Antimicrobial potential of Geobacillus sp. ZGt-1 isolated from Zara hot spring in Jordan

Piecing the puzzle of the antagonistic activity of a novel bacterial strain

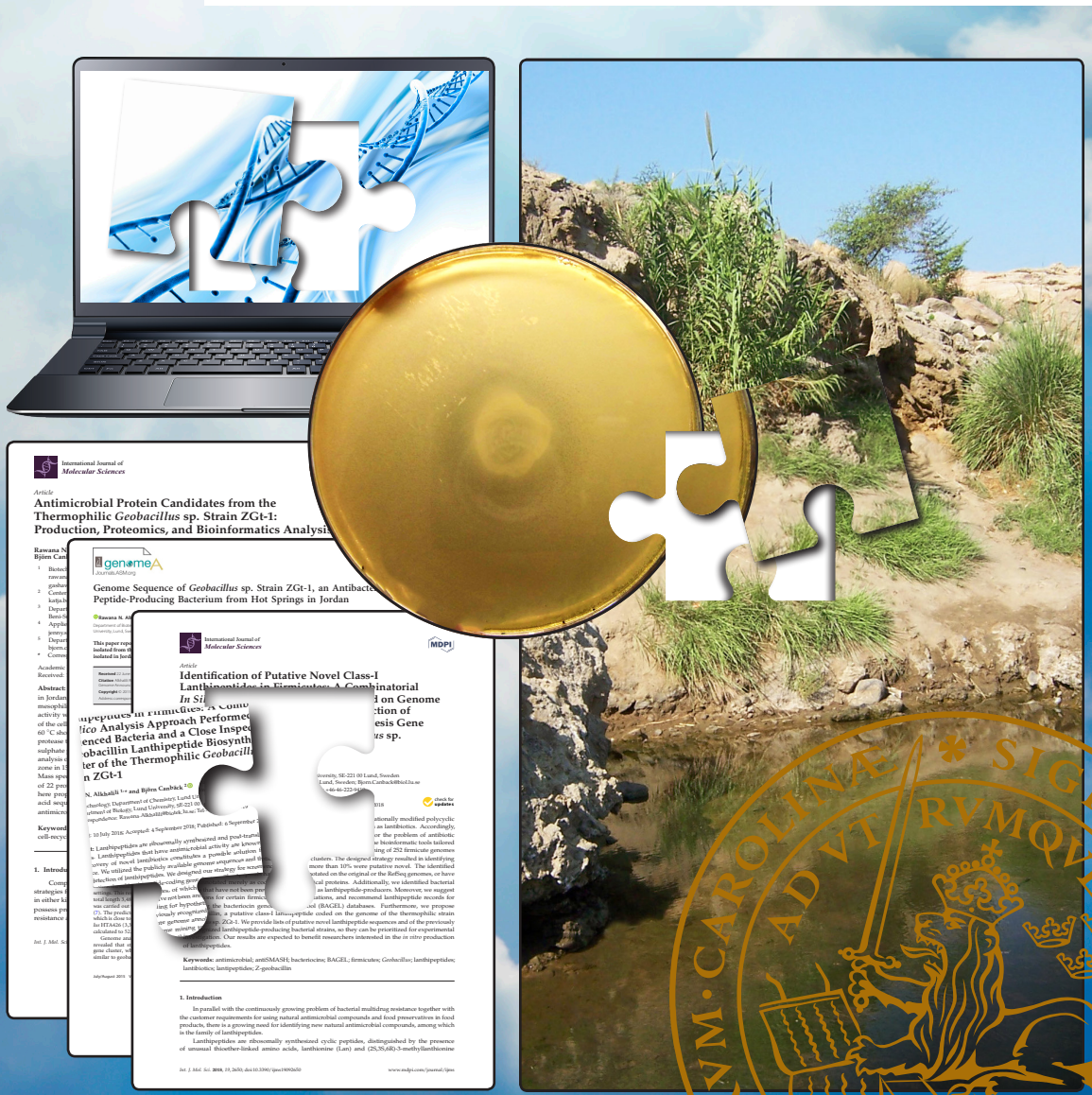
Alkhalili, Rawana

2019

Antimicrobial potential of *Geobacillus* sp. ZGt-1 isolated from Zara hot spring in Jordan

Piecing the puzzle of the antagonistic activity of a novel bacterial strain

RAWANA ALKHALILI | DIVISION OF BIOTECHNOLOGY | LUND UNIVERSITY



Antimicrobial potential of *Geobacillus* sp. ZGt-1 isolated from Zara hot spring in Jordan

Piecing the puzzle of the antagonistic activity of a novel bacterial strain

Rawana Alkhalili

LUND
UNIVERSITY

DOCTORAL DISSERTATION
by due permission of the Faculty of Engineering, Lund University, Sweden.
To be defended at the Blue Hall, Ecology Building, Sölvegatan 31, Lund, Sweden, on 6th December, at 13:15

Faculty opponent
Professor Tilmann Weber
Technical University of Denmark

Organization:
LUND UNIVERSITY
Division of Biotechnology
P.O. Box 124, SE-221 00 Lund, Sweden

Author: Rawana Alkhalili

Document name:
Doctoral Dissertation

Date of issue:
6th December, 2019

Title and subtitle:
Antimicrobial potential of *Geobacillus* sp. ZGt-1 isolated from Zara hot spring in Jordan
– Piecing the puzzle of the antagonistic activity of a novel bacterial strain

Abstract

The problem of antibiotic resistance is rising continuously, creating a warning signal that calls for finding alternatives. Not only that, but there is also a rising demand for natural food preservatives. The pharmaceutical and food industry sectors are working on fulfilling these needs.

This thesis is exploring possible solutions to these issues. The research studies presented here introduce the potential of thermophilic bacteria isolated from hot springs as a source for antimicrobial agents that could be applied in the pharmaceutical and food industries. The focus is on the bacterial strain *Geobacillus* sp. ZGt-1 that was isolated from Zara hot spring in Jordan. Experimental work and *in silico* analyses of the genome sequence of this strain revealed its antimicrobial potential. This strain grows at 60 °C and antagonizes the growth of the food spoiling thermophilic bacterium *Geobacillus stearothermophilus*. It also antagonizes the growth of the mesophilic bacteria *Bacillus subtilis* and the pathogenic *Salmonella* Typhimurium, both grown at 37 °C.

The thesis presents antimicrobial peptide and protein candidates of the strain ZGt-1. These candidates include a list of secreted proteins within the range of 10–30 kDa that are thermostable and SDS-resistant. They also include a putative novel lanthipeptide, which we identified as Z-geobacillin that is smaller than 3.5 kDa. The candidates also include toxins belonging to various families of type II toxin-antitoxin system, within the range of 3–17 kDa.

The protein candidates were produced at 60 °C by immobilizing the cells of ZGt-1 in agar beads that were cultivated in sequential batches to solve the issue of producing the proteins in liquid. The proteins were then purified and identified using a combination of proteomic and bioinformatic tools.

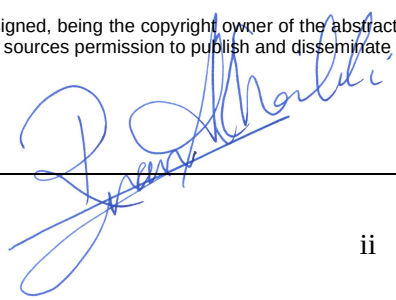
The Z-geobacillin represents the first lanthipeptide identified in a hot spring-inhabiting bacterium and is expected to be more stable than nisin. In addition to Z-geobacillin, seven putative novel class-I lanthipeptides were predicted to be produced by different firmicutes, by mining the genome sequences of all sequenced members of the firmicute phylum. Within this phylum, we also predicted the potential of 18 bacterial strains to be lanthipeptide-producers.

Type II toxin-antitoxin (TA) families of *Geobacillus* strains, which have not been well-studied, have also been covered in this thesis, and 15 putative novel toxins and antitoxins have been identified together with potentially new TA families. Moreover, a hypothesis on the regulation of gene expression of the XRE-COG2856 TA family has been proposed.

Overall, the results indicate that *Geobacillus* sp. ZGt-1 is a source of putative novel antimicrobial peptides and proteins. This study represents the first report on a *Geobacillus* strain potentially producing a group of various antibacterial peptides and proteins. The results also indicate that members of the thermophilic genus *Geobacillus,* in general*,* represent promising producers of antimicrobials.

Key words: Thermophiles, *Geobacillus*, antimicrobial potential, antimicrobial peptides, bacteriocins, lanthipeptides, toxin-antitoxin, genome sequence

Supplementary bibliographical information		**Language**: English
ISRN: LUTKDH/TKBT-19/1173-SE		ISBN 978-91-7422-710-9 (print) ISBN 978-91-7422-711-6 (pdf)
Recipient's notes	**Number of pages**: 245	



Signature _____ Date 2019-10-30

Antimicrobial potential of *Geobacillus* sp. ZGt-1 isolated from Zara hot spring in Jordan

Piecing the puzzle of the antagonistic activity
of a novel bacterial strain

Rawana Alkhalili

LUND
UNIVERSITY

Cover collage by: Rawana Alkhalili and Paula Leckius
Some of the graphics adapted from Adobe Stock.
Poem on back cover by: Rawana Alkhalili
Poem proofread by: Julienne Stewart-Sandgren and Ladaea Rylander, Lund University.

Faculty of Engineering (LTH)
Department of Chemistry, Division of Biotechnology

ISBN 978-91-7422-710-9 (print)
ISBN 978-91-7422-711-6 (pdf)
ISRN LUTKDH/TKBT-19/1173-SE

Printed in Sweden by Media-Tryck, Lund University
Lund 2019

Media-Tryck is an environmentally certified and ISO 14001:2015 certified provider of printed material. Read more about our environmental work at www.mediatryck.lu.se

MADE IN SWEDEN

To my superb and much-loved parents and brother

“Happiness and bacteria have one thing in common; they multiply by dividing!”

— *Rutvik Oza*

Table of Contents

Abstract

The problem of antibiotic resistance is rising continuously, creating a warning signal that calls for finding alternatives. Not only that, but there is also a rising demand for natural food preservatives. The pharmaceutical and food industry sectors are working on fulfilling these needs.

This thesis is exploring possible solutions to these issues. The research studies presented here introduce the potential of thermophilic bacteria isolated from hot springs as a source for antimicrobial agents that could be applied in the pharmaceutical and food industries. The focus is on the bacterial strain *Geobacillus* sp. ZGt-1 that was isolated from Zara hot spring in Jordan. Experimental work and *in silico* analyses of the genome sequence of this strain revealed its antimicrobial potential. This strain grows at 60 °C and antagonizes the growth of the food spoiling thermophilic bacterium *Geobacillus stearothermophilus*. It also antagonizes the growth of the mesophilic bacteria *Bacillus subtilis* and the pathogenic *Salmonella* Typhimurium, both grown at 37 °C.

The thesis presents antimicrobial peptide and protein candidates of the strain ZGt-1. These candidates include a list of secreted proteins within the range of 10–30 kDa that are thermostable and SDS-resistant. They also include a putative novel lanthipeptide, which we identified as Z-geobacillin that is smaller than 3.5 kDa. The candidates also include toxins belonging to various families of type II toxin-antitoxin system, within the range of 3–17 kDa.

The protein candidates were produced at 60 °C by immobilizing the cells of ZGt-1 in agar beads that were cultivated in sequential batches to solve the issue of producing the proteins in liquid. The proteins were then purified and identified using a combination of proteomic and bioinformatic tools.

The Z-geobacillin represents the first lanthipeptide identified in a hot spring-inhabiting bacterium and is expected to be more stable than nisin. In addition to Z-geobacillin, seven putative novel class-I lanthipeptides were predicted to be produced by different firmicutes, by mining the genome sequences of all sequenced members of the firmicute phylum. Within this phylum, we also predicted the potential of 18 bacterial strains to be lanthipeptide-producers.

Type II toxin-antitoxin (TA) families of *Geobacillus* strains, which have not been well-studied, have also been covered in this thesis, and 15 putative novel toxins and antitoxins have been identified together with potentially new TA families. Moreover, a hypothesis on the regulation of gene expression of the XRE-COG2856 TA family has been proposed.

Overall, the results indicated that *Geobacillus* sp. ZGt-1 is a source of putative novel antimicrobial peptides and proteins. This study represents the first report on a *Geobacillus* strain potentially producing a group of various antibacterial peptides and proteins. The results also indicate that members of the thermophilic genus *Geobacillus*, in general, represent promising producers of antimicrobials.

Popular science summary

There are continuously ongoing battles among human beings. However, we have a more severe and alarming war with totally different creatures that we cannot see with our own eyes. I am talking about "stubborn" bacteria that cause us life-threatening diseases, and at the same time, are resistant to antibiotics so we cannot get treated. You may have already heard of them as "superbugs". In fact, according to the British daily newspaper, "The Guardian", such stubborn bacteria (or superbugs) "kill far more people each year globally than terrorism" [1]. Many kinds of bacteria that used to be sensitive to antibiotics are no longer so, as they can play around the antibiotic and resist it. A report published in 2016 estimated that 10 million people will die yearly by 2050, due to antimicrobial resistance [2]!

We must act!

We need to supply our armamentarium with "weapons" that enable us to win the battle. The "weapons" have to be new, so the stubborn bacteria have not learned to avoid them, and must be effective so we can rely on them as guardians. In other words, we need a new set of antibiotics. To find a new antibiotic, we need to find new antibiotic sources. It goes as simple as this; to find a new thing that no one before you has seen, you need to explore new places.

One of the best treasuries to mine for new antibiotics is "nature". In our study, we decided to go out in nature, looking for "weapons" that can help us to win the fight against those superbugs. We chose hot springs and explored those in Jordan and found that Zara hot spring represents a potential source of what we are looking for.

The concept of the present study was based on the conflict that is ongoing between the different bacteria themselves, as a result of competition. We can actually use this conflict and turn it into our benefit. In light of what I mentioned about the importance of exploiting new places, and given that exploring hot springs as a source of antibiotics is a new approach that has not been widely exploited, there is an opportunity in hot springs for finding a potential new antibiotic that could protect us from the superbugs.

In hot springs, bacteria are exposed to harsh conditions and are in conflict with each other as they compete for limited nutrients and space. Throughout the conflict, bacteria use their own "weapons" to kill their competing bacteria. Among the "weapons" they use are different sets of substances, known as "antimicrobial substances" that are present inside the bacteria. We can take these "weapons" (or antimicrobial substances) and use them in fighting the harmful and stubborn bacteria. To be able to use these antimicrobial substances, one needs first to get hold of their source, i.e; the producing bacteria, identify their types

and their possible targets, and then get hold of the antimicrobial substances themselves.

In the present study, we collected water samples from Zara hot spring and separated bacteria from the water and grew them in the lab. The isolated bacteria are special in that they like to grow at high temperatures, way higher than the temperature of our bodies. This feature makes these bacteria safe for us since they do not prefer to grow in or on our bodies. Yet, their "weapons" can kill or limit the growth of harmful bacteria, which cause us diseases or ruin our food. Therefore, we can use these high temperature-loving bacteria as a source of "weapons". To spot and identify the "weapons" used by bacteria, one needs to have a map. For that, we used the genetic map of the bacteria that produce the "weapons". We found that those bacteria have an array of different "weapons" that could target other bacteria, which cause us diseases or spoil our dairy products and canned food. The current study provides some hopes for finding new antibiotics and also new and natural food preservatives.

Exploiting the conflicts among different kinds of bacteria may bring us the peace we are longing for in terms of antibiotic treatment. Nevertheless, one hand cannot clap alone; science and research cannot provide the ultimate solution for the problem. There is a significant act that has to be played by the society as well. Patients must avoid the misuse of antibiotics and follow the doctors' guidelines, which play a pivotal role in setting a limit to the spread of the superbugs.

Via cooperation between society and science, we can win the battle against harmful bacteria!

[1] Fong K. 2013. Antibiotic resistance: Why we must win the war against superbugs. The Guardian, March 17th.

[2] de Kraker MEA, Stewardson AJ, Harbarth S. 2016. Will 10 Million People Die a Year due to Antimicrobial Resistance by 2050? PLoS Med13(11): e1002184. doi:10.1371/journal.pmed.1002184

New Life

Since antibiotics are often overused
And the course begins and then disused
Resistant bacteria are suffused

Antibiotics are in states of divergency
Leaving the world in an emergency
Could thermophiles defeat insurgency?

Bacteria play a symphony of peace and war
Creating bunches of mysteries for us to explore
In the sparkling Zara, thermophiles' triumph is our score

From hot springs, new life may rise
Unseen creatures with hidden surprise
It's in research where the prize lies

Science radiates rays of sunshine
Enlightening the research plan design
And keeping the mind's radar always online

Thermophiles deserve our appreciation
Marvelous nation with lifelong fascination
This thesis, a tiny candle for the horizon's illumination

By Rawana Alkhalili
Lund, Spring 2019

List of papers

This thesis is based on the following publications, which will be referred to in the text by their Roman numerals. The papers are provided towards the end of this book.

I. Alkhalili, R. N., Bernfur, K., Dishisha, T., Mamo, G., Schelin, J., Canbäck, B., Emanuelsson, C. & Hatti-Kaul, R. 2016. Antimicrobial Protein Candidates from the Thermophilic *Geobacillus* sp. Strain ZGt-1: Production, Proteomics, and Bioinformatics Analysis. *Int J Mol Sci*, 17. (doi: 10.3390/ijms17081363).

II. Alkhalili, R. N., Hatti-Kaul, R. & Canbäck, B. 2015. Genome Sequence of *Geobacillus* sp. Strain ZGt-1, an Antibacterial Peptide-Producing Bacterium from Hot Springs in Jordan. *Genome Announc*, 3. (4) e00799-15. (doi: 10.1128/genomeA.00799-15).

III. Alkhalili, R. N. & Canbäck, B. 2018. Identification of Putative Novel Class-I Lanthipeptides in Firmicutes: A Combinatorial In Silico Analysis Approach Performed on Genome Sequenced Bacteria and a Close Inspection of Z-Geobacillin Lanthipeptide Biosynthesis Gene Cluster of the Thermophilic *Geobacillus* sp. Strain ZGt-1. *Int J Mol Sci,* 19. (doi: 10.3390/ijms19092650).

IV. Alkhalili, R. N., Wallenius, J. & Canbäck, B. 2019. Towards Exploring Toxin-Antitoxin Systems in *Geobacillus*: A Screen for Type II Toxin-Antitoxin System Families in A Thermophilic Genus. *Preprints*, 2019100325. (doi: 10.20944/preprints201910.0325.v1).

My contribution to the papers

The idea of exploring the antimicrobial potential of thermophiles was provided by Prof. emeritus Olle Holst.

I. I conducted the fieldwork in Jordan, collected the water samples, isolated the bacterial strains, screened for their antibacterial activity, and identified them at the molecular level. For the cultivation and protein purification experiments, I designed and planned them with GM and TD. I performed all the experiments and analyzed all the results, except for the mass spectrometry analysis (MS), which was conducted by KB. BC performed the genome annotation of the ZGt-1 strain. JS provided the mesophilic bacterial strains and supervised the work with pathogens. CE supervised the MS analysis. RHK supervised the project. I wrote the manuscript, which was then revised, edited, and approved by all authors. I submitted the manuscript.

II. Sequencing the genome of the strain was my decision, as I concluded that it was crucial for proceeding with the work. I established the collaboration with the Wellcome Trust Centre for Human Genetics in Oxford for sequencing the genome, and with BC for the genome assembly. I analyzed the assembled genome for the detection of the bacteriocin-coding genes. I carried out the process of the genome submission to the NCBI under the supervision of BC. I wrote the manuscript together with BC. The manuscript was then revised, edited, and approved by all authors. I submitted the manuscript.

III. I generated the idea of conducting this study and designed the analysis approach under the supervision of BC. BC wrote the software script and produced the initial lanthipeptide set, and I analyzed and curated the data using different software packages and databases and produced the final lanthipeptide set, which I thoroughly investigated based on available literature. All the work was done under BC supervision. I wrote the manuscript, which was then revised, edited, and approved by both authors. I submitted the manuscript.

IV. I generated the idea of conducting this study and designed the analysis approach under the supervision of BC. I and JW conducted the analyses

and curated the data under the supervision of BC. I wrote the manuscript, and JW contributed to the writing of the "Material and Method" section. The manuscript was then revised, edited, and approved by all authors. I submitted the manuscript.

Abbreviations

7TMR-HDED	7 transmembrane helices receptors-HD hydrolase; a hydrolase with a catalytic His-Asp (HD) motif, and ED stands for extracellular domain
aa	Amino acid
ABC transporters	ATP-binding cassette transporters
AbrB domain	AidB regulator domain
Amidase	*N*-acetylmuramoyl-L-alanine amidase
AMP	Antimicrobial Peptide/Protein
AMPA	Antimicrobial peptide/protein algorithm
ANN	Artificial Neural Network
antiSMASH	Antibiotics and Secondary Metabolite Analysis Shell
APD3	Antimicrobial Peptide Database
BAGEL	BActeriocin GEnome mining tooL
BCWHs	Bacterial Cell Wall Hydrolases
BSA	Bacteriocin of *Staphylococcus aureus*
c-di-AMP	cyclic-di-adenosine monophosphate
c-di-GMP	cyclic-di-guanosine monophosphate
$CAMP_{R3}$ database	The Collection of Anti-microbial Peptides database
CDD	Conservation Domain Database
COG	Clusters of Orthologous Group
DA	Discriminant analysis

DD-carboxypeptidase	Serine-type D-alanyl-D-alanine carboxypeptidase
Dha	Dehydroalanine
Dhb	Dehydrobutyrine
Doc	Death on curing
DUF	Domain of Unknown Function
DTT	Dithiothreitol
DUF	Domain of Unknown Function
Fic	Filamentation induced by Cyclic AMP
GacTA	*Geobacillus* acetyltransferase Toxin-Antitoxin
Gcn5	General control non-repressible 5
GeoAI	Geobacillin I precursor peptide
GeoB	Geobacillin I dehydratase enzyme
GeoC	Geobacillin I Cyclase enzyme
GeoGEF	Geobacillin I self-immunity ABC transporter proteins
GeoI	Geobacillin I self-immunity protein
GeoK	Geobacillin I sensor histidine Kinase protein
GeoR	Geobacillin I response Regulatory protein
GeoTI	Geobacillin I ABC Transporter protein
GNAT	Gcn5-related N-acetyltransferases
GRAS	Generally Regarded As Safe
GRAVY	Grand Average of Hydropathicity
HEPN	Higher Eukaryotes and Prokaryotes Nucleotide-binding
HTH domain	Helix-Turn-Helix domain
KAAS	KEGG Automatic Annotation Server
KEGG	Kyoto Encyclopedia of Genes and Genomes
KNTase	Kanamycin nucleotidyltransferase

LAB	Lactic acid bacteria
Lan	Lanthionine
LanP	Lanthipepttide processing protease
LPS	Lipopolysaccharide
LC-MS/MS	Liquid chromatography tandem mass spectrometry
MeLan	(2*S*,3*S*,6*R*)-3-methyllanthionine
meso-DAP	*Meso*-diaminopimelic acid
MH	Mueller Hinton culture medium
MNT	Minimal Nucleotidyltransferase
MRSA	Methicillin-Resistant *Staphylococcus aureus*
NavSS	Normalized average of aggregation propensity
NAMP	Non-Antimicrobial Peptide/Protein
NGS	Next Generation Sequencing
nt	nucleotide
NTase	nucleotidyltransferase
PCD	Programmed cell death
PG	Peptidoglycan
PGAP	Prokaryotic Genome Annotation Pipeline
Phd	Prevents host death
pI	Isoelectric point
PIN domain	PilT N-terminus domain
(p)ppGpp	Guanosine tetra or pentaphosphate
ProOpDB	Prokaryotic Operon DataBase
PSK	Post-Segregational Killing
RelBE	Relaxed BE
RF	Random Forests
RHH domain	Ribbon-Helix-Helix domain
RODEO	Rapid ORF Description and Evaluation Online genome-mining platform
SpoVT	Stage V sporulation protein T

SVM	Support Vector Machines
TA system	Toxin-Antitoxin system
TADB	Toxin-Antitoxin Database
UPF	Uncharacterized Protein Family
VapBC	Virulence associated proteins BC
VRE	Vancomycin-resistant *Enterococcus*
wHTH domain	winged Helix-Turn-Helix domain
XRE	Xenobiotic Response Element
ZgeoA	Z-geobacillin precursor peptide
ZgeoB	Z-geobacillin dehydratase enzyme
ZgeoC	Z-geobacillin cyclase enzyme
ZgeoGEF	Z-geobacillin self-immunity ABC transporter proteins
ZgeoI	Z-geobacillin self-immunity protein
ZgeoK	Z-geobacillin sensor histidine Kinase protein
ZgeoR	Z-geobacillin response regulatory protein
ZgeoT	Z-geobacillin ABC transporter protein

1. Introduction

"When I woke up just after dawn on September 28, 1928, I certainly didn't plan to revolutionize all medicine by discovering the world's first antibiotic, or bacteria killer. But I suppose that was exactly what I did."

Alexander Fleming

"It is the end of the road for antibiotics unless we act urgently".

Tom Frieden

Since the famous coincidence that led Alexander Fleming to discover penicillin in 1928, natural antimicrobials, produced mainly by bacteria and fungi, and chemically synthesized antibiotics have provided the "drug armory" with different antimicrobial therapeutic compounds over the past century (Moloney, 2016). Half a century ago, those antimicrobial compounds were considered as "miracle drugs" (Saleem, 2014). However, over time, bacteria have developed resistance against synthesized and natural antimicrobials (Saleem, 2014). The bacterial antibiotic-resistance mechanisms have outperformed the efficiency of the available antibiotics (Marinelli and Genilloud, 2014; Villa and Veiga-Crespo, 2014). More than 70% of pathogenic bacteria are resistant to most known antibiotics available in the market, and the mortality rate caused by some multi-drug resistant pathogens has reached 80% (Bérdy, 2012).

There is a need for new antibiotics, and there is a desire for chemical-free food. The rise and pervasiveness of antibiotic-resistant pathogens urge for finding antimicrobial agents with novel structures and modes of action against their targets (Marinelli and Genilloud, 2014). In addition to the need for novel antibiotics as therapeutics, there are continuous demands for preservative-free food and "food-greener additives" within the food industry. Such demands have created a critical need for finding novel antimicrobial agents that are safe to be added to food products and effective in protecting food from food-borne pathogens and food-spoilage microorganisms (Moloney, 2016; Tiwari et al., 2009; Ji, 2002).

The search for a new source of antimicrobial compounds is a challenging task calling for plotting various sets of strategies to fulfill those urgent demands

(Moloney, 2016). Although isolation of natural compounds via screening nature relies on the serendipity that poses the risk of the re-isolation of known compounds, this strategy has resulted in the discovery of diverse novel natural antimicrobial compounds from different classes of organisms (Moloney, 2016). It is more likely to discover active compounds with novel mechanisms of action when screening natural products as compared to the chemical synthesis of compounds (Bérdy, 2005). That is because nature offers a molecular diversity that exceeds the economic feasibility of chemical synthesis. Moreover, natural products are already pre-screened by nature and have been evolutionarily selected for particular biological functions and interactions (Bérdy, 2005). Nature is almost an inexhaustible source of new bioactive products (Bérdy, 2005). More novel products are assuredly available in nature, awaiting their discovery (Kingston, 2011).

The first critical step in the search for potentially novel antimicrobial compounds is the selection of the source to be screened. Unexploited sources of natural antimicrobial compounds represent an attractive starting point (Wright and Sutherland, 2007). Exploiting the biodiversity of such sources creates the chance to isolate new bioproducts of organisms, such as plants and microorganisms, or identify biosynthetic clusters of genes coding for previously non-studied natural compounds produced by known organisms (Wright and Sutherland, 2007).

Microorganisms represent the richest source of natural products that can be applied medically, veterinary or agriculturally, as therapeutics, pesticides or herbicides (Bérdy, 2005). The isolation of microorganisms from untapped ecological habitats might lead to the identification of novel microorganisms and bioproducts, some of which could have an antimicrobial activity that may help to compensate for the deficiency in the antibiotic pipeline (Wright and Sutherland, 2007). Therefore, the exploration of new places and the use of the microbial biodiversity to mine for new microbial bioproducts represent a potentially fruitful avenue. This may lead to the discovery of natural compound(s) that could constitute a solution for the problem in question (Chan et al., 2002).

The implementation of an integrative strategy that involves genomic and proteomic approaches is an essential factor for the success of the discovery process (Chan et al., 2002). Genomics and genome mining bioinformatic tools play a pivotal role in drug and other natural compound discovery. They contribute to the design of the right experimental approach needed to study the gene(s) or the bioproducts(s) of interest, and to identify the compound target(s) (Chan et al., 2002). Proteomic approaches are also needed to identify peptide bioproducts. Accordingly, a strategy that combines wet-lab experiments with dry-lab *in silico* analyses is more likely to construct a successful path that may lead to the discovery of a novel antimicrobial product.

1.1. Scope of the thesis

This thesis aims for exploring the potential of thermophilic bacteria as producers of antimicrobial substances. To fulfil this aim, water samples from hot springs in Jordan were collected and bacterial strains were isolated and screened for their antimicrobial activity. *Geobacillus* sp. ZGt-1 isolated from Zara hot spring was selected and research studies were carried out to answer different questions, as explained below.

Paper I deals with the isolation and molecular identification of *Geobacillus* sp. ZGt-1 and confirms its antagonistic activity against thermophilic and mesophilic strains. The study proved the proteinaceous nature of the antibacterial substances. The paper also presents a system that was developed for cultivating strain ZGt-1 to guarantee the production of the antibacterially active proteins. The expressed active proteins were identified using the mass spectrometry technique and bioinformatic tools. The paper concludes a list of potential antibacterially active proteins of *Geobacillus* sp. ZGt-1.

Paper II announces the genome sequence of *Geobacillus* sp. ZGt-1 and reveals the potential of the strain to produce bacteriocins. The paper suggests two putative bacteriocins of *Geobacillus* sp. ZGt-1.

Paper III deals with mining the genome sequences of firmicutes for the presence of class-I lanthipeptide-coding genes. The paper also pays a close attention to the class-I lanthipeptides of *Geobacillus* strains and their biosynthesis gene clusters, and highlights the presence of a gene coding for a putative novel lanthipeptide and suggests naming it Z-geobacillin of *Geobacillus* sp. ZGt-1. The paper concludes lists of putative novel lanthipeptide sequences of firmicutes and potential lanthipeptide-producing bacterial strains that have not been recognized previously as lanthipeptide producers.

Paper IV highlights the presence of the type II toxin-antitoxin (TA) system in *Geobacillus* strains, including *Geobacillus* sp. ZGt-1. Analysis of the genome sequences of the strains identified putative genes coding for various families of the type II TA system. The paper concludes a list of putative novel TAs and potentially new TA families, and indicates special features of the TAs of *Geobacillus* strains. The paper also proposes a hypothesis on the regulation of the gene expression of one of the TA families.

2. Natural antimicrobial products

"One touch of nature makes the whole world kin".

William Shakespeare

Natural products have been the major contributors to drugs in the history of medicine. Natural antimicrobial products are secondary metabolites that are produced by both macro- and micro- organisms (Hayek et al., 2013; Tiwari et al., 2009). Different antimicrobials produced by different species of plants, animals, fungi, and bacteria have been reported in the literature. Some of these antimicrobials have already been used in pharmaceutical and food industries as antibiotics and food biopreservatives (Hayek et al., 2013; Tiwari et al., 2009).

2.1. Antimicrobial agents of plants

Plants are constantly exposed to microbial infections; therefore, they produce antimicrobial peptides and metabolites as part of their defensive system (Sampedro and Valdivia, 2014). Flavonoids, alkaloids, terpenoids, phenolics, and plant steroids are examples on plant-derived compounds with antimicrobial properties (Saleem, 2014). Antimicrobials from fruits, vegetables, seeds, and essential oils in herbs and spices have been receiving scientific attention in the last 30 years (Hayek et al., 2013). Antimicrobials of plant origin are under research investigations and have not been commercially applied yet (Sampedro and Valdivia, 2014; Tiwari et al., 2009). They have biotechnological potential, and there are promising antimicrobials that could be applied in human medicine in the future, such as defensin peptides as antifungals (Sampedro and Valdivia, 2014).

2.2. Antimicrobial agents of animals

Animals also produce different antimicrobial agents that have evolved as part of their defense mechanisms (Tiwari et al., 2009). Many of the animal antimicrobial agents are peptides and proteins. One of the typical examples is the lysozyme,

which is a bacteriolytic enzyme, produced commercially from the egg-white of hens (Tiwari et al., 2009). It breaks down the bacterial cell wall and can be used for extending the shelf-life of different food products as it is active against a broad range of food spoilage microorganisms (Tiwari et al., 2009). Another typical example is the lactoperoxidase, an enzyme naturally found in milk and active against bacteria and fungi (Tiwari et al., 2009). Other than antimicrobial peptides, animals have antimicrobial lipids, such as milk lipids that are active against some bacteria (Tiwari et al., 2009). Some animals, such as crustaceans and arthropods have a polysaccharide in their exoskeleton called chitosan which also has antibacterial and antifungal activities (Tiwari et al., 2009). Antimicrobial compounds were also isolated from amphibians and different terrestrial vertebrates (Bérdy, 2005). However, the potential of the antimicrobial compounds of animal origin has not been exploited yet and investigating its industrial potential is still in its infancy (Tiwari et al., 2009).

2.3. Antimicrobial agents of bacteria and fungi

Bacteria and fungi, as a source of natural products, give better chances for a competent scale-up of the natural product research compared to plant and animal sources (Bérdy, 2005). Bacterial and fungal metabolites represent a rich source of potential new therapeutic drugs (Yarbrough et al., 1993).

Natural products derived from bacteria and fungi form 47% of all known bioactive natural products, and out of these products, 84% are antimicrobials (Piso, 2014)

In a variety of environments, fungi and bacteria coexist, and therefore, they compete and defend their existence by producing an array of different antimicrobial substances (Essig et al., 2014). Fungi constitute a valuable source of natural products, including antimicrobials (Awan et al., 2017). Many of the essential commercialized antibiotics are derived from fungal compounds (Awan et al., 2017). A famous one is the β-lactam antibiotic; penicillin, which is produced by some species of *Penicillium* isolated from corn, wheat, barley, flour, walnuts, and meat (Awan et al., 2017; Laich et al., 2002). Together with the successful clinical application of the bacterial gramicidin (see below), the clinical application of penicillin in human therapeutics in the early 1940s sat the official beginning of the Antibiotic Era (Villa and Veiga-Crespo, 2014). After that, the discovery of new antimicrobials from bacteria and fungi and applying them clinically proceeded but varied (Bérdy, 2005). From the early 1990s, the number of discovered antimicrobials of fungal origin was increasing continuously and reached to more than 50% by the year 2000, as compared to the discovered antimicrobials of bacterial origin during that period (Bérdy, 2005). By the year

2001, the number of newly discovered antimicrobials almost leveled off (Bérdy, 2005). At the clinical level, less than 1% of the discovered fungi-derived antimicrobials have been applied in individual therapy (Bérdy, 2005). Currently, fungi are again gaining more attention for their natural bioproducts and for the possibility of identifying novel antimicrobials to be clinically applied (Awan et al., 2017; Essig et al., 2014; Bérdy, 2005).

Bacteria produce a variety of compounds that are active against other microbes, including other competing bacterial strains (Tiwari et al., 2009). The most bountiful producers of antimicrobials have been actinomycetes. Actinomycetes are Gram-positive bacteria belonging to the phylum Actinobacteria. Within Actinobacteria, the most prominent producers have been *Streptomyces* species. *Streptomyces* species produce 39% of all microbial products, and 73% of those products have antimicrobial activity (Olano et al., 2014).

Bacterial products with antimicrobial activity include proteinaceous ones, such as bacteriocins, and lytic enzymes; such as bacterial amidases (Borysowski and Górski, 2009), and non-proteinaceous ones such as polyketides, organic acids, and hydrogen peroxide (Saleem, 2014; Tiwari et al., 2009). Non-proteinaceous antimicrobial bioproducts are beyond the scope of this thesis.

The first-ever clinically applied antibiotic was of bacterial origin, and it was released in 1939 (Kelkar and Chattopadhyay, 2007). It was gramicidin, which was produced by *Bacillus brevis* isolated from soil (reviewed by (Kelkar and Chattopadhyay, 2007)). Thanks to the clinical success of gramicidin, research on the clinical application of penicillin was keyed up (Kelkar and Chattopadhyay, 2007). The discovery of gramicidin was followed by the discovery and clinical application of streptomycin, chloramphenicol, tetracycline, and macrolides, all of which were produced from *Streptomyces* (Bérdy, 2005). More antibiotics were discovered then, and in the 1950s and 1960s, 70% of the discovered antibiotics were from *Streptomyces*. In the 1970s and 1980s, the importance of non-*Streptomyces* actinomycetes was flourishing as antibiotic-producers, since they contributed by 25-30% of all discovered antibiotics (Bérdy, 2005). However, the pharmaceutical interest in antibiotics derived from bacteria had only slightly increased in recent years (Bérdy, 2005).

In addition to the role of antimicrobial products as pharmaceuticals, they have potential in the food industry. Nisin, a bacteriocin that was isolated from lactic acid bacteria, is already applied in food as a biopreservative. Nisin is the only natural antimicrobial peptide licensed by the US Food and Drug Agency (FDA) to be used in food (Tiwari et al., 2009). However, since nisin has a deficient activity at neutral or alkaline pH, its applications in food are limited by the pH of the food product (Martirani et al., 2002). Therefore, there is a need for finding out new bacteriocins that can be active in food products under a broader range of conditions (Martirani et al., 2002). The mesophilic *Streptomyces* species have

been the most remarkable producers of antibiotics, as mentioned above, and the mesophilic lactic acid bacteria (LAB) are well known for their production of bacteriocins. Nonetheless, other bacteria, including the extremophilic bacterial species, have also been reported as producers of antimicrobial products, as discussed below.

2.3.1. Extremophilic microorganisms as a source of antimicrobials

Extremophiles are organisms that live and thrive at the extremes of life, such as living at temperature (> 45 °C or < 15 °C), pH (> pH 8.5 or < pH 5.0), pressure (> 500 atm), or salinity (> 1.0 M NaCl), or at any other extreme condition which does not support the survival of mesophilic organisms (Podar and Reysenbach, 2006). As such, those organisms thriving in unique ecosystems have unique metabolic pathways, and their enzymes are adapted to function under extreme conditions (Coker, 2016; Podar and Reysenbach, 2006). Isolation of microorganisms from such ecosystems, especially the untapped ones, may lead to the discovery of new microbes, which could represent a promising source of natural compounds with biotechnological potential (Bérdy, 2005).

Biotechnological applications of extremophiles may involve the organisms themselves, such as the case with bioleaching or, as is the case with most applications, involve their biomolecules, such as the enzymes of extremophiles (extremozymes) or any other peptide/protein (Podar and Reysenbach, 2006). The harsh conditions under which extremozymes are adapted to function are sometimes similar to the conditions of many industrial processes (Coker, 2016). Although only few extremophiles/extremozymes have been involved in large industrial-scale production (Coker, 2016), they form a multibillion-dollar industry covering biomedical, agricultural, and different industrial sectors (Podar and Reysenbach, 2006). The most reputable example of a profitable application of an enzyme isolated from a an extremophilic organism is the Taq DNA polymerase, which was isolated from the thermophilic bacterium *Thermus aquaticus* that was isolated from a geothermal spring in Yellowstone National Park (reviewed by (Podar and Reysenbach, 2006)).

Production of antimicrobials by extremophiles is not surprising since they, as all other organisms, fight for occupying a niche space and for gaining nutrients (Coker, 2016). However, the ability of extremophiles to produce antimicrobials has not been investigated as thoroughly as that of mesophiles. Recently, more attention has been directed towards antimicrobials produced by extremophiles. Some of the extremophilic archaea and bacteria have been described as producers of antimicrobial peptides, such as halocin produced by halophilic archaea (Coker, 2016), sulfolobicin produced by species of the thermophilic and acidophilic

archaeon *Sulfolobus* (Ellen et al., 2011), thermophilins produced by different thermophilic bacterial strains of *Streptococcus thermophilus* (reviewed by (Pranckute et al., 2015)), haloduracin produced by the alkaliphilic *Bacillus halodurans* C-125 (McClerren et al., 2006), bacteriocin-like substance produced by the thermophilic *Enterococcus faecalis* K-4 (Eguchi et al., 2014), and proteinaceous inhibitory compound produced by the psychrophilic *Pedobacter* sp. BG5 (Wong et al., 2011).

2.3.2. Antimicrobial peptides produced by thermophilic bacilli

Members of the *Bacillus* group *senso lato* are known for their production of diverse antimicrobial substances of different types. Those substances include different classes of the ribosomally-synthesized bacteriocins, as well as the non-ribosomally synthesized antimicrobial peptides (Abriouel et al., 2011). Some of the antimicrobial compounds produced by *Bacillus* spp. are already clinically applied, such as gramicidins, tyrocidines, and bacitracins (reviewed by (Esikova et al., 2002).

Due to their physiological properties, such as spore formation and production of antimicrobial peptides, and their growth requirements, bacilli can grow and survive in various ecosystems; in soil, aquatic environments, food, and in vegetation (Abriouel et al., 2011). They are also capable of growing and thriving under extreme conditions (Verma et al., 2018). Antimicrobial peptides produced by members of the *Bacillus* group *senso lato* are diverse with a variety of basic chemical structures (Abriouel et al., 2011). The capability of thermophilic bacilli to produce antimicrobial peptides has also been reported, as discussed below.

Antimicrobial peptides (bacteriocins) produced by bacilli can be considered as the second most important after the bacteriocins produced by LAB (Abriouel et al., 2011). Subtilin, lichenicidin, and paenibacillin are examples on bacteriocins produced by mesophilic strains of *B. subtilis*, *B. licheniformis*, and *Paenibacillus polymyxa*, respectively (Abriouel et al., 2011). Bacillocin 490 is an example of a bacteriocin produced by a thermophilic strain *B. licheniformis* 490/5 (Martirani et al., 2002). Another thermophilic strain of *B. licheniformis* has been reported as a producer of a bacteriocin-like substance (Abdel-Mohsein et al., 2011).

Since antimicrobial peptides produced by thermophilic bacterial strains, among which are the thermophilic bacilli, are expected to be thermostable at relatively high temperatures, this feature renders them as biopreservative candidates to protect heat-treated food products (Kaunietis et al., 2017). Therefore, the interest in thermophilic bacteria as a source of antimicrobial peptides is rising (Kaunietis et al., 2017). Moreover, some bacilli species are generally recognized as safe (GRAS) by the FDA, like LAB (Martirani et al.,

2002), and this feature as well grants their antibacterial peptides the potential to be involved in the food industry, as is the case with the nisin mentioned above.

From another perspective, some thermophilic bacilli are potential food-spoiling organisms in different industries; such as dairy production, canning, juice pasteurization, sugar refining, and other industries where steps of manufacturing processes take place at temperatures (40–65 °C) (Burgess et al., 2010). Among such food-contaminating bacilli are *Geobacillus* species (Burgess et al., 2010).

Geobacillus spp. can form biofilms on the surfaces of manufacturing equipment, and consequently contaminate the food product. For example, *Geobacillus stearothermophilus* is known for contaminating milk powder and low-acid canned food, causing spoilage of the final product due to the secreted bacterial enzymes and produced acids (André et al., 2013; Burgess et al., 2010).

In fact, food-contaminating thermophiles add to the interest in antimicrobial peptides of thermophilic bacteria. Antimicrobial peptides of thermophilic bacterial strains, more likely the ones that are phylogenetically closely related to the strains causing food-spoilage, constitute a potential solution to counteract the adverse effects of those food-contaminating bacteria. *Geobacillus* spp. themselves produce a variety of antimicrobial peptides that may be exploited in antagonizing the growth of the food-spoiling geobacilli (Chapter 3).

Throughout the process of screening for and isolating natural products in general, and antimicrobials in particular, genome mining tools play a significant role; as they further boost the probability of discovering novel bioproducts (Morton et al., 2015a), as discussed below.

2.4. The significance of bioinformatic tools in the discovery of antimicrobial proteins

Owing to the development of the next-generation sequencing (NGS) technology, the availability of whole-genome sequences deposited in the public databases is increasing exponentially. As a result, the *in silico* identification of potential bioactive peptides/proteins, including antimicrobial ones, has tremendously improved (Wright and Sutherland, 2007).

There are different freely available databases representing repositories of antimicrobial peptides/proteins, where one can retrieve the sequences, physicochemical properties, and biological effects of the protein in question (Torrent et al., 2012b). Such databases are either general ones that contain proteins of different types and origins (Torrent et al., 2012b), like the UniProt (The UniProt Consortium, 2017) (www.uniprot.org) and KEGG (Kanehisa and Goto, 2000) (www.kegg.jp), or specific ones that were developed especially for

depositing antimicrobial peptides (Torrent et al., 2012b), such as the bacteriocin databases, as the antiSMASH database (Blin et al., 2017a; Blin et al., 2018), and BAGEL4 (van Heel et al., 2018) (details about these tools are given in Chapter 6), the "Antimicrobial Peptide Database" (APD3) (Wang et al., 2016), and the "Collection of Anti-Microbial Peptides" ($CAMP_{R3}$) (Waghu et al., 2016).

Importantly, there are different freely available bioinformatic tools that have been developed for the detection of antimicrobial peptide-coding genes, and for the prediction of the antimicrobial potential of peptide and protein products. Such tools can also estimate the probability of the antimicrobial potential of the peptide/protein by using computational algorithms via a user-friendly interface (Torrent et al., 2012b). Therefore, they facilitate the selection of the peptide/protein of interest to proceed with by carrying out experimental studies that could ultimately lead to the discovery of a novel antimicrobial substance (Torrent et al., 2012b). Those algorithms can be part of a database (Torrent et al., 2012b), such as the prediction algorithms available in APD3 and $CAMP_{R3}$ databases, both of which can predict the antimicrobial potential of a given peptide/protein. The algorithms can also be available independent of a database, such as the antiSMASH mining tool, which annotates and analyzes biosynthesis gene clusters of secondary metabolites, including bacteriocins, in the genomes of bacteria, fungi, and plants (Blin et al., 2019) (Chapter 6). The antiSMASH is continuously updated, and the latest version, version 5.0, has been released earlier this year (Blin et al., 2019). Another genome mining tool that predicts the antimicrobial peptide-coding genes based on the genome context analysis, but is less comprehensive than the antiSMASH though, is BAGEL4 (van Heel et al., 2018). Additionally, there is also a prediction tool, known as AMPA (antimicrobial peptide/protein algorithm) (Torrent et al., 2012b) that predicts the antimicrobial potential of a given amino acid (aa) sequence.

In addition to the production of antimicrobial peptides/proteins, bacteria produce other proteins that have antimicrobial or "toxic" effects, as a part of the bacterial "toxin-antitoxin (TA) system", which may either have a bactericidal or bacteriostatic effect (Gerdes, 2013) The potential of producing such toxins can also be investigated using general databases such as KEGG (Kanehisa and Goto, 2000), and specific databases such as TADB (Toxin-Antitoxin Database) (Xie et al., 2018). The presence of genes coding for toxins in the genome can as well be predicted *in silico* using computational algorithms such as the TA finder tool (Xie et al., 2018) (details are given in Chapter 7).

The *in silico* prediction of the potential of the (micro)organism in question to produce antimicrobial compounds, based on its genome sequence analysis, aids in designing the approaches required for the process of isolating the compounds (Moloney, 2016). In other words, genome-based prediction of the antimicrobial potential minimizes the serendipity surrounding the isolation of the peptide/protein since screening for the antimicrobials starts with "reading" the

genetic map, identifying the target; i.e., the putative gene coding for the antimicrobial peptide/protein, and then going forward with the wet-lab experiments for the production and purification of the peptide/protein of interest. Screening genome sequences and detection of the genes coding for natural compounds is facilitated by the fact that the machinery biosynthetic genes coding for and regulating the expression of the peptide/protein of interest are usually clustered together (Wright and Sutherland, 2007).

2.5. The significance of mass spectrometry-based proteomics in the identification of antimicrobial proteins

Once the antimicrobial activity has been detected and the proteinaceous nature of the antimicrobial bioproducts(s) has been confirmed using laboratory assays (**Paper I**), the aa sequence and identity of the peptide/protein can be determined using the mass spectrometry technique.

During the 1990s, mass spectrometry (MS) techniques replaced Edman degradation and have become the primary method for protein identification (Steen and Mann, 2004). Compared to the Edman degradation, the MS is more sensitive as it detects peptides available in minute amounts (down to femtomoles) in complex mixtures of biomolecules (Baldwin, 2004; Steen and Mann, 2004). The mass spectrometer can fragment peptides in a matter of seconds, while the Edman degradation was taking hours or even days (Steen and Mann, 2004). The MS does not require each protein in the mixture to be purified entirely, and the mass spectrometer can identify even the acetylated or post-translationally modified proteins, which were not possible to identify using the Edman degradation (Steen and Mann, 2004). Therefore, The MS has become the main method in the field of proteomics (Steen and Mann, 2004).

After analyzing the peptides in the mass spectrometer, the peptide-mass spectra are retrieved. The peptide and the peptide fragmentation masses (resulting from the MS/MS, for example) will be searched against a protein sequence database, using specialized database-searching software, such as the Mascot search engine, the Sequest algorithm, and the InsPect, among others (Cottrell, 2011). The matching peptides will be selected and, this in turn, helps in inferring the matching proteins. As a result, the unknown protein(s) present in the complex sample that was analyzed using the MS is/are now identified (Edwards, 2011).

The protein sequence database represents the source of peptide sequences to be matched against the mass spectra. Therefore, selecting the right database influences the specificity, accuracy, and speed of the search (Edwards, 2011). If

the selected database is missing some peptide sequences, the mass spectra of those peptides will not be matched with peptide sequences in the database, and therefore, the spectra will not be identified, and their peptides will not be observed. Consequently, proteins will not be identified when the protein sequence database is poorly chosen (Edwards, 2011). On the other hand, if the selected database is more inclusive, the search will take a longer time and could result in more false positive peptide identifications and lower statistical significance of the matched peptides (Edwards, 2011).

The best protein sequence database is the organism-specific one. Sequencing the genome of the organism in question and identifying its putative coding genes and putative proteins has a significant impact on the success of the database search, and thus the identification of the protein(s) present in the analyzed sample (Cottrell, 2011). Having a small dataset of proteins that are specific for the organism in question minimizes the probability of getting peptide matches by chance or getting false positives, and raises the confidence that the matched peptide is correct, and the inferred protein is truly present in the sample, as compared to using a database composed of protein sequences of a closely-related species (Cottrell, 2011; Edwards, 2011). Accordingly, the genome sequence plays a vital role in the proteomic analysis stage of discovering antimicrobial peptides/proteins.

For the current study, the conventional whole-cell screening, cultivation, and production experiments were combined with advanced proteomic and bioinformatic analysis tools to study the antimicrobial potential of one of the geobacilli, designated as *Geobacillus* sp. strain ZGt-1, as will be discussed in the next chapters.

3. The *Geobacillus* genus and potential applications

"A picture is gradually emerging of a genus…"

(Hussein et al., 2015)

The genus *Bacillus* used to include diverse groups of rod-shaped, Gram-positive, and aerobic or facultatively anaerobic bacteria (Ash et al., 1991). Due to the heterogeneity of the members of *Bacillus* in terms of their growth conditions, nutritional demands, metabolism, and DNA composition (Ash et al., 1991), the genus underwent extensive taxonomical reclassification at the genus taxon level (Zeigler, 2001). With the exploration of extreme environments in the 1990s by microbiologists, novel extremophilic microorganisms, including thermophilies were found out in large numbers (Zeigler, 2001). Among such thermophiles are geobacilli, which were then classified as belonging to the genus *Bacillus*. In 1991, Ash and co-workers grouped *B. stearothermophilus, B. kaustophilus* and *B. thermoglucosidasius* in a phylogenetic cluster that was distinct from other bacilli, termed as "group 5" (Ash et al., 1991). In 1994, Rainey and co-workers added other *Bacillus* species; *B. thermoleovorans*, *B. thermocatenulatus*, *B. thermodenitrificans*, *B. caldotenax*, *B. caldovelox*, and *B. caldolyticus* to group 5 (Rainey et al., 1994). In 2001, members of "group 5" of bacilli received a new classification that distinguished them from other bacilli. The new classification placed those members under the new genus "*Geobacillus*" (Nazina et al., 2001).

The vegetative cells of *Geobacillus* are rod-shaped, spore-formers, forming one endospore per cell. Cells either form short chains or occur as single separate cells, and they can be motile or non-motile (Nazina et al., 2001). The cell wall structure is Gram-positive, but the cells are Gram-stain variable (Nazina et al., 2001). Geobacillli are either aerobic or facultatively anaerobic, and they are obligately thermophilic, with growth temperature ranging between 37–75 °C, with an optimum at 55–65 °C (Nazina et al., 2001). They are neutrophilic, where growth occurs in a pH range of 6.0 to 8.5, with an optimum at pH 6.2–7.5 (Nazina et al., 2001). Most *Geobacillus* spp. do not require growth factors, vitamins, NaCl, and KCl (Nazina et al., 2001). Members of *Geobacillus* are phenotypically

and phylogenetically closely-related thermophilic bacilli with a high 16S rRNA sequence similarity (98.5–99.2%) (reviewed by (McMullan et al., 2004)).

Geobacillus, literally means earth or soil small rods (Nazina et al., 2001), are widely distributed in nature as well as in artificial environments (McMullan et al., 2004). This widespread distribution reflects the high level of environmental adaptability of *Geobacillus* members (Suzuki et al., 2012). They have been isolated from different geothermal areas located in different continents (McMullan et al., 2004). *Geobacillus* strains were isolated from oil fields, such as isolates of *G. subterraneus*, *G. uzenensis*, and *G. thermodenitrificans*, isolated from oilfields in Russia, Kazakhstan and China (Nazina et al., 2001; Feng et al., 2007). They were also isolated from hot springs, such as a strain of *G. icigianus* (Bryanskaya et al., 2015) and another strain of *G. gargensis* (Nazina et al., 2001), both of which were isolated from hot springs in Russia (Bryanskaya et al., 2015; Nazina et al., 2001). *Geobacillus* isolates, such as *Geobacillus thermodenitrificans*, and *G. thermoleovorans* were also isolated from shallow marine hydrothermal vents of the Eolian Islands (Italy) (Maugeri et al., 2001). In a study by Takami and co-workers in 1997, deep-sea mud as well represented a source of *Geobacillus*; *G. kaustophilus* strain was isolated from deep-sea mud of Mariana Trench in Japan (Takami et al., 1997). *Geobacillus* was also isolated from artificial hot environments (McMullan et al., 2004), such as a domestic heating system water, from which a strain of *G. caldoxylosilyticus* was isolated (Obojska et al., 2002). *Geobacillus* strains were also isolated from temperate soil environments from geographically scattered regions in Europe (McMullan et al., 2004).

3.1. Industrial applications of *Geobacillus*

Geobacillus has attracted the attention for its industrial potential. For example, *Geobacillus* strains represent sources of a variety of thermostable enzymes demanded in different industrial applications, such as amylases, proteases, lipases, pullanases (reviewed by (McMullan et al., 2004)). Production of exopolysaccharides by *Geobacillus* is another example of the industrial potential of members of this bacterium (reviewed by (McMullan et al., 2004)). Additionally, some geobacilli have potential applications in environmental biotechnology (McMullan et al., 2004), such as bioremediation since some species are capable of hydrolyzing hydrocarbons (Bustard et al., 2002). Furthermore, the capability of some *Geobacillus* strains to metabolize herbicides has been reported; therefore, geobacilli have potential applications in the agricultural biotechnology (McMullan et al., 2004).

As mentioned in Chapter 2, geobacilli are potential antibiotic producers. However, only a few studies have investigated the antimicrobial potential of *Geobacillus* and proved it experimentally. Garg and co-workers reported the production of geobacillin I, a lanthipeptide that is nisin-analogue, by different strains of *G. thermodenitrificans* (Garg et al., 2012). Production of a bacteriocin, toebicin 218, and a bacteriocin-like inhibitory substance by *G. toebii* strains HBB-218 and HBB-247, respectively, was reported (Özdemir and Biyik, 2012b; Özdemir and Biyik, 2012a). Pokusaeva and co-workers reported the production of bacteriocins by *G. stearothermophilus* (Pokusaeva et al., 2009). Kaunietis and co-workers reported the production of a bacteriocin by *Geobacillus* sp. strain 15 (Kaunietis et al., 2017). Novotny and Perry, (1992) reported the production of bacteriocins substances, thermoleovorins, by *B. thermoleoverans*, which is now classified as *G. thermoleovorans* (Novotny and Perry, 1992). Shafia (1968) reported the production of bacteriocin-like substances, thermocins, by *Bacillus stearothermophilus* (Shafia, 1966), which is now classified as *G. stearothermophilus*. These studies indicated that *Geobacillus* is a promising taxon in terms of its antimicrobial potential.

It is noteworthy that the majority of those antimicrobial-producing *Geobacillus* strains were isolated either from soil or oil fields. In general, the soil has been and is still the most mined source for the discovery of natural bioproducts (Giudice and Renato, 2016). As mentioned in Chapter 2, microorganisms such as *Streptomyces*, producing antimicrobial substances that have reached the pharmaceutical markets were mainly isolated from soil (Bérdy, 2005). However, frequent isolation from the same habitat may result in the re-isolation of similar metabolites (Yarbrough et al., 1993). Therefore, changing the site of isolation aids in exploiting the chemical diversity of bioproducts in nature and this, in turn, raises the probability of isolating novel antimicrobial substances, as mentioned in Chapter 1 (Bérdy, 2005).

Antimicrobial substances produced by microorganisms living in aquatic ecosystems have not been exploited as much as those of terrestrial microorganisms (Pednekar et al., 2011). Therefore, sampling an aquatic ecological niche might lead to the discovery of novel metabolites with pharmaceutical and industrial potential.

In this research work, aquatic ecosystems were sampled (Chapter 4). *Geobacillus* sp. strain ZGt-1 was isolated and explored for its potential as a producer of antimicrobial peptides and proteins via two different approaches. One of them started with wet-lab experiments, where water samples were collected from hot springs and isolates were screened for their antibacterial activity, then the selected isolate, ZGt-1, was cultivated and the produced antibacterial proteins were subjected to the MS and bioinformatic analyses, using a set of specialized tools (**Paper I**; Chapter 4). The other approach relied on sequencing the genome of the selected isolate, followed by *in silico* analyses and

screening for putative bacteriocin-coding genes, as well as putative toxin-coding genes, using another set of specialized bioinformatic tools (**Paper II, Paper III, Paper IV**; Chapters 5–7).

By isolating the antimicrobial-producing bacterium, *Geobacillus* sp. strain ZGt-1, and sequencing its genome, we hope to contribute to the understanding of the antimicrobial potential of *Geobacillus* and the identification of the different antimicrobial substances produced by this taxon, especially since among all those isolated *Geobacillus* strains that are antimicrobially active, only *G. thermodenitrificans* strains NG80-2 and DSM465, whose antimicrobial potential was proved by (Garg et al., 2012), have their genomes sequenced, as inferred from the genome sequences deposited in the NCBI till October, 2019.

In the next chapter, the isolation and cultivation of *Geobacillus* sp. ZGt-1, together with its putative antimicrobial proteins, are explained.

4. *Geobacillus* sp. ZGt-1 and its antimicrobial protein candidates

"If you don't like bacteria, you're on the wrong planet."

Stewart Brand

Microorganisms have developed their own strategies to survive and thrive, and the production of antimicrobial substances is one of these strategies, as mentioned above (Hibbing et al., 2010; Singh et al., 2012a). This microbial behaviour can be exploited for the benefit of human beings, but we also need to have our own strategies to exploit it in the best way possible. Selecting the ecological niche to be sampled is a crucial factor that may control the success of our strategies (Chan et al., 2002).

Ecological systems with special environmental conditions are expected to sustain microorganisms with special strategies that support them throughout the competition for space and resources with the other microorganisms. Terrestrial hot springs are one of these special ecosystems.

The microbial communities are diverse and plentiful in hot springs (Des Marais and Walter, 2019), while nutrients are limited. These conditions, together with the high temperature that may fluctuate over wide ranges, and the fluctuating pH (Shah and Garrett, 2013) impose challenges on the inhabiting microorganisms; thus, they are expected to have special survival and thrival strategies. Accordingly, selecting hot springs for isolating microorganisms and probing their natural products could be a good approach. Additionally, as mentioned above, it is more likely to find novel products in unexplored ecological niches (Chan et al., 2002), and since antimicrobial substances produced by microorganisms living in hot springs have not been well-exploited (Pednekar et al., 2011), selecting hot springs for isolating novel microorganisms with putatively novel antimicrobial agents represents a promising approach. This principle has been employed in this thesis, where hot springs in Jordan were sampled.

Jordan, a country located in the Middle East with an area of about 90,000 km^2, has significant differences in the climate across its regions (Sawarieh, 2008). That is due to a sharp variation in its topography over a small country area (Sawarieh,

2008). Jordan has a number of hot springs that exist along the eastern Jordan Rift Valley (Salameh, 1986). Jordan Rift Valley is the northern fault of the Dead Sea Rift, which trends nearly north–south of Jordan and has arid climate with a hot summer and a warm winter (Sawarieh, 2008). The geothermal activity in Jordan is squeezed out entirely in the form of hot springs (Sawarieh, 2008). The thermal water flows to empty into the Dead Sea (Salameh, 1986).

In the current research study, water samples were individually collected from different hot springs. The sampled hot springs were Afra, Deir Alla, Ma'een, and Zara, located in different geographical regions across Jordan. The explored bacterial isolate, *Geobacillus* sp. ZGt-1, was isolated from Zara hot spring, as described below.

4.1. Zara hot spring

Zara hot spring (32°N 36°E) is located about 50 km southwest of the capital city, Amman (Figure 1) (Salameh, 1986). It is a rather untapped ecological niche; thus, it represents a promising source of novel microorganisms with pharmaceutical and industrial potential.

Figure 1
Map of Jordan showing the location of Zara hot spring. Adapted from https://atlastours.net/jordan/

The thermal water at the depths of 2.5-3.5 km reaches a maximum temperature of 110 °C and is cooled down by mixing during its upward movement till it reaches the land surface and overflows (Figure 2) (reviewed by (Eraifej, 2006)). The water temperature at the time of sampling was 46 °C, and the pH was around 7. The isolation of bacterial strains is described below.

Figure 2
Zara hot spring (Photo taken by the author using Kodak Easy Share v530).

4.2. Isolation of bacteria

Water samples were collected from the four hot springs mentioned above, on different days. The samples were transferred to the lab in an insulator box. Upon arrival to the lab, plates with solid R2A (Reasoner's 2A) medium were inoculated with the water samples and incubated aerobically at 60 °C for two days. After two days of incubating the inoculated plates, mixed cultures of bacteria had grown. The bacterial colonies were then isolated and subjected to repeated streaking onto new R2A agar plates, followed by overnight incubation at 60 °C. This procedure resulted in 59 bacterial isolates from the four hot springs.

The isolates included *Anoxybacillus* and *Geobacillus* strains, as indicated by the 16S ribosomal ribonucleic acid (rRNA) gene sequencing (Section 4.3). *Geobacillus* sp. strain ZGt-1 and *Geobacillus stearothermophilus* strain 10 were selected for the research studies related to this thesis. Both strains, ZGt-1 and 10, were isolated from Zara hot spring.

4.3. Identification of the isolates

The isolates were identified at the molecular level by amplifying and sequencing the 16S rRNA gene. The first step towards eventually identifying the isolates was the extraction of the DNA from pure bacterial cultures, as explained briefly below.

4.3.1. DNA extraction

The DNA was extracted from pure cultures of strains ZGt-1 and 10 using a specific kit, the ZR Fungal/Bacterial DNA MiniPrep kit (Zymo Research, Orange, CA, USA) (**Paper I, Paper II**). This kit involves effective physical cell lysis, where mechanical disruption by bead beating takes place to break the cells for extracting the DNA. The mechanical lysis was then followed by DNA purification using column binding and different chemical reagents. Procuring the extracted DNA paved the way for downstream applications, such as the polymerase chain reaction (PCR) amplification of the 16S rRNA gene and sequencing it (**Paper I**), as well as sequencing of the ZGt-1 genome (**Paper II,** Chapter 5).

4.3.2. PCR amplification and sequencing of the 16S rRNA gene

After the DNA extraction, the 16S rRNA gene of each isolate was PCR-amplified (**Paper I**). The PCR protocol and cycle conditions were optimized to overcome the difficulty of amplifying high GC content sequences. The rRNAs of thermophiles have a higher GC content than those of mesophiles (Galtier et al., 1999). The GC content of 16S rRNA genes and the DNA melting temperature are strongly correlated (Kimura et al., 2006). The high GC content could hinder the full separation of the double-stranded DNA (dsDNA) helix and could cause the formation of secondary structures (Bhagya et al., 2013; Shore and Paul, 2010). These problems cause the DNA polymerase to stall and impede the primer annealing (Bhagya et al., 2013; Shore and Paul, 2010). Consequently, they result in the absence or truncation of PCR-amplified products (Bhagya et al., 2013; Shore and Paul, 2010). Therefore, GC-rich sequences are seen as “stubborn” sequences that demand special PCR reaction conditions (Bhagya et al., 2013). Accordingly, throughout optimizing the PCR reaction in the present study, different troubleshooting measures were considered. The reagents used for the PCR reaction and their concentrations, as well as the reaction conditions, were optimized successively until the 16S rRNA gene was successfully amplified. The PCR protocol and cycle conditions are described in **Paper I**.

After confirming the specific amplification of the 16S rRNA gene, the PCR products were purified and then sequenced by the ABI sequencing reaction (GATC Biotech, Konstanz, Germany) (Paper I). The sequencing of the 16S rRNA genes of the isolates was carried out in the forward and reverse directions. The length of the sequenced 16S rRNA gene of strain ZGt-1 is 1454 bp, and the GC content is 59.49%. For strain 10, the length of the sequenced 16S rRNA gene is 1415, and the GC content is 59.64%.

4.3.3. Identities of the isolates

The interpreted nucleotide sequences of the 16S rRNA genes were compared against sequences available in the GenBank by applying the BLASTn 2.3.1+, using the Megablast option on the RefSeq_RNA database (NCBI Transcript Reference Sequences). The two strains, "ZGt-1" and "10", showed to affiliate to the genus *Geobacillus*, as they showed > 99.5% identity to *Geobacillus* (**Paper I**). Strain ZGt-1 could not be assigned to a certain species (**Paper I**); therefore, it has been designated as *Geobacillus* sp. ZGt-1. On the other hand, strain 10 showed to belong to *G. stearothermophilus*; accordingly, it has been designated as *G. stearothermophilus* 10 (**Paper I**). The focus was set on *Geobacillus* sp. ZGt-1 as the potential producer of antimicrobial agents, as explained below.

4.4. Antibacterial activity of *Geobacillus* sp. ZGt-1

Testing strain ZGt-1 for its antibacterial activity was carried out by employing the known agar-deferred spot method, as described in **Paper I**. It is noteworthy here that the culture medium used for carrying out this test was Mueller Hinton (MH). MH medium was selected because it is the medium recommended by the Clinical and Laboratory Standards Institute (CLSI) for different routine susceptibility tests (Hudzicki, 2009), and it is a non-selective and non-differential medium (Merck, 2009). The brand we used (Merck) sates that the MH agar has low concentrations of thymine and thymidine. This is important because otherwise, the action of some antimicrobial agents will be inhibited. This brand also states that the levels of calcium and magnesium ions are appropriately adjusted. This is also important because the level of these cations affects the expansion of the inhibition zones (D'Amatot and Thornsberry, 1979).

For the test strains, *G. stearothermophilus* strain 10 was selected and some mesophilic strains as well (see below). The isolation of a strain of *G. stearothermophilus*, a known food-spoiling bacterium, and another closely related bacterial strain, *Geobacillus* sp. ZGt-1, from the same ecological niche of

Zara hot spring was intriguing. It triggered the interest in benefitting from the competition among closely related microorganisms inhabiting the same niche and competing for the same resources, in an attempt to define a strategy for antagonizing the growth of *G. stearothermophilus*. *G. stearothermophilus* causes spoilage of dairy products and low acid canned and ready-made vegetable- and meat-based meals (Viedma et al., 2009). Thus, finding an antimicrobial agent that inhibits its growth will present potential solutions to the food industry.

The *Geobacillus* sp. ZGt-1 was tested for antagonizing the growth of *G. stearothermophilus* strain 10. The results indicated that *Geobacillus* sp. ZGt-1 was active against *G. stearothermophilus* strain 10 (Figure 3, and Figure 1 in **Paper I**). This result was in line with our expectations since both strains were isolated from the same niche, as mentioned above.

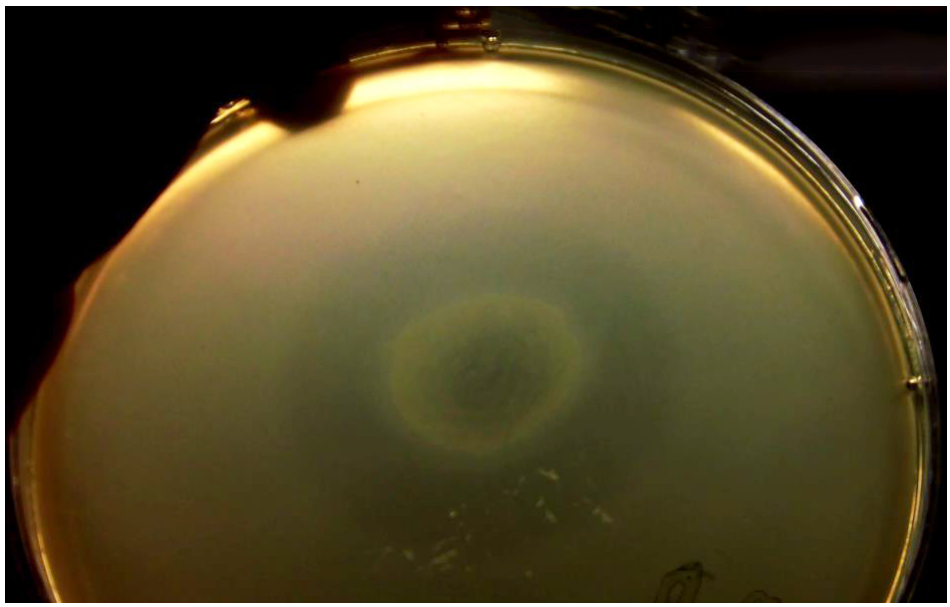


Figure 3
Antibacterial activity of *Geobacillus* sp. strain ZGt-1 against *G. stearothermophilus* strain 10.The test was done by applying the agar-deferred spot method, as explained in Paper I.

Interestingly, *Geobacillus* sp. ZGt-1 was also active, to different extents, against mesophilic bacterial strains; *Bacillus subtilis* TMB94 and the pathogenic *Salmonella enterica* serovar Typhimurium (*Salmonella* Typhimurium) CCUG 31969 (Figure 1 in **Paper I**). The differences in activity against these two strains could be ascribed to the differences in the cell wall structure between

Gram-positive and Gram-negative bacteria, as explained in **Paper I**. On the other hand,
Geobacillus sp. ZGt-1 was not active against *E. coli* 1005, *Staphylococcus aureus* NCTC 83254, *Staphylococcus epidermidis* TMB96, and *P. vulgaris* TMB02.

4.5. Production of the antibacterial substances by sequential recycling of immobilized cells of *Geobacillus* sp. ZGt-1

The major limitation for studying the antibacterial substances produced by strain ZGt-1 was correlated with the cultivation method (**Paper I**). The strain could produce the substances only when cultivated on MH agar medium, while when the cultivation was carried out in MH broth, no antibacterial activity was detected in the culture supernatant. The lack of the production of antimicrobial substances in liquid culture was not unexpected since other studies have reported this issue (Nilsen et al., 2003; Danesh et al., 2011)

To overcome this problem, cells were immobilized by entrapment in agar beads, suspended in MH broth, cultivated aerobically at 60 °C in batch fermentation mode, and recycled in sequential batches, where each batch was cultivated for 22–25 h, and with a total of 25 batches (**Paper I**). Upon the end of each batch, the culture was centrifuged, the supernatant was collected and saved, and fresh MH broth was added to the beads to start the next batch (Figure 4).

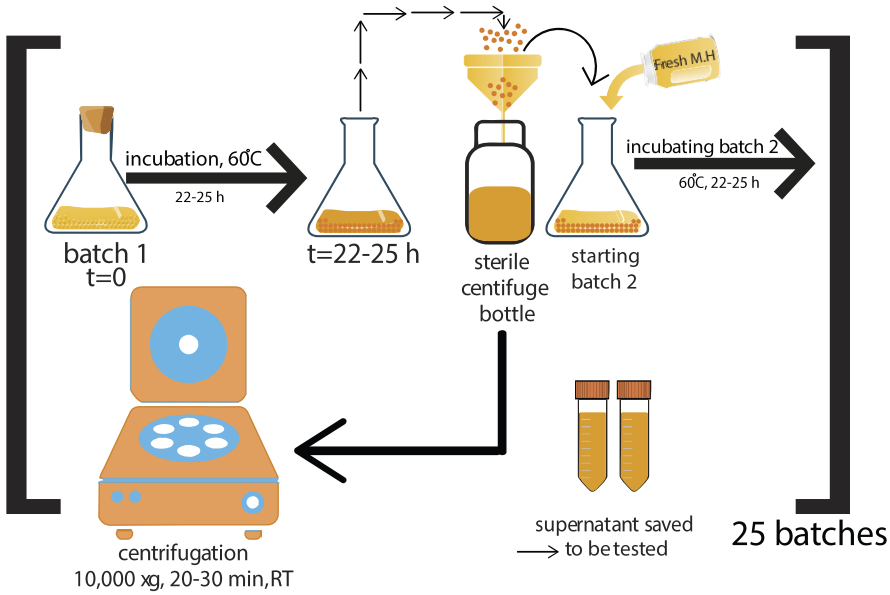


Figure 4
Schematic illustration of the steps of the sequential batch cultivation of the immobilized cells of *Geobacillus* sp. ZGt-1 for the production of the antibacterial substaces.

The cell-free supernatant collected at the end of each cultivation batch was tested for its antibacterial activity against *G. stearothermophilus* strain 10 by employing the spot-on-lawn approach. The results indicated that the antibacterial activity of strain ZGt-1 was increasing gradually over the recycled batches up to the 14th batch (Figure 2 (a–d) in **Paper I**). This increase in the antibacterial activity can be ascribed to the increase in the cell density over the cycles. Immobilization of cells increases the cells density (mass per unit volume) (Naghmouchi et al., 2008; Bertrand et al., 2001), and intensifies cell-to-cell contact and therefore, immobilized cells are more active at producing specific metabolites than those produced by free cells (Pilkington, 2005). Moreover, immobilized cells show increased tolerance to inhibitory metabolic products that might limit cell growth (Doleyres and Lacroix, 2005; Pilkington, 2005). As a result, high cell density can be achieved, and thus more metabolites, including antimicrobial substances, could be produced.

The antibacterial activity, however, started to gradually decrease after the 14th cycle until it almost disappeared by the end of the 25th cycle (Figure 2 (d–e) in **Paper I**). This decrease can be due to the mass transfer limitation problem because the diameter of the cell-agar beads of immobilized cells increases over

time due to the increase in cell density. Therefore, the internal transfer of the substrate to the entrapped cells is hampered over time (Salmon and Robertson, 1987). Consequently, it could be that the immobilized cells started going into starvation that changed their metabolic activity over time, and accordingly, the production of the antimicrobial substances was affected. Moreover, the repeated use of the beads with increased cell density created an increase in the level of free cells which were spontaneously being released from the surface of the beads over the cycles and were accumulating in the culture broth (Doleyres and Lacroix, 2005). Nutrients were readily available to those free cells as they were not subjected to the mass transfer limitation problem faced by the immobilized cells. Eventually, free cells which are already not efficient enough in metabolite production (Pilkington, 2005), outcompeted the immobilized cells.

Alternatively, this gradual decrease in the detected antimicrobial activity could be ascribed to difficulties with exporting the antimicrobial substances from the immobilized cells to the extracellular environment.

4.6. Purification of the antibacterial proteins

As explained in **Paper I**, treatment with proteinase K showed that the antibacterial substances were of proteinaceous nature. Therefore, ammonium sulphate precipitation (60% salt saturation) was applied for purifying the proteins from the cell-free supernatant collected from the sequential batches of the immobilized ZGt-1 cells. The protein precipitate was dialyzed against distilled water. The antibacterial activity of the salt-free protein precipitate was confirmed (Figure S1 in **Paper I)**. The activity was stable after heating the precipitate at 70 °C for 45 min but was lost when heated to 80 °C for 10 min (Figure S2, **Paper I**).

4.7. Identification of the antibacterial proteins

The precipitated proteins were resolved on polyacrylamide gel using the sodium dodecyl sulphate-polyacrylamide gel electrophoresis (SDS-PAGE). The protein bands associated with the antibacterial activity against *G. stearothermophilus* strain 10 were identified by applying the method of Bhunia and co-workers (Bhunia et al., 1987) (Scheme 1 in **Paper I**). The protein bands that displayed the antibacterial activity *G. stearothermophilus* 10, corresponding to 15–20 kDa (Figure 3 in **Paper I**, and Figure S3 in **Paper I**), were then excised from the gel,

trypsin-digested, and analyzed using the liquid chromatography tandem mass spectrometry (LC-MS/MS).

It is noteworthy here that the antibacterial proteins showed activity even after being subjected to SDS under reducing conditions using the dithiothreitol (DTT). Protein stability is correlated with SDS-resistance; the abundance of SDS-resistant proteins in thermophilic microorganisms is more than that in the mesophilic ones (Xia et al., 2010). Motivated by that, thermophiles might produce SDS-resistant, and duly, stable antimicrobial proteins as a lasting defence strategy that helps them thrive in the harsh conditions. Moreover, having stable proteins helps them save energy and nutrient resources, which would otherwise be dissipated on protein degradation and synthesis (Xia et al., 2010). Additionally, several antimicrobial peptides lack cysteine residues, and thus do not form disulfide bridges (Brogden, 2005). Therefore, they are insensitive to the reducing agent, DTT, used for disrupting the disulfide bonds in the SDS-PAGE. This could be the case with the antibacterial proteins of ZGt-1.

In parallel to conducting these experiments, the genome of *Geobacillus* sp. ZGt-1 was sequenced and annotated (GenBank accession no. LDPD00000000.1) (**Paper II**, Chapter 5). This helped in constructing a local database of the putative proteins of ZGt-1. It was then possible for the obtained MS/MS data to be searched specifically against the protein database of ZGt-1. This customized search reduced the probability of having false protein hits, since the best protein database is that of the strain itself, as discussed in Chapter 2. As a result, protein fragments detected by the mass spectrometer were identified.

4.7.1. Uncharacterized proteins with antibacterial potential

To identify the potential antibacterial protein candidates among the proteins detected by the MS/MS and identified by the search against the ZGT-1 local protein database, the protein molecular mass was used as a criterion. Since the SDS-PAGE analysis showed that the antibacterial proteins have a molecular mass within the range of (15–20) kDa, proteins whose mass is within the range of (10–30) kDa were considered in order to take into account the "gel shifting" phenomenon, where the observed molecular mass of the protein in the gel deviates from the formula molecular mass, a behaviour known about SDS-resistant protein (Manning and Colon, 2004; Rath et al., 2009; Rath et al., 2010). The proteins with a molecular mass of 10–30 kDa were 22 proteins and are listed in Table 1 in **Paper I**. The presence of three uncharacterized/hypothetical proteins in this list was alluring.

These three proteins of ZGt-1; 6_35, 23_543, and 4_4 were analyzed using two approaches applied in parallel, in a non-mutually exclusive way, to predict their antibacterial activity (**Paper I**). These approaches involved calculating and

comparing the physiochemical properties of these proteins (Table 1) to those of known antimicrobial peptides/proteins, and *in silico* predicting their antibacterial activity using the web-based antimicrobial peptides/proteins prediction algorithms (Table 1), in a similar way to the approach followed by (Bishop et al., 2015; Dziuba and Dziuba, 2014) with some modifications. The tools used were the AMPA (Torrent et al., 2012a), $CAMP_{R3}$ (Waghu et al., 2016), and APD3 (Wang et al., 2016) (Chapter 2).

The analysis indicated that protein 23_543 is the most likely antibacterial protein candidate, as it fulfilled all the antimicrobial potential parameters inferred from the physicochemical properties and prediction algorithms (Table 1). On the other hand, protein 6_35 fulfilled all the physicochemical properties associated with the antimicrobial potential, but the majority, not all, of the prediction algorithms predicted its potential (Table 1). Protein 4_4 fulfilled most of the physicochemical properties associated with the antimicrobial potential, and the majority of the prediction algorithms predicted its potential (Table 1).

Accordingly, although protein 23_543 is the most likely antibacterial protein candidate, proteins 6_35 and 4_4 are also possible candidates (Paper I).

Table 1.
Prediction of the antimicrobial potential of the uncharacterized proteins based on their physicochemical properties and algorithm models (From Paper I).

Physicochemical properties				
Property		**Protein Query ID**		
		6_35	**23_543**	**4_4**
Length		129	153	173
Molecular weight (kDa)		13.8927	16.8564	18.979.1
Net charge		+2	+2	+1
pI		8.80	8.61	7.72
Instability index		14.04	17.61	36.7
Aliphatic index		80.08	110.33	96.76
GRAVY index		-0.044	-0.257	-0.253
Boman index (kcal/mol)		1.12	1.57	1.19
Na^4vSS		5.3	-6.6	-2.4
Number of aggregation hot spot regions		3	6	6
Total hydrophobic ratio		40%	39%	36%
Potential of forming amphipathic helix		Yes	Yes	Yes
Number of hydrophobic residues on the same side		≥ 38	≥ 42	≥ 31
Algorithm models				
$CAMP_{R3}$ Models	SVM [§]	1.000 [1]	1.000 [1]	1.000 [1]
	RF [¤]	0.987 [1]	0.9575 [1]	0.991 [1]
	DA [#]	1.000 [1]	1.000 [1]	1.000 [1]
	ANN [†]	NAMP [2]	AMP [3]	NAMP [2]
APD3		AMP [3]	AMP [3]	AMP [3]
AMPA		NAMP [2]	AMP [3] (0.86) [1]	NAMP [2]
Summary of the fulfilled antimicrobial potential parameters inferred from the physicochemical properties and prediction algorithms				
Physicochemical properties		+ [4]	+ [4]	+/- [5]
Prediction algorithms		+/- [6]	+ [7]	+/- [5]

[1] Probability of being an antimicrobial peptide/protein.
[2] Non-antimicrobial peptide/protein.
[3] Antimicrobial peptide/protein.
[4] All the physicochemical properties required for AMPs are fulfilled by the protein.
[5] Majority of the physicochemical properties required for AMPs are fulfilled by the protein.
[6] Majority of the six used algorithms predicted the protein as antimicrobial.
[§] Machine learning algorithm; SVM stands for Support Vector Machines.
[¤] Machine learning algorithm; RF stands for Random Forest.
[#] Machine learning algorithm; DA stands for Discriminant Analysis.
[†] Machine learning algorithm; ANN stands for Artificial Neural Network.

4.7.2. Enzybiotics

As explained in **Paper I**, proteins larger than 20 kDa were also identified by the mass spectrometer. Lytic bacterial cell wall hydrolases (BCWHs) of ZGt-1 were among these detected proteins.

Antimicrobial properties of the lytic BCWHs, termed as "enzybiotics", have been gaining a special interest lately. Attributing to their bactericidal mode of action, i.e., cleavage of the cell wall peptidoglycan network at specific sites, and the low probability of developing bacterial resistance, enzybiotics represent potential antibiotic alternatives and food biopreservative (Parisien et al., 2008; Villa and Veiga-Crespo, 2010).

Bacteria constitute one source of enzybiotics besides bacteriophages; however, limited attention has been given to the enzybiotics of bacterial origin (Villa and Veiga-Crespo, 2010).

BCWHs of bacterial origin are involved in bacterial cell growth and division (Villa and Veiga-Crespo, 2010; Wyckoff et al., 2012). Moreover, they act as antibacterials by attacking the cell wall of competing bacteria leading to bacteriolysis (Parisien et al., 2008; Wyckoff et al., 2012). Among the interesting enzybiotics are the N-acetylmuramoyl-L-alanine amidase (referred to as amidase; EC:3.5.1.28) and serine-type D-alanyl-D-alanine carboxypeptidase (also known as DD-carboxypeptidase; EC:3.4.16.4).

Amidases cleave the amide bond between the glycan strands and the stem peptides of the peptidoglycan (PG) layer of the bacterial cell wall; thereby, separating the junction between the polysaccharides and the peptide moieties in the PG network (Ghuysen and Hakenbeck, 1994; Salazar and Asenjo, 2007). DD-carboxypeptidases remove the terminal D-alanine in the stem peptide of the PG layer (Ghuysen and Hakenbeck, 1994). Thus, these enzymes impair the integrity of the bacterial cell wall (Ghuysen and Hakenbeck, 1994). Therefore, they represent potential antimicrobial agents.

The mass spectrometer detected significant peptide sequences of the amidase and DD-carboxypeptidase of strain ZGt-1. The detected amidase peptide sequences corresponded to the catalytic domain of the enzyme. Those of the DD-carboxypeptidase corresponded to the catalytic domain, the supposed enzyme's binding domain, and parts of the region in between the two domains.

It cannot be confirmed whether or not these two enzybiotics were responsible for the detected antibacterial activity in the 15–20 kDa region. However, it is interesting to study the activity of these enzymes experimentally, especially because there is a lack of experimental data on the enzybiotics of thermophilic bacteria.

Overall, the results indicated that *Geobacillus* sp. ZGt-1 antagonized the growth of *G. stearothermophilus*, *B. subtilis*, and the pathogenic *S.* Typhimurium strains. They also indicated the efficiency of combining the immobilized cell technology with the cell-recycling to produce the antimicrobial proteins. The proteins produced against *G. stearothermophilus* 10 were SDS-resistant and active even when heated at 70 ºC. Three uncharacterized/hypothetical proteins of ZGt-1, proteins 23_543, 6_6, and 4_4, seemed to be associated with this antibacterial activity. Two enzybiotics, amidase and DD-carboxypeptidase of ZGt-1 might have played a role in this antagonizing activity. Furthermore, the antibacterial activity might have been the result of a synergistic action among the proteins.

5. Genome sequence of *Geobacillus* sp. ZGt-1

"Genome sequencing has changed taxonomy".

Richard Dawkins

As mentioned in Chapter 2, for the identification of the potential of a given (micro)organism to produce bioactive natural products, the genome sequence has become of great significance in guiding the identification and the downstream processing (Tracanna et al., 2017; Blin et al., 2019).

Most bacterial genomes harbor a high proportion of various bioactive natural products (Tracanna et al., 2017). Accordingly, mining such genomes may lead to the discovery of novel compounds, such as antimicrobials (Tracanna et al., 2017; Blin et al., 2019).

Analyzing the genome sequence using specialized bioinformatic tools that aid in selective mining for the product(s) of interest, provides an important dataset that helps in identifying the target(s); e.g., the putative antimicrobial coding gene(s), and the biosynthetic pathway(s) (Tracanna et al., 2017; Blin et al., 2019). These data; thus, provide a basis for conducting the research and save time and resources that could otherwise be spent while searching for potential antimicrobial products without knowing the genome sequence (Tracanna et al., 2017). After defining the potential target and its putative characteristics, the experimental design can be plotted, considering the obtained data.

In simple words, sequencing the genome and analyzing it act as the compass that directs future research work.

As explained in the previous chapter, the current research study started with the whole-cell screening approach and identification of the strains using the 16S rRNA gene, and the selection of the ZGt-1 isolate as a potential producer of antimicrobials. In parallel to the experimental work (**Paper I**), the genome of *Geobacillus* sp. ZGt-1 was sequenced in order to identify the potential antibacterial proteins, as explained in **Paper I** and Chapter 4, and to also explore whether or not strain ZGt-1 has other potential antimicrobial substances, such as bacteriocins.

Obtaining a good quality and quantity of the DNA is crucial for the success of the genome sequencing process (Dang et al., 2016). These criteria were fulfilled using the kit mentioned in Chapter 4, **Paper I** and **Paper II**. The DNA sample of strain ZGt-1 was then sent to the Wellcome Trust Centre for Human Genetics for sequencing.

The next-generation sequencing (NGS) of *Geobacillus* sp. ZGt-1 genome, assembly of the reads, and scaffolding them based on the genome sequence of *Geobacillus kaustophilus* HTA426 (NC_006510) resulted in 241 scaffolds and a genome size of ~3,5 million bp. The number of the protein-coding genes was predicted to be 3,546, which was close to the number of genes reported for strain HTA426, 3,397 genes, at the time of assembling the genome of strain ZGt-1 (**Paper II**). The features of the genome of the latter strain are summarized in Table 2.

Table 2
Features of the genome sequence of *Geobacillus* sp. ZGt-1 (LDPD01000000).

Aspect	Value
Genome size	3,483,107 bp
No. of protein-coding genes	3,546
Gene density	88%
GC content	52.2%

The assembled draft genome sequence of strain ZGt-1 was analyzed for genes coding for bacteriocins, including lanthipeptides, using the specialized bioinformatic tools, antiSMASH and BAGEL (**Paper II** and **Paper III**; this chapter and Chapter 6). As indicated in **Paper II**, the genome analysis showed that strain ZGt-1 harbors putative biosynthetic gene clusters for a lanthionine-containing bacteriocin, i.e., a lanthipeptide, and a non-lanthionine containing one (**Paper II**). The putative lanthipeptide is the focus of **Paper III.**

Moreover, analyzing the assembled genome sequence of the strain using the TA finder tool, which is specialized in identifying toxins of the type II toxin-antitoxin system (Xie et al., 2018), putative toxins that could represent potential antibacterial agents were identified (**Paper IV**; Chapter 7).

Overall, thanks to the genome sequence of strain ZGt-1 and the specialized and comprehensive software packages, like the antiSMASH and the TA finder tools, more of the antimicrobial potential of *Geobacillus* sp. ZGt-1 has been revealed, as we identified putative bacteriocins (**Paper II** and **Paper III**) and toxins (**Paper IV**), all of which have potential antimicrobial activities.

In the next chapters, the lanthipeptides (Chapter 6) and the toxins of the toxin-antitoxin system (Chapter 7) are explained in depth.

6. Bacteriocins with focus on lanthipeptides

"Bacteria represent the world's greatest success story."

Stephen Jay Gould

Production of antimicrobial peptides is presumed to be the oldest and most widespread defense strategy developed by the different living organisms from insects to plants to humans, against microorganisms (Cotter et al., 2005a). Microorganisms themselves also produce antimicrobial peptides against other competing microorganisms (Cotter et al., 2005a).

Bacteriocins are ribosomally synthesized antimicrobial peptides secreted by one bacterium to act against other bacterial members of the same species, in case of narrow spectrum bacteriocins, or against bacteria belonging to other genera, in case of wide spectrum bacteriocins. Production of bacteriocins is widely-spread among prokaryotes; it has actually been proposed that between 30–99% of the bacteria and archaea secrete at least one bacteriocin (Cotter et al., 2005a).

In 1925, the first report on the antagonistic effect of bacteriocins was described for *E. coli*, where different strains antagonized the growth of each other due to the secretion of antimicrobial peptides termed as "colicins" (Cotter et al., 2005a). Over time, the term "bacteriocins" was developed as a generic name for all kinds of gene-encoded antimicrobial peptides secreted by bacteria (Cotter et al., 2005a).

6.1. Classification of Bacteriocins

Different classifications have been proposed to classify the bacteriocins, as they include a group of different peptides and proteins with an inhibitory activity that can be sub-grouped under different categories (Cotter et al., 2005b). The classification proposed by Cotter and co-workers in 2005 groups the bacteriocins into two classes: the lanthionine-containing bacteriocins (i.e., lanthipeptides) (class-I), and the non-lanthionine-containing bacteriocins (class-II) (Cotter et al., 2005b). This classification categorizes the large, heat-labile peptidoglycan hydrolases (formerly, class-III bacteriocins) under a separate group designated as "bacteriolysins" (Cotter et al., 2005b). The current study (**Paper III**) focuses on class-I bacteriocins, lanthipeptides, as discussed below.

6.1.1. Class-I bacteriocins– Lanthipeptides

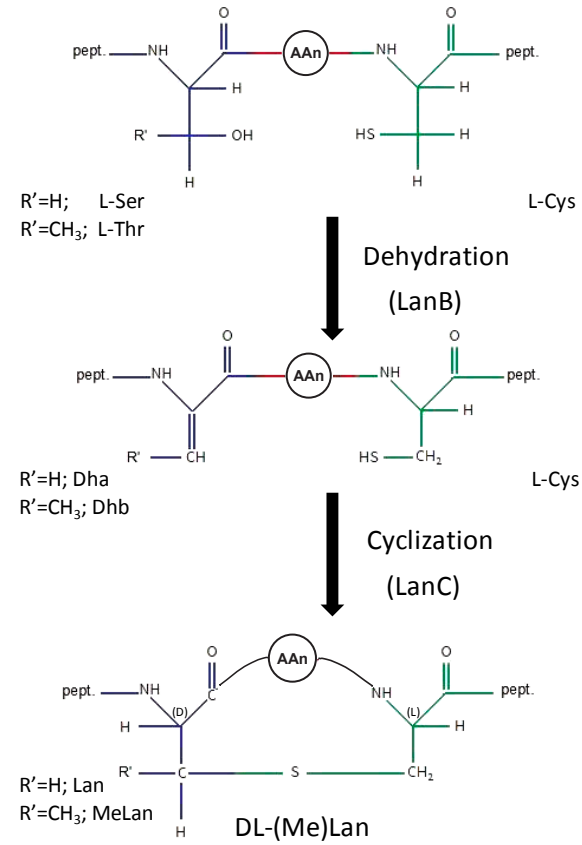
Lanthipeptides (lanthionine-containing peptides) are polycyclic ribosomally-synthesized peptides that undergo post-translational modifications resulting in the formation of unusual aa residues; dehydrated and (methyl-) lanthionine amino acids (Willey and van der Donk, 2007). This group was formerly known as lantibiotics for lanthionine-containing antibiotics. However, the term was broadened to include all lanthionine-containing peptides regardless of their antimicrobial activity (Willey and van der Donk, 2007).

Lanthipeptides with antimicrobial activity, called lantibiotics, have been gaining attention as antimicrobial agents against pathogens. They are produced by bacteria that are generally regarded as safe (GRAS), food-grade bacteria, and also by pathogenic bacterial strains themselves (Daly et al., 2010). The most famous lanthipeptide and the only marketed bacteriocin is nisin, the lantibiotic produced by the food-grade bacterium *Lactococcus lactis*. It is used as a preservative in food products to antagonize the growth of food pathogens (Willey and van der Donk, 2007).

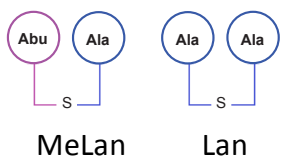
A lanthipeptide is a small peptide, coded by a structural gene, initially synthesized as a linear precursor LanA peptide containing a leader peptide and a core peptide. LanA peptide undergoes proteolytic cleavage to remove the leader peptide, and in turn, the core peptide eventually becomes the mature and active lanthipeptide (Knerr and van der Donk, 2012). Specific aa residues of the lanthipeptide undergo extensive post-translational modifications which are common to all lanthipeptides. These modifications include dehydration of certain serine and threonine residues yielding dehydroalanine (Dha) and dehydrobutyrine (Dhb), respectively (Knerr and van der Donk, 2012) (Figure 5). The dehydration reaction is followed by cyclization of the peptide, where the sulfhydryl groups of cysteine residues are added onto the Dha or Dhb to form a

lanthionine (Lan) or methyllanthionine (MeLan) bridge, respectively (Knerr and van der Donk, 2012) (Figure 5). These post-translational modifications are catalyzed by specific enzymes, as discussed in the next section.

After dehydration and cyclization, the modified peptide undergoes proteolytic cleavage to remove the leader peptide. This step takes place either inside the bacterial cell or in the exterior environment after exporting the modified peptide via transporters (McAuliffe et al., 2001; Chen and Hoover, 2003; Knerr and van der Donk, 2012). Removal of the leader peptide is critical for the activation of the antimicrobial function of the lanthipeptide (McAuliffe et al., 2001; Knerr and van der Donk, 2012).



(a)



(b)

Figure 5.
Post-translational modification reactions in lanthipeptides (a) Ser and Thr residues are dehydrated leading to the formation of Dha and Dhb residues, respectively. Dha and Dhb then undergo cyclization where the thiols of Cys residues are added onto Dha and Dhb, leading to the formation of (Me)Lan residues. AAn = The peptide sequence of a number of aa residues extending between Ser/Thr (Dha/Dhb) and Cys residues. Pept.= The rest of the lanthipeptide sequence. Adapted from (Repka et al., 2017; McAuliffe et al., 2001). (b) Symbolic illustration of the diamino acids comprising (Me)Lan residues. Abu=2-aminobutyric acaid. Adapted from (Molloy et al., 2012). Figures (a) and (b) were generated using ChemDraw.

6.2. Classification of lanthipeptides

As mentioned above, lanthipeptides represent one class of bacteriocins. Lanthipeptides themselves are also classified into different classes based on the enzymes involved in their biosynthesis. The current classification of lanthipeptides demonstrates four classes (Knerr and van der Donk, 2012). For class-I lanthipeptides, dehydration and cyclization reactions are carried out by two separate enzymes. A dedicated dehydratase (LanB) catalyzes the dehydration of serine and threonine residues, and a cyclase (LanC) catalyzes the cyclization (Arnison et al., 2013). For classes II, III, and IV lanthipeptides, dehydration and cyclization are carried out by bifunctional lanthionine synthetases; enzyme LanM for class-II, LanKC for class-III, and LanL for class-IV (Arnison et al., 2013).

In the presented study (**Paper III**), we focused on class-I lanthipeptides since this is the type of lanthipeptides that is harbored on the genome of *Geobacillus* sp. strain ZGt-1.

6.3. Mode of action of lanthipeptides

The majority of class-I and class II lanthipeptides are lantibiotics that are active against Gram-positive bacteria since their cell walls are more accessible than the Gram-negative ones (Barbosa et al., 2015). However, some strains of *E. coli*, *Helicobacter pylori*, and *Neisseria* were shown to be affected when applying high concentrations of some lantibiotics (reviewed by (Barbosa et al., 2015)). The mechanism of action of lantibiotics is believed to involve binding to lipid II, and/or pore formation (Barbosa et al., 2015; Knerr and van der Donk, 2012). Lipid II is an essential component in the bacterial cell wall since it is the monomeric peptidoglycan precursor (Barbosa et al., 2015; Brötz et al., 1998). Therefore, lantibiotics can inhibit the biosynthesis of the cell wall by binding to lipid II, and thus, interfere with the bacterial cell growth (Barbosa et al., 2015; Knerr and van der Donk, 2012). Lantibiotics can also form stable pores and disrupt the cell membrane, leading to cell death (Barbosa et al., 2015; Knerr and van der Donk, 2012). Nisin, for example, exerts both mechanisms, while mersacidin only binds to lipid II (Knerr and van der Donk, 2012). On the other hand, it seems that the mechanism of action of certain lantibiotics does not follow either of the two aforementioned common mechanisms. Cinnamycin is one of these exceptions. Cinnamycin binds specifically to the phosphatidylethanolamine (PE), one of the three main phospholipids in the bacterial cell membrane, and by doing so, it disrupts the membrane (Epand et al., 2016).

The mechanism of action of lantibiotics still needs to be further investigated in order to understand the interaction between the lantibiotics and the target cells (Barbosa et al., 2015).

6.4. Applications of lanthipeptides

With the emergence of antibiotic-resistant bacteria, lanthipeptides with antibacterial activity, i.e., lantibiotics, represent potential alternatives to conventional antibiotics. Moreover, lantibiotics represent promising food preservatives as they antagonize the growth of food-spoiling and food pathogenic microorganisms (Willey and van der Donk, 2007; Cotter et al., 2005b).

Lantibiotics have interesting features that render them attractive alternatives to the currently available antibiotics (Barbosa et al., 2015). Some lantibiotics have been proven active against pathogenic bacteria, including the drug-resistant ones, such as the methicillin-resistant *Staphylococcus aureus* (MRSA) and vancomycin-resistant *Enterococcus* (VRE). Some lantibiotics showed their inhibitory activity at nanomolar levels, while antimicrobial cationic peptides produced by eukaryotic organisms showed their inhibitory activity against pathogenic microorganisms in the micromolar concentration range (Cotter et al., 2005a). Moreover, lantibiotics have low toxicity to mammals (reviewed in (Barbosa et al., 2015)).

Previous studies revealed the significant potential of a number of lantibiotics as therapeutics. For example, in addition to its role as food biopreservative, nisin which is active against a wide range of Gram-positive and Gram-negative bacteria, has potential medical applications. It showed activity against the MRSA, VRE, and *S. pneumonia*, including penicillin-resistant strains, and it proved to be more active than vancomycin when tested against a clinical isolate of *S. pneumonia* (Cotter et al., 2005a). Additionally, nisin combined with the bacteriolytic enzyme, lysostaphin cured animals infected with mastitis (Cotter et al., 2005a). Nisin could also help in the prevention of gingivitis, plaque formation, and tooth loss (Cotter et al., 2005a). Another lantibiotic with therapeutic potential is microbisporicin (or NAI-107), which is produced by *Microbispora corallina*. It showed activity against MRSA and VRE and displayed a high potential in the treatment of nosocomial infections (Castiglione et al., 2008). Mersacidin, produced by a *Bacillus* strain and actagardine, produced by *Actinoplanes liguriae*, proved their potential in the treatment of the MRSA and VRE (Willey and van der Donk, 2007; Cotter et al., 2005a). Mutacin 1140, produced by *Streptococcus mutans*, has the potential to treat dental caries and streptococcal throat infections (reviewed in (Barbosa et al., 2015)). Moreover, the lantibiotics gallidermin, produced by *Staphylococcus gallinarum* and lacticin

3147, produced by *L. lactis* represent potential antibacterial therapeutics for acne as they showed activity against *Propionibacterium acne* (Cotter et al., 2005a).

Lanthipeptides have also been reported as having antifungal and antiviral activities, and as antinociceptive and antiallodynic (reviewed by (Repka et al., 2017)).

Lantibiotics can be used in cosmetics, such as deodorants, especially gallidermin since it is stable at the skin pH (5.4) and is active against a narrow range of bacteria, making it a specific antibacterial agent, which in turn reduces the possible side effects, such as the inhibition of the normal skin flora (Cotter et al., 2005a).

It can be clearly seen that research results are continuously providing indications on the capacity of lanthipeptides for offering at least part of the solution to the antibiotic-resistance problem and can be used as food biopreservatives, in addition to the other applications mentioned above. Thus, although nisin is the only commercialized lanthipeptide so far, more lanthipeptides are expected to attract the industrial interests within the pharmaceutical and food sectors (Cotter et al., 2005a).

In addition to the potential applications of lanthipeptides, the modifying enzymes may also have applications in peptide engineering as they could be employed for introducing modified aa into non-lanthipeptides in order to convert them into lanthipeptides (Cotter et al., 2005a).

6.5. Lanthipeptides and pathogenicity

Lanthipeptides are produced by some pathogenic bacterial strains as well, including antibiotic-resistant ones, such as *Staphylococcus*, *Streptococcus*, and *Enterococcus* strains as reviewed by (Daly et al., 2012). The antibacterially-active lanthipeptides (i.e., lantibiotics) grant the producing bacterium the benefit of out-competing the lantibiotic-sensitive microbes, which are living within the same microbial consortium and occupying an ecological niche; thus, they can establish the infection (Daly et al., 2012; Daly et al., 2010). In other words, lantibiotics produced by pathogens are proposed to promote pathogenicity (Daly et al., 2012; Daly et al., 2010). Accordingly, genome mining for lanthipeptides and identifying pathogenic bacterial strains with putative lantibiotics has a significant clinical value as it would help in illustrating the possible basis of virulence as a first step towards halting it (Daly et al., 2012).

There are several examples of lantibiotics produced by pathogens (reviewed by (Daly et al., 2012)). One such lantibiotic is BsaA2 (bacteriocin of *Staphylococcus aureus*) produced by strains of this species including MRSA strains (Daly et al., 2012; Daly et al., 2010). The antibacterial activity of BsaA2

has been experimentally verified, and it showed a broad antibacterial spectrum (Daly et al., 2010). Therefore, it is believed that BsaA2 plays a strong role in eliminating competing bacteria and thus helping *S. aureus* in occupying an ecological niche, which then aids in enhancing its pathogenicity (Daly et al., 2012; Daly et al., 2010). Accordingly, knowing the potential of the pathogenic *S. aureus* to produce BsaA2 is of significance since strategic measures can then be taken to control the bacterial invasion (Daly et al., 2012; Daly et al., 2010). It could be possible, for example, to block the production of BsaA2, or introduce non-pathogenic lantibiotic-producing strains that can out-compete *S. aureus* (Daly et al., 2010).

Correspondingly, revealing the potential of pathogenic bacterial strains to produce a lantibiotic can be a vital step towards controlling its competitiveness and thus its pathogenicity.

Our analysis indicated that each of the *S. aureus* strains (11819-97, Bmb9393, COL, MSSA476, MW2, NCTC 8325, Newman, T0131, TW20, USA300 FPR3757, USA300_TCH1516, VC40, and Z172) harbors a BsaA2-coding gene (**Paper III**), and some of these strains were also reported as potential BsaA2-producers by (Daly et al., 2010).

6.6. Culture-based or computer-based identification of bacteriocins?

Before the onset of the genome sequencing era and the development of bioinformatic analysis tools, screening for bacteriocins, in general, relied only on traditional culture-based screening (i.e., wet-lab work), as was the case with screening for any natural compound. The lab work involved isolation of microorganisms, assaying their activity of interest under different growth conditions; it is the antimicrobial activity in the case of bacteriocins, production under optimized conditions, purification of the compound of interest, and characterization of the compound using different analytical methods (Egan et al., 2018). This sequence of laborious steps was based on trial and error, which in turn led to time-consuming and costly processes (Egan et al., 2018).

With the advancement of rapid genome sequencing methods and specialized bioinformatic analysis tools, *in silico* prediction (i.e., dry-lab work) of the potential of a certain bacterial strain to produce bacteriocins in general, and lanthipeptides in particular has become feasible. It has facilitated the discovery process as it reduced the required time and cost tremendously (Egan et al., 2018).

For the *in silico* prediction of lanthipeptide production potential of a given bacterial strain whose genome sequence is already publicly available, one can refer to the genome sequence and mine it for the lanthipeptide-coding gene(s)

using the relevant bioinformatic analysis tools. Upon the identification of the putative lanthipeptide and evaluation of its novelty, a bacterial candidate can then be nominated for further work if it represents a promising subject. Afterwards, validation of the *in silico* prediction follows by applying wet-lab experiments, such as heterologous expression for the production and purification of the lanthipeptide (Egan et al., 2018). The wet-lab experiments should also include antimicrobial activity assays in order to assess if the produced peptide is a lantibiotic or merely a lanthipeptide.

On the other hand, the research could start some stages earlier where wet-lab experiments are done first and followed by *in silico* screening for the gene responsible for the detected antimicrobial activity. To raise the probability of isolating novel lanthipeptides, special ecosystems such as extreme habitats; hot springs for example, could be sampled (Chan et al., 2002). Isolation of bacterial strains and screening for their antimicrobial activity follow. Bacterial candidates can then be selected for genome sequencing. This will be followed by mining the genome for the genes coding for lanthipeptide(s) and identifying them. Afterwards, experimental validation; production and downstream processing of the lanthipeptide(s) take place.

The researcher needs to bear in mind that at the culture-based stage, where bacterial strains are screened for their antimicrobial activity, the results of this activity could include some false negatives. i.e., strains that harbor lanthipeptide-coding genes but the genes were not expressed under the selected lab conditions. As a result, some potential bacterial strains could be overlooked. The crucial role of genome analysis becomes pronounced in such cases as it can inform the researcher of the potential of the strain so as to minimize the issue of disregarding it.

Performing lab-experiments is inevitable for the production and downstream processing of the lanthipeptide of interest. However, the question is whether a researcher makes use of the already-available genome sequence data and starts directly at the *in silico* prediction stage and then proceeds with wet-lab experiments, or starts from scratch at the culture-based stage aiming for finding a novel potential microorganism, then proceeds with the *in silico* screening for the genes of interest, and follows with wet-lab experiments. In either case, the *in silico* analysis facilitates identifying the potential lanthipeptide(s) among all the other compounds produced by a given bacterial strain, helping in "finding the needle in the haystack"; thus reducing the trial and error-based procedures. It also constitutes the "road map" for designing the right methodology for the production and purification of the lanthipeptide of interest.

6.7. Bioinformatic-based discovery of novel bacteriocins

Screening for bacteriocin-producing bacteria using conventional microbiology assays, where the producer strain is tested for its inhibitory activity, is a difficult and time-consuming approach. It is controlled by different factors; such as growth conditions needed for the production of bacteriocins, detection assays, the concentration of the produced inhibitory substance, type of the indicator strains and their growth requirements, which are needed for confirming the production of bacteriocins and for demonstrating their inhibitory effect (de Jong et al., 2011). The traditional screening approach involves a comprehensive screening design, where the potential bacteriocin-producing strains are tested under a number of potential production conditions. Therefore, this design imposes challenges on the screening for and identification of potential novel bacteriocins (de Jong et al., 2011). Accordingly, following an approach that helps in limiting the number of the screened potential bacteriocin-producing strains constitutes an efficient step in the search for novel bacteriocins (de Jong et al., 2011).

Genome mining for bacteriocins, in which a genome sequence is analyzed, and the bacteriocin biosynthetic gene clusters are predicted, represents a potent approach. The continually increasing number of microbial genome sequence data deposited in the public databases, and the fact that bacteriocins are gene-encoded make bacteriocins good candidates for direct genome mining (Morton et al., 2015a; de Jong et al., 2011). This approach helps in revealing the bacterial strains which harbor the genes coding for those ribosomally-synthesized antimicrobial peptides, among others. Therefore, it aids in prioritizing the list of strains to be investigated in the lab. Accordingly, genome mining saves the time and effort needed for the discovery of novel bacteriocins (Cox et al., 2014).

Bacteriocins can be identified *in silico* from the genome based on either homology with already-identified bacteriocins, bacteriocin motifs, or based on the genes coding for the bacteriocin biosynthesis enzymes that are usually encoded in the proximity of the putative bacteriocin precursor-coding gene (Morton et al., 2015a; de Jong et al., 2011). However, screening for potential bacteriocins based on homology to the previously identified ones, using sequence homology algorithms like BLAST, has its own limitations (Morton et al., 2015a). It can help in identifying large bacteriocins (longer than 200 aa residues) because their sequences are more conserved than the small ones. However, due to the large sequence variability of bacteriocins and also the very small length of some of them (shorter than 30 aa residues), homology-based identification does not necessarily succeed in recognizing them (de Jong et al., 2011; Drissi et al., 2015). Moreover, the number of identified bacteriocins deposited in databases is relatively low compared to the significant ecological role of bacteriocins in

microbial communities (Drissi et al., 2015). Therefore, homology-based identification, which refers to the bacteriocins deposited in the databases, is likely to miss out on identifying potential novel bacteriocins that have no or low homology to the deposited ones (de Jong et al., 2011; Drissi et al., 2015). Screening for potential bacteriocins based on conserved protein motifs is more likely to succeed in identifying novel bacteriocins. However, since the aa sequences are poorly conserved among bacteriocins, identification based on the conserved motifs could also miss out on detecting potential novel bacteriocins (de Jong et al., 2011).

Conclusively bacteriocin detection methods that are based on homology with already identified bacteriocins or on conserved bacteriocin motifs are dependent on the already reported bacteriocins. They can recognize bacteriocins that are either homologous to the known ones or contain a conserved protein motif associated with known bacteriocins (de Jong et al., 2011). Therefore, mining bacterial genome sequences for the discovery of potential novel bacteriocins with low homology to known ones requires following a different approach. This approach relies on screening the genome context, as discussed below.

6.7.1. Screening genome context

Since bacteriocins represent a diversified group of peptides that are usually encoded by small and poorly conserved genes, mining the genome for them is a challenging task. This challenge can be overcome through screening genome context (de Jong et al., 2006).

Screening the genome context approach involves a comprehensive scanning for the different genes encoded within the bacteriocin gene cluster (Morton et al., 2015a; de Jong et al., 2011). The gene coding for a bacteriocin precursor peptide is usually positioned in the vicinity of its biosynthetic machinery genes, which have been found to be conserved across species (Morton et al., 2015a; de Jong et al., 2011). Therefore, the approach is based on searching for the biosynthetic machinery genes coding for the bacteriocin-synthesizing enzymes, genes involved in regulation, transport, processing of bacteriocins, or genes coding for the immunity proteins (de Jong et al., 2011).

Contrary to the genes coding for bacteriocin precursor-peptides, the other machinery genes in proximity to the coding genes are conserved across the species. This, in turn, facilitates the recognition of the ORF coding for the bacteriocin precursor peptide (de Jong et al., 2011). Screening genome context helps in identifying the unknown and non-conserved bacteriocins and the small ORFs that code for short bacteriocins, which are usually overlooked by the genome annotation tools, (Morton et al., 2015a; de Jong et al., 2011). Thus, it supports the identification of novel bacteriocins and the determination of their

class (de Jong et al., 2011). Accordingly, this method aids in the progress of the experimental work, where the approach needed for the production and purification of the potential novel bacteriocin could be designed based on the information gained from screening the genome context.

6.7.1.1. Tools for genome context screening

There are web-based tools currently available for mining the bacterial genome sequence for secondary metabolites. Among those tools are antiSMASH and BAGEL software packages, both of which allow the detection of bacteriocins encoded on the bacterial genome via annotating the genome, analyzing it, and recognizing the bacteriocin precursor-coding gene and the associated machinery genes surrounding it. Therefore, those tools lead to the automated identification of the bacteriocin biosynthetic gene clusters, and by doing so, they shorten the time and minimize the efforts of researchers interested in bacteriocin discovery (Medema et al., 2011; de Jong et al., 2011). However, antiSMASH is more advanced than the BAGEL tool.

6.7.1.1.1. antiSMASH

antiSMASH (antibiotics & Secondary Metabolite Analysis Shell) is a software pipeline for the identification of a variety of secondary metabolite biosynthetic gene clusters. It offers rapid and user-friendly analysis of the genome data input, and it is available as a web-based tool on a web server (http://antismash.secondarymetabolites.org/), and also as a stand-alone version for Linux/Unix operating system on a standard desktop computer (Medema et al., 2011).

The antiSMASH accepts individual whole genome sequences of different organisms, including bacteria, as input. It screens for gene clusters of different secondary metabolites, including bacteriocins in general and lanthipeptides in particular (Medema et al., 2011).

Since metabolites incorporate known enzyme families in their biosynthesis pathway, and the detection of those enzymes makes the identification of unknown novel metabolites possible, antiSMASH integrates the ClusterFinder algorithm, which allows analysis of the genome context leading to the detection of biosynthetic gene clusters of previously unknown secondary metabolites (Weber et al., 2015). In other words, antiSMASH analyzes the genome data input and grants *de novo* computing results (Blin et al., 2017b).

antiSMASH for the detection of lanthipeptide biosynthetic gene cluster

Detection of the lanthipeptide biosynthesis gene clusters and the lanthipeptide post-translational modifications is important for the discovery of natural products and the search for novel alternatives for antibiotics.

The antiSMASH makes use of the characteristics of ribosomally-synthesized and post-translationally modified peptides for detecting and identifying them on the queried genome. Such characteristics are (i) their relatively small size biosynthetic gene clusters, and (ii) their modification enzymes are typically coded close to the precursor peptides (Tietz et al., 2017).

The antiSMASH applies a lanthipeptide-specific analysis module for the detection of lanthipeptide biosynthetic gene clusters and the prediction of the post-translational modifications of the lanthipeptide (Blin et al., 2014). Starting from antiSMASH 2.2, this software has been offering the unique advantage of running a detailed analysis of the identified lanthipeptide gene cluster (Blin et al., 2014). In addition to the detection of the gene coding for the lanthipeptide precursor, the output generated by the antiSMASH analysis provides the researcher with the final lanthipeptide product and its possible post-translational modifications, the protease cleavage sites for cleaving the leader peptide, the biosynthetic pathway, the class of the lanthipeptide based on the neighboring genes coding for the biosynthetic enzymes, and homology to known lanthipeptide clusters (Blin et al., 2014). The output also provides the possible (Me-) Lan bridges and molecular masses of the mature lanthipeptide based on the tailoring modifications which the detected lanthipeptide could undergo (Blin et al., 2014). That detailed cluster information facilitates plotting the experimental design for the required downstream specific analysis. Moreover, since the antiSMASH output also presents the homology between the detected lanthipeptide and other known lanthipeptide clusters, it helps in de-replicating lanthipeptide discovery (Weber et al., 2015). Therefore, antiSMASH represents comprehensive software and its features make it the software of choice compared to other lanthipeptide genome mining tools (Blin et al., 2014).

At the time of preparing for **Paper III**, the antiSMASH 4 (Blin et al., 2017b) was the latest updated version. The algorithm employed by this version is from the genome-mining platform "Rapid ORF Description and Evaluation Online" (RODEO), which is employed to overcome the challenge of detecting the small and highly sequence variable ORFs coding for precursor lanthipeptides (Blin et al., 2017b). RODEO helps in the detection of small non-conserved ORFs positioned in intergenic regions that have been overlooked by annotation tools (Blin et al., 2017b). It uses a combination of heuristic scoring, the machine-learning algorithm, and motif analysis (Tietz et al., 2017). This robust combination offers an accurate and confident detection of precursor lanthipeptide(s) in the predicted gene cluster (Tietz et al., 2017). The principle of RODEO is based on rapid analysis and assessment of the genome context, which in turn helps in the identification of biosynthetic gene clusters and genes coding for ribosomally-synthesized and post-translationally modified peptides, especially class-I lanthipeptides (RODEO, 2018).

As mentioned earlier, antiSMASH is regularly updated. Currently, antiSMASH 5.0 is available, where even further comprehensive analyses can be carried out with improved runtime and upgraded visual representation (Blin et al., 2019).

6.7.1.1.2. antiSMASH database

The antiSMASH database is a comprehensive and easy-to-use resource of a collection of annotated biosynthetic gene clusters of different types of microbial secondary metabolites (Blin et al., 2017a). It represents a repository of pre-calculated antiSMASH results for the biosynthetic gene clusters of all the complete microbial genome sequences that are publicly available in the NCBI GenBank database (Blin et al., 2017a). The user can drive a simple search by browsing the database either by microorganism taxonomy or by the type of the secondary metabolite cluster (Blin et al., 2017a). An advanced search can also be tailored, where a query builder can be used to customize the search based on the user's requests (Blin et al., 2017a). The antiSMASH database was published in 2017 and is updated regularly, where all entries are re-analyzed using the latest version of antiSMASH (Blin et al., 2017a). The latest version of the database, version 2, is now available (Blin et al., 2018). This version contains a bigger dataset as it has annotations for 6200 full bacterial genomes and 18,576 bacterial draft genomes (Blin et al., 2018). Moreover, the options of the search and data export and the user interface have further improved (Blin et al., 2018).

6.7.1.1.3. BAGEL

BAGEL (BActeriocin GEnome mining tool) is a web-based software tool specialized in genome mining for bacteriocins (de Jong et al., 2006). It was the first fully automated software developed for the identification of new bacteriocin gene clusters, and it is a fast and user-friendly tool (de Jong et al., 2006). It combines different identification approaches; the search for homology with described bacteriocins in a local extensive bacteriocin database constructed for BAGEL and screening for bacteriocin motif together with screening the genome context and deducing the putative bacteriocin based on the biosynthetic machinery genes coded in the vicinity (de Jong et al., 2006). To avoid overlooking the small bacteriocin-coding ORFs which are neglected in many genome annotations, BAGEL uses a set of ORF prediction tools that help in annotating the genome data independently of GenBank annotations (de Jong et al., 2006). This results in the annotation of the genes in the genome context and the detection of the bacteriocin ORFs (van Heel et al., 2013). After a number of steps carried out "internally" by BAGEL, the output is generated (van Heel et al., 2013). The output provides the researcher with the putative bacteriocin biosynthetic gene cluster, including the class of the bacteriocin precursor peptide

and its leader peptide (de Jong et al., 2011). The output generated by BAGEL is less comprehensive compared to that of antiSMASH, since BAGEL does not predict post-translational modifications, the number of thioether bridges, or the molecular mass of the putative bacteriocin (Blin et al., 2014). BAGEL is continuously updated but at a slower pace compared to the antiSMASH, and the latest version BAGEL4 was released in the web late 2017 (unpublished yet).

In addition to the genome mining tool offered by BAGEL, it is possible for the user to only run a BLAST search. The aa sequence of the protein of interest can be blasted against any of the three bacteriocin databases in order to find homologous bacteriocins, if there is any. This helps in knowing whether the bacteriocin of interest is identical or homologous to any of the already known bacteriocins or it represents a potentially novel one.

6.7.1.1.4. BAGEL bacteriocin databases

Bacteriocin databases constructed for BAGEL webserver compiles known bacteriocin sequences that have been reported for more than one incidence. The sequences deposited in the database were retrieved from different other databases: (i) the NCBI server; (ii) the UniProt server; and (iii) the SRS server of ExPasY. Additionally, the literature search and expertise of the authors of BAGEL software were also applied to enrich the database (de Jong et al., 2011). The bacteriocin database encompasses three different databases containing small (<10 kDa) modified bacteriocins (lanthipeptides), small unmodified bacteriocins, and antimicrobial proteins >10 kDa (van Heel et al., 2013). An updated version of the databases (BAGEL4) was released in 2018.

6.8. Genome mining for class-I lanthipeptides

Due to the proved benefits of lanthipeptides as antimicrobials, together with the feasibility of genome sequencing and analysis, few studies were carried out to reveal the potential of bacterial strains as lanthipeptide producers. However, the only comprehensive study that performed *in silico* prediction of class-I lanthipeptides was published in 2010 by Marsh and co-workers (Marsh et al., 2010). The exponential increase in the publicly accessible genome sequence data and the availability of continuously upgraded relevant bioinformatic tools facilitate and validate conducting more genome-mining studies. As mentioned above, gathering different data about a putative lanthipeptide and its biosynthetic gene cluster, and the producing strain helps in avoiding the problem of re-discovery. Thus, it tremendously saves time that could otherwise be spent on wet-lab experiments (Sandiford, 2014).

In **Paper III**, we focused on the identification of class-I lanthipeptides coded on the genome sequences of firmicutes, with a special attention given to the class-I lanthipeptide, which we named as Z-geobacillin, that is coded on the genome of *Geobacillus* sp. strain ZGt-1.

6.8.1. Analysis strategy

A summary of the analysis pipeline is illustrated in Figure 6 below.
The details of the analysis are the following:

1. We downloaded all the RefSeq genome sequences of firmicutes that were available in NCBI at the time of starting the study (252 sequences in 2015).
2. We analyzed all the sequences using the latest version of antiSMASH available at that time; antiSMASH 4, which was described as a version that "provides a more sophisticated prediction and classification for class I lanthipeptides" compared to previous versions (Blin et al., 2017b)
    - antiSMASH 4 accepts whole genome sequences.
3. The set of lanthipeptides that resulted from antiSMASH 4 analysis were further analyzed using BLAST tools.
    - For the curation of antiSMASH results, each of the aa sequences of the detected lanthipeptides was analyzed using blastn, blastp, and tblastn. The BLAST search was conducted against the non-redundant and RefSeq databases.
        - This step aided in retrieving the annotations of the antiSMASH-detected lanthipeptides on the genome data presented in the NCBI.
            - In the resulting BLAST hits, the exact protein match belonging to the respective strain was examined. By doing so, we could verify the nucleotide sequence of the coding gene and its position reported by antiSMASH. We could also find the annotation of the gene product presented in the NCBI (i.e, lanthipeptide, hypothetical/ uncharacterized protein, protein other than lanthipeptide, or not annotated at all) (**Paper III**).
4. In parallel, the set of lanthipeptides that resulted from antiSMASH 4 were also analyzed using the protein BLAST tool available in BAGEL4.
    - This step aided in recognizing whether the antiSMASH-detected lanthipeptide has already been experimentally

proven as a lanthipeptide or represents a putative novel lanthipeptide.

5. In addition to the evaluation of the novelty of the detected lanthipeptide, we evaluated the novelty of the producing bacterial strain.
    - Each of the firmicute bacterial strains, whose genome harbors class-I lanthipeptide gene cluster(s) according to antiSMASH analysis, was searched against BAGEL4 databases to check if the strain has been reported as a lanthipeptide producer or is a potentially novel putative producer.
        - Knowing that about a given strain is useful as it indicates whether or not the strain represents a promising subject for future research.
6. In addition to the various analysis tools we used, we also mined the literature for class-I lanthipeptides and their firmicute producers (**Paper III**).

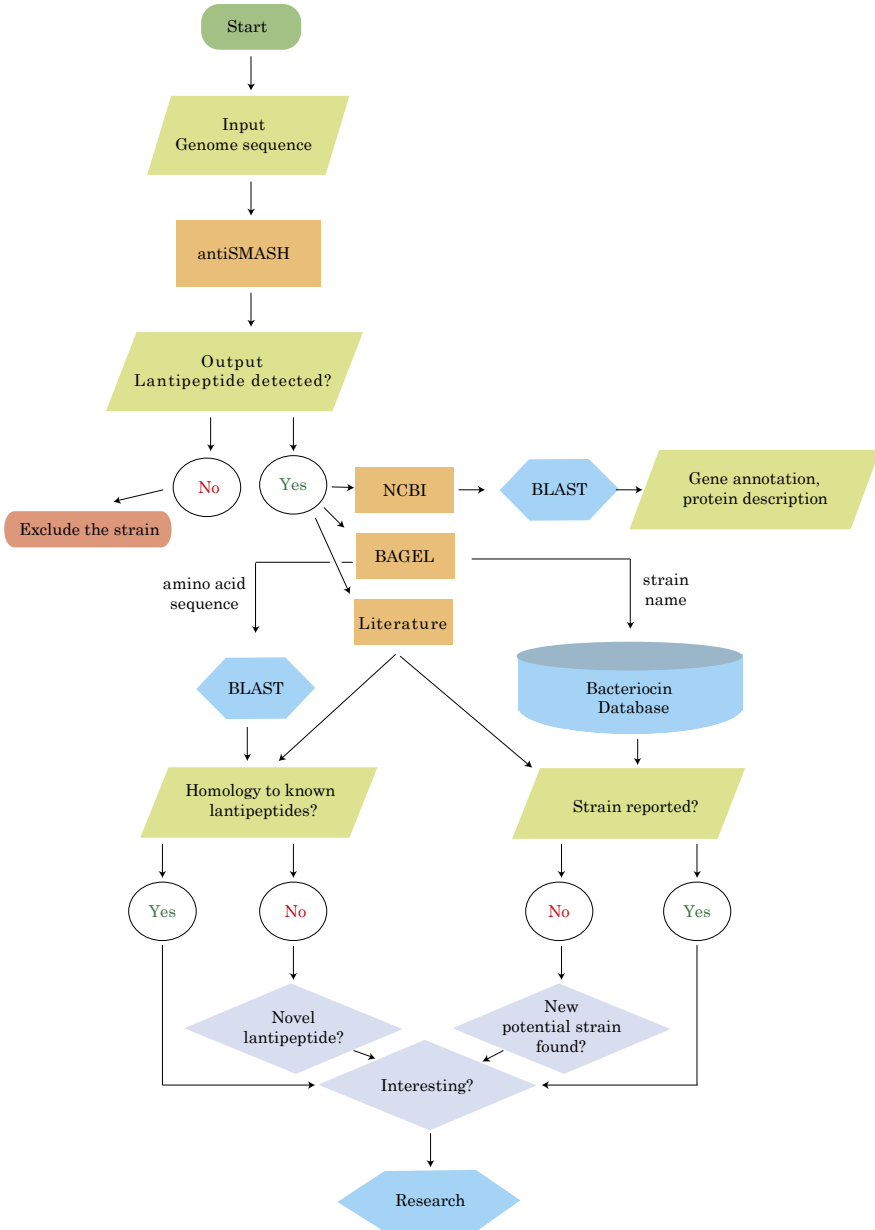


Figure 6
Workflow chart summarizing the analysis steps carried out for lanthipeptide detection (From Paper III).

6.8.2. Identification of firmicute lanthipeptides

Genome analyses indicate that lanthipeptides are widespread in bacterial genomes (Knerr and van der Donk, 2012). Firmicutes are among the primary producers of bacteriocins (Morton et al., 2015b), and they are known for the production of lanthipeptides, including lantibiotics. The majority of identified lantibiotics are those of firmicutes (Li and O'Sullivan, 2012). This does not necessarily mean that most of lantibiotics are produced by firmicutes, but it could be that the conducted research studies were biased towards this phylum (Morton et al., 2015b). Nevertheless, with the continuous generation of genome sequence data, an accompanying continuous probing for secondary metabolites, of which are lanthipeptides even among phyla that were subjected to previous research, is required.

Members of firmicutes have different industrial applications and biotechnological potential. For example, some members, such as bacilli, are considered among the most vital enzyme producers in various industries, such as food, paper, textiles, and others (Satyanarayana et al., 2012). Moreover, some members are known for the production of antimicrobial peptides (Satyanarayana et al., 2012). Among those antimicrobials are lanthipeptides, as was shown by different culture-based and genome-mining studies. On the other hand, firmicutes also include pathogenic members, and among those ones, there are pathogens that produce lanthipeptides.

In our study (**Paper III**), the genome analysis results were in line with what has been reported about firmicutes as lanthipeptide-producers and confirmed the results of previous studies. The results also revealed the lanthipeptide production potential of 40 firmicute strains that have not been associated with class-I lanthipeptides previously. Among these strains are two belonging to a bacterial species, *Streptococcus intermedius*, that none of its strains has been reported as a potential class-I producer. Furthermore, the results showed that some firmicutes code for potentially novel class-I lanthipeptides that have not been reported previously and do not show homology to any of the known small bacteriocins, as discussed further below.

Our results also showed that some firmicutes have more than one class-I lanthipeptide-coding gene (**Paper III**), such as *S. aureus* strains, which have two class-I lanthipeptide-coding genes within the same cluster, and *Bacillus thuringiensis* serovar IS5056, which has five lanthipeptide-coding genes. This is not improbable since it has been reported previously that a strain can harbor two lanthipeptide-coding genes or even more (Xin et al., 2015). For example, *B. thuringiensis* serovar *thuringiensis* strain T01001 has four lanthipeptide-coding genes within the same cluster (Xin et al., 2015). However, only one of the four genes codes for an antibacterially-active lanthipeptide, called thuricin 4A-4 (Xin et al., 2015).

Our results showed that different firmicute strains within the same species could code for identical lanthipeptides; for example, *P. polymyxa* strain M1 and strain SC2 code for identical lanthipeptides. As is the possible case with any identical peptides, the results showed that identical lanthipeptides could be coded by genes with different nt sequences, such as the genes for lanthipeptide (I) and (II) of *B. megaterium* QM B1551. Moreover, the results showed that strains of the opportunistic pathogen *S. aureus* code for the class-I lantibiotic; BsaA2, as mentioned above, and this is in agreement with previous reports (**Paper III**).

Some of the lanthipeptides identified in our study showed 100% identity to known lantibiotics, such as BacCH91, BsaA2, entianin, paenilan, subtilin, subtilomycin, suicin 90-1330, and thuricin 4A-4.

On the other hand, our analysis also resulted in lanthipeptides that showed no homology to verified small bacteriocins when their aa sequences were analyzed using BAGEL4 BLAST (**Paper III**). Displaying 0% homology is not unexpected due to the diversity in aa sequences of the precursor peptides of bacteriocins, and thus in lanthipeptides. As shown in **Paper III**, lanthipeptides of different firmicute species have a low aa sequence homology. Moreover, the study of lanthipeptides is a growing field; therefore, unless a given lanthipeptide or its homologue is deposited in a database (in this case, BAGEL4 databases), the use of BLAST homology-search tool will not retrieve homologous hits. Accordingly, relying on using BLAST alone is not enough for the prediction of novel lanthipeptides and could result in false negatives.

6.8.3. Analysis– Highlights and remarks

During the course of analyzing the antiSMASH-reported lanthipeptides using the NCBI and BAGEL4 BLAST tools, we faced cases where the lanthipeptide-coding gene reported by antiSMASH 4 was either annotated in the original genome record (GenBank record), or in the RefSeq record, but not in both; as in the case of *B. clausii* KSM-K16. The reported gene could also be annotated as coding for a non-lanthipeptide protein, or for a hypothetical/uncharacterized protein in one of the genome records, but annotated as coding for a lanthipeptide in the other record; as in the case of lanthipeptide (I) of *S. aureus* MSSA476, and the lanthipeptides of *S. aureus* ED133. Or it could also be annotated as coding for a hypothetical protein in both records, but BAGEL BLAST showed that the reported lanthipeptide is identical to an experimentally verified lanthipeptide, as in the case of the lanthipeptide of *B. subtilis* BSn5. In all these cases, our analysis confirmed that those protein records are lanthipeptides. This was of significance, especially for genes that were annotated as coding for hypothetical proteins in both genome records (**Paper III**).

We also faced cases where the antiSMASH-reported lanthipeptide was 100% identical to an experimentally verified lanthipeptide that has been reported in the literature but not in BAGEL4 databases. Therefore, our analysis could contribute to the enrichment of BAGEL4 databases (**Paper III**).

On the other hand, we faced cases where the position of the gene coding for the lanthipeptide precursor reported by antiSMASH 4 did not match with that reported by tblastn analysis and accordingly, the gene length did not correctly correspond to the length of the antiSMASH-reported lanthipeptide. This was the case with the lanthipeptide of *B. thuringiensis* serovar finitimus YBT-020, where antiSMASH 4 did not report the last nt in the stop codon (A in TAA). In other two cases, the gene reported in the RefSeq genome was longer than it should be based on the length of the antiSMASH-reported lanthipeptide as well as our manual inspection of nt sequence of the gene, as was the case with the lanthipeptide of *G. thermoleovorans* CCB_US3_UF5 and lanthipeptide (II) of *G. kaustophilus* HTA426. In these cases, we proceeded with analyzing the RefSeq genome sequences of these strains using antiSMASH 3 (**Paper III**). The antiSMASH 3 accepts fasta genome files as input and annotates the genome independently of the RefSeq genome annotation by employing the Prodigal pipeline. The analysis using antiSMASH 3 resulted in reporting the same lanthipeptide aa sequence that was reported by antiSMASH 4 analysis, but with different gene position which matched that reported by tblastn, and correctly corresponded to the length of the antiSMASH-reported lanthipeptide.

In order to confirm that the lanthipeptide reported by antiSMASH 4 and antiSMASH 3 was a true hit in each of those three cases, we analyzed the aa sequence of the lanthipeptide or its neighboring genes using InterPro. Furthermore, we analyzed the genome sequences of the three strains mentioned above using BAGEL4 genome mining tool. Both InterPro and BAGEL4 confirmed that the lanthipeptides detected by the two versions of antiSMASH were true hits. Therefore, we reported the lanthipeptide of each of those strains and the position of its coding gene as reported by antiSMASH 3, and recommended to edit the RefSeq annotation of the genes coding for the lanthipeptides of *G. thermoleovorans* CCB_US3_UF5 and lanthipeptide (II) of *G. kaustophilus* HTA426.

Additionally, lanthipeptide (II) of S. aureus NCTC 8325 also showed a discrepancy in the position of its coding gene between the position reported by antiSMASH 4 and that reported by tblastn. Similar to the case mentioned above, antiSMASH 4 did not report the last nucleotide in the stop codon (A in TAA). Without the need for antiSMASH 3, which did not report the lanthipeptide anyway, it was possible with manual inspection to recognize that it was just a mistake of missing out the last nt. Here as well, Interpro and BAGEL4 genome mining tool confirmed that the lanthipeptide reported by antiSMASH 4 was a true hit (Paper III).

6.8.4. Identification of putative novel lanthipeptides

In order to determine the novelty of a given antiSMASH-reported lanthipeptide, we sat stringent criteria that relied on referring to the annotation of the original genome as well as the RefSeq genome records (**Paper III**). The criteria also relied on BAGEL4 BLAST results and on the literature. The antiSMASH-reported lanthipeptide was considered in our study as putative novel when neither of the genome records reported the peptide as a lanthipeptide and when the lanthipeptide did not show 100% identity to any of the experimentally-verified lanthipeptides AND when it was not reported in a literature study that was based on *in silico* analysis of the genome analyzed in our study.

Out of the 69 class-I lanthipeptides which we found out by mining 252 firmicute RefSeq genome sequences, we identified seven putative novel lanthipeptides produced by five different firmicute strains; *B. thuringiensis* serovar finitimus YBT-020, *P. polymyxa* M1, *P. polymyxa* SC2, *S. aureus* NCTC 8325, and *S. intermedius* B196 (**Paper III**).

The antiSMASH-reported gene coding for each of these lanthipeptides was either unannotated in the genome records– as in the case of *S. aureus* NCTC 8325 (II), or was annotated as coding for a hypothetical protein in one record but was unannotated in the other, as in the case of the other six putative novel lanthipeptides. In order to confirm the antiSMASH prediction, BAGEL4 genome mining tool and InterPro analysis were used and did confirm the antiSMASH prediction results.

Interestingly, two of the putative novel lanthipeptides; specifically, the one of *B. thuringiensis* serovar finitimus YBT-020 and that of *S. intermedius* B196 did not show homology to any of the experimentally verified small bacteriocins reported in BAGEL4 databases. These two lanthipeptides were not reported in the literature either. The lack of homology is intriguing and indicates that exploring such lanthipeptides is of significant interest (**Paper III**).

6.9. Lanthipeptides of *Geobacillus*

Although *Geobacillus* has been attracting attention lately as an organism with potential biotechnological applications, as mentioned in Chapter 2, it has not been deeply researched for bacteriocin and – by extension – lanthipeptide production yet (Egan et al., 2018). Only few studies have reported the production of bacteriocins by *Geobacillus* species. Garg and co-workers reported the production of bacteriocins; geobacillin I and II by *G. thermodenitrificans* (Garg et al., 2012). Özdemir and Biyik, 2012 reported the production of a bacteriocin; toebicin 218 and a bacteriocin-like substance produced by two strains of *Geobacillus toebii* (Özdemir and Biyik, 2012a; Özdemir and Biyik, 2012b).

Pokusaeva and co-workers reported the production of bacteriocin-like substances, thermocins, by *G. stearothermophilus* (Pokusaeva et al., 2009). Among these reported bacteriocins, only geobacillin I and II were further characterized as lanthipeptides (Garg et al., 2012).

Egan and co-workers conducted an *in silico* screening study using BAGEL3 genome mining tool, where the publicly available genome sequences of *Geobacillus* spp. were screened for bacteriocin-coding genes (Egan et al., 2018). Genes coding for different modified bacteriocins (lanthipeptides) and unmodified ones were identified. Duly, this *in silico* analysis has revealed the potential of *Geobacillus* spp. as bacteriocin-producers yet to be exploited. However, genes coding for known bacteriocins >10 kDa were not identified on the genome sequences of *Geobacillus* spp. Nevertheless, the study indicated that *Geobacillus* represents a repertoire of bacteriocins (Egan et al., 2018). It even hinted that probably 30-99% of geobacilli produce at least one bacteriocin (Egan et al., 2018).

Among the recognized bacteriocins, Egan and co-workers identified different putative lanthipeptides. Part of the identified lanthipeptides showed 100% identity to geobacillin I (Egan et al., 2018). However, some of the lanthipeptides showed homology, but not full identity, to different bacteriocins. This indicates that lanthipeptides of *Geobacillus* spp. are diversified. Some of the identified lanthipeptides did not show homology to any of the characterized bacteriocins deposited in the Bactibase database (Egan et al., 2018). Therefore, some *Geobacillus* spp. may be producers of novel lanthipeptides. The study of Egan et al. showed that one potential lanthipeptide biosynthetic gene cluster may have more than one lanthipeptide-coding gene (Egan et al., 2018), and this is in line with our results (**Paper III**).

6.9.1. Class-I lanthipeptides of *Geobacillus* strains

In our study (**Paper III**), analyzing the completely sequenced genomes of *Geobacillus* strains and screening for class-I lanthipeptides resulted in the identification of *G. kaustophilus* HTA426 and *G. thermoleovorans* CCB_US3_UF5 as putative class-I lanthipeptide producers, in addition to *G. thermodenitrificans* NG80-2, which has already been experimentally verified as a producer of class-I and class II lanthipeptides; geobacillin I and II by Garg et al., 2012. Furthermore, *Geobacillus* sp. ZGt-1 was also identified as a putative class-I lanthipeptide-producer; Z-geobacillin (**Paper III**).

The structure of the lanthipeptide biosynthetic gene cluster in terms of the composing genes was shown to vary among *Geobacillus* spp. (Egan et al., 2018).

Our analysis of the clusters of the four *Geobacillus* strains mentioned above indicated that the clusters share the same lanthipeptide genetic makeup, where

genes coding for the lanthipeptide precursor (LanA), lanthipeptide modifying enzymes (LanBC), lanthipeptide transporter (LanT), lanthipeptide two-component response regulators (LanKR), and lanthipeptide immunity proteins; LanI and LanEFG are all present. However, the structure of the cluster is not identical among these strains. In *Geobacillus* sp. ZGt-1, *G. thermodenitrificans* strain NG80-2, and *G. thermoleovorans* CCB_US3_UF5, the genes are arranged in the lanthipeptide gene cluster in the following order; LanA, LanB, LanT, LanC, LanR, LanK, LanI, LanG, LanE, and LanF. The aa sequence of the precursor peptide (LanA) of these strains are not identical. On the other hand, *G. kaustophilus* HTA426 has two class-I LanAs (LanAI and LanAII), followed by two LanBs (LanBI and LanBII), and then followed by LanT, LanC, LanR, LanK, LanI, LanG, LanE, and LanF. LanAI is 100% identical over its entire length to LanA of *Geobacillus* sp. ZGt-1; Z-geobacillin. Moreover, the cluster of *G. kaustophilus* HTA426 has genes coding for transposases and hypothetical proteins inserted between the lanthipeptide-associated genes, and these inserted genes are more frequent than they are in the clusters of *Geobacillus* sp. ZGt-1 and *G. thermoleovorans* CCB_US3_UF5 (Paper III). Contrarily, the lanthipeptide cluster of *G. thermodenitrificans* NG80-2 is condensed, where no inserted genes have interrupted the succession of the lanthipeptide-associated genes in the cluster (**Paper III**). Our analysis indicated that none of the lanthipeptide gene clusters in these four strains has a gene coding for LanP, which is responsible for cleaving off the leader peptide. These results agree with those of Egan and co-workers, who concluded based on the *in silico* analysis, using BAGEL3, of genome sequences of different *Geobacillus* species that geobacilli lack LanP-coding genes (Egan et al., 2018). Garg and co-workers concluded that a protease coded elsewhere on the genome of *G. thermodenitrificans* NG80-2 might carry out the proteolytic cleavage of the leader peptide (Garg et al., 2012).

6.10. Z-geobacillin: A putative novel lanthipeptide of *Geobacillus* sp. ZGt-1

In addition to the identification of the seven putative novel lanthipeptides in different firmicute species mentioned above, the genome mining strategy we followed revealed one more putative novel class-I lanthipeptide. The analysis unveiled that *Geobacillus* sp. ZGt-1 harbors a complete class-I lanthipeptide biosynthesis gene cluster, and we termed the lanthipeptide as Z-geobacillin.

Revealing the presence of Z-geobacillin on the genome of strain ZGt-1 has scientific and industrial significance. With respect to scientific knowledge, Z-geobacillin is one of only few reported lanthipeptides showing that thermophilic bacteria constitute a promising source of lanthipeptides. Showing

that is of interest especially because only the thermophilic *Streptococcus thermophilus* SBT1277 (Kabuki et al., 2007) and *G. thermodenitrificans* NG80-2 (Garg et al., 2012) have been experimentally studied for their lanthipeptide production. Revealing the presence of Z-geobacillin also adds knowledge about lanthipeptides of thermopiles in general, and thermophiles isolated from hot springs in particular, especially since there are no reports on the lanthipeptide production potential of bacterial strains isolated from hot springs.

With respect to industrial applications, the significance of revealing the presence of Z-geobacillin lies in the possible fulfilment of the lack of lanthipeptides that retain their stability and antimicrobial activity over a broad range of environmental conditions. The stability and solubility of nisin, which is the only approved and commercialized bacteriocin, are highly dependent on pH (Gharsallaoui et al., 2016). Nisin is more soluble and more-antibacterially active at low pH (Gharsallaoui et al., 2016). Its thermostability decreases gradually with the increase in pH, and this leads to a decrease in the antimicrobial activity (Gharsallaoui et al., 2016). The issues of limited solubility, stability, and activity of nisin are drawbacks that impede broadening its industrial applications (Garg et al., 2012; Gharsallaoui et al., 2016). For example, food matrices that have pH values close to neutral are highly exposed to the growth of pathogenic bacteria, such as *Listeria monocytogenes*, and at the same time, the solubility and antibacterial activity of nisin are lower at such pH than at acidic pH (Gharsallaoui et al., 2016).

Consequently, it is of industrial interest to find more stable lanthipeptides. Since extremophiles, among which are thermophilic bacteria, live under harsh environmental conditions compared to the conditions where the mesophilic nisin-producing bacterium *L. lactis* lives, they are expected to offer more stable proteins (Garg et al., 2012).

Z-geobacillin is encoded on the chromosomal genome of strain ZGt-1 and shares similarities with the thermostable class-I lanthipeptide; geobacillin I produced by *G. thermodenitrificans* NG80-2, as shown below.

6.10.1. *In silico* characterization of the gene cluster of Z-geobacillin of *Geobacillus* sp. ZGt-1 and its biosynthesis pathway

Z-geobacillin precursor peptide (ZGeoA) is composed of 23-aa-leader peptide and 33-aa-core peptide. This was reported by antiSMASH, which also predicted the point at which ZGeoA is cleaved for the removal of the leader peptide and formation of the active peptide as a result. The cleavage site was predicted by antiSMASH to be ProAsn↓Ile (PN↓I) (Figure 4 in **Paper III**). This prediction is valid since it agrees with the cleavage site of geobacillin I (ProAsn↓Val) (PN↓V),

which was experimentally proved by Garg et al., 2012. Accordingly, the core peptide of Z-geobacillin starts with the aa (IleThrSer) (Figure 7)

Using antiSMASH, the complete biosynthetic gene cluster of class-I Z-geobacillin could be identified (**Paper III**).

The gene coding for the precursor lanthipeptide is a short ORF (171 base pairs) on contig 6_34 (LDPD01000000), designated as *zgeoA*. The genes coding for the enzymes required for the post-translational modifications; the lanthipeptide dehydratase-coding gene (*zgeoB*) and the lanthipeptide cyclase-coding gene (*zgeoC*) are coded downstream of *zgeoA* (Figure 3 in **Paper III**). The gene *zgeoA* codes for the leader and core peptides of Z-geobacillin. Since there is no gene coding for LanP, as mentioned above, the leader peptide is predicted to be cleaved off by a protease coded somewhere on the genome of ZGt-1, as inferred from (Garg et al., 2012; Corvey et al., 2003).

The aa sequence of the core Z-geobacillin is shown in Figure 7. The expected post-translational modifications resulting from the action of ZgeoB on the core-peptide, i.e., the possible dehydrated Ser and Thr residues that form dehydroalanine (Dha) and dehydrobutyrine (Dhb), respectively, as well as the possible thioether cross-links, resulting from the action of ZgeoC are also shown in Figure 7.

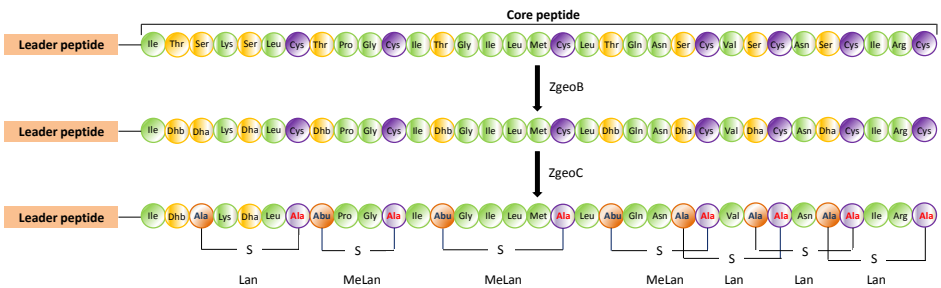


Figure 7
The aa sequence of Z-geobacillin. (a) The unmodified core peptide. (b) The core peptide after dehydration of Ser and Thr residues by ZgeoB, assuming that all these reidues undergo dehydration. (c) The core peptide after cyclization by ZgeoC and production of (Me)Lan residues. The cyclization pattern illustrated here was assumed to be the same as that of geobacillin I (Garg et al., 2012). Addition of the thiols of Cys residues to Dha and Dhb residues are shown as ”bridges” with the letter (S) standing for the sulfur of the thiol group. Abu=2-aminobutyric acaid. The graphical ilustration of the aa sequence was inspired by (Tang et al., 2015).

The modified Z-geobacillin is expected to be exported outside the producing cell of strain ZGt-1 via a transmembrane ATP-binding cassette (ABC) transporter, designated here as ZgeoT. The gene cluster of Z-geobacillin harbors *zgeoT*, which codes for the ABC transporter, downstream of *zgeoA* (Figure 3 in **Paper III**).

As is the case with the regulation of lanthipeptide biosynthesis in general (Chen and Hoover, 2003; Arias et al., 2011; Arnison et al., 2013), the biosynthesis of Z-geobacillin is expected to be regulated by the two-component regulatory system; the membrane-bound sensor histidine kinase; designated here as ZgeoK, and the cytoplasmic response regulator, designated as ZgeoR. The coding genes *zgeoK* and *zgeoR* are coded downstream of *zgeoA* (Figure 3 in **Paper III**).

Each of the lanthipeptide-producing cells needs to protect itself from the inhibitory effect of its own lanthipeptide. Therefore, these cells code for self-immunity proteins, generally designated as LanIEFG (Arnison et al., 2013). In the case of strain ZGt-1, the gene cluster of Z-geobacillin codes for *zgeoI*, *zgeoG*, *zgeoE*, and *zgeoF*. Based on what has been reported about LanI (McAuliffe et al., 2001), the putative ZgeoI is expected to be a peripheral membrane lipoprotein that would block the pore formation by Z-geobacillin. Additionally, based on what has been reported about LanIEFG (McAuliffe et al., 2001; Chen and Hoover, 2003), the putative ZgeoEFG are expected to be specialized ABC-transporters that would pump Z-geobacillin molecules which have penetrated the membrane back to the exterior environment.

6.10.2. Z-geobacillin biosynthesis pathway model

Based on previous studies on lanthieppetide biosynthesis pathway, a hypothetical model of the biosynthesis pathway of Z-geobacillin could be illustrated (Figure 8). When the putative gene, *zgeoA* is translated, this is expected to mark the start of the biosynthesis pathway of Z-geobacillin. The peptide product, the precursor peptide ZgeoA, consists of leader and core peptides. Post-translational modifications take place due to the action of ZgeoB and ZgeoC enzymes. This results in the formation of the modified precursor peptide, as recommended to be named by (Arnison et al., 2013). In this case it can be designated as "mZgeoA" that has the dehydrated residues; Dha and Dhb and has (Me)Lan bridges (thioether cross-links) and the leader peptide. This mZgeoA is expected to be exported to the exterior environment via putative transmembrane ABC transporter, ZgeoT (Figure 8). mZgeoA will not be active unless its leader peptide gets cleaved off (McAuliffe et al., 2001; Knerr and van der Donk, 2012). The proteolytic cleavage of the leader peptide of Z-geobacillin, carried out by a protease other than LanP, could take place either intracellularly before the translocation by ZgeoT, or extracellularly after exporting the modified lanthipeptide, as inferred from (McAuliffe et al., 2001; Chen and Hoover, 2003; Knerr and van der Donk, 2012).

The lanthipeptide biosynthesis needs to be regulated and this is performed by the two-component regulatory system; the membrane-bound sensor histidine

kinase and the cytoplasmic response regulator (Chen and Hoover, 2003; Arias et al., 2011). Accordingly, we expect the biosynthesis of Z-geobacillin to be regulated by the same system and in the sequence described in these two studies (Figure 8).

After the secretion of a number of Z-geobacillin molecules and when the concentration reaches a certain threshold that the membrane-bound sensor histidine kinase (corresponding to the putative ZgeoK in strain ZGt-1) senses, ZgeoK is expected to auto-phosphorylate (Figure 8). This will lead to the phosphorylation of the cytoplasmic response regulator (corresponding to the putative ZgeoR). The phosphorylated ZgeoR, in turn, should further activate the transcription of the putative *zgeoA* and *zgeoBCT*, genes leading to the expression, modification, and transportation of increasing amount of ZgeoA. Transcription of the immunity protein-coding genes, the putative *zgeoI* and *zgeoGEF*, are also activated by the phosphorylated ZgeoR. Furthermore, ZgeoR is expected to activate *zgeoR* and *zgeoK* genes further. Future experimental studies will tell more about the biosynthesis pathway of Z-geobacillin.

The Z-geobacillin biosynthetic gene cluster harbors transposon-coding genes (**Paper III**). This goes in line with what has been reported about the mobile nature of lanthipeptide clusters (Begley et al., 2009).

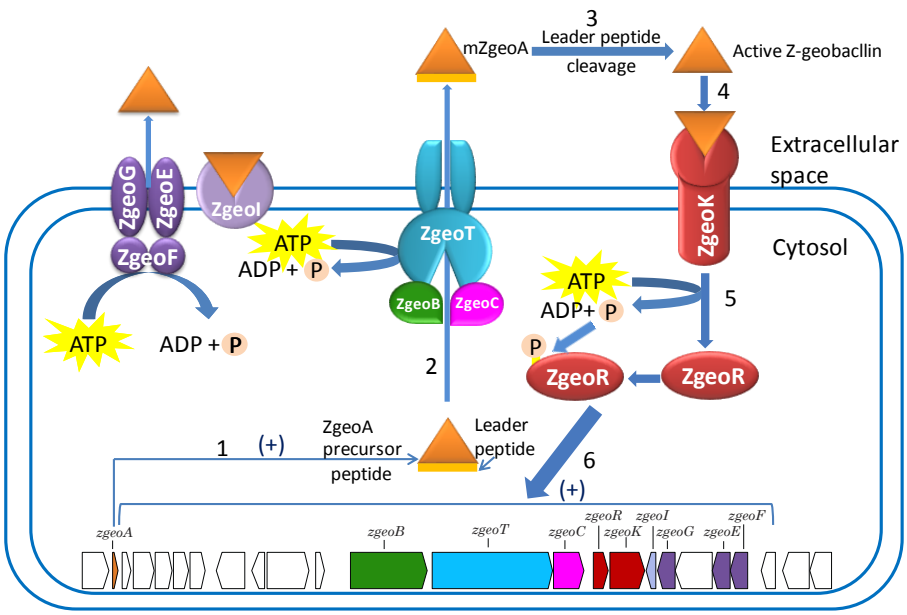


Figure 8
Schematic illustarton of the expected biosynthesis and regulation of Z-geobacillin. For sketching clarity, the diagram is illustrated with the assumption that the cleavage of the leader peptide takes place extracellularly, but it could also take place in the cytosol. Details and abbreviations are given in the text. Adapted from (Arias et al 2011; Chen and Hoover, 2003).

6.10.3. Z-geobacillin Highlights

The precursor peptide of Z-geobacillin; ZgeoA and that of geobacillin-I produced by *G. thermodenitrificans* NG80-2 are composed of 56 aa each. However, they are not 100% identical; the aa sequence of ZgeoA is 91% identical to that of geobacillin I. The aa sequences of the core peptides of the unmodified forms of these two lanthipeptides are 94% identical. This indicates that Z-geobacillin is a variant of geobacillin I.

Z-geobacillin has the potential to form seven thioether cross-links (Figure 7). This number of cross-links could be concluded from its aa sequence since it has seven Cys residues, and it was also reported by antiSMASH. Having the potential to form seven thioether cross-links is not surprising since Z-geobacillin is highly similar to geobacillin I, which itself has been experimentally proved to have seven thioether cross-links (Garg et al., 2012). On the other hand, nisin has five cross-links as it has two Cys residues less than geobacillin I and Z-geobacillin. Accordingly, Z-geobacillin is expected to have a more constrained; thus, more stable conformation compared to nisin. Since geobacillin I was proved to be more stable than nisin A at pH 7 and 8 at 37 and 60 °C (Garg et al., 2012), we expect Z-geobacillin to display similar stability.

As reported by Garg et al., 2012, seven thioether cross-links represent the highest number of bridges which has been reported for a lanthipeptide so far. To date, only geobacillin I has been proved as having seven thioether cross-links (Garg et al., 2012). This makes Z-geobacillin of interest as a lanthipeptide with the highest number of cross-links.

Based on the aa sequence of Z-geobacillin, the total number of potential dehydrations that may form is nine (Figure 7), since it has five Ser and four Thr residues. Geobacillin I has the same number of Ser and Thr, and it was experimentally confirmed that all these nine residues undergo dehydrations.
Correspondingly, assuming that all the nine Ser and Thr residues of Z-geobacillin undergo dehydration just as geobacillin I and adding the possible seven cross-links that should form, Z-geobacillin is then expected to have 16 post-translationally modified residues in the core peptide which is composed of 33 residues.

Concerning the antimicrobial spectrum of Z-geobacillin, it could be similar to that of geobacillin I, but it is not necessarily the case since Z-geobacillin is not 100% identical to geobacillin I. Moreover, *G. thermodenitrificans* NG80-2 and *Geobacillus* sp. ZGt-1 inhabited two different ecological niches; Dagang oil fields in China and Zara hot spring in Jordan, respectively. This means that the two strains were exposed to different environmental conditions, which accordingly affect their physiological traits and their survival strategies. Geobacillin I has been proved three times more active against *Streptococcus dysgalactiae* ATCC 27957 than nisin (Garg et al., 2012). It has also been proved

active against the VRE and *Bacillus anthracis* Sterne 7702 with an activity level similar to that of nisin (Garg et al., 2012). Additionally, geobacillin I has also been active against methicillin-resistant *Staphylococcus aureus* (MRSA), *Bacillus subtilis* ATCC 6633, and *L. lactis* HP but displayed lower activity than that of nisin (Garg et al., 2012). The mechanism of action of geobacillin I was shown to involve binding to lipid II as well as forming pores in the membrane (Garg et al., 2014).

These results, therefore, indicate that Z-geobacillin is a promising lanthipeptide that is worth-exploring. To the best of our knowledge, Z-geobacillin is the first lanthipeptide identified in hot spring-inhabiting bacteria. Moreover, there are no reports on antimicrobial compounds –in general– isolated from thermophilic bacteria living in hot springs. Isolation from untapped sources raises the probability of Z-geobacillin being a novel lanthipeptide.

Furthermore, considering peptide engineering, the Z-geobacillin biosynthesis enzymes, ZgeoB and ZgeoC have potential applications in introducing modified aa in non-lanthipeptides to transform them into lanthipeptides, as discussed above.

7. Type II Toxin-Antitoxin system in *Geobacillus* strains

"A noble purpose inspires sacrifice, stimulates innovation and encourages perseverance".

Gary Hamel

Organisms belonging to different kingdoms – animals, plants, fungi, and microorganisms produce toxins as a defense strategy, as a thriving strategy to help in predation, or as an establishment strategy to set up an infection in a host species (Kędzierska and Hayes, 2016).

In the previous chapters, we saw that bacteria secrete different types of peptides and/or proteins that can inhibit or kill bacterial cells other than the producing bacterium. In this chapter, we are going to see how a bacterium can produce toxins that are not secreted but kept inside the producing cell and may kill it under certain conditions. Therefore, these toxins have been described as "biological bombs" (Yarmolinsky, 1995) or "intracellular time bombs" (Kędzierska and Hayes, 2016). This type of toxins is part of a system known as the toxin-antitoxin (TA) system.

Researching the TA system is a proliferating field. Discovering the first TA loci took place in 1983 by Ogura and Hiraga, who identified a TA pair, then called "functional regions", on an *E. coli* plasmid. These "functional regions" were involved in plasmid maintenance, where one of the regions was responsible for inhibiting cell division, and the other was responsible for suppressing the inhibitory function (Ogura and Hiraga, 1983).

Attention has been directed towards TA systems only recently. Most of TA systems in bacteria were discovered over the first decade of the 2000s (Yamaguchi and Inouye, 2011). A deep understanding of this system is expected to introduce crucial biotechnological applications (Gerdes, 2013). The continuous increase in the number of sequenced genomes of prokaryotes has helped in the fast development of this field; it aided in increasing the number of discovered TA families and indicated that this highly diverse system exists in almost all bacteria and many archaea (Gerdes, 2013; Yamaguchi and Inouye, 2011).

TA systems are composed of two components; a toxin, which is a protein, and a labile antitoxin (also called "antidote") (Gerdes, 2000), which is either a protein or an RNA (Guglielmini and Van Melderen, 2011). The toxin interferes with a cellular process and causes cell death or growth attenuation under stress conditions (Page and Peti, 2016). The antitoxin neutralizes the toxin in order to protect the growth under normal growth conditions (Page and Peti, 2016). Toxins are activated under stress to either reduce the population for the benefit of the rest via programmed cell death (PCD) (Engelberg-Kulka and Glaser, 1999) as a form of "altruistic death" (Diaz-Orejas et al., 2017), or to reduce the metabolic activity in order to save resources via halting the growth or slowing it down till the conditions improve (Diaz-Orejas et al., 2017) (more details will be given further below).

TA loci may be encoded on plasmids and/or chromosomes, where some could be within mobile genetic elements, such as prophage islands (Diaz-Orejas et al., 2017). TA genes encoded on the plasmid promote the plasmid stabilization via a mechanism known as post-segregational killing (PSK) (Hayes and Kędzierska, 2014). This mechanism was first described by Gerdes and co-workers in 1986 (Gerdes et al., 1986). Daughter bacterial cells which did not inherit the plasmid are killed, due to the degradation of the labile antitoxin and the lack of its *de novo* synthesis, leading the inherited stable toxin to exert its toxic activity, causing bacteriostasis or cell death (Hayes and Kędzierska, 2014). Loss of the plasmid is considered as a stress situation that the cell responds to by activating the toxin (Diaz-Orejas et al., 2017). Consequently, the cell becomes addicted to the plasmid (De Bast et al., 2008). Therefore, the PSK is also known as the "addiction phenomenon" (De Bast et al., 2008).

Chromosomally encoded TAs as well may be associated with overcoming stress. For example, they may help in defending the cells against bacteriophages. They also help in the survival of pathogenic bacteria in eukaryotic cells during infection by inducing persistence (Diaz-Orejas et al., 2017), as shown for *Mycobacterium tuberculosis* (Ramage et al., 2009). They may also lead to biofilm formation (Diaz-Orejas et al., 2017) when bacteria are exposed to adverse factors, such as antimicrobial and DNA damage agents (Gotoh et al., 2010). Functions of Chromosomally encoded TAs are explained in section 7.3.

Currently, there are six types of TA systems. They were identified based on the nature of the antitoxin and the mechanisms used to neutralize the toxin (Diaz-Orejas et al., 2017). In types I and III, the antitoxin is a non-coding RNA, while in types II, IV—VI, it is a protein. Type II is the best described system; however, it has not been well described in thermophilic bacteria. In the TA database (TADB), the only listed geobacilli are *G. kaustophilus* HTA426 and *G. thermodenitrificans* NG80-2, and only part of their TAs are shown. For these reasons together with our interest in *Geobacillus*, type II TA system in strains of this genus was the focus of **Paper IV** and is discussed in this chapter.

7.1. General features of the type II TA system

Most members of the type II TA system have common features. Genes coding for the toxin and antitoxin share the same operon and the gene coding for the antitoxin is usually upstream of that coding for the toxin, in order to guarantee the synthesis of the antitoxin before the toxin, as a way to protect the cell (Jurėnas et al., 2017). TA transcription is autoregulated (Figure 9) (Gerdes, 2013; Page and Peti, 2016), as discussed in the next section. Another common feature is that toxins of this type are highly diversified, including even toxins of the same TA family (Jurėnas et al., 2017). Toxins have various cellular targets (Figure 9), as discussed further below.

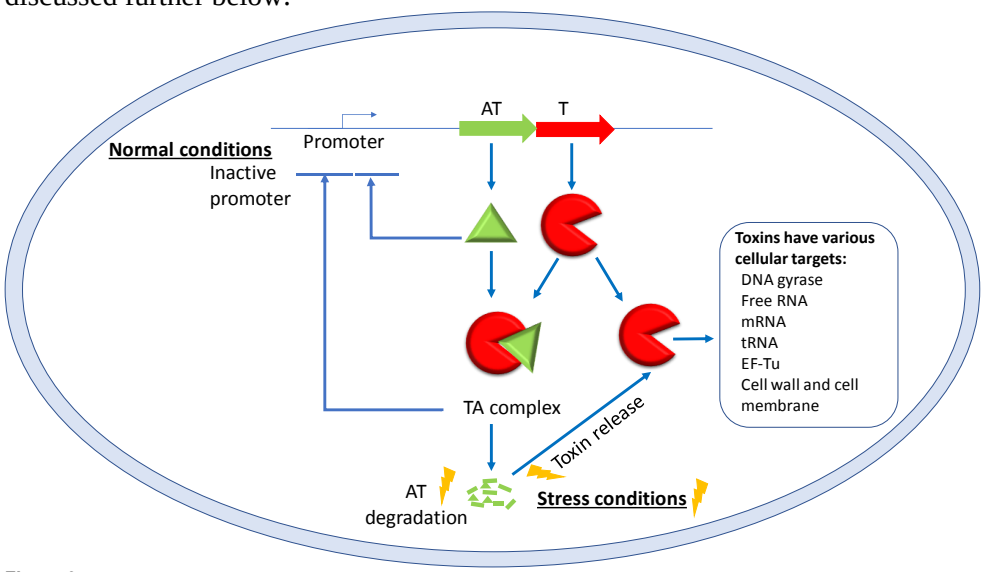


Figure 9
Schematic illustration of the regulation of type II TA system. Under normal conditions, the antitoxin (AT) neutralizes the toxin (T) by forming a complex (TA complex), and the complex and AT repress the promoter. Under stress conditions, the antitoxin is degraded by proteases, the toxin is released from the complex and interferes with a certain cellular process, depending on its target, and the promoter is de-repressed. Adapted from (Yamaguchi and Inouye, 2011; Coussens and Daines, 2016; Yang and Walsh, 2017).

7.2. Regulation of the TA transcription

In the type II TA system, the TA proteins act as transcriptional auto-repressors (Yamaguchi and Inouye, 2011). The antitoxin (AT) alone partially represses the transcription of the TA operon by binding to an operator site and blocking the RNA polymerase from binding to the promoter (Hayes and Kędzierska, 2014) (Figure 9) However, the full repression is achieved when the TA complex binds

to the operator site (Hayes and Kędzierska, 2014) (Figure 9). That is due to the following: The AT is usually composed of well and partly structured N- and C-terminal domains, respectively (Hayes and Kędzierska, 2014). The N- and C-terminal domains are involved in DNA- and toxin-binding, respectively. When the toxin (T) binds to the C-terminal domain of the AT, it restructures the latter; thus, stabilizes the AT. This then leads to a full repression (Hayes and Kędzierska, 2014). On the other hand, when the AT is degraded by proteases under stress, the promoter will be de-repressed since there is no AT and no TA complex. As a result, the T and AT will be synthesized. However, if the stress sustains, AT will continuously be degraded and T will stay active, until the stress factor is removed. As will be shown further below, there are factors that lead to the repression of the TA promoter under stress, and as a result, the AT will not be synthesized and the toxin will stay active for some time.

7.3. Possible physiological roles of type II chromosomally encoded TA families

While the plasmid-encoded TA genes are responsible for the PSK, the chromosomally encoded ones play other roles that affect the cell physiology. These roles have been debated, as some studies demonstrated that the chromosomally encoded TAs might mediate PCD, as mentioned above, while others excluded this role and showed that TAs have merely a bacteriostatic effect. Below, the various potential roles of the chromosomally encoded TAs are discussed.

7.3.1. PCD

The concept of having PCD in bacteria was described by Yarmolinsky in 1995 in a study published in Science (Yarmolinsky, 1995). He ascribed it to plasmid-born genes coding for a toxin and an "antidote" (the antitoxin) (Yarmolinsky, 1995).

One of the debated roles of the type II TA system is the trigger of PCD. The Engelberg-Kulka research group proved that, in an *E. coli* strain, the activation of the MazF toxin resulted in PCD, due to the exposure to stress factors, such as high temperatures, oxidative stress, antibiotics, or DNA damage caused by the U.V radiation or thymine starvation (Engelberg-Kulka and Glaser, 1999; Sat et al., 2003; Amitai et al., 2004). The same research group showed that the MazEF-mediated cell death is a population phenomenon that relies on the bacterial culture cell density, where death occurs in high cell density cultures, as opposed

to diluted ones (Kolodkin-Gal and Engelberg-Kulka, 2006). Accordingly, it was concluded that the MazEF-mediated cell death is a quorum sensing phenomenon and is regulated by a quorum-sensing factor, named the extracellular cell death factor (EDF), that coordinates the MazEF-mediated death when the cell is exposed to stress conditions (Kolodkin-Gal and Engelberg-Kulka, 2006). It was also shown that the MazF and the EDF were in a positive feedback loop, where the EDF activated the MazF, and the activation of MazF led to an increase in the production of the EDF, which thus increased cell death (Kolodkin-Gal and Engelberg-Kulka, 2006).

In a rather recent and independent study by the Li group, activation of the toxin, SezT, of the type II chromosomally encoded TA pair, SezAT (SsPI-1-borne Epsilon/Zeta antitoxin antitoxin), in a strain of *Streptococcus suis* had a bactericidal, rather than a bacteriostatic, effect (Yao et al., 2015b).

On the other hand, several research studies have ruled out the PCD role of the type II TA families, including the MazEF family (Fu et al., 2009; Jørgensen et al., 2009; Tsilibaris et al., 2007; Christensen-Dalsgaard and Gerdes, 2006; Christensen et al., 2003; Pedersen et al., 2002).

The Engelberg-Kulka group concluded that the production of the active EDF at specific concentrations is crucial for the PCD to take place (Kolodkin-Gal and Engelberg-Kulka, 2006). Moreover, the same group found that the PCD is a strain-dependent phenomenon since not all strains produce the EDF (Kolodkin-Gal and Engelberg-Kulka, 2006).

Conclusively, chromosomal type II TA families may cause PCD, but this is controlled by different factors. The PCD helps the bacterial population as a whole, since by reducing the number of viable cells, the surviving siblings will have access to more nutrients and may also feed on the cell debris of the dead cells (Hayes, 2003). The PCD also protects the population from the spread of bacteriophages (Hazan and Engelberg-Kulka, 2004), as will be discussed in more detail below.

7.3.2. Growth arrest under stress conditions

Studies showed that some chromosomally encoded toxins lead to the attenuation of growth, where bacteria stay viable but have a reduced metabolism and do not replicate, under unfavorable conditions (Christensen et al., 2001; Pedersen et al., 2002; Gerdes et al., 2005; Pandey and Gerdes, 2005; Wang and Wood, 2011; Coussens and Daines, 2016). In other words, TA systems promote the survival of bacteria under stress conditions, such as nutrient limitation, exposure to antibiotics, oxidative stress, unfavorable pH and temperature, bacteriophages, or host immune response (Kang et al., 2018; Coussens and Daines, 2016). Such conditions induce the degradation of antitoxins by proteases, such as Lon and

ClpXP, allowing the toxin to exert its effect (Coussens and Daines, 2016), which will then help the microorganism to save energy and aa by reducing the protein synthesis and the associated translational errors, and diminishing the accumulation of damaged DNA caused by reactive oxygen species (Coussens and Daines, 2016; Gerdes et al., 2005). Since the TA complex is already formed prior to the exposure to stress, having the toxin already synthesized will speed up the microorganism response (Coussens and Daines, 2016).

The Gerdes lab showed that the RelE toxin was activated and acted as a global inhibitor of translation in *E. coli*, when grown in aa- or glucose-limited environment or when chloramphenicol was added, without leading to cell death (Christensen et al., 2001). The same lab also showed that the MazF toxin, which acted as both translation and replication inhibitor, had only a bacteriostatic effect, as cells were able to resume their growth when the MazE antitoxin was overexpressed (Pedersen et al., 2002). Therefore, these studies excluded the programmed cell death role and emphasized that chromosomal TAs are involved in bacteriostasis as a stress management strategy. However, Tsilibaris and co-workers showed that chromosomal TAs tested by the Gerdes lab were not involved in bacteriostasis under stress conditions (Tsilibaris et al., 2007). Nonetheless, they did not rule out the potential of other chromosomal TAs in stress management (Tsilibaris et al., 2007).

Stress factors affect gene regulation and the general stress response and may trigger a switch from the planktonic growth mode to the biofilm growth mode (reviewed by (Wang and Wood, 2011)). Interestingly, studies have found that chromosomal TAs regulate biofilm formation. The group of T. K. Wood found that the MqsR toxin (motility quorum sensing regulator) of the MqsRA TA system mediated the biofilm formation in *E. coli* after the degradation of the MqsA antitoxin (Ren et al., 2004; Barrios et al., 2006; Wang and Wood, 2011), and this role was also concluded in another study (Yamaguchi et al., 2009). In addition to the MqsR, MazF, and RelE were among the toxins that were studied by the Wood lab and shown to influence biofilm formation (Kim et al., 2009).

Biofilms are involved in 80% of human bacterial chronic inflammatory and infectious diseases (reviewed by (Wang and Wood, 2011)). It should be noted that cells in biofilms eventually reach the growth arrest stage (Arnaouteli et al., 2019). For pathogens, undergoing growth arrest helps them avoid the host immune response (Coussens and Daines, 2016). Moreover, antibiotics become ineffective since most of them target the nucleic acids or proteins of replicating cells, not cells in the bacteriostasis state (Coussens and Daines, 2016). When the stress factor is removed, the microorganism can resume its growth (Coussens and Daines, 2016). Accordingly, toxins help the microorganism, and in turn the population, to survive unfavorable conditions. Clearly, the survival of pathogens has adverse consequences on the host health. Bacterial pathogens that can survive antibiotic treatment were termed as “persisters” by Bigger, 1944 (Bigger, 1944;

Coussens and Daines, 2016). Persisters may also form due to environmental factors, such as nutrient limitations, interactions among microbes, and host defense reactions, as was shown by (Mc Dermott, 1958). Persisters form mainly in biofilms and in cultures at the stationary phase (Wang and Wood, 2011). Mc Dermott, stated that microorganisms "play dead" as a survival strategy when faced by unfavorable conditions, showing that this "tactic" is not only implemented by animals (Mc Dermott, 1958). At the time of these two studies, toxin-antitoxin systems were not known; therefore, a link between them and persistence could not be established. However, this link has been suggested during the last decade, as discussed below.

Due to their growth inhibitory effect, toxins of the TA systems are capable of inducing persistence (Shah et al., 2006). Shah and co-workers indicated that some TA coding genes, such as *relE* and *mazF*, were overexpressed in *E. coli* persistent cells (Shah et al., 2006). Keren and co-workers showed that overexpression of *relE* in *E. coli* caused an increase in persister cell formation (Keren et al., 2004). The same study and another by Schumacher and co-workers also concluded that HipA toxin, of the HipBA TA family, induced *E. coli* persister cell formation (Keren et al., 2004; Schumacher et al., 2009). However, a recent study by (Goormaghtigh et al., 2018) has opposed those studies and reported the lack of a direct link between the induction of TA systems and the persister cell formation. Therefore, more studies are required in order to confirm or deny the debated role of TA systems in persistence.

Taken altogether, some studies suggest that type II chromosomal TA systems are involved in growth arrest as a stress management strategy, and thus may be involved in the associated biofilm and persister cell formation. However, carrying out more studies is demanded to confirm these roles.

7.3.3. Virulence

Expanding on the roles of type II chromosomally encoded TAs within the clinical context, some studies have reported a possible association between these TAs and virulence of pathogenic bacteria. The FitAB (fast intracellular trafficking) TA pair of *Neisseria gonorrhoeae*, which causes gonorrhea in humans, was suggested to play a role in pathogenicity (Mattison et al., 2006). FitAB controls the intracellular growth and helps the cells evade the host immune system (Mattison et al., 2006). In *Streptococcus pneumoniae*, only the highly virulent strains, as opposed to intermediately virulent and non-invasive strains, harbor the PezAT (Pneumococcal epsilon-zeta antitoxin toxin) TA pair encoded on the pathogenicity island 1 (Harvey et al., 2011). The PezAT was shown to enhance the virulence due to its role in stabilizing the pathogenicity island (see below); thus, strains lacking the PezAT had impaired virulence (Chan and Espinosa,

2016). Similar results were obtained for the PezAT homologue, SezAT TA pair of *Streptococcus suis* (Yao et al., 2015b).

7.3.4. Stabilization of mobile genome regions

TAs have been described as addiction modules since the cell becomes addicted to their presence (Yarmolinsky, 1995). Therefore, the otherwise dispensable genome regions that harbor TAs become stabilized (Yarmolinsky, 1995; Van Melderen and De Bast, 2009). TAs have also been described as selfish genes that ensure their presence in bacterial genomes by preventing the growth of TA-free progenies (Szekeres et al., 2007; Van Melderen and De Bast, 2009). This "selfish" behavior offers the benefit of preventing gene loss.

Christensen-Dalsgaard and Gerdes suggested that the chromosomal TA coding genes, *higBA*, contribute to the stabilization of the supeintegron (SI) of *Vibrio cholerae* (Christensen-Dalsgaard and Gerdes, 2006). SIs are chromosomal site-specific recombination systems, where mobile gene cassettes are integrated and expressed, and they highly influence bacterial evolution and adaptation (Labbate et al., 2009; Szekeres et al., 2007). Based on the structure of the integron and superintegron systems, instability of the mobile gene cassettes is expected; however, these systems are highly stable (Labbate et al., 2009; Szekeres et al., 2007). The stability is thought to be granted by selective pressure or by other mechanisms in the absence of selection (Labbate et al., 2009; Szekeres et al., 2007).

Szekeres and co-workers demonstrated that chromosomally encoded TAs could stabilize superintegrons (SIs) (Szekeres et al., 2007). They proved that chromosomal TA coding genes, relBE and parDE, encoded within SIs gene cassettes in *Vibrio vulnificus* counteracted the loss of large SIs (Szekeres et al., 2007). They also proved that the inclusion of these genes within a dispensable 165 kb genomic DNA fragment in *E. coli* repressed the deletion of the latter (Szekeres et al., 2007). Therefore, the RelBE and ParDE TAs can stabilize long stretches of genomic DNA (Szekeres et al., 2007). However, Tsilibaris and co-workers suggested that TAs play a role in genome stabilization only when they are encoded on the plasmid, while they might lose this role when they are integrated into the chromosome (Tsilibaris et al., 2007). More studies are certainly needed to confirm this role and to explore it for other TA families.

7.3.5. Antiaddiction modules

The term "antiaddiction module" is based on the possible cross-interaction between the homologous chromosomal- and plasmid-encoded TAs, in any given

bacterium, that may lead the prevention of the PSK (De Bast et al., 2008), as described below.

The Van Melderen group showed that in *Erwinia chrysanthemi*, the type II TA pair CcdAB (control of cell death) that is chromosomally-encoded by the *ccdAB* operon, which codes for the CcdA antitoxin and CcdB toxin that is a DNA gyrase inhibitor, interfered with the plasmid-encoded CcdAB (De Bast et al., 2008). The PSK mechanism which was supposed to be triggered in the plasmid-free progenies was abolished, because the chromosomally encoded antitoxin neutralized the plasmid-encoded toxin (De Bast et al., 2008). In other words, the chromosomally encoded TA pair gave a selective advantage to the plasmid-free cells by protecting them against the PSK (De Bast et al., 2008). This positive selection might be the reason for maintaining these chromosomal TA-coding genes (De Bast et al., 2008).

The same research group also proposed that the chromosomally encoded TAs might drive the evolution of the plasmid-encoded ones, where the selection of toxins that are not anymore homologous to the chromosomal TAs; and thus not recognized by the antiaddiction module, will take place (De Bast et al., 2008; Wilbaux et al., 2007). This will allow the coexistence of both systems (De Bast et al., 2008). In other words, the chromosomal TAs will lose their antiaddiction roles (Wilbaux et al., 2007) and may become devoid of physiological functions (Mine et al., 2009), as has been shown for the *E. coli* O157:H7 strain (Wilbaux et al., 2007; Mine et al., 2009; Van Melderen and De Bast, 2009). In this case, the chromosomally encoded TAs could be just traces of previous evolutionary events (De Bast et al., 2008).

7.3.6. Phage abortive infection

Bacteria have developed different mechanisms as protection against bacteriophages. One of them is known as phage exclusion or phage abortive infection (Abi) (Chopin et al., 2005). In Abi, the cell interrupts the development of the phage through a process that limits its spread to other cells and leads to an altruistic suicide of the infected cell (Chopin et al., 2005). As a result, the bacterial population, as a whole, survives (Chopin et al., 2005). It has been reported that some types of TA systems act as an Abi system. For example, in an *E. coli* strain, the chromosomally encoded type II MazEF TA pair was shown to inhibit the development of phage P1 (Hazan and Engelberg-Kulka, 2004).

Hazan and Engelberg-Kulka showed that when the prophage P1, which is inherited as an extrachromosomal plasmid, was in the lytic stage, this was a stress signal that the cell responded to (Hazan and Engelberg-Kulka, 2004). *E. coli* used the chromosomally encoded MazEF TA pair, the MazF toxin was activated; thus, the cell underwent PCD in order to protect the population from the spread of the

phage (Hazan and Engelberg-Kulka, 2004). The cell also underwent PCD due to an infection with virulent particles of the phage (Hazan and Engelberg-Kulka, 2004). Gerdes, 2013 is expecting bacteriostasis to take place as a defense mechanism against phages (Gerdes, 2013).

Conclusively, prophages and infecting bacteriophages may activate the TA systems, because they interfere with the transcription and translation of the bacterial cell. However, if the phage can inhibit the bacterial antitoxin-degrading proteases (Engelberg-Kulka and Glaser, 1999), neutralize the toxin, or be fast enough to lyse the cell before the activation of the toxin, the TA system will fail in protecting the bacterial population (Magnuson, 2007).

7.3.7. Abundance of TAs

Type II TA-coding genes are present on plasmids and chromosomes of bacteria and archaea (Pandey and Gerdes, 2005). Chromosomally encoded TAs are often present in numerous copies per strain (Pandey and Gerdes, 2005). Many of them are integrated within mobile genetic regions that form genomic islands, raising the probability that the horizontal gene transfer played a role in their distribution (Pandey and Gerdes, 2005; De Bast et al., 2008). Pandey and Gerdes found that chromosomally encoded TAs are abundant in free-living prokaryotes, and they are often present in "high numbers", where having eight or more TAs per strain was considered a high number according to the study (Pandey and Gerdes, 2005). Worth mentioning is that the total number of chromosomally encoded TAs could be way more than eight, as is the case in *M. tuberculosis* H37Rv, which harbors 79 TA pairs on the chromosome (Sala et al., 2014). Only few free-living prokaryotes lack TA-coding genes, while almost none of the obligate intracellular ones have any (Pandey and Gerdes, 2005). This observation was interpreted based on the stress-management function of TAs. Since free-living organisms live in continuously changing environments compared to the host-associated ones, they are in need for TAs to help them adapt to the changing conditions and overcome various stresses (Pandey and Gerdes, 2005; De Bast et al., 2008). Another interpretation was based on the TA gene stabilization function. Since obligate intracellular organisms have stable genomes as the mobile DNA is either absent or found in a much smaller proportion compared to the free-living organisms (Bordenstein and Reznikoff, 2005), the gene stability is not a critical issue. Thus, this could explain why obligate intracellular organisms do not harbor TAs (Pandey and Gerdes, 2005).

According to our analysis, the *Geobacillus* strains that we are interested in have a range of two to ten type II TA pairs, in addition to solo toxins/antitoxins harbored on the chromosome (as explained in more details below), but not on the plasmid (**Paper IV**). Having chromosomally encoded type II TAs is consistent

with the expectation for free-living organisms that need to adapt to fluctuations in the environment, especially because thermophilic bacteria are exposed to harsh conditions, including high temperatures and limited nutrients. The variation in the number of TAs per strain could be related to the ecological niche from which the strain was isolated, as explained in **Paper IV**.

7.4. Identification of type II TA systems in *Geobacillus*

With the increasing number of available microbial genome sequences and the variety of freely accessible web-based genome analysis tools, gaining insights on type II TA system families has flourished. Several genome-mining studies were carried out to identify putative TAs encoded on chromosomes, plasmids, and prophages, and indicated that TAs are widespread in prokaryotes, and they may exist in multiple copies (Pandey and Gerdes, 2005; Leplae et al., 2011; Makarova et al., 2009; Sberro et al., 2013; Rocker and Meinhart, 2016).

In our study (**Paper IV**), we made use of the TA finder tool (Xie et al., 2018) to mine the genome sequences of four *Geobacillus* strains, *G. kaustophilus* HTA426, *G. thermodenitrificans* NG80-2, *G. thermoleovorans* CCB_US3_UF5, and *Geobacillus* sp. ZGt-1. Selecting these type strains in addition to our strain of interest, ZGt-1, was based on the close similarity of strain ZGt-1 to *G. kaustophilus* HTA426, *G. thermoleovorans* CCB_US3_UF5 (**Paper I**). These strains were isolated from different ecological niches, *G. kaustophilus* strain HTA426 was isolated from the deep-sea sediment of the Mariana Trench (Takami et al., 2004), while *G. thermoleovorans* strain CCB_US3_UF5 was isolated from a hot spring in Malaysia (Muhd Sakaff et al., 2012). We also selected *G. thermodenitrificans* NG80-2, as it was isolated from a non-aquatic environment; from an oil reservoir in China (Feng et al., 2007). Accordingly, we expected these ecological differences to be reflected in the type II TA families that these strains harbor (**Paper IV**).

After retrieving the TA pairs using the TA finder prediction tool, we found that none of the plasmids harbored by these strains have a type II TA locus. All the identified type II TAs were encoded on the chromosome of every strain (**Paper IV**). Among the chromosomally encoded TAs, there were ones encoded within prophage islands. We excluded them from the presented study.

We manually curated the final TA finder-retrieved set of the chromosomally encoded TAs using the CDD (Conservation Domain Database) (Marchler-Bauer et al., 2017) and InterPro (Jones et al., 2014) domain analysis tools, and in some instances, we referred to the KEGG database (Kanehisa et al., 2018) (Figure 10).

We checked the aa sequences of each toxin/antitoxin and confirmed, using these tools, that the predicted protein harbors a toxin/antitoxin domain (Table 3). The confirmed toxin/antitoxin sequences were then subjected to further analyses for their characterization, using the Operon-Mapper tool (Taboada et al., 2018), the NCBI BLAST (Altschul et al., 1997), the KEGG database (Kanehisa et al., 2018), and the Prokaryotic Operon DataBase (ProOpDB) (Taboada et al., 2012) (Figure 10). One of the useful tools we used was the Operon-Mapper, which not only predicted the putative operon for every TA pair-coding genes, but also helped us in identifying additional toxins/antitoxins that were not predicted by the TA finder. These additional toxins/antitoxins were then subjected to the same analyses used for the TA finder-retrieved set. The Operon-Mapper tool also helped in identifying genes that share the same putative operon with the predicted TA pairs (**Paper IV**).

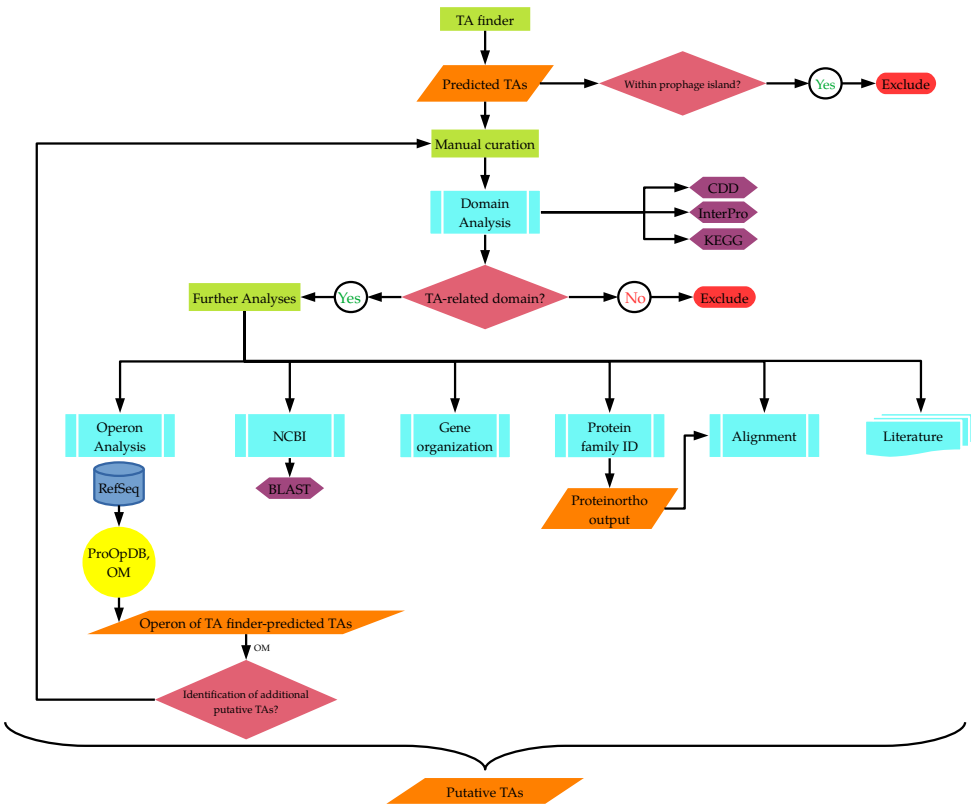


Figure 10

Workflow chart summarizing the analysis approach carried out for type II TA families identification (from Paper IV).

The final set of the resulting putative TAs was composed of 28 putative TA pairs, distributed over 8 TA families (Table 3). We also identified apparently solo (orphan) toxins and antitoxins, as shown in Table 3 below. As expected, the number of TA families varied among the 4 strains, which could be due to the differences in the ecological niches, where the strains were isolated from, as mentioned above.

Table 3

TA families, domains, and operons predicted in the study presented in Paper IV. In the last column, the term "Separate" means the toxin and antitoxin are in different operons, while "Shared" means they are in the same operon. (from Paper IV).

Strain	TA family name	Antitoxin domain	Antitoxin locus tag	Toxin domain	Toxin locus tag	TA operon
Gd †	GacTA [1]	wHTH ¤¤	GTNG_1350	GNAT ¤	GTNG_1349	Separate ‡‡
	GacT, solo toxin	N.A ±	N.A ±	GNAT ¤	GTNG_1577	Shared with GTNG_1578
	GacT, solo toxin	N.A ±	N.A ±	GNAT ¤	GTNG_1578	Shared with GTNG_1577
	GacA, solo antitoxin	HTH ¤	GTNG_1575	N.A ±	N.A ±	Shared with another protein ‡‡
	MazEF	RHH ¤¤	GTNG_0206	PemK/MazF	GTNG_0207	Shared
Gk ‡	GacTA [1]	wHTH ¤¤	GK1499	GNAT ¤	GK1498	Separate ‡‡
	MazEF (I)	AbrB/MazE ¤	GK1647	PemK/MazF ¤	GK1648	Shared
	MazEF (II)	RHH ¤¤	GK0232	PemK/MazF ¤	GK0233	Shared
	ParDE	SpoVT-AbrB (I) ¤	GK2355	ParE ¤	GK2354	Shared
	Phd-Doc	SpoVT-AbrB (II) ¤¤	GK1845	Fic/Doc ¤	GK1846	Shared
	RelBE	-/XRE #	GK3105	RelE ¤	GK3104	Shared
	VapBC	UPF[2]0175 ¤¤	GK1950	DUF[3]3368/ COG[4]2405 ¤	GK1949	Shared
	XRE-COG2856	-/HTH ǂǂ	GK3185	COG2856 ¤	GK3184	Shared with 3rd protein ‡‡
Gt §	GacTA [1]	wHTH ¤¤	*_17290	GNAT ¤	*_17280	Separate ‡‡
	MazEF (I)	MazE ¤	*_19080	PemK/MazF ¤	*_19090	Shared
	MazEF (II)	RHH ¤¤	*_2490	PemK/MazF ¤	*_2500	Shared
	MNT-HEPN (I)	NT[5]/ COG1669 ¤ (A)	*_10710	DUF[7]86 /COG2361 ¤	*_10720	Shared
	MNT solo antitoxin	NT[5]/ COG1669 ¤ (B)	Unannotated	N.A ±	N.A ±	Shared with *_10710 and *_10720
	MNT-HEPN (II)	NT[5]/KNTase ¤	*_11510	DUF[3]86 ¤/ COG2445 ¤¤¤	*_11500	Shared
	ParDE	SpoVT-AbrB (I) ¤	*_26570	ParE ¤	*_26560	Shared §§
	Phd-Doc	SpoVT-AbrB (II) ¤¤	*_21520	Fic/Doc ¤	*_21530	Shared operon
	RelBE	-/Xre #	*_34820	RelE ¤	*_34810	Shared§§
	VapBC	UPF[2]0175 ¤	*_22490	DUF[3]3368/ COG2405 ¤	*_22480	Shared
	XRE-COG2856	-/HTH ǂǂ	*_35630	COG2856 ¤	*_35620	Shared with 3rd protein ‡‡ §§
ZG ¶	GacTA	wHTH ¤¤	Contig 16_18	GNAT ¤	Contig 16_17	Separate ‡‡
	MazEF (I)	MazE ¤	Contig 16_161	PemK/MazF¤	Contig 16_162	Shared
	MazEF (II)	RHH ¤¤	Contig 4_60	PemK/MazF¤	Contig 4_61	Shared
	MNT-HEPN (I)	NT[5]/ COG1669 ¤ (A)	Contig 12_19	DUF[3]86 /COG2361 ¤	Contig 12_20	Shared
	MNT solo antitoxin	NT[5]/ COG1669 ¤ (B)	Contig 12_18	N.A ±	N.A ±	Shared with 12_19 and 12_20
	MNT-HEPN (II)	NT[5]/KNTase ¤	Contig 12_84	DUF[3]86 ¤/ COG2445 ¤¤¤	Contig 12_83	Shared
	ParDE	SpoVT-AbrB (I) ¤	Contig 23_243	ParE ¤	Contig 23_242	Shared

	Phd-Doc	SpoVT-AbrB (II) ¤¤	Contig 18_126	Fic/Doc ¤	Contig 18_127	Shared
	RelBE	-/XRE #	Contig 25_196	RelE ¤	Contig 25_195	Shared

¤ The conserved domain was inferred using CDD tool.

¤¤ The conserved domain was inferred using InterPro domain analysis tool.

¤¤¤ COG2445 was inferred using Operon-Mapper tool.

* Stands for "GTCCBUS3UF5" that is part of the locus tags in G. thermoleovorans CCB_US3_UF5.

There is no conserved domain in the antitoxin, but it is orthologous to XRE family transcriptional regulator, as shown in the KEGG Genes database/the KEGG KOALA BLAST and explained in the text.

ǂǂ There is no conserved domain in the antitoxin; however, KEGG shows that the protein motif is HTH, as explained in the text.

‡‡ Details are given in Table S3 in Paper IV.

§§ Operon prediction was based on the ProOpDB.

† *G. thermodenitrificans* NG80-2.

‡ *G. kaustophilus* HTA426.

§ *G. thermoleovorans* CCB_US3_UF5.

¶ *Geobacillus* sp. ZGt-1.

± N.A stands for "not applicable".

[1] *Geobacillus* acetyltransferase toxin-antitoxin, this TA family name was suggested in the study presented in Paper IV for *Geobacillus* strains having HTA-GNAT domain-harboring proteins.

[2] Uncharacterized Protein Family.

[3] Domain of Unknown Function, representing protein superfamily.

[4] Clusters of Orthologous Genes.

[5] Nucleotidyltransferase domain of DNA polymerase beta-like protein superfamily.

In summary, strain CCB_US3_UF5 has ten putative TA pairs, and one apparently solo putative antitoxin. Strains HTA426 and ZGt-1 have eight putative TA pairs each. While strain ZGt-1 has also one apparently solo putative antitoxin, strain HTA426 does not have solo toxins or antitoxins, as all its toxin-coding genes are paired up with neighboring antitoxin-coding ones (Table 3). The reason for the presence of a variety of TA families or the presence of more than one pair of the same TA family per strain is unknown, but there could be crosstalk among them to coordinate the cellular response to various stress conditions (Chan et al., 2012). Strain NG80-2 showed the least number of TA pairs; it showed having only two pairs, but it also has two apparently solo putative toxins, and one apparently solo putative antitoxin (Table 3).

While analyzing the aa sequences of the identified TA pairs, we found that 15 out of the 28 pairs have been annotated merely as hypothetical proteins in the NCBI. Additionally, we identified a putative antitoxin-coding gene that has been overlooked during the annotation of the genome sequence of *G. thermoleovorans* CCB_US3_UF5 (Table 4).

Table 4

Previosuly unrecognized toxins and antitoxins that have been identified in the study presented in Paper IV. These TAs have been annotated as hypothetical proteins or have not been annotated. (from Paper IV).

Strain	Genome accession no.	Putative T/AT [1]	Locus tag	Protein ID [2]
Gd †	NC_009328	RHH	GTNG_0206	WP_008881474 [3]
Gk ‡	NC_006510	AbrB	GK2355	WP_015375348 [3]
		HTH	GK3185	WP_011232655 [3]
		ParE	GK2354	WP_020278248 [3]
		RHH	GK0232	WP_011229742 [3]
		XRE	GK3105	WP_011232575 [3]
Gt §	NC_016593	AbrB	*_26570	WP_014196297 [3]
		HTH	*_35630	WP_014196828 [3]
		MNT solo antitoxin	Unannotated	WP_013146011 [4]
		ParE	*_26560	WP_014196296 [3]
		Xre	*_34820	WP_014196753 [3]
ZG ¶	LDPD00000000	AbrB	Contig 23_243	WP_015375348 [4]
		ParE	Contig 23_242	WP_020278248 [3]
		RHH	Contig 4_60	WP_011229742 [3]
		XRE	Contig 25_196	WP_082218538 [4]

† *G. thermodenitrificans* NG80-2.

‡ *G. kaustophilus* HTA426.

§ *G. thermoleovorans* CCB_US3_UF5.

¶ *Geobacillus* sp. ZGt-1.

* Stands for "GTCCBUS3UF5" that is part of the locus tags in *G. thermoleovorans* CCB_US3_UF5.

[1] Toxin/Antitoxin.

[2] Represents the RefSeq accession number of the putative toxin/antitoxin protein.

[3] Accession number belongs to the putative toxin/antitoxin protein as annotated in the RefSeq genome record of the type strain/draft genome sequence of strain ZGt-1.

[4] Accession number belongs to the NCBI blastp top hit, e-value < 10^{-20}.

7.5. Type II TA families of *Geobacillus* strains

7.5.1. GNAT-HTH (GacTA)

The GNAT-HTH TA family is composed of a toxin harboring the GNAT domain (GCN5-related N-acetyltransferases, derived initially from GCN5 (general control non-repressible 5), a histone acetyltransferase (Yeo, 2018)), and an antitoxin harboring the HTH (Helix-Turn-Helix) domain or its variants, RHH (Ribbon-Helix-Helix) (Aravind et al., 2005) or wHTH (winged Helix-Turn-Helix) (Xie et al., 2018; Hampton et al., 2018; Rivera-Gomez et al., 2017; Yamaguchi et al., 2011; Gajiwala and Burley, 2000). Some studies showed that GNAT toxin has a bacteriostatic effect (Qian et al., 2018; Cheverton et al., 2016). One study, however, reported an atypical function of GNAT, where it unexpectedly acted as a neutralizing antitoxin in *Acinetobacter baumanni* and the HTH acted as a bacteriostatic toxin (Jurėnas et al., 2017).

Our analysis indicated that each of the four *Geobacillus* strains harbors one GNAT-wHTH pair (**Paper IV**). In our study, we suggested calling this TA family "GacTA" (*Geobacillus* acetyltransferase toxin-antitoxin) in accordance with the previously reported RHH-GNAT TA families, KacAT of *Klebsiella pneumonia* (Qian et al., 2018) and TacAT of *S.* Typhimurium (Cheverton et al., 2016). However, in the case of the putative GNAT-wHTH pair in the analyzed *Geobacillus* strains, since the toxin-coding gene in this TA family precedes the antitoxin-coding one, we named it GacTA, instead of GacAT. Having the toxin-coding gene upstream of the antitoxin-coding one is considered a "reverse" order, since usually, the antitoxin-coding gene is upstream of the toxin-coding one, as mentioned above. This reverse order of genes has been reported only in few TA families; HigBA (Tian et al., 1996), MqsRA (Brown et al., 2009), and HicAB (Jørgensen et al., 2009), and has not been reported for the GNAT-HTH TA family in any of the studied prokaryotes so far. Therefore, *Geobacillus* strains seem to be the first representatives of a reverse order of the GNAT-HTH TA family. Thus, GacTA family represents another potentially "unique" TA family in addition to the HigBA, MqsRA, and HicAB, mentioned above.

The aa sequences of the GacTA TA pairs of strains CCB_US3_UF5, HTA426, and ZGt-1 are almost identical (**Paper IV**). Interestingly, in addition to harboring a putative GNAT-HTH TA pair, *G. thermodenitrificans* NG80-2 has two apparently solo putative GNAT toxins that shared an operon, and one apparently solo putative HTH antitoxin coded on the opposite DNA strand (**Paper IV**). It could be that these two putative toxins and the putative antitoxin form a three-component TA system. Toxins and antitoxins operated by different operons can, despite uncommon, interact and form a functional TA pair (Aakre et al., 2015;

Wen et al., 2014; Singh et al., 2012b). Cases of TA systems composed of three-components have been reported previously in prokaryotes (Hallez et al., 2010; Zielenkiewicz and Cegłowski, 2005). However, it has not been reported for the GNAT-HTH TA family. Additionally, in the reported three-component TA systems, there were two antitoxins and one toxin, while the GacTA family seems to have the opposite. Collectively, the GacTA family shows potential special features, such as having a reverse gene order, crosstalk among toxins and antitoxins operated by different operons, and the possibility of forming a three-component TA system composed of two toxins and one antitoxin. Carrying out experimental work will be of interest in order to prove these features.

7.5.2. MazEF

The MazEF (ma-ze means "what is it?" in Hebrew (Metzger et al., 1988) family is among the most well studied TA families (Yamaguchi and Inouye, 2011). However, it has been debated whether MazF toxin has a bacteriostatic or bactericidal effect, as mentioned above.

MazF is a ribosome-independent mRNA interferase that cleaves the mRNA at a specific RNA sequence in the absence of the ribosome (Yamaguchi and Inouye, 2011).

In the canonical type II TA system, MazE is the antitoxin that pairs with the MazF and neutralizes it. MazE mimics the RNA; it binds to the active center of the MazF and inhibits its activity (Yamaguchi and Inouye, 2011).

Beyond the canonical type II TA system, MazF may pair with other antitoxins, a phenomenon described as "mix and match" between the toxin and antitoxin superfamilies (Hayes and Van Melderen, 2011). In the *Geobacillus* strains, we found that the MazF toxins may pair with the RHH (ribbon-helix-helix), in addition to the MazE antitoxins, as discussed below.

The analysis conducted in **Paper IV** showed that strains CCB_US3_UF5, HTA426, and ZGt-1 harbor two pairs of the MazEF TA family. One is composed of the MazE-MazF composite (MazEF (I)), the MazF toxins of which have identical aa sequences in the three strains, and the same applies to the MazE antitoxins. The other MazEF TA pair is composed of the RHH-MazF composite (MazEF (II)), the MazF toxins of which also have identical aa sequences in the three strains, and the same applies to the RHH antitoxins. The putative RHH antitoxins have been annotated as hypothetical proteins by the NCBI (Table 4).

On the other hand, the strain NG80-2 has only one MazEF TA pair, composed of the RHH-MazF composite. The aa sequence of the MazF in this strain is identical to that of the MazF (II) in the three other stains, while the aa sequence of the RHH is almost identical to that of the three other stains (**Paper IV**). This

putative RHH antitoxin has also been annotated as a hypothetical protein by the NCBI (Table 4).

Overall, our results indicated that all the four strains harbor the MazEF TA family. The aa sequences of the toxins and the antitoxins of each composite are highly conserved. However, this is not the case when comparing the two composites. Although the toxin in each composite has a MazF domain, the aa sequences of the MazF toxins are different in the two composites. This finding is in line with the observations reported previously for this TA family; within one MazF protein family, the MazF toxins are diverse (Miyamoto et al., 2016). The MazEF TA pairs harbored by one strain may crosstalk to coordinate the cellular responses to environmental stresses (Tiwari et al., 2015).

7.5.3. MNT-HEPN

This TA family is composed of the MNT (Minimal Nucleotidyltransferase) protein subfamily as the antitoxin, and HEPN (Higher Eukaryotes and Prokaryotes Nucleotide-binding) protein subfamily as the toxin that has RNAse activity (Yao et al., 2015a). MNT-HEPN TA family has been reported for the first time as a novel TA pair through bioinformatic analysis by Makarova and co-workers, 2009. It is highly represented in thermophilic archaea and bacteria (Makarova et al., 2009). Our analysis indicated that two of the *Geobacillus* strains harbor this TA family, strains CCB_US3_UF5 and ZGt-1 (**Paper IV**).

Most HEPN domains harbor a catalytic motif Rx4-6H (R stands for arginine, H stands for histidine, and x stands for 4–6 of aa residues between R and H), where the residue immediately after the R is a polar aa (Jia et al., 2018). In an experimental study by Jia and co-workers, 2018, the motif Rx4-6H seemed to be responsible for the mRNAse activity of the HEPN domain-containing protein in *Shewanella oneidensis* (Jia et al., 2018).

The MNT domain-containing protein is a DNA-binding antitoxin that represses the expression of the HEPN toxin by binding to the MNT-HEPN promoter (Yao et al., 2015).

The analysis conducted in **Paper IV** indicated that two of the *Geobacillus* strains, CCB_US3_UF5 and ZGt-1, harbor two putative pairs of the MNT-HEPN TA family of different composites.

The putative MNT-HEPN (I) is composed of COG1669 (Clusters of Orthologous Group 1669), which is an MNT antitoxin-associated domain, and COG2361, which is an HEPN toxin-associated domain (Sberro et al., 2013). The aa sequences of the putative HEPN toxins in the two strains are identical and contain the catalytic motif Rx4-6H. The putative MNT antitoxins, here labeled as MNT (I), are upstream of the HEPN toxins and their aa sequences are also identical. Interestingly, the analysis conducted in **Paper IV** indicated the

presence of another putative MNT antitoxin upstream of the MNT (I), in both strains. This putative second MNT, here labeled as MNT (II), is upstream of the putative MNT (I) and shares the operon with the putative MNT-HEPN (I) TA pair. MNT (II) was annotated in the genome record of strain ZGt-1 (**Paper II**), but it was not in that of strain CCB_US3_UF5. The MNT (I) and MNT (II) have a low matching identity (**Paper IV**), which is not surprising since nucleotidyltransferase domain-harboring proteins represent a highly diverse protein superfamily (Kuchta et al., 2009). The identification of this second MNT, or the apparently solo antitoxin, shows another example of a putative three-component TA family in *Geobacillus* strains.

The putative MNT-HEPN (II) is composed of the KNTase (kanamycin nucleotidyltransferase) enzyme as the antitoxin, and the COG2445 domain as the toxin. The NTase domain of KNTases is homologous to the MNT domain (Anantharaman et al., 2013), and the COG2445 is a HEPN-associated domain (Makarova et al., 2009). The aa sequences of the putative HEPN toxins are identical between both strains and harbor the Rx4-6H motif mentioned above. They also harbor another motif, EX3KR, which is harbored by many HEPN-containing proteins (Anantharaman et al., 2013). The two putative toxins have been annotated as "DUF86 domain-containing proteins", but without stating a specific identification. DUF86 is the superfamily which the COG2445 domain belongs to (Makarova et al., 2009). Thus, the analysis conducted in (**Paper IV**) provided a more specific identification of the putative toxins. The putative gene sharing the operon with the HEPN-coding gene codes for the putative KNTase antitoxin. These aa sequences of the antitoxins are identical between both strains.

When comparing the aa sequences of the toxins of the two MNT-HEPN pairs, it can be seen that the HEPN toxins of the MNT-HEPN (I) and the MNT-HEPN (II) have a low matching identity. This is not unexpected because the HEPN-containing proteins are poorly conserved in general (Anantharaman et al., 2013).

While strains CCB_US3_UF5 and ZGt-1 have two pairs of the MNT-HEPN TA family, strains HTA426 and NG80-2 have none (**Paper IV**).

7.5.4. ParDE

In the canonical type II TA system, the ParE toxin is associated with the ParD antitoxin (Deghorain et al., 2013). ParE inhibits DNA replication by inhibiting the DNA gyrase (Yuan et al., 2010; Jiang et al., 2002). This toxin could either have a bacteriostatic effect, as is the case in *Caulobacter crescentus* (Fiebig et al., 2010), or a bactericidal effect, as is the case in *Escherichia coli* O157:H7 (Hallez et al., 2010). Muthuramalingam and co-workers, 2018 found that the ParE has a dose-dependent dual effect in *Pseudomonas aeruginosa*. At low concentrations, ParE showed a protective effect, where it protected the producing

bacterium from antibiotics that targeted the DNA gyrase (Muthuramalingam et al., 2019). On the other hand, at high concentrations, ParE became toxic to the producing cell (Muthuramalingam et al., 2018). The molecular mechanism of the switch between protection and lethality has not been revealed yet (Muthuramalingam et al., 2018).

The ParE may associate with antitoxins other than the ParD, such as antitoxins harboring the SpoVT-AbrB DNA-binding domain (SpoVT stands for Stage V sporulation protein T, and AbrB stands for AidB regulator domain). The SpoVT-AbrB-domain-harboring antitoxins belong to the superfamily of AbrB-like transcription factors (Coles et al., 2005). Members of this superfamily are homologous to the AbrB protein of *B. subtilis* (Coles et al., 2005). The AbrB protein is an essential transcriptional regulator in *B. subtilis* that regulates the cell response to stressful environments, via regulating the expression of at least 60 genes (Chan et al., 2016; Coles et al., 2005), and it is homologous to the antitoxins MazE, PIM-I, and VapB (Coles et al., 2005). SpoVT and AbrB share homology and in the type II TA family classification by Ou et al., 2013, SpoVT-AbrB domain-harboring antitoxins are described as AbrB antitoxins (Ou et al., 2013).

The analysis conducted in Paper IV indicated that three *Geobacillus* strains, CCB_US3_UF5, HTA426, and ZGt-1, have a putative ParE-AbrB TA composite. This composite has not been reported previously. We considered this TA composite as a ParDE family based on the classification by Ou and co-workers, 2013, where RelE associated with AbrB represents the RelBE TA family (Ou et al., 2013). Since the RelE and ParE belong to the same protein superfamily, it is not surprising for the ParE toxin to associate with the AbrB, and this composite may represent the ParDE TA family. Additionally, this classification considers AbrB-MazF, AbrB-PIN, and AbrB-Doc as composites of the MazEF, VapBC, and Phd-Doc TA families, respectively (Ou et al., 2013).

The analysis showed that the ParE toxins of strains HTA426, and ZGt-1 have almost identical aa sequences (Paper IV) and that they have been annotated as hypothetical proteins by the NCBI (Table 4). The aa sequences of the putative AbrB antitoxins are identical between both strains and have also been annotated as or showed 100% identity to hypothetical proteins (Table 4).

The putative ParE toxin and the AbrB antitoxin of strain CCB_US3_UF5 have also been annotated as hypothetical proteins (Table 4) but did not show 100% alignment with those of the other two strains (Paper IV).

Although three strains harbor putative ParDE TAs, strain NG80-2 does not harbor this TA family (Paper IV).

7.5.5. Phd-Doc

The Phd-Doc family is among the least distributed TA families (Garcia-Pino et al., 2013). In the canonical system, the toxin, Doc ‘‘Death on curing” is associated with the antitoxin, Phd “Prevents host death” (Garcia-Pino et al., 2013). However, the “mix and match” phenomenon applies here as well. The toxin Doc may associate with other antitoxins, and the antitoxin Phd may associate with other toxins (Garcia-Pino et al., 2013).

The Doc toxin belongs to the Fic (Filamentation induced by cyclic AMP) protein superfamily (Cruz et al., 2014) and has a bacteriostatic effect (Cruz et al., 2014; Castro-Roa et al., 2013). All Doc toxins act as kinases (Garcia-Pino et al., 2013; Cruz et al., 2014). As opposed to most of the TA system toxins, the Doc toxin does not cleave the mRNA (Garcia-Pino et al., 2013; Cruz et al., 2014). In fact, the expression of Doc leads to the stabilization of mRNA (Cruz et al., 2014; Liu et al., 2008). Doc toxicity stems from its interaction with the 30S bacterial ribosomal subunit and inactivation of the EF-Tu (translation elongation factor thermo unstable), leading to a rapid inhibition of protein synthesis and cell growth arrest (Cruz et al., 2014; Castro-Roa et al., 2013; Liu et al., 2008). The cell that undergoes dormancy due to the Doc toxicity can return to growth after reactivation of the EF-Tu (Cruz et al., 2014; Castro-Roa et al., 2013).

Other than Phd antitoxins, AbrB antitoxins may associate with Doc toxins (Coles et al., 2005; Ou et al., 2013), forming an AbrB-Doc TA composite, which is classified as the Phd-Doc TA family (Ou et al., 2013). Three of the *Geobacillus* strains analyzed in **Paper IV** harbor this composite, as discussed below.

Strains CCB_US3_UF5 and ZGt-1 have putative Doc toxins, whose aa sequences are identical, and the aa sequences of their putative AbrB antitoxins are almost identical (**Paper IV)**. Strain HTA426 also harbors a putative AbrB-Doc TA composite. The aa sequence of its putative Doc toxin is almost identical to those of the other two strains, but its putative AbrB antitoxin is shorter than the AbrB antitoxins of strains CCB_US3_UF5 and ZGt-1, and a few of its aa residues are not aligned with the sequences of these two antitoxins. On the other hand, the analysis showed that strain NG80-2 does not harbor a Phd-Doc TA family.

7.5.6. RelBE

RelBE (Relaxed BE) is one of the best-described TA families (Gerdes, 2013). In the canonical type II TA system, the toxin, RelE is associated with the antitoxin, RelB (Gotfredsen and Gerdes, 1998; Gerdes, 2013). RelE is an mRNA interferase; it cleaves the mRNA at the A-site of the ribosome (Overgaard et al., 2009). Therefore, it interferes with protein synthesis in a ribosome-dependent way (Gerdes, 2013). RelE does not cleave naked mRNA (Gerdes, 2013) and the inhibition of the protein synthesis is reversible (Pedersen et al., 2003). Thus, RelE has a bacteriostatic effect (Pedersen et al., 2003).

RelB is not the only antitoxin that associates with the RelE. The RelE toxin may associate with other antitoxins, such as the XRE (Xenobiotic Response Element) family proteins, forming the XRE-RelE composite (Ou et al., 2013).

The XRE protein family is a large family of transcriptional regulators and is among the most widespread regulators in bacteria (Novichkov et al., 2013). XRE proteins may act as antitoxins. Recently, it has been experimentally proved that the C-terminal of an XRE protein extends to block the active site of the pairing toxin (Skjerning et al., 2019). The N-terminal has an HTH DNA-binding motif that allows the protein to bind to the TA promoter and regulate the gene transcription. XRE has been experimentally validated in *B. subtils* 168 as the main protein that controls bacterial suicide (McDonnell and McConnell, 1994; McDONNELL et al., 1994), indicating its potential activity as an antitoxin.

The analysis carried out in **Paper IV** indicated that three of the analyzed *Geobacillus* strains have putative XRE-RelE TA composites.

Strains CCB_US3_UF5, HTA426 and ZGt-1 code for putative RelE toxins, whose aa sequences showed ≥ 92% identity in the three strains. The putative antitoxins are ≥ 93% identical between the three strains and are orthologous to "XRE family transcriptional regulators", but they have been annotated as or showed identity to hypothetical proteins (Table 4). On the other hand, the analysis showed that strain NG80-2 does not harbor the RelBE TA family.

7.5.7. VapBC

VapBC (Virulence associated protein) is the most widespread TA family in bacteria and archaea (McKenzie et al., 2012), but the least well-described among the different families (Ning et al., 2013; Sala et al., 2013). The canonical type II TA system, VapC toxin is associated with the VapB antiotoxin. The VapC toxin harbors a PIN domain, and the DNA-binding VapB antitoxins belong to families of transcriptional regulators, such as the AbrB, HTH, RHH, or Phd, as shown for *M. tuberculosis* which harbors 50 VapBC pairs (Ning et al., 2013, Sala et al., 2013).

The PIN (PilT N-terminus) domain is a type of pili protein originally annotated in *Myxococcus xanthus* (Ning et al., 2013) and is found in all kingdoms of life (Sala et al., 2013). The PIN domain is associated with the ribonuclease activity of the VapC toxin (Winther and Gerdes, 2011; McKenzie et al., 2012). VapC uses various mechanisms that target RNAs. For example, its toxicity may result from the degradation of mRNA transcripts, and thus it inhibits protein translation (McKenzie et al., 2012). In the pathogenic *Shigella flexneri* and *Salmonella enterica*, VapC was found to act as a tRNAse, cleaving the initiator tRNA, but its toxicicty is reversible and bacteriostatic (Winther and Gerdes, 2011). On the other hand, VapC toxicity may also stem from stable RNA binding, as is the case shown for part of the VapC toxins of *M. tuberculosis* (reviewed by Sala et al., 2013). Interestingly, other VapC toxins of *M. tuberculosis* were shown to have ribonuclease activity *in vitro* (Sala et al., 2013).

There are also PIN-like domains, such as the COG2405 domain, that act as toxins (Matelska et al., 2017). The COG2886 DNA binding proteins may associate with the PIN toxins and act as antitoxins (Ou et al., 2013).

The analysis conducted in **Paper IV** showed that two of the four *Geobacillus* strains harbor the putative VapBC family.

Strains HTA426 and CCB_US3_UF5 have putative COG2405 toxins, whose aa sequences are identical between the two strains. The strains harbor putative COG2886 antitoxins and their aa sequences are also identical. On the other hand, the analysis indicated that strains NG80-2 and ZGt-1 do not harbor the VapBC family.

7.5.8. XRE-COG2856

The XRE-COG2856 is a potential novel TA family that has not been experimentally studied yet (Makarova et al., 2009; Makarova et al., 2013). It was discovered by Makarova and co-workers in 2009 (Makarova et al., 2009) based on an *in silico* analysis and found to be abundant in the genomes of bacteria, archaea, and phages (Makarova et al., 2009; Makarova et al., 2013).

The COG2856 toxin is a metzincin Zn-dependent protease, and the antitoxin is an HTH domain-harboring protein of the XRE-family (Makarova et al., 2009). The HTH-domain is often fused with the COG2856 domain in a single protein (Makarova et al., 2009; Makarova et al., 2013). However, the *Geobacillus* strains analyzed in **Paper IV** seem to have two separate proteins forming the XRE-COG2856 TA family.

The analysis conducted in **Paper IV** showed that two of the strains, CCB_US3_UF5 and HTA426, harbor this TA family. The aa sequences of the two putative COG2856 toxins are > 99% identical between these two strains. The two putative toxins contain the conserved HEXXH motif, where X is any aa. This

motif is the Zn-binding catalytic active site of the metzincin Zn-dependent proteases, mentioned above (Makarova et al., 2009). The putative antitoxins are ~98% identical between these two strains and seem to harbor the HTH domain. Thus, they might represent putative antitoxins of the XRE-protein family.

The analysis indicated that strain ZGt-1 does not have this TA family. On the other hand, strain NG80-2 harbors the genes coding for this TA family, but they are encoded within the prophage region and will be covered in a future study.

Regulation of the xre-cog2856

The analysis conducted in **Paper IV** showed that the *xre-cog2856* genes share the operon with a third gene coding for a hypothetical protein in both strains. This hypothetical protein seems to harbor the 7TMR-HDED domain (7 transmembrane helices receptors-HD hydrolase; a hydrolase with a catalytic His-Asp (HD) motif, and ED stands for extracellular domain) (Anantharaman and Aravind, 2003; Huynh et al., 2015). This domain is expected to be involved in signal detection and transmission to the cellular machinery, which leads to triggering a response to the environmental conditions (Anantharaman and Aravind, 2003; Huynh et al., 2015). The presence of the 7TMR-HDED receptor suggests that the hypothetical protein is probably regulated by a second messenger and in turn, it regulates other proteins. This messenger is likely to be the bacterial cyclic di-3′,5′- adenosine monophosphate (c-di-AMP), since it is transmitted by the 7TMR domains (Huynh et al., 2015). This suggestion is supported by the presence of a neighboring gene coding for the enzyme "bifunctional 2',3'-cyclic-nucleotide 2'-phosphodiesterase/3'-nucleotidase (EC:3.1.4.16 3.1.3.6)". This enzyme was experimentally proved to function as a phosphodiesterase, hydrolyzing the c-di-AMP (Andrade et al., 2016). Accordingly, we could assume that the hypothetical protein sharing the operon with the putative *xre-cog2856* genes is regulated by the c-di-AMP.

The c-di-AMP signaling molecule is a recently discovered second messenger (Witte et al., 2008). It is synthesized by many bacteria and archaea (Gundlach et al., 2015), and is secreted into the extracellular space (Huynh et al., 2015, Andrade et al., 2016). This secretion could be related to stress responses (Huynh and Woodward, 2016). The existence of a protein that seems to be regulated by the c-di-AMP in the same operon with the genes coding for the XRE-COG2856 suggests a functional link between the c-di-AMP and this TA pair. The c-di-AMP is interconnected with the "stress messenger" or the "alarmone" (p)ppGpp (guanosine tetra or pentaphosphate) (Gross et al., 2006; Corrigan et al., 2015; Huynh et al., 2015; Gundlach et al., 2015; Irving and Corrigan, 2018; Zarrella et al., 2018). The (p)ppGpp mediates the stringent response that allows bacteria to adapt to stresses and promotes survival (Irving and Corrigan, 2018; Hauryliuk et

al., 2015; Gerdes et al., 2005). Therefore, it is unsurprising that the TAs are influneced by the high levels of (p)ppGpp, as explained below.

The (p)ppGpp is a global regulator of the cellular metabolism in response to changes in the environment (Gerdes et al., 2005; Srivatsan and Wang, 2008) and Wang, 2008). During starvation, the (p)ppGpp synthesizing enzyme, RelA, is activated by the uncharged tRNA at the ribosomal A-site, and the degradation of the (p)ppGpp is inhibited; thus, the (p)ppGpp concentration increases (Gerdes et al., 2005). An interaction between the (p)ppGpp, RNA polymerase (RNAP), sigma factors, and the cofactor DksA (*dnaK* suppressor) takes place (Srivatsan and Wang, 2008). The (p)ppGpp represses the transcription of stable RNAs (tRNA and rRNA) and ribosomal proteins, and induces the transcription of aa, via directing the RNAP away from the synthesis of RNAs and towards the aa biosynthesis (Gerdes et al., 2005; Srivatsan and Wang, 2008). The transcription factor Dksa helps the (p)ppGpp exert its full effects (Gerdes et al., 2005; Srivatsan and Wang, 2008). The (p)ppGpp also alters the utilization of sigma factors, where factors directing the transcription of stress-related genes, such as σ^N and σ^S, are utilized (Gerdes et al., 2005; Srivatsan and Wang, 2008). The (p)ppGpp also leads to the accumulation of the polyphosphate (PolyP) via inhibiting the exopolyphosphatase (Gerdes et al., 2005). The PolyP binds to the Lon protease, which is a DNA-binding ATP-dependent enzyme, and directs it to degrade the idling ribosomal proteins, in order to generate aa for protein synthesis (Gerdes et al., 2005). Lon is also activated by an unknown regulatory component and this leads to the degradation of antitoxins. Accordingly, toxin activation is one of the results of the high levels of (p)ppGpp (Gerdes et al., 2005).

In strains CCB_US3_UF5 and HTA426, the presence of a gene that is potentially coding for a protein possibly regulated by the c-di-AMP in the same operon with the *xre-cog2856* loci, and the interconnection between the c-di-AMP and the (p)ppGpp, and their roles in triggering the bacterial responses to stresses imply that the c-di-AMP is likely to be involved in the regulation of the *xre-cog2856* gene expression. In **Paper IV**, we have proposed a hypothesis for the potential roles of the c-di-AMP and (p)ppGpp in the regulation of the *xre-cog2856* gene expression, as discussed below.

Hypothesis — Regulation of the xre-cog2856 *expression*

As illustrated in Figure 11, when bacterial cells are experiencing stress, either the c-di-AMP levels or the (p)ppGpp levels will increase first, and the levels of the other messenger will follow since they are interconnected, as mentioned above. The synthesized c-di-AMP molecules are expected to be secreted into the extracellular space and the "hypothetical" protein of strains CCB_US3_UF5 and HTA426 is expected to sense it via the extracellular receptor domain (7TMR-HDED) mentioned above. The protein is likely to experience conformational changes then and could function as a regulatory protein that negatively regulates the expression of the adjacent *xre-cog2856* TA loci. This regulation could take place via one or more quorum sensing molecules that could inhibit the *xre-cog2856* expression, under stress conditions. Accordingly, the synthesis of the XRE-COG2856 and the "hypothetical protein" will stop. As a result, the unstable antitoxin will be degraded by the Lon protease, which is stimulated by the high level of (p)ppGpp (Hauryliuk et al., 2015), and the toxin will be activated. Consequently, the cell will undergo growth arrest. When the conditions improve, the levels of (p)ppGpp and c-di-AMP will decrease and the expectedly stable "hypothetical protein" will sense that via the receptor and unblock the expression of the *xre-cog2856* via one or more quorum sensing molecules. The cell could then resume its growth provided that the COG2856 did not have a bacteriocidal effect, or the cells had not reached the "point of no-return" described by (Amitai et al., 2004). There have not been experimental studies carried out on the XRE-COG2856 TA pair; therefore, it is still unknown whether COG2856 has a bacteriostatic effect or a bactericidal effect.

Understanding the mechanisms of action and regulation of this TA pair will broaden our knowledge on its activity as well as on the roles of the c-di-AMP and (p)ppGpp in regulating it, especially since the field of TA systems and the c-di-AMP signaling are still in their infancy.

The results of various previous studies were linked together to form this conjecture on the potential roles of c-di-AMP and (p)ppGpp in the regulation of the *xre-cog2856* expression. These studies are discussed in **Paper IV**.

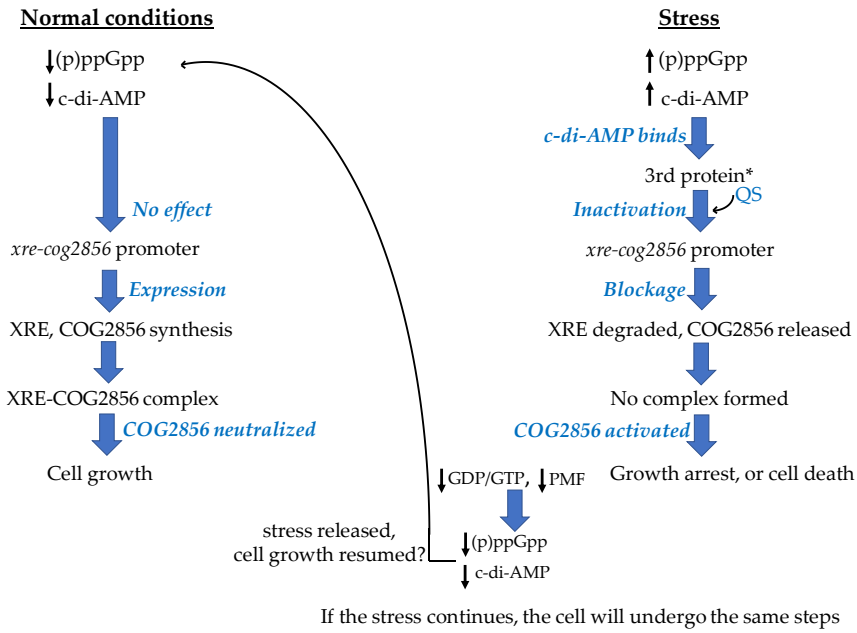


Figure 11

Shematic illustration of the hypothesis proposed in Paper IV on the regulation of *xre-cog2856* expression by the signaling messengers, (p)ppGpp and c-di-AMP. * 3rd protein represents the hypothetical protein sharing the operon with the xre-cog2856 genes. QS stands for quorum sensing molecule(s). Scheme inspired by Gross et al., 2006. (Figure from Paper IV).

7.6. Summary of the features of the *Geobacillus* type II TA families

The study in **Paper IV** was conducted in an attempt to gain a picture on the type II TA system in geobacilli, since TA systems of thermophilic bacteria in general have not been well-studied.

In total, we identified 28 putative TA pairs, distributed over 8 TA families (Table 3).

- Out of these identified pairs, 15 have either been annotated as hypothetical proteins in their genome records or have been overlooked during the genome annotation.
- We also identified apparently solo (or orphan) toxins and antitoxins, that could be part of a 3-component TA system.
- The TA families, GacTA (GNAT-HTH) and MazEF were detected in all the four strains, while other TA families were less represented.
- The GacTA family has the TA loci in a reverse order, where the toxin-coding gene precedes that of the antitoxin. Only few TA families have been reported to have this reverse order; therefore, the GacTA represents a unique TA family.
- In the same strain, some TA families were found as two pairs, whose aa sequences are diverse, such as the MazEF and the MNT-HEPN TA families.
- For each TA family, the TA pairs are, overall, highly conserved among the strains.
- *G. thermoleovorans* CCB_US3_UF5 harbors the highest number of TAs, with 10 pairs of 8 TA families. This could be because the strain was isolated from a terrestrial hot spring, where the temperatures, nutrients and pH fluctuate continuously, as mentioned in Chapter 4 (Shah et al., 2013). Thus, the cells need to be ready for adapting to the changing environment.
- *G. thermodenitrificans* NG80-2 have the least total number of TA pairs, which are also less conserved compared to the same pair in the other analyzed strains. This could be related to the spot from which the strain was isolated, an oil field (Feng et al., 2007), while the other strains were isolated from aquatic environments.
- Some *Geobacillus* strains seem to have a potentially new TA composite of the ParDE TA family, the AbrB-ParE composite.
- The XRE-COG2856 TA pair of geobacilli seems to have an uncommon pattern of unfused coding genes.

These features hint for the significance of studying the TA systems of geobacilli experimentally.

7.7. Potential applications of the type II toxin-antitoxin system in the pharmaceutical industry

Due to the demand for discovering new antimicrobial agents, exploring novel sources and alternatives is needed. TA systems represent a potential solution, since they interfere with biological processes and their targets are similar to those of antibiotics (Wen et al., 2014). TA systems could be applied as antimicrobial agents, or could be targeted by novel antimicrobial agents, as discussed below.

7.7.1. Type II toxin-antitoxin system families as antibacterial agents

Thanks to the various bacteriostatic and bactericidal effects of the toxins of the type II TA families, they represent a potential alternative to conventional antibiotics. However, there are several challenges that need to be overcome before successfully making the toxin druggable (Chan et al., 2015). These challenges are related to the features of the toxins as well as their *in vivo* behavior. For example, one toxin may have more than one cellular target (Chan et al., 2015). These targets could be in the pathogen as well as in the commensal flora and human cells (Chan et al., 2015), as discussed below. Additionally, even toxins that have structural similarity may target different cellular processes; therefore, investigating each toxin on its own has to be carried out, rather than extrapolating data (Chan et al., 2015).

In addition to the significance of characterizing the structure and function of the toxins, their *in vivo* behavior is also of critical importance, as is the case with any potential drug. For example, if a toxin or a combination of two were selected for testing their druggability, their interaction with each other and with the human cells have to be fully understood (Chan et al., 2015). Bacterial toxins might harm human cells when acting as endoribonucleases, as they cleave mRNAs regardless of their origin (Chan et al., 2015). This could be controlled by adjusting the toxin dosage (Chan et al., 2015). Bacterial toxins might also harm the commensal flora, in the same way as antibiotics do; therefore, replenishment will be needed (Chan et al., 2015). The delivery system of the toxin to the right tissue target is also an essential issue to consider (Chan et al., 2015). The intestinal absorption of the toxin, its metabolic stability, and its distribution inside the body and excretion outside it; the so-called “ADME”, must be assessed (Chan et al., 2015).

Toxins could be used as a "standalone" or in combination with conventional antibiotics, which will then require matching the toxin with the antibiotic formulation to generate a new combined drug (Chan et al., 2015).

7.7.2. Type II toxin-antitoxin system families as targets of antimicrobial agents

Type II TA systems are present in almost all pathogenic bacteria (Kang et al., 2018). Therefore, novel antimicrobial drugs that could target them and trigger the toxin to exert its bacteriostatic or bactericidal effect could be a good strategy (Kang et al., 2018).

Various approaches have been proposed to apply the "artificial activation of toxins" strategy, as discussed below:

- Artificial activation of toxins using a drug that disrupts the TA complex or prevents its formation will lead the free toxin to exert its toxic activity (Unterholzner et al., 2013; Kang et al., 2018). A peptide that outcompetes the antitoxin to prevent its interaction with the toxin and releases the latter from the TA complex, or one that mimics the region of the antitoxin which binds to the toxin could serve the purpose (Williams and Hergenrother, 2012; Kang et al., 2018).
  In a study carried out by Chopra and co-workers on the inhibition of the MoxXT TA complex formation in *Bacillus anthracis*, a designed peptide prevented the complex formation by occupying the binding interface between MoxX and MoxT. However, this peptide had a partial inhibition of the ribonuclease activity of the MoxT (Chopra et al., 2011).
  The Engelberg-Kulka group identified a quorum-sensing pentapeptide in *E.coli*, named the "Extracellular Death Factor" (EDF), that acts as an activator of the MazF toxin, which in turn triggers the PCD (Kolodkin-Gal and Engelberg-Kulka, 2006). The same group reported that the MazEF-mediated death in *E. coli* was also triggered by EDFs of other bacterial species, the Gram-negative bacterium *Pseudomonas aeruginosa* and the Gram-positive bacterium *Bacillus subtilis* (Kumar and Engelberg-Kulka, 2014). Moreover, the EDFs enhanced the endoribonucleolytic activity of the *E. coli* MazF (Kumar and Engelberg-Kulka, 2014). Therefore, the activation of the MazF could be triggered extracellularly and even enhanced by the addition of the EDF of different bacterial origins. This could constitute a novel antibacterial strategy (Kumar and Engelberg-Kulka, 2014). However, activation of toxins of TA families other than the MazEF might not be triggered by the EDFs (Chan et al., 2015).

- Activation of the antitoxin-degrading proteases; Lon or Clp could disrupt or prevent the formation of the TA complex indirectly (Unterholzner et al., 2013; Williams and Hergenrother, 2014).
- Repression of the TA operon transcription to stop the de novo synthesis of the antitoxin, by using a drug that binds to the TA promoter DNA (Williams and Hergenrother, 2014; Chan et al., 2015; Kang et al., 2018). Although this will stop the *de novo* synthesis of the toxin as well, the already synthesized toxin will be free and active, because the previously synthesized antitoxin will have been degraded due to its instability (Chan et al., 2015; Kang et al., 2018).
- Interference with the autorepression of the TA operon, as an alternative to the previous approach. Using a drug that interferes with the negative autoregulation of the TA promoter by the antitoxin and the TA complex will result in an increase in the TA complex pool (Unterholzner et al., 2013). This should then be followed by removing the drug; thus, the abundant TA complexes will negatively autoregulate the TA operon, leading to the repression of the de novo synthesis of the antitoxin. After the degradation of the latter, a relatively high level of the toxin will be released, and thus will exert its toxic activity (Unterholzner et al., 2013).
- Activation of the plasmid-encoded TAs could be applied by using a drug that interferes with the plasmid replication, causing plasmid loss, and thus leading to cell death by the PSK mechanism (Unterholzner et al., 2013).

When applying the strategy of artificial activation of toxins, several challenging issues that are also related to the features of TA pairs have to be taken into consideration. For example, toxins might be neutralized by various antitoxins (Chan et al., 2015). TA pairs could crosstalk, and copies of the same toxin within the cell could also interact, and the same applies to closely related antitoxins (Chan et al., 2015). Accordingly, a thorough understanding of the structure and function of the TA pair of interest is required (Chan et al., 2015). Defining the right TA pair to be targeted will be followed by identifying or designing a drug that activates the toxin, via any of the mentioned approaches, without inhibiting its enzymatic activity (Unterholzner et al., 2013; Williams and Hergenrother, 2014; Kang et al., 2018). There are also issues that have to be considered here, as discussed below.

As mentioned above, the dosage of the drug should be carefully adjusted to avoid harming the human cells (Chan et al., 2015). Designing a drug that activates a toxin that cannot target human cells could be a good approach (Chan et al., 2015). The ideal case would be to design a drug that specifically targets the

pathogen in question to avoid affecting the commensal flora and does not activate toxins that can harm the human cells.

Most toxins have a bacteriostatic effect that could be reversed by the antitoxins, but after exposing the pathogen to the toxin for a long time, cell death could occur, as was reported for the MazF (Amitai et al., 2004). Yet, the risk of persister cell formation is a critical issue. However, this could be solved via developing anti-persister drugs, like the PZA which is used for treating tuberculosis by acting only on non-growing persister cells.

7.7.3. Type II toxin-antitoxin system families as antiviral agents

The treatment of viral infections is one of the most difficult tasks in medicine (Shapira et al., 2012). Although vaccines are considered as the most potent way of combating viral infections, there is still no vaccine active against the human immunodeficiency virus (HIV) and hepatitis C virus (HCV) (Shapira et al., 2012). However, some studies have indicated the potential of the type II TA system as antiviral agents; particularly, the MazF toxin seems to be a good antiviral candidate.

In a study by Chono and co-workers, human T-lymphoid cells were transduced with a retroviral vector that had the *E. coli* MazF gene under the control of the HIV-1 TAR (the transactivation response) element (Chono et al., 2011). During the early stage of the HIV-1 infection, the HIV-encoded transactivator of transcription (Tat) regulatory protein binds to the TAR sequence in order to express the other viral proteins (Chono et al., 2011). The principle of this antiviral strategy using the MazF toxin was to destroy the HIV-1 transcript upon infection (Chono et al., 2011). To apply this principle, a Tat-dependent MazF expression system was constructed in a retroviral vector, where the *mazF* gene was inserted downstream of the TAR element (Chono et al., 2011). This vector was then transduced into CD4+ T-lymphoid line CEM-SS cells, which are profoundly affected by the HIV infection (Chono et al., 2011). When the transduced cells were infected with the HIV-1 virus, the Tat protein bound to the TAR sequence of the vector, which thus induced the MazF transcription (Chono et al., 2011). As a result, the active MazF cleaved the viral mRNA and blocked the virus replication (Chono et al., 2011). Interestingly, neither the cell growth nor the number of the T-lymphoid cells were affected, because the level of the induced MazF was not enough to harm the cells (Chono et al., 2011). Thus, it seems that the MazF could be used in anti-HIV gene therapy (Chono et al., 2011).

In another example, MazF was used for the removal of hepatitis C virus-infected hepatocytes (Shapira et al., 2012). The principle of this strategy relied on exploiting the specificity of the viral proteases, instead of inhibiting the viral proteins, to target only the infected cells (Shapira et al., 2012; Chan et al., 2015).

A recombinant adenoviral vector was used to deliver the MazF gene to the HCV-infected cells (Shapira et al., 2012). The constructed vector coded for a single polypeptide, where the MazF toxin and the MazE antitoxin were linked via a linker that was cleavable by one of the HCV's proteases, NS3 (Shapira et al., 2012). When the vector was delivered to the cells, the HCV-infected cells had the NS3 which cleaved the linker; thus, the MazF was freed its complex with the MazE and exerted its ribonuclease activity (Shapira et al., 2012). This resulted in the cleavage of the cellular mRNA and apoptotic cell death (Shapira et al., 2012). Interestingly, the MazF showed only a shallow level of toxicity to the healthy cells (Shapira et al., 2012). This could probably be solved in future studies.

Overall, exploiting the toxins of the type II TA system, and probably the other types as well, or targeting them seems to be a promising approach, but further in-depth studies are demanded in order to understand these systems and overcome the obstacles to the practical application of various TA families as antibacterial and antiviral agents. Using the TA systems as antibacterial agents will have broad applications in different sectors, including the food and wood industries.

8. Conclusions

The results of the studies conducted in this thesis proved the hypothesis that thermophilic bacteria living in hot springs produce antibacterial substances that can be active even against mesophilic bacteria. These studies represent the first report on antimicrobials of thermophilic bacteria living in hot springs.

The work carried out in this thesis presented two approaches for revealing the antimicrobial potential of a bacterial strain. The first is a combinatorial approach where the classical microbiology culture-based and whole-cell screening methods were applied and followed with advanced methods of proteomics and bioinformatics (**Paper I**). The second is pure *in silico* approach, where genome sequences were analyzed, and potential antimicrobial peptide/protein-coding genes were identified as promising subjects for future experimental studies (**Papers II, III,** and **IV**). The *in silico*-predicted results aim for saving time and efforts, as they may direct the future experimental work.

In summary, we isolated *Geobacillus* sp. ZGt-1 from Zara hot spring in Jordan, and the results indicated that the strain represents a potential source of antibacterial peptides and proteins, as summarized below.

- The strain is capable of secreting SDS-resistant thermostable antibacterial proteins that antagonized the growth of *G. stearothermophilus*, a dairy- and food- spoiling thermophilic bacterium. The secretion of the antibacterial proteins was enabled by immobilizing the cells and cultivating them in sequential batches with cell-recycling (**Paper I**). Using the agar gel as a matrix for cell-immobilization makes this approach applicable in the food industry (Doleyres and Lacroix, 2005).
- Three uncharacterized/hypothetical proteins of ZGt-1, proteins 23_543, 6_6, and 4_4, seemed to be associated with this antibacterial activity (**Paper I**).
- Two enzybiotics, amidase and DD-carboxypeptidase of ZGt-1 might have played a role in the antagonistic activity (**Paper I**).
- The strain also antagonized the growth of *Bacillus subtilis,* and the pathogenic *Salmonella.* Typhimurium (**Paper I**).
- The *in silico* analyses of the ZGt-1 genome indicated that the strain harbors a lanthipeptide and a non-lanthipeptide bacteriocins (**Paper II**).
- The predicted lanthipeptide of the ZGt-1 strain, Z-geobacillin, represents a putative novel one that could be more stable than the commercially available nisin (**Paper III**).

- Seven putatively novel lanthipeptides produced by firmicutes, other than *Geobacillus* sp. ZGt-1 were also identified (**Paper III)**.
- The *in silico* analyses also revealed the lanthipeptide production potential of bacterial strains that have not been recognized previously as lanthipeptide-producers (**Paper III)**.
- The *in silico* analyses revealed the potential of strains of *Geobacillus* spp. to produce various groups of the type II toxin-antitoxin system.
- Among the predicted toxin-antitoxin (TA) pairs of the analyzed *Geobacillus* strains, 15 have either been annotated as hypothetical proteins in their genome records or have been overlooked during the genome annotation; however, they have been identified as toxins/antitoxins in the study conducted in **Paper IV**.
- Among the predicted TA pairs, a new TA composite, AbrB-ParE was suggested (**Paper IV**).
- The analyzed *Geobacillus* strains showed one of the unique TA families, which we suggested to name GacTA, where the toxin-coding gene is upstream of the antitoxin-coding one.
- The *in silico* analyses also indicated that the ZGt-1 strain harbors eight TA pairs belonging to different type II TA families (**Paper IV)**.
- The gene expression regulation of the XRE-COG2856 TA family has not been understood yet. We proposed a hypothesis for the regulation of the *xre-cog2856* (**Paper IV**).

These results pave the way for future experimental studies that aim for the upstream and downstream processing of these antimicrobial substances, in order to characterize them, understand their mechanisms of action, and evaluate their applicability as food additives and pharmaceuticals.

Acknowledgments

My extensively creative and skilful supervisor; Björn Canbäck… I can never thank you enough for your brilliant, pedagogical, critical, and supportive guidance… I owe you tremendously!!

My exceptional parents and brother… I am much lucky to have you... Without the three of you, I don't know where I would be... I wish there were words that could ever express my endless love and gratitude!!

…. And a bunch of thanks to all of you!

Ann Kull, Anna Carlqvist, Annika Mårtensson, Bengt O. Bengtsson, Bo Mattiasson, Bodil Ryderheim, Carin Jarl-Sunesson, Carl Grey, Cecilia Emanuelsson, Charlotte Marklund, Christina Rengefors, Christina Wennerberg, Claes von Wachenfeldt, Dag Ahrén, David Bollivar, Diya Adawi, Ed van Niel, Elisabet Gersbro-Bülow, Eelisabeth Barane, Emma Kreuger, Emma Poaches, Eva Nordberg-Karlsson, Eva Svensson, Frans-Peder Nilson, Fredric Carlsson, Gashaw Mamo, Gerda Nilsson, Hadiye Kaddoura, Hanan Ahmad, Hanna Glad, Hussein Masoud, Jenny Schelin, Jep Agrell, Joel Wallenius, Johan Bentzer, Johan Malmström, Johanna Olsson, Jörgen Nilsson, Jóhanna Bjarney Jónsdóttir, Julienne Stewart-Sandgren, Kai Blin, Karin Rengefors, Katja Bernfur, Khalil Ahmad, Klas Flärdh, Ladaea Rylander, Lars Fredriksson, Lars Hederstedt, Linda Jansson, Lovisa Björnsson, Malin Matsson, Maria Bajuk, Maria Teresa Alvarez, Markus Ringnér, Mårten Eriksson, Martin Hedström, Martin O. Andersson, Mats Hansson, Mia Andersson, Mohammad Ibrahim, Nadia Osman, Nick Sirijovski, Nikos Tsardakas Renhuldt, Olle Holst, Olov Sterner, Osvaldo Delgado, Patrick Adlercreutz, Per Lundberg, Rajni Hatti-Kaul, Rodrigo Villagomez, Roula Nicolas Ibrahim, Ryan Oliver, Sang-Hyun Pyo, Siv Holmqvist, Sofia Essén, Sofia Mebrahtu-Wisén, Staffan Bensch, Susanne Högström, Suze Roostee, Tarek Dishisha, Torbjörn Säll, Ulrika Qvist Mathiese, Victor Ibrahim, Violeta Kaleskovska, Yusak Budi Susilo, Åsa Canbäck.

****

Abdelrahman Salah, Abdelrazek Mousa, Abir Nakhleh, Adel Abouhmad, Ahlam Mosteghanmi, Aleksandra Popovic, Allan Rasmussen, Ami Patel, André Holmberg, Arwa Abuawad, Ashwaq Aloraibi, Atena Malakpour Permlid, Aya

Fawzi, Beer Sen, Bindhu Sunil Karingannoor, Carl Jönsson, Christina Wennerberg, Christofer Ibrahim, Clara Braizat, Clara Dabour, Clara Dabour, Daniel Salas Veizaga, Deepti Sahoo, Diya Adawi, Elham Shakhtour, Emanuel Ron, Emma Wennerberg, Ewa Bukowska, Ewa Krupinska, Fayza Abed, Gabriella Ibrahim, Gaurav Dwivedi, Georgina Chavez Lizarraga, Henrik Wennerberg, Hong Chang, Humberto Itriago, Janin Schulte, Javier Linares Pastén, Jerry Jinx, Joel Wallenius, Johan Wennerberg, Johanna Volger, Jovita Janušauskaitė, Judith Matavacas, Kinga Zor, Konni Stiba, Kristin Aleklett, Kristin Kroll, Laila Shakhtour, Leonardo Cobuccio, Lesedi Lebogang, Lu Chen, Luciana Pereyra, Luis Alejandro Romero Soto, Mahmoud Sayed, Maitham Majeed, Manuel Orozco, Maria Eugenia Viloria-Cols, Marika Murto, Marisa Punzi, Markus Fröjd, Marlene Muñoz, Marwa Dabaieh, Marwaa Elhadary, Maryam Latifian, Mohamed Abdellah, Morad habaibeh, Mushtaq Ahmad, Nagwa Elnwishy, Najet Mahmoudi, Nariman Abu-Baker, Natali Sarsour, Neserin Ali, Nikolina Barchan, Oksana Zaushitsyna, Olli Andersson, Omar Khairat, Rabab Ouf, Ramin Sabet, Rana Al-Hammouz, Ravi Kiran Purama, Reza Faryar, Rita Shakhtour, Rosa Aragão, Roula Nicolas Ibrahim, Roxana Quiroga Flores, Roya Rezae, Ruth flores Jörnland, Sami Khan, Sammar Khalil, Serena Bisagni, Shaymaa Sayed, Shefaa Alasfoor, Shireen Mansour, Sindhu Matthew, Susmia Datta, Swati Aggarwal, Tim Börner, Tim Djärf, Tom Jasper, Tran Thi Thuy, Valeria Wallace, Veronica Ekdahl, Victor Ibrahim, Victor Kalbskopf, Vinardas Kelpšas, Zade Dirani, Zubaida Aldhafari

References

Aakre, C. D., Herrou, J., Phung, T. N., Perchuk, B. S., Crosson, S. & Laub, M. T. 2015. Evolving new protein-protein interaction specificity through promiscuous intermediates. *Cell,* 163**,** 594-606.

Abdel-Mohsein, H. S., Sasaki, T., Tada, C. & Nakai, Y. 2011. Characterization and partial purification of a bacteriocin-like substance produced by thermophilic Bacillus licheniformis H1 isolated from cow manure compost. *Anim Sci J,* 82**,** 340-51.

Abriouel, H., Franz, C. M., Ben Omar, N. & Galvez, A. 2011. Diversity and applications of Bacillus bacteriocins. *FEMS Microbiol Rev,* 35**,** 201-32.

Altschul, S. F., Madden, T. L., Schäffer, A. A., Zhang, J., Zhang, Z., Miller, W. & Lipman, D. J. 1997. Gapped BLAST and PSI-BLAST: a new generation of protein database search programs. *Nucleic Acids Research,* 25**,** 3389-3402.

Amitai, S., Yassin, Y. & Engelberg-Kulka, H. 2004. MazF-mediated cell death in Escherichia coli: a point of no return. *J Bacteriol,* 186**,** 8295-300.

Anantharaman, V. & Aravind, L. 2003. New connections in the prokaryotic toxin-antitoxin network: relationship with the eukaryotic nonsense-mediated RNA decay system. *Genome Biology,* 4**,** R81.

Anantharaman, V., Makarova, K. S., Burroughs, A. M., Koonin, E. V. & Aravind, L. 2013. Comprehensive analysis of the HEPN superfamily: identification of novel roles in intra-genomic conflicts, defense, pathogenesis and RNA processing. *Biology Direct,* 8**,** 15.

Andrade, Warrison a., Firon, A., Schmidt, T., Hornung, V., Fitzgerald, Katherine a., Kurt-Jones, Evelyn a., Trieu-Cuot, P., Golenbock, Douglas t. & Kaminski, P.-A. 2016. Group B Streptococcus Degrades Cyclic-di-AMP to Modulate STING-Dependent Type I Interferon Production. *Cell Host & Microbe,* 20**,** 49-59.

André, S., Zuber, F. & Remize, F. 2013. Thermophilic spore-forming bacteria isolated from spoiled canned food and their heat resistance. Results of a French ten-year survey. *Int J Food Microbiol,* 165**,** 134-43.

Aravind, L., Anantharaman, V., Balaji, S., Babu, M. M. & Iyer, L. M. 2005. The many faces of the helix-turn-helix domain: transcription regulation and beyond. *FEMS Microbiol Rev,* 29**,** 231-62.

Arias, A. A., Craig, M. & Fickers, P. 2011. *Gram-positive antibiotic biosynthetic clusters: a review,* Spain, Formatex Research Center.

Arnaouteli, S., Matoz-Fernandez, D. A., Porter, M., Kalamara, M., Abbott, J., Macphee, C. E., Davidson, F. A. & Stanley-Wall, N. R. 2019.

Pulcherrimin formation controls growth arrest of the Bacillus subtilis biofilm. *Proc Natl Acad Sci U S A,* 116**,** 13553-13562.

Arnison, P. G., Bibb, M. J., Bierbaum, G., Bowers, A. A., Bugni, T. S., Bulaj, G., Camarero, J. A., Campopiano, D. J., Challis, G. L., Clardy, J., Cotter, P. D., Craik, D. J., Dawson, M., Dittmann, E., Donadio, S., Dorrestein, P. C., Entian, K. D., Fischbach, M. A., Garavelli, J. S., Goransson, U., Gruber, C. W., Haft, D. H., Hemscheidt, T. K., Hertweck, C., Hill, C., Horswill, A. R., Jaspars, M., Kelly, W. L., Klinman, J. P., Kuipers, O. P., Link, A. J., Liu, W., Marahiel, M. A., Mitchell, D. A., Moll, G. N., Moore, B. S., Muller, R., Nair, S. K., Nes, I. F., Norris, G. E., Olivera, B. M., Onaka, H., Patchett, M. L., Piel, J., Reaney, M. J., Rebuffat, S., Ross, R. P., Sahl, H. G., Schmidt, E. W., Selsted, M. E., Severinov, K., Shen, B., Sivonen, K., Smith, L., Stein, T., Sussmuth, R. D., Tagg, J. R., Tang, G. L., Truman, A. W., Vederas, J. C., Walsh, C. T., Walton, J. D., Wenzel, S. C., Willey, J. M. & Van Der Donk, W. A. 2013. Ribosomally synthesized and post-translationally modified peptide natural products: overview and recommendations for a universal nomenclature. *Nat Prod Rep,* 30**,** 108-60.

Ash, C., Farrow, J. a. E., Wallbanks, S. & Collins, M. D. 1991. Phylogenetic heterogeneity of the genus Bacillus revealed by comparative analysis of small-subunit-ribosomal RNA sequences. *Letters in Applied Microbiology,* 13, **,** 202-206.

Awan, A. R., Blount, B. A., Bell, D. J., Shaw, W. M., Ho, J. C. H., Mckiernan, R. M. & Ellis, T. 2017. Biosynthesis of the antibiotic nonribosomal peptide penicillin in baker's yeast. *Nat Commun,* 8**,** 15202.

Baldwin, M. A. 2004. Protein Identification by Mass Spectrometry. *Issues to be Considered,* 3**,** 1-9.

Barbosa, J., Caetano, T. & Mendo, S. 2015. Class I and Class II Lanthipeptides Produced by Bacillus spp. *J Nat Prod,* 78**,** 2850-66.

Barrios, A. F. G., Zuo, R., Hashimoto, Y., Yang, L., Bentley, W. E. & Wood, T. K. 2006. Autoinducer 2 controls biofilm formation in Escherichia coli through a novel motility quorum-sensing regulator (MqsR, B3022). *J Bacteriol,* 188**,** 305-16.

Begley, M., Cotter, P. D., Hill, C. & Ross, R. P. 2009. Identification of a novel two-peptide lantibiotic, lichenicidin, following rational genome mining for LanM proteins. *Appl Environ Microbiol,* 75**,** 5451-60.

Bérdy, J. 2005. Bioactive Microbial Metabolites. *The Journal of Antibiotics,* 58**,** 1-26.

Bérdy, J. 2012. Thoughts and facts about antibiotics: Where we are now and where we are heading. *The Journal of Antibiotics,* 65**,** 385-395.

Bertrand, N., Fliss, I. & Lacroix, C. 2001. High nisin-Z production during repeated-cycle batch cultures in supplemented whey permeate using immobilized Lactococcus lactis UL719. *International Dairy Journal,* 11**,** 953-960.

Bhagya, C. H. W. M. R. C., Wijesundera Sulochana, W. S. & Hemamali, N. P. 2013. Polymerase chain reaction optimization for amplification of Guanine-Cytosine rich templates using buccal cell DNA. *Indian journal of human genetics,* 19**,** 78-83.

Bhunia, A. K., Johnson, M. C. & Ray, B. 1987. Direct Detection of an Antimicrobial Peptide of Pediococcus-Acidilactici in Sodium Dodecyl Sulfate-Polyacrylamide Gel-Electrophoresis. *Journal of Industrial Microbiology,* 2**,** 319-322.

Bigger, J. W. 1944. Treatment of staphylococcal infections with penicillin by intermittent sterilisation. *Lancet* 244**,** 497–500.

Bishop, B. M., Juba, M. L., Devine, M. C., Barksdale, S. M., Rodriguez, C. A., Chung, M. C., Russo, P. S., Vliet, K. A., Schnur, J. M. & Van Hoek, M. L. 2015. Bioprospecting the American Alligator (Alligator mississippiensis) Host Defense Peptidome. *Plos One,* 10.

Blin, K., Kazempour, D., Wohlleben, W. & Weber, T. 2014. Improved Lanthipeptide Detection and Prediction for antiSMASH. *PLoS ONE* 9.

Blin, K., Medema, M. H., Kottmann, R., Lee, S. Y. & Weber, T. 2017a. The antiSMASH database, a comprehensive database of microbial secondary metabolite biosynthetic gene clusters. *Nucleic Acids Res,* 45**,** D555-D559.

Blin, K., Pascal andreu, V., De los santos, E. L c., Del carratore, F., Lee, S. Y., Medema, M. H. & Weber, T. 2018. The antiSMASH database version 2: a comprehensive resource on secondary metabolite biosynthetic gene clusters. *Nucleic Acids Research,* 47**,** D625-D630.

Blin, K., Shaw, S., Steinke, K., Villebro, R., Ziemert, N., Lee, S. Y., Medema, M. H. & Weber, T. 2019. antiSMASH 5.0: updates to the secondary metabolite genome mining pipeline. *Nucleic Acids Research,* 47**,** W81-W87.

Blin, K., Wolf, T., Chevrette, M. G., Lu, X., Schwalen, C. J., Kautsar, S. A., Suarez Duran, H. G., De Los Santos, E. L. C., Kim, H. U., Nave, M., Dickschat, J. S., Mitchell, D. A., Shelest, E., Breitling, R., Takano, E., Lee, S. Y., Weber, T. & Medema, M. H. 2017b. antiSMASH 4.0-improvements in chemistry prediction and gene cluster boundary identification. *Nucleic Acids Res,* 45**,** W36-W41.

Bordenstein, S. R. & Reznikoff, W. S. 2005. Mobile DNA in obligate intracellular bacteria. *Nat Rev Microbiol,* 3**,** 688-99.

Borysowski, J. & Górski, A. 2009. Enzybiotics and their Potential Applications in Medicine. *Enzybiotics.* New Jersey: Wiley.

Brogden, K. A. 2005. Antimicrobial peptides: pore formers or metabolic inhibitors in bacteria? *Nat Rev Microbiol,* 3**,** 238-50.

Brötz, H., Bierbaum, G., Leopold, K., Reynolds, P. E. & Sahl, H.-G. 1998. The Lantibiotic Mersacidin Inhibits Peptidoglycan Synthesis by Targeting Lipid II. *Antimicrobial Agents and Chemotherapy,* 42**,** 154–160.

Brown, B. L., Grigoriu, S., Kim, Y., Arruda, J. M., Davenport, A., Wood, T. K., Peti, W. & Page, R. 2009. Three dimensional structure of the

MqsR:MqsA complex: a novel TA pair comprised of a toxin homologous to RelE and an antitoxin with unique properties. *PLoS Pathog,* 5**,** e1000706.
Bryanskaya, A. V., Rozanov, A. S., Slynko, N. M., Shekhovtsov, S. V. & Peltek, S. E. 2015. Geobacillus icigianus sp. nov., a thermophilic bacterium isolated from a hot spring. *International Journal of Systematic and Evolutionary Microbiology,* 65**,** 864-869.
Burgess, S. A., Lindsay, D. & Flint, S. H. 2010. Thermophilic bacilli and their importance in dairy processing. *Int J Food Microbiol,* 144**,** 215-25.
Bustard, M. T., Whiting, S., Cowan, D. A. & Wright, P. C. 2002. Biodegradation of high-concentration isopropanol by a solvent-tolerant thermophile, Bacillus pallidus. *Extremophiles,* 6**,** 319-23.
Castiglione, F., Lazzarini, A., Carrano, L., Corti, E., Ciciliato, I., Gastaldo, L., Candiani, P., Losi, D., Marinelli, F., Selva, E. & Parenti, F. 2008. Determining the structure and mode of action of microbisporicin, a potent lantibiotic active against multiresistant pathogens. *Chem Biol,* 15**,** 22-31.
Castro-Roa, D., Garcia-Pino, A., De Gieter, S., Van Nuland, N. a. J., Loris, R. & Zenkin, N. 2013. The Fic protein Doc uses an inverted substrate to phosphorylate and inactivate EF-Tu. *Nat Chem Biol,* 9**,** 811-7.
Chan, P. F., Macarron, R., Payne, D. J., Zalacain, M. & Holmes, D. J. 2002. Novel Antibacterials: A Genomics Approach to Drug Discovery. *Current Drug Targets - Infectious Disorders,* 2**,** 291-308.
Chan, W. T., Balsa, D. & Espinosa, M. 2015. One cannot rule them all: Are bacterial toxins-antitoxins druggable? *FEMS Microbiol Rev,* 39**,** 522-40.
Chan, W. T. & Espinosa, M. 2016. The Streptococcus pneumoniae pezAT Toxin-Antitoxin System Reduces beta-Lactam Resistance and Genetic Competence. *Front Microbiol,* 7**,** 1322.
Chan, W. T., Espinosa, M. & Yeo, C. C. 2016. Keeping the Wolves at Bay: Antitoxins of Prokaryotic Type II Toxin-Antitoxin Systems. *Front Mol Biosci,* 3**,** 9.
Chan, W. T., Moreno-Cordoba, I., Yeo, C. C. & Espinosa, M. 2012. Toxin-antitoxin genes of the Gram-positive pathogen Streptococcus pneumoniae: so few and yet so many. *Microbiol Mol Biol Rev,* 76**,** 773-91.
Chen, H. & Hoover, D. G. 2003. Bacteriocins and their Food Applications. *Comprehensive Reviews in Food Science and Food Safety,* 2**,** 82-100.
Cheverton, A. M., Gollan, B., Przydacz, M., Wong, C. T., Mylona, A., Hare, S. A. & Helaine, S. 2016. A Salmonella Toxin Promotes Persister Formation through Acetylation of tRNA. *Molecular cell,* 63**,** 86-96.
Chono, H., Matsumoto, K., Tsuda, H., Saito, N., Lee, K., Kim, S., Shibata, H., Ageyama, N., Terao, K., Yasutomi, Y., Mineno, J., Kim, S., Inouye, M. & Kato, I. 2011. Acquisition of HIV-1 resistance in T lymphocytes

using an ACA-specific E. coli mRNA interferase. *Hum Gene Ther,* 22**,** 35-43.
Chopin, M. C., Chopin, A. & Bidnenko, E. 2005. Phage abortive infection in lactococci: variations on a theme. *Curr Opin Microbiol,* 8**,** 473-9.
Chopra, N., Agarwal, S., Verma, S., Bhatnagar, S. & Bhatnagar, R. 2011. Modeling of the structure and interactions of the B. anthracis antitoxin, MoxX: deletion mutant studies highlight its modular structure and repressor function. *J Comput Aided Mol Des,* 25**,** 275-91.
Christensen-Dalsgaard, M. & Gerdes, K. 2006. Two higBA loci in the Vibrio cholerae superintegron encode mRNA cleaving enzymes and can stabilize plasmids. *Molecular Microbiology,* 62**,** 397-411.
Christensen, S. K., Mikkelsen, M., Pedersen, K. & Gerdes, K. 2001. RelE, a global inhibitor of translation, is activated during nutritional stress. *Proc Natl Acad Sci U S A,* 98**,** 14328-33.
Christensen, S. K., Pedersen, K., Hansen, F. G. & Gerdes, K. 2003. Toxin–antitoxin Loci as Stress-response-elements: ChpAK/MazF and ChpBK Cleave Translated RNAs and are Counteracted by tmRNA. *Journal of Molecular Biology,* 332**,** 809-819.
Coker, J. A. 2016. Extremophiles and biotechnology: current uses and prospects. *F1000Res,* 5.
Coles, M., Djuranovic, S., Soding, J., Frickey, T., Koretke, K., Truffault, V., Martin, J. & Lupas, A. N. 2005. AbrB-like transcription factors assume a swapped hairpin fold that is evolutionarily related to double-psi beta barrels. *Structure,* 13**,** 919-28.
Corrigan, R. M., Bowman, L., Willis, A. R., Kaever, V. & Grundling, A. 2015. Cross-talk between two nucleotide-signaling pathways in Staphylococcus aureus. *J Biol Chem,* 290**,** 5826-39.
Corvey, C., Stein, T., Düsterhus, S., Karas, M. & Entian, K.-D. 2003. Activation of subtilin precursors by Bacillus subtilis extracellular serine proteases subtilisin (AprE), WprA, and Vpr. *Biochemical and Biophysical Research Communications,* 304**,** 48-54.
Cotter, P. D., Hill, C. & Ross, R. P. 2005a. Bacterial Lantibiotics: Strategies to Improve Therapeutic Potential. *Current Protein and Peptide Science,* 6**,** 61-75.
Cotter, P. D., Hill, C. & Ross, R. P. 2005b. Bacteriocins: developing innate immunity for food. *Nature Reviews Microbiology,* 3**,** 777-788.
Cottrell, J. S. 2011. Protein identification using MS/MS data. *J Proteomics,* 74**,** 1842-51.
Coussens, N. P. & Daines, D. A. 2016. Wake me when it's over - Bacterial toxin-antitoxin proteins and induced dormancy. *Exp Biol Med (Maywood),* 241**,** 1332-42.
Cox, C. L., Tietz, J. I., Sokolowski, K., Melby, J. O., Doroghazi, J. R. & Mitchell, D. A. 2014. Nucleophilic 1,4-additions for natural product discovery. *ACS Chem Biol,* 9**,** 2014-22.

Cruz, J. W., Rothenbacher, F. P., Maehigashi, T., Lane, W. S., Dunham, C. M. & Woychik, N. A. 2014. Doc toxin is a kinase that inactivates elongation factor Tu. *J Biol Chem,* 289**,** 7788-98.

D'amatot, R. F. & Thornsberry, C. 1979. Calcium and Magnesium in Mueller-Hinton Agar and Their Influence on Disk Diffusion Susceptibility Results. *Current Microbiology,* 2**,** 135-138.

Daly, K. M., Cotter, P. D., Hill, C. & Ross, R. P. 2012. Lantibiotic Production by Pathogenic Microorganisms. *Current Protein and Peptide Science,* 13**,** 509-523.

Daly, K. M., Upton, M., Sandiford, S. K., Draper, L. A., Wescombe, P. A., Jack, R. W., O'connor, P. M., Rossney, A., Gotz, F., Hill, C., Cotter, P. D., Ross, R. P. & Tagg, J. R. 2010. Production of the Bsa lantibiotic by community-acquired Staphylococcus aureus strains. *J Bacteriol,* 192**,** 1131-42.

Danesh, A., Mamo, G. & Mattiasson, B. 2011. Production of haloduracin by Bacillus halodurans using solid-state fermentation. *Biotechnology Letters,* 33**,** 1339-1344.

Dang, J., Mendez, P., Lee, S., Kim, J. W., Yoon, J.-H., Kim, T. W., Sailey, C. J., Jablons, D. M. & Kim, I.-J. 2016. Development of a robust DNA quality and quantity assessment qPCR assay for targeted next-generation sequencing library preparation. *International journal of oncology,* 49**,** 1755-1765.

De Bast, M. S., Mine, N. & Van Melderen, L. 2008. Chromosomal toxin-antitoxin systems may act as antiaddiction modules. *J Bacteriol,* 190**,** 4603-9.

De Jong, A., Van Heel, A. J. & Kuipers, O. P. 2011. Genome Exploitation and Bioinformatics Tools. *In:* Drider, D. & Rebuffat, S. (eds.) *Prokaryotic Antimicrobial Peptides: From Genes to Applications.* New York, USA: Springer.

De Jong, A., Van Hijum, S. A., Bijlsma, J. J., Kok, J. & Kuipers, O. P. 2006. BAGEL: a web-based bacteriocin genome mining tool. *Nucleic Acids Res,* 34**,** W273-9.

Deghorain, M., Goeders, N., Jové, T. & Melderen, L. V. 2013. Type II Toxin-Antitoxin Loci: The ccdAB and parDE Families. *In:* Gerdes, K. (ed.) *Prokaryotic Toxin-Antitoxins.* Newcastle upon Tyne, UK: Springer.

Des Marais, D. J. & Walter, M. R. 2019. Terrestrial Hot Spring Systems: Introduction. *Astrobiology*.

Diaz-Orejas, R., Espinosa, M. & Yeo, C. C. 2017. The Importance of the Expendable: Toxin-Antitoxin Genes in Plasmids and Chromosomes. *Front Microbiol,* 8**,** 1479.

Doleyres, Y. & Lacroix, C. 2005. Technologies with free and immobilised cells for probiotic bifidobacteria production and protection. *International Dairy Journal,* 15**,** 973-988.

Drissi, F., Buffet, S., Raoult, D. & Merhej, V. 2015. Common occurrence of antibacterial agents in human intestinal microbiota. *Front Microbiol,* 6**,** 441.

Dziuba, B. & Dziuba, M. 2014. New milk protein-derived peptides with potential antimicrobial activity: an approach based on bioinformatic studies. *Int J Mol Sci,* 15**,** 14531-45.

Edwards, N. J. 2011. Protein Identification from Tandem Mass Spectra by Database Searching. *In:* Wu C., C. C. (ed.) *Bioinformatics for Comparative Proteomics. Methods in Molecular Biology (Methods and Protocols).* Humana Press.

Egan, K., Field, D., Ross, R. P., Cotter, P. D. & Hill, C. 2018. In silico Prediction and Exploration of Potential Bacteriocin Gene Clusters Within the Bacterial Genus Geobacillus. *Front Microbiol,* 9**,** 2116.

Eguchi, T., Kaminaka, K., Shima, J., Kawamoto, S., Mori, K., Choi, S. H., Doi, K., Ohmomo, S. & Ogata, S. 2014. Isolation and characterization of enterocin SE-K4 produced by thermophilic enterococci, Enterococcus faecalis K-4. *Biosci Biotechnol Biochem,* 65**,** 247-53.

Ellen, A. F., Rohulya, O. V., Fusetti, F., Wagner, M., Albers, S.-V. & Driessen, A. J. M. 2011. The Sulfolobicin Genes of *Sulfolobus acidocaldarius* Encode Novel Antimicrobial Proteins. *Journal of Bacteriology,* 193**,** 4380-4387.

Engelberg-Kulka, H. & Glaser, G. 1999. Addiction modules and programmed Cell death and antideath in bacterial cultures. *Annu. Rev. Microbiol.* , 53**,** 43–70.

Epand, R. M., Walker, C., Epand, R. F. & Magarvey, N. A. 2016. Molecular mechanisms of membrane targeting antibiotics. *Biochim Biophys Acta,* 1858**,** 980-7.

Eraifej, N. 2006. Gas Geochemistry and Isotopic Signatures in the deep Thermal waters in Jordan. *Freiberg Online Geology,* 16.

Esikova, T. Z., Temirov, Y. V., Sokolov, S. L. & Alakhov, Y. B. 2002. Secondary antimicrobial metabolites produced by thermophilic Bacillus spp. strains VK2 and VK21. *Applied Biochemistry and Microbiology,* 38**,** 226-231.

Essig, A., Hofmann, D., Munch, D., Gayathri, S., Kunzler, M., Kallio, P. T., Sahl, H. G., Wider, G., Schneider, T. & Aebi, M. 2014. Copsin, a novel peptide-based fungal antibiotic interfering with the peptidoglycan synthesis. *J Biol Chem,* 289**,** 34953-64.

Feng, L., Wang, W., Cheng, J., Ren, Y., Zhao, G., Gao, C., Tang, Y., Liu, X., Han, W., Peng, X., Liu, R. & Wang, L. 2007. Genome and proteome of long-chain alkane degrading Geobacillus thermodenitrificans NG80-2 isolated from a deep-subsurface oil reservoir. *Proceedings of the National Academy of Sciences of the United States of America,* 104**,** 5602-5607.

Fiebig, A., Castro Rojas, C. M., Siegal-Gaskins, D. & Crosson, S. 2010. Interaction specificity, toxicity and regulation of a paralogous set of ParE/RelE-family toxin-antitoxin systems. *Mol Microbiol,* 77**,** 236-51.
Fu, Z., Tamber, S., Memmi, G., Donegan, N. P. & Cheung, A. L. 2009. Overexpression of MazFsa in Staphylococcus aureus induces bacteriostasis by selectively targeting mRNAs for cleavage. *J Bacteriol,* 191**,** 2051-9.
Gajiwala, K. S. & Burley, S. K. 2000. Winged helix proteins. *Current Opinion in Structural Biology,* 10**,** 110-116.
Galtier, N., Tourasse, N. & Gouy, M. 1999. A Nonhyperthermophilic Common Ancestor to Extant Life Forms. *Science,* 283**,** 220-221.
Garcia-Pino, A., Sterckx, Y., Magnuson, R. D. & Loris, R. 2013. Type II Toxin-Antitoxin Loci: The phd/doc Family. *In:* Gerdes, K. (ed.) *Prokaryotic Toxin-Antitoxins.* Newcastle upon Tyne, UK: Springer.
Garg, N., Oman, T. J., Andrew Wang, T. S., De Gonzalo, C. V., Walker, S. & Van Der Donk, W. A. 2014. Mode of action and structure-activity relationship studies of geobacillin I. *J Antibiot (Tokyo),* 67**,** 133-6.
Garg, N., Tang, W., Goto, Y., Nair, S. K. & Van Der Donk, W. A. 2012. Lantibiotics from Geobacillus thermodenitrificans. *Proc Natl Acad Sci U S A,* 109**,** 5241-6.
Gerdes, K. 2000. Toxin-Antitoxin Modules May Regulate Synthesis of Macromolecules during Nutritional Stress. 182**,** 561–572.
Gerdes, K. 2013. Type II Toxin-Antitoxin Loci: The relBE Family. *In:* Gerdes, K. (ed.) *Prokaryotic Toxin-Antitoxins.* Newcastle upon Tyne, UK: Springer.
Gerdes, K., Christensen, S. K. & Lobner-Olesen, A. 2005. Prokaryotic toxin-antitoxin stress response loci. *Nat Rev Microbiol,* 3**,** 371-82.
Gerdes, K., Rasmussen, P. B. & Molin, S. 1986. Unique type of plasmid maintenance function: Postsegregational killing of plasmid-free cells. *Proc. Natl. Acad. Sci. USA,* 83**,** 3116-3120.
Gharsallaoui, A., Oulahal, N., Joly, C. & Degraeve, P. 2016. Nisin as a Food Preservative: Part 1: Physicochemical Properties, Antimicrobial Activity, and Main Uses. *Crit Rev Food Sci Nutr,* 56**,** 1262-74.
Ghuysen, J.-M. & Hakenbeck, R. 1994. *Bacterial Cell Wall,* Elsevier Science.
Giudice, A. L. & Renato, F. 2016. Antimicrobial Potential of Cold-Adapted Bacteria and Fungi from Polar Regions. *In:* Rampelotto, P. H. (ed.) *Biotechnology of Extremophiles.* Brazil: Springer.
Goormaghtigh, F., Fraikin, N., Putrinš, M., Hallaert, T., Hauryliuk, V., Garcia-Pino, A., Sjödin, A., Kasvandik, S., Udekwu, K., Tenson, T., Kaldalu, N. & Van Melderen, L. 2018. Reassessing the Role of Type II Toxin-Antitoxin Systems in Formation of *Escherichia coli* Type II Persister Cells. *mBio,* 9**,** e00640-18.
Gotfredsen, M. & Gerdes, K. 1998. The Escherichia coli relBE genes belong to a new toxin-antitoxin gene family. *Molecular Microbiology,* 29**,** 1065-1076.

Gotoh, H., Kasaraneni, N., Devineni, N., Dallo, S. F. & Weitao, T. 2010. SOS involvement in stress-inducible biofilm formation. *Biofouling,* 26**,** 603-611.
Gross, M., Marianovsky, I. & Glaser, G. 2006. MazG -- a regulator of programmed cell death in Escherichia coli. *Mol Microbiol,* 59**,** 590-601.
Guglielmini, J. & Van Melderen, L. 2011. Bacterial toxin-antitoxin systems: Translation inhibitors everywhere. *Mob Genet Elements,* 1**,** 283-290.
Gundlach, J., Mehne, F. M., Herzberg, C., Kampf, J., Valerius, O., Kaever, V. & Stulke, J. 2015. An Essential Poison: Synthesis and Degradation of Cyclic Di-AMP in Bacillus subtilis. *J Bacteriol,* 197**,** 3265-74.
Hallez, R., Geeraerts, D., Sterckx, Y., Mine, N., Loris, R. & Van Melderen, L. 2010. New toxins homologous to ParE belonging to three-component toxin-antitoxin systems in Escherichia coli O157:H7. *Mol Microbiol,* 76**,** 719-32.
Hampton, H. G., Jackson, S. A., Fagerlund, R. D., Vogel, A. I. M., Dy, R. L., Blower, T. R. & Fineran, P. C. 2018. AbiEi Binds Cooperatively to the Type IV abiE Toxin-Antitoxin Operator Via a Positively-Charged Surface and Causes DNA Bending and Negative Autoregulation. *J Mol Biol,* 430**,** 1141-1156.
Harvey, R. M., Stroeher, U. H., Ogunniyi, A. D., Smith-Vaughan, H. C., J., L. A. & Paton, J. C. 2011. A Variable Region within the Genome of Streptococcus pneumoniae Contributes to Strain-Strain Variation in Virulence. *PLoS ONE* 6.
Hauryliuk, V., Atkinson, G. C., Murakami, K. S., Tenson, T. & Gerdes, K. 2015. Recent functional insights into the role of (p)ppGpp in bacterial physiology. *Nat Rev Microbiol,* 13**,** 298-309.
Hayek, S. A., Gyawali, R. & Ibrahim, S. A. 2013. *In:* Méndez-Vilas, A. (ed.) *Microbial Pathogens and Strategies for Combating Them: Science, Technology and Education.* Formatex Research Center.
Hayes, F. 2003. Toxins-Antitoxins: Plasmid Maintenance, Programmed Cell Death, and Cell Cycle Arrest. *Science,* 301**,** 1496-1499.
Hayes, F. & Kędzierska, B. 2014. Regulating toxin-antitoxin expression: controlled detonation of intracellular molecular timebombs. *Toxins (Basel),* 6**,** 337-58.
Hayes, F. & Van Melderen, L. 2011. Toxins-antitoxins: diversity, evolution and function. *Crit Rev Biochem Mol Biol,* 46**,** 386-408.
Hazan, R. & Engelberg-Kulka, H. 2004. Escherichia coli mazEF-mediated cell death as a defense mechanism that inhibits the spread of phage P1. *Mol Genet Genomics,* 272**,** 227-34.
Hibbing, M. E., Fuqua, C., Parsek, M. R. & Peterson, S. B. 2010. Bacterial competition: surviving and thriving in the microbial jungle. *Nat Rev Microbiol,* 8**,** 15-25.
Hudzicki, J. 2009. Kirby-Bauer Disk Diffusion Susceptibility Test Protocol.
Hussein, A. H., Lisowska, B. K. & Leak, D. J. 2015. Chapter One - The Genus Geobacillus and Their Biotechnological Potential. In: Sariaslani, S. &

Gadd, G. M. (eds.) Advances in Applied Microbiology. Academic Press.
Huynh, T. N., Luo, S., Pensinger, D., Sauer, J. D., Tong, L. & Woodward, J. J. 2015. An HD-domain phosphodiesterase mediates cooperative hydrolysis of c-di-AMP to affect bacterial growth and virulence. *Proc Natl Acad Sci U S A,* 112**,** E747-56.
Huynh, T. N. & Woodward, J. J. 2016. Too much of a good thing: regulated depletion of c-di-AMP in the bacterial cytoplasm. *Curr Opin Microbiol,* 30**,** 22-29.
Irving, S. E. & Corrigan, R. M. 2018. Triggering the stringent response: signals responsible for activating (p)ppGpp synthesis in bacteria. *Microbiology,* 164**,** 268-276.
Ji, Y. 2002. The role of genomics in the discovery of novel targets for antibiotic therapy. *Pharmacogenomics,* 3**,** 315-323.
Jia, X., Yao, J., Gao, Z., Liu, G., Dong, Y. H., Wang, X. & Zhang, H. 2018. Structure-function analyses reveal the molecular architecture and neutralization mechanism of a bacterial HEPN-MNT toxin-antitoxin system. *J Biol Chem,* 293**,** 6812-6823.
Jiang, Y., Pogliano, J., Helinski, D. R. & Konieczny, I. 2002. ParE toxin encoded by the broad-host-range plasmid RK2 is an inhibitor of *Escherichia coli* gyrase. *Molecular Microbiology,* 44**,** 971-979.
Jones, P., Binns, D., Chang, H. Y., Fraser, M., Li, W., Mcanulla, C., Mcwilliam, H., Maslen, J., Mitchell, A., Nuka, G., Pesseat, S., Quinn, A. F., Sangrador-Vegas, A., Scheremetjew, M., Yong, S. Y., Lopez, R. & Hunter, S. 2014. InterProScan 5: genome-scale protein function classification. *Bioinformatics,* 30**,** 1236-40.
Jørgensen, M. G., Pandey, D. P., Jaskolska, M. & Gerdes, K. 2009. HicA of Escherichia coli defines a novel family of translation-independent mRNA interferases in bacteria and archaea. *J Bacteriol,* 191**,** 1191-9.
Jurėnas, D., Garcia-Pino, A. & Van Melderen, L. 2017. Novel toxins from type II toxin-antitoxin systems with acetyltransferase activity. *Plasmid,* 93**,** 30-35.
Kabuki, T., Uenishi, H., Watanabe, M., Seto, Y. & Nakajima, H. 2007. Characterization of a bacteriocin, Thermophilin 1277, produced by Streptococcus thermophilus SBT1277. *J Appl Microbiol,* 102**,** 971-80.
Kanehisa, M. & Goto, S. 2000. KEGG: Kyoto Encyclopedia of Genes and Genomes. *Nucleic Acids Research,* 28**,** 27-30.
Kanehisa, M., Sato, Y., Furumichi, M., Morishima, K. & Tanabe, M. 2018. New approach for understanding genome variations in KEGG. *Nucleic Acids Research,* 47**,** D590-D595.
Kang, S. M., Kim, D. H., Jin, C. & Lee, B. J. 2018. A Systematic Overview of Type II and III Toxin-Antitoxin Systems with a Focus on Druggability. *Toxins (Basel),* 10.

Kaunietis, A., Pranckute, R., Lastauskiene, E. & Citavicius, D. J. 2017. Medium Optimization for Bacteriocin Production and Bacterial Cell Growth of Geobacillus sp. 15 Strain. *Journal of Antimicrobial Agents,* 03.

Kędzierska, B. & Hayes, F. 2016. Emerging Roles of Toxin-Antitoxin Modules in Bacterial Pathogenesis. *Molecules,* 21.

Kelkar, D. A. & Chattopadhyay, A. 2007. The gramicidin ion channel: a model membrane protein. *Biochim Biophys Acta,* 1768**,** 2011-25.

Keren, I., Shah, D., Spoering, A., Kaldalu, N. & Lewis, K. 2004. Specialized persister cells and the mechanism of multidrug tolerance in Escherichia coli. *J Bacteriol,* 186**,** 8172-80.

Kim, Y., Wang, X., Ma, Q., Zhang, X. S. & Wood, T. K. 2009. Toxin-antitoxin systems in Escherichia coli influence biofilm formation through YjgK (TabA) and fimbriae. *J Bacteriol,* 191**,** 1258-67.

Kimura, H., Sugihara, M., Kato, K. & Hanada, S. 2006. Selective phylogenetic analysis targeted at 16S rRNA genes of thermophiles and hyperthermophiles in deep-subsurface geothermal environments. *Applied and environmental microbiology,* 72**,** 21-27.

Kingston, D. G. I. 2011. Modern natural products drug discovery and its relevance to biodiversity conservation. *Journal of natural products,* 74**,** 496-511.

Knerr, P. J. & Van Der Donk, W. A. 2012. Discovery, biosynthesis, and engineering of lantipeptides. *Annu Rev Biochem,* 81**,** 479-505.

Kolodkin-Gal, I. & Engelberg-Kulka, H. 2006. Induction of Escherichia coli chromosomal mazEF by stressful conditions causes an irreversible loss of viability. *J. Bacteriol.,* 188**,** 3420-3.

Kuchta, K., Knizewski, L., Wyrwicz, L. S., Rychlewski, L. & Ginalski, K. 2009. Comprehensive classification of nucleotidyltransferase fold proteins: identification of novel families and their representatives in human. *Nucleic Acids Res.,* 37**,** 7701-14.

Kumar, S. & Engelberg-Kulka, H. 2014. Quorum sensing peptides mediating interspecies bacterial cell death as a novel class of antimicrobial agents. *Curr Opin Microbiol,* 21**,** 22-7.

Labbate, M., Case, R. J. & Stokes, H. W. 2009. The integron/gene cassette system: an active player in bacterial adaptation. *In:* Gogarten, M. B., Gogarten, J. P. & Olendzenski, L. (eds.) *Horizontal Gene Transfer. Methods in Molecular Biology.* New York, USA: Humana Press.

Laich, F., Fierro, F. & Martin, J. F. 2002. Production of penicillin by fungi growing on food products: identification of a complete penicillin gene cluster in Penicillium griseofulvum and a truncated cluster in Penicillium verrucosum. *Appl Environ Microbiol,* 68**,** 1211-9.

Leplae, R., Geeraerts, D., Hallez, R., Guglielmini, J., Dreze, P. & Van Melderen, L. 2011. Diversity of bacterial type II toxin-antitoxin systems: a comprehensive search and functional analysis of novel families. *Nucleic Acids Res,* 39**,** 5513-25.

Li, X. & O'sullivan, D. J. 2012. Contribution of the Actinobacteria to the growing diversity of lantibiotics. *Biotechnology Letters,* 34**,** 2133-2145.
Liu, M., Zhang, Y., Inouye, M. & Woychik, N. A. 2008. Bacterial addiction module toxin Doc inhibits translation elongation through its association with the 30S ribosomal subunit. *Proceedings of the National Academy of Sciences,* 105**,** 5885-5890.
Magnuson, R. D. 2007. Hypothetical functions of toxin-antitoxin systems. *J Bacteriol,* 189**,** 6089-92.
Makarova, K. S., Wolf, Y. I. & Koonin, E. V. 2009. Comprehensive comparative-genomic analysis of type 2 toxin-antitoxin systems and related mobile stress response systems in prokaryotes. *Biol. Direct,* 4**,** 19.
Makarova, K. S., Wolf, Y. I. & Koonin, E. V. 2013. Comparative genomics of defense systems in archaea and bacteria. *Nucleic Acids Res,* 41**,** 4360-77.
Manning, M. & Colon, W. 2004. Structural basis of protein kinetic stability: Resistance to sodium dodecyl sulfate suggests a central role for rigidity and a bias toward beta-sheet structure. *Biochemistry,* 43**,** 11248-11254.
Marchler-Bauer, A., Bo, Y., Han, L., He, J., Lanczycki, C. J., Lu, S., Chitsaz, F., Derbyshire, M. K., Geer, R. C., Gonzales, N. R., Gwadz, M., Hurwitz, D. I., Lu, F., Marchler, G. H., Song, J. S., Thanki, N., Wang, Z., Yamashita, R. A., Zhang, D., Zheng, C., Geer, L. Y. & Bryant, S. H. 2017. CDD/SPARCLE: functional classification of proteins via subfamily domain architectures. *Nucleic Acids Res,* 45**,** D200-D203.
Marinelli, F. & Genilloud, O. 2014. *Antimicrobials: New and Old Molecules in the Fight Against Multi-Resistant Bacteria,* Heidelberg, Springer
Marsh, A. J., O'sullivan, O., Ross, R. P., Cotter, P. D. & Hill, C. 2010. In silico analysis highlights the frequency and diversity of type 1 lantibiotic gene clusters in genome sequenced bacteria. *BMC Genomics,* 11**,** 679.
Martirani, L., Varcamonti, M., Naclerio, G. & De Felice, M. 2002. Purification and partial characterization of bacillocin 490, a novel bacteriocin produced by a thermophilic strain of Bacillus licheniformis. *Microbial cell factories,* 1**,** 1-1.
Matelska, D., Steczkiewicz, K. & Ginalski, K. 2017. Comprehensive classification of the PIN domain-like superfamily. *Nucleic Acids Res,* 45**,** 6995-7020.
Mattison, K., Wilbur, J. S., So, M. & Brennan, R. G. 2006. Structure of FitAB from Neisseria gonorrhoeae bound to DNA reveals a tetramer of toxin-antitoxin heterodimers containing pin domains and ribbon-helix-helix motifs. *J Biol Chem,* 281**,** 37942-51.
Maugeri, T. L., Gugliandolo, C., Caccamo, D. & Stackebrandt, E. 2001. A polyphasic taxonomic study of thermophilic bacilli from shallow, marine vents. *Syst Appl Microbiol,* 24**,** 572-87.
Mc Dermott, W. 1958. Microbial Persistence. *Yale Journal of Biology and Medicine* 30.

Mcauliffe, O., Ross, R. P. & Hill, C. 2001. Lantibiotics: structure, biosynthesis and mode of action. *FEMS Microbiology Reviews,* 25**,** 285-308.
Mcclerren, A. L., Cooper, L. E., Quan, C., Thomas, P. M., Kelleher, N. L. & Van Der Donk, W. A. 2006. Discovery and *in vitro* biosynthesis of haloduracin, a two-component lantibiotic. *Proceedings of the National Academy of Sciences,* 103**,** 17243-17248.
Mcdonnell, G. E. & Mcconnell, D. J. 1994. Overproduction, isolation, and DNA-binding characteristics of Xre, the repressor protein from the Bacillus subtilis defective prophage PBSX. *Journal of Bacteriology,* 176**,** 5831-5834.
Mcdonnell, G. E., Wood, H., Devine, K. M. & Mcconnell, D. J. 1994. Genetic Control of Bacterial Suicide: Regulation of the Induction of PBSX in *Bacillus subtilis*. *Journal of Bacteriology* 176**,** 5820-5830.
Mckenzie, J. L., Robson, J., Berney, M., Smith, T. C., Ruthe, A., Gardner, P. P., Arcus, V. L. & Cook, G. M. 2012. A VapBC toxin-antitoxin module is a posttranscriptional regulator of metabolic flux in mycobacteria. *J Bacteriol,* 194**,** 2189-204.
Mcmullan, G., Christie, J. M., Rahman, T. J., Banat, I. M., Ternan, N. G. & Marchant, R. 2004. Habitat, applications and genomics of the aerobic, thermophilic genus Geobacillus. *Biochemical Society Transactions* 32**,** 214-217.
Medema, M. H., Blin, K., Cimermancic, P., De Jager, V., Zakrzewski, P., Fischbach, M. A., Weber, T., Takano, E. & Breitling, R. 2011. antiSMASH: rapid identification, annotation and analysis of secondary metabolite biosynthesis gene clusters in bacterial and fungal genome sequences. *Nucleic Acids Res,* 39**,** W339-46.
Merck 2009. Mueller-Hinton agar CLSI-Merck Microbiology Manual. 12th ed.
Metzger, S., Dror, I. B., Aizenman, E., Schreiber, G., Toone, M., Friesen, J. D., Cashel, M. & Glaserll, G. 1988. The Nucleotide Sequence and Characterizationof the *relA* Gene of *Escherichia coli*. *The Journal of Biological Chemistry* 263**,** 15699-15704.
Mine, N., Guglielmini, J., Wilbaux, M. & Van Melderen, L. 2009. The decay of the chromosomally encoded ccdO157 toxin-antitoxin system in the Escherichia coli species. *Genetics,* 181**,** 1557-66.
Miyamoto, T., Kato, Y., Sekiguchi, Y., Tsuneda, S. & Noda, N. 2016. Characterization of MazF-Mediated Sequence-Specific RNA Cleavage in Pseudomonas putida Using Massive Parallel Sequencing. *PLoS One,* 11**,** e0149494.
Molloy, E. M., Ross, R. P. & Hill, C. 2012. 'Bac' to the future: bioengineering lantibiotics for designer purposes. *Biochem Soc Trans,* 40**,** 1492-7.
Moloney, M. G. 2016. Natural Products as a Source for Novel Antibiotics. *Trends Pharmacol Sci,* 37**,** 689-701.
Morton, J. T., Freed, S. D., Lee, S. W. & Friedberg, I. 2015a. A large scale prediction of bacteriocin gene blocks suggests a wide functional spectrum for bacteriocins. *BMC bioinformatics,* 16**,** 381-381.

Morton, J. T., Freed, S. D., Lee, S. W. & Friedberg, I. 2015b. A large scale prediction of bacteriocin gene blocks suggests a wide functional spectrum for bacteriocins. *BMC Bioinformatics,* 16**,** 381.

Muhd Sakaff, M. K., Abdul Rahman, A. Y., Saito, J. A., Hou, S. & Alam, M. 2012. Complete genome sequence of the thermophilic bacterium *Geobacillus thermoleovorans* CCB_US3_UF5. *J. Bacteriol.,* 194**,** 1239.

Muthuramalingam, M., White, J. C., Murphy, T., Ames, J. R. & Bourne, C. R. 2019. The toxin from a ParDE toxin-antitoxin system found in *Pseudomonas aeruginosa* offers protection to cells challenged with anti-gyrase antibiotics. *Mol. Microbiol.,* 111**,** 441-454.

Naghmouchi, K., Fliss, I., Drider, D. & Lacroix, C. 2008. Pediocin PA-1 production during repeated-cycle batch culture of immobilized *Pediococcus acidilactici* UL5 cells. *Journal of Bioscience and Bioengineering,* 105**,** 513-517.

Nazina, T. N., Tourova, T. P., Poltaraus, A. B., Novikova, E. V., Grigoryan, A. A., Ivanova, A. E., Lysenko, A. M., Petrunyaka, V. V., Osipov, G. A., Belyaev, S. S. & Ivanov, M. V. 2001. Taxonomic study of aerobic thermophilic bacilli: descriptions of *Geobacillus subterraneus* gen. nov., sp. nov. and *Geobacillus uzenensis* sp. nov. from petroleum reservoirs and transfer of *Bacillus stearothermophilus, Bacillus thermocatenulatus, Bacillus thermoleovorans, Bacillus kaustophilus, Bacillus thermoglucosidasius* and *Bacillus thermodenitrificans* to *Geobacillus* as the new combinations *G. stearothermophilus, G. thermocatenulatus, G. thermoleovorans, G. kaustophilus, G. thermoglucosidasius and G. thermodenitrificans*. *International Journal of Systematic and Evolutionary Microbiology,* 51**,** 433–446.

Nilsen, T., Nes, I. F. & Holo, H. 2003. Enterolysin A, a cell wall-degrading bacteriocin from Enterococcus faecalis LMG 2333. *Applied and Environmental Microbiology,* 69**,** 2975-2984.

Ning, D., Liu, S., Xu, W., Zhuang, Q., Wen, C. & Tang, X. 2013. Transcriptional and proteolytic regulation of the toxin-antitoxin locus vapBC10 (ssr2962/slr1767) on the chromosome of Synechocystis sp. PCC 6803. *PLoS One,* 8**,** e80716.

Novichkov, P. S., Kazakov, A. E., Ravcheev, D. A., Leyn, S. A., Kovaleva, G. Y., Sutormin, R. A., Kazanov, M. D., Riehl, W., Arkin, A. P., Dubchak, I. & Rodionov, D. A. 2013. RegPrecise 3.0 – A resource for genome-scale exploration of transcriptional regulation in bacteria. *BMC Genomics,* 14**,** 745.

Novotny, J. F. & Perry, J. J. 1992. Characterization of bacteriocins from two strains of Bacillus thermoleovorans, a thermophilic hydrocarbon-utilizing species. *Applied and Environmental Microbiology,* 58**,** 2393-2396.

Obojska, A., Ternan, N. G., Lejczak, B., Kafarski, P. & Mcmullan, G. 2002. Organophosphonate Utilization by the Thermophile *Geobacillus*

caldoxylosilyticus T20. *Applied and Environmental Microbiology,* 68**,** 2081-2084.
Ogura, T. & Hiraga, S. 1983. Mini-F plasmid genes that couple host cell division to plasmid proliferation. *Proc. Natl. Acad. Sci.,* 80**,** 4784-4788.
Olano, C., Méndez, C. & Salas, J. A. 2014. Strategies for the Design and Discovery of Novel Antibiotics using Genetic Engineering and Genome Mining. *In:* Villa, T. G. & Veiga-Crespo, P. (eds.) *Antimicrobial Compounds: Current Strategies and New Alternatives.* Heidelberg: Springer.
Ou, H.-Y., Wei, Y. & Bi, D. 2013. Type II Toxin-Antitoxin Loci: Phylogeny. *In:* Gerdes, K. (ed.) *Prokaryotic Toxin-Antitoxins.* Newcastle upon Tyne, UK: Springer.
Overgaard, M., Borch, J. & Gerdes, K. 2009. RelB and RelE of Escherichia coli form a tight complex that represses transcription via the ribbon-helix-helix motif in RelB. *J Mol Biol,* 394**,** 183-96.

Özdemir, G. B. & Biyik, H. H. 2012a. Isolation and Characterization of a Bacteriocin-Like Substance Produced by *Geobacillus toebii* Strain HBB-247. *Indian journal of microbiology,* 52**,** 104-108.
Özdemir, G. B. & Biyik, H. H. 2012b. Isolation and characterization of toebicin 218, a bacteriocin, produced by *Geobacillus toebii* HBB-218. *African Journal of Biotechnology,* 11.
Page, R. & Peti, W. 2016. Toxin-antitoxin systems in bacterial growth arrest and persistence. *Nature Chemical Biology,* 12**,** 208.
Pandey, D. P. & Gerdes, K. 2005. Toxin-antitoxin loci are highly abundant in free-living but lost from host-associated prokaryotes. *Nucleic Acids Res,* 33**,** 966-76.
Parisien, A., Allain, B., Zhang, J., Mandeville, R. & Lan, C. Q. 2008. Novel alternatives to antibiotics: bacteriophages, bacterial cell wall hydrolases, and antimicrobial peptides. *Journal of Applied Microbiology,* 104**,** 1-13.
Pedersen, K., Christensen, S. K. & Gerdes, K. 2002. Rapid induction and reversal of a bacteriostatic condition by controlled expression of toxins and antitoxins. *Mol. Microbiol.,* 45**,** 501-510.
Pedersen, K., Zavialov, A. V., Pavlov, M. Y., Elf, J., Gerdes, K. & Ehrenberg, M. 2003. The Bacterial Toxin RelE Displays Codon-Specific Cleavage of mRNAs in the Ribosomal A Site. *Cell,* 112**,** 131-140.
Pednekar, P., Jain, R. & Mahajan, G. 2011. Anti-infective potential of hot-spring bacteria. . *J Global Infect Dis,* 3**,** 241-5.
Pilkington, P. H. 2005. Food Bioconversions and Metabolite Production. *In:* V., N. & Willaert, R. (eds.) *Applications of Cell Immobilisation Biotechnology.* Dordrecht, The Netherlands: Springer.
Piso, D. 2014. X-Ray and Neutron Scattering Foundations for the Research in Antimicrobials. *In:* Villa, T. G. & Veiga-Crespo, P. (eds.)

Antimicrobial Compounds: Current Strategies and New Alternatives. Heidelberg: Springer.
Podar, M. & Reysenbach, A. L. 2006. New opportunities revealed by biotechnological explorations of extremophiles. *Curr Opin Biotechnol,* 17**,** 250-5.
Pokusaeva, K., Kuisiene, N., Jasinskyte, D., Rutiene, K., Saleikiene, J. & Chitavichius, D. 2009. Novel bacteriocins produced by Geobacillus stearothermophilus. *Central European Journal of Biology,* 4**,** 196-203.
Pranckute, R., Kaunietis, A., Kananavičiute, R., Lebedeva, J., Kuisiene, N., Šaleikiene, J. & Čitavičius, D. 2015. Differences of antibacterial activity spectra and properties of bacteriocins, produced by Geobacillus sp. bacteria isolated from different environments. *Journal of Microbiology, Biotechnology and Food Sciences,* 05**,** 155-161.
Qian, H., Yao, Q., Tai, C., Deng, Z., Gan, J. & Ou, H. Y. 2018. Identification and characterization of acetyltransferase-type toxin-antitoxin locus in Klebsiella pneumoniae. *Mol. Microbiol.,* 108**,** 336-349.
Rainey, F. A., Fritze, D. & Stackebrandt, E. 1994. The phylogenetic diversity of thermophilic members of the genus Bacillus as revealed by 16S rDNA analysis. *FEMS Microbiology Letters,* 115**,** 205-212.
Ramage, H. R., Connolly, L. E. & Cox, J. S. 2009. Comprehensive functional analysis of Mycobacterium tuberculosis toxin-antitoxin systems: implications for pathogenesis, stress responses, and evolution. *PLoS genetics,* 5**,** e1000767-e1000767.
Rath, A., Glibowicka, M., Nadeau, V. G., Chen, G. & Deber, C. M. 2009. Detergent binding explains anomalous SDS-PAGE migration of membrane proteins. *Proceedings of the National Academy of Sciences of the United States of America,* 106**,** 1760-1765.
Rath, A., Nadeau, V. G., Poulsen, B. E., Ng, D. P. & Deber, C. M. 2010. Novel Hydrophobic Standards for Membrane Protein Molecular Weight Determinations via Sodium Dodecyl Sulfate-Polyacrylamide Gel Electrophoresis. *Biochemistry,* 49**,** 10589-10591.
Ren, D., Bedzyk, L. A., Thomas, S. M., Ye, R. W. & Wood, T. K. 2004. Gene expression in Escherichia coli biofilms. *Appl Microbiol Biotechnol,* 64**,** 515-524.
Repka, L. M., Chekan, J. R., Nair, S. K. & Van Der Donk, W. A. 2017. Mechanistic Understanding of Lanthipeptide Biosynthetic Enzymes. *Chem Rev,* 117**,** 5457-5520.
Rivera-Gomez, N., Martinez-Nunez, M. A., Pastor, N., Rodriguez-Vazquez, K. & Perez-Rueda, E. 2017. Dissecting the protein architecture of DNA-binding transcription factors in bacteria and archaea. *Microbiol.,* 163**,** 1167-1178.
Rocker, A. & Meinhart, A. 2016. Type II toxin: antitoxin systems. More than small selfish entities? *Curr Genet,* 62**,** 287-90.
Rodeo 2018. RODEO Manual. Mitchell Research Lab: The University of Illinois. Available at http://www.ripprodeo.org/manual.html

Sala, A., Bordes, P., Fichant, G. & Genevaux, P. 2013. Toxin-Antitoxin Loci in *Mycobacterium tuberculosis*. *In:* Gerdes, K. (ed.) *Prokaryotic Toxin-Antitoxins.* Newcastle upon Tyne, UK: Springer.

Sala, A., Bordes, P. & Genevaux, P. 2014. Multiple toxin-antitoxin systems in *Mycobacterium tuberculosis*. *Toxins (Basel),* 6**,** 1002-20.

Salameh, E. 1986. *Curative Water in Jordan,* Amman, Jordan, Water Research and Study Centre (WRSC)-University of Jordan.

Salazar, O. & Asenjo, J. A. 2007. Enzymatic lysis of microbial cells. *Biotechnology Letters,* 29**,** 985-994.

Saleem, M. 2014. Natural Products as Antimicrobial Agents – an Update. *In:* David A. Phoenix, Frederick Harris, a. & Dennison, S. R. (eds.) *Novel Antimicrobial Agents and Strategies.* Weinheim: Wiley-VCH Verlag Gmbh & Co.

Salmon, P. M. & Robertson, C. R. 1987. Mass-Transfer Limitations in Gel Beads Containing Growing Immobilized Cells. *Journal of Theoretical Biology,* 125**,** 325-332.

Sampedro, J. & Valdivia, E. R. 2014. New Antimicrobial Agents of Plant Origin. *In:* Villa, T. G. & Veiga-Crespo, P. (eds.) *Antimicrobial Compounds: Current Strategies and New Alternatives.* Heidelberg: Springer.

Sandiford, S. K. 2014. Advances in the arsenal of tools available enabling the discovery of novel lantibiotics with therapeutic potential. *Expert Opinion on Drug Discovery,* 9**,** 283-297.

Sat, B., Reches, M. & Engelberg-Kulka, H. 2003. The *Escherichia coli mazEF* Suicide Module Mediates Thymineless Death. *Journal of Bacteriology,* 185**,** 1803-1807.

Satyanarayana, T., Sharma, A., Mehta, D., Puri, A. K., Kumar, V., Nisha, M. & Joshi, S. 2012. Biotechnological Applications of Biocatalysts from the Firmicutes Bacillus and Geobacillus Species. 343-379.

Sawarieh, A. 2008. Thermal Water Resources in Jordan. *Geothermal Training Programme, .* 30th Anniversary Workshop ed. Reykjavík, Iceland.

Sberro, H., Leavitt, A., Kiro, R., Koh, E., Peleg, Y., Qimron, U. & Sorek, R. 2013. Discovery of functional toxin/antitoxin systems in bacteria by shotgun cloning. *Mol. Cell,* 50**,** 136-48.

Schumacher, M. A., Piro, K. M., Xu, W., Hansen, S., Lewis, K. & Brennan, R. G. 2009. Molecular mechanisms of HipA-mediated multidrug tolerance and its neutralization by HipB. *Science (New York, N.Y.),* 323**,** 396-401.

Shafia, F. 1966. Thermocins of Bacillus Stearothermophilus. *Journal of Bacteriology,* 92**,** 524-&.

Shah, D., Zhang, Z., Khodursky, A., Kaldalu, N., Kurg, K. & Lewis, K. 2006. Persisters: a distinct physiological state of E. coli. *BMC Microbiol,* 6**,** 53.

Shah, S. A. & Garrett, R. 2013. Archaeal Type II Toxin-Antitoxins. *In:* Gerdes, K. (ed.) *Prokaryotic Toxin-Antitoxins.* Newcastle upon Tyne, UK: Springer.

Shapira, A., Shapira, S., Gal-Tanamy, M., Zemel, R., Tur-Kaspa, R. & Benhar, I. 2012. Removal of hepatitis C virus-infected cells by a zymogenized bacterial toxin. *PLoS One,* 7**,** e32320.

Shore, S. & Paul, N. 2010. Robust PCR amplification of GC-rich targets with Hot Start 7-deaza-dGTP. *BioTechniques,* 49**,** 841-843.

Singh, A. P., Singh, R. B. & Mishra, S. 2012a. Microbial and Biochemical Aspects of Antibiotic Producing Microorganisms from Soil Samples of Certain Industrial Area of India- An Overview. *The Open Nutraceuticals Journal,* 5**,** 107-112

Singh, J. V., Chaudhary, R. K. & Gautam, P. 2012b. In silico Analysis of Novel Relb, Rele and Mazf Toxin-Antitoxin Homolog's from the Genome of Xenorhabdus Nematophila. *American Journal of Bioinformatics Research,* 2**,** 21-32.

Skjerning, R. B., Senissar, M., Winther, K. S., Gerdes, K. & Brodersen, D. E. 2019. The RES domain toxins of RES-Xre toxin-antitoxin modules induce cell stasis by degrading NAD+. *Mol Microbiol,* 111**,** 221-236.

Srivatsan, A. & Wang, J. D. 2008. Control of bacterial transcription, translation and replication by (p)ppGpp. *Curr Opin Microbiol,* 11**,** 100-5.

Steen, H. & Mann, M. 2004. The ABC's (and XYZ's) of peptide sequencing. *Nat Rev Mol Cell Biol,* 5**,** 699-711.

Suzuki, H., Murakami, A. & Yoshida, K.-I. 2012. Counterselection System for Geobacillus kaustophilus HTA426 through Disruption of *pyrF* and *pyrR*. *Applied and Environmental Microbiology,* 78**,** 7376-7383.

Szekeres, S., Dauti, M., Wilde, C., Mazel, D. & Rowe-Magnus, D. A. 2007. Chromosomal toxin-antitoxin loci can diminish large-scale genome reductions in the absence of selection. *Mol Microbiol,* 63**,** 1588-605.

Taboada, B., Ciria, R., Martinez-Guerrero, C. E. & Merino, E. 2012. ProOpDB: Prokaryotic Operon DataBase. *Nucleic Acids Res,* 40**,** D627-31.

Taboada, B., Estrada, K., Ciria, R. & Merino, E. 2018. Operon-mapper: a web server for precise operon identification in bacterial and archaeal genomes. *Bioinformatics,* 34**,** 4118-4120.

Takami, H., Inoue, A., Fuji, F. & Horikoshi, K. 1997. Microbial flora in the deepest sea mud of the Mariana Trench. *FEMS Microbiology Letters,* 152**,** 279-285.

Takami, H., Takaki, Y., Chee, G. J., Nishi, S., Shimamura, S., Suzuki, H., Matsui, S. & Uchiyama, I. 2004. Thermoadaptation trait revealed by the genome sequence of thermophilic Geobacillus kaustophilus. *Nucleic Acids Res.,* 32**,** 6292-303.

Tang, W., Dong, S. H., Repka, L. M., He, C., Nair, S. K. & Van Der Donk, W. A. 2015. Applications of the class II lanthipeptide protease LicP for sequence-specific, traceless peptide bond cleavage. *Chem Sci,* 6**,** 6270-6279.

The uniprot consortium 2017. UniProt: the universal protein knowledgebase. *Nucleic acids research,* 45**,** D158-D169.

Tian, Q. B., Hayashi, T., Murata, T. & Terawaki, Y. 1996. Gene Product Identification and Promoter Analysis of *hig* Locus of Plasmid Rts1. *Biochem. Biophys. Res. Commun.,* 225**,** 679-684.

Tietz, J. I., Schwalen, C. J., Patel, P. S., Maxson, T., Blair, P. M., Tai, H. C., Zakai, U. I. & Mitchell, D. A. 2017. A new genome-mining tool redefines the lasso peptide biosynthetic landscape. *Nat Chem Biol,* 13**,** 470-478.

Tiwari, B. K., Valdramidis, V. P., O'donnell, C. P., Muthukumarappan, K., Bourke, P. & Cullen, P. J. 2009. Application of natural antimicrobials for food preservation. *J Agric Food Chem,* 57**,** 5987-6000.

Tiwari, P., Arora, G., Singh, M., Kidwai, S., Narayan, O. P. & Singh, R. 2015. MazF ribonucleases promote Mycobacterium tuberculosis drug tolerance and virulence in guinea pigs. *Nat Commun,* 6**,** 6059.

Torrent, M., Di Tommaso, P., Pulido, D., Nogues, M. V., Notredame, C., Boix, E. & Andreu, D. 2012a. AMPA: an automated web server for prediction of protein antimicrobial regions. *Bioinformatics,* 28**,** 130-1.

Torrent, M., Victoria Nogues, M. & Boix, E. 2012b. Discovering New In Silico Tools for Antimicrobial Peptide Prediction. *Current Drug Targets,* 13**,** 1148-1157.

Tracanna, V., De Jong, A., Medema, M. H. & Kuipers, O. P. 2017. Mining prokaryotes for antimicrobial compounds: from diversity to function. *FEMS Microbiology Reviews,* 41**,** 417-429.

Tsilibaris, V., Maenhaut-Michel, G., Mine, N. & Van Melderen, L. 2007. What is the benefit to Escherichia coli of having multiple toxin-antitoxin systems in its genome? *J Bacteriol,* 189**,** 6101-8.

Unterholzner, S. J., Poppenberger, B. & Rozhon, W. 2013. Toxin-antitoxin systems: Biology, identification, and application. *Mob Genet Elements,* 3**,** e26219.

Van Heel, A. J., De Jong, A., Montalbán-López, M., Kok, J. & Kuipers, O. P. 2013. BAGEL3: automated identification of genes encoding bacteriocins and (non-)bactericidal posttranslationally modified peptides. *Nucleic Acids Research,* 41**,** W448-W453.

Van Heel, A. J., De Jong, A., Song, C., Viel, J. H., Kok, J. & Kuipers, O. P. 2018. BAGEL4: a user-friendly web server to thoroughly mine RiPPs and bacteriocins. *Nucleic Acids Res,* 46**,** W278-W281.

Van Melderen, L. & De Bast, M. S. 2009. Bacterial Toxin–Antitoxin Systems: More Than Selfish Entities? *PLoS Genet* 5**,** e1000437.

Verma, J. P., Jaiswal, D. K., Krishna, R., Prakash, S., Yadav, J. & Singh, V. 2018. Characterization and Screening of Thermophilic Bacillus Strains for Developing Plant Growth Promoting Consortium From Hot Spring of Leh and Ladakh Region of India. *Frontiers in microbiology,* 9**,** 1293-1293.

Viedma, P. M., Abriouel, H., Ben Omar, N., Lopez, R. L., Valdivia, E. & Galvez, A. 2009. Inactivation of Geobacillus stearothermophilus in

canned food and coconut milk samples by addition of enterocin AS-48. *Food Microbiol,* 26**,** 289-93.
Villa, T. G. & Veiga-Crespo, P. 2010. *Enzybiotics: antibiotics enzymes as drugs and therapeutics.*, Willey, Hoboken.
Villa, T. G. & Veiga-Crespo, P. 2014. *Antimicrobial Compounds: Current Strategies and New Alternatives,* Heidelberg, Springer.
Waghu, F. H., Barai, R. S., Gurung, P. & Idicula-Thomas, S. 2016. CAMPR3: a database on sequences, structures and signatures of antimicrobial peptides. *Nucleic Acids Res,* 44**,** D1094-7.
Wang, G., Li, X. & Wang, Z. 2016. APD3: the antimicrobial peptide database as a tool for research and education. *Nucleic Acids Res,* 44**,** D1087-93.
Wang, X. & Wood, T. K. 2011. Toxin-antitoxin systems influence biofilm and persister cell formation and the general stress response. *Appl Environ Microbiol,* 77**,** 5577-83.
Weber, T., Blin, K., Duddela, S., Krug, D., Kim, H. U., Bruccoleri, R., Lee, S. Y., Fischbach, M. A., Muller, R., Wohlleben, W., Breitling, R., Takano, E. & Medema, M. H. 2015. antiSMASH 3.0-a comprehensive resource for the genome mining of biosynthetic gene clusters. *Nucleic Acids Res,* 43**,** W237-43.
Wen, Y., Behiels, E. & Devreese, B. 2014. Toxin–Antitoxin systems: their role in persistence, biofilm formation, and pathogenicity. *Pathogens and Disease,* 70**,** 240-249.
Wilbaux, M., Mine, N., Guerout, A. M., Mazel, D. & Van Melderen, L. 2007. Functional interactions between coexisting toxin-antitoxin systems of the ccd family in Escherichia coli O157:H7. *J Bacteriol,* 189**,** 2712-9.
Willey, J. M. & Van Der Donk, W. A. 2007. Lantibiotics: Peptides of Diverse Structure and Function. *Annual Review of Microbiology,* 61**,** 477-501.
Williams, J. J. & Hergenrother, P. J. 2012. Artificial activation of toxin-antitoxin systems as an antibacterial strategy. *Trends Microbiol,* 20**,** 291-8.
Winther, K. S. & Gerdes, K. 2011. Enteric virulence associated protein VapC inhibits translation by cleavage of initiator tRNA. *Proceedings of the National Academy of Sciences,* 108**,** 7403-7407.
Witte, G., Hartung, S., Buttner, K. & Hopfner, K. P. 2008. Structural biochemistry of a bacterial checkpoint protein reveals diadenylate cyclase activity regulated by DNA recombination intermediates. *Mol Cell,* 30**,** 167-78.
Wong, C. M. V. L., Tam, H. K., Alias, S. A., González, M., González-Rocha, G. & Domínguez-Yévenes, M. 2011. Pseudomonas and Pedobacter isolates from King George Island inhibited the growth of foodborne pathogens. *Polish Polar Research,* 32**,** 3-14.
Wright, G. D. & Sutherland, A. D. 2007. New strategies for combating multidrug-resistant bacteria. *Trends Mol Med,* 13**,** 260-7.

Wyckoff, T. J., Taylor, J. A. & Salama, N. R. 2012. Beyond growth: novel functions for bacterial cell wall hydrolases. *Trends in Microbiology,* 20**,** 540-547.
Xia, K., Zhang, S. J., Solina, B. A., Barquera, B. & Colon, W. 2010. Do Prokaryotes Have More Kinetically Stable Proteins Than Eukaryotic Organisms? *Biochemistry,* 49**,** 7239-7241.
Xie, Y., Wei, Y., Shen, Y., Li, X., Zhou, H., Tai, C., Deng, Z. & Ou, H. Y. 2018. TADB 2.0: an updated database of bacterial type II toxin-antitoxin loci. *Nucleic Acids Res,* 46**,** D749-D753.
Xin, B., Zheng, J., Xu, Z., Song, X., Ruan, L., Peng, D. & Sun, M. 2015. The Bacillus cereus group is an excellent reservoir of novel lanthipeptides. *Appl Environ Microbiol,* 81**,** 1765-74.
Yamaguchi, Y. & Inouye, M. 2011. Regulation of growth and death in *Escherichia coli* by toxin-antitoxin systems. *Nat Rev Microbiol,* 9**,** 779-90.
Yamaguchi, Y., Park, J. H. & Inouye, M. 2009. MqsR, a crucial regulator for quorum sensing and biofilm formation, is a GCU-specific mRNA interferase in Escherichia coli. *J Biol Chem,* 284**,** 28746-53.
Yamaguchi, Y., Park, J. H. & Inouye, M. 2011. Toxin-antitoxin systems in bacteria and archaea. *Annu Rev Genet,* 45**,** 61-79.
Yang, Q. E. & Walsh, T. R. 2017. Toxin–antitoxin systems and their role in disseminating and maintaining antimicrobial resistance. *FEMS microbiology reviews,* 41**,** 343-353.
Yao, J., Guo, Y., Zeng, Z., Liu, X., Shi, F. & Wang, X. 2015a. Identification and characterization of a HEPN-MNT family type II toxin-antitoxin in *Shewanella oneidensis*. *Microb Biotechnol,* 8**,** 961-73.
Yao, X., Chen, T., Shen, X., Zhao, Y., Wang, M., Rao, X., Yin, S., Wang, J., Gong, Y., Lu, S., Le, S., Tan, Y., Tang, J., Fuquan, H. & Li, M. 2015b. The chromosomal SezAT toxin-antitoxin system promotes the maintenance of the SsPI-1 pathogenicity island in epidemic Streptococcus suis. *Mol Microbiol,* 98**,** 243-57.
Yarbrough, G. G., Taylor, D. P., Rowlands, R. T., Crawford, M. S. & Lasure, L. L. 1993. Screening Microbial Metabolites for New Drugs Theoretical and Practical Issues. *The Journal of Antibiotics,* 46**,** 535-544.
Yarmolinsky, M. B. 1995. Programmed Cell Death in Bacterial Populations. *Science,* 267**,** 836-837.
Yeo, C. C. 2018. GNAT toxins of bacterial toxin-antitoxin systems: acetylation of charged tRNAs to inhibit translation. *Mol. Microbiol.,* 108**,** 331-335.
Yuan, J., Sterckx, Y., Mitchenall, L. A., Maxwell, A., Loris, R. & Waldor, M. K. 2010. Vibrio cholerae ParE2 poisons DNA gyrase via a mechanism distinct from other gyrase inhibitors. *J Biol Chem,* 285**,** 40397-408.
Zarrella, T. M., Metzger, D. W. & Bai, G. 2018. Stress Suppressor Screening Leads to Detection of Regulation of Cyclic di-AMP Homeostasis by a Trk Family Effector Protein in *Streptococcus pneumoniae*. *Journal of Bacteriology,* 200**,** e00045-18.

Zeigler, D. 2001. Bacillus Genetic Stock Center Catalog of Strains: The Genus Geobacillus - Introduction and Strain Catalog. 7th ed. Columbus, Ohio: The Ohio State University.

Zielenkiewicz, U. & Cegłowski, P. 2005. The Toxin-Antitoxin System of the Streptococcal Plasmid pSM19035. *Journal of Bacteriology*, 187**,** 6094-6105.

New Life

Since antibiotics are often overused
And the course begins and then disused
Resistant bacteria are suffused

Antibiotics are in states of divergency
Leaving the world in an emergency
Could thermophiles defeat insurgency?

Bacteria play a symphony of peace and war
Creating bunches of mysteries for us to explore
In the sparkling Zara, thermophiles triumph is our score

From hot springs, new life may rise
Unseen creatures with hidden surprise
It s in research where the prize lies

Science radiates rays of sunshine
Enlightening the research plan design
And keeping the mind s radar always online

Thermophiles deserve our appreciation
Marvelous nation with lifelong fascination
This thesis, a tiny candle for the horizon's illumination

Printed by Media-Tryck, Lund 2019 NORDIC SWAN ECOLABEL 3041 0903

LUND UNIVERSITY

ISBN: 978-91-7422-710-9
ISRN: LUTKDH/TKBT-19/1173-SE
Division of Biotechnology
Faculty of Engineering, LTH
Lund University