Universiteit Leiden
The Netherlands

Expanding the chemical space of antibiotics produced by Paenibacillus and Streptomyces

Machushynets, N.V.

Chapter 9

General discussion

Bacteria are invaluable sources of natural products (NPs), owing to their remarkable capacity to produce a plethora of both primary and specialized metabolites. As a result, bacterial fermentation has long been used to manufacture NPs that have pharmaceutical and agrochemical value. The extensive array of bioactivities exhibited by bacterial natural products is a result of the evolutionarily diversified metabolism, with a main portion still waiting to be discovered (Gavriilidou *et al.*, 2022). The limited rate of antibiotic discovery, coupled with the escalating threat of antibiotic resistance, underscores the pressing need for innovative approaches to fully harness the diversity of bacterial natural products (Silver, 2011, Lewis, 2013).

In this thesis, approaches to aid the discovery of novel antibiotics from bacteria are presented, focusing on new strategies from the perspectives of microbiology, chemistry and bioinformatics. This study encompasses exploring the chemical diversity of specialized metabolites from *Paenibacillus* and *Streptomyces*, eliciting the production of bioactive natural products, identifying novel biosynthetic gene clusters (BGCs), and devising strategies to prioritize bioactive molecules for subsequent isolation, structure elucidation, or organic synthesis.

GNPS molecular networking for the discovery of new bioactive compounds

Global Natural Products Social molecular networking (GNPS MN) is a powerful tool to process and analyze tandem mass spectrometry (MS^2) data (Wang *et al.*, 2016). GNPS MN has been successfully applied to assist in the discovery of new microbial compounds (Nguyen *et al.*, 2016, Wu *et al.*, 2019, Männle *et al.*, 2020) to support strain prioritization, culture conditions, and extraction methods for drug discovery purposes (Crüsemann *et al.*, 2017, Floros *et al.*, 2016), microbial interaction investigations (Watrous *et al.*, 2012, Vallet *et al.*, 2017) and chemical diversity profiling (Purves *et al.*, 2016, Nguyen *et al.*, 2013). Additionally, NP prioritization strategies have been applied based on the combination of molecular networking with different sources of information (e.g., genomic, taxonomic, geographic, bioactivity, and spectral data) (Olivon *et al.*, 2017, Nothias *et al.*, 2018, Olivon *et al.*, 2020, Crüsemann *et al.*, 2017). In **Chapter 3**, we introduce an early-stage prioritization method that combines feature-based GNPS MN and MassQL to prioritize peptides that contain basic amino acids (2,4-diaminobutyric acid, lysine, ornithine, histidine, and arginine) with potential antimicrobial properties. Basic amino acid signatures were detected and mapped into massive multi-informative MN acquired over an extract library of 227 *Paenibacillus* isolates from the Auburn University Plant-Associated Microbial strain collection. This led to the discovery of new paenilipoheptins and new bacitracin congeners, designated as paenitracins (Figure 1).

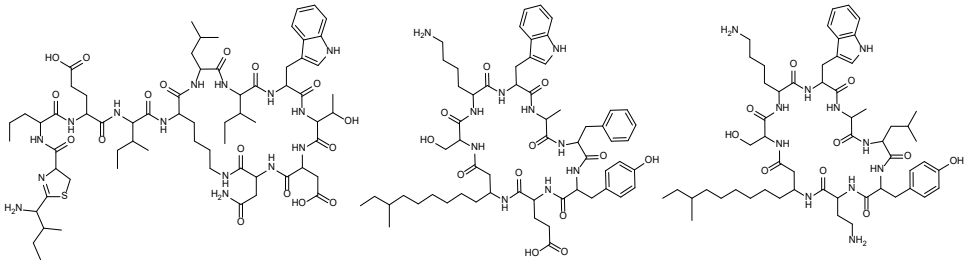

Figure 1. Overview of the compounds discovered by the MassQL MN integrative approach (Chapter 3).

Paenitracin differs from the known bacitracin A at three amino acid positions featuring amino acids that are unique among all bacitracins identified so far, namely Leu7, Trp9, and Thr10 (Suleiman *et al.*, 2017). Despite its significant structural divergence, paenitracin exhibits potent bioactivity against Gram-positive pathogens, including vancomycin-resistant *Enterococcus faecium* E155. Paenitracin is produced by *Paenibacillus* species, as opposed to previously reported bacitracins that have been isolated from *Bacillus* (Johnson *et al.*, 1945, Stone & Strominger, 1971). We used an adopted GNPS MN approach to provide insight into chemical structures and diversity of *Paenibacillus* specialized metabolites and to facilitate the discovery of new natural products.

Chemical elicitors in combination with metabolomics

In our exploration of natural product diversity, metaphorically we are likely still "looking through a keyhole", which only offers a restricted view of the entire chemical landscape of bacterial specialized metabolites. Bacteria do not produce their whole spectrum of specialized metabolites under routine laboratory conditions. Many BGCs are activated only under specific environmental conditions, and the cognate signals are likely missing under laboratory culturing conditions (Demain & Fang, 2000). Understanding the signals and the corresponding regulatory mechanisms will allow the development of eliciting approaches, which are essential for natural product discovery.

Streptomycetes have a complex lifestyle that begins with the germination of a spore, which grows out to form a multicellular network known as the vegetative mycelium. In response to stressors like nutrient depletion, a portion of the mycelium is autolytically degraded, releasing nutrients into the environment for the development of aerial hyphae and spores (Jakimowicz & van Wezel, 2012, Flärdh & Buttner, 2009). Antibiotic production by streptomycetes is linked to the growth phase (Rigali *et al.*, 2008, Barka *et al.*, 2016). Therefore, the modulation of growth conditions plays a critical role in shaping the metabolic output and, consequently, the diversity of bioactive compounds synthesized by these bacteria (van der Heul *et al.*, 2018, Urem *et al.*, 2016). Novel quinazolinone A and B were discovered following

alterations of the growth conditions during cultivation of *Streptomyces* sp. MBT27 (**Chapter 4**). We discovered that the production of quinazolinone alkaloids was induced in response to glycerol. The chemical structures of quinazolinones A and B and the glycerol-dependent production suggest that glycerol participates in the construction of the ring system (Figure 2). Our data indicate that quinazolinones A and B form a new sub-branch in the family of quinazolines. Identification and an in-depth analysis of the biosynthetic gene cluster should reveal the precise biosynthetic pathway for these novel molecules.

Moreover, carbon sources had a significant effect on the antimicrobial activity of *Streptomyces* sp. MBT27 (**Chapter 5**). Antimicrobial activity was most pronounced when the culture medium was supplemented with glycerol + mannitol, glucose, glycerol or fructose. Statistical analysis showed that the bioactivity correlated with elevated production of actinomycins, the first antibiotics ever discovered in an actinomycete (Katz, 1967). Importantly, carbon utilization not only affected the overall production levels of specialized metabolites, but also contributed to the chemical diversity of the actinomycins, whereby production of the novel compound actinomycin L was observed in the cultures grown with 1% glucose, 1% and 2% glycerol, 1% mannitol +1 % glycerol (Figure 2). The isolation of actinomycin L in such a well-studied family of NPs underlines that new molecules can be discovered even within extensively studied microbes and compound classes.

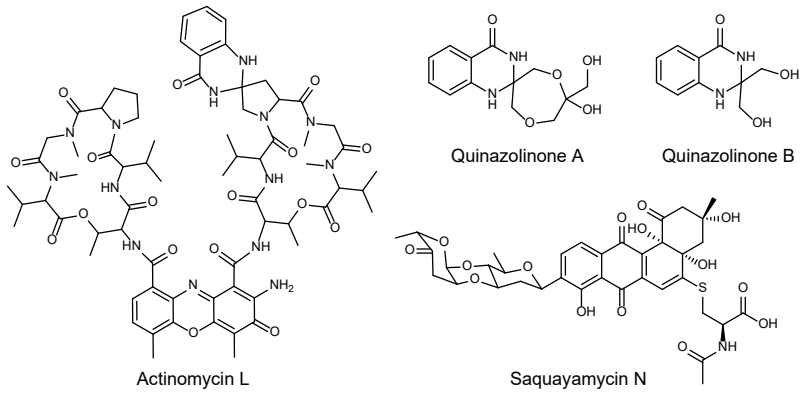

Figure 2. Overview of the compounds discovered by the elicitation techniques (Chapters 4, 5, and 6).

To enhance early-stage prioritization of bioactive compounds, we developed an analytical platform for the rapid and efficient identification and dereplication of bioactive metabolites based on nanofractionation (nanoRAPIDS) (**Chapter 6**). NanoRAPIDS combines at-line high-resolution nanofractionation coupled to LC-MS/MS and featured-based molecular networking (FBMN) in a single analysis. The nanoRAPIDS platform allows the identification of low-abundance molecules against a background of abundantly present known antibiotics.

We used the platform to correlate antimicrobial bioactivities to compounds in extracts of *Streptomyces* sp. MBT84 grown in the presence or absence of catechol as the elicitor (van Bergeijk *et al.*, 2022). The majority of the bioactivity peaks were associated with mass features annotated as angucyclines, which predominantly contributed to the increased bioactivity of the crude extract of *Streptomyces* sp. MBT84 when grown with catechol (van Bergeijk *et al.*, 2022). This afforded the isolation of a novel angucycline, saquayamycin N. The combination of various elicitation strategies with multi-omics approaches and nanoRAPIDS platform offers a powerful workflow to prioritize bioactive metabolites for isolation and assist in 'decrypting' cryptic BGCs.

Global genome mining as a way to access hidden chemical diversity

Genome mining methodologies complement the capabilities of analytical chemistry techniques by accessing specialized metabolites that may not be produced under any tested growth conditions (Sayed *et al.*, 2020, Gavriilidou *et al.*, 2022, Lloyd Karen *et al.*, 2018). Genome mining-guided natural product discovery involves detecting novel BGCs, predicting the chemical structures of the molecules of interest and estimating their bioactivity based on these structures (Medema & Fischbach, 2015). The exploration of bacterial nonribosomal peptide synthetases (NRPSs) through global-genome mining has shown a substantial genomic potential within Firmicutes, with the majority of NRPSs remaining unexplored (Li *et al.*, 2018). *Paenibacillus* spp. are important producers of nonribosomal peptides (NRPs), including polymyxins, octapeptins, pelgipeptins, fusaricidins, tridecaptins, and cilagicins (Cochrane & Vederas, 2016). These NRPs exhibit a broad range of activities, ranging from antibacterial and antifungal to anticancer and antiviral. To gain insight into the chemical diversity of NRPs and discover new potentially bioactive natural products, we analyzed 785 complete publicly available *Paenibacillus* genomes using BiG-SCAPE (Navarro-Muñoz *et al.*, 2020). (**Chapters 7** and **8**). Further in silico analysis using PARAS facilitated the prediction of amino acid compositions for peptide scaffolds, leading to the discovery of ten novel tridecaptin BGCs with variations in amino acid composition compared to known scaffolds.

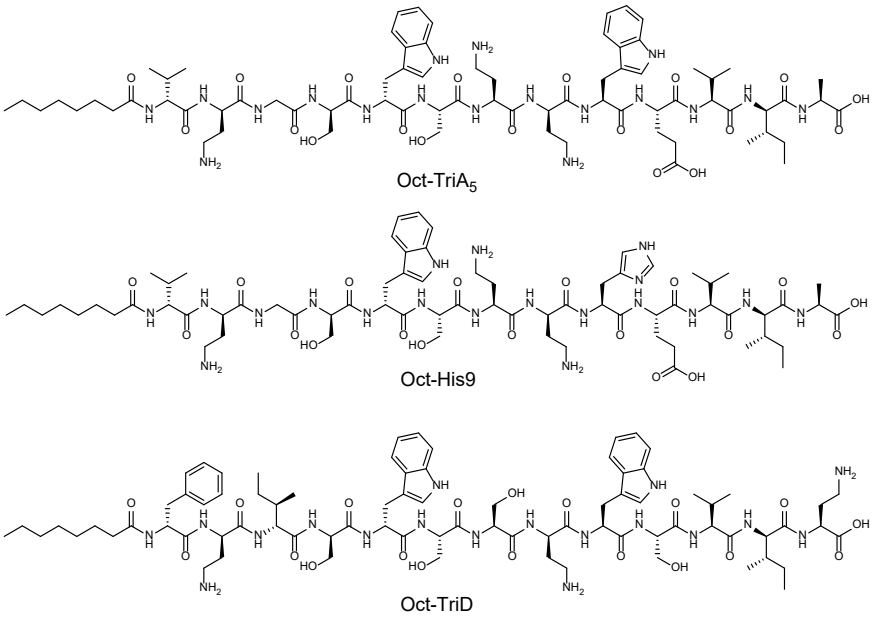

Figure 3. Syn-BNPs discovered by global genome mining of *Paenibacillus* NRPs.

We identified significant diversification within the tridecaptin family, whereby chemically synthesized congeners demonstrated distinct bioactivity profiles compared to known tridecaptins (**Figure 3**). Tridecaptin A_5 was discovered through global genome mining, identified by metabolomic studies, and produced synthetically. Considering the striking difference in the spectrum of activity between known tridecaptin A_1 (active against Gram-negatives) and new tridecaptin A_5 (active against both Gram-positives and Gram-negatives), we were able to pinpoint a specific residue (position 9) responsible for altering the antibacterial spectrum and cytotoxicity profile of these compounds. Single amino acid substitution in position 9 led to the discovery of Oct-His9, which efficiently inhibited colistin-resistant *Pseudomonas aeruginosa* and showed lower hemolytic and cytotoxic activity compared to Oct-$TriA_1$. Global genome mining also revealed Oct-TriD, that contained 6 substitutions as compared to Oct-$TriA_1$ and showed significant bioactivity against *E. faecium*. Thus, it is important to continue to explore tridecaptins towards the development of effective antibacterial agents. The approach of synthetic-bioinformatic natural product (syn-BNP) represents a cutting-edge methodology at the intersection of computational biology, synthetic chemistry, and microbiology (Chu *et al.*, 2020). Integrating single amino acid substitution strategies into syn-BNP development enables researchers to systematically explore the structure-activity relationships of natural products, facilitating the rational design of analogs with enhanced therapeutic potential.

Chimeric biosynthesis of tridecaptins and actinomycin

Analysis of *Paenibacillus* genomes using BiG-SCAPE revealed that some of the tridecaptin BGCs differ not only by the predicted amino acid composition but also by their architecture. These "truncated BGCs" contain the 10-modular NRPS *triD* gene but lack the gene encoding the terminal 3-modular TriE synthase. Suggestively, the truncated tridecaptin-like BGC always co-localizes either with a full-length BGC containing all 13 modules or with another truncated tridecaptin-like BGC that encodes a 3-modular TriE synthase. Considering the widespread co-occurrence of full-length and truncated BGCs, we hypothesized that both tridecaptin BGCs are functional and participate in the biosynthesis of so-called "chimeric tridecaptins". Through MS-based metabolomic analysis of 227 plant-associated *Paenibacillus* spp. we detected the chimeric tridecaptin B in the crude extracts of seven *Paenibacillus* isolates. Genome sequencing of one such isolate, *Paenibacillus* spp. JJ-1683, identified both full-length and truncated tridecaptin BGCs. Therefore, we conclude that tridecaptin B is a product of collaborative biosynthesis, where the NRPS of the truncated cluster provides the first ten amino acids, while the full-length cluster contributes the last three.

Another instance of collaborative biosynthesis is seen in the production of actinomycin L (**Chapter 5**). The biosynthesis of actinomycin L requires genes outside of the actinomycin cluster for the production of anthranilamide, one of the precursors of actinomycin L. The feeding experiments convincingly showed that actinomycin L is formed through the reaction of anthranilamide with the 4-keto group on the proline residue in the pentapeptide lactone prior to the condensation of the pentapeptide lactones into actinomycin L. The enzyme responsible for converting anthranilic acid to anthranilamide remains to be identified. Chimeric biosynthesis is a mechanism of bacterial and fungal secondary metabolism, which has the potential to increase further the structural diversity of encoded molecules and subsequent biological activity.

Conclusions and future perspectives

Integration of diverse data types, such as metabolomics, genomic, transcriptomic, and phenotypic data, provides a holistic insight into the antibiotic-producing potential of bacteria and to prioritize the antibiotics they produce. To make new scientific breakthroughs in this important field, we need to find ways to explore undiscovered natural product diversity. Artificial intelligence (AI) techniques are a very promising addition to the available tools and enable combining information from multiple sources to uncover complex relationships between molecular structures, biological functions, and bioactivity (Mullowney *et al.*, 2023). AI developments are expected to cause a revolution in the discovery of bioactive NPs and the design of NP-inspired drugs and their future development.

Besides finding novel chemical entities, microbes are biofactories for the development of exogenous metabolites, peptides, and proteins through recombinant DNA

technology (Pham *et al.*, 2019). The future steps in metabolic engineering should involve genetic manipulations of NP biosynthetic powerhouses like S*treptomyces* (Whitford *et al.*, 2021), so as to develop novel and complex NPs in a more sustainable manner.

This research has demonstrated that a combined use of metabolomic and genomic approaches is an effective way to discover new NPs. By integrating genomic insights with metabolomics analyses, we have discovered a rich diversity of chemical structures and bioactivities even within known families of molecules, including quinazolinones, actinomycins, saquayamycins, tridecaptins, paenilipoheptins, and bacitracins. This strategy not only facilitates the identification of previously unexplored natural products but also offers insights into their biosynthetic pathways and potential applications.

By harnessing the collective power of AI technologies, natural product chemistry, synthetic biology, and genome mining strategies, we can unlock the vast reservoir of microbial biodiversity and pave the way for the development of next-generation antibiotics to combat emerging infectious diseases and address the growing threat of antimicrobial resistance.