CHAPTER 6.1

Summary and concluding remarks
Natural products

Natural products are compounds synthesized by living beings that have a great variety of functions for the source organism and for human use. One example is that of secondary metabolites. These are defined as molecules that are not essential for the producing organism but that may provide an advantage for their survival. Pigments, chelators and bioactive molecules (e.g. antibiotics) are some of the most important groups of molecules that are part of secondary metabolism. Actinomycetes are the main bacterial source of secondary metabolites, with *Streptomyces* as one of the most important producers. Most of the antibiotics that are now being used in medicine have been isolated from *Streptomyces*. The fast development of resistance to antibiotics in pathogenic bacteria is a well-known problem that has been extensively described and that needs to be addressed. From 1962 to 2000 not a single new antibiotic class has been introduced for treatment of infectious diseases. The last class commercialized is that of lipopeptides in the early 2000s, but this chemical class was already known from 1952. The discovery of new classes of molecules to fight the growing number of resistant microorganisms is therefore currently a priority. Great efforts are being made in further exploring the secondary metabolite richness that nature provides.

Activation of putative biosynthesis gene clusters (BGCs)

Different tools and strategies have been developed to identify gene clusters that may encode pathways for the biosynthesis of novel antibiotic compounds in the growing number of available microbial genomes. With the ongoing development of bioinformatic tools and genome sequencing it is becoming evident that the number of microbial natural products currently known is just the tip of the iceberg of the real repertoire available. This was also observed for the genome sequences of
streptomycetes, bacteria traditionally used for the production of secondary metabolites, showing the potential to produce many unknown compounds with putative antimicrobial properties \cite{11, 12}. Also, strains of genera not well-known for the production of secondary metabolites, e.g. \textit{Rhodococcus} or \textit{Mycobacterium} \cite{13}, harbour a great number of BGCs in their genomes. Most of these BGCs have remained cryptic (their products are unknown) and different techniques have been developed aiming to activate them, including heterologous expression of BGCs, redirecting precursors by inactivating known biosynthetic pathways, manipulating regulatory systems or building improved biosynthesis pathways by synthetic biology approaches \cite{14-17} (Chapter 1).

In the work reported in this PhD thesis we explored different techniques to activate BGCs in \textit{Streptomyces clavuligerus}, a well-known producer of antibiotic compounds (Chapter 2 and 3). We also performed a deep bioinformatic analysis of the surprisingly large repertoire of cryptic secondary gene clusters found in the genome sequences of rhodococci and Mycobacteria (Chapter 4). This included the $\gamma$-butyrolactone signalling system which is well-known for its regulatory role in streptomycetes secondary metabolism \cite{18}. This cluster had been previously predicted in four strains of rhodococci \cite{13} but we show that it extends throughout the genus. We characterized it in more detail in \textit{Rhodococcus jostii} RHA1 (Chapter 5).

**Heterologous production of indigoidine**

\textit{S. clavuligerus} is known for the synthesis of clavulanic acid, a $\beta$-lactamase inhibitor that is co-formulated with amoxicillin to produce the well-known drug Augmentin \cite{19, 20}. Clavulanic acid is one of the main secondary metabolites produced by \textit{S. clavuligerus}. Its biosynthesis pathway is partially shared with the one of clavam antibiotics, such as alanyl-clavam \cite{21}. The wild type strain is also known to produce trace amounts of
holomycin and tunicamycin-like antibiotics MM 19290. Genome analysis showed that this strain contains an impressive number of cryptic BGCs. We targeted the cryptic BGC for the biosynthesis of the blue pigment indigoidine. This pigment is produced by diverse bacteria and it is thought to have anti-oxidative and antimicrobial properties. It is however not known whether indigoidine is the final product of the biosynthesis pathway and the function of most genes in this BCG are unknown. The cryptic gene cluster of S. clavuligerus encodes a homologue of the nonribosomal peptide synthetase needed for the production of the blue pigment indigoidine (IndC). Interestingly, IndC of S. clavuligerus has an extra domain representing a 4-oxalocrotonate tautomerase (4-OT)-like enzyme, of unknown function. This tautomerase-like enzyme is usually encoded by a separate gene in the indigoidine biosynthesis gene cluster (indD). In Chapter 2 we show that such a fusion of putative indC and indD genes occurs more widespread, and is present in at least 30 different bacterial strains. Heterologous expression of the S. clavuligerus indC(D) gene plus flanking genes in several Streptomyces host strains did not result in blue pigment production. Only expression of indC(D) alone, controlled by a strong promoter, resulted in production of indigoidine in Streptomyces coelicolor, R. jostii, and Escherichia coli, showing for the first time that IndC(D) from S. clavuligerus indeed encodes an indigoidine synthetase. We also show that separate expression of the S. clavuligerus IndC, lacking the tautomerase domain, yields more indigoidine than expression of the complete indC(D) gene. To further study the activity of the tautomerase domain in IndC(D), a truncated S. clavuligerus gene, encoding only the IndD domain, was successfully expressed in E. coli and purified as an active enzyme, catalyzing a promiscuous Michael-like addition reaction but none of the canonical reactions described for 4-OT tautomeras. The physiological function of this tautomerase-like enzyme is still unknown.
but its fusion to IndC may ensure the same expression levels, guaranteeing that both enzymes are in close proximity for a possible fast reaction before the indigoidine monomers are condensed to form indigoidine, or even as a way to select for an inactive tautomerase. Chapter 2 shows that secondary metabolites from \textit{S. clavuligerus} can be produced in different heterologous hosts, including \textit{Rhodococcus}, an actinomycete with faster growth than \textit{Streptomyces}. A range of molecular tools are available for rhodococci making these strains highly interesting as host for heterologous expression (and engineering) of \textit{Streptomyces} secondary metabolite gene clusters.

\textbf{Six derivatives of antibiotics MM 19290 were isolated and partially characterized from \textit{S. clavuligerus} strain Δ7}

Aiming to activate or upregulate secondary metabolite biosynthesis pathways in \textit{S. clavuligerus} ATCC 27064, we constructed a mutant strain that is fully blocked in the synthesis of its main secondary metabolite clavulanic acid, \textit{S. clavuligerus} strain Δ7. The mutations introduced are known to block also the synthesis of holomycin. \textit{S. clavuligerus} strain Δ7 was found to produce bioactive compounds. After extraction and several purification steps, we were able to analyse the bioactive compounds by mass spectrometry (MS) and MS/MS. This resulted in the putative identification and characterization of 6 different compounds with masses equivalent to those of different derivatives of tunicamycin, and the structurally related corynetoxins and streptovirudines. Tunicamycin-like compounds named MM 19290 are known to be produced only in trace amounts by the wild type strain \textit{S. clavuligerus} ATCC 27064 and no structural information was available yet. Different derivatives of tunicamycin show different activity. Tunicamycins are able to inhibit N-glycosylation which makes them highly toxic, but they are used in various ways in research and medicine. New derivatives of these antibiotics
may show different activities that could be further exploited. The results presented in this work contribute to our knowledge on the diversity of tunicamycin-related compounds synthesized by \textit{S. clavuligerus} ATCC 27064 (Chapter 3).

**The great potential for secondary metabolite production of \textit{Rhodococcus}**

Rhodococci are well-known for their capacity to degrade a wide range of (aromatic) compounds \textsuperscript{35-38} but hardly explored for their ability to produce secondary metabolites. Recently it has become evident that rhodococci encode a large number of unexplored putative BGCs \textsuperscript{13}. The genus \textit{Rhodococcus} is closely related to that of \textit{Mycobacterium}, which comprises feared pathogens such as \textit{Mycobacterium tuberculosis} and \textit{Mycobacterium leprae}. To further explore the specialized metabolic potential of \textit{Rhodococcus}, we performed a computational analysis of 20 available rhodococcal genomes, and compared these with those of several \textit{Mycobacterium} strains as well as that of \textit{Amycolicicoccus subflavus}. Most of the predicted BGCs from these strains lack any homology with gene clusters responsible for synthesis of known natural products. We identified five gene clusters that are shared among all strains of these three genera studied. Their analysis led to identification of PKS13 as the last condensation step in mycolic acid synthesis; this is an essential biosynthetic activity for these genera and already used as target to combat pathogenic \textit{Mycobacterium} strains \textsuperscript{39}. Essential enzymes involved in cell wall synthesis, in iron-sulphur cluster biosynthesis or vitamin K12 synthesis, were also found in these shared clusters. We were not able to assign a function to two of the universally shared gene clusters but their presence in the genomes of all these strains suggests that they likely also have an essential role. These shared clusters thus may be responsible for the production of interesting novel compounds and/or
provide targets to combat the pathogenic strains. Mutational inactivation of these clusters may generate weaker strains that can be used as live vaccines.

We also studied the predicted gene clusters that only are present in closely related strains (clade-specific) and therefore appear in one branch of the phylogenetic tree, and the clusters that are only present in single strains. Strain-specific clusters are interesting targets for secondary metabolite production, since there is less risk that the final product is known already from a different species; in many studies this has been problematic and many compounds have been re-discovered. By studying the conservation patterns of the predicted BGCs across the different species, we hypothesized about their biological roles and prioritized them for a future functional characterization. We also observed that from all rhodococci studied, *R. jostii* RHA1 and three strains from *Rhodococcus opacus* contain the highest number of BGCs in their genomes, even higher than the predicted number in the well-known secondary metabolite producers *S. coelicolor* or *S. clavuligerus*, making them highly interesting strains to study in more detail (Chapter 4).

**Identification and characterization of the γ-butyrolactone system from *R. jostii* RHA1**

In view of the great number of cryptic secondary metabolite gene clusters present in rhodococci, it is highly relevant to gain insight into their secondary metabolism regulation. Our further analysis focussed on *R. jostii* RHA1, one of the most interesting *Rhodococcus* strains. One of the most conserved BGCs in rhodococci is a homologue of the extensively studied quorum-sensing system for γ-butyrolactones, known to participate in the activation of secondary metabolism in the *Streptomyces* genus. The data shows that *R. jostii* RHA1 synthesises γ-butyrolactones, the first report of the production of these molecules.
outside the *Streptomyces* genus. Extracted γ-butyrolactones from *R. jostii* RHA1 (RJBs) were shown to be able to interact with receptor protein ScbR of *S. coelicolor* using a previously developed reporter system \(^{44}\). This indicates that inter-genera communication occurs via these RJB molecules. LC-MS analysis of the same extracts using synthetic standards of the *S. coelicolor* γ-butyrolactones \(^{45}\) identified an RJB molecule with identical mass and elution time as 6-dehydro SCB2, the predicted precursor of *S. coelicolor* γ-butyrolactone SCB2. We furthermore identified the *R. jostii* RHA1 key RJB biosynthesis gene *gbla*. Following deletion of the *gbla* gene, no RJB production could be detected whereas higher RJB levels were synthesized when *gbla* was overexpressed. We were also able to detect for the first time the synthesis of diffusible molecules with antimicrobial activity from *R. jostii* RHA1 (Chapter 5). The only known secondary metabolite BGCs from this strain are involved in the formation of siderophores. The detected antimicrobial molecules thus may be novel and further increase our interest in this strain. Knowing the relevant role that γ-butyrolactones have in *Streptomyces* secondary metabolism \(^{46}\), RJBs could be used as tool for the awakening of cryptic secondary metabolic gene clusters in rhodococci.

The work in this thesis provides new insights in the activation of cryptic secondary metabolite gene clusters in *S. clavuligerus* and *R. jostii* RHA1 and shows in detail the potential that the genus *Rhodococcus* has to produce novel natural products. Both *Streptomyces* and *Rhodococcus* are highly interesting genera to explore for novel secondary metabolites, native ones or following expression of foreign clusters. They encode a surprisingly high number of cryptic BGCs in their genomes \(^{11-13}\) and techniques for DNA manipulation are available for both genera \(^{47, 48}\). *Streptomyces* has historically been used for this purpose and has been extensively studied for the production of bioactive compounds \(^7\).
Furthermore, great effort has been made to unravel the regulatory networks in *Streptomyces* species resulting in a better understanding of gene expression that will surely be key to the current efforts being dedicated to activate cryptic gene clusters.\cite{41,49-52} *Rhodococcus* however shows a great potential to produce secondary metabolites that are unexplored.\cite{13} The genus is extensively studied for its great catabolic potential; and molecular tools are available to perform genetic manipulation.\cite{37,48} The presence of the common \(\gamma\)-butyrolactone system between both genera indicates that this system is a promising tool for activation of the secondary metabolism in both genera. In order to improve our chances of discovering novel antimicrobials that can be used against resistant bacterial strains, we need to apply a combination of different approaches. First, the use of bioinformatics tools to detect putative BGCs in newly sequenced genomes, serving to choose those strains that may be the source of interesting new compounds and to select the most promising cryptic gene clusters for further study (Chapter 4). The next step is to select a suitable host strain for activation of these gene clusters, which also will define which of the different strategies will be followed to induce the production of these compounds (Chapters 1 - 3). The analysis of the regulatory systems of donor and hosts strains is also vital to understand why these gene clusters are not active and how their expression can be triggered (Chapter 5). This thesis therefore gives insight in all the necessary steps needed to attempt and perform the activation of cryptic BGCs.