Universiteit Leiden
The Netherlands

Expanding the chemical space of antibiotics produced by Paenibacillus and Streptomyces

Machushynets, N.V.

Chapter 1

Thesis outline

Nature sustains seemingly unlimited resources of natural products (NPs), which are metabolites derived from living organisms, such as plants, animals, and microorganisms. Microorganisms are considered nature's medicine makers, producing biomolecules with varying functional activities such as antibiotics, antifungals, anticancer agents and immunosuppressants. About 70 % of clinically used antibiotics are produced by bacteria, with the majority originating from Actinobacteria (Katz & Baltz, 2016). Typically, these compounds are products of specialized metabolism and play a role in the survival of bacteria.

The evolution and widespread distribution of antibiotic resistance have made diseases that were once easily treatable challenging again. Unfortunately, the increase in antibiotic resistance is paralleled by a lack of success in discovering new antibacterial drugs (Brown & Wright, 2016). A significant challenge in antibiotic discovery is the repeated rediscovery of known compounds, often referred to as "low-hanging fruits" (Kolter & van Wezel, 2016). It has been estimated that only 3% of the natural products potentially encoded in bacterial genomes have been experimentally characterized (Gavriilidou *et al.*, 2022). This imbalance between the enormous potential and the limited exploration emphasizes the urgency for innovative strategies for antibiotic discovery.

There are two main approaches to antibiotic discovery: the traditional "top-down" method and the more recent "bottom-up" strategy known as genome mining (Figure 1) (Luo *et al.*, 2014). In the "top-down" approach, compounds are discovered from microorganisms without prior knowledge of their genetic makeup. Conversely, genome mining involves the systematic analysis of microbial genomes to identify biosynthetic gene clusters (BGCs) responsible for the biosynthesis of natural products. Genome mining facilitates prioritizing and subsequent activation or chemical synthesis of the predicted product. The goal of the work described in this thesis was to uncover previously undiscovered natural products with potent bioactivities. The thesis is divided into two sections based on the methodology employed for antibiotic discovery. The first section (**Chapters 3–6**) covers "top-down" approaches and the second section (**Chapters 7–8**) utilizes "bottom-up" strategies.

The research involves both chemistry and biology, e.g. (1) screening and prioritization of bacterial extract libraries to discover new bioactive compounds employing MS-based metabolomics and Global Natural Products Social molecular networking (GNPS MN); (2) activating silent BGCs in actinomycetes; (3) detection and isolation of bioactive compounds from a complex mixture; (4) identifying the structures of compounds by using spectroscopic techniques like nuclear magnetic resonance (NMR) or mass spectrometry (MS) and, where needed, complemented by X-ray crystallography; (5) organic synthesis of the discovered compounds and subsequent derivatization to enhance the bioactivity properties (6) testing antimicrobial activity of discovered compounds; (7) linking the compound to the biosynthetic gene clusters.

Chapter 2 provides a review of recent developments in the antibiotic discovery field. This review discusses (1) the contemporary challenges in antibiotic discovery; (2) advances in genome mining and in translating the genetic code into compounds; (3) strategies to expand the known chemical space and to activate the cryptic BGCs; (4) advances in LC-MS-based metabolomics for the prioritization and subsequent identification of bioactive compounds.

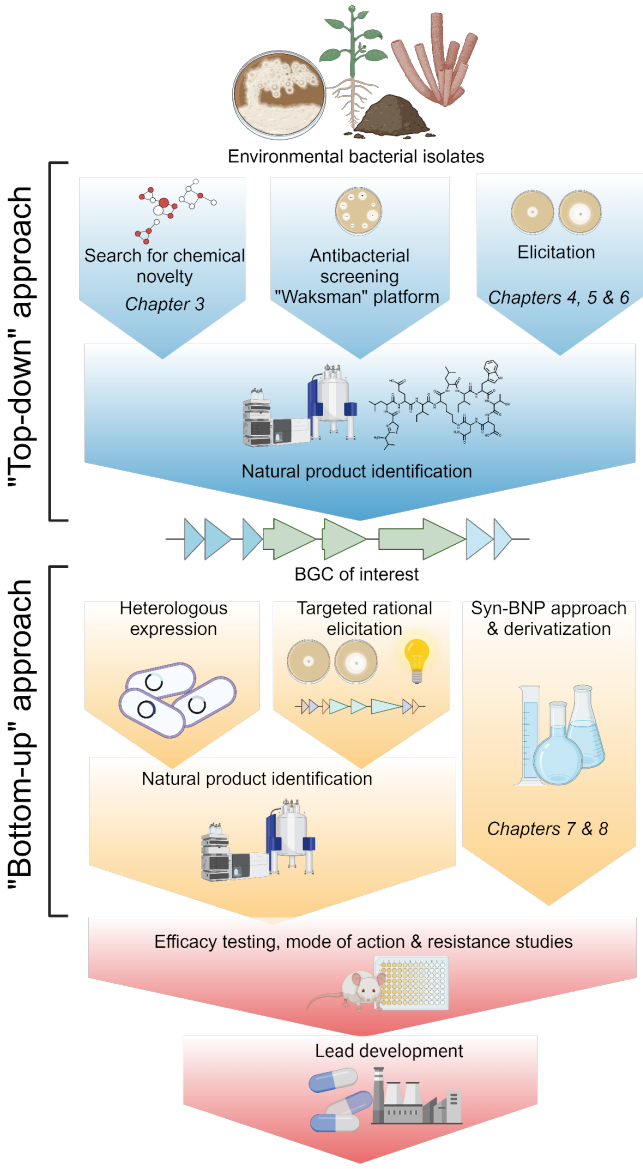

Figure 1. Summary of the major steps and processes in the antibacterial drug discovery pipeline.

Chapter 3 shows the application of MS-based metabolic profiling to facilitate the discovery of novel microbial NPs. We used MassQL-enhanced molecular networks (MN) to profile nonribosomal peptides (NRPs) produced by *Paenibacillus*. We propose an NP prioritization approach based on GNPS MN and MassQL that effectively targets basic amino acid-containing peptides within massive multi-informative MN. Compounds of interest are thereby highlighted based on the detection of the basic amino acid signatures and mapping the sub-structure information onto an MN acquired over an extract library of 227 *Paenibacillus* plant-associated isolates. This led to the discovery of new bacitracin congeners, designated as paenitracins, with potent bioactivity against Gram-positive pathogens, including vancomycin-resistant *Enterococcus faecium* E155.

In **Chapters 4** and **5**, different culture media were tested to increase secondary metabolite production and antimicrobial activity in *Streptomyces* sp. MBT27. LC-MS and molecular networking were subsequently used to profile the metabolomes elicited through these culture condition perturbations compared to the respective control experiments. LC-MS data were statistically analyzed using multivariate data analyses to differentiate chemical profiles and correlate specific *m/z* values to the bioactivity of the crude extracts. The up-scale fermentation in the resulting optimized culture media facilitated the isolation of new compounds from quinazolinone (**Chapter 4**) and actinomycin families (**Chapter 5**). The biosynthetic pathways of new compounds were proposed, and the corresponding BGC was identified through bioinformatics analysis.

Early-stage compound dereplication is critical to the discovery of bioactive natural products. In **Chapter 6,** we present an efficient analytical platform to allow the discovery of low-abundance compounds at medium throughput, based on at-line nanofractionation combined with molecular networking and bioactivity assays, which we call nanoRAPIDS. NanoRAPIDS encompasses analytical scale separation of natural extracts, post-column nanofractionation followed by high-throughput bioassay, automated mass spectrometry identification, and GNPS MN for dereplication and analog search. Proof of concept was obtained for families of iturins and surfactins in the extracts from *Bacillus*, whereby we readily identified compounds that act against Gram-positive and Gram-negative bacteria and fungi. Then, nanoRAPIDS was applied to address the eliciting role of catechol in the production of bioactive compounds, highlighting the individual mass features that explained the increased bioactivity in the crude extract.

In **Chapter 7,** we report the genome-guided discovery of new lipopeptide antibiotics tridecaptin A_5 and tridecaptin D. Through bioinformatic-guided organic synthesis and comprehensive antibacterial assays, we demonstrate that synthetic analogs Oct-TriA_5 and Oct-TriD exhibit unusual bioactivities not commonly associated with tridecaptins. The change in the antibacterial spectrum of Oct-TriA_5 could be explained solely by a Phe to Trp substitution compared to Oct-TriA_1, while Oct-TriD contained six substitutions. Screening

of tridecaptin analogs substituted at position 9 identified Oct-His9 as a potent congener with exceptional efficacy against colistin-resistant *Pseudomonas aeruginosa* and reduced hemolytic and cytotoxic properties. These findings underscore the significant potential of tridecaptin analogs as viable alternatives in combatting antibiotic-resistant bacterial pathogens.

Moreover, global genome mining of *Paenibacillus* nonribosomal peptide synthetase (NRPS) BGCs uncovered tridecaptin-like BGCs with unusual architecture. **Chapter 8** illustrates biosynthetic "crosstalk" among two gene clusters enabled de novo synthesis of chimeric tridecaptin B_1. We hypothesized that the truncated tridecaptin-like BGC in *Paenibacillus* sp. JJ-1683 might be involved in the chimeric biosynthesis of tridecaptin B_1 by biosynthesizing the first ten amino acids, upon which the last three are incorporated by the TriE synthase encoded by the full-length tridecaptin BGC. This result underscores chimeric biosynthesis as a strategy for the dynamic production of similar compounds with distinct chemistries and bioactive profiles.

Finally, **Chapter 9** presents a general discussion of the thesis, which also includes a summary of the most important data and observations and provides perspectives for future antibiotic discovery from microbial sources.