genetic studies in the native conditions were performed, as long as no genetic tools for *Actinoplanes* were available at the time. The first steps of acarbose biosynthesis were already well researched, it was determined that the starting precursor of the biosynthetic pathway 2-*epi*-5-*epi*-valiolone is derived from *sedo*-heptulose-7-phosphate. Further conversion of 2-*epi*-5-*epi*-valiolone to the final cyclitol moiety was studied by testing multiple enzymatic mechanisms such as dehydration, reduction, epimerization, and phosphorylation [142]. The bisoynthesis of the deoxysugar part of acarbose was suggested to be carried out via a separate biosynthetic branch by AcbA, AcbB and acbV proteins. The cyclization of two products nucleoside diphosphate-1-*epi* -valienol 7-phosphate and dTDP-4-amino-4,6-dideoxy-D-glucose was suggested to be carried out by AcbS protein. However, further formation of acarbose or other longer chain molecules intra- or extracellularly was described as unclear.

Soon after the suggestion of the first model, the new role of acarbose as a possible "carbophore" and a cyclic metabolism of acarbose was proposed. The cycle included the combination of an active export of acarbose, combined with dephosphorylation and conversion by α-1,4-glucosyltransferases (e.g. AcbD), re-uptake by the specific importer, rephosphorylation and deglucosylation of acarviosyl-containing molecule. The role of acarbose as a tool in a blocking of enzymes of other bacteria was widely accepted. However, here the new multifunctional role of acarbose as a "sink" for C-sources and simultaneously a transport vehicle was suggested [128]. The next model of

acarbose biosynthesis was suggested in [131], further building on the ideas of the first model. Due to the first successful examination of intra- and extracellular proteome of *Actinoplanes* in this work, *acb* gene cluster proteins, and other related to the acarbose biosynthesis, were identified. For example, AcbD protein was identified in both intra- and extracellular proteomes, however being more abundant in second case. Therefore, a role for AcbD as an extracelullar acarviosyltransferase was further supported. Importantly, new AglEFG/MsiK α-glucoside, namely maltose/sucrose/trehalose, import system and MalEFG acarviosyl-metabolite import system, were suggested in addition.

The model of [131] was further developed in [132]. In this work, qualitative study of proteome of *Actinoplanes*, was combined with HPLC and LCMS measurements of supernatant samples, collected from *Actinoplanes* grown in various Carbon source minimal media. Interestingly, proteins related to acarbose biosynthesis, were detected in bigger or smaller amounts in all media tested. This proves, that glucose has no negative effect on acarbose production, as it was suggested earlier. Further on, LCMS measurements of supernatant samples, revealed presence of all four expected acarviosyl components in all conditions tested. Based on the results of the study, a model for intracellular formation of major and minor metabolites was suggested. The overview on this model can be seen from Figure 5.

Further ideas regarding modeling of acarbose metabolism were developed and discussed in PhD thesis of Dr. Vera Ortseifen [97]. The inspiration for these models stems from the examination of the proteoms and metaboloms of two industrial producer strains in comparison to the wild type. The model of acarbose metabolism in a wild type strain, grown in maltose minimal media, which was suggested by Dr. Vera Ortseifen is represented in Figure 6.

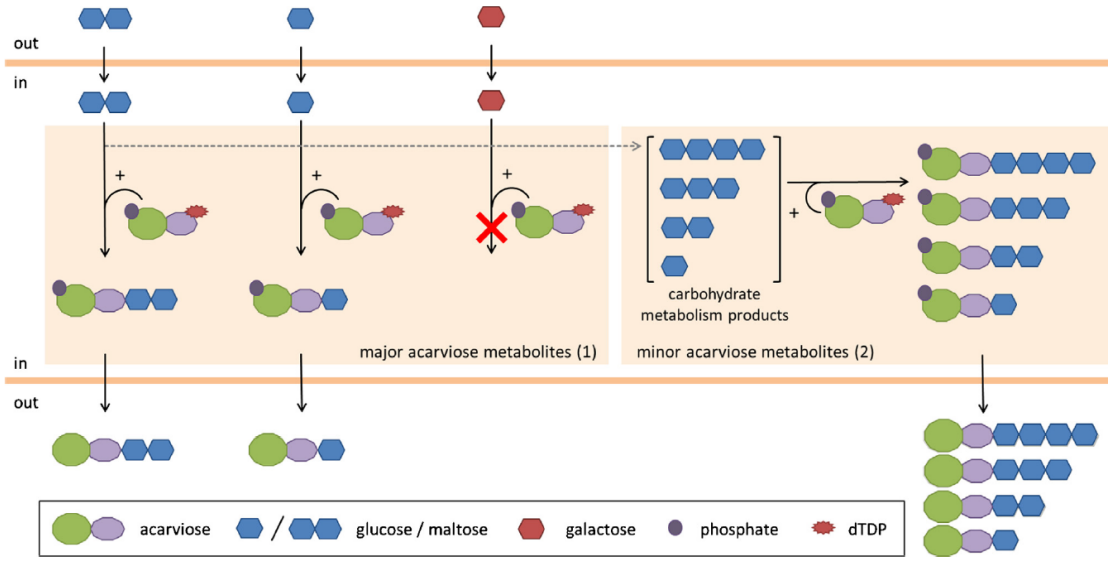

Fig. 5. Model for the formation of acarviosyl metabolites by *Actinoplanes* sp. SE50/110, as reported and discussed in [132]. Copyright 2014 by Elsevier B.V. The figure is reprinted with permission.

1.8.1 Regulation of acarbose metabolism in *Actinoplanes* sp. SE50/110

At the time only limited information exists regarding the regulation of acarbose production in *Actinoplanes* sp. SE50/110. MalR regulator, its functions and involvement in acarbose metabolism was described in Master thesis of Julian Droste [27]. Knockout mutant *Actinoplanes δmalR*, which was characterized in work of Julian Droste, was generated in this work through ReDirect procedure, which is discussed in details in Results section. MalR was identified as a possible regulator of acarbose metabolism mainly through *in silico* comparison of regulators of acarbose biosynthetic gene clusters in streptomycetes with the *Actinoplanes* genome. Gene *malR* is situated in immediate proximity to the putative maltose/maltodextrin ABC transporter *malEFG* operon. Therefore, it was logically suggested to be involved in the regulation of this operon, rather than *acb* cluster. In work of Julian Droste, *Actinoplanes δmalR* strain was characterized through cultivation experiments, genome-wide

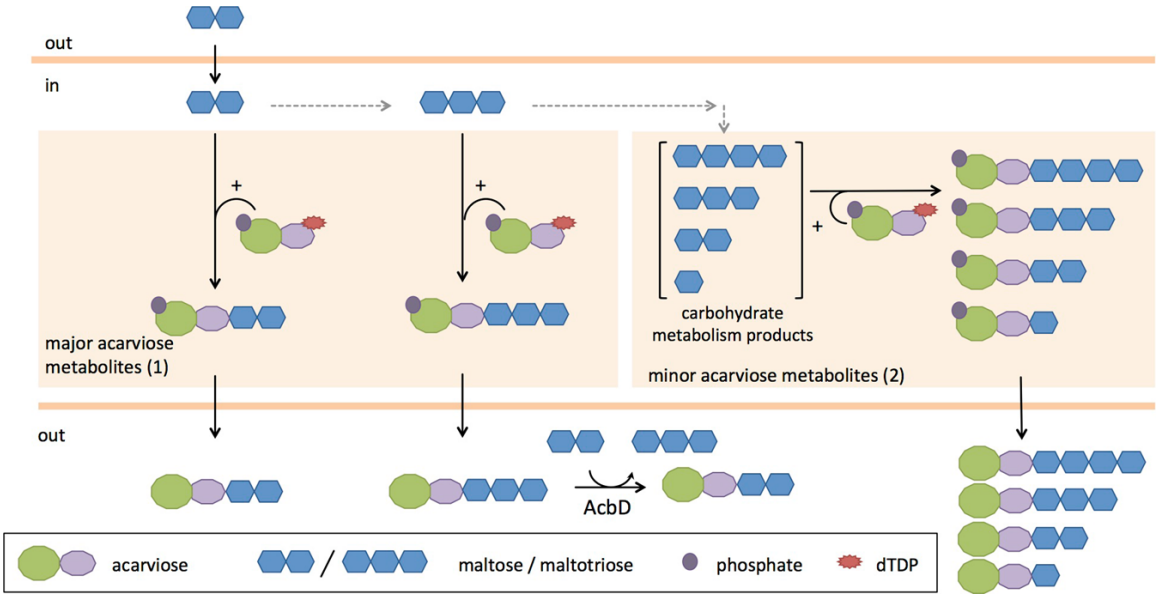

Fig. 6. Model for the formation of acarviosyl metabolites by *Actinoplanes* sp. SE50/110, grown in Mal-MM media, as reported and discussed in [97]. The figure is reprinted with authors permission.

microarrays, which were verified with qRT-PCR. From the results of cultivation experiments, it was shown that the knockout of *malR* gene has only limited influence on acarbose production. In order to show the binding of MalR protein to the region of *acb* cluster, electrophoretic mobility shift assays were performed. It was elucidated that MalR protein is able to bind to the region between two genes from *acb* cluster *acbE* and *acbD*. It was also suggested, that MalR protein acts as a repressor of these two genes, thus mildly influencing acarbose production. In this work, one more putative regulator of acarbose metabolism, namely CadC (ACPL_212) is described. More information on this protein can also be found in [123].

1.9 Genetic engineering methods developed for Actinobacteria

Genetic engineering methods, which are traditionally developed for Actinobacteria, can be roughly divided into two categories: native gene cloning systems or studies

of the genes and genetic clusters in heterologous conditions. Methods, which belong to both of these categories were proven to be very successful in past, depending on the task or particular bacteria, for which it was applied.

Typically, heterologous expression of gene or gene clusters is the option, which is explored first. It is usually defined as "expression of a gene or a set of genes from one organism in a different species" [39]. This method can be used for various purposes, some of them are discussed here. The most widespread use is probably to study the function of individual genes, which are present in a gene cluster of interest. In this case, genes are cloned into expression vectors and than transferred to so-called "model organisms", bacteria, for which genetic methods are already established and whose metabolism is studied and understood to a great extent. Than, protein, expressed from gene of interest, is purified and studied in regard of its biochemical and enzymatic properties. These experiments help to clearly define the function and role of particular proteins in biosynthetic pathways. Similar studies were conducted for several genes from *acb* gene cluster, e.g. *acbD* gene [51]. Other frequent use of this technique is to demonstrate that an entire biosynthetic gene cluster has been cloned and some defined amount of genes is sufficient for a synthesis of a particular metabolite [82]. It is also possible to analyze the whole route of biosynthesis of particular natural product in heterologous host, especially when host is more genetically exploited than the original source. In this case, whole biosynthetic gene cluster is cloned, typically as a part of cosmid or fosmid, and transferred to the host. Than, single mutant strains, deficient in all of the genes of this cluster, are created and studied [44]. Another widely used application of this method includes production of metabolites at higher levels than in the natural producer. As it is well known at the moment, genomes of Actinobacteria possess a huge potential towards production of various vast number of secondary metabolites, which are frequently encoded in the genome as a cryptic gene clusters [39, 10]. Majority of the identified cryptic gene clusters are not expressed at significant levels in their natural hosts, in order to be detected in a screening procedure. Therefore, expression of such clusters in heterologous

hosts, frequently together with genetic manipulation of pathway-specific regulatory genes of this cluster and adoption of conditions, most favorable for the production of this particular compound, lead to the higher metabolite production [38, 39]. It is also possible to exploit possibility of creation of new compounds by combinatorial biosynthesis, for example by combining genes from different metabolic clusters, which could originate from several different bacteria. Such a genetic combination can be than expressed in heterologous host and lead to the creation of new, previously uncharacterized compounds [4]. Study of genes or genetic clusters in heterologous conditions is a powerful genetic tool, which was proven to be useful on multiple examples of industrially important producers [141, 14].

Development of the native gene cloning systems typically includes number of steps and has multiple pros and cons in comparison to the method of heterologous expression, which was discussed above. The comparatively biggest amount of techniques, methods and protocols, developed for streptomycetes in this direction so far, is included into [68]. These include, but are not limited to: selection of an appropriate cultivation media, which supports the optimal production level and rate of desired product; application of the suitable DNA transfer technique, most typically, electroporation, protoplast transformation or intergeneric conjugation; choice of the plasmid-based vector system, integrative or replicative [68]; development of the reporter systems, which can be used to study the transcription or translation of various genes or combination of genes [93].

Both of the groups of the methods have its advantages and disadvantages, which have to be considered carefully in order to make a sensible decision. Heterologous expression of genes and clusters is relatively fast, straightforward and is known to achieve high results. However, not all of the genetic clusters can be expressed in heterologous conditions, due to multiple reasons, e.g. considerable differences between host and natural producer strain, not optimal conditions of the production. From the other side, establishment of the native gene cloning system is usually labour and time consuming, because typically several different approaches are being tested in order

to find the most suitable one. But once such a system is established and profoundly tested, it can be used for a great variety of genetic experiments.

In recent years a string of novel genetic methods, which do not or only partially belong to the abovementioned groups, has risen. Among them are site-specific recombinases, genetic controlling elements, endonucleases and the CRISPR-Cas9 system [14]. Here, we discuss only the later due to the limitation of space, for a more comprehensive and full review please refer to [14].

Till the moment, CRISPR-Cas9 is the only adaptive immune system in prokaryotes known to researchers. In this system, small guide RNAs are employed for sequence specific interference with invading nucleic acids [54]. CRISPR array by itself comprises a genomic sequence that consists of short repetitive elements, or repeats, which are separated by unique sequences. These sequences, or spacers, originate from mobile genetic elements such as bacteriophages, transposons or plasmids. Such a region is usually flanked by a set of *cas* genes encoding the Cas proteins [14, 54]. The advantage of the CRISPR-Cas9 system, as an genetic engineering technique, is the introduction of a double strand break in a desired target sequence with a great flexibility and precision that is mediated by a customized single guide RNA. CRISPR-Cas9, as a major genetic engineering technique, was adopted in a wide variety of organisms. Actinobacteria, mainly streptomycetes were among those, for whom this method was applied early after its discovery. Targeted chromosomal deletions, ranging from 1 bp to 30 kb, in four *Streptomyces* species have been generated using the engineered CRISPR-Cas9 system [14, 19]. In addition, based on the catalytically disabled Cas9 protein, it was possible to efficiently control gene expression in a model Actinobacteria strain *S. coelicolor* [14, 122].

For the members of genus *Actinoplanes* only several known examples of establishment of native gene cloning systems are available at the moment. Gene cloning systems, based on the methods, previously used and described for steptomycetes were established for *A. teichomyceticus*, *A. friuliensis* and *A. missouriensis* [48, 59, 58, 96, 127, 64]. Genetic manipulations of the members of this genus remains a challenging

task, which requires further studies. *Actinoplanes* sp. SE50/110 became the first representative of the genus, for whom CRISPR-Cas9 was applied as a genetic engineering tool, as apart of this work, which is discussed below [137].

1.10 Goals and objectives of the study

Large amount of data regarding genomics, transcriptomics, proteomics and metabolomics of *Actinoplanes* sp. SE50/110 was acquired in recent years. Several models of acarbose metabolism were formulated, taking this data into account [142, 131, 132, 97]. However, none of models gives a clear and transparent explanation of all steps of acarbose metabolism, its regulation and acarbose transport. Most of the steps of acarbose biosynthesis pathway are suggested from theoretical considerations and were not yet confirmed experimentally. Roles of many proteins, which belong to the *acb* cluster, in acarbose metabolism, e.g. AcbI, AcbJ, AcbQ, AcbD, are unclear. Functions of the limited amount of proteins were characterized through heterologous expression or through biochemical and/or enzymatic studies. Therefore, the main goal of this PhD project is to build a system of genetic tools and methods, that can be used to further clarify the mechanisms and steps of acarbose metabolism. Availability of such system allows for a possibility to gain a reliable genetic proof for the functions of *acb* cluster gene products, which can fulfill the information gap and give the opportunity to formulate complete acarbose metabolism model.

It is important to mention, that existing acarbose producer strains, which were reportedly developed by mutagenesis, seem to reach the limit of possible acarbose production rates by now [110]. In order to develop producer strains via rational design approaches, functioning genetic engineering methods are urgently needed. It would allow not only to improve the amounts of synthesized acarbose, but also to avoid the co-production of other undesired acarviosyl- metabolites. It would greatly facilitate the industrial production of acarbose while omitting lengthy purification steps.

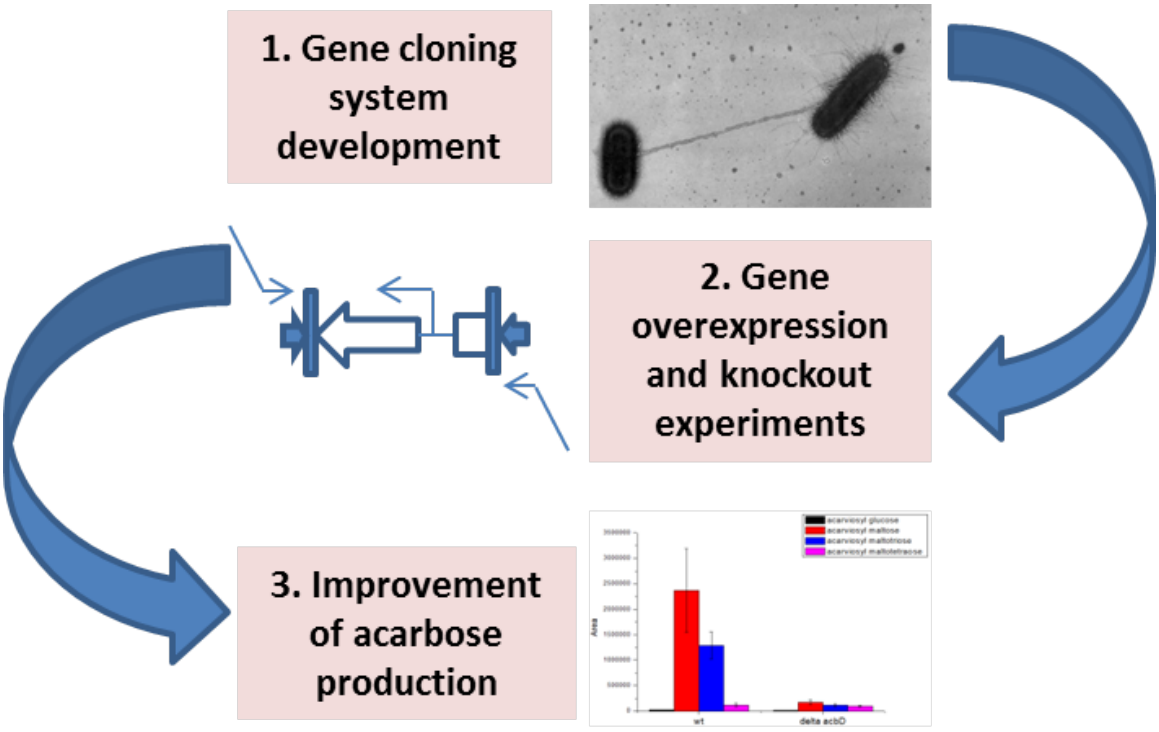

Fig. 7. General overview on the project, presented in PhD thesis.

In order to achieve these goals, experiments can be carried out either in native or heterologous conditions. Both of the approaches are considered and tested in the course of this study. In case of successful development of the native genetic system, several other goals are considered. Among them: investigation of acarbose regulatory mechanisms; targeted correction and enhancement of acarbose biosynthesis through genes overexpression and knockout experiments. Genetic studies are accomplished in close conjunction to proteomic, transcriptomic and genomic studies, which are conducted by PhD students in "Actinoplanes" working group. The general overview regarding the idea and goals of this project can be gained from Figure 7. It can be shortly summarized as follows. The successful development of the gene cloning system or heterologous expression of the *acb* gene cluster is used to perform various genetic experiments, e.g. knockout or overexpressions of *acb* cluster genes. The results of

these experiments are going to be used in further rational improvement of acarbose production by *Actinoplanes* production strain. As a second line of improvements, manipulation of putative acarbose regulatory genes is considered.

2. MATERIALS AND METHODS

2.1 Strains, media, growth conditions and reagents

To maintain and isolate plasmid DNA, *Escherichia coli* DH5α MCR strain was used as a host. *E. coli* ET12567 pUZ8002 (*dam*-13::*Tn*9, *dcm*-6, *hsdM*, *hsdS*) [68] was used for intrageneric conjugation with *Actinoplanes* sp. SE50/110 (ATCC 31044) and all *Streptomyces* strains, utilized in heterologous expression of *acb* gene cluster. Details regarding *Streptomyces* strains and other heterologous hosts can be found in Table 1. The site-specific integration vectors, pIJ6902 (7.4 kb) and pSET152 (5.7 kb), pSOK804 (5.5 kb), pRT801 (5.2 kb), contain φC31, VWB and φBT1 *int* and *attP* genetic regions respectively. All of the vectors contain *oriT* of RK2, as well as an apramycin resistance gene for selection in actinomycetes and *E. coli* [12]. These plasmids do not contain the replicative functions and can be maintained in recipient strains only in the chromosomally integrated state.

Replicative plasmids pKC1139 (6.5 kb) [68], pKC1218 (5.5 kb) [68] and pSOK101 (7.1 kb) [144] are based on pSG5, SCP2* and pIJ101 replicons respectively. Similar to the integrative vectors, all of the replicative plasmids contain *oriT* of RK2 and an apramycin resistance gene for selection in actinomycetes and *E. coli*. Plasmids pSET152, pSOK804, pRT801, pIJ6902, pKC1139, pSOK101, pKC1218 were received from B. Ostash (Ivan Franko National University of Lviv, Ukraine).

pSETGUS (7.7 kb) vector is based on pSET152, contains an apramycin resistance gene for selection in actinomycetes and *E. coli*, φC31 *int* and *attP* genetic regions and *gusA* gene, cloned under *tipA* promoter. pSETGUS was received from A. Luzhetskyy (Saarbrücken, Germany) [93]. All pGUS-based vectors, which contained promoter regions, introduced into *Actinoplanes* sp. SE50/110 in this study, were received from L. Horbal (Saarbrücken, Germany) [58].

All cultivations of *E.coli* strains were done in LB medium, prepared as described in [68]. The growth conditions were 37 °C and 180 rpm in a GFL shaking incubator 3032 (GFL, Burgwedel, Germany). All *Actinoplanes* strains were grown on soy flour medium (SFM; 20 g/L soy flour, 20 g/L mannitol, 20 g/L agar, tapped water to 1 L; pH adjusted to 8.0 prior to autoclaving; autoclaved twice), oatmeal medium (OM; oatmeal flour 38g/L, agar 10 g/L; tapped water to 1 L; pH adjusted to 8.0 prior to autoclaving; autoclaved twice). SFM plates, used for setting *Actinoplanes-E.coli* or *Streptomyces-E.coli* conjugations, were additionally dried at 37 °C before the experiment to assure absence of water on a plate surface. Minimal, complete, R2, R5 media were prepared as described in [68]. Bennett agar medium was prepared as described in [8]. Liquid cultivations were done in NBS medium (10 g/L glucose, 4 g/L peptone, 4g/L yeast extract, 1 g/L MgSO4·7H_2O, 2 g/L KH_2PO_4, 5.2 g/L K_2HPO_4·3H_2O) and maltose minimal medium (prepared as described in [131]. Liquid cultivations of *Actinoplanes* strains were performed at 28 °C and 140 rpm (in a GFL shaking incubator 3032) in baffled polycarbonate flasks (Corning, Corning, NY, USA). When needed, chloramphenicol (25 μg/mL) or kanamycin sulphate (50 μg/mL) were added as a selection markers for *E. coli* and apramycin sulphate (50 μg/mL) was used for the selection of *E. coli*, *Actinoplanes* sp. SE50/110 and streptomycetes exconjugants.

Soy flour was manufactured by Sobo Naturkost (Cologne, Germany); oatmeal flour was purchased from a local farmer market. Set of paper discs used for determination of antibiotic spectra was purchased from Himedia (India). All chemicals and reagents were purchased from Sigma-Aldrich (St. Louis, MO, USA), VWR (Radnor, PA, USA), and Merck (Darmstadt, Germany), if not stated otherwise.

All restriction endonucleases, used for the verification of plasmids, were purchased from NEB (Ipswich, MA, USA). All polymerase chain reaction (PCR) reactions were set up with Phusion High-Fidelity PCR Master Mix (GC Buffer; NEB, Ipswich, MA, USA).

2.2 DNA constructs and primer sequences, used in this study

List and description of all DNA constructs, made in the course of this work can be found in Table 13.

List and description of all DNA primers, used in this study can be found in Table 12. To create, construct, analyze and picture DNA sequences, Geneious R9 software (http://www.geneious.com [66]) was used. Additionally, to select primers for qRT-PCR, Clone Manager Professional Version 9 (Scientific and Educational Software, Denver (USA)) was utilized. In order to test DNA sequences, e.g. selected primers, BLAST search against genomic sequence of *Actinoplance* sp. SE50/110 at default settings was performed (http://blast.ncbi.nlm.nih.gov/Blast.cgi).

2.3 Creation of constructs via Gibson assembly

Gibson assembly was used to create plasmid constructs, generally as described in [37]. In the preparation for this procedure, vector and insert fragments were synthesized separately via PCR with specially selected pairs of primers Table 12. Than, the purified vector and insert were mixed with the Gibson assembly mixture. Equimolar amount of purified DNA of both vector and insert, with a total volume of 5 μL, were incubated with 15 μL of the Gibson mixture for 1 hour at 50 °C. The Gibson mixture was prepared in-house as described in [37].

2.4 Gel electrophoresis and isolation of DNA fragments from agarose gel

Gel electrophoresis is used to separate DNA fragments after PCR reactions, plasmid isolations or endonuclease reaction. 1% agarose gels were used typically to separate DNA fragments after PCR reaction. Higher percentage gels were used to separate small sized fragments. The DNA fragments are separated for 20 minutes with a tension of 100 mV and a maximum current of 400 mA. All gels are stained in an in-house prepared ethidium bromide solution for 10 minutes. Gels are visualized with

UV light by the gel documentation work station AlphaImager (ProteinSimple, California (USA)). After gel electrophoresis, DNA bands were directly cut from the gel and transferred to 1.5 ml Eppendorf tubes. Further, the DNA fragments were isolated from the gel with the MinElute Gel Extraction Kit by QIAGEN. The DNA was eluted with 10 μL of distillated water. The final DNA concentration was measured with NanoDrop (NanoDrop products, Wilmington (USA)).

2.5 Determination of a spore count

Determination of a spore count was regularly done as described in [99] with minor modifications. Strains were grown on SFM for six days. Spore suspensions were collected using sterile water and filtered through non-adsorbent cotton wool. Then 10-fold dilutions of the suspensions were prepared down to 10^{-6} spore titer. 0.1 ml of all dilutions was usually plated. Colonies on each plate were counted after six days of incubation.

2.6 Preparation of scanning electron images (SEM) of *Actinoplanes* sp. SE50/110 colonies

Single *Actinoplanes* colonies were grown for six days on SFM medium plates. The colonies were cut out with the layer of agar and were mounted on glass cover slips. These were placed in a self-build chamber over 4% osmium tetroxide solution for 12 hours. The fixed specimens were lyophilized, fixed to SEM carriers with double sided tape and covered with a 30 nm layer of gold in a BAL-TEC sputter coater (SCD 005). Prepared colonies were analyzed in an S-450 scanning electron microscope (Hitachi) at 15 kV.

2.7 Determination of antibiotic resistance spectra

Antibiotic disc diffusion method was used to study the antibiotic resistance profiles of *Actinoplanes*, according to the instructions of the disc manufacturer (Himedia, India). Spore suspensions were prepared as described above. Then 10-fold dilutions of the suspensions were prepared down to 10^{-6} spore titer and 0.1 ml of all dilutions was usually plated on SFM plates. Discs were placed in sets of four on freshly plated lawns of *Actinoplanes*. Growth inhibition zones were measured after five days of incubation.

To define specific survival rates of *Actinoplanes* sp. SE50/110 on selected antibiotic concentrations, freshly prepared spore suspensions were used to make stepwise dilutions as described above and 0.1 ml of all dilutions was usually plated on SFM plates with and without antibiotics. Colonies on each plate were counted after six days of growth and reference (100%) survival percentage was calculated from control plates (absence of the antibiotic).

2.8 Preparation, transformation and regeneration of protoplasts of *Actinoplanes* sp. SE50/110. Electroporation of mycelia and protoplasts of *Actinoplanes* sp. SE50/110

Electroporation of both, protoplasts and mycelia cells, was conducted in this work. Mycelia cells were grown in resource-rich NBS liquid medium to logarithmic growth phase (14-18 hours, since the beginning of cultivation). Then, the protocol, established for *S. coelicolor* and *S. lividans* was followed [68]. In this protocol, 2 kV electric pulse (10 kV/cm $^{-1}$) is applied. Different field strengths ranging from 5 to 12.5 kV/cm $^{-1}$ were tested as described in [81] for *Nocardia* sp. CS682. Subsequently, plates were incubated for 5 days at 28 °C and then analyzed.

Protocols for protoplasts preparation, fusion and regeneration were taken from [85, 68]. For protoplast generation, cells were grown in V, VM or YEME media [85, 68]. After harvesting mycelia from cultivatin solution, procedure, developed

in [85] for protoplast generation was generally followed. It involves treatment of mycelia with a combination of enzymes, hen egg-white lysozyme (HEWL) and *S. globisporus* mutanolysin in final concentrations 5 and 0.018 mg/mL respectively. To avoid formation of foam, non-ionic detergent Pluronic was added to a growth media in the final concentration of 100 mg/mL. Phase contrast microscope was used to examine the progress of protoplast formation. Protoplasts were frozen at -80 °C or directly used for electroporation.

2.9 *E. coli* – *Actinoplanes* and *E. coli* – *Streptomyces* conjugation procedures

Protocols, described in [68, 59, 48] were used for the development of *E. coli* – *Actinoplanes* conjugation protocol. For setting conjugation matings, freshly grown plates of *Actinoplanes* sp. SE 50/110 (SFM media, 5-6 days, 28 °C), which were showing signs of sporulation, e.g. change of a lawn colour from deep orange to white, were used. Surface of plates was washed with 2.5 mL of sterile water or LB and carefully scratched with sterile cotton-tipped swabs. The spore/mycelia suspensions of all plates were mixed and separated as 1 mL of suspension per one sterile tube. Culture of the donor *E. coli* ET12567 (pUZ8002), containing different vectors, was grown in the presence of 50 mg apramycin/L, 25 mg chloramphenicol/L and 50 mg kanamycin/L to an OD_{600} of 0.4 – 0.6. To remove the antibiotics, cells were washed once with an equal volume of sterile water or LB, and finally resuspended in 1 mL of LB. Both host and donor cells were centrifuged shortly at 8,000 x g for two minutes. The supernatant was discarded and the remained pellets were mixed in correlation *Actinoplanes*:*E. coli* = 2:1. Mixtures were carefully resuspended by gentle pipetting up and down and plated on fresh SFM plates. SFM plates, used for conjugation were usually additionally dried for several hours before usage at 37 °C. After 20 - 22 hours of incubation, plates were overlaid with 1 mL of sterile water, containing 1 mg of apramycin and 1 mg of fosfomycin and allowed to dry. After seven days of further

incubation *Actinoplanes* colonies were transferred to fresh SFM plates, supplemented with apramycin. Subsequently, exconjugants were purified from *E. coli* ET12567 with the help of two approaches. Single selected *Actinoplanes* colonies were repeatedly re-transferred to new SFM plates until no *E. coli* cells could be visually observed on the surface. Further, single colonies of *Actinoplanes* were used for inoculation in NBS liquid medium. If still present, the faster growing *E. coli* would overgrow *Actinoplanes* sp. SE50/110, which could be easily detected through visual inspection. In another approach, single sporulating *Actinoplanes* colonies were inoculated in liquid maltose minimal media, supplemented with 50 mg/L of apramycin, fosfomycin and acarbose and incubated at 28 °C for 48 hours (160 rpm). 0.1 mL of grown cultures was used for step wise dilution with sterile water down to approximately 10^{-6} cell count and plated on a series of SFM plates, which were further incubated for 4-6 days. Single colonies of *Actinoplanes* exconjugants were selected and inoculated in NBS media to check the absence of *E. coli* cells as described above. Purified colonies of *Actinoplanes* were deposited as glycerol cultures for long time storage at -80 °C.

The frequency of exconjugant occurrence was calculated as a ratio of the number of exconjugants to the titer of recipient strain spores, as described in [94].

All *Escherichia coli – Streptomyces* conjugations were set as described in [68].

2.10 Screening of *Actinoplanes* colonies after conjugation and their characterization

After conjugation, selected colonies of *Actinoplanes* exconjugants were tested with the help of PCR. For this purpose, genomic DNA was isolated and two sets of primers were used, binding to internal regions of *aac3(IV)* gene and *oriT* region of all vectors respectively (Table 12).

PCR products were purified with QIAquick Gel Extraction Kit (Qiagen) and submitted for in-house Sanger sequencing. The phenotypic characterization of *Actinoplanes* exconjugants was done by plating them on SFM plates.

The stability of plasmid inheritance was studied as described in [59] with minor modifications. After 5 passages at 28 °C under non selective conditions the exconjugants were plated onto SFM and allowed to sporulate. The spores were harvested and plated out in the presence and absence of apramycin. For each type of exconjugant, 300 colonies were checked for apramycin resistance.

2.11 Screening of *Streptomyces* M1152+ and M1154+ colonies after conjugation and their characterization

After conjugation at lest nine colonies of each exconjugant strain were selected and transferred to a new SFM plate with apramycin. All colonies were tested with colony PCR, using primers to internal regions of *acbD* gene (Table 12). PCR products were purified with QIAquick Gel Extraction Kit (Qiagen) and submitted for in house Sanger sequencing. Purified colonies of all verified *Streptomyces* exconjugant strains were deposited as glycerol cultures for long time storage at -80 °C.

2.12 HPLC measurements of acarviosyl-containing metabolites

For determination of acarbose biosynthetic activity, all strains were grown in triplicates in parallel (liquid minimal media, 140 rpm). Liquid minimal media with different Carbon sources was inoculated with freshly prepared spore suspensions. For this purpose, firstly *Actinoplanes* sp. SE50/110 was plated at SFM and incubated for six days until showing clear signs of sporulation. Spore suspensions, prepared as described above, were mixed and used for inoculation in correspondence of 1 mL per flask. Every 48 hours, 1 mL of growing cultures was collected and centrifuged down at 20,000 x g for 2 minutes. Supernatant was carefully separated from the pellet and stored at -20°C. The rest of the pellet was washed twice with distillated water and dried at 70°C. Dried pellets were weighed and used to estimate a cell dry weight. Acarbose measurements were carried out via high performance liquid chromatography (HPLC) as described in [131].

2.13 LC-MS measurements of acarviosyl-containing metabolites and their analysis

LC-MS measurements of acarviosyl metabolites were done following protocol, described in [132] with modifications. For all LC-MS based metabolite measurements, 10mL of *Actinoplanes* sp.SE50/110 or *Streptomyces* M1152+ and M1154+ supernatants were harvested by centrifugation at 14,000×g for 15 min at -4°C (Eppendorf, Hamburg, Germany). Metabolites were extracted by solid phase extraction using 3 mL Chromabond Easy columns (Macherey–Nagel, Düren, Germany). In order to calibrate them, columns were conditioned with 3 mL of methanol and 3 mL of water subsequently. Then, 9 mL of each supernatant was applied, followed by washing twice with mixture of 3 mL water/methanol (95:5). Acarviosyl metabolites were eluted using 3 mL of methanol. For LC–MS measurements 10 μL of each sample were injected into the LC–MS system. LC-ESI-QTOF conditions of all conducted measurements and data processing conditions were identical to [132]. Data, which was obtained in measurements, was analyzed using the Compass software (Bruker Daltonics, Bremen, Germany). Identification and relative quantification of metabolites was conducted by targeted generation of base peak chromatograms for expected m/z-values and subsequent peak integration.

2.14 Plasmid and genomic DNA isolation

For plasmid isolation *E. coli* DH5α cultures, carrying respective plasmids, were cultivated in LB media at 37 °C for 12 hours. Further on, GeneJET Plasmid Mniprep Kit (Thermo Fisher Scientific) was used accordingly to manufacturer advises. Isolated plasmid DNA was dissolved in water and stored at -20°C. For the isolation of genomic DNA a conventional method described in [68] was used.

2.15 DNA sequencing, sequence and insertion analysis

For library preparations of three derived *Actinoplanes* exconjugants, a Nextera DNA Sample Prep Kit (Illumina, San Diego, CA, USA) was used. Resequencing was performed on Illumina MiSeq System (Illumina, San Diego, CA, USA) using the MiSeq Reagent v3 Kit (Illumina, San Diego, CA, USA). The paired-end sequencing run produced reads with a length of 2x300 bp. In the next step, the read quality was checked with the program FASTQC [6]. After sequencing and processing of the raw data, *de novo* assemblies were performed using the GS De Novo Assembler Version 2.8 (Roche) with default settings. Sequencing of *Actinoplanes*Δ*acbD* and *Actinoplanes*Δ*cadC* strains was performed in a similar manner. To identify the vector sequences of pSET152, pSOK804 and pRT801, the contigs were filtered by applying r2cat (Husemann & Stoye, 2009) and *Actinoplanes* sp. SE50/110 (Genbank Acc. No. CP003170) as reference genome. Identified vector sequences were imported into the software platform GenDB [89] for automatic annotation. Results were manually refined as described recently in [28, 135]. The specific integration sites of vector plasmids were determined by applying *in silico* finishing [135]. All sequencing data were deposited at the European Nucleotide Archive (ENA) under Bioproject PRJEB12637. The vector sequences of pSOK804 and pRT801 were deposited in Bioproject PRJEB12637. For all pairwise global sequence alignments EMBOSS Needle software was used [88]. The alignment was created with MUSCLE 3.8.31 program, additional analysis was done by software described in [24].

2.16 Isolation of RNA and assessment of its quality

The same protocol was used for RNA isolation from all *Actinoplanes* and *Streptomyces* strains. RNA isolation was done with the use of RNeasy mini kit (Qiagen, Hilden, Germany), following the directions of manufacturer. Samples for RNA isolation were collected from growing cultures at exponential phase of growth, for each particular cultivation depending on the growth curve, generally at 48 or 72 hours of

growth in three biological and two technical replicates. Frozen cell pellets were suspended in 800 μL RLT buffer and transferred to 2 mL lysing matrix tubes (0.1 mm spherical silica beads, MP Biomedicals, Santa Ana, California, USA). The disruption of cell suspension was carried out in a homogenizer (FastPrep FP120, Thermo Fisher Scientific, Waltham, MA, USA) for two times for 20 seconds at speed setting 6.5. In between of two homogenization steps, samples were incubated on ice for 1 minute. The cell suspension was immediately centrifuged for 3 minutes at 13,000 x g and 4 °C. The supernatant was used for RNA extraction using a Qiagen RNeasy mini kit in combination with an RNase-free DNase kit (Qiagen, Hilden, Germany) for the purification of RNA samples from residual DNA. In order to test for the presence of DNA in RNA samples, PCR reaction with primers binding to the genomic regions of *Actinoplanes* sp. SE50/110 was used. Quality and quantity of the RNA was analyzed with a NanoDrop 1000 spectrometer (Peqlab, Erlangen, Germany) and an Agilent RNA 6000 Pico kit run on an Agilent Bioanalyzer 2100 (Agilent Technologies, Santa Clara, CA, USA).

2.17 DNA microarrays

DNA microarrays performed according to the protocol developed by Dr. Armin Neshat and Timo Wolf. The description of the protocol, quoted here, originates from unpublished research paper of T. Wolf and J. Droste (Droste and Wolf, 2017; submitted to BMC Genomics on 26.01.2017). Whole genome oligonucleotide microarrays were designed with eArray (Agilent Technologies, Santa Clara, CA, USA) and subsequently ordered from Agilent Technologies, Santa Clara, CA, USA. All steps, including sample preparation, cDNA synthesis and labeling, microarray hybridization and washing, scanning and feature extraction, were carried out as described in the manual by the manufacturer (Agilent, 2012). Two-Color Microarray-Based Prokaryote Analysis FairPlay III Labeling (Version 1.4, Agilent) kit was used with the several adjustments, which were originally tested and proposed by Dr. Armin Neshat. The

mix for hybridization was prepared with 330 ng of each labeled cDNA and 11 μL gene expression blocking agent. The cDNA blocking mix was filled up to 55 μL with water and mixed with 55 μL Hi-RPM hybridization buffer. 100 μL of the hybridization mix were used for the hybridization of each array (quoted here, as described in unpublished research paper Droste and Wolf, 2017; submitted to BMC Genomics on 26.01.2017). Washing of the microarrays was performed with water and acetonitrile, as described in a manual. As an adjustment, two wash cycles were performed instead of one. In addition, sterile distilled water was used to wash the racks, magnet stirring bars and other equipment. The slides were scanned in the Agilent DNA Microarray Scanner. Feature extraction was performed with the Agilent Feature Extraction Software Version 10.7.3.1 (Agilent Technologies, Santa Clara, CA, USA). The analysis of the microarray data was performed with EMMA2 [26]. During analysis of data, a p-value of 0.05 was used as a cut-off for significance. The significance of results was further confirmed with qRT-PCR, when needed.

2.18 Reverse transcription quantitative PCR (qRT-PCR)

qRT-PCR was applied for relative mRNA quantification of single genes. For qRT-PCR, three biological and two technical replicates of each sample are measured. RNAse-free water is used as the negative control in these measurements. Primers were designed to amplify 100 to 150 bp of intragenic regions of genes of interest. A list of all primers used for qRT-PCR tests can be found in Table 12. A SensiFast SYBR No-Rox One-Step Kit (Bioline, London, UK) and 96 well lightcycler plates (Sarstedt, Nümbrecht, Germany) were used for measurements in a LightCycler 96 System (Roche, Mannheim, Germany). All RNA samples were diluted to 200 μg/μL prior to measurement. 1 μL of each RNA sample was mixed with 19 μL mix containing 1 μL of primers, 0.2 μL reverse transcriptase, 0.4 μL RNase inhibitor, 10 μL reaction mix and 7.4 μL 5M Betain. The program was designed as follows: reverse transcription was performed at 45 °C for 20 minutes, followed by 2 minutes at 95

°C, three step amplification (95 °C 5 s, 60 °C 10 s, 72 °C 10 s, 60 cycles). Melting profile was generated in process of measurement. The LightCycler 96 V1.1 software was used for the analysis of results. At first, inspection of control measurements and melting curves was done. The relative RNA amount was normalized on total RNA (200 ng) and calculated as 2-ΔCq. ΔCq is calculated as the difference of the mean Cq in the mutant strain compared to the control strain.

2.19 RNA-sequencing

All RNA samples are tested for the presence of DNA with the help of PCR and quantified with Agilent 2100 Bioanalyzer. Ribosomal RNA was removed with the help of the Ribo-Zero Magnetic Kit from Illumina Inc. (San Diego, USA) as described in the protocol Part 15065382 RevA from 2014. Ribosomal RNA-depleted samples are further purified by ethanol precipitation according to the advises, listed in the protocol from the Ribo-Zero Magnetic Kit. The RNA pellet is air dried at room temperature for 5 minutes. Library preparation is performed according to the manufacturer's protocol (Illumina, 2012). All ingredients are provided by the Illumina TruSeq Stranded mRNA LT Sample Prep Kit from Illumina (San Diego, USA). The quality of the library is further examined with Bioanalyzer Agilent Technologies DNA-chip. The library is diluted to 4 nM and afterwards used for sequencing. RNA sequencing is carried out with the Illumina MiSeq platform.

2.20 Plate-based assessment of GUS activity

Assay was performed as described in [93]. Strains, carrying various constructs, based on pGUS and pSETGUS, were grown for six days on SFM plates at 28°C. To detect the GUS activity directly on plates, we used the chromogenic substrate XGluc, which forms a blue precipitate, 5,5-dibromo-4,4-dichloro-indigo, after hydrolysis by GUS enzyme. Surface of each colony tested was carefully scratched with sterile toothpick and a drop of 0.1 M XGluc solution was placed on the top center of the scratched

surface. Plates were incubated in darkness at a room temperature for 30 minutes, until the formation of dark blue colour.

2.21 Redirect-related methods

The cosmid library, used to create ReDirect-based gene knockouts, was received from Dr. Paul-Bertram Kamp with the help of Dr. Susanne Schneiker-Bekel. The cosmid clones were constructed using the CopyControltm Cloning System of Biozym Scientific GmbH (Hessisch Oldendorf, Germany) on the base of pCC1BAC vector. The Cloning-Ready Vector was cleaved at Eco72I site; *E. coli* TransforMay EPI300 cells were used for transformation. More information regarding CopyControl^{TM} Cloning System can be found under www.biozym.com.

E.coli BW25113/pKD46 strain (F-, δ(araD-araB)567, δlacZ4787(::rrnB-3), λ-, rph-1, δ(rhaD-rhaB)568, hsdR514, pKD46) was received from Yale CGSC Coli genetic stock center. All ReDirect procedures were carried out as described in [46].

2.22 Phenotypic microarrays of BioLog system

The set up, execution of the experiment and analysis of the data was done as described in [113] and http://www.biolog.com/. For the analysis, each strain was grown on SFM media for six days, until clear signs of sporulation were visible. The surface of each plate was washed with 3 mL of sterile water and spore suspension was collected. Further, spore suspensions were diluted with water, measured on turbidimeter and transferred to the PM plates. PM1, PM2A, PM3B, PM4A plates were tested in two technical and three biological replicates for each strain. OmniLog incubator/reader OmniLog PM Software were used as up to advise of manufacturer.

2.23 Analysis of the extracellular proteome of *Actinoplanes*sp. SE50/110 and its derivatives with gel-based proteomics

All analysis was following the methods described and developed in [131] and [97]. Supernatants from 50 mL cultures of *Actinoplanes* strains were obtained by separating cells by centrifugation. After the separation supernatants were lyophilized and proteins were extracted with phenol, precipitated with methanol and washed with 70 v/v% ethanol. Purified proteins were rehydrated in 8M urea and 0.2% CHAPS and quantified with a modified Bradford assay (modified from Bradford, 1976) using Roti-Nanoquant (Carl Roth, Karlsruhe,Germany) as described in [131]. For isoelectric focusing 450 μg of protein was used. IPG strips (Immobiline DryStrips, GE Healthcare, Solingen, Germany) were than used with previously prepared proteins. For the separation in the second dimension a ProteanII IXL Cell (Biorad Laboratories, München, Germany) and 10.5% Tris-tricine polyacrylamide gels were used. Proteins were stained with Coomassie and digested with trypsin. Protein digests were measured with MALDI-TOF-MS. Proteins were identified with PMF and MS/MS for which the ultrafleXtreme mass spectrometer (Bruker Daltonics, Bremen, Germany), the BioTools (Bruker Daltonics, Bremen, Germany) software and the Mascot algorithm were used [97].

3. RESULTS AND DISCUSSION

3.1 Heterologous expression of *acb* gene cluster in a set of specially designed host strains

Heterologous expression of secondary metabolite clusters is thought to be the fastest and the most efficient way to analyze biosynthetic pathways and to increase respective metabolite production [38, 39]. Heterologous expression of biosynthetic clusters also may be necessary when the natural host is poorly accessible by genetic means or has a prolonged growth period [14]. Earlier, an attempt to express *acb* cluster in *Streptomyces lividans* TK23 and *S. lividans* 1326 was made [121, 143]. As a result, acarbose production by *S. lividans* transformants was detected in small amounts, possibly due to the absence of the essential regulatory mechanisms. Interestingly, it was shown, that acarbose production in heterologous conditions can be induced by maltooligosaccharides. This finding could hint at the regulator gene, which is located outside the cluster and which is present in both *Actinoplanes* and *S. lividans* ([129]; unpublished results, personal communication). As long as acarbose production was not achieved at the satisfactory levels, we decided to modify and repeat the experiment carried out by Dr. Holger Thomas.

Typically, such experiments consist of several steps: 1) cloning of a gene cluster in a suitable transfer vector; 2) selecting an appropriate host strain for expression; 3) introducing and propagating the gene cluster in the chosen host; 4) culturing the recombinant strain under the appropriate growth conditions; 5) analyzing the biosynthetic activity of the host with and without introduced cluster.

3.1.1 Transfer of *acb* cluster through *E. coli* - *Streptomyces* conjugation and verification of exconjugant strains

The first step of the experiment is significantly simplified since *acb* cluster (approximately 32 kbp) has been already cloned into a cosmid cos6.2 [129]. Firstly, standard in house Sanger DNA sequencing of the cosmid was performed, in order to check if all *acb* genes are intact. No significant mutations were found within cluster. The cosmid by itself is based on the vector pOJ436 (11,1 kb), it contains gene of an apramycin resistance, lacZα with MCS for cloning DNA, φC31 phage *int* gene, *attP* sequence and therefore is able to integrate site-specifically at the φC31 attachment sites in the host genome [12]. At the moment, there is a broad choice of possible heterologous hosts within *Streptomyces* genus. Usually, they possess such traits as: simplicity of genetic manipulations, fast and dispersed growth, high sporulation rate, lack of their own secondary metabolites biosynthesis mainly due to the large genome deletions. In this study, we decided to use a set of several strains with various changes in a primary and secondary metabolism. The short overview of the strains, which were used in this experiment is given in Table 1.

Heterologous host	Distinctive characteristics	Source
Streptomyces coelicolor M1146	1) clusters of actinorhodine, undecylprodigiosyn, Ca-dependent antibiotic were removed from the genome. 2) cloning techniques readily available; 3) doesn't contain marker genes;	Prof Mervyn Bibb, [38]
S. coelicolor M1152	1), 2), 3) same as above; 4) harbors point mutation of *rpoB* gene, that is proven to enhance antibiotic production.	Prof Mervyn Bibb, [38]

Heterologous host	Distinctive characteristics	Source
S. coelicolor M1154	1), 2), 3) same as above; 4) harbors point mutation of *rpoB* and *rpsL* genes, that are proven to enhance antibiotic production.	Prof Mervyn Bibb, [38]
S. coelicolor ΔpfkA*2*	Harbors deletion of gene coding for phosphofructokinase, that results in increased Carbon flux through the pentose phosphate pathway.	Dr Jeroen Siebring, [15]
S. venezuelae ATCC10712	1) fast growth and sporulation (within 2 days); 2) cloning techniques are readily available; 3) doesn't contain marker genes.	Ivan Franko National University of Lviv (IFNUL), [119]
S. thermospinisporus NRRL-B24318	1) is able to grow at high temperatures ($\geq$40°C); 2) carboxydotrophic (can utilize *CO* and CO_2 as sole carbon sources).	Ivan Franko National University of Lviv (IFNUL), [70]
S. avermitilis SUKA4	All biosynthetic gene clusters for endogenous major metabolites are removed from the genome.	Prof Haruo Ikeda, [72]

Table 1: List of heterologous hosts, used for *acb* gene cluster expression.

S. coelicolor is the best genetically characterized actinomycete, which genome has been sequenced [11] and annotated to a high standard [1]. The large array of tools are available for genetic manipulations in this strain [68, 29, 30, 46]. *S. coelicolor*

M145 was used for the construction of M1146, M1152 and M1154 strains, developed specifically for the heterologous expression of secondary metabolite gene clusters. The actinorhodin, prodiginine, calcium dependent antibiotic (CDA), and the cryptic Type I polyketide (*cpk*, SCO6269–6288) gene clusters were deleted from both strains, thus removing all detectable antibiotic activity and reducing potential competition for precursors between endogenous and cloned biosynthetic pathways. Both M1152 and M1154 contain the *rpoB* (encoding the β-subunit of RNA polymerase) mutation, while M1154 contains additionally *rpsL* (encoding ribosomal protein S12) mutation reported previously to increase levels of secondary metabolite production. Also, other derivatives of *S. coelicolor* were used in this study. These are the strains with the changes being introduced into the primary metabolism. Secondary metabolites are synthesized in dedicated biosynthetic routes, but precursors and co-factors are derived from the primary metabolism. High level production of antibiotics in streptomycetes therefore usually requires engineering of the primary metabolism. Deletion of *pfkA*2 (SCO5426) (putative phosphofructokinase) in *S. coelicolor* resulted in a higher production of the pigmented antibiotics actinorhodin and undecylprodigiosin. It was shown, that the *pfkA*2 deletion strain had an increased carbon flux through the pentose phosphate pathway, largely because of accumulation of glucose 6-phosphate and fructose 6-phosphate. Furthermore general genome-scale metabolic model simulations, that were made in this study [15], have shown that decreased phosphofructokinase activity leads to an increase in pentose phosphate pathway flux and in flux to pigmented antibiotics and pyruvate. It is already known, that the precursor of acarbose, 2-epi-5-epi-valiolone is obtained from the cyclization of sedoheptulose-7-phosphate, which is directly supplied by a pentose phosphate pathway [116]. Consequently, the increased Carbon flux through the pentose phosphate pathway should lead to the increased acarbose production by transformant strains. *S. venezuelae* has recently entered the arena of *Streptomyces* genetics as a novel model for fundamental studies and heterologous expression experiments. This strain can be characterized by fast, dispersed growth under broad range of conditions, and simplicity of genetic manip-

ulations [105]. *S. venezuelae* was already used for the expression of one of the most famous aminoglycosides, kanamycin, which cluster was cloned from *S. kanamyceticus*. The obtained transformant strain was characterized as a fast and effective kanamycin producer in a comparison to the wild type strain [119]. Finally, thermophilic carboxydotrophic *S. thermospinisporus* (NRRL-B24318) was used in effort to evaluate a possibility to use a chemolitotrophic bacteria for antibiotic production [70].

3.1.2 Characterization of biosynthetic patterns of *S. coelicolor* M1152 and M1154 exconjugant strains in maltose, maltotriose and maltose/maltotriose liquid minimal media

For the acarbose biosynthesis measurements, M1152, M1154, M1152+, M1154+ and *Actinoplanes* sp. SE50/110 strains were grown in parallel in a maltose-, maltotriose- and maltose/maltotriose- containing liquid minimal media. Samples to measure acarbose production were taken at 48, 72 and 96 hours in triplicates. Measurements were carried out with via high performance liquid chromatography (HPLC) and liquid chromatography-mass spectrometry (LC-MS). All samples of analyzed exconjugants have shown no sign of acarbose production, while wild type of *Actinoplanes* sp. SE50/110 was characterized by typical amounts of acarbose, e.g. 0.16±0.02 g/L of acarbose at 72 hours of cultivation (Figure 8) in case of maltose minimal media (Mal-MM) cultivations.

To reach a higher level of sensitivity, all samples were further analyzed by LC-MS method. However, still no trace of acarviosyl-maltose was found in all collected M1152+ and M1154+ samples (Figure 9). As a control in all measurements wild types of M1152, M1154 and *Actinoplanes* sp. SE50/110 were used.

To verify whether *acb* gene cluster is expressed in cells of *S. coelicolor* strains, RNA was isolated from M1152+, M1154+ and *Actinoplanes* sp. SE50/110 samples grown in Mal-MM, MT-MM or maltose/maltotriose minimal media at 72 hours of cultivation. As a verification method quantitative reverse transcription PCR (qRT-

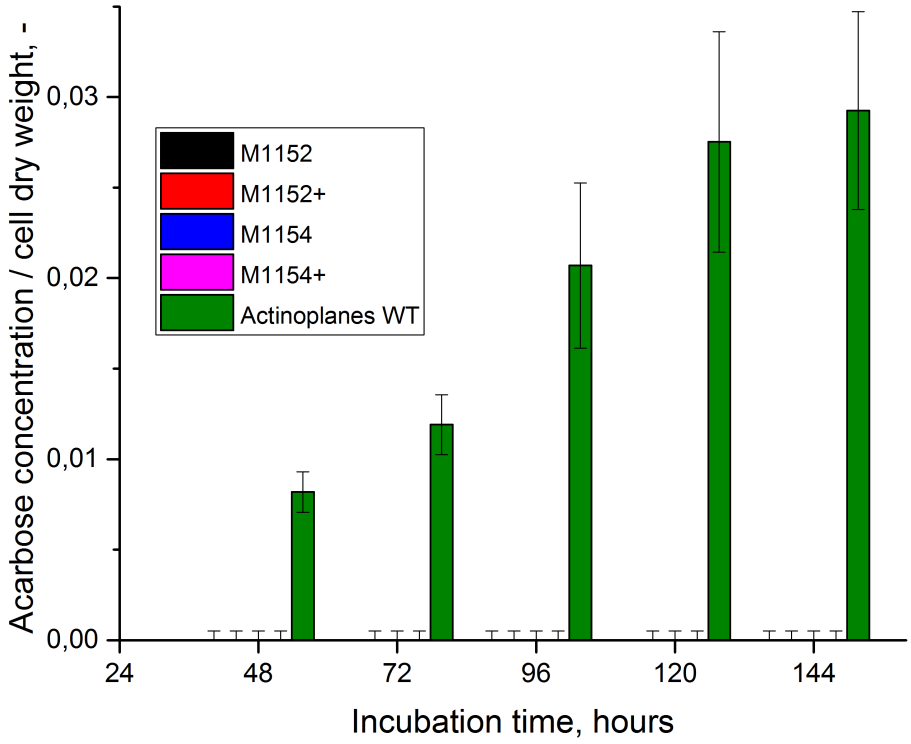

Fig. 8. HPLC measurements of acarbose production by *Actinoplanes* sp. SE50/110 (*Actinoplanes* WT), M1152, M1152+, M1154, M1154+ all grown in parallel in Mal-MM. The calculated concentration of acarbose is normalized to the respective detected dry weight for each time point and strain respectively. Mean values and standard deviations are shown for three biological replicates.

PCR) with primers to genes *acbD*, *acbA* and *acbE* (Table 12) was performed. No quantifiable amounts of RNA of respective genes have been found in all samples tested. These findings point out, that most probable reason for the deficiency in acarbose production in M1152+, M1154+ lies in the non-expression of *acb* gene cluster.

In order to better understand above mentioned results, several explanations have to be taken into account: 1) *Actinoplanes* sp. SE50/110 is genetically similar to

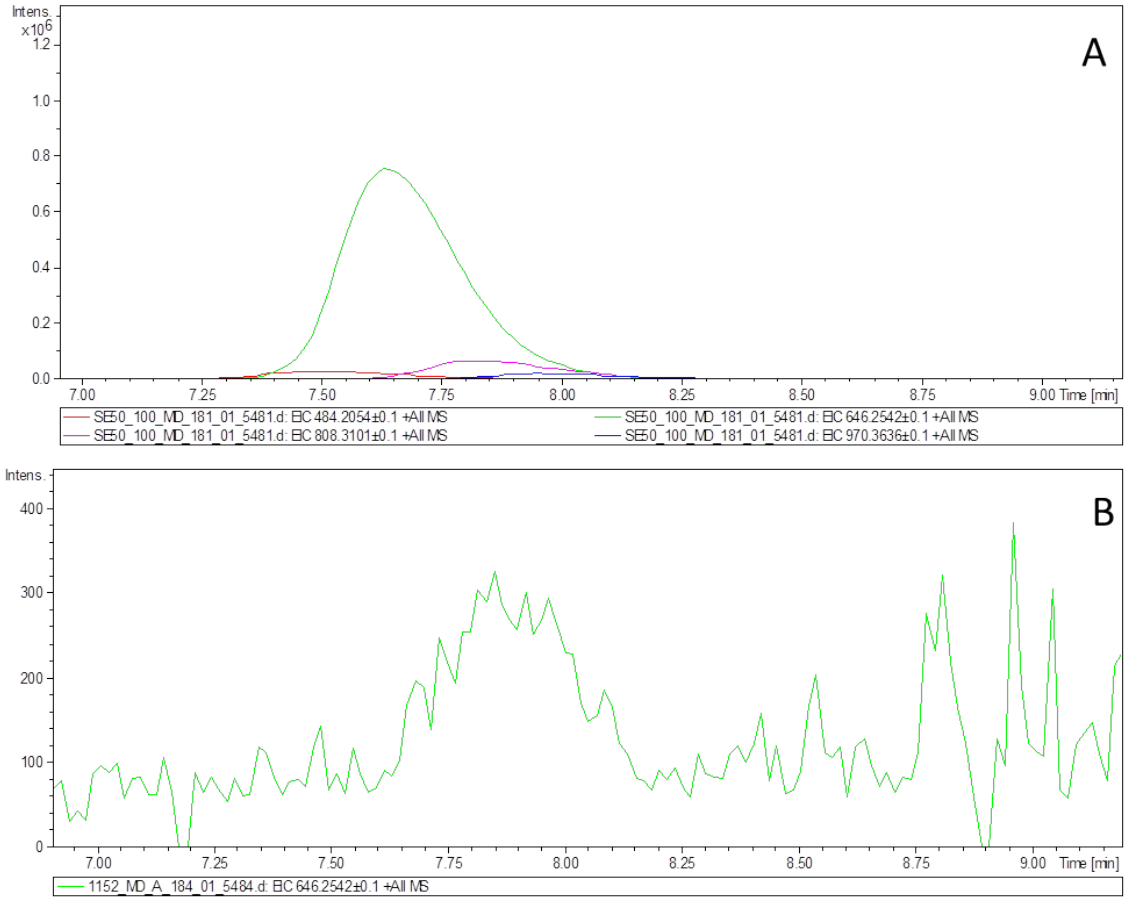

Fig. 9. Extracted ion chromatograms of detected acarviosyl metabolites in the culture supernatants of *Actinoplanes* wild type (A) and 1152+ strain (B), both cultivated in Mal-MM. Samples are taken in exponential growth phase at 72 hours of cultivation. All samples are measured in three biological replicates. Components are depicted as follows: acarviosyl-glucose in orange, acarviosyl-maltose in green, acarviosyl-maltotriose in pink and acarviosyl-maltotetraose in blue. M/z ratios: acarviosyl-glucose 484.2054 +/- 0.1, acarviosyl-maltose 646.2542 +/- 0.1, acarviosyl-maltotriose 808.3101 +/- 0.1, acarviosyl-maltotetraose 970.3636 +/- 0.1.

S. coelicolor, both belong to class Actinobacteria. However, the possibility exists, that the later lacks crucial genes, necessary for acarbose production, e.g. coding for putative regulators of acarbose metabolism; 2) *acb* cluster, cloned into cosmid, could

lack genes, responsible for acarbose biosynthesis or transport, therefore leading to the abolishment of acarbose production; 3) conditions, required for acarbose biosynthesis were not reached in this setup, e.g. additional adjustments of cultivation media, temperature have to be made.

Generally, heterologous expression of secondary metabolite clusters is regarded as fast and effective way to obtain high levels of production of desired metabolite [38, 39]. It has proven to be highly successful in a range of application cases [119, 141]. However, in some of them, due to the various reasons, this methodology failed or turned out to be not applicable. Instead of following the route of further adjustment of heterologous expression, which could potentially be time consuming and inefficient, we decided to switch to the development of the native gene cloning system for *Actinoplanes* sp. SE50/110. It is widely recognized, that genetic engineering work in case of members of genus *Actinoplanes* is complicated to conduct and for many strains was never established [126, 129]. From the other side, once established, these techniques give a possibility to answer a wide range of questions regarding strain growth, metabolism and morphology, to conduct any type of genetic experiment, gene overexpression or gene knockout, of one or several targets simultaneously. Moreover, it enables the development of stable acarbose overproduction strains, which are rationally designed and suitable for industrial use. Till now, industrial derivatives of *Actinoplanes* sp. SE50/110 were constructed without use of genetic engineering with other techniques, which seem to reach their limitations.

3.2 Development of genetic engineering toolkit for *Actinoplanes* sp. SE50/110

Typically by the development of the "gene cloning system" researchers mean, the establishment of all necessary tools and conditions to perform genetic engineering experiments in a particular strain. To understand better, what is usually hiding behind this term lets break it into necessary steps, which will be followed in the next chap-

ters: 1) selection of the solid growth media; 2) search for marker genes, suitable for selection in particular conditions (antibiotic resistance genes); 3) choice of the suitable DNA transfer system; 4) search for integrative and replicative vectors; 5) application of the developed system via gene knockouts, overexpressions and complementations (Figure 10).

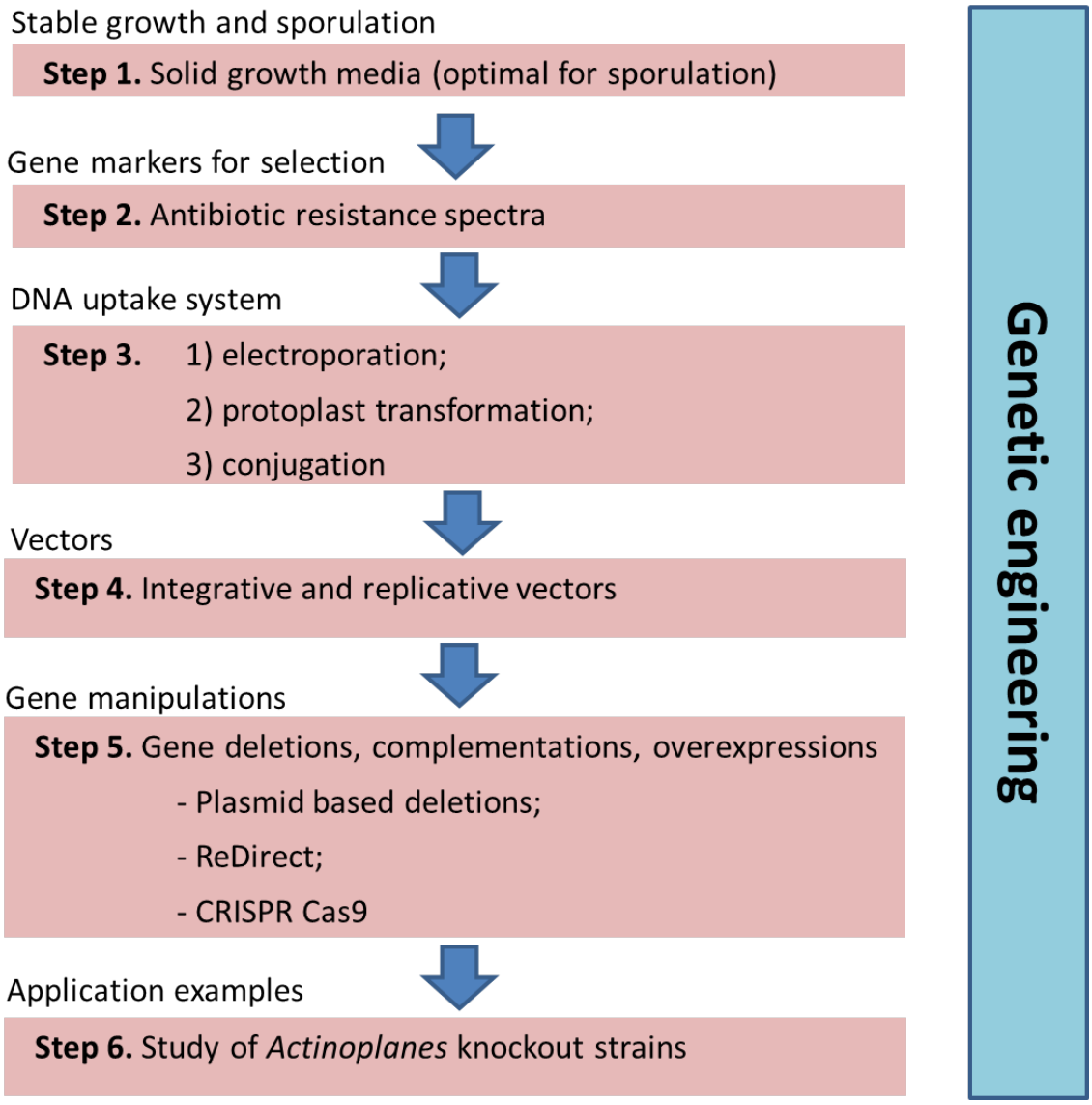

Fig. 10. Schematic representation of main steps, involved in the development of gene cloning system for *Actinoplanes* sp. SE50/110.

3.2.1 Selection of appropriate culture conditions for growth and sporulation in *Actinoplanes* sp. SE50/110

Data, described in section 3.2 and partly in sections 3.3 and 3.4 of "Results and Discussion" chapter was published in [43] and [137].

In order to establish gene cloning system, replicable and stable growth of the target strain is needed. Therefore, as the first step in developing gene cloning system for *Actinoplanes* sp. SE50/110, various solid media were tested for suitability for its growth and sporulation. High amount of spores is known to be crucial for the successful conjugation procedure, which is one of the clear options for DNA transfer to *Actinoplanes* sp. SE50/110 and is going to be implemented in next experiments. For following tests, several different media were selected commonly being used for Actinobacteria (Table 2) [68, 8].

Media	MM	CM	R2	R5	Bennett agar	LB	SFM	OM
Growth	-	-	+/-	+/-	+	+/-	+	+
Sporulation/ Sporulation rate	-	-	-	-	-	-	+/− $2 \pm 0.5 \times 10^7$	+/− $3 \pm 0.5 \times 10^6$

Table 2
List of solid media, tested for *Actinoplanes* sp. SE50/110 growth and sporulation. MM – minimal medium, CM – complete medium, LB – Luria-Bertani medium, SFM – soya flour medium, OM – oatmeal medium [68, 8]; "+" – "strong" growth, "+/-" "weak" growth "-" – no growth.

These in detail were: minimal, complete, R2, R5, LB and Bennett media as defined media and soya flour (SFM) and oatmeal medium (OM) with a non-defined composition. Among these media, only SFM and OM were found to be suitable for both a sustainable strain growth and sporulation. Notably, *Actinoplanes* sp.

SE50/110 was not able to grow on the other solid media tested, even on R2 and R5 media, especially designed for obtaining and regenerating protoplasts. Interesting observation was made regarding growth of *Actinoplanes* sp. SE50/110 on Bennett agar media: actively growing colonies were surrounded by the white halo, which increases in its diameter proportionally to the size of the colony during incubation time. It could hint on other secondary metabolites or extracellular substances being produced and secreted by strain. To verify whether *Actinoplanes* sp. SE50/110 produces acarbose and/or other secondary metabolites, while being grown on SFM, OM and Bennett agar media, a simple test was made with the help of commonly used so called "test cultures" *Sarcina lutea*, *Bacillus cereus*, and *E. coli*. Equal bricks of *Actinoplanes* sp. SE50/110 lawn, grown on three different media, were placed on freshly plated "test cultures" and incubated for 12 hours. It is clearly evident, that *Actinoplanes* sp. SE50/110, grown on SFM, can inhibit growth of *S. lutea* and *B. cereus*, but not *E. coli* (Figure 11A, B, C). To verify, whether observed effects are caused by acarbose, we made similar test with pure solution of acarbose in three different rising concentrations being soaked into filter paper, and placed onto freshly plated "test cultures". In this case, acarbose clearly inhibits growth of *E.coli*, but not other two strains (Figure 11D, E, F). These results are in good agreement with previous studies, showing inhibitory effect of acarbose on *E. coli* growth [16]. It also indicates,that while *Actinoplanes* is potentially synthesizing other metabolites, which posses inhibitory activities, it is probably not able to produce acarbose in any of the media tested.

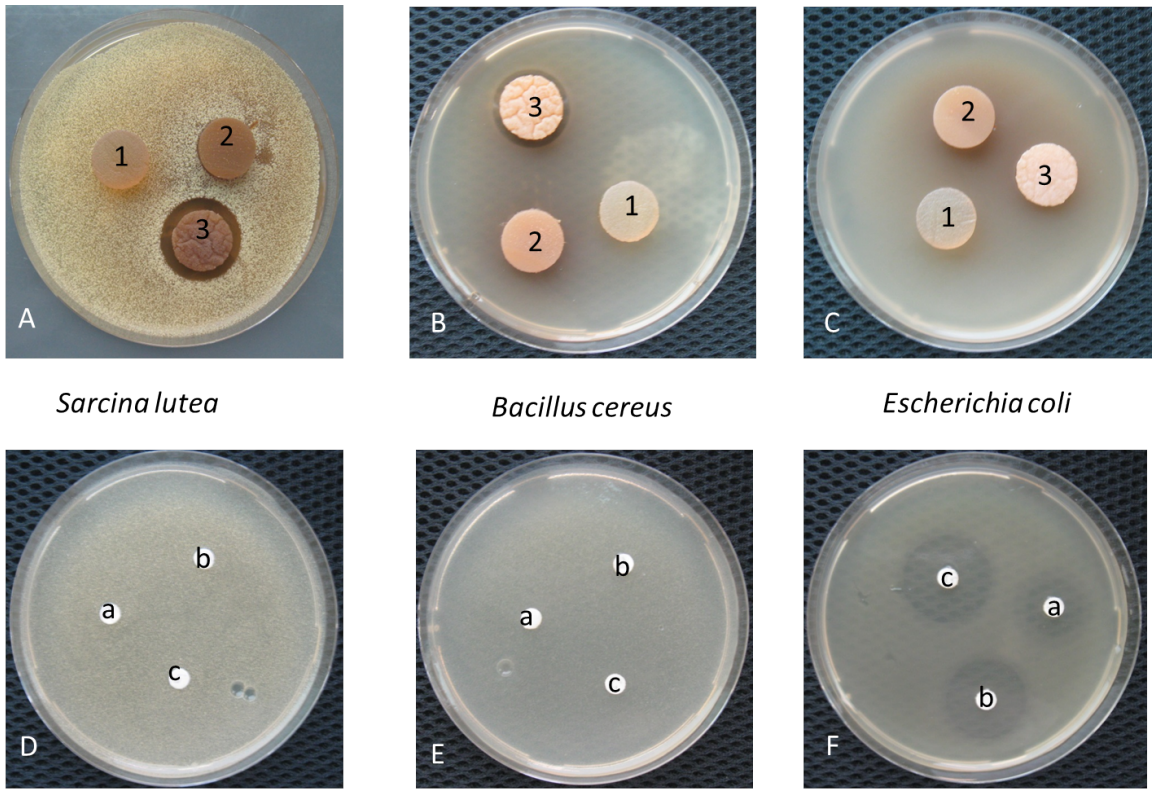

Fig. 11. Test of *Actinoplanes* sp. SE50/110 activity against "test cultures". *S. lutea*, *B. cereus* and *E. coli* are plated onto LB plates A and D; B and E; C and F respectively. Cubes of *Actinoplanes* sp. SE50/110 lawn are freshly cut out of 1 – Bennett agar media, 2 – OM media, 3 - SFM media; a, b and c – round pieces of filter paper, soaked in acarbose solution and put onto lawn of the "test cultures" in concentration 1 g/L, 10 g/L and 20 g/L respectively.

To define spore titers on various media, lawns of *Actinoplanes* sp. SE50/110 were harvested and stepwise tenfold dilutions were plated for counting of colony forming units. Estimated spore counts reached from $2 \pm 0.5 \times 10^7$ to $3 \pm 0.5 \times 10^6$ spores per mL for SFM and OM media, respectively (Table 2). Since SFM media was found to be favorable for sporulation, it was used in following experiments.

Colonies of *Actinoplanes* sp. SE50/110 on SFM media have a rough structure, well-defined contours and mostly form a "flower"-shaped structures (Figure 12A). Substrate mycelia have a dark yellow to orange, rarely brown, coloring, while the background of the plate exhibits dark brown pigmentation, probably due to the melanin formation. Members of genus *Actinoplanes* are frequently reported to gain a yellow or orange colouring due to the pigment of unknown origin, which is being accumulated by the substrate mycelia. This feature can be regarded as a typical and distinctive for the members of genus *Actinoplanes* and other related genera [102].

After 3-5 days of incubation, lawns of *Actinoplanes* sp. SE50/110, grown on SFM or OM media, typically showed distinctive signs of sporulation, such as increased hydrophobicity (water drops are stable, while placed on the top of mycelia) and change of colony colour from orange to white. To test, whether these observations are correct, scanning electron microscopic studies of *Actinoplanes* sp. SE50/110 single colonies were performed. Pictures were made from freshly grown SFM plates with different magnification. It is clearly seen (Figure 12B, C and D), that *Actinoplanes* sp. SE50/110 is forming globular and round shaped sporangia-like structures, which are placed onto thin sporangiophores. No developed aerial mycelia is present in these pictures. Interestingly, there is a thin layer of non-defined material, which covers many of the colonies pictured.

To test, whether *Actinoplanes* is able to sporulate in liquid cultures, multiple similar pictures were made using samples from liquid minimal media. In liquid media, *Actinoplanes* typically forms aggregation of clustered cells, which accumulate with during incubation, similarly to other Actinobacteria. In this case, all microscopic pictures did not show any sign of sporangia or sporangiophores (Figure 12E and F).

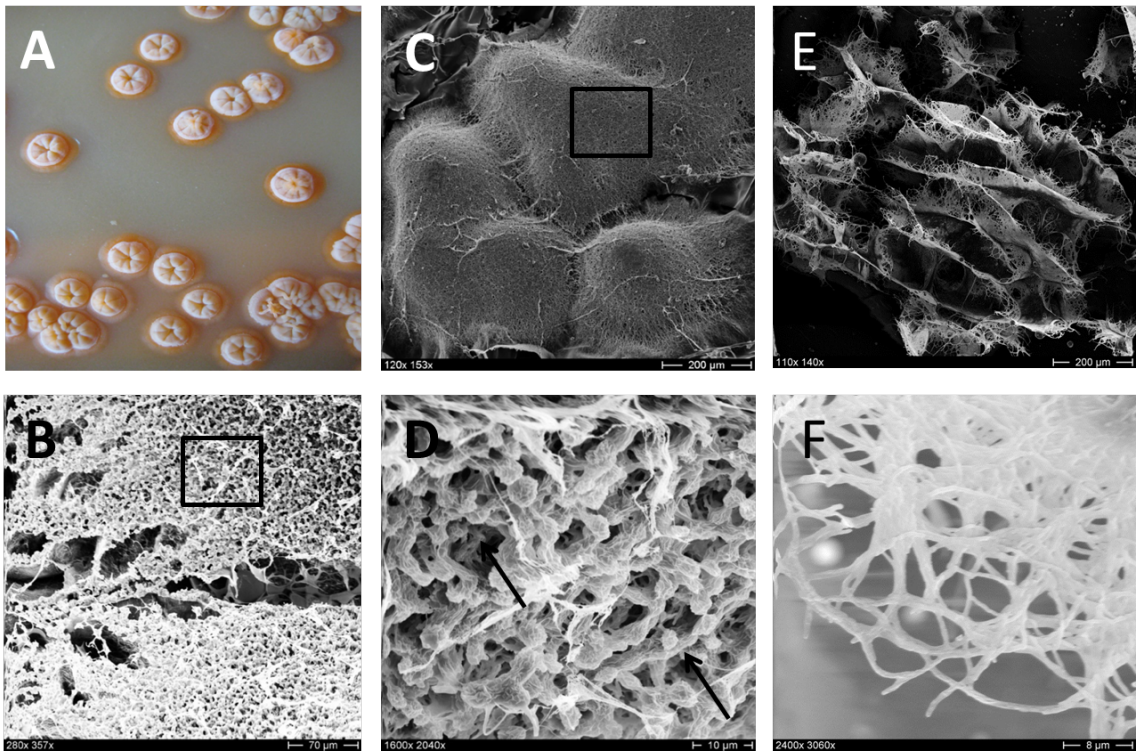

Fig. 12. Morphological features of *Actinoplanes* sp. SE50/110 colonies grown on SFM medium (A, B, C and D) and maltose minimal liquid medium (E and F). B, C, D, E and F – scanning electron microphotographs. A – "flower" shaped, round single and joined colonies; B – several tandemly joined colonies. Region, identified by square is magnified and shown in C; C – colony surface, abundantly covered with sporangia-like structures. Region, identified by square is magnified and shown in D; D – colony surface with typical structures, presumably round sporangia, substrate mycelia and sporangiophores, which are indicated with arrows in a left center part and right corner of the picture respectively. E - microphotographs of a single cell cluster; F - magnification of the edges of a single cell cluster. Modified from figure in [43]. Copyright 2016 by the Elsevier B.V. Adapted and reprinted with permission.

These observations are in a good agreement with other studies, which have shown similar structures for other members of Actinoplanaceae [102, 126, 124]. As for well studied genus *Streptomyces*, only minor amount of its members are able to sporulate

during liquid cultivations [68]. Typically, members of the genus *Actinoplanes* form motile spores, which are called zoospores [126]. Once released from the submerged sporangia, zoospores are able to swim in water in order to colonize distantly located environments [126, 124].

3.2.2 Antibiotic resistance spectra of *Actinoplanes* sp. SE50/110

To select appropriate antibiotic resistance marker genes, it is important to know, whether they would be suitable for a particular strain. Therefore, as a second step in development of a gene cloning system for *Actinoplanes* sp. SE50/110, natural antibiotic resistance/sensitivity pattern of this strain was studied. To gain a general overview on this question, fast and relatively uncomplicated disc diffusion method [99] with a standard set of 18 antibiotics was used. It includes antibiotics with various modes of action, such as β-lactams, aminoglycosides, macrolides, tetracyclines, polymyxins. To see all results of this test, please refer to Table 3.

Actinobacteria are typically resistant to β-lactams (ampicillin) and sensitive to aminoglycosides (streptomycin, kanamycin, gentamycin) and macrolides (erythromycin). The same is true for *Actinoplanes* sp. SE50/110. Some distinctive characteristics can be also highlighted in this case: resistance to low concentrations of spectinomycin, cefpirome, colistin and trimethoprim. Resistance to thiostrepton and spectinomycin was earlier shown for another member of genus *Actinoplanes*, *A. friuliensis* [50]. Typically for actinomycetes, the most frequently used antibiotics are: ampicilin, apramycin, kanamycin, hygromycin, viomycin and thiostrepton. There is a set of antibiotic resistance cassettes, which were developed specially for streptomycetes, are readily available and cover most of above mentioned antibiotics [68]. These cassettes are used not only in routine cloning experiments as part of the vectors, but also in systems, developed for gene knockout procedures, e.g. ReDirect system [46, 47, 45]. Therefore, these antibiotics were tested additionally, in order to determine a specific sensitivity of *Actinoplanes* sp. SE50/110 to them, in concentrations routinely used for

Number	Name	Abbreviation	Concentration, μg/disc	Growth inhibition zone, mm
1.	Erythromycin	Ery	10	43
2.	Ampicillin	Amp	25	0
3.	Rifampicin	Rif	30	30
4.	Bacitracin	Bac	10 units	26
5.	Chloramphenicol	Chl	10	15
6.	Streptomycin	Str	10	30
7.	Norfloxacin	Nor	5	21
8.	Polymyxin B	Pol	300 units	10
9.	Tetracycline	Tet	30	36
10.	Spectinomycin	Sp	10	0
11.	Gentamicin	Gen	10	15
12.	Linkomycin	Lin	15	14
13.	Cefpirome	Cef	30	0
14.	Colistin	Col	10	0
15.	Kanamycin	Kan	30	33
16.	Nalidixic acid	Nal	30	0
17.	Vancomycin	Van	5	12
18.	Trimethoprim	Trm	5	0

Table 3
Antibiotic resistance spectra of *Actinoplanes* sp. SE50/110 strain, studied with a disc diffusion method on SFM media. Modified from table in [43].

selection. Also, fosfomycin and nalidixic acid were included in this set, as a potential selection agents in *E.coli* - *Streptomyces* conjugation procedures, that we planed to implement later. In this test, tenfold dilutions of *Actinoplanes* sp.SE50/110 spore

suspension were grown on different antibiotic concentrations and compared to the media without antibiotic supplement (Table 4).

Name	Concentration, μg/μl	Survival, %	Possibility of application as a marker
Ampicillin	100	0	+
Apramycin	50	0	+
Hygromycin	50	0	+
Kanamycin	50	0	+
Nalidixic acid	100	0	Not considered
Fosfomycin	50 100	98 100	Not considered
Thiostrepton	50 100	82 69	- -
Viomycin	30	0	+

Table 4
Antibiotic resistance spectra of *Actinoplanes* sp. SE50/110 strain after six days of growth on SFM plates. Modified from table in [43]. Copyright 2016 by the Elsevier B.V. Adapted and reprinted with permission.

As it is evident from Table 4, all antibiotics, except for thiostrepton can be used for selection in case of *Actinoplanes* sp. SE50/110, since the strain shows a high level of resistance towards it. Interestingly, *Actinoplanes* sp. SE50/110 is sensitive to nalidixic acid,therefore this antibiotic cannot be used for selection against *E. coli* in conjugation experiments. However, as an alternative in this case, fosfomycin can be utilized, since *Actinoplanes* sp.SE50/110 is resistant to high concentrations of this antibiotic.

3.2.3 Choice of DNA uptake system for *Actinoplanes* sp. SE50/110

After selection of solid growth media, which supports sporulation of *Actinoplanes* sp. SE50/110, and also necessary tests on it antibiotic resistance/sensitivity profile, we move on to the choice of the proper DNA uptake system. There are many ways in which DNA can be introduced into cells of *Actinoplanes* sp. SE50/110, judging from accumulated practices for actinomycetes and related genera. The most common and widely used are: 1) electroporation of mycelia and protoplasts; 2) transformation of protoplasts; 3) conjugation with other actinomycetes or intergeneric conjugation between *E. coli* and *Actinoplanes* sp. SE50/110 [68, 9]. These three techniques were implemented for *Actinoplanes* sp. SE50/110 and outcome of these experiments is discussed in following chapters. As other possibility of DNA transfer, transduction with phages can be noted. However, due to the limited amount of wide host-range generalized phages, this technique is rarely implemented and was not applied for *Actinoplanes* sp. SE50/110 [68].

All tests regarding DNA uptake techniques were performed with the one copy integrative vector pSET152, as the most frequently used streptomycetes vector, based on a broad host range integration system [12]. It is also important to note, that Actinobacteria are known to posses methylation-specific restriction systems [68, 9]. In order to circumvent it, all DNA used in these experiments, was passed through methylation deficient host *E. coli* ET12567 (pUZ8002) strain [80].

3.2.3.1 Protoplast transformation and electroporation of *Actinoplanes* sp. SE50/110

Electroporation as a method of DNA transfer is relatively the simplest and the fastest among other available options. It was successfully established for streptomycetes and also for other rare Actinobacteria [81, 79]. Electroporation involves application of short, high voltage pulse, to cells mixed with DNA in salt depleted buffer. This pulse results in creation of pores in cell wall and subsequent DNA up-

take [68]. Field strength and pulse duration are important variants in this procedure. Importantly also, two kinds of electroporation are possible: of whole mycelial cells or protoplasts. At first, we attempted to electroporate mycelia cells, freshly grown in resource-rich NBS liquid medium to logarithmic growth phase (14-18 hours, since the beginning of cultivation). Then, the protocol, established for *S. coelicolor* and *S. lividans* was followed [68]. In this protocol, 2 kV electric pulse (10 kV/cm $^{-1}$) is applied. After 5 days of incubation at 28 °C, which is a typical time, sufficient for *Actinoplanes* sp. SE50/110 growth, there were no viable transformants in all replicates tested, even when post-electroporation incubation time was extended from 3 to 5 hours. It was shown earlier, that conditions for electroporation can vary and are highly strain-specific. Therefore, we also tested different field strengths ranging from 5 to 12.5 kV/cm $^{-1}$, as described in [81] for *Nocardia* sp. CS682. Unfortunately, change of the field strength in this case did not lead to the improvement in electroporation efficiency. Therefore, the focus of the study was shifted to the development of protoplast transformation and conjugation procedures.

Protocols for protoplasts preparation, fusion and regeneration, similarly to electroporation, are very frequently sophisticated and genus- or strain-specific. It can probably be explained by interspecies variation of the cell wall composition [85]. Thats why, we decided to apply protocols, which were developed specifically for the members of genus *Actinoplanes* [85] and other related rare Actinobacteria [86]. In order to achieve high transformation rate, at first strain has to show abundant growth in protoplast production medium. For this purpose, we tested V and VM media, rich in starch and sucrose, as suggested in [85] and YEME media, which is traditionally used for streptomycetes [68]. Notably, *Actinoplanes* was not able to grow in YEME media, however showed dispersed growth in V and VM media, which were used further. After harvesting mycelia from cultivation solution, procedure, developed in [85] for protoplast generation was followed. It involves treatment of mycelia with a combination of enzymes, hen egg-white lysozyme (HEWL) and *S. globisporus* mutanolysin in final concentrations 5 and 0.018 mg/mL respectively. For similar pro-

cedure in streptomycetes, typically only HEWL, which is N-acetyl-β-D-muramidase is used [68]. However, mutanolysin, another muramidase, was shown to be crucial to receive protoplasts in "rare" actinomycetes. To avoid formation of foam, non-ionic detergent Pluronic was added to a growth media in the final concentration of 100 mg/mL. Protocol, described in [86] suggests 1 to 2 days of enzyme treatment in order to generate an appropriate amount of protoplasts, however in case of *Actinoplanes*, we observed formation of protoplasts and spheroplasts already after 3 hours of digestion (Figure 13). At this time of lysis still huge aggregations of mycelia are visible, therefore we also tested prolonged incubation up to 6 and 24 hours. Under phase contrast microscope, it is clearly visible, that these aggregations of mycelia disappear after longer exposure to enzyme mixture (Figure 13).

Amount of protoplasts at all time points, which were studied, can be regarded as high in comparison to other *Actinoplanes* strains: 8.4 x 10^7, 6.4 x 10^7 and 4.55 x 10^7 at 3, 6 and 24 hours respectively.

To test protoplast transformation, the procedure described in [68] for streptomycetes, was followed. It involves utilization of PEG 1000 in combination with electroporation technique. Unfortunately, after multiple attempts, no viable transformants were observed on regeneration media. In this case, MP3 and M103 from [86] and R2YE [68] were tested as possible regeneration media. R2YE is a rich media, specifically designed for regeneration of streptomycetes protoplasts, but interestingly it did not support growth of *Actinoplanes* colonies. On MP3 and M103 two to three colonies per plate were observed after transformation procedure. However, these colonies did not grow after transfer to a new media with selective antibiotic, which possibly hints, that they were rather mutants, than transformants.

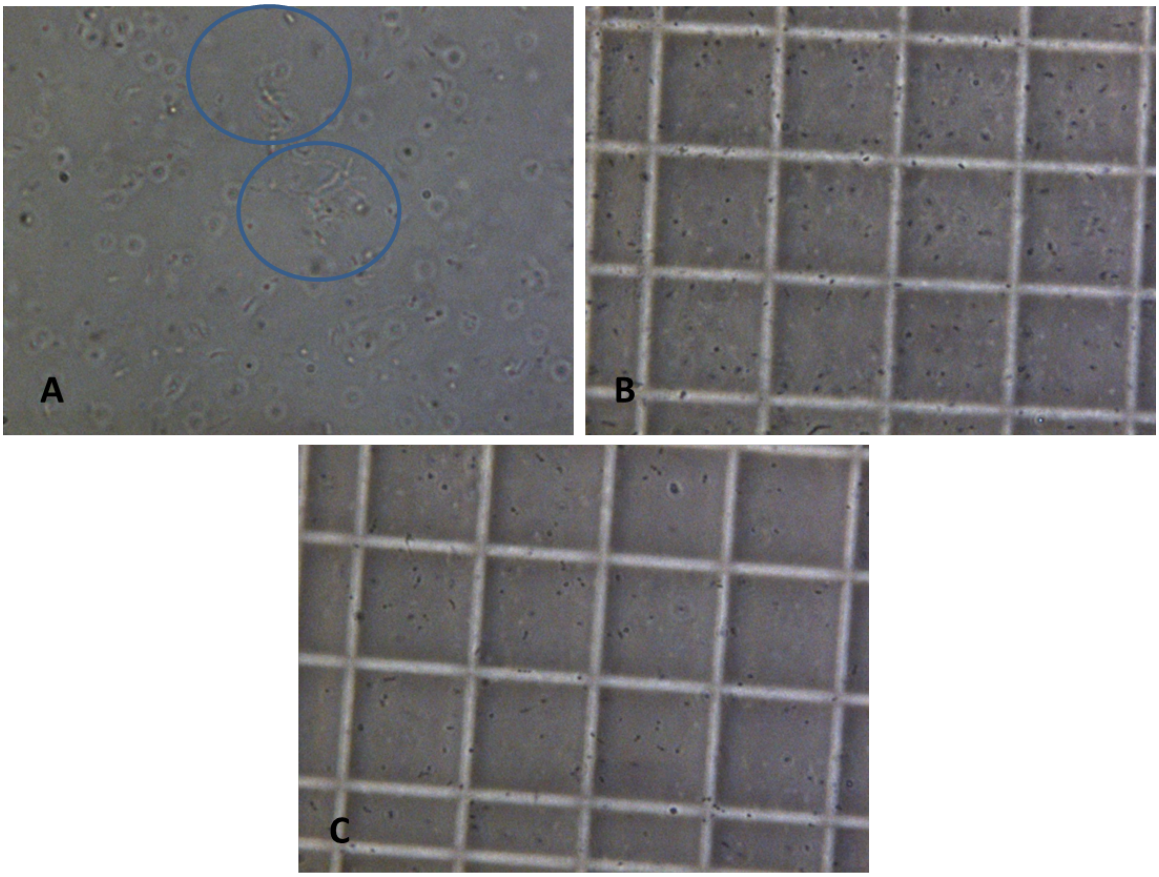

Fig. 13. Microscopic picture of *Actinoplanes* sp. SE50/110 protoplasts after A - 3 hours, B - 6 hours and C - 24 hours of lysis. Mycelia aggregations are being highlighted with blue circles.

3.2.3.2 Adaptation and optimization of the conjugation procedure between *E. coli* and *Actinoplanes* sp. SE50/110

As protoplast transformation and electroporation, following traditional, established protocols, were shown to be non effective for *Actinoplanes* sp.SE50/110 and the conjugation procedure was already successfully applied to some strains of the genus *Actinoplanes* (e.g. *A. teichomyceticus, A. friuliensis*; [48, 59, 96, 50] we decided to adapt reported protocols for *Actinoplanes* sp. SE50/110.

At the beginning, we followed procedures, described in [68] and [59]. In further experiments, several protocol steps were modified in connection to specific traits of *Actinoplanes* sp. SE50/110.

It is frequently reported, that selection of an appropriate conjugation media, is able to greatly influence the successful conjugal transfer in Actinobacteria. Generally, higher sporulation titers, result in higher transformation frequencies. That is why, media, which support abundant sporulation are directly influencing the outcome of the mating experiment. Previously in this work, two solid media supporting growth and sporulation were identified, SFM and OM. Both of them were tested in conjugation experiments and showed conjugation frequencies of $3.0 \pm 0.5 \times 10^{-3}$ and $2.5 \pm 0.5 \times 10^{-6}$ for SFM and OM respectively, thus confirming our assumptions. Therefore, we continue to use SFM media for all conjugation matings.

Media, used for conjugation, are frequently supplemented with 10 mM $MgCl_2$, in order to increase conjugation frequency [68]. Functions and role of $MgCl_2$ in this process are not studied yet. For *A. teichomyceticus*, an optimal concentration of $MgCl_2$ in media was found to be 40 mM [48]. We have tested addition of 10 mM $MgCl_2$ to SFM and OM media, but found no influence on conjugation frequencies.

Spores, after harvesting from plates, are typically being subjected to "heat shock", a short incubation at 50 °C - 55 °C [68]. Spore germination, which can be induced by heating, was reported to increase conjugation rates [87]. As it was discussed previously, spores of almost all representatives of genus *Actinoplanes* are motile [126, 124] and are expected to be sensitive to a "heat shock" procedure [48, 59]. For *Actinoplanes* sp. SE50/110, we observed that incubation of freshly harvested spore suspension for 15 min at 50°C led to significant decrease in the amount of vital spores (from $3 \pm 0.5 \times 10^7$ to $2 \pm 0.5 \times 10^3$). Therefore, this step was not included into finalized conjugation protocol.

Incubation time for *Actinoplanes-E. coli* conjugation plates has to be extended up to 20-22 hours to receive exconjugant colonies, since no colonies were observed after 12-18 hours of incubation, in comparison to *A. teichomyceticus* as described in

[48]. As long as *Actinoplanes* sp. SE50/110 was found to be sensitive to nalidixic acid, only fosfomycin and apramycin were used in selection against *E. coli* cells and added to an overlay solution. After 4-6 days of incubation, following an overlay of the agar plates with the respective antibiotic mixture, single exconjugant colonies were transferred to fresh apramycin containing SFM plates. However, resulting *Actinoplanes* exconjugants frequently contained *E. coli* cells. In order to "clean" cultures of *Actinoplanes* exconjugants, we took advantage of acarbose effect on *E. coli* growth. As mentioned before and as described in [16], acarbose can affect growth of *E. coli*, while grown on media with maltose as the sole source of carbon and energy. Probably, due to its structural similarity to maltotetraose, acarbose is being recognized as a substrate by the maltose/maltodextrin system of *E. coli*. Single colonies of *Actinoplanes* were transferred to liquid maltose minimal media, supplemented with 50 mg/L of apramycin, fosfomycin and acarbose and cultivated for 48 hours at 28 °C. Cultures were stepwise diluted with sterile water and plated on apramycin containing SFM plates. Single colonies of "*E. coli* – free" *Actinoplanes* exconjugants were transferred to another media and used for further tests and for making glycerol long storage cultures. A general overview on the developed *Actinoplanes* - *E. coli* intergeneric conjugation procedure is depicted in Figure 14. The final protocol can be found in the Methods chapter.

3.3 Application of the developed gene cloning system for the genetic engineering of *Actinoplanes* sp. SE50/110

As a next step in establishment of an effective system for genetic engineering in *Actinoplanes* sp. SE50/110, we explored a possibility to transfer set of various integrative and replicative vectors, commonly used for Actinobacteria. Successful application of these vectors, opens a door to a wide range of genetic experiments, namely gene knockouts, overexpressions and complementations of mutant strains. This chapter also includes an exploration of the effects of studied vectors on *Actinoplanes* physiol-

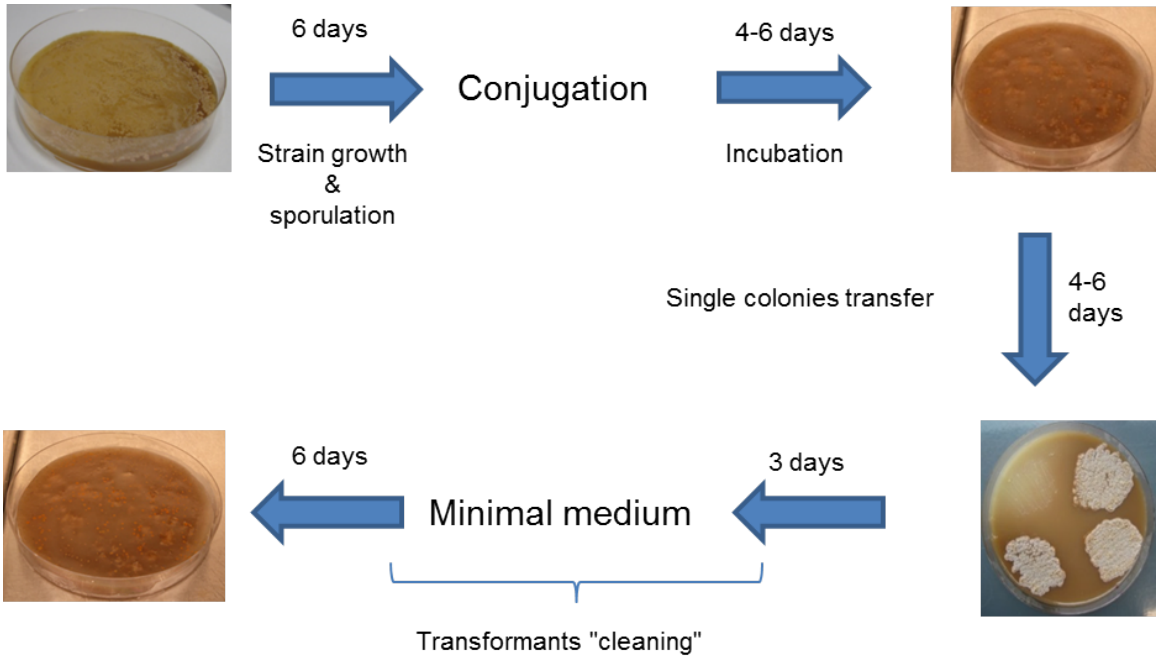

Fig. 14. General overview of *E.coli-Actinoplanes* conjugation procedure. Modified from figure in [43]. Copyright 2016 by the Elsevier B.V. Adapted and reprinted with permission.

ogy, morphology and biosynthetic activity. Further on, an application of integrative vectors to study activity of heterologous promoters in *Actinoplanes* is described.

3.3.1 Transfer of a set of actinophage-based integrative vectors to *Actinoplanes* sp. SE50/110 and characterization of exconjugant strains

Previously, integrative vectors were shown to be a powerful tool for genetic engineering in at least two strains of genus *Actinoplanes* [50, 59]. Also, an attempt on genetic modification of *Actinoplanes* sp. SE50/110 with the help of integrative vector is noted in [50] as part of unpublished results. Since there was no clear data available on which vectors are applicable for *Actinoplanes* sp. SE50/110, we decided to include different integration systems into this experiment. As the first and the most

obvious candidate, pSET152 vector was chosen, as in our previous experiments it was shown to be applicable for *Actinoplanes* sp. SE50/110. In addition, three integrative vectors: pIJ6902, pSOK804, pRT801 based on three different phage integration systems: φC31, VWB and φBT1 were chosen. All of them are widely used for the representatives of *Streptomyces* genus and were also shown to be applicable for *A. teichomyceticus* [59]. All of these plasmids contain the apramycin resistance marker gene *aac*(3)IV, which was selected because of the high sensitivity of *Actinoplanes* sp. SE50/110 to apramycin.

As mentioned before, methylation deficient *E. coli* ET12567 (pUZ8002), carrying integrative vectors was used as donor strain. In all sets of matings, exconjugant colonies with high conjugation frequencies were received for all vectors tested (Table 5). Authenticity of exconjugant colonies was checked with the help of PCR, utilizing primers to internal parts of the vector (*oriT* regions and *aac*(3)IV).

Number	Vector name	Integrative system	Transformation frequency	Reference
1.	pSET152	φC31	$3.3 \pm 0.5 \times 10^{-3}$	[12]
2.	pIJ6902	φC31; contains promoter *tipAp*	$3.8 \pm 0.5 \times 10^{-3}$	[61]
3.	pRT801	φBT1	$4.5 \pm 0.2 \times 10^{-3}$	[42]
4.	pSOK804	VWB	$2.4 \pm 0.4 \times 10^{-3}$	[112]

Table 5
Overview of integrative vectors used in conjugation experiments. Modified from table in [43]. Copyright 2016 by the Elsevier B.V. Reprinted with permission.

It is known, that integrative vectors are able to exist as extrachromosomal non-replicating forms in mycelia. Recently, pSET152 vector was detected extrachromosomally in cells of *S. ghanaensis* and other Actinobacteria [100, 68]. It was suggested, that some of its copies, which are tandemly integrated in the chromosome, are able

to excise at a very low frequency [78, 94]. To verify such possibility, total DNA of *Actinoplanes* colonies bearing pSET152, pIJ6902, pSOK804 or pRT801 was used to transform *E. coli* in order to check for the appearance of apramycin-resistant colonies. Since no such colonies were received in all cases, we can conclude that these vectors can exist only integrated in the *Actinoplanes* genome.

For all further implementations of a gene-cloning system such as gene overexpression experiments, complementation of mutants, it is important to verify the stability of integrative vectors in cells under non-selective conditions. This test was carried out in accordance to [59]. To study the stability of plasmid inheritance, exconjugant colonies were grown for five passages at 28 °C without application of an antibiotic. For each type of vector, 300 colonies were checked for apramycin resistance at each round and this trait was found to be 100% stable for all studied vectors.

It is known, that some vectors can affect the morphology and/or the secondary metabolism of certain actinomycetes. For example, the presence of pSET152 in the cells of *A. teichomyceticus* was shown to significantly increase teicoplanin production [59]. Therefore, colony morphology, spore titers and acarbose production levels of all exconjugants were assessed. Regarding morphology, no changes were observed by visual inspection in case of all exconjugant strains (Figure 15). Also, sporulation of exconjugant strains is not changed considerably. For example, to compare pSET152 carrying strain with the wild type strain, titers changed from $4 \pm 0.5 \times 10^7$ to $3 \pm 0.2 \times 10^7$ respectively.

It is of great importance to know, whether particular vector is able to influence metabolism of the host, especially secondary metabolites production. Therefore, we cultivated strains, carrying vectors in parallel with wild type in maltose minimal media and measured their respective acarbose production. As it is evident from Figure 16, there was only little influence of all studied vectors on strain growth and on acarbose production.

To finish the characterization of integrative vectors, their copy number and more importantly integration site in the genome of *Actinoplanes* have to be determined.

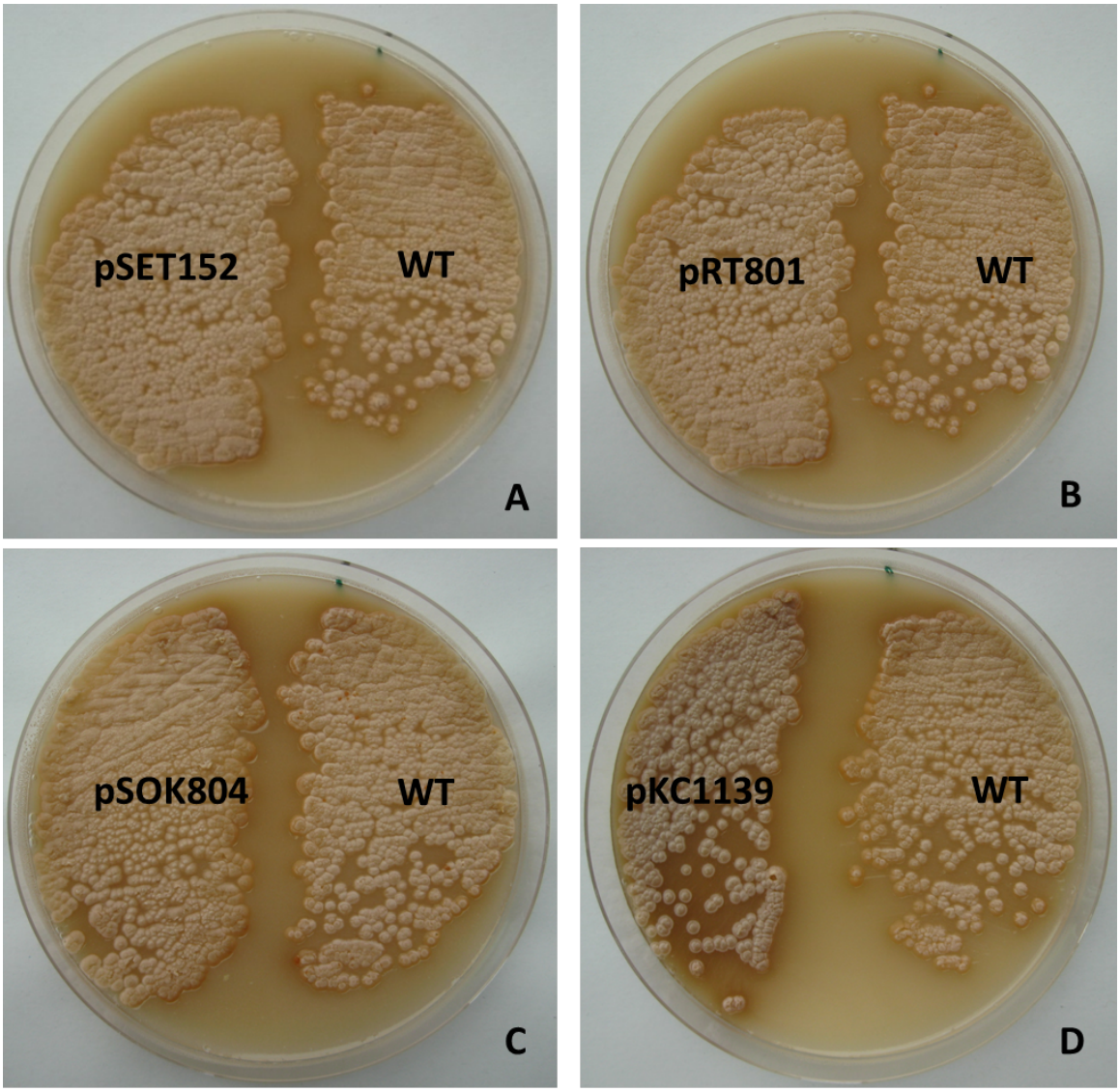

Fig. 15. Morphology comparison of *Actinoplanes* sp. SE50/110 and *Actinoplanes* pSET152 (A), pRT801 (B), pSOK804 (C), pKC1139 (D). All plates are grown for six days on SFM media. Modified from figure in [43]. Copyright 2016 by the Elsevier B.V. Reprinted with permission.

Integration of a whole vector sequence into a gene coding sequence, or a region potentially involved in the regulation of cellular functions, can be detrimental to cell metabolism. Specific sequences of *attB* sites for three actinophage systems, which

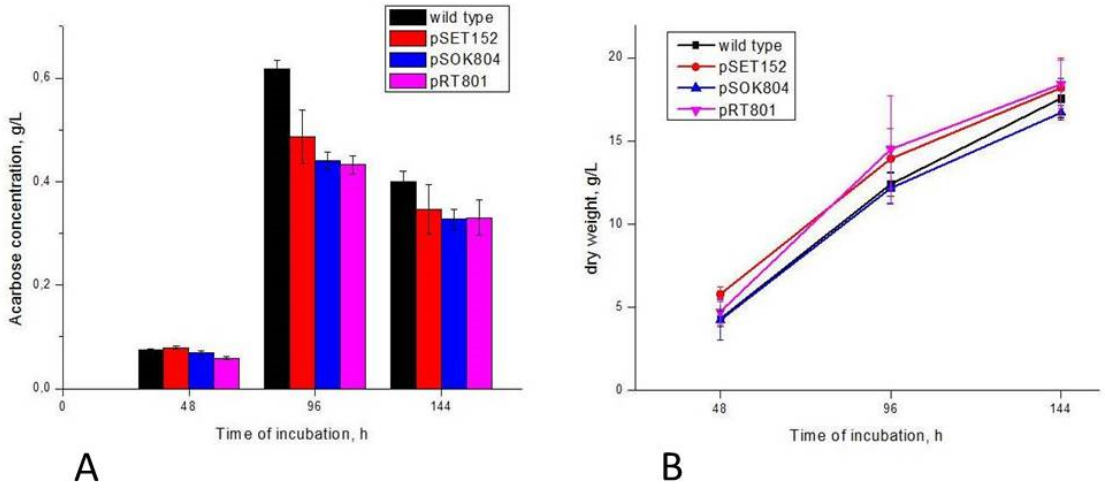

Fig. 16. Acarbose concentration for cultivations of *Actinoplanes* sp. SE50/110 wild type without a plasmid and the pSET152, pSOK804, pRT801 exconjugant strains, grown in Mal-MM. A – recorded values for acarbose concentration, g/L; B – recorded values for cell dry weight, g/L. Mean values and standard deviations are shown for three biological replicates. Modified from figure in [43]. Copyright 2016 by the Elsevier B.V. Adapted and reprinted with permission.

were used in this study, were not previously characterized for members of genus *Actinoplanes* and are generally poorly studied. Therefore, in order to determine them, genomes of pSET152, pSOK804 and pRT801 *Actinoplanes* exconjugants were completely sequenced and integration sites were studied by *in silico* finishing [49, 135, 136]. In order to ensure, that all the possible variations of insertion sites are included, DNA of three independently received exconjugants, for each vector respectively, was mixed before sequencing. Colonies, containing pIJ6902, were excluded of these experiment, as long as pIJ6902 and pSET152 are both based on φC31 integrative system.

3.3.2 Characterization of pSET152 (φC31-based), pSOK804 (VWB-based) and pRT801 (φBT1-based) integration sites in *Actinoplanes* sp. SE50/110 genome

An interesting feature of the *Streptomyces* phage integration systems is that each system was found to posses a unique *attB* site, and the individual *attB* sites are located in the unrelated genes [10]. It gives a possibility to construct vectors, which can be used simultaneously for one particular strain, in order to overexpress genes of interest. However, this question is poorly studied in other genera of Actinobacteria, especially so called "rare" actinomycetes. In *Actinoplanes* sp. SE50/110, the $attB^{\varphi C31}$ site is located within gene ACPL_6602. This gene is annotated as coding for a putative pirin homolog, likewise in other Actinobacteria, such as *S. coelicolor*, *A. teichomyceticus*. The role of pirin in bacteria is poorly understood, however it is clearly influencing diverse biological processes. In eukaryotes, pirin was reported to act mainly as a transcriptional cofactor [114].

Integration of vectors which have an *attP* site into the *attB* locus of recipient bacterial genome is happening via the integrase (*int*) function. Importantly, φC31 bacteriophage possesses a large serine recombinase, which require no additional phage or host functions for site-specific integration. Integration is unidirectional and happens in the absence of any additional factors [10]. The efficiency of integration process, using the bacteriophage φC31 *att/int* system, was reported to be highly dependent on the homology of *attB* sites. Sequences of *attB* sites which show higher homology to those of *S. lividans*, result in higher conjugation efficiencies [48]. The nucleotide sequence of *Actinoplanes* sp. SE50/110 $attB^{\varphi C31}$ site was found to be most closely related to that of *A. teichomyceticus* (88.2% identity; 45 out of 51 identical nucleotides) and *S. scabies* (88.2% identity; 43 out of 51 identical nucleotides) (Figure 17). The *Actinoplanes* sp. SE50/110 $attB^{\varphi C31}$ site is also highly related to those from *S. coelicolor*, *S. lividans, S. ambofaciens, S. aureofaciens* (80.4% identity; 41 out of 51 matches) and *S. ghanaensis* (82.4% identity; 42 out of 51 matches). Thus, notable

accordance between *attB* sites across genera could be observed. Specifically important is the conservation of an *attB* "core sequence", where the crossover between *attB* and *attP* takes place. This site consists of 5'-TT-3' sequence [20].

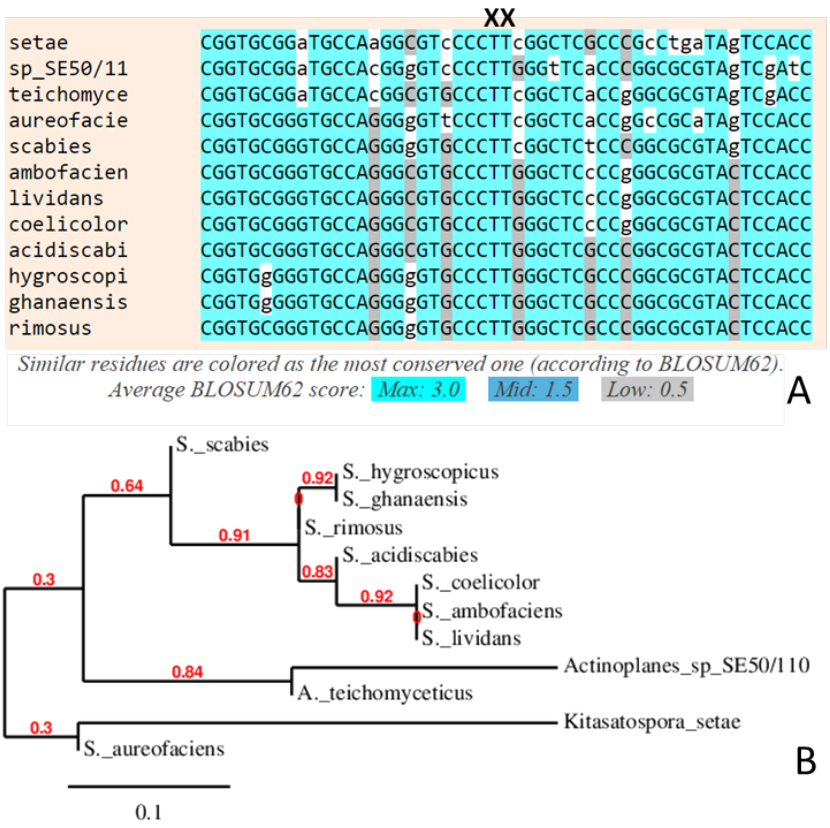

Fig. 17. Alignment of the *attB* site sequences (A) and phylogenetic tree among *Actinoplanes* sp. SE50/110 and other actinomycetes: setae - *Kitasatospora setae* KM-6054; sp. SE50/110 - *Actinoplanes* sp. SE50/110; teichomyceticus - *A. teichomyceticus* NBRC13999; aureofaciens – *S. aureofaciens* ATCC10762; scabies – *S. scabies* NRRL B-16523 B16523; ambofaciens – *S. ambofaciens* ATCC23877; lividans – *S. lividans* 66 TK64; coelicolor – *S. coelicolor* A3(2); acidiscabies – *S. acidiscabies*; hygroscopicus – *S. hygroscopicus* NRRL5491; ghanaensis – *S. ghanaensis* ATCC14672; rimosus – *S. rimosus* R7. The alignment was created with MUSCLE 3.8.31 program and by software described in [24]; core sequence is marked with "XX" sign. Modified from figure in [43]. Copyright 2016 by the Elsevier B.V. Adapted and reprinted with permission.

It was reported, that *S. lividans* and *S. coelicolor* both possess three highly active pseudo*attB* sites in their genome, which can serve as additional integration sites, apart from original [20]. The fourth pseudo*attB* site for the φC31 integration system was recently identified in the genome of *S. albus* [13]. In case of *Actinoplanes* sp. SE50/110, no additional sites were detected.

The vector pRT801 is based on the *attP/attB* system of another temperate bacteriophage related to φC31, φBT1 [42]. In *S. coelicolor* φBT1 integrates into a gene annotated to code an integral membrane protein that is unrelated to the pirin-like gene used for φC31 integration [10]. However, *attB* sites for φBT1 are comparatively poorly studied; there are no known sequences in *Actinoplanaceae* or other rare actinomycetes [42, 10]. The best studied *attB* site of *S. coelicolor* lies within gene *SCO*4848, homologes of which exist in the *Actinoplanes* sp. SE50/110 genome. Recently, it was demonstrated, that disruption of *SCO*4848 generates a delay in spore germination of *S. coelicolor*. It is probably due to the fact, that *SCO*4848 is co-transcribed with *SCO*4849, as long as spore germination phenotype is complemented by *SCO*4849 [40]. However, in *Actinoplanes* sp. SE50/110 the $attB^{\varphi BT1}$ site is located in an intergenic space, between genes ACPL_7730 coding for a dihydrolipoamide dehydrogenase and ACPL_7731 coding for a Gamma-glutamylcyclotransferase. Due to the evidence, insertion does not cause any influence on sporulation of *Actinoplanes*. No other pseudo-*attB* sites for φBT1 were detected.

In order to broaden the spectrum of genetic tools applicable for *Actinoplanes* sp. SE50/110, the pSOK804 vector [125], that is based on the VWB bacteriophage integration system, was applied. Previously, pSOK804 was successfully applied for gene cloning systems in various representatives of the actinomycetes: *A. teichomyceticus*, *S. sioyaensis*, *S. venezuelae*, *S. ghanaensis* [125, 100, 59]. However, only two $attB^{\mathrm{VWB}}$ sites have been characterized so far in *S. venezuelae* ETH14603, which is a natural host for VWB phage, and in *S. ghanaensis* [125, 100]. Nothing is known about the genomic context and nature of $attB^{\mathrm{VWB}}$ loci in other strains [100].

Both *S. venezuelae* and *S. ghanaensis* $attB^{VWB}$ sites are embedded into a putative $tRNA^{Arg}$ (AGG) gene [100]. Although such gene exists in the genome of *Actinoplanes* sp. SE50/110, the $attB^{VWB}$ site is not located in the $tRNA^{Arg}$ gene, but in the coding region of the hypothetical protein ACPL_7821. These results suggest a variability among $attB^{VWB}$ across species, however to make any conclusions, more information regarding $attB^{VWB}$ in other representatives of Actinoplanaceae is needed. In the genome of *Actinoplanes* no other pseudo-*attB* sites for VWB were detected.

3.3.3 Application of integrative vectors for genetic engineering in *Actinoplanes* sp. SE50/110

All integrative vectors studied, pSET152, pSOK804 and pRT801 have different, single and unique integration sites in the genome of *Actinoplanes*. More importantly, all vectors do not significantly influence metabolism or morphology of the strain. Therefore, all of them can be used for genetic engineering of *Actinoplanes*, separately or in a combined manner. Therefore, in this chapter we describe our first attempts to apply integrative vectors for overexpression experiments and for the development of suitable reporter system. Application of integrative vectors as a part of gene knockout experiments will be discussed in a next chapter.

3.3.3.1 Adaptation of the GUS reporter system to determine suitable heterologous promoters for use in *Actinoplanes* sp. SE50/110

After the successful development of a genetic engineering toolkit and the application and characterization of integration sites of different integrative vectors, we would like to investigate the applicability of the reporter system for genetic studies in *Actinoplanes* sp. SE50/110. Reporter genes enable the visualization of multiple biological processes in living cells. These genes, fused to regulatory sequences, and being introduced into a biological system, are able to provide a signal, which can be quantified [93].

At the moment the broad choice of reporter genes, applicable for actinomycetes species exists, including *xylE* (encoding catechol 2,3-dioxygenase), *cat* (encoding chloramphenicol acetyltransferase), *neo* (encoding neomycin phosphotransferase), *luxAB* and *luc* (encoding luciferases) and *gfp* (encoding green fluorescent protein - GFP) [68]. At the moment, only three reporter systems: GFP, XylE and newly described GUS system, were applied for studies in *Actinoplanes* [127, 58]. XylE system was previously proven to be not suitable for studies in *A. teichomyceticus* [58]. GFP system is often characterized as less sensitive than other systems, due to background fluorescence of all materials used in assays and absence of enzymatic signal amplification. In addition, the assay may suffer from UV-induced toxicity and photobleaching; in several systems, the formation of GFP aggregates was shown to cause cytotoxicity [63]. GUS system, on the other hand, is claimed by its developers to be highly sensitive due to the stability and high specific activity of the GUS enzyme, which does not need any cofactors and is tolerant to the most commonly used chemicals and assay conditions. Also, it is expected, that most streptomycetes do not carry any endogenous GUS activity [93]. Therefore, we decided first to apply GUS reporter system and, in case if it is proven to be ineffective, test utilization of GFP system next.

To verify, whether this system is applicable for *Actinoplanes*, pSETGUS vector, a derivative of pSET152 [93], was transferred to *Actinoplanes* sp. SE50/110 through conjugation mating with *E. coli* as described before. It contains the *gusA* gene, coding for the β-glucuronidase enzyme, under the control of the "thiostrepton-inducible" *tipA* promoter. Promoter *tipA* has proved highly successful for the overexpression of streptomycete genes previously [68]. It requires the TipAL protein for induction [55], which is not present in all streptomycetes. Genome of *Actinoplanes* sp. SE50/110 contains a gene, ACPL 2423 which shows homology to *tipAL*, therefore *tipA* promoter is expected to be inducible in this strain as well.

Exconjugant strains of *Actinoplanes* were verified with the help of PCR as it is described previously for pSET152 vector. To prove the activity of the *gusA* gene,

exconjugant strains were grown on SFM media together with wild type strain and *Actinoplanes* pSET152. On the sixth day of growth, drops of 0.1 M XGluc solution were placed onto the surface of lawns (Figure 18). In all *Actinoplanes* pSETGUS exconjugants formation of a blue colour (5,5-dibromo-4,4-dichloroindigo) was observed under these conditions, while colours of the wild type strain and its pSET152-carrying strains were unchanged. This fact not only proves the presence of transferred vectors in *Actinoplanes* cells, but also shows an absence of any background β-glucuronidase enzymatic activity in the wild type. As it was expected from studies in *A. teichomyceticus* [58], the "thiiostrepton-inducible" *tipA* promoter is suitable for gene expression experiments in *Actinoplanes* sp. SE50/110. It is also clearly seen, that thiostrepton is not necessarily needed for induction of this promoter, as long as it was not added to media, prior to test. However, in order to identify a specific response of *tipA* promoter to induction by thiostrepton, studies have to be carried out in liquid culture and quantified.

In order to widen a spectrum of promoters, which can be used in *Actinoplanes* sp. SE50/110, we further studied a set of constructs, containing various promoter sequences, namely: 1) pGUSPErmE; 2) pSETPwblAgusA; 3) pSETPcdaRgusA; 4) pSETPmoeE5; 5) pSETPAmgusA. These constructs contain promoter regions of the following genes: the apramycin resistance gene *aac(3)IV* from *Klebsiella pneumoniae* [12], the mutated variant of the promoter of the erythromycin resistance gene from *Saccharopolyspora erythrea ermE** [127], a regulatory genes *cdaR* and *wblA* from *S. coelicolor* [107, 32], *moeE5* gene, NDP-hexose epimerase coding gene from *S. ghanaensis* [98]. All plasmids used here were kindly provided by A. Luzhetskyy and L. Horbal (Saarbrücken, Germany) and previously studied in *A. teichomyceticus* [58]. Before transfer with *E. coli* - *Actinoplanes* intergeneric conjugation, all constructs were verified, as advised by L. Horbal, with restriction endonucleases, [58].

Judging from the plate-based assessment of GUS activity of all exconjugant strains (Figure 18A), all promoters, which were studied, are active in cells of *Actinoplanes* and can be used for further studies.

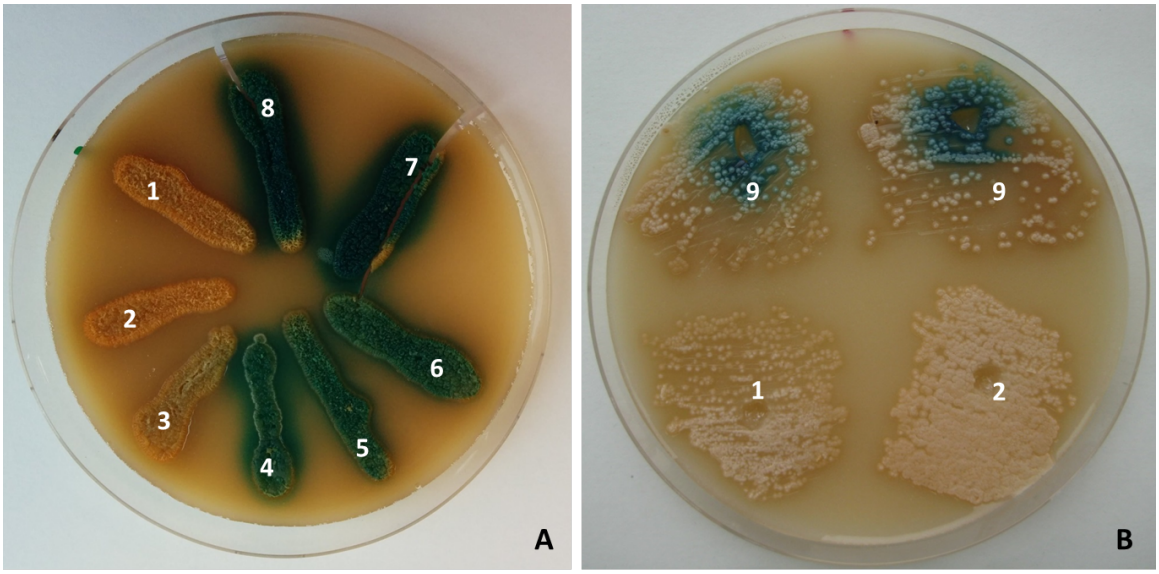

Fig. 18. Lawns of *Actinoplanes* sp. SE50/110 and its exconjugants with drops of 0.1 M XGluc solution, placed on the top center of the surface of mycelia. A - a test of a set of heterologous promoters in *Actinoplanes* sp. SE50/110; B - "proof-of-the-principle" test of GUS reporter system. 1 - *Actinoplanes* sp. SE50/110; 2 - *Actinoplanes* pSET152; 3 - *Actinoplanes* pGUS; 4 - *Actinoplanes* pGUSPErmE; 5 - *Actinoplanes* pSETPwblAgusA; 6 - *Actinoplanes* pSETPcdaRgusA; 7 - *Actinoplanes* pSETPmoeE5; 8 - *Actinoplanes* pSETPAmgusA; 9 - *Actinoplanes* pSETGUS. All constructs, used in this experiment, are kindly provided by A. Luzhetskyy and L. Horbal and described in [93] and [58]. Modified from figure in [43]. Copyright 2016 by the Elsevier B.V. Adapted and reprinted with permission.

Results of this experiment pave a way of further possible applications of GUS reporter system for studies in *Actinoplanes* sp. SE50/110, e.g. transcriptional or translational gene fusions, fast selection of knockout strains or further experiments.

3.3.3.2 Overexpression of *acb* gene cluster in *Actinoplanes* sp. SE50/110

One of the easiest and the most straightforward ways to increase production of a secondary metabolite is connected to the overexpression of genes, which constitute

the respective biosynthetic cluster [14]. Therefore, as one of the first applications of integrative vectors for genetic engineering in *Actinoplanes* sp. SE50/110, we decided to duplicate the entire *acb* gene cluster in its cells. For this purpose we used cosmid 6.2, earlier utilized in experiments for heterologous expression of *acb* cluster. As discussed earlier, this cosmid is based on the vector pOJ436 (11,1 kb), it contains gene of an apramycin resistance, lacZα with MCS for cloning DNA, φC31 phage *int* gene, *attP* sequence and is able to integrate site-specifically at the φC31 attachment sites in the genome. We expect it to integrate specifically at one site in the genome, as described in this work previously for pSET152.

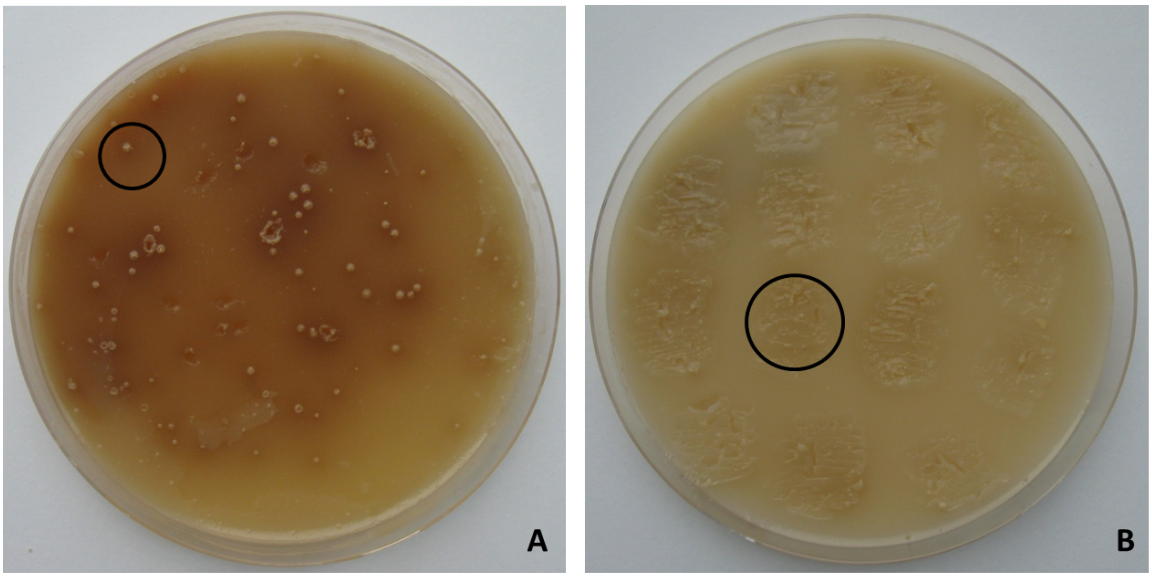

Fig. 19. Colonies of *Actinoplanes* sp. SE50/110 exconjugants on the overlaid SFM mating plates after six days of growth (A) and their respective clones, on the SFM with apramycin, after six days of growth (B).

In this experiment, *E.coli* - *Actinoplanes* conjugation matings were set in five replicates following the procedure described earlier. In all of the matings, colonies of exconjugants were detected soon after overlaying, similar to pSET152-carrying strains. However, exconjugant colonies appeared to be not viable, after being trans-

ferred to SFM media, supplied with apramycin (Figure 19). As all multiple attempts to perform this experiment gave the same result, it is logical to conclude that it is not possible to overexpress *acb* gene cluster in this particular experimental set up. There are following possible explanations for this phenomena: 1) additional copies of *acb* gene cluster are harmful for the survival of *Actinoplanes* sp. SE50/110; 2) insertion of a big DNA segment into chromosome causes instability of the genome, therefore leading to instability of exconjugants.

3.3.4 Transfer of a set of replicative vectors to *Actinoplanes* sp. SE50/110 and characterization of exconjugant strains

Replicative plasmids are recognized as invaluable tools to study secondary metabolism of Actinobacteria, and their application to *Actinoplanes* would offer a greater flexibility in future genetic manipulations. Furthermore, it creates a possibility to rationally construct stable industrial overproducer strains. With this in mind, we set out to test several well known replicative streptomycetes vectors as possible vehicles for DNA transfer into *Actinoplanes* sp. SE50/110. Till so far, successful transfer of four integrative vectors, which are based on three different integration systems, has been demonstrated here. All of this vectors are being maintained as a single copy insertions into genome. To offer a higher flexibility, three replicative plasmids, which are expected to vary in their copy number and stability were further applied: pKC1218 (SCP2* replicon, low copy number), pKC1139 (pSG5 replicon, moderate copy number), pSOK101 (pIJ101 replicon, high copy number). All of the vectors carry apramycin resistance gene, which was shown to be effective for *Actinoplanes* earlier, and similarly to integrative vectors, are all bifunctional, capable to replicate in both *E. coli* and Actinobacteria. This type of vectors is very attractive as it is easier to manipulate DNA in *E. coli* than in any other bacteria. Some of these vectors, however, can suffer instability in host bacteria. This has been a serious problem, particularly with pIJ101 derivatives [68]. pSG5 replicon-based plasmids are known to

be temperature-sensitive and occur at a copy number of 20-50 per cell. Replication of such plasmids stops above 34 °C, resulting in gradual plasmid loss [68]. In comparison to pIJ101-based plasmids, pSG5 offers a broader host range and stability features.

Number	Vector name	Replicon	Transformation frequency	Reference
1.	pKC1139	pSG5 replicon	$2.6 \pm 0.6 \times 10^{-3}$	[68]
2.	pSOK101	pIJ101 replicon	–	[144]
3.	pKC1218	SCP2* replicon	–	[68]

Table 6
Overview of all replicative vectors, used in this study.

The same as before, methylation deficient *E. coli* ET12567 (pUZ8002), carrying replicative vectors was used as donor strain in all conjugation experiments. In all sets of matings, exconjugant colonies with high conjugation frequencies were received only for pKC1139 vector, but not for other Section 3.3.4. Authenticity of exconjugant colonies was checked with the help of PCR, utilizing primers to internal parts of the pKC1139 vector - *oriT* regions and *aac*(3)IV. As an additional step of verification, pKC1139 was isolated from *Actinoplanes* mycelia cells and mapped with restriction endonucleases.

In a similar experiment, set up for *A. teichomyceticus*, no exconjugants were received for pSOK101 vector [59]. It can be explained by the fact, that pIJ101 replicon is unable to function in members of genus *Actinoplanes*. It was also noted earlier, that even in streptomycetes this replicon is not always efficient [68]. Also, plasmids with a very high copy number are shown to be often segregationally unstable.

Efficiency of pKC1218E transfer to *A. teichomyceticus* was characterized by authors as extremely low [59]. As in case of *Actinoplanes* sp. SE50/110, absence of exconjugants can also be explained by instability of SCP2* replicon.

Further on, verification of possible influences of pKC1139 vector on *Actinoplanes* sp. SE50/110 morphology and acarbose biosynthesis was carried out. Similarly to previous experiments, there is no visual change of sporulation patterns or morphology of *Actinoplanes* pKC1139 strain in comparison to wild type Figure 15D.

As it can be seen from Figure 20, presence of pKC1139 plasmid does not change significantly growth of *Actinoplanes* or its acarbose biosynthesis.

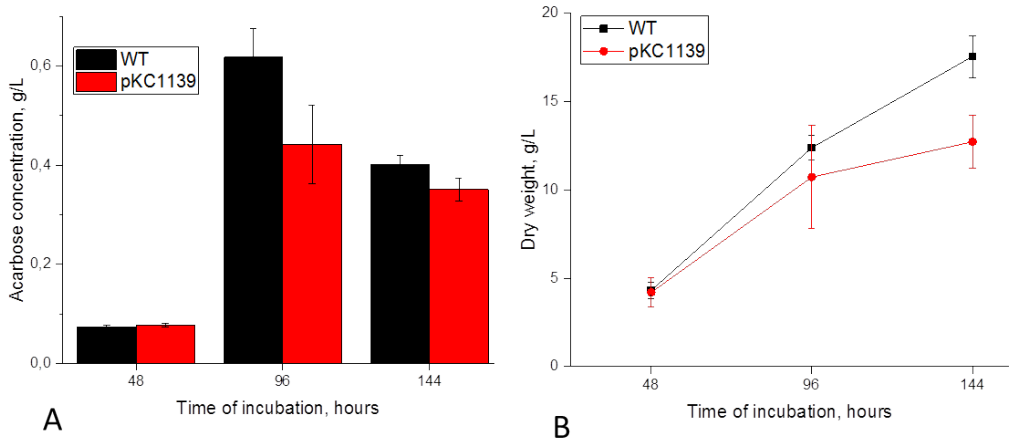

Fig. 20. Acarbose concentration for cultivations of *Actinoplanes* sp. SE50/110 wild type without a plasmid and the pKC1139 exconjugant strain, grown in Mal-MM. A – recorded values for acarbose concentration, g/L; B – recorded values for cell dry weight, g/L. Mean values and standard deviations are shown for three biological replicates.

To examine the stability of pKC1139 vector in cells of *Actinoplanes*, a plate-based test, described here earlier for integrative vectors, was conducted with minor modifications (for details please refer to Materials and Methods chapter). Liquid cultures of *Actinoplanes* pKC1139 strains were incubated in parallel at 28 °C and 37 °C for three subsequent generations. To asses the number of apramycin-resistant colonies in each round for both conditions, at least 200 colonies were transferred to SFM with and without antibiotic. Stability was determined as a ratio of apramycin-

resistant colonies to all colonies, which were tested. Respectively for 37 °C and 28 °C, stability of vector has ranged from 97.3% and 97.5% in first generation, 86.7% and 95.3% in second generation and 82.6% and 94.2% in third generation.

Judging from these results, by contrast to streptomycetes, pKC1139 is stable in cells of *Actinoplanes* even in case of exposure to an elevated temperature culture conditions. This fact could be regarded as beneficial for a multiplication of particular genes in *Actinoplanes* and their stable maintenance in multiple copies during several rounds of incubation. Nevertheless, it can be detrimental for application of pKC1139 to gene knockout experiments. Both of these applications are discussed in details in following chapters.

3.4 Development of techniques for the construction of gene knockouts in *Actinoplanes* sp. SE50/110

Methods regarding gene replacement and gene disruption in actinomycetes can be divided into two major branches: methods to create DNA constructs and to deliver them into respective Actinobacteria strain. Later were discussed in this work in previous chapters. Regarding creation of knockout DNA constructs, several possibilities are usually considered for Actinobacteria: 1) insertional inactivation via a single crossover; 2) insertional inactivation via double crossing over; 3) insertional inactivation using an in-frame deletion [68]; 4) excision of targeted DNA fragment via the action of site-specific recombinases (SSRs) and related techniques [92, 52, 46]; 5) precise editing of DNA by means of CRISPR-Cas9 related techniques [14]. At the moment, however, those methods are frequently combined and given names as particular "systems" or "technologies". One of them, the ReDirect technology, which was formerly developed for gene knockouts in *S. coelicolor* [46], is applied here. Apart from that, in this chapter, multiple possibilities for the generation of gene knockouts are step-by-step tested for their applicability in *Actinoplanes*, beginning from the

most frequently used techniques to the application of complex systems, e.g. ReDirect and CRISPR-Cas9.

3.4.1 Knockout experiments on the base of integrative and replicative vectors

As the first target in gene knockout experiments, *acbD* gene was selected. This gene belongs to *acb* cluster, coding for a putative acarviosyltransferase enzyme. Due to the previous high interest in this enzyme and general belief, that it plays a significant role in acarbose biosynthesis, it was used as the first and proof-of-principle target in this work.

In this method, a 1 kb DNA fragment internal to the *acbD* gene was cloned into a multiple cloning cite of suicide pKGLP2 vector, which contains hygromycin resistance gene and *gusA* gene [93], both suitable for the selection in *Actinoplanes* sp. SE50/110, as it was shown in earlier experiments. The final construct pKGLP2acbD was transferred to *E. coli* ET12567 (pUZ8002) strain after verification. This construct is supposed to integrate into the chromosome by a single homologous crossover. As a result, chromosome should contain an integrated copy of the vector flanked by two mutant alleles of the gene, one truncated at the 3' end and the other at the 5' end. The selection for exconjugants is carried out with addition of hygromycin into overlaying solution. No *Actinoplanes* sp. SE50/110 exconjugant clones were obtained in these conjugation experiments after five separate conjugation experiments. The scheme of this experiment can be found in Figure 21 A.

Further on, similar experiment was conducted using pKC1139 as a cloning vector. Temperature-sensitive plasmids, such as the pSG5-derived vector pKC1139, are frequently used for knockout experiments in Actinobacteria [68]. The presence of temperature-sensitive replicon is necessary in this case for the final selection of plasmid-lacking knockout exconjugants. In this case, second approach, insertional inactivation via double crossing over, was applied. In this method, antibiotic resis-

tance gene is put into a cloned copy of the target gene. Created construct, carried on a cloning vector, is then used to replace the chromosomal copy of the target gene via two crossovers, one on either side of the marker. It typically results in a stable mutation that can not be reverted [68]. To achieve it, 8 kb fragment, containing *acbD* gene was cloned into pKC1139hyg plasmid. This vector contains hygromycin resistance gene instead of apramycin resistance gene, as in original pKC1139. Complete ORF of *acbD* gene was substituted for a resistance cassette, based on aac(3)IV gene, coding for apramycin resistance gene, via ReDirect procedure, which will be discussed in details here later. The resulting construct was verified by PCR and restriction endonucleases analysis and further transferred to *E. coli* ET12567 (pUZ8002) strain. A detailed scheme of this experiment can be found in Figure 21 B. This approach also did not result in successful disruption of *acbD* gene, hence no exconjugants were detected in all mating experiments.

Possibly, the failure to apply these two approaches, with utilization of integrative and replicative vectors, could be explained by the low recombination frequencies between homologous regions of pKGLP2acbD and pKC1139acbDaac(3)IV and chromosome of the wild type. This could occur, as the overlapping sequences in both cases were not large enough for recombination to be executed successfully. This problem can be solved while using cosmids, which harbour large genomic DNA sequences of the wild type.

3.4.2 Application of ReDirect technology for gene knockouts in *Actinoplanes* sp. SE50/110

The next step in constructing gene knockout mutants of *Actinoplanes* was to apply ReDirect procedure. The principle of this procedure is described in [46] and a general overview of ReDirect as modified and applied here, can be found in Figure 22.

The strategy for ReDirect can be characterized as PCR-targeting, with an open reading frame of the targeted gene within respective cosmid, being replaced with

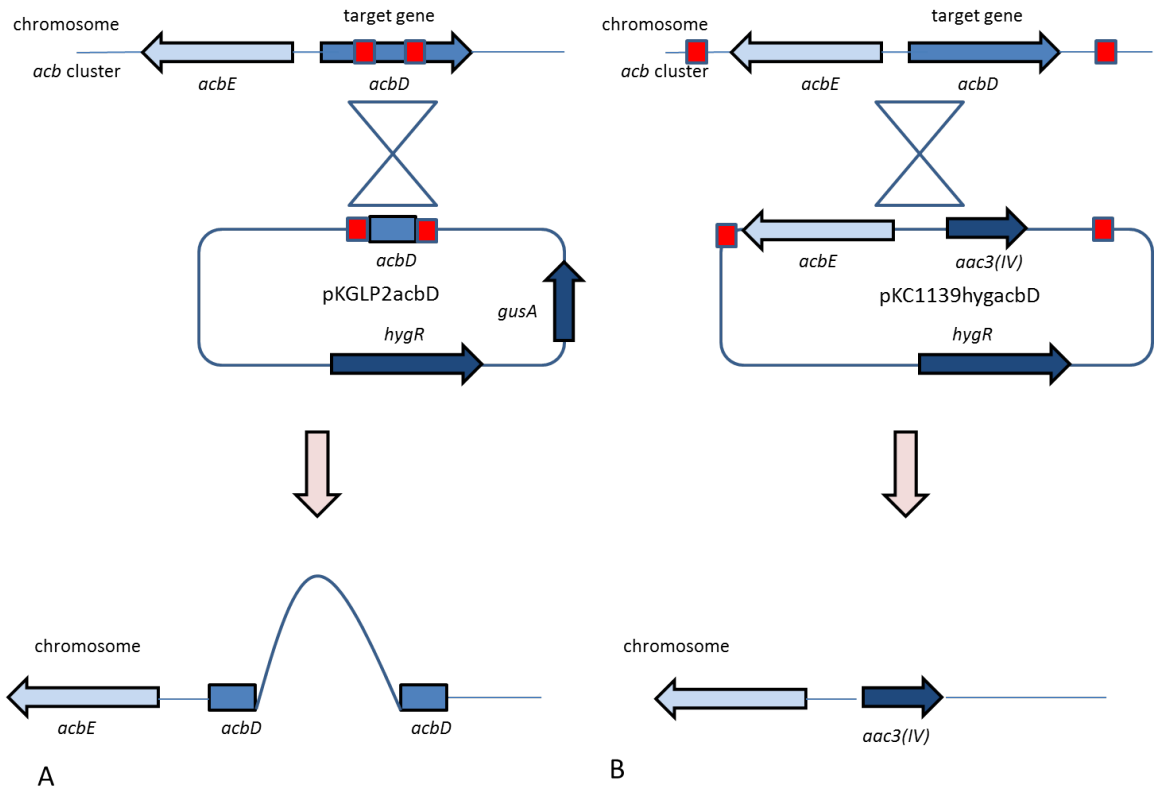

Fig. 21. General scheme, depicting construction of *Actinoplanes* knockout mutants with the use of integrative (A) and replicative (B) vectors.

antibiotic resistance marker. This marker is generated by PCR, as a part of antibiotic resistance cassette, and contains 39 bp extensions, homologous to the sequence before the beginning and the end of the targeted gene, including its start and stop codons. Such cassette typically includes *oriT* region, which allows the whole construct to be transferred by conjugation mating later on, and two FRT sites serving as FLP recombination targets. These sites can be utilized later to excise the marker from the genome of the mutant strain Figure 22 1b. All steps of PCR targeting are conducted solemnly in the cells of specifically constructed strain *E. coli* BW25113, which contains a pKD46 plasmid. This plasmid carries lambda recombination system genes, necessary for the exchange of the targeted gene to a resistance cassette. These

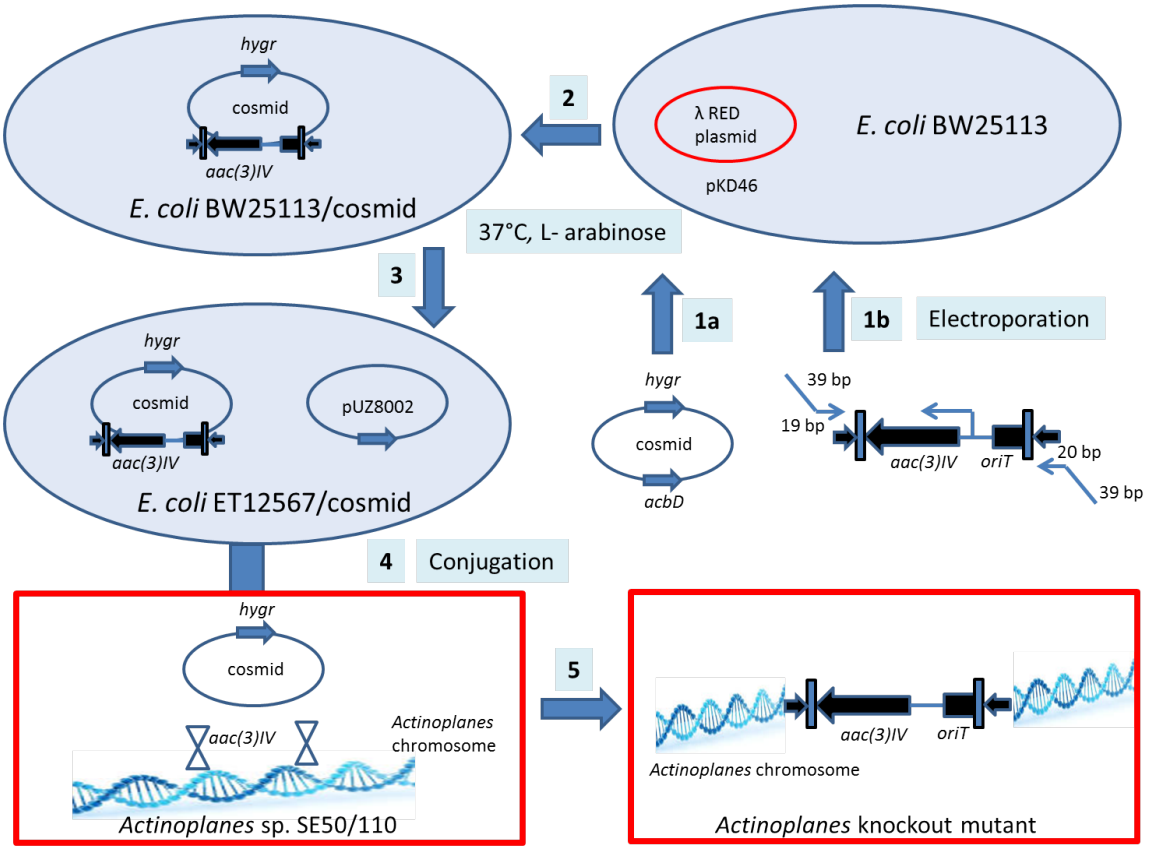

Fig. 22. General scheme, depicting construction of knockout mutant of *Actinoplanes* with ReDirect system. A procedure is executed as follows: 1a - electroporation of the *Actinoplanes* cosmid, which contains gene of interest, to *E. coli* BW25113/pKD46; 1b - electroporation of PCR-synthesized antibiotic resistance cassette to *E. coli* BW25113/pKD46/cosmid; 2 - induction of recombination with L-arabinose and subsequent incubation at 37 °C to eliminate pKD46 plasmid; 3 - electroporation of cosmid, which contains disrupted gene to *E. coli* ET12567/pUZ8002 strain; 4 - conjugation between *E. coli* ET12567/pUZ8002/cosmid and *Actinoplanes* sp. SE50/110; 5 - selection of *Actinoplanes* knockout strain.

genes can be induced by addition of L-arabinose to the cultivation media. As the first step, cosmid which contains a large part of a target strain genome and gene of interest within it, is transferred to *E. coli* BW25113 through electroporation Figure 22 1a. Than, the resulted strain is electroporated with the previously synthesized linear

PCR product Figure 22 1a. The addition of L-arabinose and temperature being elevated to 37 °C, allow the recombination to occur, and simultaneously to eliminate the lambda recombination plasmid from cells of *E. coli* BW25113 Figure 22 2. At this step, cosmid is being isolated from *E. coli* BW25113 and the knockout of gene of interest is being verified with PCR. After the confirmation of successful knockout, cosmid is being electroporated to *E. coli* ET12567/pUZ8002 strain. The passage of the cosmid DNA through non-methylating *E. coli* ET12567/pUZ8002 strain is necessary to circumvent a potent methyl-specific restriction system of *Actinoplanes* sp. SE50/110. During the typical *E.coli-Actinoplanes* conjugation mating, cosmid is being introduced to *Actinoplanes* cells. On this step, exconjugants are being selected in respect to the resistance marker, which was used for the gene knockout in antibiotic resistance cassette. To guarantee the occurrence of the second crossover and loss of the cosmid backbone, exconjugants are selected in parallel by their sensitivity to an antibiotic, resistance to which is conferred by marker gene of the cosmid backbone. For each knockout, at least three selected exconjugants are verified with the help of PCR and, in case of need, with genome sequencing.

In all ReDirect-mediated experiments the advantage was made of the existent cosmid library, formerly created in the Center of Biotechnology for the sequencing of *Actinoplanes* sp. SE50/110 strain, on the base of pCC1BAC cosmid. At first, a general list of future possible knockout targets was established, which included: 1) *acb* cluster genes; 2) possible regulators of acarbose metabolism; 3) genes of other clusters of secondary metabolites, e.g. cACPL4 cluster genes; 4) other genes, possibly involved in the metabolism of acarbose, e.g. *cgt* gene.

Clones of *E. coli*, which contain cosmids harboring potential genes of interest, were verified first for their viability on agar plates with addition of chloramphenicol. Next, colony PCR was performed with each respective viable *E. coli* clone to verify, whether they contain genes of interest, typically with primers to the internal regions of the gene. Altogether, 55 different library clones were checked in the course of this study. Then, the most optimal for the knockout cosmid for each par-

ticular target gene was selected, judging from the genomic region, covered in the cosmid. List of cosmids used for gene knockouts in this study is given in Table 14. According to the antibiotic resistance spectra of the wild type strain, as studied in a previous experiments, chloramphenicol resistance gene, embedded into pCC1BAC cosmid, can not be used for the selection in *Actinoplanes*, therefore it was substituted with more suitable hygromycin resistance gene hyg^{r}. To achieve it, PCR-targeting was performed for each cosmid. To check the correctness of ReDirect, isolated cosmids were tested with PCR, utilizing primers, adjacent to the chl^{r} gene. Received PCR products were sequenced with the help of in-house sequencing facility. Next, verified cosmids were electroporated again to *E. coli* BW25113/pKD46. Individual antibiotic resistance cassettes for each gene of interest were synthesized by PCR with the help of specially designed "long primers", as advised in [46] and electroporated into *E. coli* BW25113/pKD46 clones with respective cosmids. To confirm successful PCR-targeting of target gene, colony PCR was performed with five colonies of *E. coli* BW25113/pKD46 for each target gene, grown on agar plates with addition of apramycin and hygromycin. Cosmids were once more isolated from positive colonies and tested with PCR. PCR products, containing regions around targeted genes were sequenced to prove the correctness of PCR-targeting.

Details, regarding origin and construction of cosmid library, also primers used to verify and construct gene knockouts, can be found in Materials and methods section of this work.

As an example, knockout of *acbD* gene is described in details here. Disruption of *acbD* was performed on a $aut_f2p0002h06.x$ cosmid (from here noted as h06). In the existent library three cosmid clones in total were found to contain parts of *acb* cluster, however h06 cosmid contained the majority of the *acb* genes Figure 25. PCR-targeting procedure was performed twice on a h06 cosmid 1) chloramphenicol resistance gene chl^{r} was substituted for hygromycin resistance gene hyg^{r} (h06hyg construct); 2) *acbD* gene was substituted for apramycin resistance gene *aac*(3)IV (h06hygaac(3)IV construct). After both procedures constructs were checked with PCR. In the first case,

colony PCR with TG15/TG16 primer pair was set to confirm the gene substitution Figure 24A. In the second case, TG17/TG18 primer pair was used in PCR to confirm the second gene substitution Figure 24B. The schematic representation of ReDirect knockout cassette for h06hygaac(3)IV construct can be found in Figure 23.

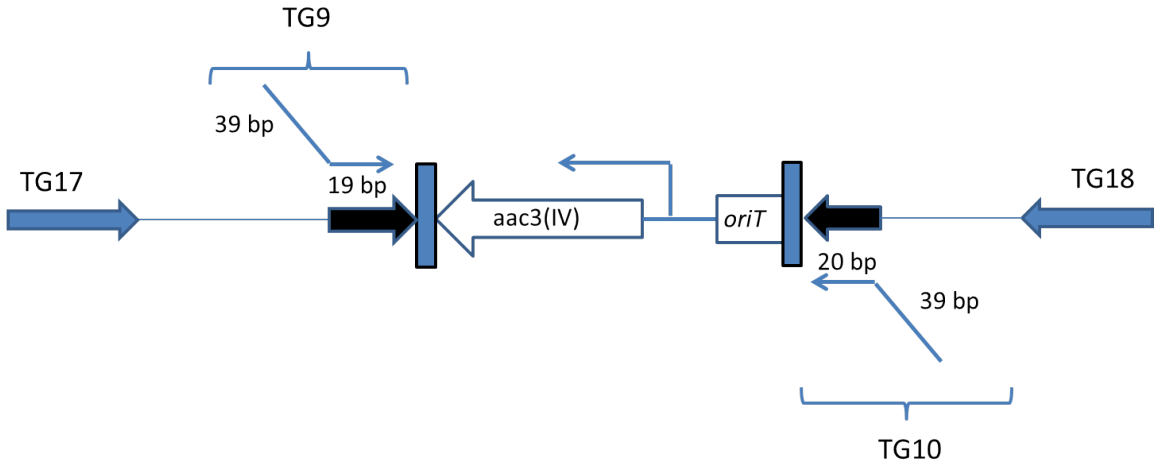

Fig. 23. The schematic representation of the construction of ReDirect knockout cassette for h06hygaac(3)IV construct. Primers TG9 and TG10 were used to amplify the ReDirect knockout cassette. These primers consist of two parts: 39 bp, which are taken from genomic sequence and 19 or 20 bp (marked with black arrows), which are standard in all ReDirect primers. Primers TG17 and TG18 were used to verify the final construction product via PCR (marked with blue arrows).

After verification h06hygaac(3)IV construct was transferred to *Actinoplanes* sp. SE50/110 cells via conjugation. Clones, obtained after conjugation, were transferred to SFM plates with hygromycin and apramycin. Strains No. 44, 49, 59, and 66, which have shown resistance to apramycin and sensitivity to hygromycin, were selected for further studies Figure 24D. For a confirmation, total DNA was isolated from all of the above mentioned strains and PCR products, containing *acbD* gene were sequenced.

Additionally, single knockouts of other *acb* cluster genes were performed on the same cosmid in parallel Figure 25, Table 14. *E. coli* BW Δ*acbU*, Δ*acbS*, Δ*acbR*,

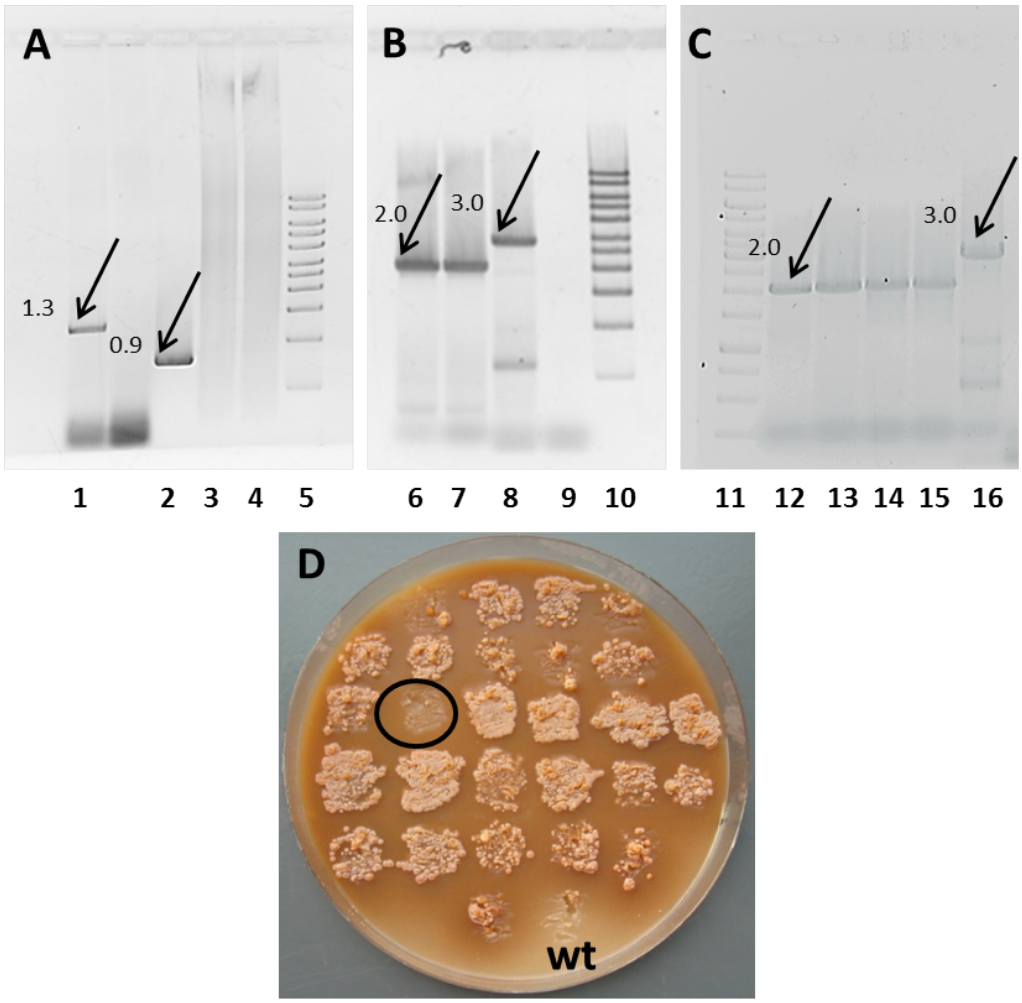

Fig. 24. Selection of the colonies during ReDirect with the help of PCR (A, B, C) and antibiotic sensitivity tests (D) in order to construct *Actinoplanes* Δ*acbD* strain; A, B and C - PCR products, separated in agarose gel; PCR reaction is set with TG15/16 (A) and TG17/18 (B, C) primer pairs. In case of A, B and C - DNA isolated from *E. coli* BW h06hyg, *E. coli* BW h06hygaac3(IV) and *Actinoplanes* Δ*acbD* strains respectively, is used as the matrix for PCR reaction. D - SFM plate with addition of hygromycin; selection of colonies of *Actinoplanes* Δ*acbD* mutant after conjugation procedure is carried out on SFM plates with addition of hygromycin and apramycin.

Δ*acbP*, Δ*acbI*, Δ*acbJ*, Δ*acbO*, Δ*acbK*, Δ*acbM*, Δ*acbL*, Δ*acbC*, Δ*acbE*, Δ*acbA*, Δ*acbB* and Δ*acbQ* single knockout mutants were constructed and verified as de-

scribed above for the *acbD* gene knockout. In addition to *Actinoplanes* Δ*acbD*, two more knockout strains *Actinoplanes* Δ*acbA* and *Actinoplanes* Δ*acbB* were constructed. As a part of an effort to study other metabolic gene clusters of *Actinoplanes*, several genes of cACPL4 cluster were deleted via ReDirect. Namely, *E. coli* BW and *Actinoplanes* Δ*ACPL*_6141, Δ*ACPL*6142-53, Δ*ACPL*6139-40, Δ*ACPL*6142-53 knockout strains were constructed. However, due to the time limitation of this study, these and other knockouts, created in the course of this study, further are being studied in *Actinoplanes* working group by my colleagues. Recently, mutant strain *Actinoplanes* Δ*malR* strain was characterized in [27] and a research paper, currently unpublished (Droste and Wolf; 2017, unpublished; submitted to BMC Genomics on 26.01.2017).

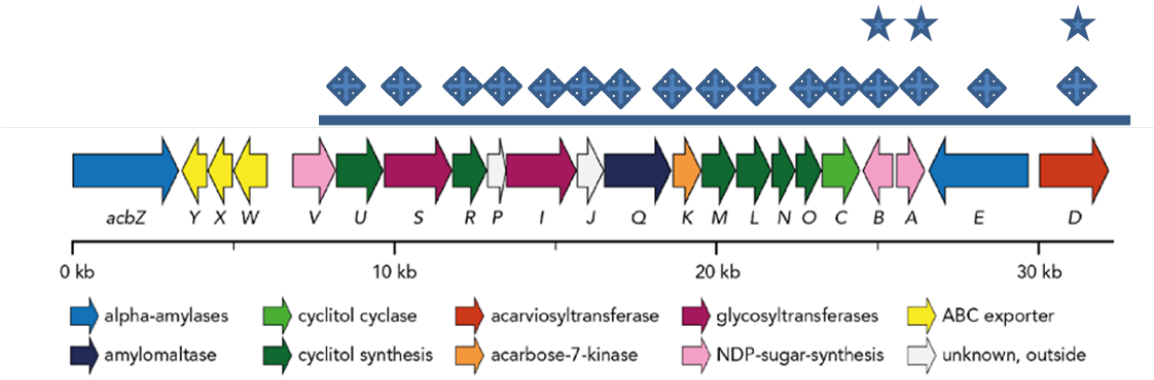

Fig. 25. Depiction of *acb* cluster genes, for which ReDirect knockout constructs were generated. A region, which belongs to $aut_f2p0002h06.x$ cosmid is depicted with a blue line. A single knockouts, performed in BW 25113/pKD46 are depicted as squares above respective genes. Single knockout constructs, successfully transferred to *Actinoplanes* sp. SE50/110 are shown as stars above respective *acb* genes. Modified from figure in [110], the Open Access article distributed under the terms of the Creative Commons Attribution License (http://creativecommons.org/licenses/by/2.0).

In this study, additionally to *Actinoplanes* Δ*acbD*, knockout of a putative regulator of acarbose metabolism, ACPL_212 *cadC* gene is described. Disruption of *cadC*

was performed on a $aut_f1p0005h12.x$ cosmid (from now called h12). In the cosmid library, used in this study, three cosmid clones in total were found to contain *cadC* gene. However, only h12 cosmid clone was found to be viable (Table 14). PCR-targeting procedure was performed twice, similar to as described previously, on h12 cosmid 1) chloramphenicol resistance gene chl^r was substituted for hygromycin resistance gene hyg^r (h12hyg construct); 2) *cadC* gene was substituted for apramycin resistance gene *aac*(3)IV (h12hygaac(3)IV construct). Both constructs were checked with PCR, similar to as described for the knockout of *acbD* gene. After verification h12hygaac(3)IV construct was transferred to *Actinoplanes* sp. SE50/110 cells via conjugation. Clones, obtained after conjugation, were transferred to SFM plates with hygromycin and apramycin. However, neither in first nor in second or third round of selection on plates, any knockout colonies were identified. Therefore, to facilitate the second crossover, spores of each selected three strains after conjugation were inoculated in liquid media for two days at 28 °C. The culture was plated after each round at SFM plates without antibiotic supplement and let to sporulate. Spores were harvested and inoculated again into liquid culture. After three rounds of cultivation in liquid minimal media, culture was plated again on SFM media and 126 colonies were verified again on SFM plates, supplemented with apramycin and hygromycin. Three strains, numbered 11, 50 and 65, which have shown resistance to apramycin and sensitivity to hygromycin, were selected for further studies. For a confirmation, total DNA was isolated from all of the above mentioned strains and PCR products, containing *cadC* gene were sequenced. Analysis of the *Actinoplanes* $\Delta cadC$ strain phenotype is discussed in chapters below.

As a continuation of work on the methods and systems, useful for the knockout, silencing or in general manipulation of genes in *Actinoplanes* sp. SE50/110, CRISPR-Cas9 technology was adapted. This work was conducted in collaboration with Timo Wolf and other colleagues and published in [137].

In comparison to other methods of genetic engineering CRISPR-Cas9 is easy to apply and can be great in its labor and time saving. Application of CRISPR-Cas9

system requires a working system for the transfer of genetic material and system of vectors, which can be transferred to the host. Earlier in this work, *E. coli* - *Actinoplanes* conjugation system was established and verified. Also, a set of various vectors was applied. In case of CRISPR-Cas9 type of system, a replicative vector, typically with a temperature sensitive replicon is needed. As long as applicability of pKC1139 vector was already shown in this work, the similar vector, pCRISPomyces-2, carrying a *cas*9 gene, codon optimized for streptomycetes was selected. Plasmid pCRISPomyces-2, similarly to pKC1139 possesses a pSG5 temperature sensitive replicon, which omits its replication at elevated temperatures and leads to the plasmid evacuation from growing cells. However, as it is shown in this work earlier, pKC1139 is less sensitive to the temperature in case of *Actinoplanes* sp. SE50/110. Therefore, in future research more options have to be exploited in order to find alternative to this vector.

Further on the published protocol of application of CRISPR-Cas9 for streptomycetes system was followed [19]. Functionality of CRISPR-Cas9 system was than shown by the scarless deletion of the gene encoding the tyrosinase MelC, which putatively catalyzes the formation of the dark pigment eumelanin. It was than shown, that the *Actinoplanes*Δ*melC*2 mutant is no longer capable of producing dark brown pigment, probably eumelanin, neither while cultivated in the liquid growth media, nor solid growth plates. Genomes of several strains of knockout mutants were sequenced in order to rule out possible off-target effects and presence of other mutations.

3.5 Application of the developed techniques for the knockout of the acarviosyl-transferase encoding gene *acbD* from *acb* gene cluster in *Actinoplanes* sp. SE50/110

The *acbD* knockout strain was constructed and used as the first proof-of-principle target to demonstrate the usefulness of the developed gene cloning system in this work. The decision to study the effects of knockout of this particular gene, and not

any other of *acb* cluster, stems from multiple reasons. Firstly, the product of this gene, so called acarviosyl transferase or ATase, was already produced, purified, sequenced and studied heterologously in *S. lividans* [51]. It was determined, that AcbD does not posses neither α-amylase, nor cyclodextrin glycosyltransferase (CGTase) activities, as it might be suggested from its predicted domain structure. Bioinformatic analysis of the putative structure of AcbD protein reveals four domains, two α-amylase, one starch-binding and one which belong to IPT superfamily of proteins (Figure 26). However, AcbD exhibits solely transglycosylation activity, specifically on donor substrates, which possess acarviosyl moiety. It was also shown to accept a wide range of substrates, but none of them was as effective as maltose itself [51].

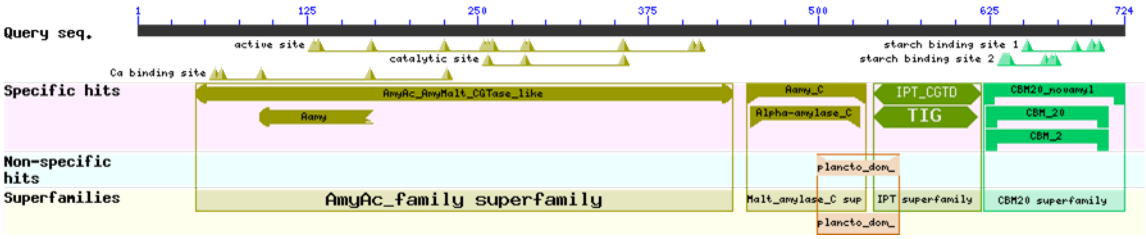

Fig. 26. Bioinformatic prediction of the putative domain structure of AcbD protein, performed in correspondence to [83]. Protein superfamilies are assigned due to the Pfam search [31]. Starch and ion binding sites are marked with triangles.

On the base of these studies possible ecological role of acarbose as a "carbophore" was postulated [129]. Later on, with the further accumulation of proteomic and transcriptomic data regarding characterization of *Actinoplanes* growth in various media, further suggestions regarding the function of AcbD appeared [131, 97]. This study gives us a possibility to partly verify these, previously formulated assumptions. Secondly, *acbD* gene is situated on the border of the *acb* cluster and is expressed as a single transcription unit. It makes this gene the perfect target for gene manipulations, as long as its knockout should not cause any polar effects on the neighboring genes of the cluster.

3.5.1 Characterisation of *Actinoplanes* Δ*acbD* strain phenotype

The complete workflow to construct *acbD* mutant strain is discussed in details in previous section.

The whole genome sequence of mutant strain 49 was determined, similarly as its described earlier in this work for *Actinoplanes* exconjugants strains. It was confirmed, that the substitution of *acbD* gene is carried out as according to the procedure of ReDirect. Importantly, absence of any other major mutations in genome of *Actinoplanes* sp. SE50/110 Δ*acbD* 49 was confirmed.

To characterize the phenotype of *acbD* mutant strain, a set of experiments was set up: 1) growth studies in minimal liquid media with various carbon sources and analysis of metabolite production in these media; 2) BioLog phenotypic microarrays; 3) transcriptomic and proteomic studies. Additionally, complementation and overexpression of *acbD* gene were carried out and comparison of its effects on acarbose production was carried out in minimal liquid media with various carbon sources.

3.5.1.1 Characterization of *Actinoplanes* Δ*acbD* strain growth and acarbose production in minimal media, supplemented with various carbon sources

In order to get the first impression regarding the effects of mutation of *acbD* gene on acarbose production, the small scale cultivation experiment was set up. Both *Actinoplanes* Δ*acbD* 49 strain and wild type strain were cultivated in minimal media with addition of various carbon sources, namely Cpur (mixture of maltose and maltotriose), maltose and glucose. As it is clearly seen from the growth curve, growth of 49 strain was similar to a growth of wild type strain in most of the cases (see Supplementary material, Figure 44). However, utilization of glucose as a sole Carbon source, seemed to considerably delay the growth of mutant in comparison to the wild type.

To characterize the amount of acarbose, produced by *Actinoplanes* Δ*acbD* 49 strain, HPLC measurements were made every 24 hours, the last one at the stationary growth phase of 144 hours. Samples were taken from cultures grown in Cpur and maltose minimal media. Three biological replicates were made for each strain. On glucose, as a sole Carbon source, no acarbose could have been detected neither in case of wild type, nor 49 mutant. As a negative control in this experiment, media without any Carbon source was used. In this case only background growth of both strains was observed.

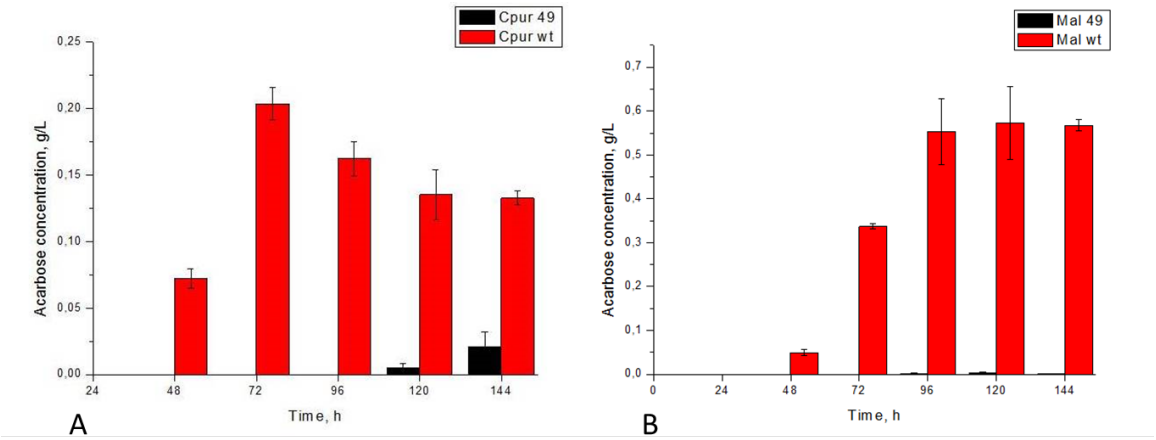

Fig. 27. Concentration of acarbose, produced by *Actinoplanes* sp. SE50/110 and *Actinoplanes* Δ*acbD* 49 strains grown in a minimal media, supplemented with Cpur (A) and maltose (B). Mean values and standard deviations are shown for three biological replicates. All calculation were done as described in [132].

As it is clearly seen from the results of HPLC measurements (Figure 27), 49 strain can be characterized as deficient in acarbose biosynthesis in a comparison to wild type. In all measurements *Actinoplanes* Δ*acbD* 49 produced low amounts of acarbose, which can be detected only in the stationary phase of growth. In case of maltose supplemented minimal media, amount of acarbose detected was even lower than in maltose/maltotriose supplemented minimal media. To clarify the picture of

acarbose biosynthesis by 49 mutant strain, LCMS measurements of samples from 4 various media: glucose, maltose, Cpur and glucose/maltose were performed.

In order to perform LCMS measurements and verify the previous observation regarding acarbose biosynthesis, separate cultivation experiment was performed. The samples were taken at 144 hours of cultivation, after the onset of stationary phase. The growth curve for each of setups in this cultivation can be found in Supplementary material, Figure 45. The growth of *Actinoplanes* Δ*acbD* 49 in this cultivation was similar to previous, with a significant growth delay in glucose minimal media.

As it is seen from the results of these measurements, *Actinoplanes* Δ*acbD* 49 strain is capable of producing all minor acarviosyl- components, such as acarviosyl-glucose, -maltotriose and -maltotetraose, and acarbose or acarviosyl-maltose, as well. However, amount of all of these compounds is significantly lowered in a comparison to a wild type strain (Figure 28).

Regarding biosynthetic pattern of the mutant, it is significantly different from that of the wild type. In glucose minimal media, no acarviosyl-glucose, but a trace amounts of acarviosyl-maltose were detected. It is very distinctive as in comparison to the wild type, which is producing almost solely acarviosyl-glucose in this media (Figure 28A). When maltose or a mixture of maltose and glucose was used, acarviosyl-maltotriose, not acarviosyl-maltose, was the main product synthesized by Δ*acbD* 49 strain (Figure 28B and C). In case of Cpur, a mixture of maltose and maltotriose, being a sole Carbon source in cultivation media, similarly mutant and the wild type were able to produce a mixture of acarviosyl-maltose and acarviosyl-maltotriose (Figure 28D).

To verify whether this particular phenotype is specific to *Actinoplanes* Δ*acbD* 49 mutant, or a part of a general trend, three other independently received mutants, 44, 59 and 66 were cultivated in maltose minimal media, in parallel to wild type Figure 29. All of these strains were previously checked for the knockout of *acbD* gene with PCR and respective genetic regions, surrounding *acbD* gene, were sequenced with the help of in-house sequencing facility. All strains included into this experiment were shown to be deficient in acarbose production, similarly to *Actinoplanes* Δ*acbD* 49

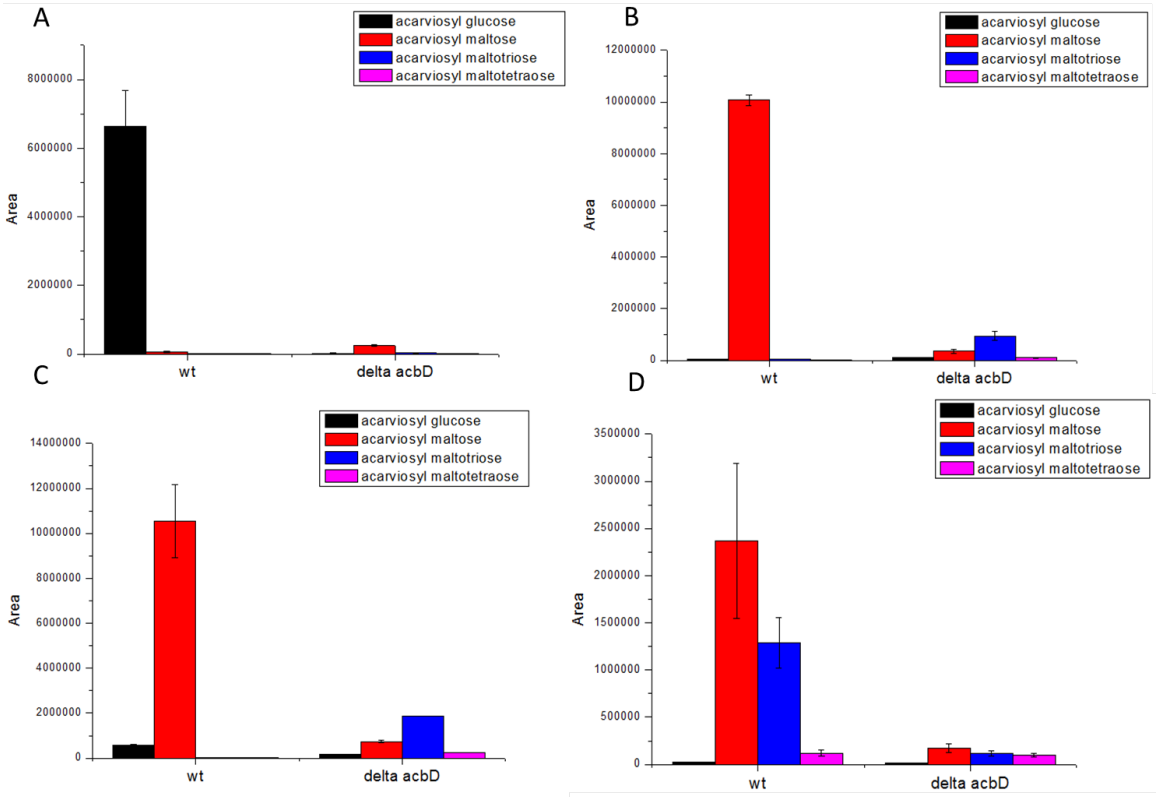

Fig. 28. Biosynthetic activity of *Actinoplanes* $\Delta acbD$ 49 strain and wild type, measured with LCMS; samples are taken at 144 hours cultivation time. Minimal media was supplemented with: A – glucose, B – maltose, C – maltose/glucose, D - Cpur; The presented relative peak areas (Area) of acarviosyl metabolites were obtained by normalizing the peak areas of the respective extracted ion chromatograms to culture cell dry weights. Mean values and standard deviations are shown for three biological replicates. All calculation were done as described in [132].

strain. As in previous experiment, low amounts of acarbose were detected only in the stationary phase of growth. These results, together with whole genome sequencing of *Actinoplanes* $\Delta acbD$ 49 strain, directly indicate that the observed deficiency in acarbose production of $\Delta acbD$ strain is not connected to any handling mistakes during strain construction or any untargeted genetic modifications or rearrangements of the

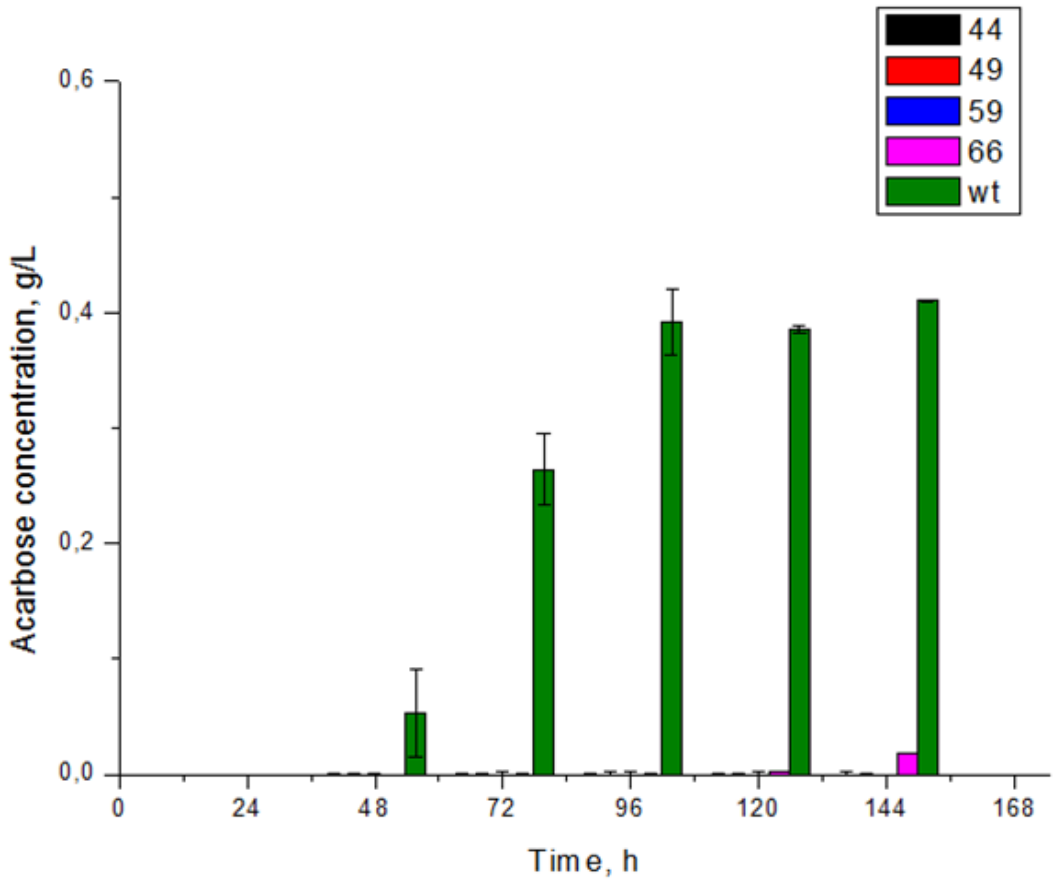

Fig. 29. Concentration of acarbose, produced by *Actinoplanes* sp. SE50/110 and *Actinoplanes* 44, 49, 59, 66 Δ*acbD* strains, grown in a minimal media supplemented with maltose. Mean values and standard deviations are shown for three biological replicates. All calculation were done as described in [132].

genome. An additional way to confirm this conclusion and to exclude possibilities of any other unwanted influences on the phenotype of *Actinoplanes* Δ*acbD* 49 strain is to analyze the phenotype of Δ*acbD* complementation strain.

In order to gain more insight into phenotype of *Actinoplanes* Δ*acbD* mutant, technique of BioLog phenotypic microarrays was used.

3.5.1.2 Analysis of Δ*acbD* strain phenotype with the help of BioLog phenotypic microarrays

Previous phenotypic characterization of *Actinoplanes* Δ*acbD* mutant has left unanswered questions regarding other possible effects of the mutation on the growth of strain in other conditions, e.g. consumption of other Carbon Nitrogen, Sulphur sources. As due to the technical limitations of shake flask cultivations and a subsequent lengthy handling, potential mutant strain auxotrophy or altered growth, can usually be characterized only for a limited amount of factors. For example, in this study, growth of *Actinoplanes* Δ*acbD* 49 strain was previously verified in minimal media with four different carbon sources. However, because of the long cultivation periods, up to 144 hours or longer, and mycelial type of growth of *Actinoplanes*, further testing in shake flask cultures proves to be practically complicated. However, several known research methods, such as BioLog phenotypic microarrays, offer a possibility of multiple parallel testing of various growth factors over a long period of time. Phenotype microarray (PM) technology is used for the metabolic profiling of microbial cells within 96-well plate system. The PM assay allows for cells to be assessed for utilization of various nutrients. The assay utilizes a redox sensitive tetrazolium dye which becomes irreversibly reduced upon detection of cellular metabolic output. Detection process is synchronized with a colour change [113]. For each particular cell in the plate, the kinetic graphs of cell respiration, recorded through the reduction of the dye, are detected. Later on, respective graphs of mutant and wild type are overlayed, and the differences are detected.

In this study, to characterize the growth of *Actinoplanes* Δ*acbD* 49 strain, following sets of compounds were tested: 1) PM1 and PM2 (Carbon utilization); 2) PM3B (Nitrogen sources); 3) PM4A (Phosphorus and Sulphur sources); 4) PM5 (Nutrient Supplements). The full list of each plate components can be found at http://www.biolog.com/. In all tests, 23 compounds were found to cause higher respiration rates for the *Actinoplanes* Δ*acbD* mutant strain and 15 - lower. Quite

noticeably, maltose, glucose, maltotriose and mannitol were causing higher respiration rates. This effect can be associated with the deficiency of *Actinoplanes* $\Delta acbD$ regarding acarbose production. The full lists of compounds, that were found to be utilized differentially by $\Delta acbD$ mutant and wild type can be found in Supplementary material Table 15 and Table 16.

3.5.1.3 Characterization of extracellular proteome of *Actinoplanes* $\Delta acbD$ 49 mutant in comparison to the wild type

In order to verify, whether knockout of *acbD* gene did not change the composition or abundance of *Actinoplanes* extracellular proteins, studies of extracellular proteome were conducted. Extracellular proteins were isolated from cultures of wild type and $\Delta acbD$ 49 mutant, both grown in maltose minimal media at 72 hours time point, the late exponential phase of growth. All preparations and analysis of extracellular proteomes were done in accordance to the protocols, methods and conditions described and developed for the analysis of the extracellular proteins of *Actinoplanes* by Dr V. Ortseifen [97]. After phenol-mediated extraction, proteins were quantified. Later on, after isoelectric focusing, samples in triplicates were loaded onto two dimensional SDS PAGE gel electrophoresis. After the end of separation, gels were stained with Coomassie Briliant Blue and all of the visible spots were manually picked for following MALDI-TOF measurements (Matrix Assisted Laser Desorption/Ionization).

Results of the conducted 2D gels are represented in Figure 30. In total 126 and 122 protein spots were picked for a wild type and *Actinoplanes* $\Delta acbD$ 49 mutant strains respectively. Of this spots, 19 were classified and identified and 3 as uncertain by the BioTools Software using the Mascot search engine for wild type and 19 and 2 for the mutant strain. Importantly, in the respective position, occupied by AcbD protein on the 2D gels of the wild type, other protein ACPL_882 was identified in the case of the mutant. This fact confirms the knockout of the *acbD* gene and absence

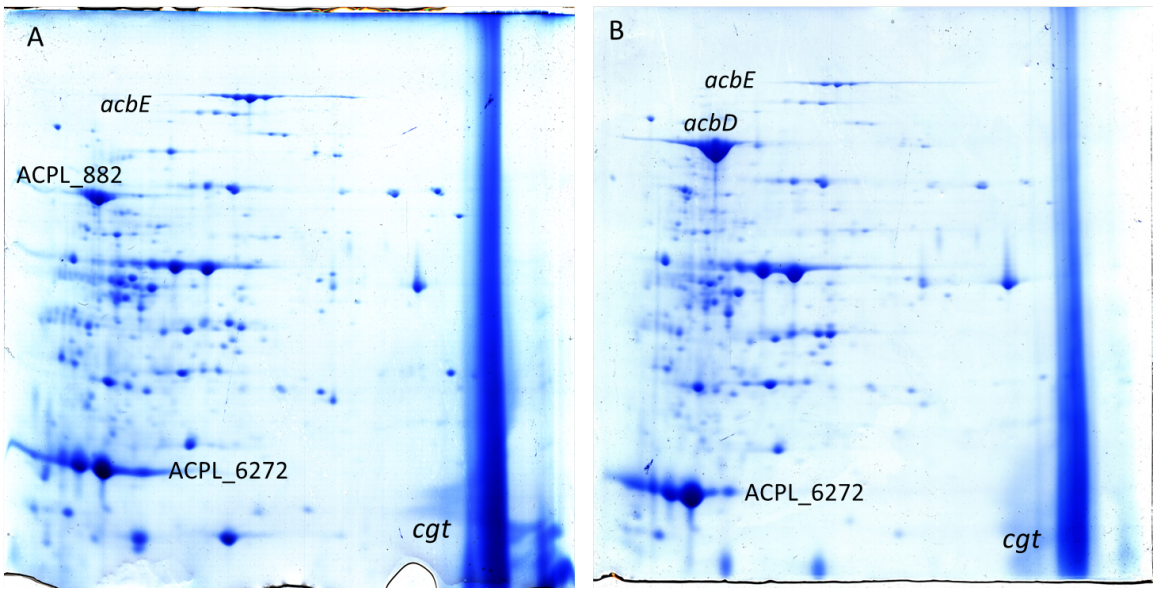

Fig. 30. Comparison of 2D gels, representing extracellular proteomes of *Actinoplanes* Δ*acbD* mutant (A) and *Actinoplanes* sp. SE50/110 (B), grown in maltose supplemented minimal media, 72 hours of growth. Prominent protein spots are highlighted.

of its protein product in the extracellular proteome of the mutant. Otherwise, no significant differences between wild type and the mutant strain were found.

3.5.1.4 Transcriptional characterization of *acbD* knockout strain

To verify whether the expression of *acb* cluster genes, or any other genes, which could be directly or indirectly linked to acarbose metabolism or its regulation, was influenced by the mutation of *acbD* gene and caused the observed phenotype, further transcriptomics-related studies were performed. For these tests RNA samples were isolated from 72 hours grown in maltose minimal media cultures of wild type and *Actinoplanes* Δ*acbD* 49 mutant strain.

As *acbD* gene is located near the border of the *acb* cluster, the possibility of the polar effect on the other genes, due to its knockout, is quite minimal. However, in

order to exclude such possibility and to test the expression of other *acb* genes, at first a qRT-PCR test was conducted. In this test, transcription of *acbD*, *acbE* and *acbZ* was checked. These two particular genes were chosen for this experiment, because of their position on two opposite sides of *acb* cluster (Figure 25). As a control for comparison and normalization, a house keeping gene *rpoB* was used.

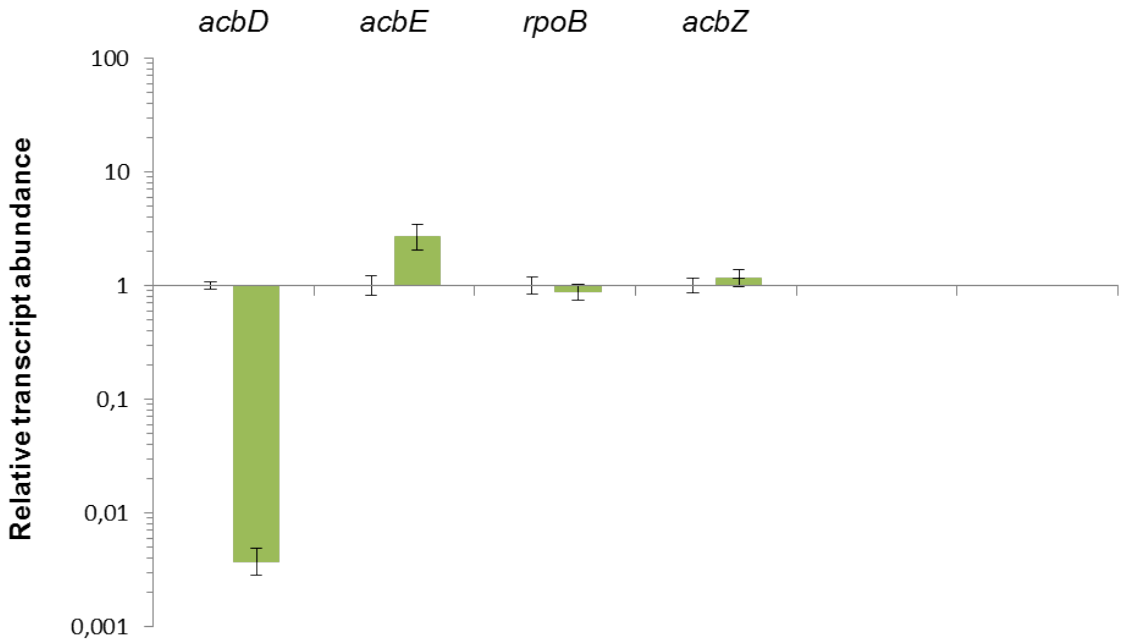

Fig. 31. qRT-PCR measurements of *Actinoplanes* Δ*acbD* 49 RNA samples in comparison to the wild type; samples are isolated from cultures, grown in maltose minimal media, 72 hours of incubation.

As it is seen from Figure 31 there are nearly no transcription signals from *acbD* gene, as expected, due to its open reading frame being completely substituted by a marker gene. From the results of qRT measurements it is clear, that expression of *acbE* and *acbZ* genes is not lowered significantly. Moderately elevated levels of *acbE* expression can be explained by the transcription of *acbE* from the strong promoter of *aac*(3)IV gene, coding for apramycin resistance. This gene was used previously in the knockout of *acbD* gene, as a part of ReDirect substitution cassette.

These results are indicating that there are no significant changes in transcription levels of genes, located in the immediate proximity to *acbD* gene. To verify expression levels of all *acb* cluster genes, approach of comparative microarrays was found to be more suitable, as it allows parallel testing and gives an overview on expression of multiple genes and genetic clusters in *Actinoplanes* genome.

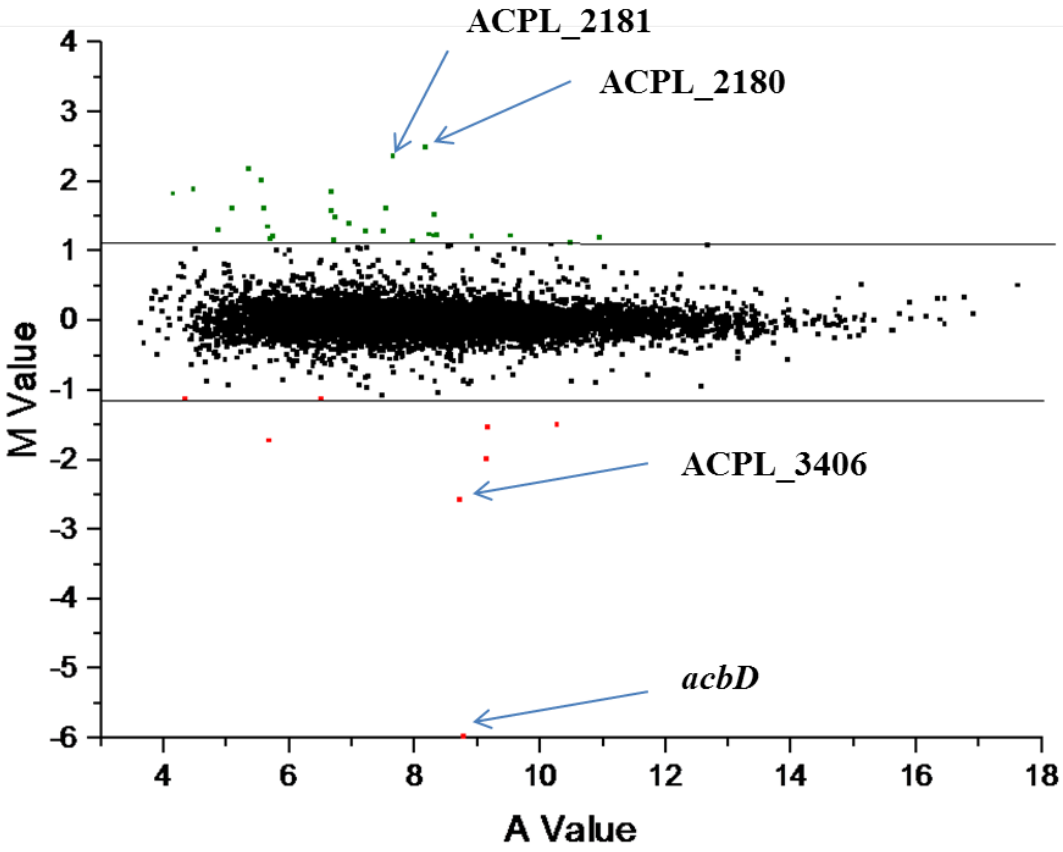

Fig. 32. Plot of the results of whole genome comparative microarrays, performed for the comparison of *Actinoplanes* sp. SE50/110 and *Actinoplanes* Δ*acbD* 49 RNA samples, isolated from cultures grown in maltose minimal media; 72 hours of incubation.

Whole genome comparative microarrays for *Actinoplanes* sp. SE50/110 and *Actinoplanes* Δ*acbD* 49 were prepared using the same RNA samples as in previous exper-

iment, but here three replicates were combined for each strain (Figure 32). As to the protocols, previously established by T. Wolf and Dr A. Neshat, Agilent oligonucleotide microarrays were constructed, consisting of a total of 42,225 features and representing 8,239 genes of *Actinoplanes* sp. SE50/110 genome. Previous adaptation of microarray hybridization procedure was done by T. Wolf as a part of his PhD thesis. The log2 (fold change) cut off (M-value) for a false discovery rate of 0.01 was determined as 1.1 and -1.1, respectively.

List of all significantly differentially expressed genes can be found in Table 7 and Table 8 respectively. Altogether, there were thirty genes, which transcripts were shown to be significantly more abundant and seven genes – less in a mutant strain in comparison to the wild type strain. Importantly, *acb* cluster genes do not belong to this list, with one exception of *acbE* gene, putatively being highly expressed due to the secondary effects as in agreement with previous qRT-PCR measurements.

Among differentially expressed genes, there are no presumable acarbose metabolism regulatory genes, which expression pattern may have been changed in *Actinoplanes* Δ*acbD* 49 mutant strain. These data leads to the suggestion, that *acbD* gene is probably not involved in acarbose biosynthesis regulation and also excludes possibilities of any technical or handling mistakes, connected to *acbD* gene knockout. As long as transcription of neighbor *acb* genes is not changed, there are no visible polar effects of *acbD* knockout.

No	Gene number	Gene name	Gene product	M value	A value
1	ACPL_2180	Not specified (NS)	Putative uncharacterized transporter ygaY;	2.48	8.17
2	ACPL_2181	NS	3-oxoacyl-[acyl-carrier-protein] reductase EC=1.1.1.100;	2.36	7.66
3	ACPL_6840	NS	L-arabinose transport system permease protein;	2.16	5.35
4	ACPL_6839	NS	sn-glycerol-3-phosphate transport system permease protein;	2.01	5.56
5	ACPL_3377	NS	Phytochrome-like protein;	1.88	4.47
6	ACPL_2176	tdh	L-threonine 3-dehydrogenase (EC:1.1.1.103);	1.86	6.68
7	ACPL_6296	NS	hypothetical protein	1.82	4.15
8	ACPL_7905	NS	K02408 flagellar hook-basal body complex protein;	1.61	5.09
9	ACPL_6841	NS	sn-glycerol-3-phosphate-binding periplasmic protein;	1.60	7.54
10	ACPL_6298	ftsK	DNA translocase ftsK;	1.60	5.59
11	ACPL_2175	NS	Short-chain specific acyl-CoA dehydrogenase. mitochondrial;	1.58	6.66
12	ACPL_1608	NS	Uncharacterized protein;	1.52	8.31
13	ACPL_2614	NS	Flotillin-2;	1.49	6.74
14	ACPL_0630	NS	Pre-neck appendage protein;	1.39	6.95
15	ACPL_2542	NS	Inner membrane amino-acid ABC transporter permease protein yecS;	1.35	5.66
16	ACPL_6297	NS	hypothetical protein predicted by Glimmer/Critica;	1.30	4.87
17	ACPL_4522	NS	hypothetical protein predicted by Glimmer/Critica;	1.28	7.21

No	Gene number	Gene name	Gene product	M value	A value
18	ACPL_7899	NS	Flagellin A;	1.28	7.50
19	ACPL_7489	lysK	K05831 acetyl-lysine deacetylase;	1.24	8.23
20	ACPL_5265	NS	hypothetical protein predicted by Glimmer/Critica;	1.23	8.35
21	ACPL_7493	rimK	K05827 lysine biosynthesis protein LysX;	1.22	9.54
22	ACPL_7488	NS	K00615 transketolase	1.22	8.30
23	ACPL_7903	NS	Flagellar basal-body rod protein flgB;	1.21	5.74
24	ACPL_2173	NS	K00120	1.21	5.74
25	ACPL_5150	NS	hypothetical protein predicted by Glimmer/Critica;	1.21	8.91
26	ACPL_3683	acbE	Alpha-amylase EC=3.2.1.1;	1.19	10.95
27	ACPL_7904	flgC	Flagellar basal-body rod protein flgC;	1.18	5.70
28	ACPL_6838	aglA	alpha amylase catalytic region;	1.16	6.71
29	ACPL_4523	NS	RNA polymerase sigma factor sigM;	1.13	7.97
30	ACPL_7495	NS	hypothetical protein	1.11	10.48

Table 7: Genes, which transcripts were shown to be significantly more abundant, in case of $\Delta acbD$ 49 strain/wild type comparative microarray experiment.

No	Gene number	Gene name	Gene product	M value	A value
1	ACPL_4391	*cad*	short-chain dehydrogenase/reductase SDR;	-1,14	6,51
2	ACPL_0571	NS	Probable actinorhodin transporter;	-1,50	10,27
3	ACPL_2291	NS	alcohol dehydrogenase;	-1,53	9,16
4	ACPL_3408	NS	Uncharacterized protein;	-1,73	5,68
5	ACPL_3914	NS	Uncharacterized protein;	-1,98	9,15
6	ACPL_3406	NS	hypothetical protein;	-2,58	8,71
7	ACPL_3684	*acbD*	acarviose transferase (ATase)	-5,97	8,78

Table 8: Genes, which transcripts were shown to be significantly less abundant, in case of $\Delta acbD$ 49 strain/wild type comparative microarray experiment.

Interestingly, there are several genes, which are differentially expressed in this experiment and are clustered in the genome of *Actinoplanes* sp.SE50/110. Several such clusters can be distinguished: 1) sugar transport related genes; 2) genes related to flagellar formation and movement; 3) amino acid metabolism related genes (lysine); 4) nitrate/sulphate transport and metabolism genes. At the moment it is not clear, what are the reasons for such clustering and for higher abundance of transcripts of these particular genes.

An additional way to confirm the absence of any secondary effects, possibly caused by the knockout of *acbD* gene, is to analyze the phenotype of Δ*acbD* complementation strain. The construction of this strain is described in a next chapter.

3.5.2 Construction of *acbD* complementation and overexpression strains in *Actinoplanes* sp. SE50/110

In order to carry out complementation of this mutant strain, a derivative of integrative vector pSET152 was constructed via substitution of apramycin resistance gene *aac*(3)IV with hygromycin resistance gene with the help of Gibson assembly. Further on, the wild type copy of *acbD* gene, including the putative promoter region of *acbD* gene was cloned into pSET152hyg. The cloning was done similarly to pSET152hyg, with Gibson assembly and the correctness of both constructs was verified with in-house Sanger sequencing. Details regarding these constructs can be found in Materials and methods section.

Both constructs were transferred to *E. coli* ET12567 (pUZ8002); in both cases six independent transformant colonies were verified with colony PCR, utilizing primers to *oriT* region of pSET152 and primers to internal regions of *acbD* gene. After verification, both constructs were transferred to *Actinoplanes* sp. SE50/110 and to *Actinoplanes* sp. SE50/110 Δ*acbD* 49 strains via standard conjugation procedure as described before. In each of four types of matings, six colonies of *Actinoplanes* exconjugants were first selected for their resistance to hygromycin, or as in case of *Actinoplanes* sp. SE50/110 Δ*acbD* 49 derivatives, for their hygromycin/apramycin resistance. Further on, all selected strains were verified with PCR, using the same primers, as used for the verification of constructs earlier. All selected colonies were preserved for following studies as glycerol cultures.

Strains *Actinoplanes* pSET152hyg and 49 pSET152hyg were used in further studies as controls, in order to verify the possible impact of integrative vector. Strain 49 pSET152hygacbD was used to study the complementation of *acbD* mutation. Com-

plementation strains typically are constructed in order to verify whether the observed phenotype is caused solely by the substitution of the target gene. These studies give a possibility to exclude other effects on phenotype, which can be caused e.g. by polar effects of the gene substitution or any undesired mutations in the genome. Strain *Actinoplanes* pSET152hygwt was used as an overexpression of *acbD* gene in the wild type strain in order to study the effect of additional copy of *acbD* gene on acarbose production.

3.5.3 Characterization of *Actinoplanes* Δ*acbD*, its complementation and *acbD* overexpression strains growth and acarbose production in minimal media, supplemented with various Carbon sources

TO gain more information regarding the effects of the *acbD* mutation, *acbD* overexpression and Δ*acbD* complementation strains were studied in the same conditions as a mutant and a wild type strain. All of the constructed strains, were tested in minimal media, supplemented with glucose, maltose, maltotriose and Cpur in parallel with wild type and *Actinoplanes* Δ*acbD* 49. Growth curves related to all cultivations of this experiment can be found in Supplementary material (Figure 46). In all experiments HPLC measurements were performed every 24 hours until 144 hours of growth, except of Glu-MM, as long as no acarbose can be routinely detected in this media. LCMS measurements were performed with supernatants collected from all media at the end of cultivation in the stationary phase of growth. All measurements were performed in three biological replicates.

In case of cultivations in Glu-MM (Figure 33), it is evident from LCMS measurements, that *Actinoplanes* Δ*acbD* strain is not capable to synthesize acarviosyl-glucose, which is the major component produced by wild type in this media. However, in contrary to wild type, the mutant is producing low amounts of acarviosyl-maltose in these conditions. Noticeably, introduction of empty integrative vector pSET152 leads to lower production of acarviosyl-glucose in the wild type. This effect was not

established and detected before. Therefore, the expression of the wild type copy of the *acbD* gene in the *Actinoplanes* Δ*acbD* strain seems to restore the wild type levels of acarviosyl-glucose, while being compared to the wild type with pSET152 empty vector. In the same fashion, the additional copy of the *acbD* gene does not cause the change of acarviosyl-glucose production, while compared to the wild type with pSET152 empty vector.

In case of cultivations in Mal-MM, *Actinoplanes* Δ*acbD* 49 strain, as it was shown before, is not able to produce acarviosyl-maltose, which is a typical characteristic of the wild type. However, it was shown to produce low amounts of acarviosyl-maltotriose (Figure 34B). In this cultivations, no effect of the presence of integrative vector pSET152 on the wild type acarviosyl-maltose production was detected. Importantly, introduction of the wild type copy of the *acbD* gene into *Actinoplanes* Δ*acbD* 49 strain led to the restoration of the near to the wild type levels of acarviosyl-maltose production. Similarly to the previous cultivation, introduction of the single additional copy of the *acbD* gene to the wild type strain does not change neither the pattern of metabolites, which are being synthesized nor their relative amounts. Strain *Actinoplanes* pSET152hygacbD, similarly to the wild type, produces acarviosyl-maltose as its main metabolite.

HPLC measurements, conducted for this cultivation, give a possibility to estimate the amount of acarviosyl-maltose being synthesized in relation to the amount of biomass accumulated in a dynamic fashion (Figure 34A). The amounts of acarviosyl-glucose are typically increasing in the media throughout the cultivation, as it can be also observed in this particular experiment. However, low levels of acarviosyl-maltose, which are being accumulated by the *Actinoplanes* Δ*acbD* 49 strain, remain unchanged during all cultivation period. Complementation and overexpression strains behave in a similar fashion to the wild type.

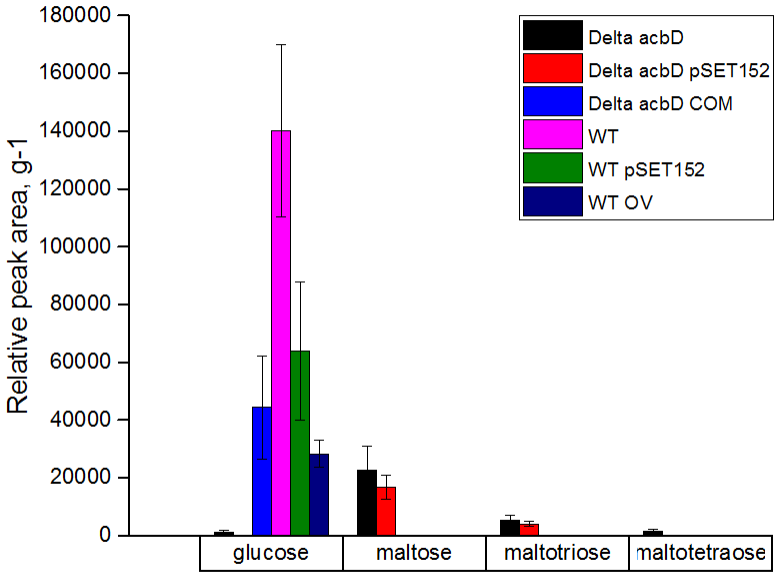

Fig. 33. Biosynthetic activity of *Actinoplanes* Δ*acbD* 49 strain, *Actinoplanes* Δ*acbD* 49 complementation strain, *acbD* overexpression strain, wild type and respective control strains in Glu-MM, measured with LCMS method; strains are marked as follows: Delta acbD - *Actinoplanes* Δ*acbD* 49 strain; Delta acbD pSET152 - *Actinoplanes* Δ*acbD* pSET152; Delta acbD COM - *Actinoplanes* Δ*acbD* pSET152acbD; WT - *Actinoplanes* sp. SE50/110; WT pSET152 - *Actinoplanes* pSET152; WT OV - *Actinoplanes* pSET152acbD. Mean values and standard deviations are shown for three biological replicates. Calculations were performed as in [132].

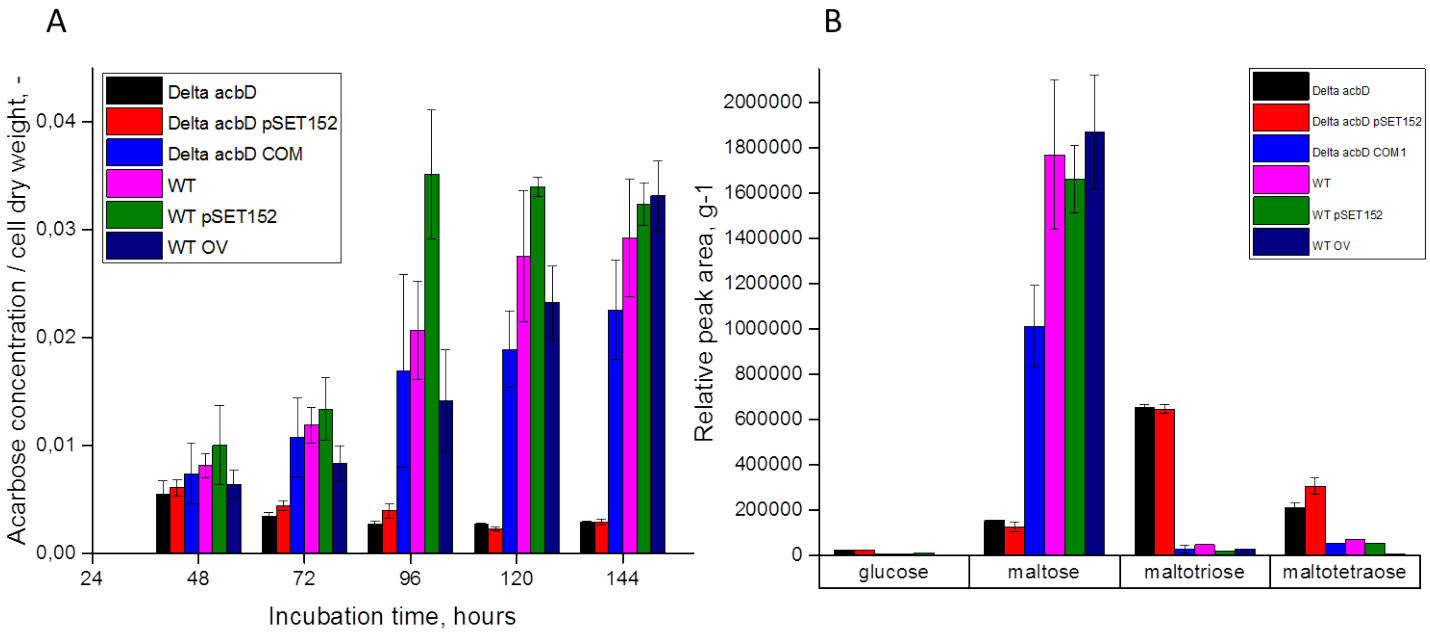

Fig. 34. Acarbose production of *Actinoplanes* Δ*acbD* 49 strain, *Actinoplanes* Δ*acbD* 49 complementation strain, *acbD* overexpression strain, wild type and respective control strains in Mal-MM, measured with HPLC (A) and LCMS (B) methods; strains are marked as follows: Delta acbD - *Actinoplanes* Δ*acbD* 49 strain; Delta acbD pSET152 - *Actinoplanes* Δ*acbD* pSET152; Delta acbD COM - *Actinoplanes* Δ*acbD* pSET152acbD; WT - *Actinoplanes* sp. SE50/110; WT pSET152 - *Actinoplanes* pSET152; WT OV - *Actinoplanes* pSET152acbD. The presented relative peak areas of acarviosyl metabolites were obtained by normalizing the peak areas of the respective extracted ion chromatograms to culture cell dry weights. Mean values and standard deviations are shown for three biological replicates. Samples for LCMS measurements were collected in stationary phase of growth. Calculations were performed as in [132].

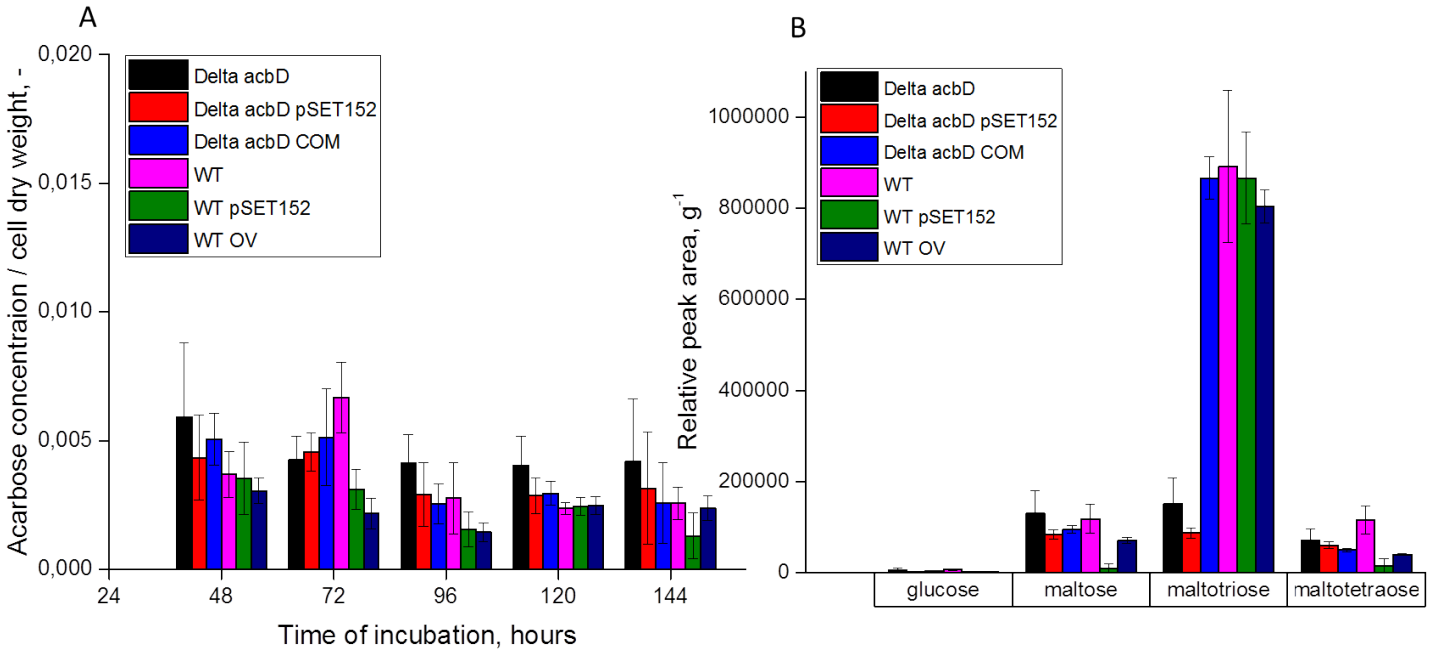

Fig. 35. Acarbose production of *Actinoplanes* Δ*acbD* 49 strain, *Actinoplanes* Δ*acbD* 49 complementation strain, *acbD* overexpression strain, wild type and respective control strains in MT-MM, measured with HPLC (A) and LCMS (B) methods; strains are marked as follows: Delta acbD - *Actinoplanes* Δ*acbD* 49 strain; Delta acbD pSET152 - *Actinoplanes* Δ*acbD* pSET152; Delta acbD COM - *Actinoplanes* Δ*acbD* pSET152acbD; WT - *Actinoplanes* sp. SE50/110; WT pSET152 - *Actinoplanes* pSET152; WT OV - *Actinoplanes* pSET152acbD. The presented relative peak areas of acarviosyl metabolites were obtained by normalizing the peak areas of the respective extracted ion chromatograms to culture cell dry weights. Mean values and standard deviations are shown for three biological replicates. Samples for LCMS measurements were collected in stationary phase of growth. Calculations were performed as in [132].

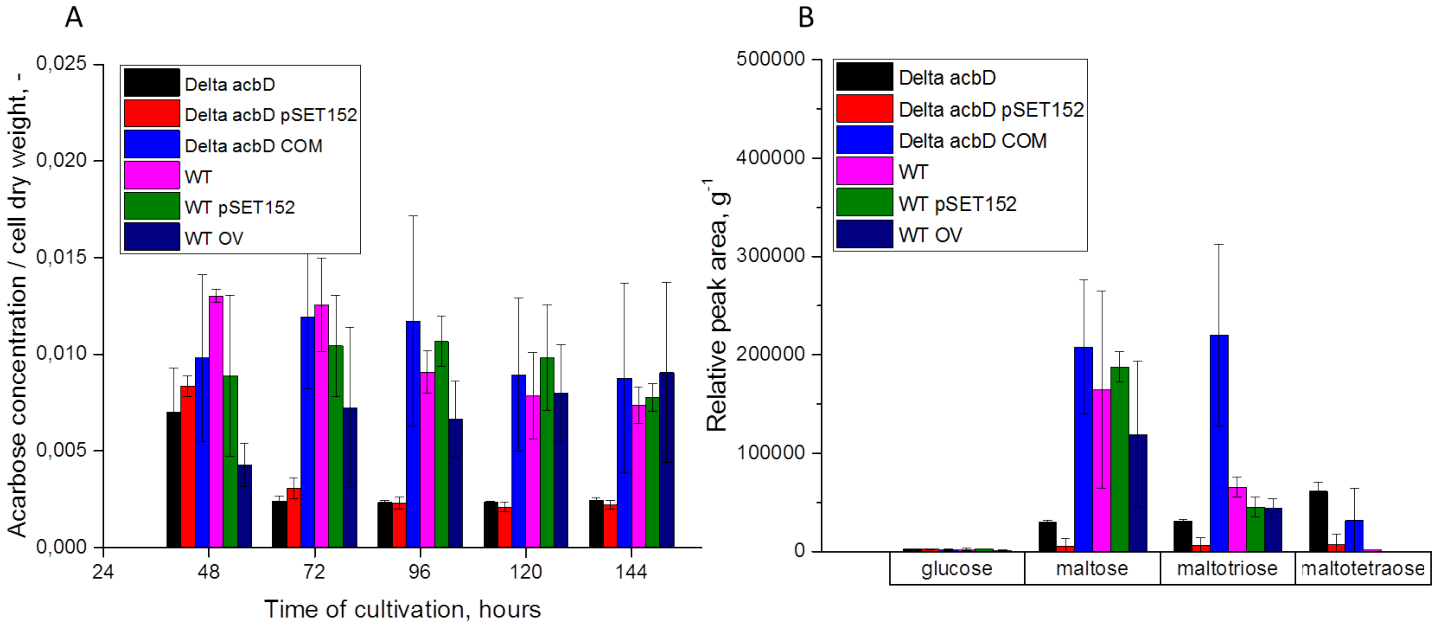

Fig. 36. Acarbose production of *Actinoplanes* Δ*acbD* 49 strain, *Actinoplanes* Δ*acbD* 49 complementation strain, *acbD* overexpression strain, wild type and respective control strains in Cpur-MM (maltose/maltotriose mixture), measured with HPLC (A) and LCMS (B) methods; strains are marked as follows: Delta acbD - *Actinoplanes* Δ*acbD* 49 strain; Delta acbD pSET152 - *Actinoplanes* Δ*acbD* pSET152; Delta acbD COM - *Actinoplanes* Δ*acbD* pSET152acbD; WT - *Actinoplanes* sp. SE50/110; WT pSET152 - *Actinoplanes* pSET152; WT OV - *Actinoplanes* pSET152acbD. Relative peak areas of acarviosyl metabolites were obtained by normalizing the peak areas of the extracted ion chromatograms to cell dry weights. Mean values and standard deviations are shown for three biological replicates. Samples for LCMS measurements were collected in stationary phase of growth. Calculations were performed as in [132].

In case of MT-MM cultivations, the major component detected for the wild type in LCMS measurements was acarviosyl-maltotriose, as its been expected from previous studies [132]. Other components, acarviosyl-maltose and acarviosyl-maltotetraose, were also present in supernatant, but in lower quantities. In this cultivations, no effect of the presence of integrative vector pSET152 on the wild type production of acarviosyl- metabolites was detected. The biosynthetic pattern of *Actinoplanes* Δ*acbD* 49 strain was however quite distinct to that of the wild type (Figure 35B). Acrviosyl-maltose, -maltotriose and -maltotetraose all were detected in the supernatant of *Actinoplanes* Δ*acbD* 49, but in a distinctively low amounts. As it is seen from HPLC measurements, amount of acarviosyl-maltose, synthesized by *Actinoplanes* Δ*acbD* 49 strain is almost equal to that of the wild type and remains unchanged throughout the whole cultivation (Figure 35A).

As it was shown in previous cultivations, introduction of the wild type copy of the *acbD* gene into *Actinoplanes* Δ*acbD* 49 strain led to the restoration of the wild type phenotype in the mutant strain. Namely, of the biosynthetic pattern - the main component synthesized is acarviosyl-maltotriose and also the amounts of acarviosyl-maltose, as it is detected in HPLC measurements (Figure 35A and B). Introduction of the single additional copy of *acbD* gene, however, does not change neither pattern nor amounts of the metabolites, synthesized by the wild type strain.

In case of Cpur-MM cultivations, acarviosyl-maltose and acarviosyl-maltotriose were detected in the supernatant of the wild type strain, with acarviosyl-maltose clearly prevailing over acarviosyl-maltotriose. The introduction of pSET152 into the wild type did not lead to the change of the pattern or amount of metabolites, synthesized by the wild type strain (Figure 36A and B). *Actinoplanes* Δ*acbD* 49 strain did not show a clear major component in these cultivations, while acarviosyl-maltose, -maltotriose and -maltotetraose all detected in low amounts. The amount of acarviosyl-maltose, synthesized by this strain was low as compared to the wild type and other strains and did not change during the period of the cultivation (Figure 36A).

The complementation strain *Actinoplanes* Δ*acbD* pSET152acbD was able to accumulate both acarviosyl-maltose and -maltotriose with the first compound being accumulated in a similar fashion to the wild type. However, the amount of acarviosyl-maltotriose, being accumulated by the complementation strain, is higher in comparison to the wild type. *Actinoplanes* pSET152acbD strain, which contain a single additional copy of the *acbD* gene, was able to accumulate similar pattern and amounts of metabolites as the wild type strain (Figure 36 A and B).

The summarized results of all of the cultivations, conducted to characterize *Actinoplanes* Δ*acbD* 49 strain in comparison to the complementation and overexpression strains, can be seen in Table 9.

3.5.4 Influence of acquired information regarding the phenotype of *Actinoplanes* Δ*acbD* 49 strain on existing models of acarbose production

This sub-chapter is designed to summarize and discuss the results of the study of the *Actinoplanes* Δ*acbD* 49 strain and influence of this information on the previously suggested models of acarbose metabolism. The overview of acarbose metabolism models can be found in Introduction section of this work.

In the studies, conducted in this work *Actinoplanes* Δ*acbD* 49 strain was constructed via ReDirect procedure, which included the substitution of an ORF of the *acbD* gene with apramycin resistance antibiotic cassette. The correct construction of the mutant was verified independently with several methods such as PCR and whole genome sequencing. Later confirmed absence of any mutations or other genetic rearrangements in the genome of the mutant. Knockout of the *acbD* gene was also confirmed by studies of extracellular proteome of *Actinoplanes* Δ*acbD* 49 strain, which showed clear absence of AcbD protein. Further on, qRT-PCR measurements of the samples isolated from maltose minimal media, have shown absence of the *acbD* gene transcript.

The phenotype of *Actinoplanes* Δ*acbD* 49 strain was studied with the help of BioLog microarrays and cultivation studies in minimal media, supplied with various carbon sources. It was revealed, that *Actinoplanes* Δ*acbD* 49 strain is producing significantly lowered amounts of acarbose in comparison to the wild type. However, it was shown to retain the ability to synthesize all spectra of acarviosyl-metabolites, which can be measured at the moment, in a similar lowered manner to acarbose.

By conducting whole genome comparative microarrays between *Actinoplanes* Δ*acbD* 49 and wild type samples, both cultivated in Mal-MM, it was shown, that the expression of *acb* gene cluster is not affected by the knockout of the *acbD* gene. Therefore, lowered amount of the acarviosyl- metabolites, synthesized by a mutant strain, is not caused by alteration in the expression of *acb* genes or any other genes related to the acarbose biosynthetic pathway.

To verify whether the displayed particular phenotype of *Actinoplanes* Δ*acbD* 49 strain is not caused by any secondary effects, Δ*acbD* complementation strain was constructed. This strain included the wild type copy of the *acbD* gene, carried on the single copy integrative vector pSET152. The same genetic construct was introduced into the wild type strain in order to verify the effect of the additional copy of *acbD* gene on acarviosyl-metabolite production. All strains, including control strains retaining only empty vectors, were cultivated in minimal media, supplemented with glucose, maltose, maltotriose and the mixture of maltose/maltotriose. In order to make the results of this experiment easier to perceive, the simplified version of the results of this experiment is depicted in the Table 9. This Table includes only the main components, synthesized by each strain in each particular media.

Several important conclusions can be drawn from this experiment. Firstly, the wild type copy of *acbD* gene, introduced as a part of pSET512, complements the phenotype of the *Actinoplanes* Δ*acbD* 49 strain. It is visible from all of the cultivations conducted, while biosynthesis of each particular main acarviosyl- compound, specific to each medium tested, is restored to the levels, close to those of the wild type (fig. 33, Figure 34, Figure 35, Figure 36, Table 9). These results confirm, that the observed

phenotype of *Actinoplanes* Δ*acbD* 49 strain is caused only by knockout of the *acbD* gene itself and not any other secondary effects. Secondly, *Actinoplanes* Δ*acbD* 49 strain displays a distinctive pattern of metabolite biosynthesis in comparison to the wild type. For example, acarviosyl-glucose is the main compound, typically detected for the wild type in Glu-MM. But for *Actinoplanes* Δ*acbD* 49 strain only low amounts of acarviosyl-maltose can be detected in this media. This holds true also for other tested media Table 9. The other related interesting conclusion is regarding the low amounts of all metabolites, and in particular acarviosyl-maltose, being synthesized by *Actinoplanes* Δ*acbD* 49 strain, which are not dependent on the type of media tested. Thirdly, the expression of the additional copy of *acbD* gene does not cause significant changes in acarviosyl-metabolite production.

In order to make any conclusions or suggestions regarding the function of AcbD protein in acarbose metabolism, it is helpful to reflect onto which expectations were laid out previously in formulated models, namely [142, 128, 131, 132, 97], and verify, whether the results, obtained in this study, can confirm them.

The first model of acarbose biosynthesis [142, 128] was giving AcbD protein an important role of extracellular glucosyltransferase, responsible for the high abundance of various acarviosyl-metabolites, containing different amount of glucose units. This suggestion was mostly based on heterologous expression and enzymatic studies of AcbD [51], which have shown, that this protein does not possess α-amylase activities. This model did not speculate regarding intra- or extracellular production of acarviosyl- metabolites, due to the lack of data.

In a view of results of this study, it is possible to assume, that AcbD function does not restrict only to creating the multiple variants of acarviosyl- metabolites, while *Actinoplanes* Δ*acbD* 49 strain is still capable of producing at least four different acarviosyl- homologues. The question of lowered amount of all metabolites, produced by *Actinoplanes* Δ*acbD* 49 strain can also be not answered only by this model alone.

Strains	Media			
	Glu-MM	Mal-MM	MT-MM	Mal/MT-MM
WT	-glucose	-maltose	-maltotriose	-maltose
Δ*acbD*	-maltose	-maltotriose	-maltose/ -maltotriose	-maltose/ -maltotriose/ -maltotetraose
Δ*acbD* COM	-glucose	-maltose	-maltotriose	-maltose/ -maltotriose
WT OV	-glucose	-maltose	-maltotriose	-maltose

Table 9: The major acarviosyl-compounds, detected for the wild type strain, *Actinoplanes* Δ*acbD* strain and their respective derivatives, cultivated in liquid minimal media, supplemented with various Carbon sources. Strains are marked as follows: Δ*acbD* - *Actinoplanes* Δ*acbD* 49 strain; Δ*acbD* COM - *Actinoplanes* Δ*acbD* pSET152acbD; WT - *Actinoplanes* sp. SE50/110; WT OV - *Actinoplanes* pSET152acbD. Data regarding *Actinoplanes* pSET152 and *Actinoplanes* Δ*acbD* 49 strain is not shown.

Next model, which builds on the previous one, was suggested in [131]. It largely retains all of the components of the previous model and adds several suggestions. In a meanwhile, it became known from crystallization studies, that supposed acarbose importer AcbHFG does not belong to the acarbose cluster [76]. From studies of extra- and intracellular proteome, conducted in [131], new candidate carbohydrate transporters were suggested. Interestingly, AcbD was found to be present both intra- and extracellularly, but with a higher abundance in a later case. This was taken as a support of extracellular function of AcbD, namely acarviosyltransferase function.

AcbD was once again claimed to be responsible for the multiple forms of acarviosyl-metabolites.

Next model of acarbose biosynthesis [132] is further building on ideas of previous ones and includes scheme of the biosynthesis of minor acarviosyl compounds. The difference in the formation of major and minor acarviosyl compounds is seen in the origin of the variable saccharide unit. In the case of major metabolites, it is provided with the medium, imported inside the cell and used in biosynthesis. While, in case of minor compounds, the source of saccharide unit originates intracellularly. The explanation of the formation of minor acarviosyl compounds, suggested in this model, holds in case of analysis of *Actinoplanes* $\Delta acbD$ 49 strain phenotype. Knockout of *acbD* gene does not stop the formation of minor metabolites. However, sharp decrease in the amounts of major metabolite in case of *Actinoplanes* $\Delta acbD$ 49 is beyond the explanation of this model.

The work conducted by Dr. Vera Ortseifen, further suggests regarding putative modeling of acarbose metabolism and AcbD function in it, were made [97]. These suggestions are of high interest, but producer strains which were used in this study are not available to general public. Therefore, it is not possible to perform genetic experiments to verify this model. However, it is possible to comment on the predictions, which were made by Dr. Vera Ortseifen for the wild type strain according to the formulated model. The model, built in this work, predicted that in case of knockout of *acbD* gene a pattern of synthesized acarviosyl metabolites would shift. To be more precise, acarviosyl-maltose being the major metabolite in case of cultivations in MT-MM and acarviosyl-maltotriose being the major metabolite in case of Mal-MM cultivations of *acbD* knockout strain. The later effect can be observed in the results, which were obtained in this work. However, no clear effect is observed in MT-MM, while similar amounts of acarviosyl-maltose and acarviosyl-maltotriose are being synthesized by the knockout strain. In addition, postulated models do not explain or predict a sudden drop in the amounts of acarbose, synthesized by the *Actinoplanes* $\Delta acbD$ strain.

In order to conclude and to summarize all data together. From enzymatic characterization of AcbD, done in [51], several suggestions regarding putative functions of this enzyme were initially made. AcbD was proposed to act as an enzyme responsible for the final step of the acarbose biosynthesis intracellularly or the last step, connected with the export of the compound. On the other hand, extracellularly, it was supposed to act as an acarviosyltransferase. These suggestions see further support in a light of results of this study. The sharp decrease of amount of major acarviosyl metabolite in each respective media signs, that the major metabolites are not formed or formed in limited amounts in absence of AcbD protein. As it was suggested in [132], minor acarviosyl metabolites are formed intracellularly, by utilising saccharides provided from carbohydrate metabolism, probably by the action of one of the Acb proteins, e.g. AcbI or AcbJ. It is possible to assume, that the minor metabolites are formed intracellularly by the action of one of the candidate proteins, either AcbI, AcbJ, AcbQ or AcbS. At the same time, major component is formed extracellularly or intracellularly by the action of AcbD. As long as it was proven, that AcbD can be detected in both conditions and in general most of the Acb proteins are present in the close proximity to the membrane, both conditions are possible.

Another suggestion to explain results of the study of the *Actinoplanes* $\Delta acbD$ strain, could be connected to the disruption of the acarviosyl metabolites transfer to the extracellular space in absence of AcbD protein or secondary effects on the proteome of *Actinoplanes*. However, as long as no effective techniques to verify intracellular acarviosyl-compounds concentration was developed so far, this matter can not be verified.

However, these suggestions remain to be mainly speculations, partially supported with the examination of *Actinoplanes* $\Delta acbD$ strain phenotype, while yet too lees information is available on the actual "machinery" of acarbose metabolism in the wild type strain. Further knockouts of genes from *acb* gene cluster, study of their metabolome, proteome and transcriptome will shed more light on this issue. There-

fore, this application example shows well the actual significance of the development of the native gene cloning system for *Actinoplanes* sp. SE50/110.

3.6 Application of the developed techniques for the knockout of the ACPL_212 (*cadC*) gene, coding for the putative regulator of acarbose biosynthesis

Gene *cadC* was chosen as a target for gene knockout, based on previous experiments, conducted in *Actinoplanes* working group. Anna Köpfer in her Master thesis project [73] has identified through regulator fishing experiments a number of proteins, which can bind to *acb* gene cluster, and are putatively involved in regulation of acarbose biosynthesis. With the help of gel-free regulator fishing, Anna predicted, that ACPL_212 protein can bind upstream of *acbA*, *acbB*, *acbV* and *acbW* genes, also mentioning the possibility of ACPL_212 being a global regulator, controlling not only acarbose biosynthesis, but also other cell processes Therefore, cadC gene became one of our first targets for gene disruption experiments. Gene ACPL_212 has at least three putative paralogs in the genome of *Actinoplanes*: ACPL_1828 (115 bits/57% identity), ACPL_5258 (89.7 bits/47% identity), ACPL_4644 (87 bits/51% identity), which can be identified through protein BLAST search. Interestingly, ACPL_4276 gene, also annotated as *cadC* was also shown to bind to the *acb* cluster in the same work [73].

Gene name	Putative product	Closest homologous protein/organism	Max score / identity, %
ACPL_210	Hypothetical protein	Hypothetical protein / *A. awajinensis*	266 / 51%
ACPL_211	Hypothetical protein	Hypothetical protein / *A. awajinensis*	204 / 70%
ACPL_8395	Hypothetical protein	Hypothetical protein / *A. awajinensis*	88.2 / 81%

Gene name	Putative product	Closest homologous protein/organism	Max score / identity, %
ACPL_212 (*cadC*)	Transcriptional regulator of ArsR family	Transcriptional regulator / *A. friuliensis*	246 / 99%
ACPL_213 (f*ucA*)	Class II aldolase	Class II aldolase / *A. missouriensis*	416 / 90%

Table 10: Genes located upstream and downstream of ACPL_212, and their putative products. (BLAST search was performed against non-redundant protein sequences database with default parameters).

As a first step, it is important to explore the genetic "neighborhood" of a gene of interest (Table 10). ACPL_212 is located in a close connection to four other genes, predictably being transcribed in the same direction. Genes ACPL_210, ACPL_8395 and ACPL_211 are annotated as hypothetical proteins. Further *in silico* analysis of these genes did not reveal any additional information regarding the putative function of their products. A short gene ACPL_8395 was predicted and annotated by Patrick Schwientek ([109]) as a result of analysis of RNA sequencing data. Interestingly, only one homologous gene was identified in case of ACPL_8395 in all publicly available databases. Gene ACPL_213 or *fucA* is presumably coding for class II aldolase enzyme. This assumption is in good agreement with BLAST search results and a fact, that this gene seems to be conserved across *Actinoplanes* genus.

3.6.1 *in silico* analysis of the putative structure of CadC protein

Knowledge of a gene and protein sequences gives a possibility to predict a possible function of a protein. Judging from various protein structure prediction software tools and search in databases (Pfam [31], Interpro[62], NCBI CDD [84]), we expect CadC

to be a transcriptional regulator of ArsR family. It contains winged-helix type HTH motif and possibly can bind to DNA molecules. Also, it harbors putative dimerization site with four ion binding sites (Zn2+ or other ions) per dimer. The ArsR-type HTH domain is known to be involved in stress-response to heavy metal ions. This family of prokaryotic metal-sensing transcription repressors is named after *E. coli arsR*, an arsenic-responsive repressor of the *ars* operon for arsenate reductase and metal ion extrusion, and after *Synechococcus* PCC 7942 *smtB*, a Zn(II)-responsive repressor of the *smtA* gene for a Zn sequestering metallothionein. ArsR/SmtB-like repressors of metal resistance operons specifically bind to the operator/promoter and seem to dissociate from the DNA in the presence of metal ions, permitting transcription of proteins involved in metal-ions efflux and/or detoxification [17]. However, in several cases members of ArsR family proven to act not only as repressors, as it is described for cadmium resistance operon regulator CadC in *Staphylococcus aureus* ATCC12600 [56].

Phyre2 software predicts the model of the secondary structure of CadC to be depicted as in Figure 37, containing two β strands and 5 α helixes. Multiple alignment for this model can be found in Figure 37B. In this case, alignment is built on the base of Hidden Markov Models (HMM), which are built based on the homologues, detected with primary PSI BLAST search. Than HMM model is scanned against the library of HMMs of proteins, whose structures were solved experimentally.

Interestingly, other hits, non-related to metal homeostasis were also found through Pfam search. One of them, was to the sugar-specific transcriptional regulator Trmb. However having high sequence coverage, it has a high E value of 0.0083. It can be explained by the fact, that members of this family of proteins are mostly characterized in evolutionary distantly related to *Actinoplanes* organisms, e.g. hyperthermophilic archaea *Pyrococcus furiosus*. Members of the TrmB family act as global transcriptional regulators for the activation or repression of sugar ABC transporters and central sugar metabolic pathways. It includes glycolytic, gluconeogenic, and other metabolic

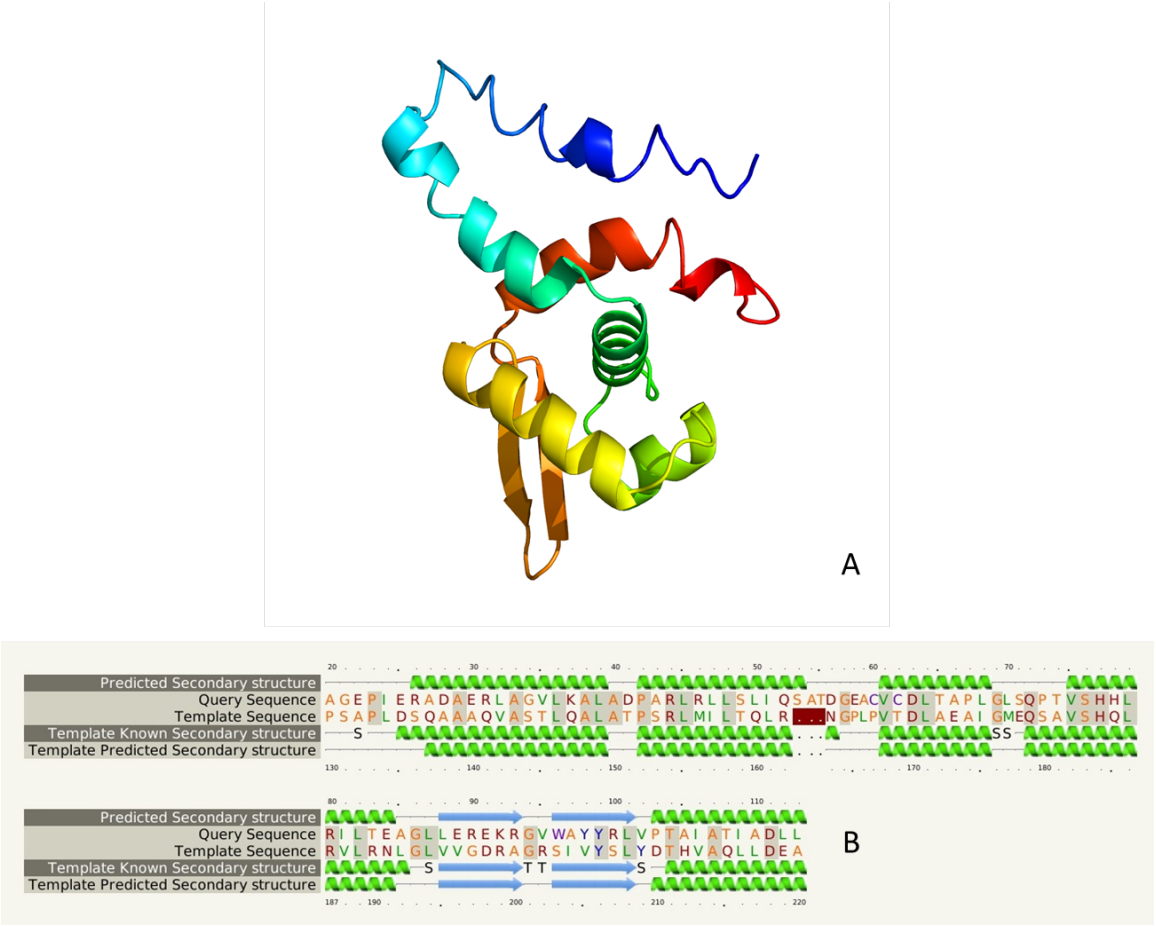

Fig. 37. Prediction of secondary structure of CadC protein: model of putative CadC structure and the alignment, which was used to build it (B). Image of the model is coloured by rainbow, respectively from N to C terminus of the protein. Altogether, 83% of residues were modeled at >90% confidence. Model dimensions (Angstrem): X:44.830 Y:36.064 Z:48.605. Model is determined with 99.9% confidence. Model and alignment are built by software, described in [67].

pathways. [69]. As this family is newly described, there is only limited experimental data regarding non-archeal family members.

3.6.2 Construction of *Actinoplanes* Δ*cadC* strain, its complementation and respective overexpression strains

Work regarding characterization of phenotype of *Actinoplanes* Δ*cadC*, respective complementation strains and *cadC* overexpression strains, was done with help of Claudia Tran, in course of her Master Thesis project [123].

The construction of *Actinoplanes* Δ*cadC* strain is described in details in previous section. In order to ensure the correct knockout of *cadC* gene, in-house whole genome sequencing was performed. As a result, absence of any genetic rearrangements, except of ReDirect origin, was confirmed. For details regarding genome sequencing of *Actinoplanes* Δ*cadC* please refer to [123].

In order to begin work on construction of the complementation or overexpression strains, it is important to find the correct position of the respective promoter region. RNA seq data is proven to be helpful in such experiments. As it is evident from RNA seq data, previously acquired in *Actinoplanes* working group and kindly provided for this study by Patrick Schwientek, multiple transcription start sites are possible (Figure 38A). As long as transcriptional activity is generally high in this region, in comparison to the overall genome level, the single transcription start site for *cadC* can not be assigned only on the base of the existent data (Figure 38B).

As long, as it was difficult to predict the exact promoter region of *cadC* gene, two constructs containing varying in length genetic regions upstream of *cadC*, gene were created. The first of them contains only short region, immediately upstream of ACPL_212 (further called cadC A), and the second one including full ORF of ACPL_8395 (cadC B). Later on, those two constructs were cloned into pSET152 and pSET152hyg. Following strains were made using these newly constructed plasmids: 1) WT pSET152cadC A (further called OV1). It contains *cadC* gene under predicted "native promoter" in a single copy integrative pSET152 plasmid; 2) WT pSET152cadC B (OV2). This strain contains *cadC* gene under a "long native promoter" region, including ORF of ACPL_8395 in a single copy integrative pSET152

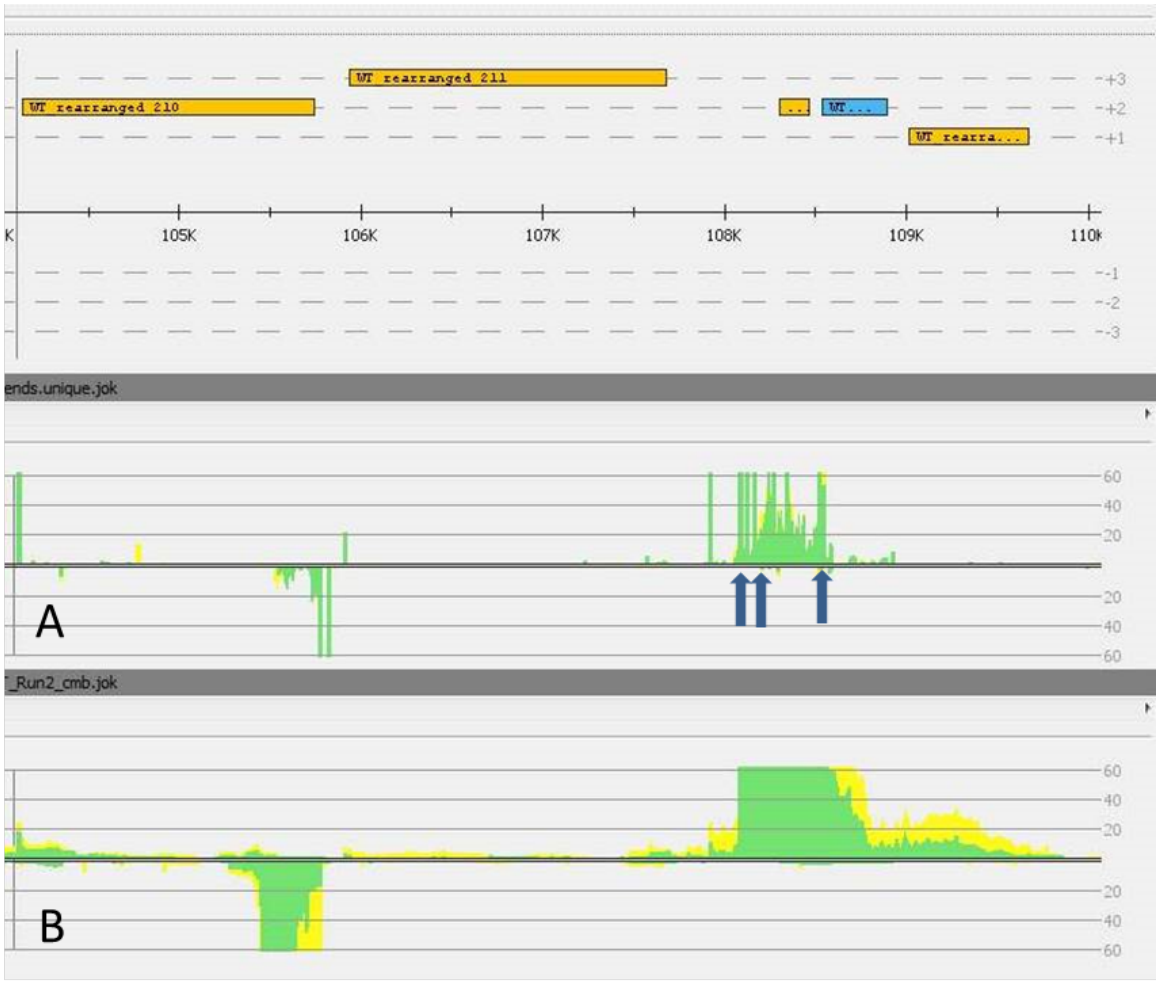

Fig. 38. An overview of the transcriptional activity in the ACPL_210 – ACPL_213 genetic region. Visualized by ReadXplorer 2.1.0; RNA sequencing data is kindly provided by Patrick Schwientek. A – sequencing of enriched 5prime ends of primary transcripts; B – whole transcriptome RNA sequencing data. Gene *cadC* is highlighted in blue colour, multiple possible transcription start sites are indicated with blue arrows.

plasmid; 3) deltacadC pSET152cadCA (COM1). This strain is constructed in a way, similar to OV1, containing *cadC* under "native promoter" in an integrative plasmid pSET152; 4) deltacadC pSET152cadCB (COM2) is constructed in a way, similar to OV2, containing *cadC* under "long native promoter" in an integrative plasmid

pSET152. Genomes of both of the complementation strains, COM1 and COM2 have been sequenced in order to verify the correct construction of these strains [123].

In order to have a control strains for comparison, pSET152hyg was transfered through conjugation to both wild type and *Actinoplanes* Δ*cadC* strain. Exconjugant strains were verified with colony PCR, using specific primers to bind to pSET152, described in this work before.

No.	Name (as in text)	Genetic construct	Full strain name
1.	Δ*cadC* 31(41)	Deletion of *cadC* gene in *Actinoplanes* sp. SE50/110 strain	*Actinoplanes* Δ*cadC* 31(41)
2.	WT pSET152*cadC* A (OV1)	Overexpression of *cadC* gene with a putative native promoter in a single copy plasmid	*Actinoplanes* pSET152*cadC*A
3.	WT pSET152*cadC* B (OV2)	Overexpression of a region, containing ACPL_8395 and *cadC* in a single copy plasmid	*Actinoplanes* pSET152*cadC*B
4.	WT pSET152	Wild type strain with empty vector pSET152	*Actinoplanes* pSET152
5.	WT pIJ6902*cadC* (OV3)	Overexpression of *cadC* in a single-copy plasmid under a strong promotor (*tipA*)	*Actinoplanes* pIJ6902*cadC*
6.	WT pKC1139E*cadC* (OV4)	Overexpression of *cadC* gene in a 30 copy plasmid under a strong promotor *ermE*	*Actinoplanes* pKC1139EE*cadC*
7.	WT pKC1139EE	Wild type strain with empty vector pKC1139EE	*Actinoplanes* pKC1139EE

No.	Name (as in text)	Genetic construct	Full strain name
8.	Δ*cadC* pSET152*cadC* A (COM 1)	Complementation of Δ*cadC* strain with the construct *cadC A* in a single copy-plasmid	*Actinoplanes* Δ*cadC* pSET152*cadC*A
9.	Δ*cadC* pSET152*cadC* B (COM2)	Complementation of Δ*cadC* strain with the ACPL_8395 and *cadC* in a single copy plasmid	*Actinoplanes* Δ*cadC* pSET152*cadC*B
10.	Δ*cadC* 31(41) pSET152	Δ*cadC* strain with empty vector pSET152	*Actinoplanes* Δ*cadC* 31(41) pSET152

Table 11: List of all strains, constructed in a course of *Actinoplanes* Δ*cadC* mutant study. All created constructs were sequenced to verify the absence of possible point mutations.

Two additional overexpression strains were constructed in the course of this study: WT pIJ6902cadC (OV3) and WT pKC1139EcadC (OV4). The first strain contains *cadC* under the strong promoter *tipA* in a single copy pSET152-based vector. Promoter *tipA* was earlier verified and proven to function in *Actinoplanes* sp. SE50/110. OV4 strain contains *cadC* gene under strong, constitutive *ermE* promoter in a multicopy pKC1139 replicative plasmid. This strain was constructed with the intention to study the influence of the multicopy presence of *cadC* in the cell. Details regarding primers used for cloning and construction can be found in Materials and methods section; also, please refer to Table 11 and [123].

3.6.3 Characterization of *Actinoplanes* Δ*cadC* strain phenotype

To study the phenotype of *Actinoplanes* Δ*cadC* strain, a similar systematic approach as in case of *Actinoplanes* Δ*acbD* strain was applied. The most attention

was paid to the influence of *cadC* knockout on morphology and acarviosyl-metabolite biosynthesis in minimal liquid media. To better understand the effects of the knockout on the expression of other genes, RNA-sequencing of samples, isolated from maltose minimal media was applied.

3.6.3.1 Morphology of *Actinoplanes* Δ*cadC* strain in comparison to the wild type

Colonies of *Actinoplanes* Δ*cadC*, obtained after conjugation, were highly distinct in their appearance in comparison to the wild type. Therefore, more attention was paid to study its morphology. In a course of liquid media cultivations, it became evident, that knockout mutant is not able to synthesize dark brown pigment, previously described as putative eumelanin [137]. Especially stark difference is observed in complex liquid media (Figure 39D). Also this effect can be seen while growing on agar plates, e.g. SFM or LB media (Figure 39A and B). Interestingly, even in media, which does not support dark pigment formation in wild type, as Bennett complex medium or maltose minimal medium, the colour of mutant's mycelia is comparatively lighter (Figure 39C). As per its phenotype, *Actinoplanes* Δ*cadC* seems to be very similar to previously described *Actinoplanes* Δ*melC* strain [137]. This resemblance is even more evident, when two strains grow together with wild type on the same LB plate (Figure 40C)

As another aspect of its specific morphology, *Actinoplanes* Δ*cadC* seems to be unable to sporulate on SFM, as its surface is not changing colour during the growth period. Also the lawn surface does not possess hydrophobic qualities of wild type, while tested with drops of water. In order to study this matter with more detail, scanning electron microphotographs of all strains, grown in parallel on SFM, were made Figure 41.

In order to verify, whether the specific morphology phenotype of *Actinoplanes* Δ*cadC* knockout can be complemented and whether the overexpression of *cadC* will

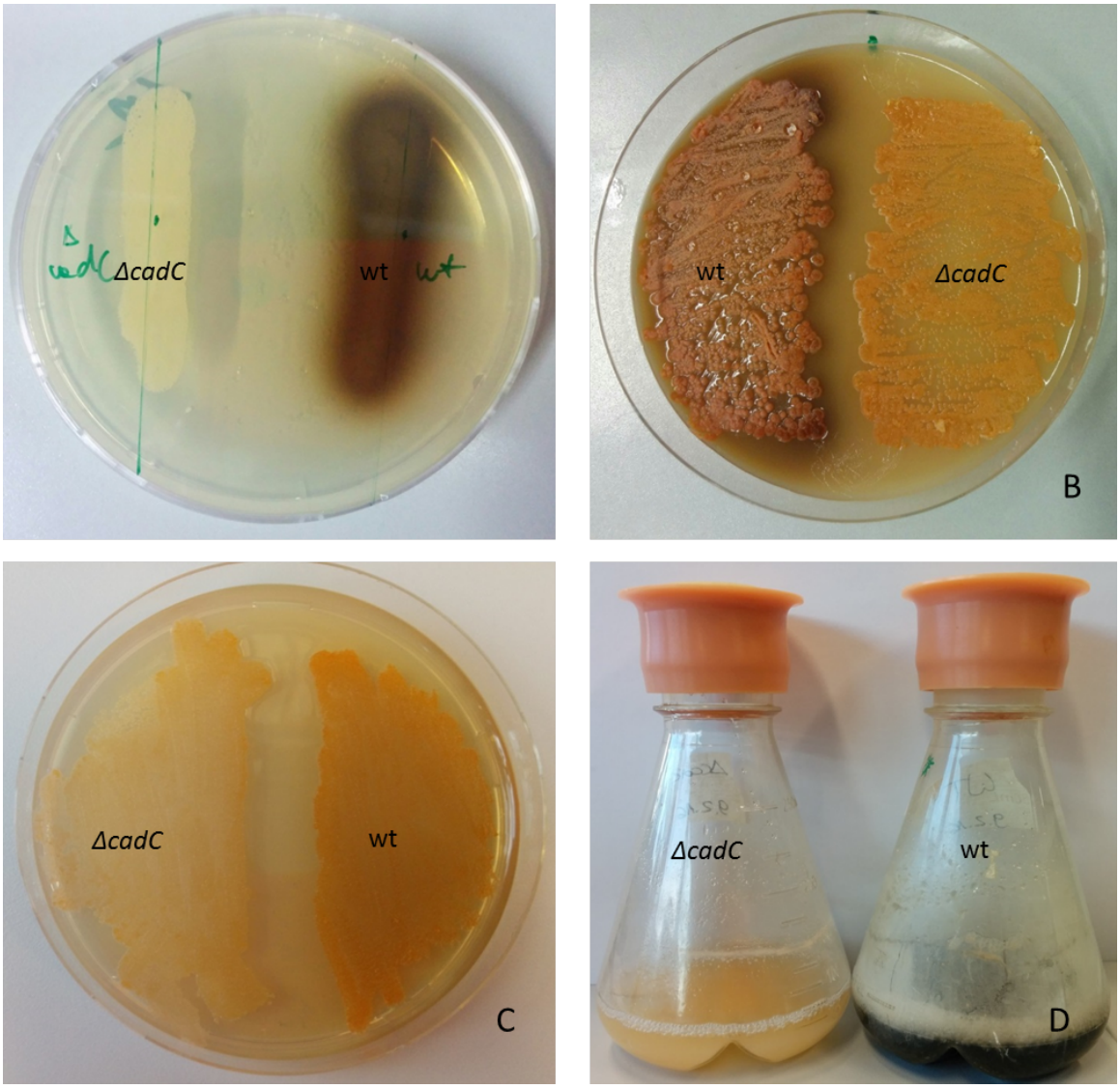

Fig. 39. Growth of *Actinoplanes* Δ*cadC* and wild type strains on LB (A), SFM (B) and Bennett (C) solid media, six days of growth; NBS liquid media (D), two days of growth.

cause any visible changes, all constructed strains were cultivated together in complex liquid media and on SFM plates (Figure 41A and B). As it is evident from Figure 41A, introduction of construct cadCB to *Actinoplanes* Δ*cadC* knockout strain leads to the restoration of pigment production, while it is not the case for cadCA construct. This

also holds true for the growth on SFM plates. As it is seen from Figure 41B, COM1 strain is similar in its phenotype to the knockout strain, while COM2 resembles wild type strain. Introduction of the empty plasmid vector does not seem to influence the phenotype of either wild type strain or mutant. All constructed overexpression strains, OV1, OV2, OV3 and OV4 are similar in their phenotype to the wild type strain.

To better understand the causes of the changes in the observable phenotype of *Actinoplanes* Δ*cadC*, scanning electron microphotographs of its colonies, grown on SFM, were made as it was previously described for the wild type strain in this work Figure 41. Typically for the wild type strain, globular or oval sporangia are formed on thin sporangiophores, stemming directly from the substrate mycelia Figure 41A. However, *Actinoplanes* Δ*cadC* strain seems to lack the sporangia, with only few of them occasionally dispersed along the lawn of sporangiophores, that could have been observed during this experiment Figure 41B. Similar phenotype is also observed for *Actinoplanes* Δ*cadC* pSET152 strain Figure 41C. COM1 and COM2 strains, both carrying wild type copies of the *cadC* gene, but with promoter regions of different length, seem to resemble knockout mutant in phenotype (Figure 41D and E) It is surprising, however, as COM2 was expected to have a similar architecture of morphological structures as wild type, because of visual signs observed on SFM plates. It is possible to assume, that sporangia were not detected in case of COM2, because of delay in time of its sporulation in regard to the wild type. All samples of colonies for the preparation of scanning electron images were collected on the sixth day of *Actinoplanes* growth. COM2 strain seems to require longer incubation time to begin to sporulate, as its observed on SFM plates.

An additional characterization of morphological phenotype of all strains was done with the help of fluorescent stereomicroscopy in Master thesis of Claudia Tran. Results of this experiment can support suggestion described above [123]. It was shown, that after longer incubation period, namely eight days, surface of COM2 colonies does

resemble that of the wild type, but not the knockout strain. Incubation of the COM1 strain for eight days or longer, does not change its phenotype.

Interestingly, OV1 strain, which contains one additional copy of *cadC* gene, resemble the wild type in its morphology Figure 41F. The same holds true for other overexpression strains, OV2, OV3 and OV4 (data not shown).

Results of the phenotypic analysis of all strains point to the direct or indirect influence of CadC on pigment production and morphological differentiation of *Actinoplanes*. However, it is quite typical for pleiotropic regulators of Actinobacteria to be involved in regulation of several processes. Complementation of the knockout mutant by the wild type copy of *cadC* gene under the long promoter version, partly restores the wild type phenotype in knockout strain. This can be taken as a proof of CadC protein being involved in these processes.

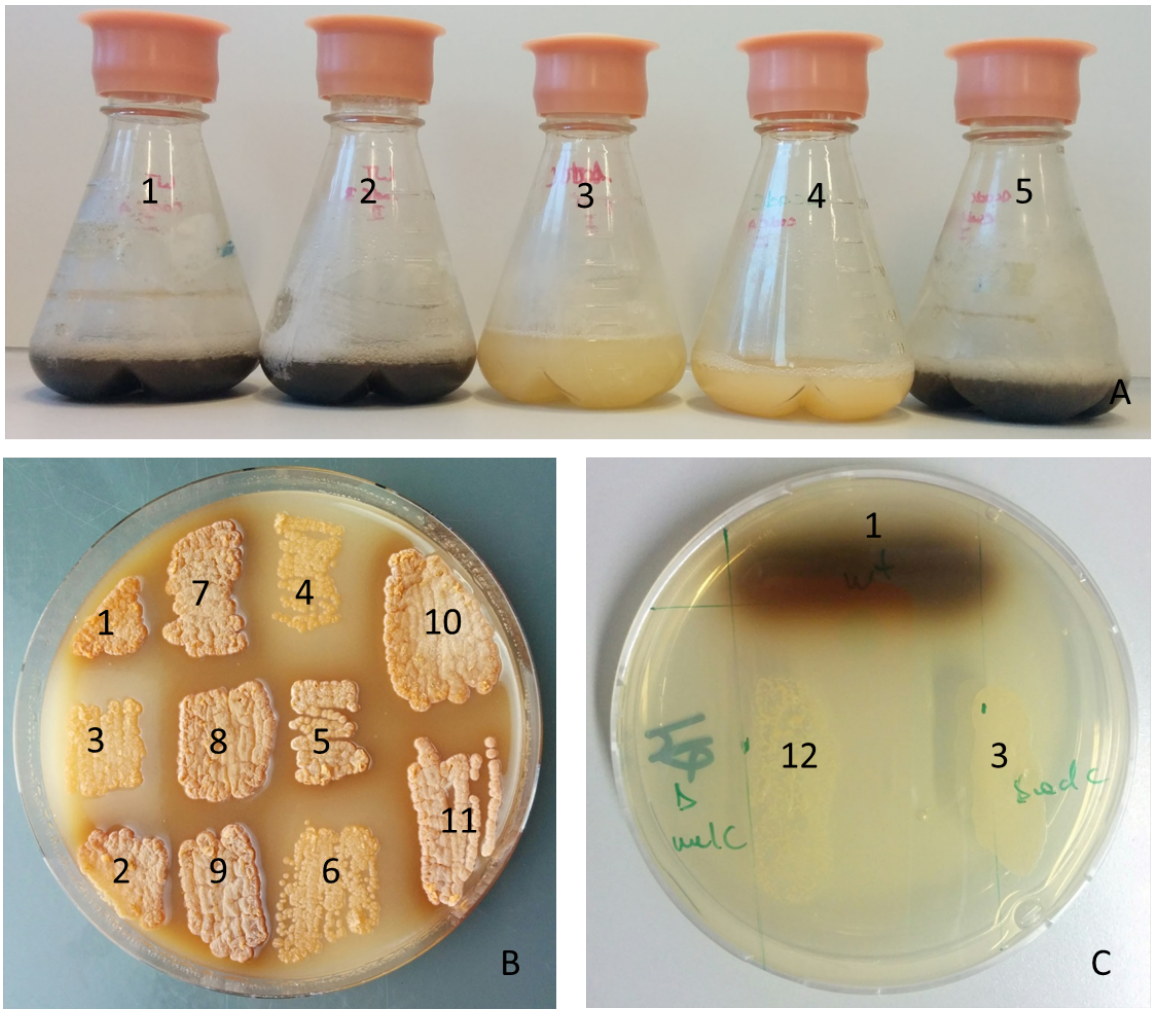

Fig. 40. Pictures, depicting morphological features of *Actinoplanes* sp. SE50/110 (1), *Actinoplanes* pSET152 (2), *Actinoplanes* Δ*cadC* (3), *Actinoplanes* Δ*cadC* pSET152 (6), *Actinoplanes* Δ*cadC* pSET152cadCA (4), *Actinoplanes* Δ*cadC* pSET152cadCB (5), *Actinoplanes* pSET152cadCA (7), *Actinoplanes* pSET152cadCB (8), *Actinoplanes* pIJ6902cadC (9), *Actinoplanes* pKC1139cadC (10) and *Actinoplanes* pKC1139 (11) strains grown in complex liquid media (A) and on SFM media (B). Figure C represents comparison of the wild type (1), *Actinoplanes* Δ*cadC* (3) and *Actinoplanes* Δ*melC* (12) strains, grown on LB plate. Strains are grown for two days in liquid media, 8 days on SFM plates and 5 days on LB plates. Pictures were made in collaboration with C. Tran. Picture B is reprinted with a permission of C. Tran ([123]).

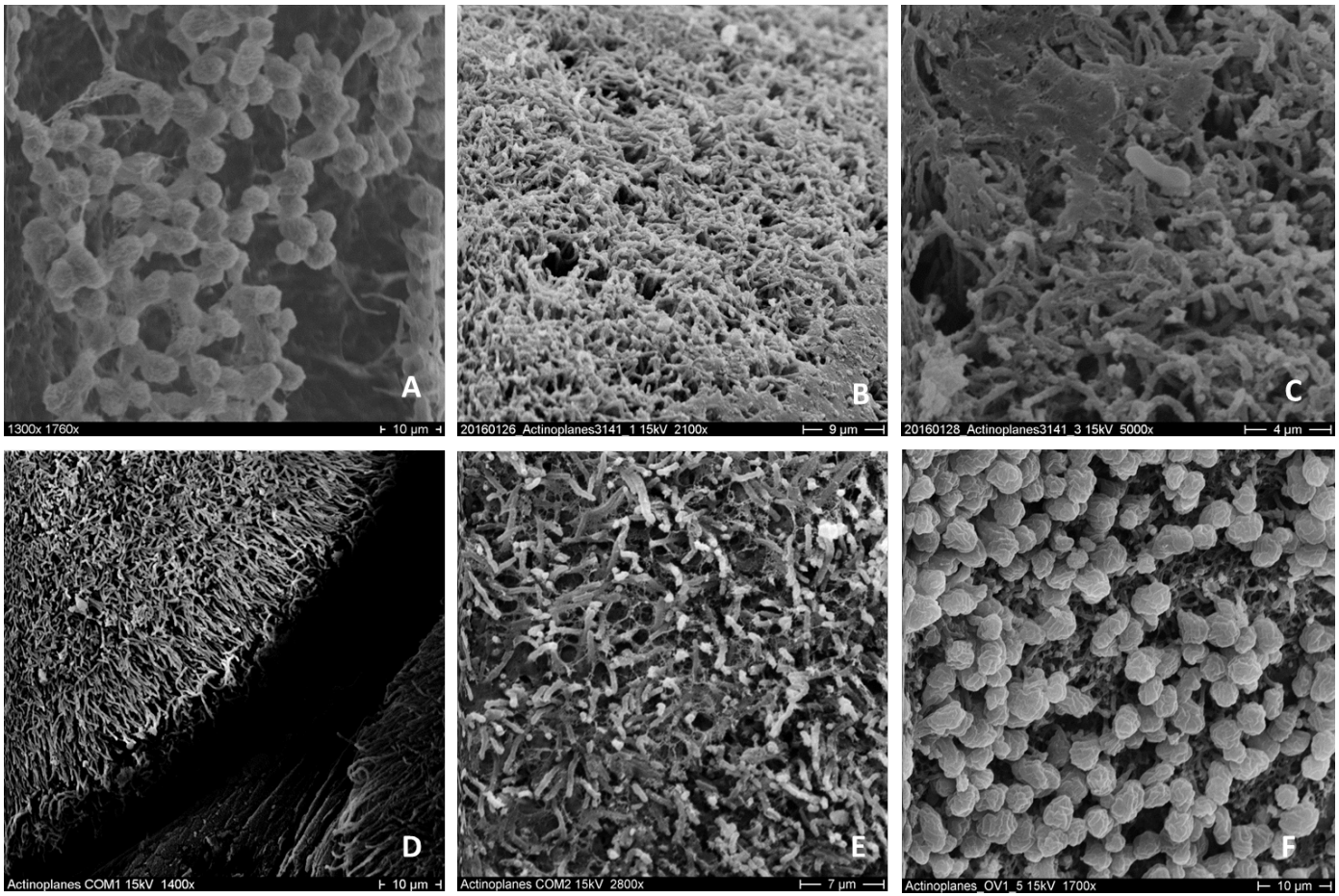

Fig. 41. Scanning electron microphotographs, depicting morphological features of *Actinoplanes* sp. SE50/110 (A), *Actinoplanes* Δ*cadC* (B), *Actinoplanes* Δ*cadC* pSET152 (C), *Actinoplanes* Δ*cadC* pSET152cadCA (D), *Actinoplanes* Δ*cadC* pSET152cadCB (E) and *Actinoplanes* pSET152cadCA (F) colonies grown on SFM media for six days. Pictures were made in collaboration with Dr. H. Bednarz, Prof. K. Niehaus and C. Tran. Picture E is reprinted with permission of C. Tran ([123]).

3.6.3.2 Characterization of *Actinoplanes* Δ*cadC* strain ability to produce acarviosyl-compounds in comparison to the wild type

Characterization of *Actinoplanes* Δ*cadC* strain ability to synthesize acarviosyl-compounds was done in collaboration with Claudia Tran and was analyzed in detail in her Master thesis [123]. Here, the summary of conducted experiments and their results is presented with the kind allowance of author.

At the beginning, HPLC measurements were carried out to quantify acarbose, produced both by the wild type and knockout mutant strain, while grown in maltose minimal media. Strikingly, knockout mutant strain was producing only small amounts of acarbose, close to its detection limit, during all cultivation procedure. It was calculated, that *Actinoplanes* sp. SE50/110 is able to produce up to 0.6 g/L acarbose after 140 h of growth while the maximal acarbose amount, detected for *Actinoplanes* Δ*cadC* strain after 72 h of growth was 0.03 g/L. In order to verify, whether *Actinoplanes* Δ*cadC* strain holds an ability to synthesize all acarviosyl metabolites, LCMS measurements of samples, collected in a stationary phase (144 hours) in a previous cultivation, were conducted. When maltose was used as a C-source, all acarviosyl-metabolites were detected in a supernatant of *Actinoplanes* Δ*cadC*, however in a considerably lowered quantities. In order to verify, whether this effect of *cadC* knockout can be neutralized with the reintroduction of the wild type copy of the *cadC* gene, two previously created complementation strains COM1 and COM2 were also analyzed regarding their metabolite biosynthetic activities. It was shown, that COM2 strain shows a partial restoration of the acarbose biosynthesis, while COM1 strain produces acarbose in amounts, close to the *Actinoplanes* Δ*cadC* strain [123]. These results are in agreement with previously described morphological characterization of all strains, while COM2 strain was shown to be similar in its phenotype to wild type and COM1 to the knockout mutant. It also signals a potential necessity of ACPL_8395, a gene located immediately upstream of *cadC*, for its expression. Also, it shows that ob-

servable changes in *Actinoplanes* Δ*cadC* strain phenotype are rather due to primary than to the secondary effects.

The results of this section signal about possible involvement of CadC into the direct or indirect regulation of acarbose metabolism. These results are in a good accordance with previous experiments. It was already shown, that CadC is binding to the regions of *acb* cluster [73], which can be also assumed from the *in silico* analysis of the protein structure, showing a putative HTH motif in CadC protein, typical for DNA binding proteins, in particular transcriptional regulators. Altogether it leads to the possibility of CadC binding the regions of *acb* cluster and therefore regulating its expression.

However, more information and experimental data is needed to clarify the involvement of CadC in regulation of cell processes. In order to study the effects of *cadC* knockout on the expression of other genes, RNA sequencing and other related transcriptomic studies were conducted. The description of these experiments can be found in next chapter.

3.6.3.3 Characterization of *Actinoplanes* Δ*cadC* strain transcriptome with the help of RNA-seq technology

As it is evident from previous experiments, described above, *Actinoplanes* Δ*cadC* strain differs significantly in its visual, morphological phenotype, and also in its ability to produce acarviosyl- metabolites, while cultivated in typical for *Actinoplanes* minimal media with different C-sources. In order to verify, whether these and other changes in *Actinoplanes* phenotype are linked to changes in gene expression, rather than other effects, RNA sequencing studies of *Actinoplanes* Δ*cadC* were conducted. This powerful technique gives a possibility to gain a unique insight into the expression of an overwhelming majority of genes in a particular condition of interest. In this case, it also gives a possibility to discover other changes, possibly caused by the knockout of *cadC*, which are not obvious, or not possible to learn from observation,

e.g. changes of transcription of other regulatory genes. Importantly, these studies can give a strong hint about the processes of acarbose metabolism regulation and their correlation.

RNA samples for this experiment were isolated from cultures, harvested at the exponential growth phase (48 hours of growth), grown in the maltose liquid minimal media. After the isolation, RNA quality was assessed by PCR to eliminate the DNA impurities. For the further library preparation procedures, three biological replicates, both for wild type and for *Actinoplanes* Δ*cadC* strain respectively, were mixed in two samples. In this case, two following Illumina sequencing runs were performed instead of typical six. This experimental set up is helpful in order to reduce high experimental cost, however it does not allow to conduct meaningful statistical analysis of the resulting RNA sequencing data in order to identify significantly differentially expressed genes. Otherwise, it gives a good outlook on the changes of transcriptome of the knockout mutant strain. In this case, to confirm the significance of the changes, observed in results, typically other quantitative methods are used, e.g. qRT PCR. Details regarding the experimental set up of this experiment can be found in Materials and methods section of this thesis.

As a result of the sequencing run, 3,466,738 and 5,029,123 reads were generated for *Actinoplanes* Δ*cadC* and wild type respectively. The quality assessment of raw reads and read trimming was done by FastQC [5] and using the script, kindly provided by Dr. Daniel Wibberg (unpublished data; personal communication). Following mapping of all reads to the reference genome of *Actinoplanes* sp. SE50/110 was done with bowtie2 [74]. As a result of mapping, of all reads, generated for the knockout strain, 3,466,738 (100.00%) were paired; of these: 64,397 (1.86%) were aligned concordantly 0 times, 3,330,935 (96.08%) aligned concordantly exactly 1 time, 71,406 (2.06%) aligned concordantly more than 1 time; an overall alignment rate was 98.83%. Of all reads, generated for the wild type, 5,029,123 (100.00%) were paired; of these: 77,939 (1.55%) aligned concordantly 0 times, 4,795,440 (95.35%) aligned concordantly exactly 1 time, 155,744 (3.10%) aligned concordantly more than 1 time;

an overall alignment rate was 99.16%. The data was then imported to ReadXplorer 2.1.0 [53] for analysis and visualization. DESeq2 tool [77] was used further, as an integrated feature of ReadXplorer, in order to detect differentially expressed genes. Results of DESeq2 analysis were sorted according to the log2Fold Change value, only results with $\{x \mid |x| > 1.4\}$ were considered for further analysis. In this case, however $p_{adjusted}$ value can not be calculated due to the absence of replicates; a p-value cut-off of 0.05 was therefore applied.

Among all of the considered putatively differentially expressed genes, of a high interest are genes of *acb* cluster (Figure 42). As it is evident from results, all genes of this cluster are expressed in a lower manner, except of *acbD, acbE and acbZ* genes, whose expression seems to be not effected. These results were confirmed with qRT PCR, which provides quantitative confirmation (Figure 43). As it is evident from qRT PCR results, relative amounts of *acbV* transcripts are significantly lowered in comparison to the wild type. At the same time, no significant difference is observed in case of *acbD* and *acbE*. Gene *rpoB* was used as a control for comparison. It is also interesting to note, that the expression of ACPL_213, a neighboring gene of *cadC* is relatively elevated. This effect can be explained by the effect of *aac*3(IV) gene promoter region on the expression of neighboring genes, which was observed previously in case of *acbD* gene. However, relative expression of ACPL_8395 is not significantly influenced in comparison to the wild type.

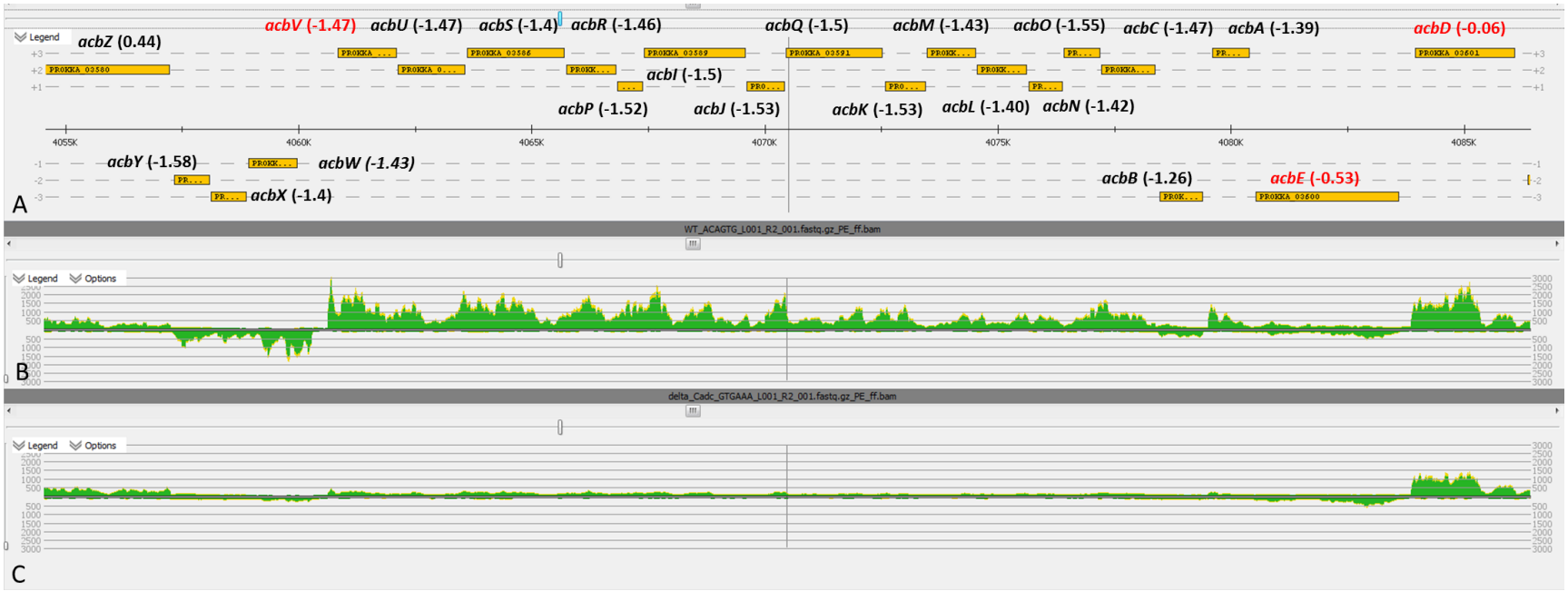

Fig. 42. An overview of the transcriptional activity in the *acb* cluster genetic region from RNA sequencing data of *Actinoplanes* Δ*cadC* strain and wild type strain, visualized by ReadXplorer 2.1.0; A – section, depicting the location of *acb* cluster genes. Each gene is marked by yellow rectangle. Gene names are indicated together with the calculated log2Fold Change value in parenthesis. For genes, which names are highlighted in red, additional qRT PCR experiments were performed; B – whole transcriptome sequencing data of *Actinoplanes* sp. SE50/110; C – whole transcriptome sequencing data of *Actinoplanes* Δ*cadC* strain.

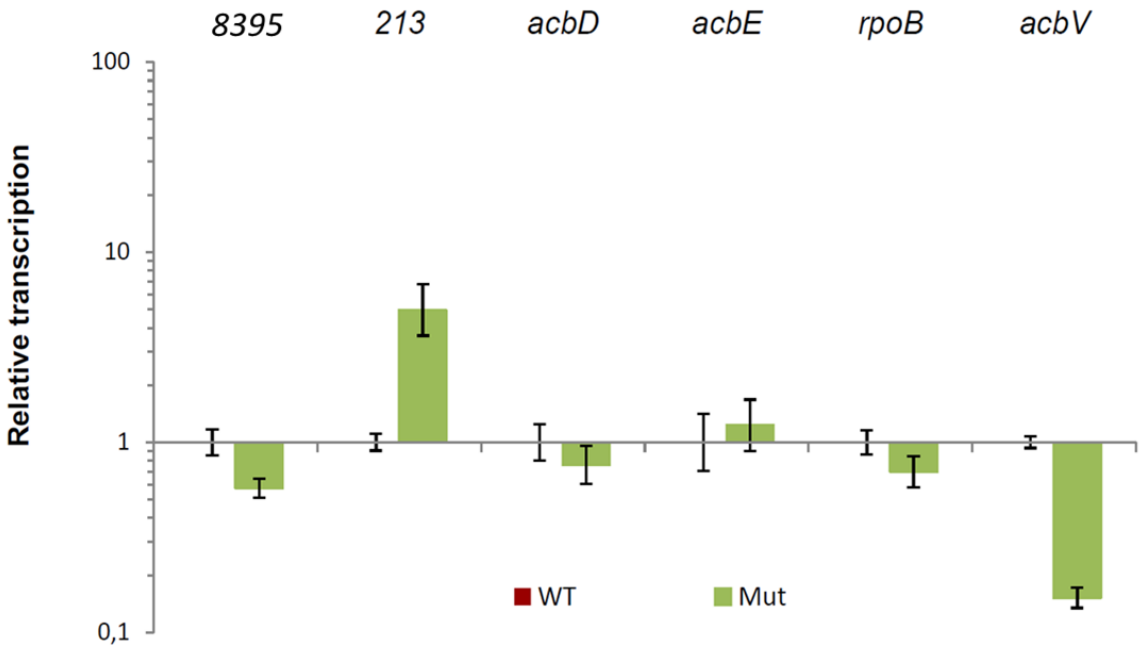

Fig. 43. qRT-PCR measurements of *Actinoplanes* Δ*cadC* RNA samples in comparison to the wild type; samples are isolated from cultures, grown in maltose minimal media, 72 hours of incubation. Names of genes are as follows: 8395 - ACPL_8395, 213 - ACPL_213, *acbD, acbE, rpoB, acbV*.

Several other interesting results can be highlighted from RNA sequencing data. For example, gene *melC*2 (-1.62 log2Fold Change value), coding for putative enzyme tyrosinase, potentially involved in synthesis of melanin, was also among the genes, for which amount of transcripts was significantly lowered in comparison to the wild type. It was evident from the RNA sequencing data, that there is a strong antisense transcription in the region of the *melC*2 ORF, which could cause this effect. The antisense transcription was verified with qRT-PCR in [123]. This effect could explain characteristically pale colour of *Actinoplanes* Δ*cadC* colonies in comparison to the dark brown pigmented wild type colonies. This suggestion is further reinforced by the considerable morphologic similarity of *Actinoplanes* Δ*cadC* colonies and *Actinoplanes* Δ*melC* colonies, which was described previously.

It is interesting to point out the comparatively lowered amount of transcript of putative transcriptional regulator *deoR* or ACPL_6374 (-1.74 log2Fold Change value). As per its genomic position, it is likely to form operon with ACPL_6367 - ACPL_6373 genes, which products are forming a putative phosphoenolpyruvate sugar phosphotransferase system or PTS system. All of these genes also are characterized by a lowered transcript abundance in comparison to the wild type. Further *in silico* prediction of the function of ACPL_6374 attribute it a function of a putative transcriptional regulator of sugar metabolism DeoR. Pfam search predicts both presence of HTH domain and DeoR sensor domain. Proteins of this family posses a wide range of functions, among which is an ability to regulate the transport of various sugars, e.g. fructose, glucose, and sucrose [35]. The results of RNA sequencing were further verified and confirmed with qRT-PCR experiment, which shows significantly lowered amount of ACPL_6374 transcript abundance in comparison to the wild type (data not shown). However, till now existence of PTS transport systems and their regulation is not proven or studied in case of *Actinoplanes* sp. SE50/110, therefore it is possible only to assume this function of ACPL_6367 - ACPL_6373 operon. RNA sequencing data, which was received in this experiment was further verified with whole genome microarray analysis of the knockout mutant, wild type and complementation strains in work of Claudia Tran [123].

Results, received in RNA sequencing analysis are in good agreement with previous phenotypic characterization of *Actinoplanes* Δ*cadC* mutant. The dramatic decrease in acarbose production of *Actinoplanes* Δ*cadC* strain in comparison to the wild type is seemingly caused by the lower abundance of the *acb* cluster genes transcripts. CadC regulator is the second regulator of *Actinoplanes* being studied so far in detail. MalR regulator, which was studied earlier, was shown to influence the transcription of several genes from *acb* gene cluster, namely *acbD* and *acbE*, by binding to the intergeneric region between two genes [27]. Interestingly, knockout of *cadC* does not influence expression of these genes, but the rest of the *acb* cluster. Therefore, it is possible to imagine, that these both regulators possess different functions in acarbose

metabolism and complement each other to a certain extent. CadC regulator seems to posses pleiotropic functions in metabolism of *Actinoplanes* sp SE50/110, while its knockout causes not only abolishment of acarbose production, but drastic changes in pigment biosynthesis and its morphology. From the *in silico* prediction, ACPL_212 is expected to belong to the ArsR family, which are known to be involved in a stress-response to heavy metal ions. However, this assumption did not get a firm support in a light of results described above. It is difficult to speculate, what is the exact function and the mechanism of CadC action. It is possible to assume, that this regulator possesses the function, similar to TrmB family of regulators, mainly involved in sugar transport regulation. This hypothesis is further supported by the fact, that knockout of *cadC* causes lower transcript abundance of several genes, putatively involved in sugar transport and the regulation of this processes, e.g. ACPL_6367 - ACPL_6373 operon and ACPL_6374. Also, regulation of *acb* gene cluster expression falls well into this picture, as acarbose was thought to be involved in the harnessing pool of carbohydrates as a possible "carbophore" molecule. However, it is important to stress, that these speculations need further experimental confirmation to be verified. The fact of binding of CadC protein and its exact position in *acb* cluster can be established with the help of electrophoretic mobility shift assays. This experiment also gives a possibility to test a wide range of possible ligand molecules, which can be bound by CadC. This and further work on the topic of acarbose metabolism regulation will be further conducted in Actinoplanes working group.

4. CONCLUSIONS

In the recent years, using transcriptomics, proteomics and metabolomics a large amount of data on acarbose, a secondary metabolite of *Actinoplanes* sp. SE50/110, was acquired. And on these grounds, a number of putative models of acarbose metabolism and its transport were suggested. In order to verify the validity of such models and to supplement them with the additional information regarding steps, which are not clear or do not have any, direct or indirect confirmations yet, a suitable genetic engineering system was urgently required. Until now, the possibility to conduct genetic research of the wild type strain did not exist, due to the lack of available tools and techniques for *Actinoplanes* sp. SE50/110. As a first option to study the functions of Acb proteins, heterologous expression of the whole *acb* cluster in an Actinobacteria host, was explored. Despite successful transfer of the whole *acb* cluster to the specially designed strains of *S. coelicolor*, no acarbose biosynthesis was detected. Further studies have shown, that the reason for this effect probably lies in the low to none transcription of *acb* cluster in heterologous conditions. It could be speculated that the causes for such effect can stem from multiple reasons, e.g. possible absence of regulatory genes in genome of *S. coelicolor*, which are necessary for the metabolism of acarbose or conditions being non optimal to assure the production of acarbose. Hence the study was focused towards development of the native genetic engineering system, that can allow for the determination for exact reason for such an effect.

As a main result of this work, the development of full system of methods, which allow the genetic manipulations to be conducted in native conditions, is described. Gene cloning system was developed in a set of consecutive steps. As a first step, growth of *Actinoplanes* sp. SE50/110 was studied in various solid media. SFM media was identified as an optimal media for the growth and sporulation of *Actinoplanes* sp. SE50/110. For the first time, it was confirmed, that wild type strain is able to

abundantly produce globular sporangia on a thin "sporangiophore"-like structures without formation of the aerial mycelia, while grown on SFM plates. Therefore, this media was adopted for further experiments. As a second step, antibiotic resistance spectra of *Actinoplanes* sp. SE50/110 was studied. This information was further used in selection of appropriate marker genes, used as a part of vectors or for gene knockout experiments. As a third step, three types of DNA transfer systems were tested: protoplast transformation, electroporation and *E. coli* - *Actinoplanes* intergeneric conjugation. The best results were achieved with the *E. coli* - *Actinoplanes* intergeneric conjugation system, which was further optimized for the transfer of DNA molecules to *Actinoplanes* sp. SE50/110. Three different actinophage-based integrative vectors were transferred to *Actinoplanes* sp. SE50/110 via conjugation procedure. Studies of exconjugant clones have shown that all vectors are integrated in single copy per-genome, are inherited stable over multiple generations and do not influence *Actinoplanes* morphology or acarbose production. These results show that all of them can be used for further genetic experiments. Further on, integration sites for φC31, VWB and φBT1 bacteriophages in the *Actinoplanes* sp. SE50/110 genome were described and characterized for the first time.

Furthermore, three replicative vectors, based on different replicons were tested for their usage in *Actinoplanes* sp. SE50/110. Vector pKC1139, based on pSG5 temperature sensitive replicon, was shown to have no influence neither on morphology of *Actinoplanes* nor on its metabolite production. It was shown to be stably inherited by cells even at elevated temperatures. Therefore, it can be used for genetic engineering of *Actinoplanes*. As a fourth step, various techniques for constructing gene knockouts in *Actinoplanes* sp. SE50/110 were tested. At first, gene disruption was attempted with "suicide" vector via single crossover and in parallel with replicative pKC1139 vector via double crossover. However, both of these approaches did not succeed. This could be due to the low recombination frequencies between homologous regions of the vectors, used for gene knockouts, and chromosome of the wild type strain. Such an effect can be caused by the overlapping sequences in both

cases being not large enough for recombination to be executed successfully. In order to enlarge the overlapping sequences between disruption construct and chromosome, cosmid library of *Actinoplanes* sp. SE50/110 genome was further used for building gene knockouts. A ReDirect procedure, originally designed for streptomycetes, was than utilized for knockout construction. ReDirect deploys PCR-targeting of the ORF of gene of interest, situated within a cosmid, being replaced with antibiotic marker cassette. Such cassette is being generated by PCR and includes the *oriT* region, which allows it to be transferred to cells of *Actinoplanes* via conjugation procedure. With the help of ReDirect, knockout constructs of majority of *acb* cluster genes were generated. Furthermore, knockouts within cACPL4 gene cluster and of putative regulatory genes of acarbose were generated. Three *Actinoplanes* knockout mutants, deficient in genes *acbA, acbB and acbD* respectively, were constructed in the course of this work. However, due to the time limitation of the study, phenotypes of only two strains *Actinoplanes*Δ*acbD*, knockout of the gene coding for putative acarviosyltransferase enzyme, and *Actinoplanes*Δ*cadC*, knockout of the gene coding for a putative regulator of acarbose metabolism, were studied in detail. Knockout libraries, generated in this work, will be used further to study metabolism of acarbose and its regulation in *Actinoplanes* sp. SE50/110.

The correct construction of both mutant strains was verified by whole genome sequencing. No transcription signals from these genes were detected either in microarray experiments nor qRT-PCR measurements. Absence of AcbD protein in extracellular proteome of *Actinoplanes* was confirmed in independent experiment. For both knockout mutants, complementation and overexpression strains, based on previously studied integrative and replicative vectors, were constructed. In order to characterize the biosynthetic abilities of *Actinoplanes*Δ*acbD* strain, cultivations of all of the strains were conducted in liquid minimal media, supplemented with various C-sources, namely glucose, maltose, maltotriose and Cpur. HPLC and LCMS measurements were conducted to characterize the amount and spectra of acarviosyl metabolites, accumulated in supernatant in these media. As a main result, signifi-

cantly lowered amounts of acarbose were produced by *Actinoplanes*Δ*acbD* strain in all media in comparison to the wild type. However, at the same time, mutant was shown to retain the ability to synthesize all minor acarviosyl metabolites, which are possible to measure with current techniques. Whole genome comparative microarrays study of RNA samples, isolated from cultures grown in Mal-MM have shown, that expression of *acb* cluster is not significantly changed in comparison to the wild type strain. Therefore, the observed phenotype of the mutant is probably not caused by the changes in gene expression. The complementation of the knockout mutant with the wild type copy of the *acbD* gene, leads to the restoration of levels of acarviosyl metabolites biosynthesis, which are comparable of those of the wild type strain. At the same time, overexpression of the wild type strain copy of *acbD* gene does not lead to a significant changes in amounts of acarviosyl metabolites, synthesized by the wild type strain or their pattern. These facts could be used to state, that the phenotype of *Actinoplanes*Δ*acbD* strain is not dependent on a any secondary effects, but most probably is a direct display of the lack of function of AcbD protein. These results have lead to the revision of previously suggested acarbose metabolism models, none of which apparently can fully explain the observed phenotype of *Actinoplanes*Δ*acbD* strain. It is possible to speculate, that AcbD, apart from the functions, which were predicted before, is directly involved in the last step of acarviosyl metabolite biosynthesis, intracellularly or extracellularly, as it was previously suggested in [51]. However, in order to clarify the role of AcbD in acarbose metabolism, further studies are needed. Characterization of the phenotypes of other *acb* gene knockout strains, together with double- or triple- knockout combinations of these genes, will give a possibility to formulate the complete model of acarbose metabolism.

*Actinoplanes*Δ*cadC* strain was characterized by its apparent differences in morphology and pigment biosynthesis in comparison to the wild type, which was confirmed through the preparation of SEM images. Regarding its biosynthetic capabilities, only low amounts of acarbose were detected in case of Mal-MM cultivations. RNA sequencing, coupled with subsequent qRT-PCR analysis, have shown that all

acb cluster genes, but three (*acbD, E, Z*) were expressed in significantly lowered manner in the mutant strain. Therefore, gene *cadC* became the first known regulator of acarbose metabolism to influence the absolute majority of *acb* genes. Future studies will help to clarify the exact role and function of this regulator. As a continuation of work involving the application of novel methods of gene knockouts and genetic engineering, CRISPR-Cas9 system was applied. It was used to delete the gene coding for putative tyrosinase *melC*2. The constructed mutant possessed typical for such mutations phenotype, characterized by the loss of colonies brown colouring. No undesired genetic rearrangements were detected in genome of the mutant strain. Therefore, it is possible to conclude, that CRISPR-Cas9 system can be successfully applied for gene knockout experiments.

In order to further broaden the set of genetic tools, available for *Actinoplanes* sp. SE50/110, applicability of GUS reporter system was studied. It was first tested by transferring the pSETGUS vector, which is based on the integrative vector pSET152, to *Actinoplanes* sp. SE50/110. In pSETGUS the *gusA* gene is controlled by the *tipA* promoter, which was formerly characterized as thiostrepton-inducible. By adding XGluc onto the colonies of *Actinoplanes* sp. SE50/110, it was possible to show that the *gusA* gene is expressed even in the absence of thiostrepton addition. Furthermore, the set of promoters, namely *ermEp, cdaRp, wblAp, moeE5p and aac3(IV)p* were tested for their applicability for the genetic engineering of *Actinoplanes* sp. SE50/110 with the help of GUS system. Therefore, it was proven, that *gusA*-based vectors can be used as a reporter system for the measurement of successful gene transfer and expression in *Actinoplanes* sp. SE50/110.

The developed gene transfer system evidently offers multiple possibilities to help answering fundamental theoretical and practical questions regarding acarbose production. Namely, formulation and verification of the complete model of acarbose and its regulation as well as the development and optimization of acarbose overproducing strains is now possible by the use of the developed genetic engineering system.

ACKNOWLEDGMENTS

This work was carried out in Center of Biotechnology (CeBiTec), Bielefeld University, Germany in the group of Prof. Dr. Jörn Kalinowski.

I am sincerely grateful to:

my supervisors, Prof. Dr. Jörn Kalinowski and Prof. Dr. Alfred Pühler for their supervision of my PhD thesis and for the great possibility to perform my work in the Center of Biotechnology. Moreover, for their constant support and motivation, for fruitful discussions and inspiration.

Prof. Dr. Karsten Niehaus and Dr. Hanna Bednarz for their guidance and help on multiple occasions and for the cooperation in the course of my work.

Dr. Susanne Schneiker-Bekel for her daily help in my work, enthusiastic encouragement throughout my Ph.D. program and for her critical and insightful comments on research papers and this thesis.

Claudia Tran for the possibility to conduct research together and to supervise her Master thesis.

Büsra Gülücü for our work together on a project in a course of her Bachelor thesis.

Dr. Iris Brune and Dr. Karina Brinkrolf for their help in all struggles of my daily life, for a cheerful encouragement and invaluable support throughout my PhD.

All current and former members of *Actinoplanes* and Coryne group. Particularly, Dr. Sergej Wendler and Dr. Vera Ortseifen for joint work on many projects, for their friendship and warm working atmosphere. Dr. Almut Mentz, Dr. Arwa Al-Dilaimi, Dr. Armin Neshat, Dr. Tobias Busche, Dr. Robert Kulis-Horn, Dr. Markus Persicke, Dr. Christian Rückert, Julian Droste, Lena Shaffert for being wonderful colleagues and friends.

Manuel Wittchen for his help with analysis of RNA-sequencing data and for his valuable advices.

Dr. Daniel Wibberg for his help in analysis of sequencing data on multiple occasions.

Dr. Till Zemke for a co operation on multiple projects in a course of my studies, as well as for a thoughtful revision and comments on manuscripts.

Dr. Udo Wehmeier for his revision of this thesis and for constant support and advice throughout my PhD.

Anika Winkler and Katharina Hanuschka for their help in Sanger and Illumina sequencing.

Gudrun Horstmann, Silvia Mai and Dr. jur. Werner Glenewinkel for their friendship and constant support. For being an amazing neighbors and friends.

My friends Sherzod Hakimov, Ekin Yildiran, Stefania Colella, Oresta Vasyliv and Olga Tsypik for always being there for me.

My parents Petro and Galyna Gren, Satish and Sheela Yadav, husband Ankit Kariryaa, my whole family for making my life meaningful.

I would like to acknowledge financial support from the Graduate Cluster Industrial Biotechnology (CLIB2021) at Bielefeld University, Germany. The Graduate Cluster is supported by Bielefeld University and the Ministry of Innovation, Science and Research (MIWF) of the federal state North Rhine-Westphalia, Germany.

Erklärung

Hiermit erkläre ich, dass ich die hier vorliegende Dissertation eigenständig und ohne unerlaubte Hilfe angefertigt habe. Ich versichere, dass ich keine anderen als die angegebenen Quellen sowie Hilfsmittel verwendet habe. Die Stellen der Arbeit, die anderen Quellen im Wortlaut oder dem Sinn nach entnommen wurden, sind durch Angaben der Herkunft kenntlich gemacht. Die Dissertation wurde weder in der vorgelegten noch in ähnlicher Form bei einer anderen Institution eingereicht. Ich bewerbe mich hier erstmalig um den Doktorgrad der Naturwissenschaften.
Gedruckt auf alterungsbeständigem Papier ° ° ISO 9706.

Tetiana Gren

APPENDIX

A. SUPPLEMENTARY MATERIAL

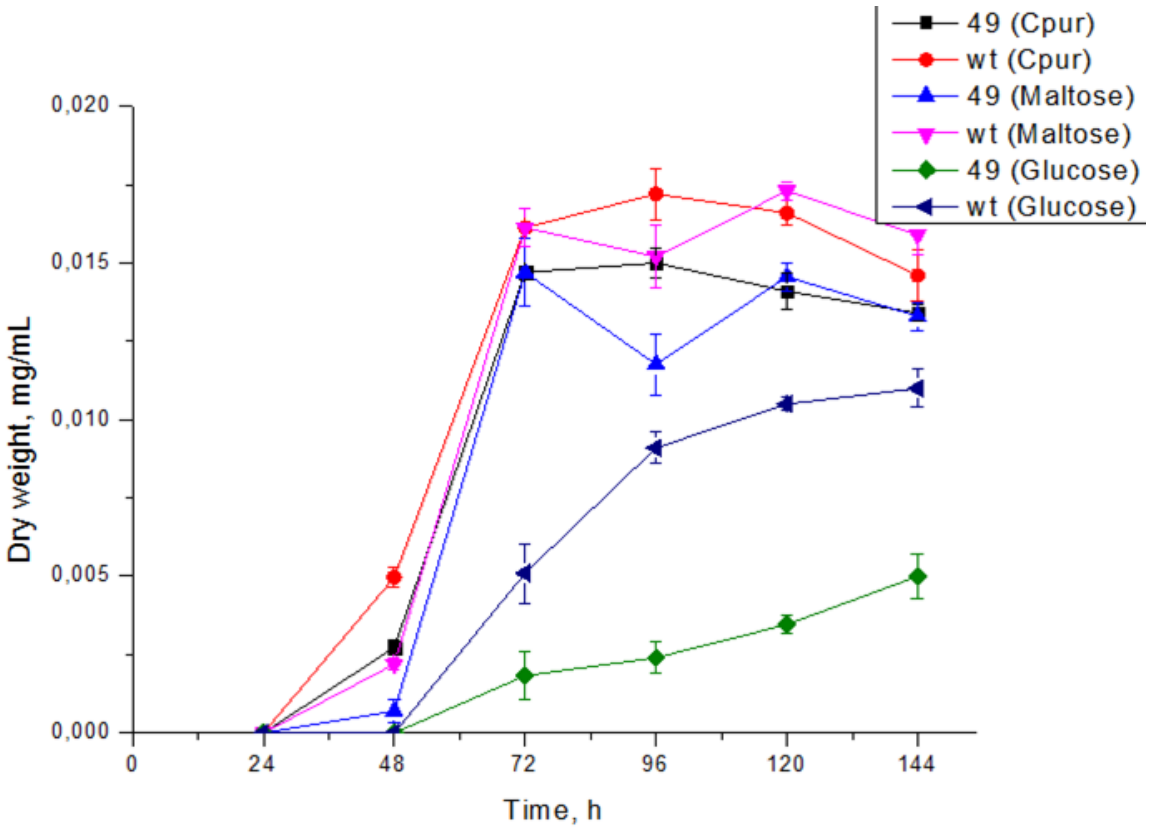

Fig. 44. Growth of *Actinoplanes* $\Delta acbD$ 49 strain and wild type in minimal media, supplemented with glucose, maltose, and Cpur. Mean values and standard deviations are shown for three biological replicates.

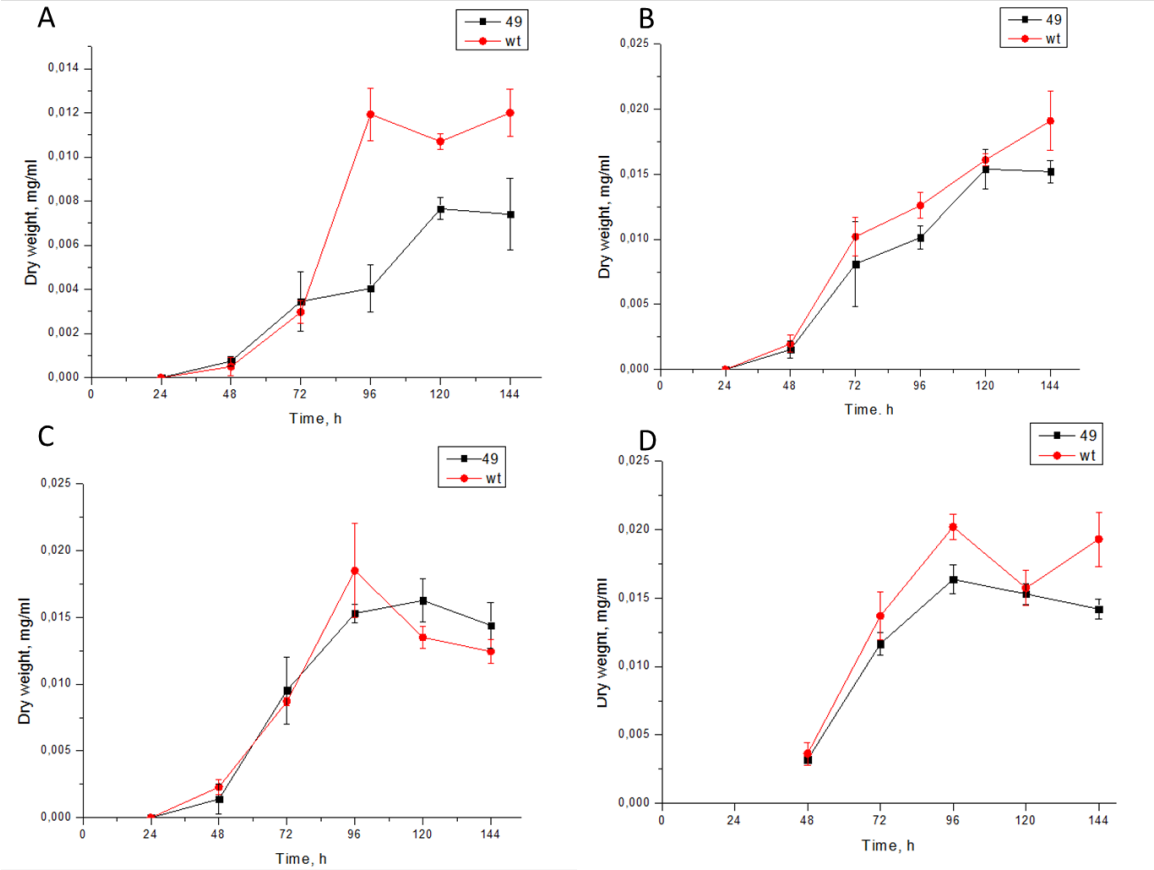

Fig. 45. Growth of *Actinoplanes* Δ*acbD* 49 strain and wild type in minimal media, supplemented with: A – glucose, B – maltose, C – maltose/glucose, D - Cpur. Mean values and standard deviations are shown for three biological replicates

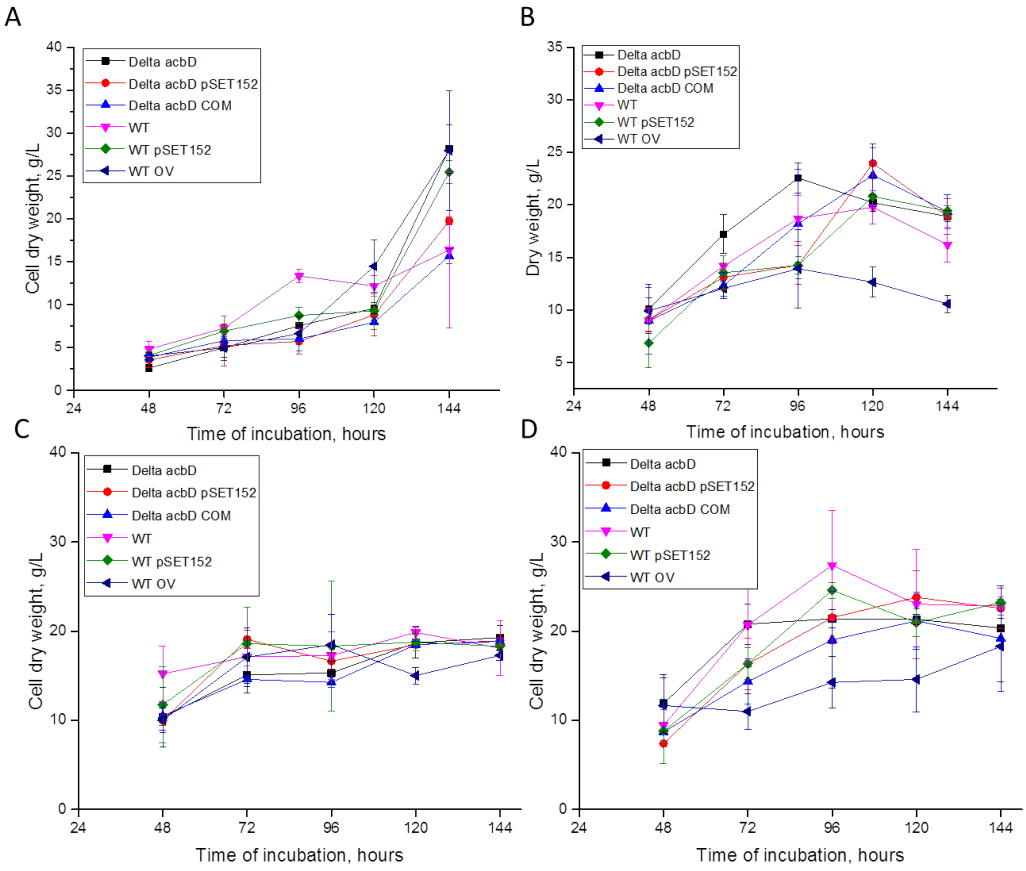

Fig. 46. Growth of *Actinoplanes* Δ*acbD* 49 strain, *Actinoplanes* Δ*acbD* 49 complementation strain, *acbD* ovrexpression strain, wild type and respective control strains in minimal media, supplemented with: A – glucose, B – maltose, C – maltotriose, D - Cpur; strains are marked as follows: Delta acbD - *Actinoplanes* Δ*acbD* 49 strain; Delta acbD pSET152 - *Actinoplanes* Δ*acbD* pSET152; Delta acbD COM - *Actinoplanes* Δ*acbD* pSET152acbD; WT - *Actinoplanes* sp. SE50/110; WT pSET152 - *Actinoplanes* pSET152; WT OV - *Actinoplanes* pSET152acbD. Mean values and standard deviations are shown for three biological replicates.

No	Name	Sequence	No of bases	Remarks
1	aac3(IV)fwd	GTGCCGTTGATCGTGCTATG	20	Used to test the presence of pSET152, pSOK804, pIJ6902, pRT801 and pKC1139 in cells of *E. coli* or *Actinoplanes*
2	aac3(IV)rev	TCCAACGTCATCTCGTTCTCC	21	Same as above
3	oriTfwd	GGATCGGTCTTGCCTTGCTC	20	Same as above
4	oriTrev	CAGGTCGACGGATCTTTTCC	20	Same as above
5	T52-qRT_acbA_fwd	CCGGCTTCTACCTCTACGAC	20	Used for qRT-PCR
6	T53-qRT_acbA_rev	TGGTTGACCTCGGTGATCTC	20	Same as above
7	T008-acbD-com-fwd	GACTCTAGAGGATCCGCGGCCGCGCGCGATTCCATCGTGGACCCTCTCTC	50	Used to generate pSET152hygacbD construct
8	T009-acbD-com-rev	AAACAGCTATGACATGATTACGAATTCGATTGGTGGTCTACCTGCTCCTG	50	Same as above
9	T022-cadC_com_fwd_1	GACTCTAGAGGATCCGCGGCCGCGCGCGATAAGTTAGCCGGTTCGTTCCT	50	Used to generate pSEThyg152cadCA construct

No	Name	Sequence	No of bases	Remarks
10	T023-cadC_com_rev_1	AAACAGCTATGACATGATTACGAATTCGATCCCACGTTCCAGTGACGCCA	50	Same as above
11	T024-cadC_com_fwd_2	GACTCTAGAGGATCCGCGGCCGCGCGCGATCCATGCGGAACCACGCCTAG	50	Used to generate pSET152hygcadCB construct
12	T025-cadC_com_rev_2	AAACAGCTATGACATGATTACGAATTCGATGCGTCCCTCCTCGAATCGCC	50	Same as above
13	T026-212_cf_tipA	GCGTCAGAGAAGGGAGCGGAAAGTTAGCCGGTTCGTTCCT	40	Used to generate pIJ6902cadC construct
14	T027-212_cR_tipA	AGGCATGCAAGCTTCATATGGCGTCCCTCCTCGAATCGCC	40	Same as above
15	T028-6902_212_f	GGCGATTCGAGGAGGGACGCCATATGAAGCTTGCATGCCT	40	Same as above
16	T029-6902_212_r	AGGAACGAACCGGCTAACTTTCCGCTCCCTTCTCTGACGC	40	Same as above

No	Name	Sequence	No of bases	Remarks
17	T030-212_qrt_fwd	CTGCTGAGCCTTATCCAGT	19	Used for qRT-PCR
18	T031-212_qrt_rev	GACCATCGCCGACCTTCTGA	20	Same as above
19	T032-213_qrt_fwd	GCTGCATCTCGCTCTGTACC	20	Same as above
20	T033-213_qrt_rev	CACTCGGCGCAGATAGAACG	20	Same as above
21	T034-211_qrt_fwd	TCGAGAGCTCGACACGGATG	20	Same as above
22	T035-211_qrt_rev	GTAAGGCCGACTTCGATCGC	20	Same as above
23	TG27-acbA_RR_fwd	GAGCCATTTACGGGGCTTGCTGATATGGTC GGTCACGTGATTCCGGGGATCCGTCGACC	59	Used for the amplification of the *acbA*-containing cassette for ReDirect
24	TG28-acbA_RR_rev	CTTCGCCGCCCGGGCCGGTCACCCTCCGCG GGCCGCTCATGTAGGCTGGAGCTGCTTC	58	Same as above

No	Name	Sequence	No of bases	Remarks
25	TG29-acbB_RR_fwd	CGGGCACCCATTGGCGGCATTGCTACAATCCCGGCGATGATTCCGGGGATCCGTCGACC	59	Used for the amplification of the *acbB*-containing cassette for ReDirect
26	TG30-acbB_RR_rev	GGAAGCGACCGGCGTCCGTCCGCCCACCGGTTGCCGTGATGTAGGCTGGAGCTGCTTC	58	Same as above
27	TG31-acbC_RR_fwd	GCTGACGTGAGTGGTGTCGAGACGGTAGGGGTGCACGCGATTCCGGGGATCCGTCGACC	59	Used for the amplification of the *acbC*-containing cassette for ReDirect
28	TG32-acbC_RR_rev	CCGGTCGCTTCCGGCGATCGGCGTCCGCGGCCCGAGCTATGTAGGCTGGAGCTGCTTC	58	Same as above
29	TG33-acbE_RR_fwd	GGCTAGCGTCCCAGCGATCGCGAGAGAGGGTCCACGATGATTCCGGGGATCCGTCGACC	59	Used for the amplification of the *acbE*-containing cassette for ReDirect
30	TG34-acbE_RR_rev	AGGAACCGGCCCGGTCCGTGGTCGCCGCAGCGCCGGTCATGTAGGCTGGAGCTGCTTC	58	Same as above

No	Name	Sequence	No of bases	Remarks
31	TG35-acbl_RR_fwd	ATCGTGCACGTGGTGAGCTTCGCCTTCGAAGCCCTCGCCATTCCGGGGATCCGTCGACC	59	Used for the amplification of the *acbl*-containing cassette for ReDirect
32	TG36-acbl_RR_rev	GTGCCGATGCCGTCCATGTCAGCGGCCCCCGGCCCCGCATGTAGGCTGGAGCTGCTTC	58	Same as above
33	TG37-acbJ_RR_fwd	CGGTCGGGGCCGGGGGCCGCTGACATGGACGGCATCGGCATTCCGGGGATCCGTCGACC	59	Used for the amplification of the *acbJ*-containing cassette for ReDirect
34	TG38-acbJ_RR_rev	CGGCATCCGTCGTGGTGGTCATCGCAGGCTTGCCTTTCGTGTAGGCTGGAGCTGCTTC	58	Same as above
35	TG39-acbK_RR_fwd	CCTGACCTCGCTGACCCGCACCTTTCAAGGAGATCCATGATTCCGGGGATCCGTCGACC	59	Used for the amplification of the *acbK*-containing cassette for ReDirect
36	TG40-acbK_RR_rev	GGTGGTGCGGTGGCCGCTTCACGCCGCCGGCCGCAGCAGTGTAGGCTGGAGCTGCTTC	58	Same as above

No	Name	Sequence	No of bases	Remarks
37	TG41-acbL_RR_fwd	GGAAGCGTTGGTCGGGCGATGAGCCGGCACCGCGCGATCATTCCGGGGATCCGTCGACC	59	Used for the amplification of the *acbL*-containing cassette for ReDirect
38	TG42-acbL_RR_rev	CCAGAGTCCCGCTCATCGATGATCCTCTCTGACATTGACTGTAGGCTGGAGCTGCTTC	58	Same as above
39	TG43-acbM_RR_fwd	GCGGCCGGCGGCGTGAAGCGGCCACCGCACCACCCGGTCATTCCGGGGATCCGTCGACC	59	Used for the amplification of the *acbM*-containing cassette for ReDirect
40	TG44-acbM_RR_rev	GCACGATCGCGCGGTGCCGGCTCATCGCCCGACCAACGCTGTAGGCTGGAGCTGCTTC	58	Same as above
41	TG47-acbO_RR_fwd	CCGGCCGGTGCGGACGTGACCTGCCGGGTGGGGCTGACCATTCCGGGGATCCGTCGACC	59	Used for the amplification of the *acbO*-containing cassette for ReDirect
42	TG48-acbO_RR_rev	CTACCGTCTCGACACCACTCACGTCAGCTTCCCTTTCTCTGTAGGCTGGAGCTGCTTC	58	Same as above

No	Name	Sequence	No of bases	Remarks
43	TG49-acbP_RR_fwd	GGTCACCACCGGAGCCCGGCGATGACCGGCGCCGTCCGGATTCCGGGGATCCGTCGACC	59	Used for the amplification of the *acbP*-containing cassette for ReDirect
44	TG50-acbP_RR_rev	AGGCGAAGCTCACCACGTGCACGATTCGGCTCCTTCCGGTGTAGGCTGGAGCTGCTTC	58	Same as above
45	TG51-acbQ_RR_fwd	AAACCGAAAGGCAAGCCTGCGATGACCACCACGACGGATATTCCGGGGATCCGTCGACC	59	Used for the amplification of the *acbQ*-containing cassette for ReDirect
46	TG52-acbQ_RR_rev	GACATGGATCTCCTTGAAAGGTGCGGGTGAGCGAGGTGATGTAGGCTGGAGCTGCTTC	58	Same as above
47	TG53-acbR_RR_fwd	CGTCTGGGACGAGGTGACGGCATGAGCACGGGCGTACGGATTCCGGGGATCCGTCGACC	59	Used for the amplification of the *acbR*-containing cassette for ReDirect
48	TG54-acbR_RR_rev	CGGCCCGGACGGCGCCGGTCATCGCCGGGCTCCGGTGGTTGTAGGCTGGAGCTGCTTC	58	Same as above

No	Name	Sequence	No of bases	Remarks
49	TG55-acbS_RR_fwd	CCTACCGGCCGCCGGTCACCGGGGAGAGCTGAGCACATGATTCCGGGGATCCGTCGACC	59	Used for the amplification of the *acbS*-containing cassette for ReDirect
50	TG56-acbS_RR_rev	GCACCGCCCGTACGCCCGTGCTCATGCCGTCACCTCGTCTGTAGGCTGGAGCTGCTTC	58	Same as above
51	TG57-acbU_RR_fwd	GACGGCATGACACCCCGGCCGGTCAGCACGATCGACGTGATTCCGGGGATCCGTCGACC	59	Used for the amplification of the *acbU*-containing cassette for ReDirect
52	TG58-acbU_RR_rev	AAGTACGTCTCGATGATGTGCATGTGCTCAGCTCTCCCCTGTAGGCTGGAGCTGCTTC	58	Same as above
53	ACPL_6141_del1	AGCAGTATGCAGTCTTCCCCGTCGCGTGGCGGCCGGATGATTCCGGGGATCCGTCGACC	59	Used for the amplification of the ACPL_6141 containing cassette for ReDirect
54	ACPL_6141_del2	GACAATTTCCCGTACGCGATGATCACCCCTTCCCGGTCATGTAGGCTGGAGCTGCTTC	58	Same as above

No	Name	Sequence	No of bases	Remarks
55	ACPL6139-40del1	CGAGACAGTGACAACGGGCTTCCGGGAAGGAAGAACATGATTCCGGGGATCCGTCGACC	59	Used for the amplification of the ACPL6139-40 containing cassette for ReDirect
56	ACPL6139-40del2	CGGGGAAGACTGCATACTGCTCGGTACCCCCTACCGTTATGTAGGCTGGAGCTGCTTC	58	Same as above
57	ACPL6142-53del1	TGATTCCCGCATTTTTGGCAGCGCTCTGGAGGAATTGTGATTCCGGGGATCCGTCGACC	59	Used for the amplification of the ACPL6142-53 containing cassette for ReDirect
58	ACPL6142-53del2	CGGAACGGGACGTGCGCGAGAACGGGAGCGGCATACATGTGTAGGCTGGAGCTGCTTC	58	Same as above
59	cACPL4_test1	GGCGCGACGATCCTGATCTC	20	Used to verify ReDirect constructs within cACPL4 cluster
60	cACPL4_test2	CGCGCACCCATGCGGATTAC	20	Same as above
61	TG5-acbD_fwd_6	AAATCTAGACGGAGAAGGTGACGGTGG	27	Used to amplify 6 kbp fragment, which contains *acbD* gene
62	TG6-acbD_rev_6	AAAGAATTCGCTGAGGACCAGGCGAAGAC	29	Same as above

No	Name	Sequence	No of bases	Remarks
63	TG7-acbD_fwd_8	AAATCTAGATCGAGTTTCGCCGGTTTCAC	29	Used to amplify 8 kbp fragment, which contains *acbD* gene
64	TG8-acbD_rev_8	AAAGAATTCTGGGCACCTTCGACATCCTC	29	Same as above
65	TG9-acbDRR_fwd	TGCCCTTCCTCGTCATCCCTTCACAAGGAG AAGCTCGTGATTCCGGGGATCCGTCGACC	59	Used for the amplification of the *acbD*-containing cassette for ReDirect
66	TG10-acbDRR_rev	GGGCCCGGTCACGGCCGGGCCCTCCGCACG GTGGCGTCATGTAGGCTGGAGCTGCTTC	58	Same as above
67	TG11-catRRHg_fwd	GAGTTATCGAGATTTTCAGGAGCTAAGGAA GCTAAAATGCCCGTAGAGATTGGCGATCCC	60	Used to substitute cml^r gene for hyg^r gene in original cosmid library clones via ReDirect
68	TG12-catRRHg_rev	AGGCGTTTAAGGGCACCAATAACTGCCTTA AAAAAATTACAGGCGCCGGGGGCGGTGTC	59	Same as above

No	Name	Sequence	No of bases	Remarks
69	TG15-cat_ver_fwd	TATTCAGGCGTAGCAACCAG	20	Used to verify substitution of cmlr to hygr gene in original cosmid library via ReDirect procedure
70	TG16-cat_ver_rev	TGATCGGCACGTAAGAGGTTC	21	Same as above
71	TG17-acbD_ver_fwd	AAATCTAGATCCATCGTGGACCCTCTCTC	29	Used to verify the correct substitution of *acbD* gene for aac3(IV) via ReDirect
72	TG18-acbD_ver_rev	AAAGAATTCTGGTGGTCTACCTGCTCCTG	29	Same as above
73	T054-cadC_ermE_fwd	CCGGTTGGTAGGATCCAGCGAAGTTAGCCGGTTCGTTCCT	40	Used to generate pKC1139EcadC construct
74	T055-cadC_ermE_rev	CATATGGCGGCCGCCTGCAGCGCCAGGCAGTGCCCATCGA	40	Same as above
75	T056-EE_fwd	CTGCAGGCGGCCGCCATATGCATCCTAGGC	30	Used to generate pKC1139EcadC construct

No	Name	Sequence	No of bases	Remarks
76	T057-EE_rev	CGCTGGATCCTACCAACCGGCACGATTGTC	30	Same as above
77	T40-rpoB_fwd	GACCAGCGTCAAGGTCATCC	20	Used for qRT PCR
78	T41-rpoB_rev	GACCAACCGATCGCCTTCAG	20	Same as above
79	T92-pSET152_hyg_GOI_fwd	TCGTGCTTTCAGCTTCGATG	20	Used for PCR reactions to examine if the insert was successfully cloned into the pSET152hyg and pIJ6902 vector.
80	T93-pSET152_hyg_GOI_rev	AGGTTTCCCGACTGGAAAGC	20	Same as above
81	KT57-qRT-acbV fwd	ACCAGCAGTACGCCGTTGTC	20	Used for qRT PCR
82	KT57-qRT-acbV rev	ACATCGCCGAGAAACCCAAC	20	Same as above
83	T209-qRT_acbD_fwd	GCCAGCAACATCAAAGTG	18	Used for qRT PCR
84	T210-qRT_acbD_rev	TGGTAGCGATCGTTGAAG	18	Same as above

No	Name	Sequence	No of bases	Remarks
85	7-7449_fwd	CGAGCAGCTCTACCTGGTGATGC	23	Used for qRT-PCR
86	15-7449_rev	GCTTGTTCTTGAAGATCGGCGTCAG	25	Same as above
87	acbD_VO_fwd	AACCCGGCGAACAACAATCC	20	Used to amplify intenal fragment of *acbD* gene
88	acbD_VO_rev	CAAGTTCGTCAAGGTCGCGG	20	Same as above
89	ACPL_212_RR_fwd	CACCCATGCAGGAGAATCGATGCAGTTGGT GGAAGGATGATTCCGGGGATCCGTCGACC	59	Used for the amplification of the ACPL_212 containing cassette for ReDirect
90	ACPL_212_RR_rev	CCCCACGTTCCAGTGACGCCAGGCAGTGCC CATCGATCATGTAGGCTGGAGCTGCTTC	58	Same as above
91	TG85-5091_RR_fwd	TGCAATCAGCGGTGAGGATCATCAAAGGGG ACTGTCATGATTCCGGGGATCCGTCGACC	59	Used for the amplification of the ACPL 5091 containing cassette for ReDirect
92	TG86-5091_RR_rev	TATCTGAGCCATATCCCTCGGCGCGACCAC TGCGGATCATGTAGGCTGGAGCTGCTTC	58	Same as above

No	Name	Sequence	No of bases	Remarks
93	TG91-6402_RR_fwd	TTAACGCGCTAATCCGTATCTGACGGGAGTATTTTCATGATTCCGGGGATCCGTCGACC	59	Used for the amplification of the ACPL 6402 containing cassette for ReDirect
94	TG92-6402_RR_rev	GCCCTTCGAGCTGCGTCGTCATTACTTGGTGATCGCGGCTGTAGGCTGGAGCTGCTTC	58	Same as above
95	TG93-6403_RR_fwd	ACGGCTCGGCACTACAGTGACAACATGCGTGCACGGATGATTCCGGGGATCCTGCGACC	59	Used for the amplification of the ACPL 6403 containing cassette for ReDirect
96	TG94-6403_RR_rev	GCAAGAAGTTGCAGCAAGTTTCCCGGGCGGTGCGGGTCATGTAGGCTGGAGCTGCTTC	58	Same as above

Table 12: List of all primers used in this study

No	Name of construct	Primers, used for construction	Comments
1	pKGLP2acbD	acbD_VO_fwd and acbD_VO_rev	Vector pKGLP2 which contains an internal 1.9 fragment of *acbD* gene
2	pKC1139hygacbD	TG7-acbD_fwd_8 and TG8-acbD_rev_8	Vector pKC1139hyg, that contains 8 kbp region with *acbD* gene
3	pKC1139hygacbDaac3(IV)	TG9-acbDRR_fwd and TG10-acbDRR_rev were used to perform ReDiret on pKC1139hygacbD.	pKC1139hygacbD with *acbD* gene being substituted for aac3(IV) with the help of ReDirect procedure
4	h06hyg	TG11-catRRHg_fwd and TG12-catRRHg_rev were used to perform ReDiret on h06; TG15-cat_ver_fwd and TG16-cat_ver_rev were used to verify this construct.	h06hyg with cml^r gene being substituted for hyg^r gene via ReDirect procedure
5	h06hygaac(3)IV	TG9-acbDRR_fwd and TG10-acbDRR_rev were used to perform ReDiret on h06hyg. TG17-acbD_ver_fwd and TG18-acbD_ver_rev were used to verify this construct	h06hyg with *acbD* gene being substituted for aac3(IV) with the help of ReDirect procedure
6	h06hygaac(3)IVacbA	TG27-acbARR_fwd and TG28-acbARR_rev were used to perform ReDiret on h06hyg.	h06hyg with acbA gene being substituted for aac3(IV) with the help of ReDirect procedure

No	Name of construct	Primers, used for construction	Comments
7	h06hygaac(3)IVacbB	TG29-acbBRR_fwd and TG30-acbBRR_rev were used to perform ReDiret on h06hyg.	h06hyg with acbB gene being substituted for aac3(IV) with the help of ReDirect procedure
8	h06hygaac(3)IVacbC	TG31-acbCRR_fwd and TG32-acbCRR_rev were used to perform ReDiret on h06hyg.	h06hyg with acbC gene being substituted for aac3(IV) with the help of ReDirect procedure
9	h06hygaac(3)IVacbE	TG33-acbERR_fwd and TG34-acbERR_rev were used to perform ReDiret on h06hyg.	h06hyg with acbE gene being substituted for aac3(IV) with the help of ReDirect procedure
10	h06hygaac(3)IVacbI	TG35-acbIRR_fwd and TG36-acbIRR_rev were used to perform ReDiret on h06hyg.	h06hyg with acbI gene being substituted for aac3(IV) with the help of ReDirect procedure
11	h06hygaac(3)IVacbJ	TG37-acbJRR_fwd and TG38-acbJRR_rev were used to perform ReDiret on h06hyg.	h06hyg with acbJ gene being substituted for aac3(IV) with the help of ReDirect procedure
12	h06hygaac(3)IVacbK	TG39-acbKRR_fwd and TG40-acbKRR_rev were used to perform ReDiret on h06hyg.	h06hyg with acbK gene being substituted for aac3(IV) with the help of ReDirect procedure

No	Name of construct	Primers, used for construction	Comments
13	h06hygaac(3)IVacbL	TG41-acbLRR_fwd and TG42-acbLRR_rev were used to perform ReDiret on h06hyg.	h06hyg with acbL gene being substituted for aac3(IV) with the help of ReDirect procedure
14	h06hygaac(3)IVacbM	TG43-acbMRR_fwd and TG44-acbMRR_rev were used to perform ReDiret on h06hyg.	h06hyg with acbM gene being substituted for aac3(IV) with the help of ReDirect procedure
15	h06hygaac(3)IVacbO	TG47-acbORR_fwd and TG48-acbORR_rev were used to perform ReDiret on h06hyg.	h06hyg with acbO gene being substituted for aac3(IV) with the help of ReDirect procedure
16	h06hygaac(3)IVacbP	TG49-acbPRR_fwd and TG50-acbPRR_rev were used to perform ReDiret on h06hyg.	h06hyg with acbP gene being substituted for aac3(IV) with the help of ReDirect procedure
17	h06hygaac(3)IVacbR	TG53-acbRRR_fwd and TG54-acbRRR_rev were used to perform ReDiret on h06hyg.	h06hyg with acbR gene being substituted for aac3(IV) with the help of ReDirect procedure
18	h06hygaac(3)IVacbQ	TG51-acbQRR_fwd and TG52-acbQRR_rev were used to perform ReDiret on h06hyg.	h06hyg with acbQ gene being substituted for aac3(IV) with the help of ReDirect procedure

No	Name of construct	Primers, used for construction	Comments
19	h06hygaac(3)IVacbS	TG55-acbSRR_fwd and TG56-acbSRR_rev were used to perform ReDiret on h06hyg.	h06hyg with acbS gene being substituted for aac3(IV) with the help of ReDirect procedure
20	h06hygaac(3)IVacbU	TG57-acbURR_fwd and TG58-acbURR_rev were used to perform ReDiret on h06hyg.	h06hyg with acbU gene being substituted for aac3(IV) with the help of ReDirect procedure
21	h12hyg	TG11-catRRHg_fwd and TG12-catRRHg_rev were used to perform ReDiret on h12; TG15-cat_ver_fwd and TG16-cat_ver_rev were used to verify this construct.	h12 with cml^r gene being substituted for hyg^r gene via ReDirect procedure
22	h12hygaac(3)IV	ACPL_212_RR_fwd and ACPL_212_RR_rev were used to perform ReDirect on h12hyg.	h12hyg with cadC gene being substituted for aac3(IV) with the help of ReDirect procedure
23	b02hyg	TG11-catRRHg_fwd and TG12-catRRHg_rev were used to perform ReDiret on b02; TG15-cat_ver_fwd and TG16-cat_ver_rev were used to verify this construct.	b02 with cml^r gene being substituted for hyg^r gene via ReDirect procedure

No	Name of construct	Primers, used for construction	Comments
24	b02hygaac(3)IV6141	ACPL_6141_del1 and ACPL_6141_del2 were used to perform ReDirect on b02hyg. cACPL4_test1 and cACPL4_test2 were used to verify this construct	b02hyg with ACPL_6141 gene being substituted for aac3(IV) with the help of ReDirect procedure
25	b02hygaac(3)IV 6142-53	ACPL6142-53del1 and ACPL6142-53del2 were used to perform ReDirect on b02hyg. cACPL4_test1 and cACPL4_test2 were used to verify this construct	b02hyg with ACPL_6142-53 genes being substituted for aac3(IV) with the help of ReDirect procedure
26	b02hygaac(3)IV 6139-40	ACPL6139-40del1 and ACPL6139-40del2 were used to perform ReDirect on b02hyg. cACPL4_test1 and cACPL4_test2 were used to verify this construct	b02hyg with ACPL_6139-40 genes being substituted for aac3(IV) with the help of ReDirect procedure
27	h10hyg	TG11-catRRHg_fwd and TG12-catRRHg_rev were used to perform ReDiret on h10; TG15-cat_ver_fwd and TG16-cat_ver_rev were used to verify this construct.	h10h with cml^r gene being substituted for hyg^r gene via ReDirect procedure

No	Name of construct	Primers, used for construction	Comments
28	h10hygaac(3)IVmalR	TG93-6403_RR_fwd and TG94-6403_RR_rev were used to perform ReDirect on h10hyg	h10hyg with malR gene being substituted for aac3(IV) with the help of ReDirect procedure
29	h10hygaac(3)IVmalE	TG91-6402_RR_fwd and TG92-6402_RR_rev were used to perform ReDirect on h10hyg	h10hyg with malE gene being substituted for aac3(IV) with the help of ReDirect procedure
30	a01hyg	TG11-catRRHg_fwd and TG12-catRRHg_rev were used to perform ReDiret on a01; TG15-cat_ver_fwd and TG16-cat_ver_rev were used to verify this construct.	a01 with cml^r gene being substituted for hyg^r gene via ReDirect procedure
31	a01hygaac(3)IVcgt	TG85-5091_RR_fwd and TG86-5091_RR_rev were used to perform ReDirect on h10hyg	a01hyg with *cgt* gene being substituted for aac3(IV) with the help of ReDirect procedure
32	pSET152hygacbD	T008-acbD-com-fwd and T009-acbD-com-rev were used to synthesize an insert	pSET152hyg with a fragment, which contains a wild type copy of *acbD* gene
33	pSET152cadCA	T022-cadC_com_fwd_1 and T023-cadC_com_rev_1 are used to generate insert, which contains *cadC*	pSET152hyg, which contains a copy of *cadC* gene

No	Name of construct	Primers, used for construction	Comments
34	pSET152cadCB	T024-cadC_com_fwd_2 and T025-cadC_com_rev_2 are used to generate insert, which contains ACPL_8935 and *cadC*	pSET152hyg, which contains a copy of *cadC* gene
35	pIJ6902cadC	T026-212_cf_tipA and T027-212_cr_tipA used to create an insert; T028-6902_212_f and T029-6902_212_r are used to szntesize vector	pIJ6902, which contains *cadC* gene under *tipA* promoter
36	pKC1139EcadC	T054-cadC_ermE_fwd and T055-cadC_ermE_rev are used to create insert; T056-EE_fwd and T057-EE_rev are used to synthesize the vector	pKC1139, which contains *cadC* under *ermE* promoter

Table 13: Description of all DNA constructs, created in the course of this work

Number	Full cosmid name	Gene/ region of interest	Genomic region, nt	Status (in library)	*E. coli* BW knockout mutants	*Actinoplanes* knockout mutants
1	aut_f2p0002g07.y	*acb* cluster	4018677-4057082	Viable	-	-
2	aut_f2p0002h06.x	*acb* cluster	4061812-4095358	Viable	Δ*acbU*, Δ*acbS*, Δ*acbR*, Δ*acbP*, Δ*acbI*, Δ*acbJ*, Δ*acbK*, Δ*acbM*, Δ*acbL*, Δ*acbO*, Δ*acbC*, Δ*acbB*, Δ*acbA*, Δ*acbE*, Δ*acbD*	Δ*acbA*, Δ*acbB*, Δ*acbD*
3	aut_f2p0002f07.x	*acb* cluster	4066802-4106011	Viable	-	-
4	aut_f1p0002c04.x	cACPL4	-	Viable	-	-
5	aut_f1p0002f08.y	cACPL4	-	Viable	-	-
6	aut_f1p0002d04.y	cACPL4	6716292-6758434	Viable	-	-

Number	Full cosmid name	Gene/ region of interest	Genomic region, nt	Status (in library)	*E. coli* BW knockout mutants	*Actinoplanes* knockout mutants
7	aut_f1p0002b02.y	cACPL4	6729418-6773292	Viable	Δ*ACPL*_6141 (dtxR) Δ*ACPL*6142-53 Δ*ACPL*6139-40	Δ*ACPL*_6141 (dtxR) Δ*ACPL*6142-53 Δ*ACPL*6139-40
8	aut_f1p0005g11.y	*cadC* (ACPL_212)	-	Not viable	-	-
9	aut_f1p0005h12.x	*cadC* (ACPL_212)	-	Viable	Δ*cadC*	Δ*cadC*
10	aut_f2p0002e06.y	*cadC* (ACPL_212)	-	Not viable	-	-
11	aut_f1p0002g04.x	*cgt* (ACPL_5091)	-	Viable	-	-
12	aut_f1p0005a01.y	*cgt*(ACPL_5091)	-	Viable	Δ*cgt*	-
13	aut_f1p0002h10.x	*malE* (ACPL_6402), *malR* (ACPL_6403)	-	Viable	Δ*malE*, Δ*malR*	Δ*malR*

Number	Full cosmid name	Gene/ region of interest	Genomic region, nt	Status (in library)	*E. coli* BW knockout mutants	*Actinoplanes* knockout mutants
14	aut_f1p0001a01.y	*malR* (ACPL_6403)	-	Not viable	-	-
15	aut_f1p0001f10.x	*malR* (ACPL_6403)	-	Not viable	-	-
16	aut_f1p0000g02.y	pACPL4	-	Not viable	-	-
17	aut_f2p0003d07.y	pACPL4	6774485-6808718	Not viable	-	-
18	aut_f1p0003f12.x	*tetR* (ACPL_1950)	2047193-2090740	Viable	-	-
19	aut_f1p0004.a07.x	*tetR* (ACPL_6747)	7405923-7439453	Not viable	-	-

Table 14: List of all genetic constructs, generated through ReDirect procedure in this study

Plate type	Mode of action	Substance	Difference
PM01	C-Source, amino acid	L-Proline	5933,58
PM01	C-Source, carbohydrate	D-Mannitol	5774,42
PM01	C-Source, carbohydrate	a-D-Glucose	6779,42
PM01	C-Source, carbohydrate	Maltose	5990,42
PM01	C-Source, carbohydrate	Maltotriose	6973,08
PM01	C-Source, carbohydrate	D-Cellobiose	5972,5
PM02	C-Source, carbohydrate	L-Arabitol	5693,11
PM02	C-Source, carbohydrate	Arbutin	6854,67
PM02	C-Source, carbohydrate	Gentiobiose	6422,33
PM02	C-Source, carbohydrate	D-Melezitose	11115,22
PM02	C-Source, carbohydrate	Palatinose	7146,22
PM02	C-Source, carbohydrate	D-Raffinose	12514,33
PM02	C-Source, carbohydrate	Salicin	5992,28
PM02	C-Source, carbohydrate	L-Sorbose	5136,78
PM02	C-Source, carbohydrate	Turanose	5695,61
PM02	C-Source, carbohydrate	Xylitol	5377,72
PM02	C-Source, carboxylic acid	L-Tartaric Acid	5217,06
PM02	C-Source, polymer	Dextrin	9175,83
PM02	C-Source, polymer	Glycogen	6032,61
PM03	N-Source, other	Guanine	5688,92
PM05	Nutritional supplement	a-Ketobutyric Acid	13818,67
PM05	Nutritional supplement	Tween 40	12935
PM05	Nutritional supplement	Tween 80	18111,92
PM04	P-Source, organic	Cytidine 3‘,5‘-Cyclic Monophosphate	7503,58

Table 15: Analysis of $\Delta acbD$ strain phenotype with the help of phenotypic microarrays of BioLog system. Phenotypes Gained - Faster Growth /Resistance

Plate type	Mode of action	Substance	Difference
PM01	C-Source, amino acid	Ala-Gly	-12134,5
PM02	C-Source, amino acid	L-Arginine	-6937,83
PM02	C-Source, amino acid	L-Ornithine	-5553,5
PM01	C-Source, carbohydrate	Thymidine	-5508,08
PM01	C-Source, carboxylic acid	Acetic Acid	-14552,25
PM01	C-Source, carboxylic acid	a-Ketobutyric Acid	-8057,08
PM02	C-Source, carboxylic acid	g-Amino-N-Butyric Acid	-5359,5
PM01	C-Source, ester	Methylpyruvate	-18282,5
PM02	C-Source, ester	D-Lactic Acid Methyl Ester	-5252,89
PM03	N-Source, amino acid	L-Cysteine	-17690
PM03	N-Source, amino acid	N-Phthaloyl-L-Glutamic Acid	-6044,75
PM03	N-Source, other	Guanosine	-8619,25
PM03	N-Source, other	Xanthine	-6068,42
PM03	N-Source, other	Alloxan	-6368,83
PM03	N-Source, other	D,L-a-Amino-Caprylic Acid	-18698,17
PM04	S-Source, organic	L-Djenkolic Acid	-6865,83

Table 16: Analysis of $\Delta acbD$ strain phenotype with the help of phenotypic microarrays of BioLog system. Phenotypes Lost - Slower Growth / Sensitivity

LIST OF REFERENCES

LIST OF REFERENCES

[1] StrepDB - The *Streptomyces* annotation server. http://strepdb.streptomyces.org.uk/ accessed on 17.02.2017.

[2] K. G. Alberti and P. Z. Zimmet. Definition, diagnosis and classification of diabetes mellitus and its complications. Part 1: diagnosis and classification of diabetes mellitus provisional report of a WHO consultation. *Diabet. Med.*, 15(7):539–553, Jul 1998.

[3] K. G. M. Alberti, P. Zimmet, and J. Shaw. International diabetes federation: a consensus on type 2 diabetes prevention. *Diabetic Medicine*, 24(5):451–463, 2007.

[4] S. Alt, N. Burkard, A. Kulik, S. Grond, and L. Heide. An artificial pathway to 3, 4-dihydroxybenzoic acid allows generation of new aminocoumarin antibiotic recognized by catechol transporters of *E. coli*. *Chemistry & biology*, 18(3):304–313, 2011.

[5] S. Andrews. FastQC: A quality control tool for high throughput sequence data. *Babraham Bioinformatics*, page 1, 2010.

[6] S. Andrews et al. Fastqc: A quality control tool for high throughput sequence data. *Reference Source*, 2010.

[7] A. D. Association et al. 2. classification and diagnosis of diabetes. *Diabetes care*, 38(Supplement 1):S8–S16, 2015.

[8] R. M. Atlas. *Handbook of microbiological media*. CRC press, 2010.

[9] R. H. Baltz. Genetic manipulation of antibiotic producing *Streptomyces*. *Trends in microbiology*, 6(2):76–83, 1998.

[10] R. H. Baltz. *Streptomyces* temperate bacteriophage integration systems for stable genetic engineering of actinomycetes (and other organisms). *Journal of Industrial Microbiology and Biotechnology*, 39(5):661–672, 2012.

[11] S. D. Bentley, K. F. Chater, A. M. Cerdeno-Tarraga, G. L. Challis, N. R. Thomson, K. D. James, D. E. Harris, M. A. Quail, H. Kieser, D. Harper, A. Bateman, S. Brown, G. Chandra, C. W. Chen, M. Collins, A. Cronin, A. Fraser, A. Goble, J. Hidalgo, T. Hornsby, S. Howarth, C. H. Huang, T. Kieser, L. Larke, L. Murphy, K. Oliver, S. O'Neil, E. Rabbinowitsch, M. A. Rajandream, K. Rutherford, S. Rutter, K. Seeger, D. Saunders, S. Sharp, R. Squares, S. Squares, K. Taylor, T. Warren, A. Wietzorrek, J. Woodward, B. G. Barrell, J. Parkhill, and D. A. Hopwood. Complete genome sequence of the model actinomycete *Streptomyces coelicolor* A3(2). *Nature*, 417(6885):141–147, May 2002.

[12] M. Bierman, R. Logan, K. O'Brien, E. T. Seno, R. N. Rao, and B. E. Schoner. Plasmid cloning vectors for the conjugal transfer of DNA from *Escherichia coli* to *Streptomyces* spp. *Gene*, 116(1):43–49, jul 1992.

[13] B. Bilyk and A. Luzhetskyy. Unusual site-specific DNA integration into the highly active pseudo *attB* of the *Streptomyces albus* J1074 genome. *Applied microbiology and biotechnology*, 98(11):5095–5104, 2014.

[14] O. Bilyk and A. Luzhetskyy. Metabolic engineering of natural product biosynthesis in Actinobacteria. *Curr. Opin. Biotechnol.*, 42:98–107, Apr 2016.

[15] I. Borodina, J. Siebring, J. Zhang, C. P. Smith, G. van Keulen, L. Dijkhuizen, and J. Nielsen. Antibiotic overproduction in *Streptomyces coelicolor* A3(2) mediated by phosphofructokinase deletion. *J. Biol. Chem.*, 283(37):25186–25199, Sep 2008.

[16] C. Brunkhorst, C. Andersen, and E. Schneider. Acarbose, a pseudooligosaccharide, is transported but not metabolized by the maltose-maltodextrin system of *Escherichia coli*. *J. Bacteriol.*, 181(8):2612–2619, Apr 1999.

[17] L. S. Busenlehner, M. A. Pennella, and D. P. Giedroc. The Smtb/ArsR family of metalloregulatory transcriptional repressors: structural insights into prokaryotic metal resistance. *FEMS microbiology reviews*, 27(2-3):131–143, 2003.

[18] W. Caspary and S. Graf. Inhibition of human intestinal α-glucosidehydrolases by a new complex oligosaccharide. *Research in Experimental Medicine*, 175(1):1–6, 1979.

[19] R. E. Cobb, Y. Wang, and H. Zhao. High-efficiency multiplex genome editing of *Streptomyces* species using an engineered CRISPR/Cas system. *ACS synthetic biology*, 4(6):723–728, 2014.

[20] P. Combes, R. Till, S. Bee, and M. C. M. Smith. The *Streptomyces* genome contains multiple pseudo-*attB* sites for the φC31-encoded site-specific recombination system. *Journal of bacteriology*, 184(20):5746–52, 2002.

[21] J. N. Couch. *Actinoplanes*, a new genus of the Actinomycetales. *J Elisha Mitchell Sci Soc*, 66:87–92, 1950.

[22] J. N. Couch. Some new genera and species of the Actinoplanaceae. *J. Elisha Mitchell Sci. Soc*, 79(53-70):3, 1963.

[23] A. L. Demain. Importance of microbial natural products and the need to revitalize their discovery. *Journal of industrial microbiology & biotechnology*, 41(2):185–201, 2014.

[24] A. Dereeper, V. Guignon, G. Blanc, S. Audic, S. Buffet, F. Chevenet, J.-F. Dufayard, S. Guindon, V. Lefort, M. Lescot, et al. Phylogeny. fr: robust phylogenetic analysis for the non-specialist. *Nucleic acids research*, 36(suppl 2):W465–W469, 2008.

[25] J. J. DiNicolantonio, J. Bhutani, and J. H. O'Keefe. Acarbose: safe and effective for lowering postprandial hyperglycaemia and improving cardiovascular outcomes. *Open heart*, 2(1):e000327, 2015.

[26] M. Dondrup, S. P. Albaum, T. Griebel, K. Henckel, S. Jünemann, T. Kahlke, C. K. Kleindt, H. Küster, D. Linke, Burkhard Mertens and, V. Mittard-Runte, H. Neuweger, K. J. Runte, A. Tauch, F. Tille, A. Pühler, and A. Goesmann. EMMA 2 – a MAGE-compliant system for the collaborative analysis and integration of microarray data. *BMC Bioinformatics*, 10(1):50, 2009.

[27] J. Droste. Charakterisierung des Transkriptionsregulators MalR in *Actinoplanes* sp. SE50/110, 2015.

[28] F. Eikmeyer, A. Hadiati, R. Szczepanowski, D. Wibberg, S. Schneiker-Bekel, L. M. Rogers, C. J. Brown, E. M. Top, A. Puhler, and A. Schluter. The complete genome sequences of four new IncN plasmids from wastewater treatment plant effluent provide new insights into IncN plasmid diversity and evolution. *Plasmid*, 68(1):13–24, Jul 2012.

[29] M. Fedoryshyn, L. Petzke, E. Welle, A. Bechthold, and A. Luzhetskyy. Marker removal from actinomycetes genome using Flp recombinase. *Gene*, 419(1-2):43–47, Aug 2008.

[30] M. Fedoryshyn, E. Welle, A. Bechthold, and A. Luzhetskyy. Functional expression of the Cre recombinase in actinomycetes. *Appl. Microbiol. Biotechnol.*, 78(6):1065–1070, Apr 2008.

[31] R. D. Finn, P. Coggill, R. Y. Eberhardt, S. R. Eddy, J. Mistry, A. L. Mitchell, S. C. Potter, M. Punta, M. Qureshi, A. Sangrador-Vegas, et al. The Pfam protein families database: towards a more sustainable future. *Nucleic acids research*, 44(D1):D279–D285, 2016.

[32] K. Fowler-Goldsworthy, B. Gust, S. Mouz, G. Chandra, K. C. Findlay, and K. F. Chater. The Actinobacteria-specific gene *wblA* controls major developmental transitions in *Streptomyces coelicolor* A3(2). *Microbiology*, 157(5):1312–1328, 2011.

[33] W. Frommer, B. Junge, L. Mueller, D. Schmidt, and E. Truscheit. New enzyme inhibitors from microorganisms [drug plants]. *Planta Medica*, 1979.

[34] W. Frommer, W. Puls, D. Schafer, and D. Schmidt. Glycoside-hydrolase enzyme inhibitors, Apr. 8 1975. US Patent 3,876,766.

[35] L. Gaigalat, J.-P. Schlüter, M. Hartmann, S. Mormann, A. Tauch, A. Pühler, and J. Kalinowski. The DeoR-type transcriptional regulator SugR acts as a repressor for genes encoding the phosphoenolpyruvate: sugar phosphotransferase system (PTS) in *Corynebacterium glutamicum*. *BMC molecular biology*, 8(1):104, 2007.

[36] U. Ghani. Re-exploring promising α-glucosidase inhibitors for potential development into oral anti-diabetic drugs: Finding needle in the haystack. *European journal of medicinal chemistry*, 103:133–162, 2015.

[37] D. G. Gibson, L. Young, R.-Y. Chuang, J. C. Venter, C. A. Hutchison, and H. O. Smith. Enzymatic assembly of DNA molecules up to several hundred kilobases. *Nature methods*, 6(5):343–345, 2009.

[38] J. P. Gomez-Escribano and M. J. Bibb. Engineering *Streptomyces coelicolor* for heterologous expression of secondary metabolite gene clusters. *Microb Biotechnol*, 4(2):207–215, Mar 2011.

[39] J. P. Gomez-Escribano and M. J. Bibb. *Streptomyces coelicolor* as an expression host for heterologous gene clusters. *Meth. Enzymol.*, 517:279–300, 2012.

[40] N. Gonzalez-Quiñonez, M. T. López-García, P. Yagüe, B. Rioseras, A. Pisciotta, R. Alduina, and Á. Manteca. New ϕbt1 site-specific integrative vectors with neutral phenotype in *Streptomyces*. *Applied microbiology and biotechnology*, 100(6):2797–2808, 2016.

[41] M. Goodfellow, L. Stanton, K. Simpson, and D. Minnikin. Numerical and chemical classification of *Actinoplanes* and some related actinomycetes. *Microbiology*, 136(1):19–36, 1990.

[42] M. A. Gregory, R. Till, and M. C. M. Smith. Integration site for *Streptomyces* phage BT1 and development of site-specific integrating vectors. *Journal of Bacteriology*, 185(17):5320–5323, sep 2003.

[43] T. Gren, V. Ortseifen, D. Wibberg, S. Schneiker-Bekel, H. Bednarz, K. Niehaus, T. Zemke, M. Persicke, A. Puhler, and J. Kalinowski. Genetic engineering in *Actinoplanes* sp. SE50/110 - development of an intergeneric conjugation system for the introduction of actinophage-based integrative vectors. *J. Biotechnol.*, 232:79–88, Aug 2016.

[44] B. Gust. Cloning and analysis of natural product pathways. *Methods in enzymology*, 458:159–180, 2009.

[45] B. Gust, G. Chandra, D. Jakimowicz, T. Yuqing, C. J. Bruton, and K. F. Chater. λ red-mediated genetic manipulation of antibiotic-producing *Streptomyces*. *Advances in applied microbiology*, 54:107–128, 2004.

[46] B. Gust, T. Kieser, and K. Chater. REDIRECT technology: PCR-targeting system in *Streptomyces coelicolor*. *The John Innes Centre, Norwich, United Kingdom*, 2002.

[47] B. Gust, S. O'Rourke, N. Bird, T. Kieser, and K. Chater. Recombineering in *Streptomyces coelicolor*. *Norwich: The John Innes Foundation*, 2003.

[48] H.-S. Ha, Y.-I. Hwang, and S.-U. Choi. Application of conjugation using phiC31 *att/int* system for *Actinoplanes teichomyceticus*, a producer of teicoplanin. *Biotechnology letters*, 30(7):1233–8, jul 2008.

[49] S. Heinl, D. Wibberg, F. Eikmeyer, R. Szczepanowski, J. Blom, B. Linke, A. Goesmann, R. Grabherr, H. Schwab, A. Puhler, and A. Schluter. Insights into the completely annotated genome of *Lactobacillus buchneri* CD034, a strain isolated from stable grass silage. *J. Biotechnol.*, 161(2):153–166, Oct 2012.

[50] E. Heinzelmann, S. Berger, O. Puk, B. Reichenstein, W. Wohlleben, and D. Schwartz. A glutamate mutase is involved in the biosynthesis of the lipopeptide antibiotic friulimicin in *Actinoplanes friuliensis*. *Antimicrob. Agents Chemother.*, 47(2):447–457, Feb 2003.

[51] M. Hemker, A. Stratmann, K. Goeke, W. Schröder, J. Lenz, W. Piepersberg, and H. Pape. Identification, cloning, expression, and characterization of the extracellular acarbose-modifying glycosyltransferase, AcbD, from *Actinoplanes* sp. strain SE50. *Journal of bacteriology*, 183(15):4484–4492, aug 2001.

[52] S. Herrmann, T. Siegl, M. Luzhetska, L. Petzke, C. Jilg, E. Welle, A. Erb, P. F. Leadlay, A. Bechthold, and A. Luzhetskyy. Site-specific recombination strategies for engineering actinomycete genomes. *Applied and environmental microbiology*, 78(6):1804–1812, 2012.

[53] R. Hilker, K. B. Stadermann, D. Doppmeier, J. Kalinowski, J. Stoye, J. Straube, J. Winnebald, and A. Goesmann. ReadXplorer–visualization and analysis of mapped sequences. *Bioinformatics*, 30(16):2247–2254, aug 2014.

[54] F. Hille and E. Charpentier. CRISPR-Cas: biology, mechanisms and relevance. *Phil. Trans. R. Soc. B*, 371(1707):20150496, 2016.

[55] D. Holmes, J. Caso, and C. Thompson. Autogenous transcriptional activation of a thiostrepton-induced gene in *Streptomyces lividans*. *The EMBO journal*, 12(8):3183, 1993.

[56] A. J. Hoogewerf, L. A. V. Dyk, T. S. Buit, D. Roukema, E. Resseguie, C. Plaisier, N. Le, L. Heeringa, and D. A. V. Griend. Functional characterization of a cadmium resistance operon in *Staphylococcus aureus* ATCC12600: CadC does not function as a repressor. *Journal of basic microbiology*, 55(2):148–159, 2015.

[57] D. A. Hopwood. *Streptomyces in nature and medicine: the antibiotic makers*. Oxford University Press, 2007.

[58] L. Horbal, A. Kobylyanskyy, O. Yushchuk, N. Zaburannyi, A. Luzhetskyy, B. Ostash, F. Marinelli, and V. Fedorenko. Evaluation of heterologous promoters for genetic analysis of *Actinoplanes teichomyceticus* producer of teicoplanin, drug of last defense. *Journal of Biotechnology*, 168(4):367–372, 2013.

[59] L. Horbal, N. Zaburannyy, B. Ostash, S. Shulga, and V. Fedorenko. Manipulating the regulatory genes for teicoplanin production in *Actinoplanes teichomyceticus*. *World Journal of Microbiology & Biotechnology*, 28(5):2095–2100, 2012.

[60] R. Hu, Y. Li, Q. Lv, T. Wu, and N. Tong. Acarbose monotherapy and type 2 diabetes prevention in eastern and western prediabetes: an ethnicity-specific meta-analysis. *Clinical therapeutics*, 37(8):1798–1812, 2015.

[61] J. Huang, J. Shi, V. Molle, B. Sohlberg, D. Weaver, M. J. Bibb, N. Karoonuthaisiri, C. J. Lih, C. M. Kao, M. J. Buttner, and S. N. Cohen. Cross-regulation among disparate antibiotic biosynthetic pathways of *Streptomyces coelicolor*. *Mol. Microbiol.*, 58(5):1276–1287, Dec 2005.

[62] S. Hunter, R. Apweiler, T. K. Attwood, A. Bairoch, A. Bateman, D. Binns, P. Bork, U. Das, L. Daugherty, L. Duquenne, et al. InterPro: the integrative protein signature database. *Nucleic acids research*, 37(suppl 1):D211–D215, 2009.

[63] H. Hutter. Fluorescent reporter methods. *C. elegans: Methods and Applications*, pages 155–173, 2006.

[64] M.-S. Jang, A. Fujita, S. Ikawa, K. Hanawa, H. Yamamura, T. Tamura, M. Hayakawa, T. Tezuka, and Y. Ohnishi. Isolation of a novel plasmid from *Couchioplanes caeruleus* and construction of two plasmid vectors for gene expression in *Actinoplanes missouriensis*. *Plasmid*, 77:32–38, 2015.

[65] L. Katz and R. H. Baltz. Natural product discovery: past, present, and future. *Journal of industrial microbiology & biotechnology*, 43(2-3):155–176, 2016.

[66] M. Kearse, R. Moir, A. Wilson, S. Stones-Havas, M. Cheung, S. Sturrock, S. Buxton, A. Cooper, S. Markowitz, C. Duran, T. Thierer, B. Ashton, P. Meintjes, and A. Drummond. Geneious Basic: an integrated and extendable desktop software platform for the organization and analysis of sequence data. *Bioinformatics*, 28(12):1647–1649, Jun 2012.

[67] L. A. Kelley, S. Mezulis, C. M. Yates, M. N. Wass, and M. J. Sternberg. The Phyre2 web portal for protein modeling, prediction and analysis. *Nature protocols*, 10(6):845–858, 2015.

[68] T. Kieser, J. Bibb, M. Buttner, K. Chater, and D. Hopwood. Practical *Streptomyces* genetics, The John Innes Foundation, Norwich, UK. 2000.

[69] M. Kim, S. Park, and S.-J. Lee. Global transcriptional regulator TrmB family members in prokaryotes. *Journal of Microbiology*, 54(10):639–645, 2016.

[70] S. B. Kim and M. Goodfellow. *Streptomyces thermospinisporus* sp. nov., a moderately thermophilic carboxydotrophic streptomycete isolated from soil. *International journal of systematic and evolutionary microbiology*, 52(4):1225–1228, 2002.

[71] F. E. Koehn and G. T. Carter. The evolving role of natural products in drug discovery. *Nature reviews Drug discovery*, 4(3):206–220, 2005.

[72] M. Komatsu, K. Komatsu, H. Koiwai, Y. Yamada, I. Kozone, M. Izumikawa, J. Hashimoto, M. Takagi, S. Omura, K. Shin-ya, et al. Engineered *Streptomyces avermitilis* host for heterologous expression of biosynthetic gene cluster for secondary metabolites. *ACS synthetic biology*, 2(7):384–396, 2013.

[73] A. Köpfer. Analyse der transkriptionellen Regulation des *acb*-Genclusters in *Actinoplanes* sp. SE50/110 mittels Regulator Fishing, 2012.

[74] B. Langmead and S. L. Salzberg. Fast gapped-read alignment with bowtie 2. *Nature methods*, 9(4):357–359, 2012.

[75] H. Lechevalier, M. P. Lechevalier, and N. N. Gerber. Chemical composition as a criterion in the classification of actinomycetes. *Advances in applied microbiology*, 14:47–72, 1971.

[76] A. Licht, H. Bulut, F. Scheffel, O. Daumke, U. F. Wehmeier, W. Saenger, E. Schneider, and A. Vahedi-Faridi. Crystal structures of the bacterial solute receptor AcbH displaying an exclusive substrate preference for β-D-galactopyranose. *Journal of molecular biology*, 406(1):92–105, 2011.

[77] M. I. Love, W. Huber, and S. Anders. Moderated estimation of fold change and dispersion for RNA-seq data with DESeq2. *Genome biology*, 15(12):550, 2014.

[78] A. Luzhetskiĭ, B. Ostash, and V. Fedorenko. Interspecies conjugation of *Escherichia coli-Streptomyces globisporus* 1912 using integrative plasmid pSET152 and its derivatives. *Genetika*, 37(10):1340–1347, 2001.

[79] Z. Ma, J. Liu, X. Shentu, Y. Bian, and X. Yu. Optimization of electroporation conditions for toyocamycin producer *Streptomyces diastatochromogenes* 1628. *Journal of basic microbiology*, 54(4):278–284, 2014.

[80] D. J. MacNeil, K. M. Gewain, C. L. Ruby, G. Dezeny, P. H. Gibbons, and T. MacNeil. Analysis of *Streptomyces avermitilis* genes required for avermectin biosynthesis utilizing a novel integration vector. *Gene*, 111(1):61–68, Feb 1992.

[81] S. Maharjan, D. Koju, H. C. Lee, J. C. Yoo, and J. K. Sohng. Metabolic engineering of *Nocardia* sp. CS682 for enhanced production of nargenicin a1. *Applied biochemistry and biotechnology*, 166(3):805–817, 2012.

[82] F. Malpartida and D. Hopwood. Molecular cloning of the whole biosynthetic pathway of a *Streptomyces* antibiotic and its expression in a heterologous host. 1984.

[83] A. Marchler-Bauer, Y. Bo, L. Han, J. He, C. J. Lanczycki, S. Lu, F. Chitsaz, M. K. Derbyshire, R. C. Geer, N. R. Gonzales, et al. CDD/SPARCLE: functional classification of proteins via subfamily domain architectures. *Nucleic Acids Research*, page gkw1129, 2016.

[84] A. Marchler-Bauer, M. K. Derbyshire, N. R. Gonzales, S. Lu, F. Chitsaz, L. Y. Geer, R. C. Geer, J. He, M. Gwadz, D. I. Hurwitz, et al. CDD: NCBI's conserved domain database. *Nucleic acids research*, page gku1221, 2014.

[85] G. L. Marcone, L. Carrano, F. Marinelli, and F. Beltrametti. Protoplast preparation and reversion to the normal filamentous growth in antibiotic-producing uncommon actinomycetes. *The Journal of Antibiotics*, 63(2):83–88, feb 2010.

[86] G. L. Marcone, L. Foulston, E. Binda, F. Marinelli, M. Bibb, and F. Beltrametti. Methods for the genetic manipulation of *Nonomuraea* sp. ATCC 39727. *Journal of industrial microbiology & biotechnology*, 37(10):1097–103, 2010.

[87] P. Mazodier, R. Petter, and C. Thompson1. Intergeneric conjugation between *Escherichia coli* and *Streptomyces* species. *Journal of Bacteriology*, 171(6):3583–3585, jun 1989.

[88] H. McWilliam, W. Li, M. Uludag, S. Squizzato, Y. M. Park, N. Buso, A. P. Cowley, and R. Lopez. Analysis tool web services from the EMBL-EBI. *Nucleic acids research*, 41(W1):W597–W600, 2013.

[89] F. Meyer, A. Goesmann, A. C. McHardy, D. Bartels, T. Bekel, J. Clausen, J. Kalinowski, B. Linke, O. Rupp, R. Giegerich, and A. Puhler. GenDB–an open source genome annotation system for prokaryote genomes. *Nucleic Acids Res.*, 31(8):2187–2195, Apr 2003.

[90] G. Mikhailets and M. Polyak. Acarbose (review). *Pharmaceutical Chemistry Journal*, 23(1):82–87, 1989.

[91] W. Minas, J. E. Bailey, and W. Duetz. Streptomycetes in micro-cultures: Growth, production of secondary metabolites, and storage and retrieval in the 96–well format. *Antonie Van Leeuwenhoek*, 78(3):297–305, 2000.

[92] M. Myronovskyi and A. Luzhetskyy. Genome engineering in actinomycetes using site-specific recombinases. *Appl. Microbiol. Biotechnol.*, 97(11):4701–4712, Jun 2013.

[93] M. Myronovskyi, E. Welle, V. Fedorenko, and A. Luzhetskyy. Beta-glucuronidase as a sensitive and versatile reporter in actinomycetes. *Applied and environmental microbiology*, 77(15):5370–83, aug 2011.

[94] M. Myronovskyy, B. Ostash, I. Ostash, and V. Fedorenko. A gene cloning system for the siomycin producer *Streptomyces sioyaensis* NRRL-B5408. *Folia Microbiologica*, 54(2):91–96, 2009.

[95] G. Ning. Decade in review type 2 diabetes mellitus: At the center of things. *Nature Reviews Endocrinology*, 11(11):636–638, 2015.

[96] S. Nolden, N. Wagner, R. Biener, and D. Schwartz. Analysis of RegA, a pathway-specific regulator of the friulimicin biosynthesis in *Actinoplanes friuliensis*. *J. Biotechnol.*, 140(1-2):99–106, Mar 2009.

[97] V. Ortseifen. Genombasierte Modellbildung zur Biosynthese von Acarviostatin-Metaboliten in drei *Actinoplanes* sp. SE50/110-Stämmen, 2016.

[98] B. Ostash, E. H. Doud, C. Lin, I. Ostash, D. L. Perlstein, S. Fuse, M. Wolpert, D. Kahne, and S. Walker. Complete characterization of the seventeen step moenomycin biosynthetic pathway. *Biochemistry*, 48(37):8830–8841, 2009.

[99] B. Ostash, T. Gren, Y. Hrubskyy, S. Tistechok, S. Beshley, V. Baranov, and V. Fedorenko. Cultivable actinomycetes from rhizosphere of birch (*Betula pendula*) growing on a coal mine dump in Silets, Ukraine. *Journal of basic microbiology*, 54(8):851–7, aug 2014.

[100] B. Ostash, R. Makitrinskyy, S. Walker, and V. Fedorenko. Identification and characterization of *Streptomyces ghanaensis* ATCC14672 integration sites for three actinophage-based plasmids. *Plasmid*, 61(3):171–5, 2009.

[101] B. Ostash, A. Saghatelian, and S. Walker. A streamlined metabolic pathway for the biosynthesis of moenomycin A. *Chemistry & biology*, 14(3):257–267, 2007.

[102] F. Parenti and C. Coronelli. Members of the genus *Actinoplanes* and their antibiotics. *Annual Reviews in Microbiology*, 33(1):389–411, 1979.

[103] W. Piepersberg. Molecular biology, biochemistry and fermentation of aminoglycoside antibiotics. In *Biotechnology of Antibiotics, Second Edition,*, pages 81–163. CRC Press, 1997.

[104] W. Piepersberg, P. Diaz-Guardamino Uribe, A. Stratmann, H. Thomas, U. Wehmeier, and C. Zhang. Developments in the biosynthesis and regulation of aminoglycosides. *Microbial secondary metabolites: Biosynthesis, genetics and regulation*, pages 1–26, 2002.

[105] S. T. Pullan, G. Chandra, M. J. Bibb, and M. Merrick. Genome-wide analysis of the role of GlnR in *Streptomyces venezuelae* provides new insights into global nitrogen regulation in actinomycetes. *BMC Genomics*, 12:175, 2011.

[106] W. Puls and U. Keup. Influence of an α-amylase inhibitor (BAY d 7791) on blood glucose, serum insulin and NEFA in starch loading tests in rats, dogs and man. *Diabetologia*, 9(2):97–101, 1973.

[107] N. J. Ryding, T. B. Anderson, and W. C. Champness. Regulation of the *Streptomyces coelicolor* calcium-dependent antibiotic by *absA*, encoding a cluster-linked two-component system. *Journal of bacteriology*, 184(3):794–805, 2002.

[108] D. D. Schmidt, W. Frommer, B. Junge, L. Muller, W. Wingender, E. Truscheit, and D. Schafer. α-Glucosidase inhibitors. New complex oligosaccharides of microbial origin. *Naturwissenschaften*, 64(10):535–536, Oct 1977.

[109] P. Schwientek, A. Neshat, J. Kalinowski, A. Klein, C. Rückert, S. Schneiker-Bekel, S. Wendler, J. Stoye, and A. Pühler. Improving the genome annotation of the acarbose producer *Actinoplanes* sp. SE50/110 by sequencing enriched 5'-ends of primary transcripts. *Journal of biotechnology*, 190:85–95, 2014.

[110] P. Schwientek, R. Szczepanowski, C. Rückert, J. Kalinowski, A. Klein, K. Selber, U. F. Wehmeier, J. Stoye, and A. Pühler. The complete genome sequence of the acarbose producer *Actinoplanes* sp. SE50/110. *BMC genomics*, 13(1):1, 2012.

[111] P. Schwientek, S. Wendler, A. Neshat, C. Eirich, C. Ruckert, A. Klein, U. F. Wehmeier, J. Kalinowski, J. Stoye, and A. Puhler. Comparative RNA-sequencing of the acarbose producer *Actinoplanes* sp. SE50/110 cultivated in different growth media. *J. Biotechnol.*, 167(2):166–177, Aug 2013.

[112] O. N. Sekurova, T. Brautaset, H. Sletta, S. E. Borgos, . M. Jakobsen M, T. E. Ellingsen, A. R. Str?m, S. Valla, and S. B. Zotchev. In vivo analysis of the regulatory genes in the nystatin biosynthetic gene cluster of *Streptomyces noursei* ATCC 11455 reveals their differential control over antibiotic biosynthesis. *J. Bacteriol.*, 186(5):1345–1354, Mar 2004.

[113] A. Shea, M. Wolcott, S. Daefler, and D. A. Rozak. Biolog phenotype microarrays. *Microbial Systems Biology: Methods and Protocols*, pages 331–373, 2012.

[114] P.-C. Soo, Y.-T. Horng, M.-J. Lai, J.-R. Wei, S.-C. Hsieh, Y.-L. Chang, Y.-H. Tsai, and H.-C. Lai. Pirin regulates pyruvate catabolism by interacting with the pyruvate dehydrogenase E1 subunit and modulating pyruvate dehydrogenase activity. *Journal of bacteriology*, 189(1):109–118, 2007.

[115] A. Stratmann. Identifizierung eines acarbose-biosynthesegenclusters in *Actinoplanes* sp. und charakterisierung ausgewählter enzyme des acarbose-stoffwechsels, 1997.

[116] A. Stratmann, T. Mahmud, S. Lee, J. Distler, H. G. Floss, and W. Piepersberg. The AcbC protein from *Actinoplanes* species is a C7-cyclitol synthase related to 3-dehydroquinate synthases and is involved in the biosynthesis of the α-glucosidase inhibitor acarbose. *Journal of Biological Chemistry*, 274(16):10889–10896, 1999.

[117] P. J. Szaniszlo and H. Gooder. Cell wall composition in relation to the taxonomy of some Actinoplanaceae. *Journal of bacteriology*, 94(6):2037–2047, 1967.

[118] G.-Y. Tan and T. Liu. Rational synthetic pathway refactoring of natural products biosynthesis in Actinobacteria. *Metabolic Engineering*, 2016.

[119] L. P. Thapa, T. J. Oh, H. C. Lee, K. Liou, J. W. Park, Y. J. Yoon, and J. K. Sohng. Heterologous expression of the kanamycin biosynthetic gene cluster (pSKC2) in *Streptomyces venezuelae* YJ003. *Appl. Microbiol. Biotechnol.*, 76(6):1357–1364, Oct 2007.

[120] The National Institute of Diabetes and Digestive and Kidney Diseases, U.S. Department of Health and Human Service. Symptoms and Causes of Diabetes, 2017. https://www.niddk.nih.gov/health-information/diabetes/overview/symptoms-causes accessed on 17.02.2017.

[121] H. Thomas. Acarbose-metabolismus in *Actinoplanes* sp. se50/110, 2002.

[122] Y. Tong, P. Charusanti, L. Zhang, T. Weber, and S. Y. Lee. CRISPR-Cas9 based engineering of actinomycetal genomes. *ACS synthetic biology*, 4(9):1020–1029, 2015.

[123] C. Tran. CadC, a potential regulator protein in *Actinoplanes* sp. SE50/110, 2017.

[124] K. Uchida, M.-S. Jang, Y. Ohnishi, S. Horinouchi, M. Hayakawa, N. Fujita, and S.-I. Aizawa. Characterization of *Actinoplanes missouriensis* spore flagella. *Applied and Environmental Microbiology*, 77(7):2559–2562, apr 2011.

[125] L. Van Mellaert, L. Mei, E. Lammertyn, S. Schacht, and J. Anné. Site-specific integration of bacteriophage VWB genome into *Streptomyces venezuelae* and construction of a VWB-based integrative vector. *Microbiology (Reading, England)*, 144 (Pt 1:3351–8, 1998.

[126] W. G. Wade. The genus *Eubacterium* and related genera. In *Prokaryotes Vol. 4*, pages 823–835. Springer, 2006.

[127] N. Wagner, C. Oßwald, R. Biener, and D. Schwartz. Comparative analysis of transcriptional activities of heterologous promoters in the rare actinomycete *Actinoplanes friuliensis*. *Journal of biotechnology*, 142(3):200–204, 2009.

[128] U. Wehmeier and W. Piepersberg. Biotechnology and molecular biology of the α-glucosidase inhibitor acarbose. *Applied microbiology and biotechnology*, 63(6):613–625, 2004.

[129] U. F. Wehmeier. The biosynthesis and metabolism of acarbose in *Actinoplanes* sp. SE50/110: a progress report. *Biocatalysis and Biotransformation*, 21(4-5):279–284, 2003.

[130] S. Wendler, D. Hürtgen, J. Kalinowski, A. Klein, K. Niehaus, F. Schulte, P. Schwientek, H. Wehlmann, U. F. Wehmeier, and A. Pühler. The cytosolic and extracellular proteomes of *Actinoplanes* sp. SE50/110 led to the identification of gene products involved in acarbose metabolism. *Journal of Biotechnology*, 167(2):178–189, aug 2013.

[131] S. Wendler, D. Hurtgen, J. Kalinowski, A. Klein, K. Niehaus, F. Schulte, P. Schwientek, H. Wehlmann, U. F. Wehmeier, and A. Puhler. The cytosolic and extracellular proteomes of *Actinoplanes* sp. SE50/110 led to the identification of gene products involved in acarbose metabolism. *J. Biotechnol.*, 167(2):178–189, Aug 2013.

[132] S. Wendler, V. Ortseifen, M. Persicke, A. Klein, A. Neshat, K. Niehaus, S. Schneiker-Bekel, F. Walter, U. F. Wehmeier, J. Kalinowski, and A. Pühler. Carbon source dependent biosynthesis of acarviose metabolites in *Actinoplanes* sp. SE50/110. *Journal of Biotechnology*, 191:113–120, 2014.

[133] S. Wendler, A. Otto, V. Ortseifen, F. Bonn, A. Neshat, S. Schneiker-Bekel, F. Walter, T. Wolf, T. Zemke, U. F. Wehmeier, M. Hecker, J. Kalinowski, D. Becher, and A. Puhler. Comprehensive proteome analysis of *Actinoplanes* sp. SE50/110 highlighting the location of proteins encoded by the acarbose and the pyochelin biosynthesis gene cluster. *J Proteomics*, 125:1–16, Jul 2015.

[134] S. Wendler, A. Otto, V. Ortseifen, F. Bonn, A. Neshat, S. Schneiker-Bekel, T. Wolf, T. Zemke, U. F. Wehmeier, M. Hecker, J. Kalinowski, D. Becher, and A. Puhler. Comparative proteome analysis of *Actinoplanes* sp. SE50/110 grown with maltose or glucose shows minor differences for acarbose biosynthesis proteins but major differences for saccharide transporters. *J Proteomics*, 131:140–148, Jan 2016.

[135] D. Wibberg, J. Blom, S. Jaenicke, F. Kollin, O. Rupp, B. Scharf, S. Schneiker-Bekel, R. Sczcepanowski, A. Goesmann, J. C. Setubal, R. Schmitt, A. Puhler, and A. Schluter. Complete genome sequencing of *Agrobacterium* sp. H13-3, the former *Rhizobium lupini* H13-3, reveals a tripartite genome consisting of a circular and a linear chromosome and an accessory plasmid but lacking a tumor-inducing Ti-plasmid. *J. Biotechnol.*, 155(1):50–62, Aug 2011.

[136] D. Wibberg, J. Blom, C. Ruckert, A. Winkler, A. Albersmeier, A. Puhler, A. Schluter, and B. E. Scharf. Draft genome sequence of *Sinorhizobium meliloti* RU11/001, a model organism for flagellum structure, motility and chemotaxis. *J. Biotechnol.*, 168(4):731–733, Dec 2013.

[137] T. Wolf, T. Gren, E. Thieme, D. Wibberg, T. Zemke, A. Puhler, and J. Kalinowski. Targeted genome editing in the rare actinomycete *Actinoplanes* sp. SE50/110 by using the CRISPR/Cas9 System. *J. Biotechnol.*, 231:122–128, Aug 2016.

[138] World Health Organization. Global report on diabetes, 2016. http://www.who.int/diabetes/global-report/en/ accessed on 17.02.2017.

[139] G. D. Wright. Something old, something new: revisiting natural products in antibiotic drug discovery. *Canadian journal of microbiology*, 60(3):147–154, 2014.

[140] Y. Xu, L. Wang, J. He, Y. Bi, M. Li, T. Wang, L. Wang, Y. Jiang, M. Dai, J. Lu, et al. Prevalence and control of diabetes in chinese adults. *Jama*, 310(9):948–959, 2013.

[141] S. Yin, Z. Li, X. Wang, H. Wang, X. Jia, G. Ai, Z. Bai, M. Shi, F. Yuan, T. Liu, et al. Heterologous expression of oxytetracycline biosynthetic gene

cluster in *Streptomyces venezuelae* WVR2006 to improve production level and to alter fermentation process. *Applied Microbiology and Biotechnology*, pages 1–10, 2016.

[142] C.-S. Zhang. Biosynthesis of the C7-cyclitol moiety of acarbose in *Actinoplanes* species SE50/110. 7-O-phosphorylation of the initial cyclitol precursor leads to proposal of a new biosynthetic pathway. *Journal of Biological Chemistry*, 277(25):22853–22862, jun 2002.

[143] C.-S. Zhang. Genomic analysis of secondary metabolite producing actinomycetes: Acbm is a 2-epi-5-epi-valiolone 7-kinase, 2002.

[144] S. Zotchev, K. Haugan, O. Sekurova, H. Sletta, T. E. Ellingsen, and S. Valla. Identification of a gene cluster for antibacterial polyketide-derived antibiotic biosynthesis in the nystatin producer *Streptomyces noursei* ATCC 11455. *Microbiology (Reading, Engl.)*, 146 (Pt 3):611–619, Mar 2000.