VILNIUS UNIVERSITY

ARNOLDAS KAUNIETIS

IDENTIFICATION, HETEROLOGOUS BIOSYNTHESIS AND CHARACTERIZATION OF NOVEL BACTERIOCINS FROM THERMOPHILIC BACTERIA

Doctoral dissertation

Biomedical sciences, biology (01B)

Vilnius, 2018

The dissertation was prepared during the period of 2013-2017 at Vilnius University (Lithuania). Part of this research work was carried out during the traineeship (from 01-11-2015 till 01-05-2017) at the University of Groningen (the Netherlands).

Doctoral advisors:

- Prof. Dr. Donaldas J. Čitavičius (Vilnius University, biomedical sciences, biology – 01B. From 01-10-2013 till 31-01-2016).
- Prof. Dr. Lilija Kalėdienė (Vilnius University, biomedical sciences, biology – 01B. From 01-02-2016 till 30-09-2017).

VILNIAUS UNIVERSITETAS

ARNOLDAS KAUNIETIS

NAUJŲ BAKTERIOCINŲ IŠ TERMOFILINIŲ BAKTERIJŲ IDENTIFIKAVIMAS, HETEROLOGINĖ BIOSINTEZĖ BEI CHARAKTERIZAVIMAS

Daktaro disertacija

Biomedicinos mokslai, biologija (01B)

Vilnius, 2018

Disertacija rengta 2013-2017 metais Vilniaus universitete (Lietuva). Dalis šio darbo tyrimų atlikti stažuotės laikotarpiu (nuo 01-11-2015 iki 01-05-2017) Groningeno universitete (Nyderlandai).

Doktoranto vadovai:

- Prof. dr. Donaldas J. Čitavičius (Vilniaus universitetas, biomedicinos mokslai, biologija – 01B. Nuo 2013-10-01 iki 2016-01-31).
- Prof. dr. Lilija Kalėdienė (Vilniaus universitetas, biomedicinos mokslai, biologija – 01B. Nuo 2016-02-01 iki 2017-09-30).

Contents

1 Abbreviations

AMP - antimicrobial peptide

NRPS - nonribosomal peptide synthase

NRP - nonribosomal peptide

LAB - lactic acid bacteria

RiPP - ribosomally synthesized and post-translationally modified peptide

rSAM - radical S-adenosylmethionine

SKF - sporulation killing factor

PTS - phosphoenolpyruvate:sugar phosphotransferase system

NMR - nuclear magnetic resonance spectroscopy

LAP - linear azol(in)e-containing peptide

SDS-PAGE - sodium dodecyl sulfate polyacrylamide gel electrophoresis

DNA - deoxyribonucleic acid

RNA - ribonucleic acid

MIC - minimum inhibitory concentration

Man - mannose

GlcNAc - N-acetylglucosamine

ORF - open reading frame

BLIS - bacteriocin-like inhibitory substance

ATP - adenosine triphosphate

MALDI-TOF - matrix assisted laser desorption ionization - time of flight

MS - mass spectrometry

IAA - iodoacetamide

GGDC - genome-to-genome distance calculator

dDDH - digital DNA-DNA hybridization

DDH - DNA-DNA hybridization

BLAST - basic local alignment search tool

MCS - multicloning site

PCR - polymerase chain reaction

UDP - uracil DNA glycosylase

RBS - ribosome binding site

RP-HPLC - reverse phase - high pressure liquid chromatography

LC-ESI-MS - liquid chromatography-electrospray ionization - mass spectrometry

MS/MS - tandem mass spectrometry

GC-MS - gas chromatography - mass spectrometry

TCEP - tris(2-carboxyethyl)phosphine

NB - nutrient broth

Glc - glucose

AU - arbitrary unit

CBE - cell free extract

AEC - anion-exchange chromatography

CEC - cation-exchange chromatography

HIC - hydrophobic interaction chromatography

IFE - isoelectric focusing

IPG - isoelectric focusing polyacrylamide gel

LB - lysogeny broth

TFA - trifluoroacetic acid

ACN - acetonitrile

IPTG - Isopropyl β-D-1-thiogalactopyranoside

IMAC - immobilized metal affinity chromatography

CAM - carboxyamidomethyl

CFU - colony forming unit

RPM - revolutions per minute

ABC - ATP-binding cassette

2 Introduction

A variety of organisms, such as plants, animals, insects, bacteria or fungi, produce and secrete bioactive compounds with antimicrobial properties as a part of their innate immunity. Among these compounds a group of low molecular weight proteins or peptides are defined as antimicrobial peptides (AMPs). They have attracted a big interest as potential alternatives to conventional antibiotics or antimicrobial compounds that could be applied in various fields of industry (Nguyen et al., 2011; Molchanova et al., 2017; Liu et al., 2017).

Bacteriocins are AMPs synthesized by ribosomes and produced by bacteria. They are active against other bacteria, which are usually closely related to the producer. These peptides exhibit considerable diversity with respect to their size, structure, mechanism of action, inhibitory spectrum, immunity mechanisms, and targeted receptors (Hegarty et al., 2016). In the era of emergence of antibiotic resistant bacteria (WHO, 2015), bacteriocins have been suggested as a potential alternative to antibiotics in clinics and veterinary settings, also as food preservatives against spoilage and pathogenic microorganisms (Lubelski et al., 2008; Cotter et al., 2013; Lagha et al., 2017).

Mesophilic *Bacillus* bacteria and thermophilic *Geobacillus* sp. or *Aeribacillus pallidus* bacteria species are able to survive industrial pasteurization and form biofilms within pipes of stainless steel equipment in a diary food production (Gopal et al., 2015). Bacteriocins as natural products could help to avoid spore forming thermophilic bacteria contamination in some industry processes. Moreover, thermophilic bacteria were shown to have great potential for applications in biofuel production (Ahmad et al., 2017; Jiang et al., 2017). In these industrial processes contamination of other bacteria is undesirable, and bacteriocins as heat-stable products could be used for biocontrol.

Studies on bacteriocins have a great interest. It discloses our knowledge and understanding about these compounds, reveales their diversity. Research on

bacteriocins may reveal novel antibacterial peptides or proteins, it helps to understand their mode of actions, biosynthesis pathways, post-translational modifications or other valuable properties. New knowledge in this field will facilitate applycation of bacteriocins in industry of food, veterinary or medicine. There are many microorganizms, which were not extensively studied for bacteriocins before. For this reason, they could be a good source for the new ones. Moreover, new bacteriocins can have novel post-translational modifications which provide valuable properties for the peptides. Enzymes performing these modifications could be applied in synthesis of new bacteriocins or other bioactive peptides with improved properties.

2.1 Aim of the study

Identification, biosynthesis and characterization of novel bacteriocins of gram-positive thermophilic endo-spore forming bacteria.

Tasks

- To identify genes of novel bacteriocins encoded in thermophilic *Aeribacillus pallidus* 8 and *G. stearothermophilus* 15 by genome mining approach.
- To purify and determine the amino acid sequence of novel bacteriocins produced by thermophilic *Geobacillus stearothermophilus* 15.
- To perform heterologous expression and characterization of new bacteriocins.
- To screen and evaluate the substrate specificity of identified post-translational modification enzymes and the possibility to apply them for heterologous biosynthesis of novel bacteriocins in *E. coli*.

2.2 Novelty and significance of this work

Detailed studies on bacteriocins of thermophilic bacteria are very scarce. Most of them are described only obscurely and their amino acid sequences are not revealed.

This study describes novel bacteriocins from two thermophilic endo-spore forming bacteria *A. pallidus* 8 and *G. stearothermophilus* 15. A novel bacteriocin of glycocin subclass was identified and found to be encoded in strain 8. It is post-translationally modified by an S-linked glucose on a specific Cys residue. Such glycosylation is a very rare phenomenon among bacteria. This study reports the first successful case of heterologous glycocin biosynthesis and also shows that the leader processing is not absolutely necessary for activity, but glycosylation is. Moreover, for the first time the characterized biosynthetic machinery of glycocin was applied to produce other novel glycocins. These findings expand the knowledge about the small class of glycocins and enables efficient screening for novel glycocins.

Additionally, there has been identified a novel high molecular weight bacteriocin produced by strain 15. This antibacterial protein belongs to the IIIrd class of bacteriocins which has not been studied extensively to date. The novel bacteriocin has no sequence similarity to any other known bacteriocins. Further studies on the novel bacteriocin could expand the knowledge about the IIIrd class of bacteriocins. It enables efficient screening for the novel high molecular weight bacteriocins.

2.3 Thesis statement

- The research on antibacterial compounds produced by thermophilic bacteria and bacteriocin mining in genomes of these bacteria reveals new bacteriocins.
- Novel post-translationally modified glycocin identified in thermophilic bacteria exposes high stability and extremely low minimal inhibitory concentration. Another novel identified unmodified high-molecular-weight bakteriocin shares no sequence similarity to any known bacteriocin.
- First time ever, it was demonstrated that glycocins can be synthesized in a heterologous host by the expression of gene cluster of their biosynthetic machinery.

- Moreover, enzymes of glycocin post-translationally machinery can by applied for heterologous biosynthesis of other novel glycocins.

3 Literature review

3.1 Bacteriocins - the antimicrobial peptides

In bacteria AMPs are synthesized by ribosomes (Perez et al., 2014) or by nonribosomal peptide synthases (NRPSs). Nonribosomal peptides (NRPs) synthesized by NRPSs are produced by the fusion of amino acids, which is achieved by large multimodular enzymes – NRPSs. Non canonical amino acids, fatty acids and α-hydroxy acids have been also identified as building blocks for NRPs (Strieker et al., 2010). Meanwhile, the ribosomally synthesized AMPs derived from bacteria are named *bacteriocins* (Cotter et al., 2005). The first description of bacteriocin mediated inhibition was reported by A. Gratia (1925), when antagonism between strains of *Escherichia coli* was first discovered. Since then they were called 'colicins', to reflect the original producer organism (Jack et al., 1995; Cotter et al., 2005). Colicins are comparatively large antibacterial proteins (> 10 kDa) and are the prototype bacteriocins of gram-negative bacteria. The term 'microcin' was introduced to separate group of gram-negative antibacterial peptides with a size < 10 kDa from colicins with a higher molecular mass (Asensio et al., 1976; Gillor et al., 2004).

In contrast to bacteriocins of gram-negative bacteria, bacteriocins of gram-positive bacteria are usually small peptides (<10 kDa). Many of them are synthesized ribosomally as precursor peptides and then subsequently are modified post-translationally to generate their antibacterial activity. Post-translational modification serves to confer specific chemical properties that could not be obtained by a peptide synthesis alone. Furthermore, post-translational modifications can be used as a mechanism to control the activation of the antibacterial activities of the bacteriocin, and thus exert a level of control and host immunity (Perez et al., 2014).

They are active against other bacteria, either in the same species (narrow spectrum), or across genera (broad spectrum) and, as with the host defense peptides, cell signaling mechanisms can be involved, too (Cotter et al., 2005).

There is speculation that the primary function of bacteriocins in their natural ecological environments is not the killing other bacteria. It is assumed that natural concentrations of bacteriocins in their usual environments are relatively low. There are many examples in literature, showing a variety of regulatory functions of bacteriocins, including their own production. From an evolutionary point of view, it would be logical that bacteriocins prevent from settling down of intruder cells by forming biofilms in a bacteriocin-producer niche. In this case the function of bacteriocins, when they are produced in low amounts, could be the signaling or repelling other bacteria species rather than killing them, when they are produced in high amounts (Chikindas et al., 2017).

3.2 Classification of bacteriocins

From the time when the first bacteriocins were discovered and characterized, information about their variety, mode of action, synthesis, secretion and genetic organization was expanding. Regarding their increasing diversity, Klaenhammer (1993) proposed the first classification scheme for bacteriocins. At that time, the highest interest was focused on bacteriocins produced by lactic acid bacteria (LAB). Regarding the knowledge in this field, it was suggested to define bacteriocins into four classes:

I class. Lantibiotics, small membrane-acting peptides (< 5 kDa) containing the unusual amino acids like lanthionine, 3-methyllanthionine or dehydrated residues.

II class. Small heat-stable, non-lanthionine containing membrane-acting peptides (< 10 kDa) which were subsequently divided into smaller groups.

III class. Large heat-labile proteins (> 30 kDa).

IV class. Bacteriocins composed of protein and lipid or carbohydrate moieties required for activity.

Ribosomally synthesized and post-translationally modified peptides (RiPPs) are present in a variety of life forms, not only in bacteria. Genome sequencing efforts facilitated identification and expanded the diversity of these natural products. RiPPs, as natural products with a great variety of structures, have attracted attention of different researcher's communities and the nomenclature is different between these groups of investigators, because they have proposed the universal nomenclature of RiPPs. Bacteriocins, as ribosomally synthesized peptides and which are post-translationally modified, can be grouped and classified according to the universal nomenclature of RiPPs, too (Arnison et al., 2013). This classification was often included in the new bacteriocin classification schemes.

Increasing number of new bacteriocins and their classes reported among gram-negative and gram-positive bacteria required revision of the classification. Since the classification of bacteriocins by Klaenhammer was published, researchers in their new classification schemes updated the class of lantibiotics. Usually, they always proposed to group all modified bacteriocins – ribosomally synthesized antimicrobial peptides with post-translational modifications, to the class I. After all, it included the part of the universal nomenclature for RiPPs. New classifications included class II, composed of bacteriocins with no modifications. Both classes were subdivided into subclasses and some of them differed among various classification schemes. The designation of other large, heat-labile or complex bacteriocins with lipid or carbohydrate moieties and cyclic bacteriocins was controversial between different classification schemes. Moreover, it was argued that the designation of bacteriocins should be retained for peptide antimicrobials (Cotter et al., 2005; Jack et al., 1995; Cotter et al., 2013).

The latest bacteriocin classification (**Fig. 1**) was proposed by Alvarez-Sieiro et al. (2016) and it is designed for LAB produced bacteriocins because the majority of them are synthesized by this group of bacteria.

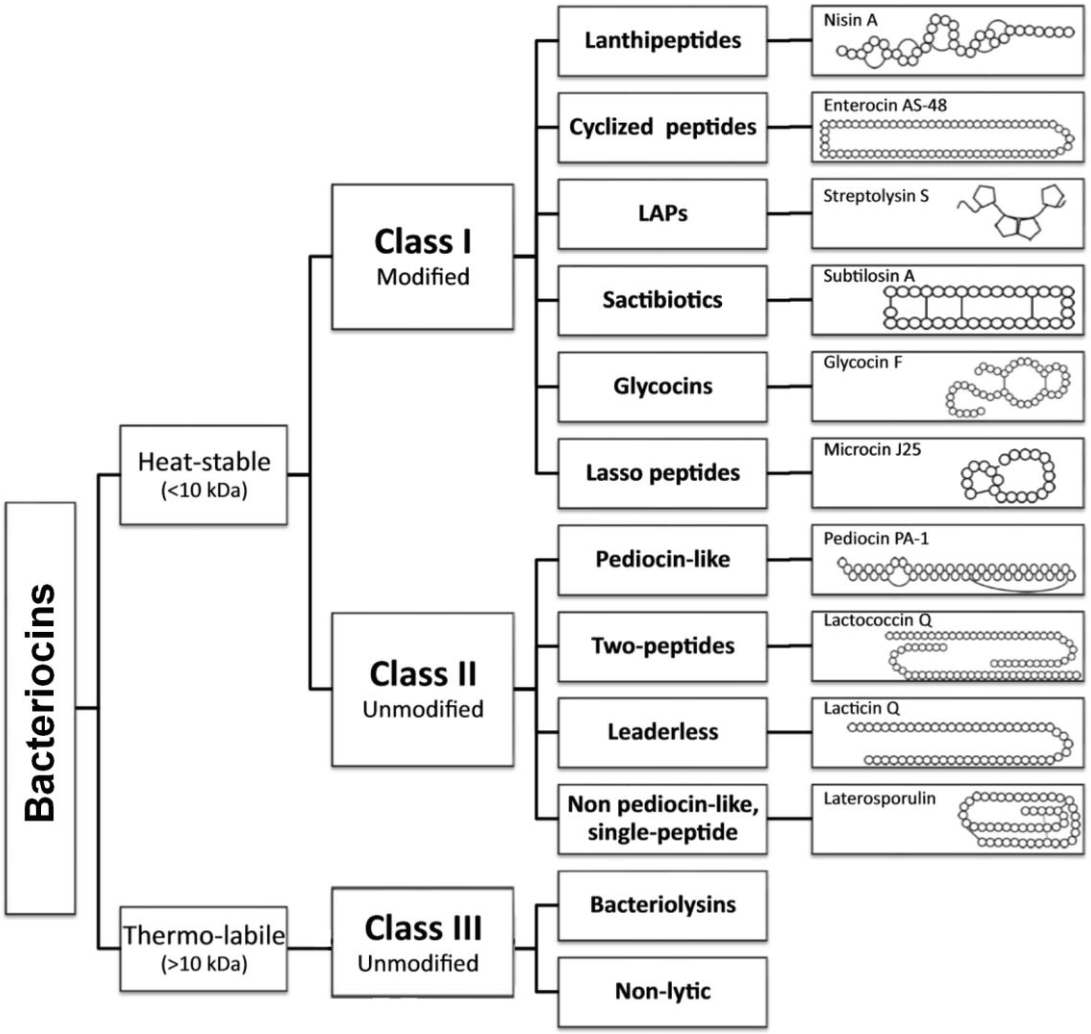


Figure 1. Classification scheme for bacteriocins and their structures (Alvarez-Sieiro et al. 2016).

In addition, LAB produced bacteriocins have great potential and interest in application. Despite that, as authors note, this classification is also applicable for other gram-positive bacteria bacteriocins. As in others, it incorporates part of the universal nomenclature for RiPPs. Based on this scheme they are grouped into three major classes and divided to subclasses:

- **Class I – post-translationally modified bacteriocins.** It encompasses smaller than 10 kDa peptides which undergo post-translational modifications during synthesis (RiPPs). Modifications provide uncommon amino acids and structures that imply on properties of the molecules. These bacteriocins

consist of a leader peptide sequence fused with a core peptide. The leader serves as recognition sequence for modification enzymes, transport or keeping the peptide inactive.

- **Class II – unmodified bacteriocins.** This class groups less than 10 kDa bacteriocins that have no modifications. Thus, they do not require enzymes for their maturation, except a peptidase for the leader sequence cleavage and/or bacteriocin transporter.
- **Class III – large unmodified bacteriocins.** These are larger than 10 kDa, with no modifications and which have bacteriolytic or non-lytic mode of action.

Because of quite big diversity of bacteriocins within class I and II, the description and overview of antibacterial peptides presented in the following sections is focused only on some of the subclasses.

3.2.1 Class I - post-translationally modified bacteriocins

3.2.1.1 Lanthipeptides

It includes antimicrobial peptides with unusual amino acids, such as lanthionine and/or methyl-lanthionine. Lanthipeptides generally are encoded in operons containing genes of enzymes which are involved in the maturation process. Based on these enzymes lanthipeptides are divided into four types. Type I and type II are considered as lantibiotics which exhibits antimicrobial activity, meanwhile type III and type IV have no known antimicrobial activity (Alvarez-Sieiro et al., 2016).

Nisin is the best studied bacteriocin produced by *Lactococcus lactis*. It is type I lantibiotic. Its biosynthetic gene cluster contains 11 genes responsible for biosynthesis and immunity. Lantibiotic maturation encompasses few enzymatic reactions (**Fig. 2**).

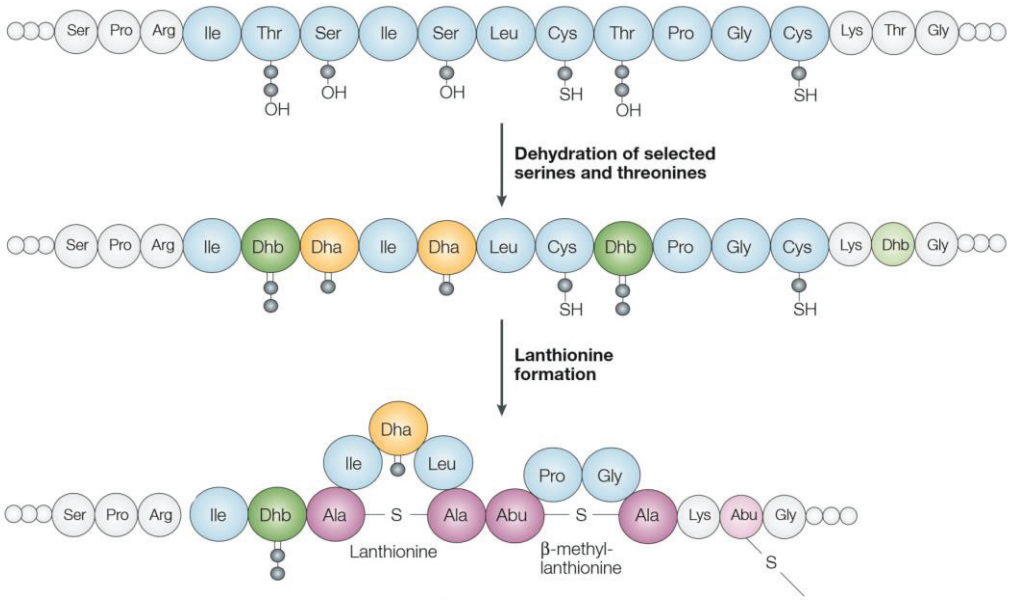


Figure 2. Synthesis and structure of lantibiotics. Modified residues are formed by enzymatic dehydration of Ser (resulting dehydroalanine, Dha) or Thr (resulting dehydrobutyrine, Dhb). Their condensation with the sulphydryl group of a Cys forms bridges and results new lanthionine or β-methyl-lanthionine residues. New bonds form rings within the lantibiotic (Cotter et al., 2005).

Nisin precursor NisA is modified by NisB which dehydrates Ser and Thr amino acids. The cyclase NisC links thiol group of a Cys to the dehydrated residues lanthionine or methyl-lanthionine and generates cyclic structures in the peptide. After modifications, the precursor is transported outside the cell by NisT and the leader sequence is cleaved by NisP protease resulting an active and mature bacteriocin. Producer protects itself against nisin by proteins NisI and NisFEG (Lubelski et al., 2009; Al Khatib et al., 2014; Plat et al., 2017). Type II lantibiotics are modified by bifunctional enzyme, which performs dehydration of Ser and Thr and links them to thiol groups of Cys (Shimafuji et al., 2015).

3.2.1.2 Head-to-tail linked peptides

This type of bacteriocins have covalently linked N- and C-termini by a peptide bond, generating a circular molecule. All of them contain 4 or 5 α-helical segments. There is no C-terminal extension in circular bacteriocins. The leader sequence is cleaved off during the maturation (Alvarez-Sieiro et al., 2016). Usually they are produced

by *Firmicutes* bacteria and have a broad-spectrum antimicrobial activity. This group of bacteriocins can inhibit common food-borne pathogens, like *Clostridium* spp. and *Listeria* spp. (Gabrielsen et al., 2014). Circular bacteriocins are known to possess higher stability than general linear bacteriocins (Masuda et al., 2012).

To date the most well characterized head-to-tail linked bacteriocin is enterocin AS-48 produced by *Enterococcus faecalis*. Its synthesis is determined by 10 genes: *as-48A* which encodes the precursor; *as-48B* encodes a putative cyclase; *as-48C* – DUF95 protein, related to immunity and bacteriocin production; *as-48C1D* are considered as ABC transporter genes; *as-48D1* encodes immunity protein; genes *as-48EFGH* encodes an additional ABC transporter, which is also immunity related (van Belkum et al., 2011; Grande-Burgos et al., 2014).

Gabrielsen et al. (2014) suggested to separate circular bacteriocins into two subgroups based on a physicochemical characteristics and level of sequence identity. Subgroup I includes enterocin AS-48 (Martínez-Bueno et al., 1998), uberolysin (Wirawan et al., 2007), lactocyclicin Q (Sawa et al., 2009), carnocyclin A (van Belkum et al., 2010), leucocyclicin Q (Masuda et al., 2011), garvicinML (Gabrielsen et al., 2012), circularin A (Kemperman et al., 2013), amylocyclin (Scholz et al., 2014) and enterocin NKR-5-3B (Himeno et al., 2015). Peptides of subgroup I have several positively charged amino acid residues and high isoelectric points (~ 10). Most of them have low sequence similarity and have few conserved residues within peptides of this group. Subgroup II comprise butyrivibriocin AR10 (Kalmokoff et al., 2003), gassericin A (Ito et al., 2009) and acidocin B (Acedo et al., 2015). These peptides have high sequence similarity and neutral or low isoelectric points (Gabrielsen et al., 2014).

3.2.1.3 Sactibiotics

According to the RiPPs nomenclature proposed by Arnison et al. (2013) they can be also designated as sactipeptides. Sactibiotics harbor thioether links between sulfur and α-carbon of amino acid residues. Intramolecular thioether bonds are

formed by radical S-adenosylmethionine (rSAM) enzymes. They generate sactionine linkages by cross-linking the sulfur atom of Cys to the α-carbon of an acceptor amino acid. These linkages differ from the Cys-Ser/Thr bonds found in lanthipeptides, were Cys residues are linked to β-carbons of dehydrated residues. There are known five sactipeptides to date: subtilosin A, thurincin H, sporulation killing factor (SKF) and two-component thuricin CD. All of them are considered as sactibiotics, they have antibacterial activity and are produced by *Bacillus* spp. bacteria (Mahanta et al., 2017).

Subtilosin A is 35 amino acid residues long and head-to-tail linked peptide. It contains three thioether bonds linking three Cys residues and α-carbons of two Phe and Thr residues. The biosynthetic gene cluster consists of precursor gene *sboA* and processing genes *albABCDFG*, which are responsible for the post-translational modifications and export of mature peptide (Flühe et al., 2012).

Thurincin H is antimicrobial peptide of 31 amino acid. It possesses sulfur to α-carbon bonds between four Cys and two Thr, Ser or Asn. It is assumed that these links facilitate shape, stability and antibacterial properties of the peptide. rSAM and other proteins encoded in the gene cluster facilitates formation of thioether bridges and maturation of the bacteriocin. Gene cluster of thurincin H encodes ThnA – precursor peptide of 40-residues, ThnB – rSAM enzyme, ThnP – protease, ThnDET – ABC transporter, ThnR – transcriptional regulator, and ThnI is encoding unknown function protein (Sit et al., 2011; Wieckowski et al., 2015).

The SKF is a peptide of 26 amino acids. It has cyclic structure, contains a sactionine linkage between Cys and Met and has a disulfide bond. Gene cluster of SKF is comprised of *skfA* – a precursor gene, *skfB* – gene of rSAM enzyme, *skfC* – gene of a putative protease, *skfH* – gene of a putative thioredoxin, *skfEF* – genes coding proteins for the export and immunity, and *skfG* is gene coding unknown function protein (González-Pastor et al., 2003; Flühe et al., 2013).

Thuricin CD is bacteriocin composed of two peptides Trn-α and Trn-β (49 amino acids and 47 amino acids, respectively). Both peptides are post-translationally

modified and each of them have three sulfur to α-carbon bridges. The linkages are between Cys and Thr, Ser and Asn amino acid residues. Gene cluster of thuricin CD, aside from the two structural genes for Trn-α and Trn-β, encodes proteins TrnFG which are likely ABC transporters responsible for export of the bacteriocin outside the cell, TrnCD which are two rSAM enzymes, and protein TrnE – a putative peptidase that may play a role in self-immunity to thuricin CD (Rea et al., 2010; Sit et al., 2011).

3.2.1.4 Glycocins

It is post-translationally glycosylated bacteriocins (**Fig. 3**). The sugar moieties are linked to side chains of Cys, Ser or Thr residues. A glycocin can be regarded as being 'glycoactive' when a sugar moieties are essential for the antimicrobial activity (Norris and Patchett 2016). Only few of them are identified and reported to date: sublancin 168, produced by *B. subtilis* (Oman et al., 2011); glycocin F, produced by *Lactobacillus plantarum* (Stepper et al., 2011); ASM1 (homologous to glycocin F), produced by *Lb. plantarum* (Hata et al., 2009; Norris and Patchett 2016); enterocin F4-9, produced by *E. faecalis* (Maky et al., 2015) and thurandacin, encoded in *B. thuringiensis* which was identified by genomic data mining and chemoenzymatically synthesized *in vitro* (Wang et al., 2013).

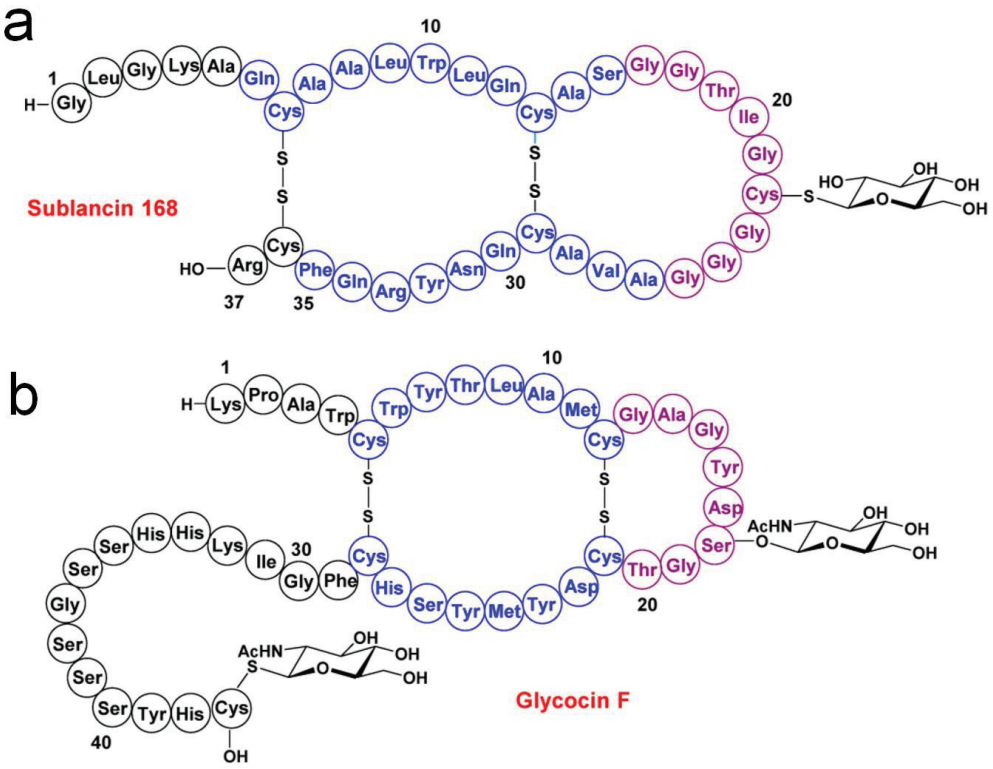


Figure 3. Glycocin structures. (**a**) Sublancin 168, (**b**) glycocin F. Helical regions are in blue, loop regions in magenta (Garcia de Gonzalo et al., 2014).

Not so many studies have been done on these glycosylated antimicrobial peptides. The understanding of growth inhibition is obscure. Studies have shown that phosphoenolpyruvate:sugar phosphotransferase system (PTS) is a factor affecting glycocin F and sublancin 168 antibacterial activities (Garcia De Gonzalo et al., 2015; Norris and Patchett 2016). It is known that sublancin 168 does not affect the integrity of the cell membrane and acts bactericidic (Garcia De Gonzalo et al., 2015). Meanwhile, glycocin F and enterocin F4-9 have a bacteriostatic effect (Stepper et al., 2011; Maky et al., 2015). The nuclear magnetic resonance spectroscopy (NMR) structures of sublancin 168, glycocin F and ASM1 were solved identifying two antiparallel α-helices stabilized by two disulfide bonds between Cys residues (Venugopal et al., 2011; Garcia De Gonzalo et al., 2014; Norris and Patchett 2016). Synthetic machinery of most well studied glycocin – sublancin 168 is encoded in gene cluster containing *sunA, sunS, sunT, bdbA* and *bdbB* genes. Precursor peptide,

SunA, is modified by S-glycosyltransferase, SunS – a unique enzyme which forms very unusual β-S-linkage between Cys residue and glucose. SunA glycosylation by SunS have been confirmed *in vitro* by chemoenzymatical synthesis of mature sublancin 168. Based on SunT sequence similarity to bacteriocin ABC transporter/peptidase it is assumed that SunT cleaves the leader sequence and transports the core peptide outside the cell. Two thiol-disulfide oxidoreductases, BdbA and BdbB, might be responsible for disulfide bonds formation in the peptide (Paik et al., 1998; Dorenbos et al., 2002; Oman et al., 2011). In addition, it was confirmed that the same gene cluster is coding immunity protein SunI (Dubois et al., 2009). Similar genetic organization was found in gene clusters coding putative synthetic machineries of glycocin F (Stepper et al., 2011), thurandacin (Wang et al., 2013) and enterocin F4-9 (Maky et al., 2015).

3.2.1.5 Linear azol(in)e-containing peptides (LAPs)

Alvarez-Sieiro et al. (2016) included this subclass of post-translationally modified bacteriocins in their classification. These peptides have various combinations of heterocyclic rings of thiazole and (methyl)oxazole. The rings are derived from Cys, Ser or Thr residues by ATP-dependent cyclodehydration and subsequent flavin mononucleotide dependent dehydrogenation. Usually inactive LAPs precursors are processed by heterotrimeric complex of synthetase which is composed of dehydrogenase (B) and cyclodehydratase (C/D) (Arnison et al. 2013).

3.2.1.6 Lasso peptides

It is composed of peptides with unique topology. These bacteriocins have a macrolactam ring from 7 to 9 amino acid residues and a linear C-terminal tail. The ring is generated between the N-terminal α-amino group and Asp or Glu side chain located at positions 7, 8, or 9 of the amino acid sequence. The C-terminal tail is threaded through the ring resulting a ’lasso’ structure (Hegemann et al., 2015) (**Fig. 4**).

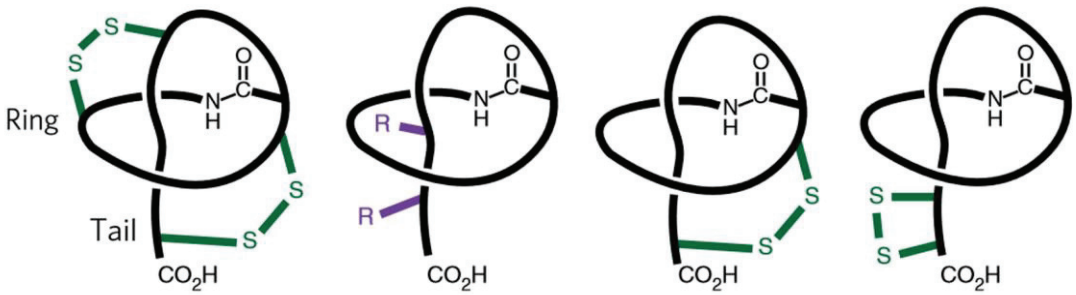


Figure 4. Lasso peptide structures (Tietz et al., 2017).

Sometimes the structure is stabilized by disulfide bonds. One of the earliest bacteriocins of this subclass is microcin J25 (Hegemann et al., 2015). Its biosynthetic gene cluster *mcjABCD* necessary for production and immunity was found encoded in a plasmid (Solbiati et al., 1998).

3.2.2 Class II - unmodified bacteriocins

3.2.2.1 Pediocin-like bacteriocins

It is a subclass of unmodified antibacterial peptides with a highly hydrophilic, charged and conservative N-terminal region, containing typical Tyr-Gly-Asn-Gly-Val/Leu sequence, which is called “pediocin box”. Usually, N-terminus has a disulphide bond linkage as well as leader sequence with a double Gly motif for a proteolytic cleavage. Gene clusters coding for pediocin-like bacteriocins, in addition to a precursor gene, have an ABC transporter, immunity and accessory genes for translocation of the bacteriocin. Some of this type bacteriocins are secreted by the general sec-dependent export system (Cui et al., 2012).

3.2.2.2 Two-peptide bacteriocins

It is a subclass of unmodified bacteriocins composed of two peptides, which are acting as a single unit, but not synergistically. For optimal antibacterial activity both peptides should be in equal ratio. Typically, they are cationic and contain hydrophobic and/or amphiphilic segments. Precursors of these bacteriocins are encoded in two genes. The producer is coding genes for precursors, secretion and

immunity. In addition, it is coding an accessory protein with unknown function (Nissen-Meyer et al., 2010).

3.2.2.3 Leaderless bacteriocins

This subclass differs from all the rest bacteriocins that they have no leader sequence. They do not undergo any post-translational maturation and bacteriocin activation processes (Masuda et al., 2012). Some producers are coding unit of few genes of high homology leaderless bacteriocins which are near each other (Ovchinnikov et al., 2016). There is lack of information about biosynthesis mechanism of this type antibacterial peptides. It is not very clear how they are recognized by the transporter protein, how they are secreted, and does the producer organism express the immunity protein (Masuda et al., 2012). Studies on leaderless bacteriocin lacticin Q suggest that an ABC transporter with accessory proteins may facilitate bacteriocin secretion and immunity (Iwatani et al., 2013).

3.2.2.4 Non-pediocin-like and single-peptide bacteriocins

To this subclass are grouped the rest bacteriocins that do not fit to previous subclasses of unmodified bacteriocins. It compiles heterologous and unrelated linear bacteriocins (Alvarez-Sieiro et al., 2016).

3.2.3 Class III – large-molecular-weight unmodified bacteriocins

They are large-molecular-weight and heat-labile unmodified antimicrobial proteins. This group contains bacteriolytic and non-lytic bacteriocins (Alvarez-Sieiro et al., 2016). They were not very extensively studied comparing to lantibiotics or other antimicrobial peptides.

Enterolysin A is bacteriolytic high molecular weight heat-labile bacteriocin produced by *E. faecalis*. This bacteriocin is 34.5 kDa protein with a sec-dependent signal peptide of 27 amino acids which is cleaved after Val-Asn-Ala residues. Theoretical pI of enterolysin A is 9.2 (Nilsen et al., 2003). It was cloned and expressed in *E. coli* and it was demonstrated that recombinant and purified bacteriocin had 8.5 times lower specific activity per mg of protein, compared to a

native enterolysin A. It was explained by probably improper folding or generation of multimer forms of the protein (Nigutová et al., 2008).

One of the most well characterized class III bacteriocin is zoocin A, which is produced by *Streptococcus zooepidemicus*. It is encoded by gene *zooA* and its product corresponds to the 285 amino acid protein with signal peptide sequence. Biologically active with cleaved signal peptide mature zoocin A has a molecular mass of 27.9 kDa. It is bacteriolytic enzyme with two domains: N-terminus contains endopeptidase and C-terminus has substrate recognition domain. The producer is encoding immunity protein Zif which changes composition of the host's cell wall and prevents producer's cell wall lysis by zoocin A. Cloning and expression of *zooA* in *E. coli* resulted in a secretion of active zoocin A (Simmonds et al., 1997; Beatson et al., 1998; Chen et al., 2013).

Helveticin J produced by *Lb. helveticus* is non-lytic bactericidal protein with a narrow spectrum inhibitory activity. It is active against bacteria closely related to the producer within *Lactobacillaceae* family. Helveticin J molecular weight, resolved by SDS-PAGE analysis, is 37 kDa (Joerger and Klaenhammer 1986). DNA sequence coding the protein was found in the genome. Heterologous expression in *E. coli* of helveticin J was not successful. Meanwhile, heterologous synthesis of the same bacteriocin in *Lb. acidophilus* yielded active recombinant bacteriocin (Joerger and Klaenhammer 1990).

Dysgalacticin is produced by *S. dysgalactiae* subsp. *equisimilis*. It is non-lytic bacteriocin and has a narrow spectrum of antimicrobial activity against human pathogen *S. pyogenes*. Dysgalacticin is 21.5 kDa protein with pI of 4.7. Its structural gene, *dysA*, encodes a 220 amino acids protein with a signal peptide for a sec-dependent transport system. This bacteriocin was successfully expressed in heterologous host *E. coli,* while purified dysgalacticin was active. Predicted secondary structure of dysgalacticin by bioinformatics tools suggested that the protein has a relatively unstructured N-terminal region, and a primarily helical C-

terminal segment. Furthermore, dysgalacticin contains a disulfide bond which is important for bacteriocin bioactivity (Heng et al., 2006).

3.3 Leader sequences of bacteriocins

Leader sequences of bacteriocins have multiple roles in post-translational modifications, export, and immunity. Different biosynthetic enzymes of bacteriocins recognize and interact with different segments of the leader sequence. It is necessary or, in some cases, it strongly facilitates post-translational modifications of the core peptide and is necessary as a receptor for bacteriocin transporters for the transfer of bacteriocin outside the cell. Usually, leader sequences in C-terminus have double glycine motif, Gly-Gly, which composition is critical for the leader cleavage by proteases of bacteriocin biosynthetic machineries (Roy et al., 1998; Sprules et al., 2004; Kotake et al., 2008; Patton et al., 2008). It is known that bacteriocin leader *in vivo* is cleaved after the complete maturation. It results in an active bacteriocin. The leader may be cleaved also in not mature precursor but with lower efficiency (Lagedroste et al., 2017). Bioinformatics analysis and experiments show that leaders usually tend to form α-helices. Interestingly, some bacteriocin leaders form a secondary structure in trifluorethanol, but not in aqueous solutions (Sprules et al., 2004). There are speculations that leader binding with modification proteins may induce functional structure changes in the leader sequence (Khusainov et al., 2013).

Some studies on lantibiotics, e.g. labyrinthopeptin A2, have shown that a conservative sequence in N-terminus of the leader, which forms α-helix, is the most important for the binding modification enzymes (**Fig. 5**) (Müller et al., 2011). Changes in other regions of the leader, like C-terminus, may also impact the binding of some enzymes and processivity of post-translational modifications. Studies on nisin and other bacteriocin leaders have shown a mutational freedom in the C-terminus, which does not influence post-translational modifications in the core peptide. There is only one exception – the introduction of a helix breaking Pro in

the C-terminus modulates the processivity of modification enzymes (Plat et al., 2017). There have been shown that bacteriocin precursors without leader may be modified by modification proteins, but with a decreased enzyme activity (Levengood et al., 2007).

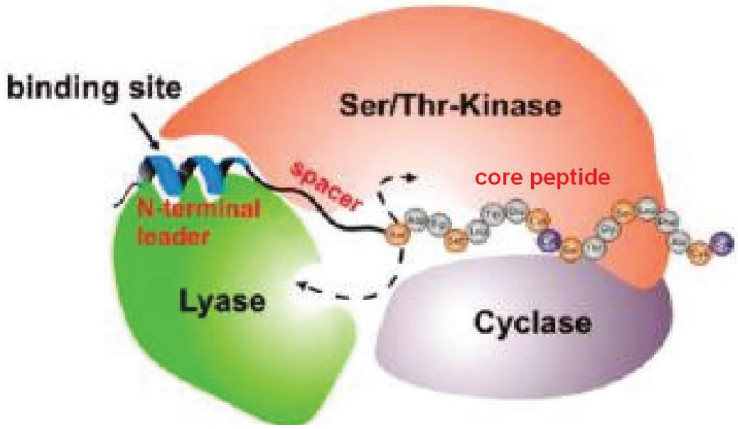


Figure 5. Model for labyrinthopeptin A2 and its leader peptide binding the complex of modification enzymes (Müller et al., 2011).

Studies on unmodified bacteriocin of lasso peptide subclass, Microcin J25, confirmed that the function of the leader is a recognition element that docks the precursor in one or both of bacteriocin maturation proteins. But it also suggested that the second to a last amino acid in the in C-terminus of the leader may bind to a shape-selective pocket within maturation proteins. It is assumed that the penultimate Thr residue in the leader of Microcin J25 may be the sole or at least the primary recognition element in lasso peptide leaders (Pan et al., 2012).

3.4 Mode of actions

Bacteriocins have diverse mechanisms of action: some of them act at the cell envelope and other act within the cell (**Fig. 6**). Lantibiotics and some class II bacteriocins target lipid II, which is one of the key components in biosynthesis of the bacterial cell envelope – peptidoglycan. Other bacteriocins form pores in the cell membrane, which results in a loss of membrane potential and cell death. Bacteriocins that have antibacterial activity against gram-negative bacteria usually

first recognize receptors localized in the outer membrane and then pass through the outer membrane. Later, bacteriocins bind to other receptor in the inner membrane, which transports them inside the cell. Then, they can kill the cell by interfering with DNA, RNA or protein metabolism (Cotter et al., 2013). A group of class III bacteriocins (*bacteriolysins*) function through the cell-wall hydrolysis resulting in lysis of sensitive cells (Cotter et al., 2005).

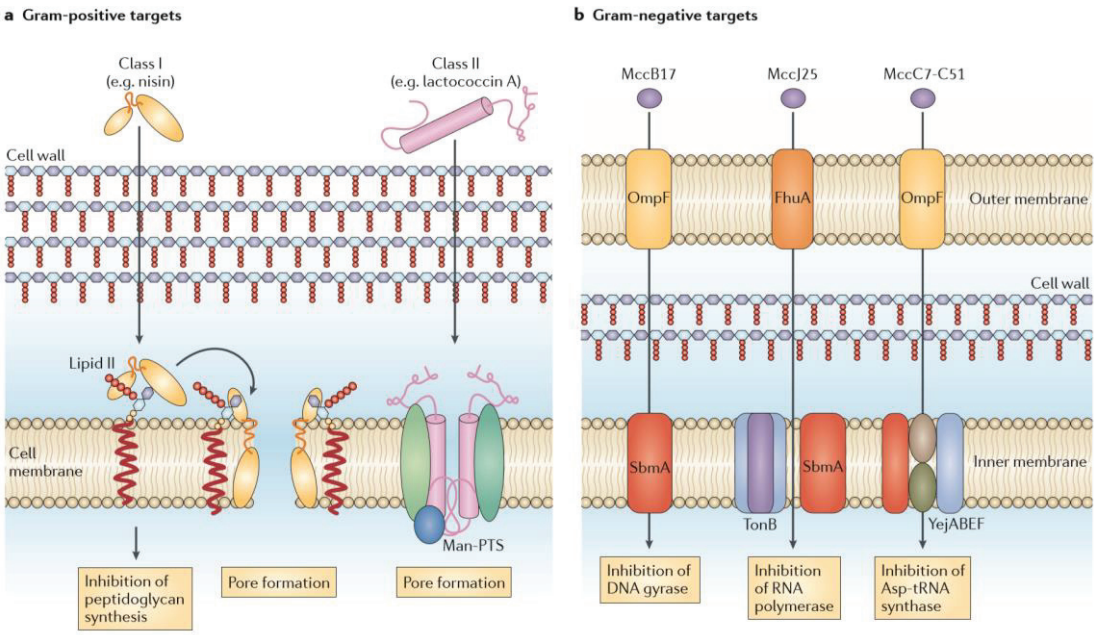


Figure 6. Different mechanisms of action of bacteriocins. (**a**) Bacteriocins that act against gram-positive bacteria, usually target the cell envelope. (**b**) Other bacteriocins that act against gram-negative bacteria usually target them by interfering with DNA, RNA and protein metabolism (Cotter et al., 2013).

Bacteriocin producing cells are protected from their own bacteriocin by the expression of an immunity protein. In general, the immunity protein acts specifically toward its related bacteriocin. It is assumed that bacteriocin gets locked onto the receptor by its immunity protein by forming a complex, which is settled only when the bacteriocin acts from the outside. In this way, the immunity protein might interrupt the subsequent steps that lead to the cell death (Ríos Colombo et al., 2017).

Best-studied lantibiotic nisin and some other lantibiotics form complex between the lanthionine rings (A and B) and the pyrophosphate of lipid II (**Fig. 7**). The effect

depends on the combination of the length of the peptide and the thickness of the lipid bilayer. Nisin-like bacteriocins kill bacteria primarily by forming pores together with the lipid II (**Fig. 6**). Other bacteriocins that are short, cannot span a membrane bilayer to form pores. They have a pyrophosphate cage for interaction with the lipid II and removes it from its functional location subsequently inhibiting the biosynthesis of peptidoglycan (Hsu et al., 2004; Hasper et al., 2006).

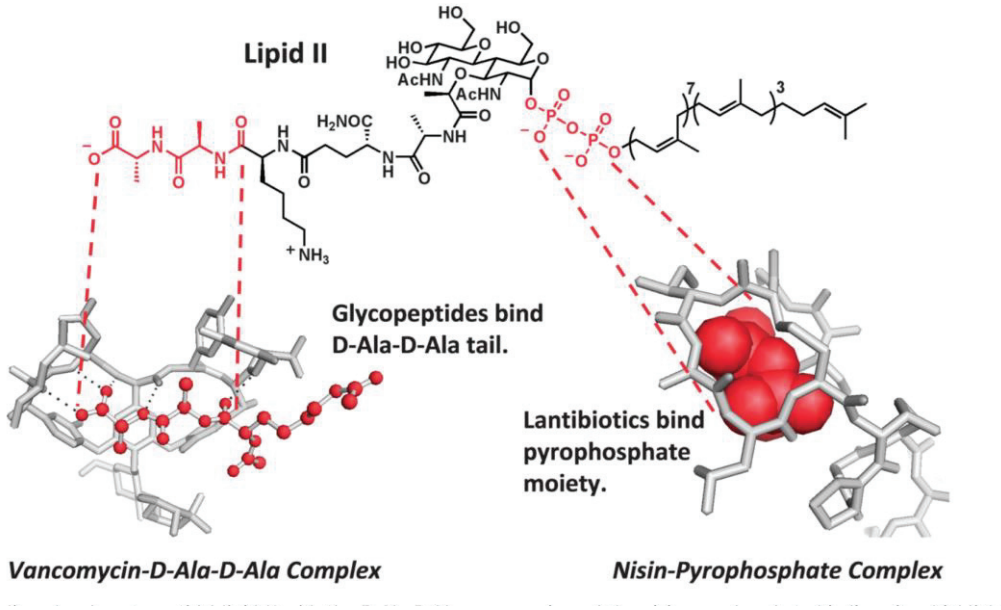


Figure 7. Vancomycin (NRP) and nisin (RiPP) binding to the lipid II. Structure of disaccharyl-pentapeptidyl-lipid II with the D-Ala-D-Ala-vancomycin and the nisin-pyrophosphate binding sites highlighted (Walsh and Wencewicz 2014).

Nisin exhibits antibacterial activity at very low concentrations. Its minimum inhibitory concentration (MIC) for *Corynebacterium* spp. GH17 and *L. lactis* subsp. *cremoris* HP is 39 nM, for *S. salivarius* K12 and NU10 strains – 79 nM (Barbour et al., 2016).

A two-component bacteriocin Thuricin CD belongs to the newly designated sactibiotic subclass. Its precise mechanism of action has not been elucidated. There are studies which show that both peptides (Trnα and Trnβ) are membrane-acting and cause a collapse of the membrane potential of target cells. It has been suggested that peptides of thuricin CD insert into the membrane and form pores, which leads

to the permeabilization, the flux of ions across the membrane and its depolarization, which eventually results in the cell death. Thuricin CD has a narrow spectrum antibacterial activity and a specific receptor may be present in only a relatively small number of sensitive target species. Because of that it is likely that lipid II is not a receptor that bounds the peptides and the receptor is absent in non-target bacteria species (Mathur et al., 2017). There are determined MIC values of thuricin CD, for *C. difficile* strains they are > 125 nM (Mathur et al., 2016), for *L. monocytogenes* – 2.5 μM (Mathur et al., 2014).

Enterocin AS-48, a circular bacteriocin produced by *E. faecalis* strains, is active against most of tested gram-positive bacteria. Gram-negative bacteria are also inhibited by this bacteriocin. Nevertheless, gram-negative bacteria is less sensitive to enterocin AS-48 than gram-positive bacteria. It is explained by the protective effect of outer membrane of the bacteria. The bacteriocin inserts into the membrane and causes its permeabilization, what leads to the cell death. It have been shown that enterocin AS-48 tends to form aggregates and dimers in aqueous solutions. Studies suggest that the bacteriocin can change molecular dimer association stages from the water-soluble to the membrane-bound at the membrane surface (**Fig. 8**). It could permit enterocin AS-48 molecules to insert into the bacterial cytoplasmic membrane. Moreover, it is assumed that enterocin AS-48 does not need a receptor in the membrane (Grande-Burgos et al., 2014).

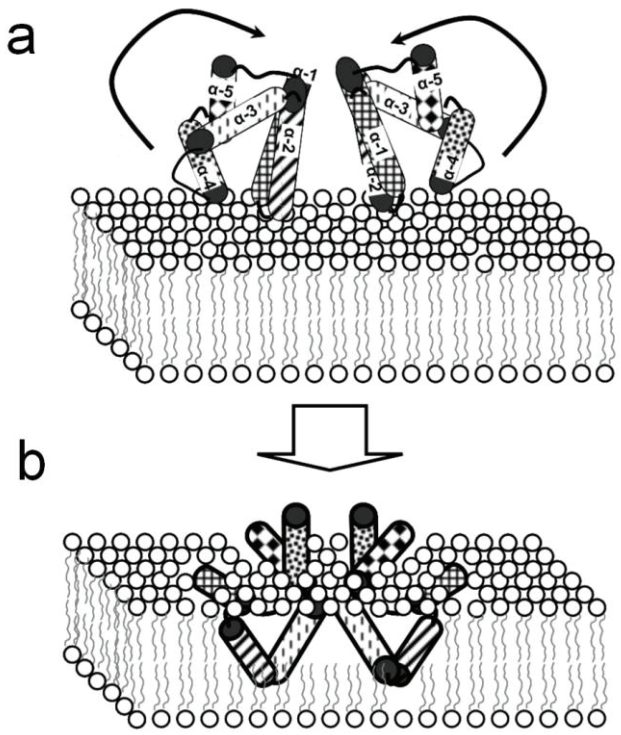


Figure 8. Proposed mechanism for transition of enterocin AS-48. (**a**) water-soluble stage (**b**) membrane-bound stage (Grande-Burgos 2014).

Garvicin Q, GarQ, is an antibacterial peptide of non-pediocin-like and single-peptide bacteriocins subclass produced by *L. garvieae*. It has a wide antimicrobial spectrum against some species from *Bacillus*, *Enterococcus*, *Lactobacillus*, *Lactococcus* and *Pediococcus* genera (Tosukhowong et al., 2012). Studies have shown that GarQ uses Man-PTS subunits IID and IIC as a receptor. It is assumed that the N-terminal part of the bacteriocin initially interacts with certain amino acids of the Man-PTS IID subunit. This results in the interaction of the C-terminal part of the bacteriocin with Man-PTS IIC subunit. The interaction between GarQ and the receptor might result in a structural changes in Man-PTS. It may lead the IIC subunit (permease) to open as a pore, which might result to the leakage of solutes across the membrane and disruption of the membrane integrity. Man-PTS is also known as a receptor for pediocin-like and lactococcin-like (**Fig. 6**, member of the non-pediocin-like/single-peptide subclass) bacteriocins. It was shown that despite recognizing the same receptor on the same host cells, GarQ and lactococcin-like group bacteriocins,

differ in their mode of interaction with the Man-PTS. This difference may involve specific interactions between the bacteriocin and individual amino acids located in the Man-PTS subunits IIC and IID. There were determined MIC values of GarQ for *L. garvieae* which is > 18.36 nM, for *L. lactis* > 4.5 nM (Tymoszewska et al., 2017). Microcin B17, MccB17, is a post-translationally modified antibacterial peptide produced by *E. coli.* It inhibits growth, a rapid decline in DNA replication, and induction of SOS response in sensitive *E. coli* (Heddle et al., 2001). An outer membrane porin protein OmpF is important for the translocation of microcin B17 across the outer membrane and protein SbmA – across the inner membrane into the bacterial cells (**Fig. 6**) (Laviña et al., 1986). The primary target of MccB17 is gyrase (type II DNA topoisomerase) that introduces supercoils into the DNA. Bacteriocin induces gyrase-dependent DNA cleavage (Heddle et al., 2001). Determined MIC value of MccB17 for *E. coli* BL21(DE3) is 60 nM (Shkundina et al., 2014).

Another an antibacterial peptide microcin J25 (MccJ25) with an unusual lasso structure is produced by *E. coli* (Bayro et al., 2003). MccJ25 enters sensitive cells across the outer membrane receptor FhuA, and then it passes the inner membrane through SbmA protein (**Fig. 6**) (Salomón et al., 1995). The bacteriocin inhibits transcription by binding to the nucleotide uptake channel of bacterial RNA polymerase (Bayro et al., 2003).

It is assumed that members of glycocin subclass have different mode of actions. Glycocin F, ASM1 and enterocin F4-9 are bacteriostatic and their antibacterial activity depends on the presence of the sugar in the peptide. Meanwhile sublancin 168 and thurandacin A are bactericidal and glycosylation of the peptide is not necessary for the activity. It was observed that a GlcNAc-specific PTS transporter confers sensitivity to glycocin F. The mechanism(s) by which glycocin F and GlcNAc transporter complexes might induce a rapid and sustained bacteriostasis are not known. In case of sublancin 168 it was shown that large mechanosensitive channel MscL may serve as a gate of entry for the bacteriocin to the cell (Kouwen et al., 2009). It was indicated that PTS is affecting the sensitivity to sublancin 168.

Specifically, cytoplasmic proteins HPr and glucose transporter PtsG are implicating in sensitivity to sublancin 168. HPr transfers phosphate group to PtsG (permease), which subsequently phosphorylates incoming sugar. Point mutation in HPr protein, which prevents binding and transfer of the phosphate group, increases the resistance of the sensitive strain. There have been speculations that sublancin 168 may be phosphorylated upon its entry into the cell (Garcia De Gonzalo et al., 2015). Determined MIC values of sublancin 168 for methicillin-resistant *Staphylococcus aureus* is 15 μM (Wang et al., 2017) for *B. megaterium* and *B. subtilis* – 1.29 μM (Paik et al., 1998), for *B. cereus* ATCC 10987 – 0.88 μM (Ji et al., 2015), for *E. faecalis* ATCC 29212 – 1.8 μM, for *S. aureus* ATCC 25923 – 1.75 μM, for *S. agalactiae* ATCC 27956 – 0.54 μM and for *S. pyogenes* ATCC 19615 – 0.2 μM (Ji et al., 2015). Determined MIC values of enterocin F4-9 for *E. mundtii* JCM 8731 – 0.94 μM, for *E. faecalis* JCM 5803, *B. coagulans* JCM 2257, *E. coli* JM109 – 1.87 μM, for *E. durans* NBRC 100479 – 15 μM (Maky et al., 2015).

Some class III bacteriocins, like thuricin, megacin A-216 and megacin A-19213, produced by *Bacillus* sp. bacteria exhibit phospholipase A activity (Favret and Yousten 1989; Von Tersch and Carlton 1989; Kiss et al., 2008). In case of megacin A-216, observations indicated that it acts by impairing the cell membrane integrity of a sensitive strain. It is assumed that this bacteriocin damages the membrane permeability barrier (Kiss et al., 2008). Studies on dysgalacticin propose that it exerts the bactericidal activity by binding to the glucose- and/or mannose-PTS permease resulting in the inhibition of glucose transport. Results of the research on dysgalacticin suggest that it kills target cells by: inhibiting glucose uptake by putative binding to the glucose-phosphotransferase system (PTS) and/or mannose-PTS; and causing a loss of K^+ ions and dissipation of the membrane potential (Swe et al., 2009). Another high molecular weight bacteriocin zoocin A has a narrow activity spectrum. It functions enzymatically as an endopeptidase and cleaves peptidoglycan cross-links of the cell (Simmonds et al., 1996). Class III bacteriocin enterolysin A is active against a wide range of gram-positive bacteria. It degrades

the cell wall and kills the cells by solubilization of peptidoglycan units of cell wall. In contrast to zoocin A, it cleaves the interpeptide bonds within the peptidoglycan units at two locations (Khan et al., 2013).

3.5 Genome mining for bacteriocins

The availability of genome sequences facilitates identification of genes coding for known or novel putative precursors of bacteriocins and other natural antibacterial products. Moreover, genome mining may reveal their biosynthetic pathways. Bacteriocins with post-translational modifications are well suited for genome mining. The gene of bacteriocin precursor may be found directly in the genome if the sequence of the peptide has been revealed previously. Genes encoding bacteriocin precursor peptides may be used as queries for searching in genomes, resulting in analogs of known compounds. Meanwhile, genes encoding the key enzymes for bacteriocin biosynthesis are also used, it may result in more distant analogs of known products In addition, their short biosynthetic pathways make them suitable for heterologous expression. Genome mining allows identification of new compounds and their biosynthetic genes encoded by the "silent or cryptic clusters" when they are not produced under laboratory conditions by the native producer (Arnison et al., 2013).

Unfortunately, the genome mining approach typically does not provide any information if the peptide has antibacterial activity. Moreover, compounds with a new structural scaffolds are difficult to discover, because neither their biosynthetic genes nor their precursor genes are known. There may be cases when conclusions from bioinformatics alone about biosynthesis systems of bacteriocins are not completely reliable. In particular, sequence similarity of precursor peptides does not guarantee that the final products belong to the same family of bacteriocins. Proximity between genes for precursor and modification enzyme does not guarantee a target/substrate relationship. The genes encoding a precursor peptide, its maturation enzyme(s), and its export transporter can all be far apart in a genome.

Despite these limitations, genome mining is becoming a strategy for screening and identification of new compounds with exciting bioactivities (Arnison et al., 2013). For novel bacteriocins identification and their characterization or classification genome-mining and annotation tools have been developed. These tools are free-access web-based BAGEL3 (van Heel et al., 2013) and anti-SMASH (Medema et al., 2011) tools which are capable of identifying bacteriocin biosynthetic gene cluster or other known secondary metabolite compound. Moreover, there is available bacteriocin repository BACTIBASE (Hammami et al., 2010).

BAGEL3 (http://bagel.molgenrug.nl) enables the identification of gene clusters for bacteriocin biosynthesis through a knowledge-based database. It takes advantage that accessory genes encoding proteins needed for the processing, modification, transport, regulation and/or immunity are commonly located in near a putative bacteriocin gene. Small open reading frames (ORFs) are often omitted during automated annotation efforts, especially when their product sequences do not show strong homology with those of already described peptides, hampering a direct mining approach. It uses DNA sequences as input instead of annotated genomes, making it less dependent on the ORFs predictions. BAGEL3 uses an identification approach that combines direct mining for the gene and indirect mining via context genes. A typical BAGEL3 search on a genome sequence results in a set of putative bacteriocin gene clusters. These are ranked according to the presence of significant features in the amino acid sequences and their genomic context (van Heel et al., 2013).

A comprehensive antiSMASH (http://antismash.secondarymetabolites.or) pipeline is capable of identifying biosynthetic gene clusters covering the whole range of known secondary metabolite compound classes: polyketides, non-ribosomal peptides, terpenes, aminoglycosides, aminocoumarins, indolocarbazoles, bacteriocins, nucleosides, beta-lactams, butyrolactones, siderophores, melanins and others. It aligns the identified gene clusters to their nearest relatives from a database containing all other known gene clusters, and integrates or cross-links all previously

available secondary-metabolite specific gene analysis methods in one interactive view (Medema et al., 2011).

3.6 Bacteriocins of thermophilic bacteria

Researchers are using term bacteriocin-like inhibitory substances (BLISs) when the sequence of bacteriocin is not revealed. The first BLISs produced by various thermophilic bacteria *G. stearothermophilus* strains (formerly known as *B. stearothermophilus*) were named thermocins and described by Shafia in 1966. Later on, Yule and Barridge in 1976 investigated one of them – thermocin 10 produced by *G. stearothermophilus* NU-10 strain. By gel filtration it was determined that the compound was about 20 kDa. Chemical composition analysis showed its proteinaceous nature with small quantities of associated carbohydrate. In 1983 Fikes and others tried to investigate thermocin 10 in more detail and reported that it inhibits RNA and DNA synthesis, cell motility, oxygen uptake and ATP synthesis. Sharp and others in 1979 described thermocin 93 produced by *G. stearothermophilus* RS93. They partially purified it from supernatant and by gel filtration indicated that it is about 13.5 kDa compound with antibacterial activity only against other closely related thermophilic bacteria. Both thermocins were heat-stable and their activity decreased after treatment with proteases, confirming their proteinaceous nature.

Another study was done on BLIS produced by *G. thermoleovorans*. Strains S-II and NR-9 isolated from mud and water samples in the United States produced thermoleovorin-S2 and thermoleovorin-N9, respectively. Partially purified from the culture supernatant BLISs were examined by gel chromatography. It was determined that thermoleovorin-S2 is 42 kDa and thermoleovorin-N9 – 36 kDa compound. Inhibitory activity of both thermoleovorins was decreased by proteases. They were stable at 60°C and only thermoleovorin-S2 was more stable at 70°C and 80°C temperature (Novotny and Perry 1992).

Geobacillus sp. HBB-247 which was isolated from a thermal pool in Turkey secreted BLIS which was partially purified for characterization. It was estimated that it is stable up to 60°C, sensitive to proteolytic enzymes and effective against *E. faecalis*, *E. avium, Listeria* sp., *C. pasteurianum*, *Cellulomonas fimi* and some thermophilic *Geobacillus* spp. and *Anoxybacillus* spp. strains. Tricine-SDS-PAGE analysis revealed that the molecular weight of the BLIS was about 38 kDa (Başbülbül Özdemir and Biyik 2012). Another one *Geobacillus* sp. HBB-218 isolate from soil collected around the thermal spring produced BLIS which was purified by chromatography methods and named toebicin 218. It was sensitive to some proteases, stable at wide range of pH values and heat-stable. Tricine-SDS-PAGE analysis identified it as 5.5 kDa compound. It showed inhibition of a wide range of gram-positive bacteria including some clinically important *E. faecalis* and *Micrococcus luteus*. Moreover, it was more active against thermophilic *Geobacillus* spp. and *Anoxybacillus* spp. bacteria (Başbülbül Özdemir and Biyik 2012).

Geobacillus sp. ZGt-1, isolated from a hot spring in Jordan, produced BLIS of molecular mass 15-20 kDa. The antibacterial activity was lost after treatment with proteases and it was stable after heating up to 70°C (Alkhalili et al., 2016).

G. stearothermophilus 17, 30, 31 and 32A strain isolated from oil wells in Lithuania produced BLISs that were characterized. Partially purified antibacterial compounds were analyzed by SDS-PAGE and their molecular masses were revealed: 6.9 kDa, 5.6 kDa, 7.1 kDa and 7.2 kDa, respectively. They were sensitive to proteolytic enzymes and heat-stable. BLISs produced by strains 17, 30, and 32A had disulfide bonds which were necessary for antibacterial activity (Pokusaeva et al., 2009).

It is known that BLISs are also produced by thermophilic *B. licheniformis* bacteria. Proteinaceous nature antibacterial agents produced by stains 490/5, H1, T6-5 and P40 were partially purified and their molecular masses of 2, 3.5 kDa, 20 kDa and 105 kDa were determined, respectively (Cladera-Olivera et al., 2004; Korenblum et al., 2005; Abdel-Mohsein et al., 2011; Martirani et al., 2012).

Taking into account that bacteria can produce more than one different antibacterial agents and that these BLISs were not completely purified, these data should be evaluated critically.

The availability of sequenced bacterial genomes prompted the search of new gene-encoded peptide natural products based on a gene discovery approaches. The attention was drawn to *G. thermodenitrificans* NG80-2 which was isolated from an oil fields in China. Sequenced genome of this strain revealed two gene clusters encoding putative lantibiotic biosynthetic machineries (**Fig. 9**).

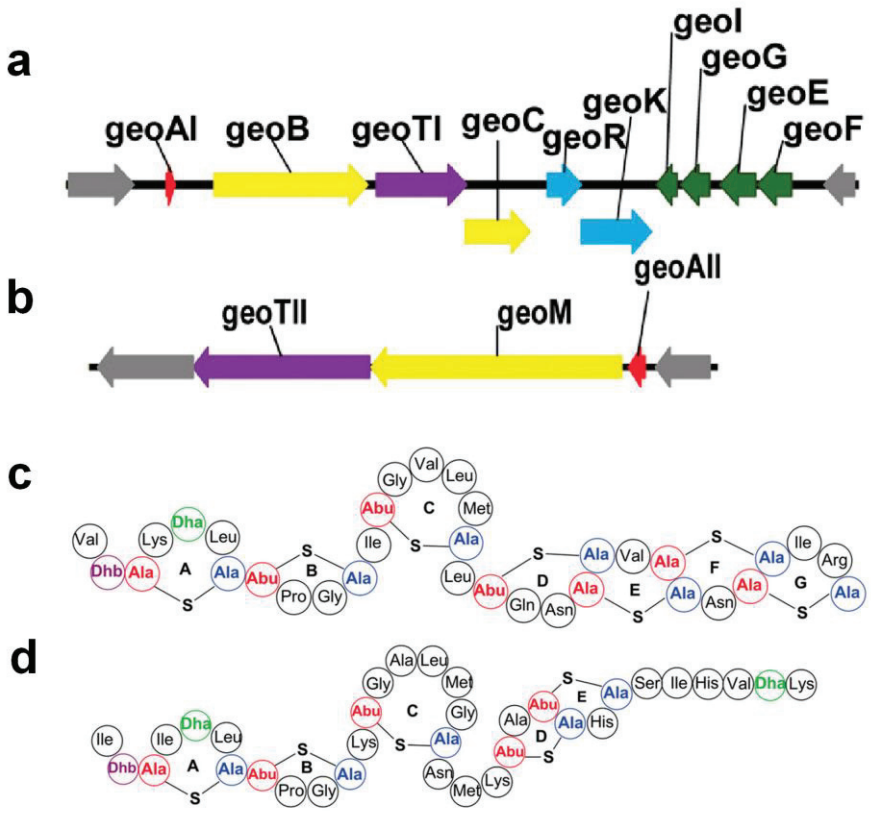


Figure 9. Gene clusters of (**a**) geobacillin I and (**b**) geobacillin II biosynthetic machineries. The gene in red is for the precursor peptide, in yellow for modification enzymes, in purple – transporter/ protease, in blue – regulatory proteins, in green – immunity proteins. Structures of (**c**) geobacillin I and (**d**) nisin with ring topology (Garg et al., 2012).

One of them was coding novel class I lantibiotic termed geobacillin I. It was obtained by heterologous expression in *E. coli* and subsequent purification. Characterization of the bacteriocin revealed that it is an analog of nisin. The

antimicrobial spectrum of geobacillin I was generally similar to that of nisin A, with increased activity against *S. dysgalactiae*. Furthermore, the compound was more stable than nisin A at pH 7 and 8 and at high temperatures. In addition to geobacillin I, the genome of *G. thermodenitrificans* NG80-2 also contained a class II lantibiotic biosynthetic gene cluster. The corresponding compound was produced in *E. coli*, and had a ring topology different than that of any known lantibiotic. Interestingly, geobacillin II demonstrated antimicrobial activity only against *Bacillus* spp. strains. Both geobacillins were smaller than 5 kDa peptides. Seven *Geobacillus* spp. strains were screened for production of geobacillins using whole-cell MALDI-TOF mass spectrometry and five of them produced geobacillin I, but none – geobacillin II (Garg et al., 2012). To date geobacillins are the best studied and characterized bacteriocins from the thermophilic bacteria. Geobacillin I mutants were synthesized to determine mode of action and to compare it to nisin. It was revealed that geobacillin I binds to lipid II and forms pores in the membrane as nisin, but detailed mechanism of pore formation may differ from that of nisin (Garg et al., 2014). Factors that influence stereochemistry and enzymatic activity of geobacillin II modification enzyme, GeoM, were investigated by synthesis and co-expression of sequence modified precursors and maturation enzyme GeoM (Garg et al., 2016).

3.7 Application of bacteriocins

Pathogens have emerged that are resistant to a single, and subsequently multiple, antibiotics. Moreover, there is a shortage of new families of antibiotics that could potentially compensate for resistance to existing antibiotics. As a consequence, there is a need for the development of new antimicrobials that can be used in clinical settings. Alternatives that have been investigated include plant derived compounds, bacteriophages and phage lysins, RNA-based therapeutics, probiotics, and antimicrobial peptides from a variety of sources (Cotter et al., 2013).

Bacteriocins have many properties which suggest them as viable alternatives to antibiotics. It includes their potency (as determined *in vitro* and *in vivo*), low

toxicity, the availability of both broad- and narrow-spectrum peptides, the possibility of *in situ* production by probiotics and fact that these peptides can be bioengineered (Cotter et al., 2013).

However, to date, the primary focus for the use of these bacteriocins has been on animal, rather than human, health. Existing commercial examples include the use of thiostreptonin to treat dermatological indications in domestic animals and the use of nisin as the active agent in the mastitis prevention product Wipe Out (ImmuCell Corporation) (Cotter et al., 2013). Nisin is suggested as a dietary supplement for broiler chickens. Studies on nisin demonstrated that it exerts a modulating effect on the microbial ecology of the gastrointestinal tract in broiler chickens and improves their performance significantly (Józefiak et al., 2013).

A lack of sufficient investment has been a significant problem with respect to the medical application of bacteriocins. However, there are several bacteriocins developed for human applications. These bacteriocins include many thiopeptides, such as LFF571, as well as Mu1140-S – a synthetic form of the lantibiotic mutacin1140 (Oragenics) and NVB302 – a semi-synthetic derivative of the lantibiotic deoxyactagardine B. In addition to these synthetic bacteriocins, the natural producer of mutacin1140 has been engineered for use in preventing tooth decay (SMaRT Replacement Therapy). Microbisporicin is being developed as an injectable therapy to control multidrug-resistant gram-positive pathogens (Cotter et al., 2013).

Bacteriocins could help to avoid spore forming thermophilic bacteria contamination in some industry processes. In diary food production mesophilic *Bacillus* sp., thermophilic *Geobacillus* spp. or *A. pallidus* species are able to survive industrial pasteurization and form biofilms within pipes of stainless steel equipment (Gopal et al., 2015). Recently, thermophilic bacteria were shown to have great potential for application in biofuel production (Ahmad et al., 2017; Jiang et al., 2017). In these industrial processes contamination of other bacteria is undesirable, and bacteriocins as heat-resistant natural products could be used for biocontrol.

Nisin A containing products such as 'Nisaplin' (Danisco) are commercially available and can be used as food preservatives. It has been shown that combinations of Nisaplin-citric acid/cinnamaldehyde could significantly reduce and eradicate *L. monocytogenes* biofilms. Moreover, nisin A could be substituted with bioengineered and more potent its version M21A. Combinations of these antimicrobials allows the use of lower concentrations of the individual additives in the food. Such combinations could enhance the microbial safety of foods and provide consumers with natural alternative to artificial preservatives (Smith et al., 2016).

Studies on enterocin CRL35 produced by *E. faecium* have shown its antiviral properties. It was reported that bacteriocin inhibited viral multiplication of Herpes simplex (HSV) type 1 and 2. It inhibits late stages of replication *in vitro* (Wachsman et al., 2003).

It has been shown that bacteriocins may have anticancer activity. It is assumed that they might interact with a membrane of a cancer cell. It is predominantly negatively charged in contrast to the normal membranes of mammalian cells. This difference could facilitate selective interaction with membranes of cancer cells. Moreover it is assumed that they have higher membrane fluidity as compared to normal cells and this could also facilitate destabilization of the membrane (Kaur and Kaur 2015).

Some studies on bacteriocins have been projected on their use in veterinary and agriculture. Bacteriocin Bac IH7 produced by *B. subtilis* can promote a plant growth. It enhanced tomato and muskmelon germination, increased shoot weight, height and root lengths. It may also serve as a biocontrol for seed borne pathogen *Alternaria solani* (Hammami et al., 2011).

Another bacteriocin thuricin 17 promotes plant growth. It functions as a bacterial signal compound, enhancing plant growth in legumes and non-legumes. In *Arabidopsis thaliana* and *Glycine max* bacteriocin increased level of phytohormones (Subramanian and Smith 2015).

4 Methods

4.1 Growth media and cultivation conditions for thermophilic bacteria

Nutrient broth (NB) medium (1% tryptone, 0.5% meat extract, 0.5% NaCl) was used as a routine cultivation medium for thermophilic bacteria, unless otherwise stated. NB supplemented with glycerol (20% (v/w)) was used as the stock culture medium for the thermophilic bacteria strains and stored at -70 or -80°C temperature. The stock cultures were always refreshed and harvested on NB agar (1.5% (w/v)) medium in an incubator at 55-60°C temperature overnight. For a bacteriocin production and extraction, strain 15 was grown in NB media supplemented with salts: KCl – 0.07 μmol/L and $MgCl_2$ – 1.1 μmol/L.

4.2 Phenotypic analysis of strain 15

Gram staining and KOH test were performed as described by Chandra et al. (2011) with some modifications. Safranin was used instead of carbol-fuchsin in Gram-staining. Formation of endo-spores of the strain was determined by Schaeffer-Fulton staining technique (Mormak and Casida, 1985).

4.3 16S rDNA sequencing

DNA sequence coding for 16S rRNA was amplified by PCR method described bellow. Primers 27F, 1492R (**Supplementary Table 1**) and genomic DNA of strains 8 or 15 were included in the PCR mixtures. The resulting DNA products were sent for sequencing. The obtained sequences were analyzed by BLAST method in NCBI database.

4.4 Determination of antibacterial activity spectrum of bacteria strains 8 and 15

The antagonistic interactions of the strains 8 and 15 were tested against other closely related thermophilic bacteria by agar overlay method. The tested strains were grown in the middle of Petri dish with NB agar medium overnight at 60˚C to produce inhibition substances. After incubation, melted NB agar medium with inoculated other strain (final OD (600 nm) ~0.04) was poured on the top of the grown culture. Plates were incubated at 60˚C for 12 h.

4.5 Genomic DNA sequencing and bioinformatics analysis

Genomic DNA of bacteria strains 8 and 15 were submitted for the whole genome sequencing service in the GATC Biotech (Germany). The isolated DNA was sheared to 500-bp fragments for the next generation sequencing libraries. The libraries were 250-base paired-end sequenced on an Illumina HiSeq 2000. Obtained raw sequencing data was submitted to the Velvet server for a de novo paired-end assembly on each genome, resulting in the draft genome sequences in contigs. The RAST server was subsequently used to annotate the genomes. The genome sequences of strains 8 and 15 have been deposited into the NCBI database in https://www.ncbi.nlm.nih.gov/genome under the accession numbers LVHY00000000 and LVHZ00000000, respectively.

Identification of strains was performed by the Genome-to-Genome Digital Calculator (GGDC) for a digital DNA-DNA hybridization (dDDH) analysis in http://ggdc.dsmz.de/home.php.

For genome analysis of strain 15 towards other *Geobacillus* spp*., Parageobacillus* spp. and *Caldibacillus* sp. genomes was compared. The genomes of *G. stearothermophilus* ATCC 12980^T (JYNW00000000.1), *G. kaustophilus* NBRC 102445^T (BBJV00000000.1), *G. thermoleovorans* KCTC 3570^T (CP014335.1), *G. thermodenitrificans* KCTC3902^T (CP017694.1), *G. subterraneus* KCTC 3922^T

(CP014342.1), *G. thermocatenulatus* KCTC 3921^T (CP018058.1), *G. jurassicus* NBRC 107829^T (BCQG01000000), *G. zalihae* NBRC 101842^T (BCPV01000000), *G. thermoleovorans* DSM 15325 (CP017692.1), *G. galactosidasius* DSM 18751^T (NDYL01000000), *G. uzenensis* BGSC 92A1^T (NEWL01000000), *G. icigianus* DSM 28325^T (JPYA01000000), *G. genomospecies* 1 Et2/3 (JYCF00000000), *G. genomospecies* 1 Et7/4 (JYBP01000000), *G. genomospecies* 2 PSS1 (JPOI00000000), *G. genomospecies* 3 JF8 (CP006254.2), *P. thermoglucosidasius* DSM 2542^T (CP012712.1), *P. toebii* NBRC 107807^T (BDAQ01000000), *P. caldoxylosilyticus* NBRC 107762^T (BAWO01000000), *P. thermantarcticus* DSM 9572^T (FOJS01000001), *P. genomospecies* 1 NUB3621 (CM002692.1) and *C. debilis* DSM 16016^T (ARVR01000000) were used for comparison.

For the analysis genome of strain 8 towards other *Aeribacillus* genomes was compared. The genomes of *A. pallidus* KCTC3564^T (CP017703.1), *A. pallidus* 8m3 (LWBR00000000) and *A. pallidus* GS3372 (JYCD00000000) were used for comparison.

Genomic DNA was submitted for a bacteriocin mining to the BAGEL4 server in http://bagel4.molgenrug.nl/. For a BLAST analysis peptide or protein sequences were submitted to the NCBI database in https://blast.ncbi.nlm.nih.gov/Blast.cgi.

The presence and location of signal peptide cleavage sites in amino acid sequences was predicted by SignalP 4.1 server in http://www.cbs.dtu.dk/services/SignalP. A secondary structure of peptides was predicted using web tool PSIPRED in http://bioinf.cs.ucl.ac.uk/psipred.

4.6 Bacteriocin activity assays

A colony of sensitive strain *P. genomospecies* 1 NUB36187 (BGSC 9A11) was spread on NB agar plate and incubated overnight at 60°C. Next morning, using an inoculum loop, all biomass from the plate was spread on a fresh NB agar plate and incubated for 4 hours at 60°C. After the incubation (before cells starts to form spores), all biomass from the plate was washed with NB medium and the cell

suspension adjusted to OD (600 nm) of 1. It was inoculated to the melted NB agar medium (55°C) in ratio 1:100 and thoroughly mixed. 15 mL of the resulting cell suspension was dispersed in a Petri plate and left to solidify. For a well diffusion assay, wells were cut out in the solid medium in the Petri plate and filled with 50 µL of serial twofold dilutions of samples. Thus, the arbitrary unit of antibacterial activity per milliliter (AU/mL) was defined as $2^n \times 1$ mL / V. V defines sample volume (mL). 2^n defines the titer of the reciprocal highest dilution that resulted in inhibition of the indicator lawn. (Daba et al., 1991).

For a spot on lawn assay, samples of 10 µL were spotted on the solid medium in the Petri dish. Later on, the plate was incubated at 60°C (37°C for *Bacillus* spp.) overnight. *P. genomospecies* 1 NUB36187 (BGSC 9A11) was used as sensitive strain for all routine experiments.

4.7 DNA amplification by PCR

A single PCR mix included Phusion HF Buffer (Thermo Scientific), dNTPs (0.5 mM each), $MgCl_2$ (1.5 mM), *PfuX7* DNA polymerase (homemade), primers (0.5 µM each) and DNA template (1 ng/µL). Target DNA was PCR amplified by 30 cycles of denaturing (94°C for 30 sec), annealing (5°C lower then T_m, for 30 sec), and extending (68°C for 1 min per 1 kbp). Amplifications were confirmed by 1% or 2% agarose gel electrophoretic analyses. Primers used in this work are listed in **Supplementary Table 1**.

4.8 DNA restriction, cloning, ligation, transformation and plasmid isolation

Strain 8 was grown in a Bacto Brain Heart Infusion medium (BD Diagnostics) and strain 15 was grown in NB medium at 55°C, 250 RPM. The genomic DNA was extracted with the GenElute Kit (Sigma-Aldrich) according to the manufacturer's recommendations. DNA digestion was performed with restriction endonucleases purchased from Thermo Scientific and according to the manufacturer's

recommendations. Amplified or digested DNA was always cleaned or extracted from agarose gel with NucleoSpin Gel and PCR Clean-up Extraction Kit (Macherey-Nagel), unless stated otherwise. The cleaned DNA insert and vector were ligated with T4 DNA Ligase (Thermo Scientific), unless stated otherwise, according to the manufacturer's recommendations. The ligation products were transformed to *E. coli* TOP10 (Invitrogen) cells by electroporation. Cells were plated on Lysogeny Broth (LB) agar plates with appropriate antibiotics and grown at 37°C overnight. Several colonies were picked and tested by colony PCR, to confirm if the insert is present in the vector. Colonies with the correct insert were inoculated to LB medium with the appropriate antibiotic. The cultures were grown at 37°C overnight, and plasmids were isolated using NucleoSpin Plasmid Extraction Kit (Macherey-Nagel). For gene expression, plasmid DNA was transformed to *E. coli* BL21(DE3) (Invitrogen) by electroporation. Cells were plated on LB agar plates with appropriate antibiotics and grown at 37°C overnight. Several colonies were picked and inoculated to LB medium with the appropriate antibiotic. The cultures were grown at 37°C overnight, mixed with glycerol to the end concentration of 20% and stored at -80°C for further use.

4.9 Molecular cloning of pallidocin biosynthetic gene cluster

The gene cluster encoding the pallidocin biosynthetic machinery *pal* was amplified by PCR as a single unit using F-PalA-USER and R-PalA-USER (**Supplementary Table 1**) as primers and *A. pallidus* 8 genomic DNA as template. The pBAD24 vector was PCR amplified using F-pBAD-USER and R-pBAD-USER (**Supplementary Table 1**) as primers and pBAD24 plasmid as template. The resulting DNA products were ligated by ligase-free UDG-mediated cloning, using USER Enzyme (New England Biolabs) according to the manufacturer's protocol and directly transformed to *E. coli* TOP10 cells by electroporation. Propagated plasmid was isolated and the sequence of the insert in the plasmid was confirmed

by DNA sequencing. The insert *pal*, starting with a start codon for the PalA, was cloned into the MCS after the arabinose promoter and RBS of the pBAD24 vector.

4.10 Overexpression of the gene cluster of the pallidocin biosynthetic machinery and purification of secreted pallidocin

E. coli BL21(DE3) cells, transformed with the expression vector pBAD24 coding *pal*, was inoculated to 50 mL of LB-ampicillin medium. The culture was grown at 37 °C for 16 h and inoculated to 1 L of M9-ampicillin minimal medium ($Na_2HPO_4 \times 7H_2O$ – 12.8 g/L, KH_2PO_4 – 3 g/L, NaCl – 0.5 g/L, NH_4Cl – 1 g/L, $MgSO_4$ – 0.24 g/L, $CaCl_2$ – 0.11 g/L and 4% (V/V) glycerol in MiliQ water) in ratio 1:100. The cells were grown at 37°C to OD (600 nm) of 0.6-0.7. Arabinose was added to a final concentration of 2 mM and the culture was incubated at 37°C for 16 h. The synthesized pallidocin was secreted to the medium. Cells were harvested by centrifugation at 10,000×g for 15 min at 4°C. Supernatant was collected, immediately filtered through a 0.45 μm filter and uploaded on Econo-Column chromatography column, 2.5×30 cm (Bio-Rad) filled with 50 g of Amberlite XAD16N hydrophobic polyaromatic resin (Sigma-Aldrich) which was previously equilibrated with deionized water. After sample loading, the column was washed with 500 mL deionized water. Elution was performed with 250 mL of 100% methanol. The eluate was collected, diluted with deionized water (1:3 ratio) and lyophilized in a freeze-dryer. Pellets were dissolved in 100 mL of 50 mM lactic acid buffer (pH 4.5) and filtered through a 0.45 μm filter. Then, it was loaded on NGC system (Bio-Rad) equipped with HiTrap SP HP 5 mL cation exchange column (GE Healthcare Life Sciences), which was previously equilibrated with 50 mM lactic acid buffer (pH 4.5). The column was then washed with 50 mM lactic acid buffer (pH 4.5) and elution performed with 50 mM lactic acid buffer containing 300 mM NaCl (pH 4.5). The eluate was mixed with TFA to the end-concentration of 0.1% and loaded on RP-HPLC system (Agilent) equipped with Jupiter Proteo, C-12, 250×10 mm column (Phenomenex). Before the run, the column was equilibrated in

5% of solvent B (solvent A = MiliQ water with 0.1% TFA, solvent B = ACN with 0.1% TFA). Bacteriocin was eluted by an increase of solvent B up to 60% over 80 min with a flow rate of 2 mL/min. Elution fractions were tested for antibacterial activity against *P. genomospecies* 1 NUB36187 using a drop on a lawn assay. Active fractions were lyophilized, pellets dissolved in a solution containing 6 M guanidine-HCl and 0.1% TFA, then submitted for the 2nd round of purification on Jupiter Proteo, C-12, 250×4.6 mm column (Phenomenex). Before the run, the column was equilibrated in 5% of solvent B. Bacteriocin was eluted by an increase of solvent B up to 60% over 80 min with a flow rate of 1 mL/min. Elution fractions were tested for antibacterial activity against *P. genomospecies* 1 NUB36187 using a spot on lawn assay. Elution fractions with antibacterial activity were analyzed by MALDI-TOF MS, lyophilized to dryness and stored at −80°C until further use.

4.11 Molecular cloning of *palA-his*, *palS*, *palT* and *palST* genes

The gene *palA* was amplified by PCR using F-PalA-BspHI and R-PalA-HindIII (**Supplemenatry Table 1**) as primers and *A. pallidus* 8 genomic DNA as template. The gene *palS* was amplified by PCR using F-PalS-In-Fusion and R-PalS-In-Fusion (**Supplementary Table 1**) as primers and *A. pallidus* 8 genomic DNA as template. The gene *palT* was amplified by PCR using F-PalT-In-Fusion and R-PalT-In-Fusion (**Supplementary Table 1**) as primers and *A. pallidus* 8 genomic DNA as template. The DNA region containing two genes *palS* and *palT* was amplified by PCR using F-PalS-In-Fusion and R-PalT-In-Fusion (**Supplementary Table 1**) as primers and *A. pallidus* 8 genomic DNA as template.

DNA product *palA* was double digested with BspHI and HindIII, meanwhile a pRSFDuet-1 vector – with NcoI and HindIII. Cleaned DNA insert and vector were ligated by conventional cloning and transformed to *E. coli* TOP10 cells by electroporation. Propagated plasmid was isolated and the sequence of the insert in the construct (pRSFDuet-1-palA) was confirmed by DNA sequencing.

A site-directed mutagenesis approach was used to introduce His7-tag sequence (GGHHHHHHH) in the C-terminus of the PalA peptide. A new construct pRSFDuet-1 coding *palA-his* was generated by PCR amplification using phosphorylated primers (5'-end): F-PalA-His and R-PalA-His (**Supplementary Table 1**). Construct pRSFDuet-1-*palA* was used as template. The PCR product was cleaned, ligated and transformed to *E. coli* TOP10 cells. Propagated plasmid was isolated and the His7-tag coding sequence in the resulting construct (pRSFDuet-1-*palA-his*) was confirmed by DNA sequencing.

pBAD24 vector was double digested with NcoI and PstI, run on agarose electrophoresis and extracted from the gel. DNA products *palS*, *palT* and *palST* were ligated with double digested pBAD24 vector using Quick-Fusion Cloning Kit (Bimake) according to manufacturer's recommendations. After the ligation, mixtures were diluted with MiliQ in ratio 1:5 and transformed to *E. coli* TOP10 cells by electroporation. Propagated plasmids were isolated and the sequence of the insert in the resulting constructs (pBAD24-palS, pBAD24-palT and pBAD24-palST) were confirmed by DNA sequencing.

The gene *palA-his* was cloned in MCS after the IPTG promoter and RBS of the pRSFDuet-1 vector. Genes *palS/palT/palST* were cloned in MCS after the arabinose promoter and RBS of the pBAD24 vector.

4.12 *palA-his* gene co-expression with *palS*, *palT* or *palST*, and purification of synthesized peptides

E. coli BL21(DE3) cells with transformed vectors: pRSFDuet-1 coding *palA-his* and pBAD24 coding *palS/palT/palST,* was inoculated to LB-ampicillin-kanamycin medium and grown at 37°C for 16 h. The next day it was inoculated to 100 mL of LB-ampicillin-kanamycin medium in ratio 1:100 and grown at 37°C to OD (600 nm) of 0.6-0.7. Arabinose and IPTG were added to the final concentration of 1 mM of each and the culture was continued to be incubated for 4 hours at 37 ˚C. After the induction, cells were harvested by centrifugation at 7,000×g for 15 min at 4°C

and resuspended in 5 mL of binding buffer (20 mM NaH_2PO_4, 500 mM NaCl, pH 7.4).

Obtained cell suspension was sonicated on ice for 20 min using VCX 130 Sonicator with cycle 10s ON and 10s OFF with an amplitude of 75%. The highest concentration of the produced peptide was found in the insoluble fraction of lysed cells. Cell debris was removed by centrifugation at 15,000×g for 20 min at 4°C. The supernatant was discarded and pellet of insoluble fraction was resuspended and sonicated at the same conditions in 5 mL of binding buffer with 6 M guanidine-HCl. The sample was filtered through a 0.45 μm filter and applied to NGC system (Bio-Rad) equipped with HisTrap FF 1 mL (GE Healthcare Life Sciences) immobilized metal affinity chromatography (IMAC) column previously equilibrated in binding buffer with 4 M guanidine-HCl. After sample application, the column was washed with binding buffer containing 4 M guanidine-HCl. The peptide was eluted with elution buffer (20 mM NaH_2PO_4, 500 mM NaCl, 500 mM imidazole, pH 7.4) containing 4 M guanidine-HCl.

The fraction containing the eluted peptide was further purified by RP-HPLC system (Agilent) equipped with Jupiter Proteo, C-12, 250×10 mm column (Phenomenex). The eluate was mixed with TFA to reduce the pH to 2-3 and loaded on the column equilibrated in 5% of solvent B. The peptides were eluted by an increase of solvent B up to 60% over 60 min with a flow rate of 2 mL/min. All fractions were tested for antibacterial activity against *P. genomospecies* 1 NUB36187 using a drop on lawn assay. Also, elution fractions were analyzed by MALDI-TOF MS.

4.13 Co-expression of glycocin precursor genes with *palS*, *palT* or *palST*, and purification of synthesized peptides

Synthetic genes of sublancin 168 (*sunA-his*), glycocin F (*gccF-his*), enterocin F4-9 (*enfA4-9-his*), hypothetical peptide 1 (*hyp1-his*) and hypothetical peptide 2 (*hyp2-his*) were ordered in GenScript and synthesized with optimized codons for *E. coli* in a pUC57 vector. Genes were coding glycocin precursors with leaders and His6-

tag sequences in the C-terminus of the peptide. Every gene was recloned to the pRSFDuet-1 vector. The new pRSFDuet-1 vectors coding glycocin precursors (*sunA-his/gccF-his/enfA4-9-his/hyp1-his/hyp2-his)* were co-expressed with the pBAD24 vector coding *palS/palT/palST* in *E. coli* BL21(DE3) cells. The co-expression was performed in the same approach as was co-expression of pRSFDuet-1-palA-his with pBAD24-pasS, described before. The highest yield of synthesized peptides was found in the insoluble fraction of cell lysate. The peptides were purified from the insoluble fraction, analyzed for antibacterial activity and by MALDI-TOF MS in the same approach as pallidocin precursor described above.

4.14 Cloning of *core_hyp1-his* and *core_hyp2-his* genes and their co-expression with *palS*

A site-directed mutagenesis approach was used to engineer *core_hyp1-his* and *core_hyp2-his* genes coding for Hyp1-His and Hyp2-His peptides without leader sequences and with a His6-tag in the C-terminus. New construct pRSFDuet-1 coding *core_hyp1-his* was generated by PCR amplification using phosphorylated (5'-end) primers: F-Hyp1-NisP and R-Core_Hyp-His (**Supplementary Table 1**) and plasmid pRSFDuet-1-hyp1-his as template. New construct pRSFDuet-1 coding *core_hyp2-his* was generated by PCR amplification using phosphorylated (5'-end) primers: F-Hyp2-NisP and R-Core_Hyp-His (**Supplementary Table 1**) and plasmid pRSFDuet-1-hyp2-his as template. The amplified vectors were ligated and transformed to *E. coli* TOP10 cells by electroporation. Propagated plasmids were isolated and sequences of inserts in new constructs (pRSFDuet-1-core_hyp1-his and pRSFDuet-1-core_hyp2-his) were confirmed by DNA sequencing.

E. coli BL21(DE3) cells was co-expressed with pRSFDuet-1 vector coding *core_hyp1-his/core_hyp2-his* and with pBAD24 vector coding *palS*. The co-expression was induced for 7 h in the same approach as was co-expression of pRSFDuet-1-palA-his with pBAD24-pasS, described before. Synthesized peptides were purified from the insoluble fraction and tested for antibacterial activity in the

same method as pallidocin precursor described before. Also, it was analyzed by MALDI-TOF MS and LC-ESI-MS.

4.15 Molecular engineering of *his-Xa-palA* gene to introduce a Factor Xa cleavage site and a His-tag (N-terminus) in the PalA peptide

The gene *palA* was amplified by PCR using F-PalA-BamHI and R-PalA-HindIII[2] as primers (**Supplementary Table 1**) and *A. pallidus* 8 genomic DNA as template. PCR product *palA* and vector pRSFDuet-1 were double digested with BamHI and HindIII. Resulting products were cleaned, ligated by conventional cloning and transformed to *E. coli* TOP10 cells by electroporation. Propagated plasmid was isolated and the sequence of the insert in the construct (pRSFDuet-1-his-palA) was confirmed by DNA sequencing. The insert, *palA* gene, was introduced after the His6-tag coding sequence of the vector.

A site-directed mutagenesis approach was used to install a Factor Xa proteolytic cleavage site directly to the N-terminal part of the PalA peptide. The primers were designed to contain nucleotide sequences necessary to encode the amino acids – IEGR in place of four wild-type peptide residues – LQGS. Plasmid pRSFDuet-1 coding *his-Xa-palA* was generated by PCR amplification using phosphorylated (5'-end) primers: F-PalA-Xa and R-PalA-Xa (**Supplementary Table 1**) and construct pRSFDuet-1-his-palA as template. The amplified vector was ligated and transformed to *E. coli* TOP10 cells by electroporation. Propagated plasmid was isolated and the sequence of the insert in the construct (pRSFDuet-1-his-Xa-palA) was confirmed by DNA sequencing.

4.16 Synthesis and purification of the pre-His-Xa-PalA-Glc peptide

E. coli BL21(DE3) cells transformed with vectors pRSFDuet-1 coding *his-Xa-palA* and pBAD24 coding *palS* were grown in LB-ampicillin-kanamycin medium at 37°C for 16 h. Next day it was inoculated to 200 mL of fresh LB-ampicillin-

kanamycin medium in ratio 1:100 and grown at 37°C to OD (600 nm) of 0.6-0.7. Arabinose and IPTG were added to the final concentration of 1 mM of each and the culture continued to be incubated at 37°C for 4 h. After induction, cells were harvested by centrifugation at 7,000×g for 15 min at 4°C and resuspended in 5 mL binding buffer.

The cell paste suspended in binding buffer was sonicated on ice for 20 min (as described previously). Lysed cellular debris was removed by centrifugation at 15,000×g for 20 min at 4°C. The supernatant was discarded and the pellet containing the insoluble fraction was resuspended in 5 mL of binding buffer with 6 M guanidine-HCl. The sonication was repeated at the same conditions, the sample was filtered through a 0.45 μm filter and applied to NGC system (Bio-Rad) equipped with HisTrap FF 1 mL (GE Healthcare Life Sciences) IMAC column. The column was previously equilibrated in binding buffer with 4 M guanidine-HCl. After sample application, the column was washed with binding buffer containing 4 M guanidine-HCl. The peptide was eluted with elution buffer containing 4 M guanidine-HCl.

The fraction containing the eluted peptide was further purified on RP-HPLC system (Agilent) equipped with Jupiter Proteo, C-12, 250×10 mm column (Phenomenex). The eluate was mixed with TFA to reduce pH to 2-3 and loaded on the column equilibrated in 5% of solvent B. The peptides were eluted by an increase of solvent B up to 60% over 60 min with a flow rate of 2 mL/min. Elution fractions were tested for antibacterial activity against *P. genomospecies* 1 NUB36187 using a drop on lawn assay. Also, eluted fractions were analyzed by MALDI-TOF MS. Eluted fractions containing antibacterial activity and glycosylated pallidocin precursor with leader (pre-His-Xa-PalA-Glc) were lyophilized by a freeze-dryer.

Pellets were dissolved in 5 mL solution of 6 M guanidine-HCl and 0.1% TFA, and loaded for the 2nd round of purification on RP-HPLC system (Agilent) equipped with Jupiter Proteo, C-12, 250×4.6 mm column (Phenomenex). Before the run, the column was equilibrated in 5% of solvent B. The peptides were eluted by an

increase of solvent B from 35% up to 50% over 60 min with a flow rate of 1 mL/min. Elution fractions were tested for antibacterial activity against *P. genomospecies* 1 NUB36187 using a drop on lawn assay. Also, elution fractions were analyzed by MALDI-TOF MS. Elution fractions containing antibacterial activity and glycosylated pallidocin precursor with leader (pre-His-Xa-PalA-Glc) were lyophilized by a freeze-dryer.

4.17 Leader cleavage of pre-His-Xa-PalA-Glc peptide and core peptide purification

Pellets of purified glycosylated pallidocin precursor pre-His-Xa-PalA-Glc were dissolved in 1 mL solution of 6 M guanidine-HCl and loaded on gel filtration column PD-10 (GE Healthcare Life Sciences). The buffer exchange of the sample was performed according to manufacturer's recommendations. The column was previously equilibrated with 50 mM tris-HCl, 100 mM NaCl, pH 7.5 buffer. After the sample was loaded, eluate was collected and mixed with $CaCl_2$ to the end concentration of 2 mM. 10-20 μL of Factor Xa peptidase (enzyme conc. – 1 mg/mL, NEB) was added to 0.5 mL of previously gel filtrated peptide solution (peptide conc. – 0.5 mg/mL). The mixture was stored at room temperature for 3-6 h. After sample treatment with the peptidase, 4 mL of 6 M guanidine-HCl and TFA, to quench the pH to 2-3, were added. Then, sample was loaded on RP-HPLC system (Agilent) equipped with Jupiter Proteo, C-12, 250×4.6 mm column (Phenomenex) equilibrated in 5% of solvent B. The peptide was eluted by an increase of solvent B from 20% up to 60% over 80 min with a flow rate of 1 mL/min. Elution fractions were tested for antibacterial activity against *P. genomospecies* 1 NUB36187 using a drop on lawn assay. Also, elution fractions were analyzed by MALDI-TOF MS. Elution fractions containing active and glycosylated pallidocin core peptide PalA-Glc (mature pallidocin) were lyophilized by freeze-dryer. Pellets were stored at -20°C or dissolved in MiliQ containing 50% ACN and 0.1% TFA solution for further use. The peptide concentration was measured with NanoPhotometer N60 (Implen)

at 280 nm wavelength. The yield of synthesized pallidocin from 200 mL of bacterial culture was ~30 μg.

4.18 Iodoacetamide (IAA) assay for detection of free cysteines

To detect the presence of free Cys thiols in the peptide, an IAA assay was used. For the detection of free Cys residues in a native peptide, reactions contained 100 mM tris-HCl (pH 8.3), 40 mM IAA and the peptide. For the detection of free Cys residues in a reduced peptide, reactions contained 100 mM tris-HCl (pH 8.3), 10 mM tris(2-carboxyethyl)phosphine-hydrochloride (TCEP), 40 mM IAA and the peptide. All reactions were in 1 mL total volume and incubated for 2 hours at room temperature in the dark. The reaction mixtures were quenched with TFA to pH <4. Samples were loaded on RP-HPLC system (Agilent) equipped with a Jupiter Proteo, C-12, 250×4.6 mm column (Phenomenex) equilibrated in 5% of solvent B. Peptides were eluted by an increase of solvent B from 20% up to 60% over 40 min with a flow rate of 1 mL/min. Elution fractions containing peptides were tested for antibacterial activity against *P. genomospecies* 1 NUB36187 using a drop on lawn assay and further analyzed by MALDI-TOF MS. Free cysteines were determined by the presence or absence of thiol modification (+57 Da), carboxyamidomethyl (CAM).

4.19 Effect of pH on pallidocin stability

50 mM buffer solutions of KCl-HCl (pH 2), citric acid-sodium citrate (pH 4), phosphate (pH 6), tris-HCl (pH 8) and sodium carbonate-sodium bicarbonate (pH 10) were prepared for the following assay. Pallidocin was dissolved in MiliQ water containing 50% ACN and 0.1% TFA to the end concentration of 1 ng/μL. 13.5 μL of each buffer solution was mixed with 1.5 μL of pallidocin solution and the resulting mixtures were stored for 3 hours at room temperature. After the incubation, 15 μL of 500 mM tris-HCl (pH 7.5) buffer solution and 120 μL of NB

medium were added to the mixtures. Next, serial two fold dilutions with NB medium were made for each mixture. Bacteriocin activity was tested for each sample by agar well diffusion assay.

4.20 Effect of temperature on pallidocin stability

Pallidocin was dissolved in MiliQ water containing 50% ACN and 0.1% TFA to the end concentration of 1 ng/μL. 150 μL of NB medium was mixed with 1.5 μL pallidocin solution and stored at room temperature for 24 hours, 10 days and 30 days. Also, samples were stored at different temperatures: 60°C, 90°C for 3 hours and autoclaved for 15 min at 121°C. After the incubation, serial two fold dilutions with NB medium were made for each mixture. Bacteriocin activity was tested for each sample by agar well diffusion assay.

4.21 MIC determination for pallidocin

Minimum inhibitory concentrations (MIC) were determined by the method described by Wiegand et al. 2008 with some modifications. One colony of a sensitive strain was picked from NB agar plate, inoculated to liquid NB medium and grown at 55°C, 250 RPM. Bacteria culture was cultivated to OD (600 nm) of 0.1. Then, the culture was diluted with NB medium till concentration of 10×10^5 CFU/mL. OD correlation to CFU/mL of every strain was determined previously, before the experiment. 150 μL of fresh NB medium was mixed with 5 μL of pallidocin solution (1 ng/μL in 50% ACN and 0.1% TFA) and serial two fold dilutions with NB medium were made. 75 μL of resulting diluted pallidocin mixtures were transferred to a 96 well plate and mixed with previously prepared cell suspension of sensitive strain. The end volume of the mixture in the well – 150 μL and the end concentration of the sensitive strain in it – 5×10^5 CFU/mL. Positive controls – mixture of NB medium with sensitive strain, and negative controls –

mixture of NB medium with pallidocin, were prepared and dispersed in the same 96 well plates. Plates were incubated in a shaker for 18 hours at 55˚C and 250 RPM.

4.22 Determination of the sugar modification of pallidocin

The presence of a sugar moiety on Cys25 of pallidocin core peptide was confirmed via acid-catalyzed methanolysis and derivatization of the sugar, analysis by gas chromatography-mass spectrometry (GC-MS), and comparison to derivatized sugar standards. Purified pallidocin was lyophilized in a glass tube, internal standard mannitol and 0.5 mL of methanolic HCl 1 M (Sigma-Aldrich) were added. The mixture was heated at 85°C overnight. The reaction was cooled at room temperature and neutralized by adding solid silver carbonate (Sigma-Aldrich) till pH 7. For re-N-acetylation, 2 drops of acetic anhydride (Sigma-Aldrich) was added, mixed and stored overnight at room temperature in the dark. Next day, the reaction mixture was centrifuged at 1,000×g for 2 min. Supernatant was transferred to a new glass tube. 0.5 mL of methanol was added to the silver salt pellets, mixed and centrifuged again. Supernatant was collected and pooled with the first one. The procedure was repeated twice. Collected supernatant was evaporated in a vacuum evaporator. Tri-methylisation of the sugar was performed by adding 0.3 mL of silylation reagents – pyridine:hexamethyldisilazane:trimethyl-chlorosilane = 5:1:1. The mixture was incubated at room temperature for 30 min. Sugar standards mix: D-mannose, D-galactose, D-glucose, N-acetylgalactosamine and N-acetylglucosamine were also treated and derivatized using the conditions described above.

The derivatized sugar of pallidocin and sugar standards mix were analyzed individually by GCMS-QP2010 Plus system (Shimadzu) equipped with a Zebron ZB-1HT column, L = 30 m × I.D. = 0.25 mm × df = 0.25 μm (Phenomenex). The temperature gradient was from 140°C to 240°C at 4°C per min. The carrier gas was helium and the flow rate set at 2.0 mL/min.

4.23 Mass spectrometry analysis

Matrix-assisted laser desorption/ionization time-of-flight mass spectrometry (MALDI-TOF-MS) was carried out on a Voyager-DE-Pro (Applied Biosystems). 1-μL of RP-HPLC purified sample was spotted, dried on the target plate. Subsequently, 1 μL of matrix solution (5 mg/mL α-cyano-4-hydroxycinnamic acid (Sigma-Aldrich) dissolved in 50% ACN and 0.1% TFA) was spotted on the top of the sample plate. Data were analyzed with "Data Explorer" software version 4.0.0.0 (Applied Biosystems).

RP-HPLC purified glycocins were submitted for high resolution mass spectrometry (LC-ESI-Q-MS and MS/MS) analysis at the Interfaculty Mass Spectrometry Center (IMSC) of the University Medical Center Groningen (UMCG). The LC-MS/MS analyses of peptides were performed on a Dionex Ultimate 3000 nano-LC system equipped with a column Acclaim Pepmap, 75 mm × 150 mm (Thermo Scientific) coupled to a Q-Exactive Plus quadrupole-Orbitrap mass spectrometer (Thermo Scientific). A linear gradient of 2–60% (A: 99.9% MiliQ and 0.1% formic acid; B: 99.9% ACN and 0.1% formic acid) in 50 min at a flow rate of 300 nL/min was employed. The mass spectrometer was operated in data-dependent mode, automatically switching between MS and MS/MS acquisition for the five most abundant doubly and triply charged ions with a minimal signal of 2,500 in a given MS spectrum. Full-scan MS spectra were acquired from *m/z* 300 to 1,700 in the Orbitrap spectrometer with a resolution of 60,000. The five most intense ions which met the set criteria were then isolated for fragmentation in the linear ion trap, with a dynamic exclusion of 10 s. Protein sequences analysis was performed on Xcalibur platform (Thermo Scientific).

Geo26 produced by strain 15 was submitted for mass spectrometry analysis at Proteomic Center of Vilnius University. Liquid chromatography (LC) separation of trypsin cleaved bacteriocin was performed with a nanoAcquity UPLC system (Waters Corporation, Wilmslow, UK). Peptides were loaded on a reversed-phase

trap column PST C18, 100A° 5 µm, 180 µm × 20 mm (Waters Corporation) and subsequently separated on HSS-T3 C18 1.8 µm, 75 µm × 250 mm analytical column (Waters Corporation) in 60 min linear gradient (A: 99.9% MiliQ and 0.1% formic acid; B: 99.9% ACN and 0.1% formic acid) at a flow rate of 300 nL/min. The analytical column temperature was kept at 40°C. Bacteriocin analyzed by Synapt G2 HDMS mass spectrometer using MS/MS method. Protein sequences analysis was performed on ProteinLynx Global SERVER platform.

4.24 Protein purification by ammonium sulfate precipitation

A colony of strain 15 from NB agar medium plate was picked and inoculated to NB medium and incubated at 200 RPM, 55°C. After overnight growth, the culture was inoculated to a fresh NB medium supplemented with the salts in ratio 1:100 and propagated at the same conditions. After 7 h cultivation, the culture was centrifuged for 15 min at 10,000×g, 4°C. The supernatant was saturated up to 80% with ammonium sulfate and stirred for 2-3 hours, then centrifuged for 30 min at 10,000×g. The supernatant was discarded and the pellet was dissolved in 1/10 volume of the starting bacterial culture in buffer A (25 mM tris-HCl, pH 8). The dialysis was performed with SnakeSkin Dialysis Tubing, 3.5K MWCO (Thermo Fisher Scientific) in the same buffer according to the manufacturer's recommendations. The obtained crude bacteriocin extract (CBE) was used for further analysis or stored at -20°C.

4.25 Protein analysis by Glycine-SDS-PAGE

Purified proteins produced by strain 15 were analyzed by Glycine-SDS-PAGE in 15% acrylamide separating gel. After the electrophoresis, the gel was sliced into 2 pieces. One half of the gel was immersed in a fixation solution (25% isopropanol, 10% glacial acetic acid and 65% H_2O) for 1 h and stained with PageBlue Protein Staining Solution (Thermo Fisher Scientific) according to the manufacturer's

recommendations. The other half of the gel was also immersed in the fixation solution for 2-3 h and washed with distillated water for 10-15 min 4-5 times. The gel was placed into a Petri dish and overlaid with NB agar medium inoculated with the indicator strain (the final OD (600 nm) of 0.1). When the agar medium was solid, the dish was incubated in an incubator for 12-18 h at 55°C. After the incubation, stained gel was aligned with gel, which was overlaid with the indicator strain, to determine the approximate molecular mass of a bacteriocin.

4.26 Protein analysis by Tricine-SDS-PAGE

Purified proteins produced by strain 15 were analyzed by Tricine-SDS-PAGE in 16% acrylamide separating gel described by Schägger (2006). After the electrophoresis the gel was sliced into 2 pieces. One half of the gel was stained by silver staining described by Chevallet et al. (2006). The other half of the gel was immersed in the fixation solution (25% isopropanol, 10% glacial acetic acid and 65% H_2O) for 2-3 h and washed with distillated water for 10-15 min 4-5 times. The gel was placed into a Petri dish and overlaid with NB agar medium inoculated with the indicator strain (the final OD (600 nm) of 0.1). When the agar medium was solid, the dish was incubated in an incubator for 12-18 h at 55°C. After the incubation, stained gel was aligned with gel, which was overlaid with the indicator strain, to determine the approximate molecular mass of a bacteriocin.

4.27 Effects of enzymes on activity of the CBE

The effects of enzymes: pepsin, α-amylase and lipase (0.1 mg/mL), on bacteriocin activity (CBE preparations after purification) were tested by treatment with enzymes for 2 hours at 37°C. CBE preparations without enzymes were used as controls. After treatment, the residual inhibitory activity against indicator strain *P. genomospecies* 1 NUB36187 (BGSC 9A11) was determined by agar well diffusion assay.

4.28 Effect of temperature on Geo18 and Geo26 stability

Partially purified Geo18 and Geo26 preparations (CBE) were stored at different temperatures: 60˚C, 70˚C, 80˚C, 90˚C, 100˚C for 2 hours or autoclaved for 15 min at 121˚C. After the incubation, serial two fold dilutions with NB medium were made for each mixture. Bacteriocin activity was tested for each sample with agar well diffusion assay.

4.29 Chromatographic purification of bacteriocins produced by strain 15

After isolation of the CBE, it was filtered through a 0.45 μm filter and loaded on anion-exchange column, UNO Q6, 12×53 mm, 6 mL (Bio-Rad) equilibrated in the buffer A (25 mM tris-HCl, pH 8). A flow throw fraction was collected for a further use. Proteins were eluted by linear gradient increase of the elution buffer B (25 mM tris-HCl, pH 8 and 1 M NaCl) up to 50% over 10 column volumes with a flow rate of 2 mL/min. All fractions were tested for bacteriocin activity by a spot on lawn assay. Active elution fractions were further dialyzed in SnakeSkin Dialysis Tubing, 3.5K MWCO in the buffer C (25 mM sodium-citrate, pH 3.8). Dialyzed sample was centrifuged at 15.000×g for 10 min and the supernatant was loaded on cation-exchange column, UNO S6, 12×53 mm, 6 mL (Bio-Rad) equilibrated in the buffer C. Proteins were eluted by linear gradient increase of the buffer D (25 mM sodium-citrate, pH 3.8 and 1 M NaCl) up to 50% over 10 column volumes with a flow rate of 2 mL/min. All fractions were tested for bacteriocin activity by a spot on lawn assay. Active elution fractions were dialyzed in SnakeSkin Dialysis Tubing, 3.5K MWCO in the buffer E (15 mM MES pH 6.5 and 2 M ammonium sulfate). Dialyzed sample was centrifuged at 15.000×g for 10 min and the supernatant was loaded on Bio-Scale MT5 column, 10×64 mm (Bio-Rad) packed with Macro-Prep (methacrylate-based) hydrophobic interaction chromatography (HIC) resin (Bio-Rad) previously equilibrated in the buffer E. Proteins were eluted by linear gradient increase of the elution buffer F (15 mM MES, pH 6.5) up to 100% over 10 column

volumes with a flow rate of 2 mL/min. All fractions were tested for bacteriocin activity by a spot on lawn assay. Active elution fractions were dialyzed in SnakeSkin Dialysis Tubing, 3.5K MWCO in the buffer A or concentrated with chloroform/methanol precipitation and analyzed in SDS-PAGE.

The flow throw fraction, after CBE load on the anion-exchange column, was further dialyzed with SnakeSkin Dialysis Tubing, 3.5K MWCO in the buffer G (25 mM MES pH 6.5). The sample was centrifuged at 15.000×g for 10 min and the supernatant was loaded on cation-exchange column, UNO S6, 12×53 mm, 6 mL (Bio-Rad) equilibrated in the buffer G (25 mM MES, pH 6.5). Proteins were eluted by linear gradient increase of the elution buffer H (25 mM MES, pH 6.5 and 1 M NaCl) up to 50% over 10 column volumes with a flow rate of 2 mL/min. All fractions were tested for bacteriocin activity by a spot on lawn assay. Active elution fractions were dialyzed in SnakeSkin Dialysis Tubing, 3.5K MWCO in the buffer A or concentrated by chloroform/methanol precipitation and analyzed by isoelectro focusing.

4.30 The isoelectro focusing analysis of bacteriocins produced by strain 15

Purified proteins obtained by the ion-exchange chromatography were concentrated by methanol/chloroform precipitation. The pallets were dissolved in an IPG sample buffer containing 7 M urea, 1 M thiourea, 4% CHAPS, 0.5% IPG Buffer pH 3-10 (GE Healthcare Life Sciences), 40 mM DTT and 0.002% bromophenol blue. Next, 1 mL of samples (1 µg of protein per sample) was loaded on two ReadyStrip IPG Strips, 24 cm, pH 3–10 (Bio-Rad) for passive re-hydration over night at room temperature. The third strip was used for analysis of IEF Standards (Bio-Rad) andwas prepared in the same way. Isoelectric focusing (IEF) was performed with Multiphor II device (GE Healthcare) using following steps: 300 V, 2 mA, 5W (300 V/h); 1000 V 2 mM, 5 W (1000 V/h); 2000 V, 2 mA, 5 W (2000 V/h); 3500 V, 2 mA, 5 W (70 kV/ h).

After the analysis, the IPG strip with the standards and one of the strips with the sample, were washed in a fixation solution (25% isopropanol, 10% glacial acetic acid and 65% H_2O) for 20 min. After the fixation, strip with the IEF standards was stained with PageBlue Protein Staining Solution (Thermo Scientific) according to the manufacturer's recommendations. Another strip with the sample, after the fixation, was placed in a Petri dish and overlaid with NB agar medium, which was inoculated with an indicator strain (the final OD (600 nm) of 0.1). When the medium was solid, the plate was incubated in a thermostat at 55°C for 12-18 h.

After the incubation, strip overlaid with the indicator strain and strip with the stained standards were aligned to determine the bacteriocin location and pI based on the inhibition zone. The other IPG strip with the same sample part of the strip corresponding inhibition zone of bacteriocin was cut. Part of the third strip, where the location of the bacteriocin was determined by the inhibition zone, was cut and immersed in an SDS-PAGE sample buffer for 15 min. Further, it was analyzed in a second dimension electrophoresis (SDS-PAGE). Protein band of interest corresponding the molecular weight of the bacteriocin was cut from the gel and submitted for mass spectrometry analysis.

4.31 Cloning of the *geo26* gene coding for geobacillin 26 produced by strain 15

Gene sequence of *geo26* was amplified by PCR using primers: F-Geo26 and R-Geo26 **(Supplementary Table 1)** as primers and strain 15 genomic DNA as template. The PCR product was digested with PciI and HindIII. pBAD24 and pNZ8048 vectors were double digested with NcoI and HindIII, run on agarose gel electrophoresis and extracted from the gel. Digested PCR product was ligated with double digested vectors. Construct pBAD24-geo26 was propagated in *E. coli* TOP10 and construct pNZ8048-geo26 in *L. lactis* pNZ9000. The sequence of the insert in resulting constructs was confirmed by DNA sequencing. Gene *geo26* was

cloned into the MCS after the arabinose promoter and RBS of the pBAD24 vector and into the MCS after the nisin promoter and RBS of the pNZ9000 vector.

A site-directed mutagenesis approach was used to introduce His6-tag sequence (MHHHHHHH-) in the N-terminus of the Geo26 protein. A new construct pBAD24 coding *his-geo26* was generated by PCR amplification using phosphorylated primers (5'-end): F-His-Geo26 and R-His-Geo26 (**Supplementary Table 1**). Construct pBAD24-geo26 was used as template. The PCR product was cleaned, ligated and transformed to *E. coli* TOP10 cells. Propagated plasmid was isolated and the His6-tag coding sequence in the resulting construct (pBAD24-his-geo26) was confirmed by DNA sequencing.

4.32 Expression of the *geo26* gene and antibacterial activity screening of Geo26 protein

E. coli BL21(DE3) cells with transformed pBAD24 vector coding *geo26* was inoculated to LB-ampicillin medium and grown at 37°C for 16 h. The next day it was inoculated to two flasks of 20 mL of LB medium (no ampicillin) in ratio 1:100 and grown at 37°C to OD (600 nm) of 0.6-0.7. Arabinose was added to one of the flasks to the final concentration of 2 mM. Flasks with bacterial cultures were continued to be incubated for 4 hours at 37°C. After the incubation, 5 mL of cell culture (induced and noninduced-control) was harvested by centrifugation at 7,000×g for 15 min at 4˚C. Supernatant was collected and cells were resuspended in 0.3 mL of buffer containing 50 mM tris-HCl pH 8. Cell suspensions were mixed with a glass beads (0.1 mm diameter) and mechanically disrupted in a pulsing vortex mixer. The cell lysates were centrifuged at 16,000×g for 5 min. Cell culture supernatant and the supernatant of cell lysate were collected and tested for antibacterial activity by agar well diffusion assay.

Expression of gene *geo26* in *L. lactis* was performed using NICE Expression System for *L. lactis* (MoBiTec GmbH) according to manufacturer's recommendations. 2 ng/mL of nisin was used for gene induction. Cells and

supernatant were collected for antibacterial screening after 4 h induction. The cell lysis and activity screening were performed as described above for *E. coli*.

4.33 Expression of the *his-geo26* gene and purification of the His-Geo26 protein

E. coli BL21(DE3) cells with transformed pBAD24 vector coding *his-geo26* was inoculated to LB-ampicillin medium and grown at 37°C for 16 h. The next day it was inoculated to 100 mL of LB-ampicillin medium in ratio 1:100 and grown at 37°C to OD (600 nm) of 0.6-0.7. Arabinose was added to the final concentration of 2 mM and the culture was continued to be incubated for 4 h at 37˚C. After the induction, cells were harvested by centrifugation at 7,000×g for 15 min at 4˚C and resuspended in 5 mL of binding buffer (20 mM NaH_2PO_4, 500 mM NaCl, pH 7.4). Obtained cell suspension was sonicated on ice for 20 min using VCX 130 Sonicator with cycle 10s ON and 10s OFF with an amplitude of 75%. The presence of the protein was checked in soluble and in insoluble fractions of the cell lysate. Cell debris was removed by centrifugation at 15,000×g for 20 min at 4°C. The supernatant was loaded on NGC system (Bio-Rad) equipped with HisTrap FF 1 mL (GE Healthcare Life Sciences) immobilized metal affinity chromatography (IMAC) column previously equilibrated in binding buffer. After sample application, the column was washed with binding buffer and elution was performed with elution buffer (20 mM NaH_2PO_4, 500 mM NaCl, 500 mM imidazole, pH 7.4).

The pellet of insoluble fraction was resuspended and sonicated at the same conditions in 5 mL of binding buffer with 6 M guanidine-HCl. The sample was filtered through a 0.45 μm filter and applied to NGC system (Bio-Rad) equipped with HisTrap FF 1 mL (GE Healthcare Life Sciences) immobilized metal affinity chromatography (IMAC) column previously equilibrated in binding buffer with 4 M guanidine-HCl. After sample application, the column was washed with binding buffer containing 4 M guanidine-HCl. The elution was performed with elution buffer containing 4 M guanidine-HCl.

5 Results

5.1 Partial identification of strains 8 and 15

In this work thermophilic bacteria strains 8 and 15 were used as a putative bacteriocin producers. Strains were derived from the culture collection of the Microbiology and Biotechnology Department of Vilnius University. The strain 15 was previously isolated from Lithuanian oil wells, and the strain 8 from the subsurface above oil pools. Strain 8 was previously identified as gram-positive endo-spore forming thermophilic bacteria by colleague (Pranckutė, doctoral dissertation, 2017). In this study strain 15 was identified as gram-positive endo-spore forming thermophilic bacteria as well.

Partial identification of strains 8 and 15 were performed by 16S rDNA BLAST analysis in NCBI database. Analysis showed that 16S rDNA sequence of the strain 15 has 99% identity to 16S rDNA sequences of variaous *Geobacillus stearothermophilus* strains. Meanwhile analysis of 16S rDNA sequence of the strain 8 showed that it has 99-100% identity to 16S rDNA sequences of various *Aeribacillus pallidus* strains. Based on these results strain 8 belongs to the genus of *Aeribacillus*, and the stain 15 – *Geobacillus*.

5.2 Identification of antibacterial activity spectrum of strains 8 and 15

The antibacterial activity of strains 8 and 15 was determined (**Table 1**) by agar overlay method. It was observed that the strain 8 produces antibacterial substance(s) that is active against 12 of 13 tested thermophilic strains: *Parageobacillus* spp., *Geobacillus* spp., *Anoxybacillus* sp. Meanwhile, the strain 15 produces substance(s) that is active against 6 of 13 tested strains: *Parageobacillus* spp., *Geobacillus* spp. and *Anoxybacillus* sp. strains as well.

Table 1. Antibacterial activity spectrum of strains 8 and 15 against thermophilic *Parageobacillus* spp. *Geobacillus* spp. *Anoxybacillus* sp. and *Aeribacillus* sp. strains. Antibacterial activity observed (+); antibacterial activity not observed (-); observed very low antibacterial activity (±). The antagonistic interactions tested by agar overlay method.

Indicator strains	Antibacterial activity producer	
	strain 8	strain 15
P. toebi DSM 14590^T	+	-
P. thermoglucosidns DSM 2542^T	+	-
P. genomospecies 1 NUB36187 (BGSC 9A11)	+	+
G. thermodenitrificans DSM 465^T	+	-
G. thermocatenulatus DSM 15378	+	-
G. thermoleovorans DSM 5366^T	+	±
G. thermoleovorans DSM 15325^T	+	-
G. stearothermophilus DSM 22^T	+	±
G. jurassicus DSM 15726^T	+	+
G. subterraneus DSM 13552^T	+	+
G. thermoleovorans HTA426	+	-
A. tepidamans DSM 16325^T	+	±
A. pallidus DSM 3670^T	-	-
Strain 15	+	-
Strain 8	-	-

This results show that both strains: 8 and 15, produces antibacterial substances that is active against closely related thermophilic bacteria. Strain 8 produces broad-spectrum antibacterial agents and strain 15 produces antibacterial agents which exhibits narrow-spectrum of activity.

5.3 Genomic DNA sequencing

Strains 8 and 15 were submitted for the whole genome sequencing service in the GATC Biotech (Germany) on an Illumina HiSeq 2000. Obtained raw sequencing data was submitted to the Velvet server for a *de novo* assembly on each genome, resulting in the draft genome sequences in contigs. The RAST server was subsequently used to annotate the genomes. The genome sequences of strains 8 and

15 have been deposited into the NCBI database under the accession numbers LVHY00000000 and LVHZ00000000, respectively.

5.4 Identification of strains

Genomic tools for the thermophilic genus *Geobacillus* phylogenomic re-assessment was successfully used in previous studies (Aliyu et al., 2016; Burgess et al., 2017). Regarding this, the Genome-to-Genome Distance Calculator (GGDC), for a digital DNA-DNA hybridization (dDDH) analysis (Meier-Kolthoff et al., 2014) of the strains 8 and 15 was applied in this study. The probability of DDH > 70% indicates that strain 8 belongs to the same species as *A. pallidus* KCTC 3564^T (prob. 95.5%) (**Table 2**). The genome G+C% content difference between strains 8 (38.8%) and KCTC 3564^T (39.3%) is 0.55. According to the calculator's recommendations, the difference between two strains of the same species may be no more than 1%. It should be noted that the genus of *Aeribacillus* contains only one species.

Table 2. Phylogenomic assessment of strain 8 by GGDC.

Query genome	Reference genome	Formula 2		G+C% difference
		DDH	Prob. DDH >= 70%	
Strain 8	*A. pallidus* KCTC 3564^T	**89.1**	**95.5**	**0.55**
	A. pallidus 8m3	90.9	96.1	0.1
	A. pallidus GS3372	74.5	85.09	18.6

Phylogenomic analysis of the strain 15 by GGDC revealed that the highest probability of DDH > 70% is between strain 15 and *G. stearothermophilus* ATCC 12980^T (prob. 64.07%), and *G. thermoleovorans* DSM 15325 (prob. 77.36%) (**Table 3**). The rest *(Para)Geobacillus* spp. and *Caldibacillus debili*s bacteria did not show high probability of dDDH with the strain 15. The genome G+C% content difference between the strain 15 and strains ATCC 12980^T and DSM 15325 is 0.7 and 0.1, respectively. It should be noted that *G. thermoleovorans* DSM 15325

(Dinsdale et al., 2011) used to be named *G. lituanicus* DSM 15325^T species (Kuisienė et al., 2004). Analysis of *G. thermoleovorans* DSM 15325 genome by GGDC revealed that the highest probability of DDH > 70% is between strain DSM 15325 and *G. stearothermophilus* ATCC 12980^T (prob. 83.79%) (**Table 4**). Genome G+C% content difference between the strains is 0.8.

Table 3. Phylogenomic assessment of strain 15 by GGDC.

Query genome	Reference genome	Formula 2		G+C% difference
		DDH	Prob. DDH >= 70%	
Strain 15	*G. stearothermophilus* ATCC 12980^T	**63.8**	**64.07**	**0.7**
	G. kaustophilus NBRC 102445^T	40.1	2.76	0.38
	G. genomosp. 1 Et2/3	34.1	0.49	0.57
	G. genomosp. 1 Et7/4	34.1	0.49	0.68
	G. thermoleovorans KCTC 3570^T	39.2	2.23	0.02
	G. thermodenitrificans KCTC3902^T	27.2	0.03	3.26
	G. thermodenitrificans subsp. *thermodenitrificans* DSM 465^T	27.1	0.03	3.32
	G. subterraneus KCTC 3922^T	33.9	0.46	0.17
	G. thermocatenulatus KCTC 3921^T	38.1	1.66	0.48
	G. jurassicus NBRC 107829^T	32.6	0.29	0.15
	G. zalihae NBRC 101842^T	42.3	4.69	0.48
	G. thermoleovorans DSM 15325	**69.5**	**77.36**	**0.1**
	G. galactosidasius DSM 18751^T	25.4	0.01	10.8
	G. uzenensis BGSC 92A1^T	33.5	0.4	0.14
	G. genomosp. 3 JF8	32.5	0.28	0.5
	G. icigianus DSM 28325^T	29	0.07	0.34
	P. thermoglucosidasius DSM 2542^T	21.1	0	8.48
	P. toebii NBRC 107807^T	22.2	0	10.22
	P. caldoxylosilyticus NBRC 107762^T	21.5	0	8.45
	C. debilis DSM 16016^T	20.8	0	0.74
	P. thermantarcticus DSM 9572^T	22.8	0	8.7
	P. genomosp. 1 NUB3621	20.6	0	7.99
	G. genomosp. 2 PSS1	31.9	0.22	0.03

Table 4. Phylogenomic assessment of *G. thermoleovorans* DSM 15325^T by GGDC.

Query genome	Reference genome	Formula 2		G+C difference
		DDH	Prob. DDH >= 70%	
G. thermoleovorans DSM 15325^T	*G. stearothermophilus* ATCC 12980^T	**73.5**	**83.79**	**0.8**
	G. kaustophilus NBRC 102445^T	39.2	2.21	0.27
	G. genomosp. 1 Et2/3	33.3	0.37	0.46
	G. genomosp. 1 Et7/4	33.8	0.44	0.58
	G. thermoleovorans KCTC 3570^T	38.3	1.76	0.13
	G. thermodenitrificans KCTC 3902^T	27.3	0.03	3.16
	G. thermodenitrificans subsp. *thermodenitrificans* DSM 465^T	26.9	0.02	3.22
	G. subterraneus KCTC 3922^T	33.5	0.4	0.07
	G. thermocatenulatus KCTC 3921^T	39.5	2.38	0.38
	G. jurassicus NBRC 107829^T	32.6	0.29	0.05
	G. zalihae NBRC 101842^T	42.6	4.97	0.38
	G. galactosidasius DSM 18751^T	26.8	0.02	10.69
	G. uzenensis BGSC 92A1^T	33.1	0.34	0.04
	G. genomosp. 3 JF8	32.9	0.32	0.61
	G. icigianus DSM 28325^T	28.8	0.06	0.24
	P. thermoglucosidasius DSM 2542^T	21.5	0	8.37
	P. toebii NBRC 107807^T	23.2	0	10.12
	P. caldoxylosilyticus NBRC 107762^T	21.6	0	8.34
	C. debilis DSM 16016^T	22.5	0	0.64
	P. thermantarcticus DSM 9572^T	23.4	0	8.6
	P. genomosp. 1 NUB3621	21.3	0	7.88
	G. genomosp. 2 PSS1	31.2	0.17	0.14

Based on the phylogenomic assessment by GGDC it can be confirmed that the strain 8 can be identified as *A. pallidus* 8, meanwhile the strain 15 is closely related to *G. stearothermophilus* species and can be designated as *G. stearothermophilus* 15. Moreover, *G. thermoleovorans* DSM 15325 should be redesignated to *G. stearothermophilus* DSM 15325.

5.5 Genome mining for bacteriocins

Genomes of *A. pallidus* 8 and *G. stearothermophilus* 15 were processed with the BAGEL4 tool for identification of gene clusters coding for bacteriocins. Genome analysis of the strain 15 observed no genes of any known or putative bacteriocin or its biosynthetic machinery. Meanwhile, genome analysis of the strain 8 identified four putative bacteriocin biosynthesis gene clusters (**Fig. 10**). Proteins encoded in these gene clusters have sequence similarity to proteins of biosynthetic machineries of known bacteriocins. Three gene clusters are coding for a putative sactipeptides and one for a putative glycocin. Two of these sactipeptide gene clusters include small open reading frames which might encode putative precursor peptides. Thesse precursors have no sequence similarity to any known and described bacteriocin precursors. The third putative sactipeptide gene cluster and the putative glycocin gene cluster encode precursors which have sequence similarity to known and described bacteriocins.

This gene cluster coding for the putative glycocin biosynthetic machinery was named *pal*, and it was predicted between base pairs 2,768-7,572 in contig with accession number: LVHY01000134.1. It contains five genes, which were named *palA*, *palS*, *palT*, *paldbA* and *paldbB* (**Fig. 11**). These genes encode for proteins, which are similar to the sublancin 168 biosynthetic machinery encoded in various *Bacillus* sp. bacteria. BLAST analysis revealed that *palA* is coding for the 61 amino acid pre-peptide (**Fig. 12**) with 39% sequence similarity to the sublancin 168 precursor (NP_390031.1). We named this putative new bacteriocin pallidocin.

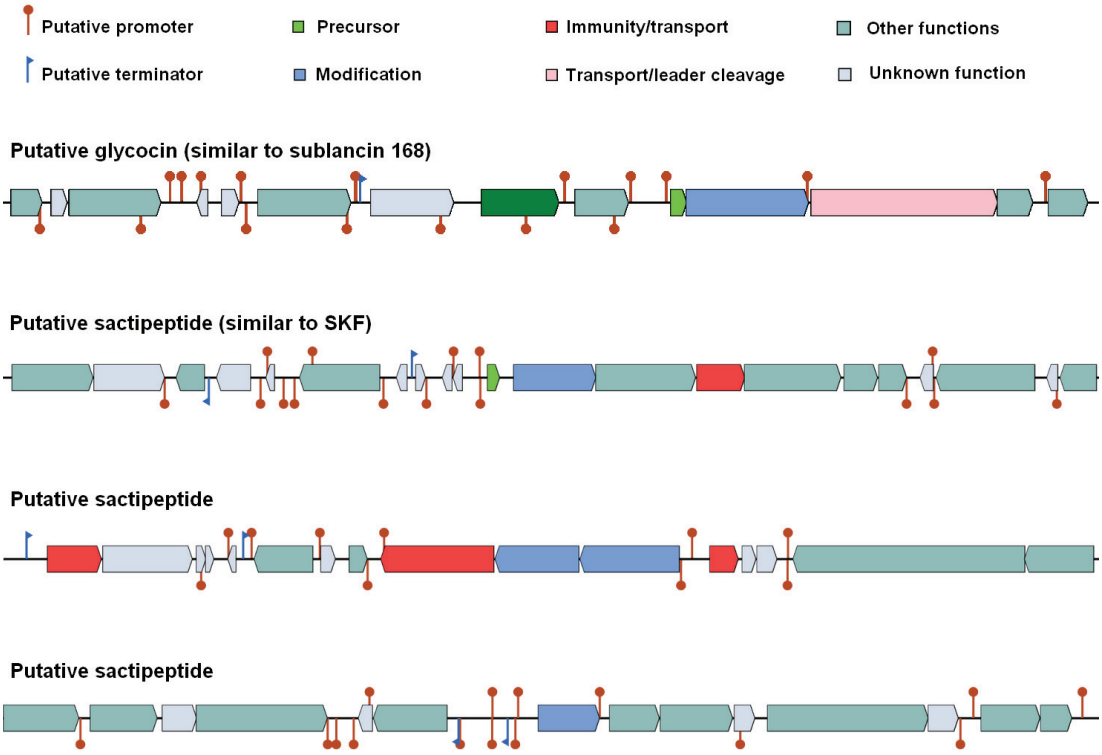


Figure 10. Bacteriocin mining in *A. pallidus* 8 genome using BAGEL4.

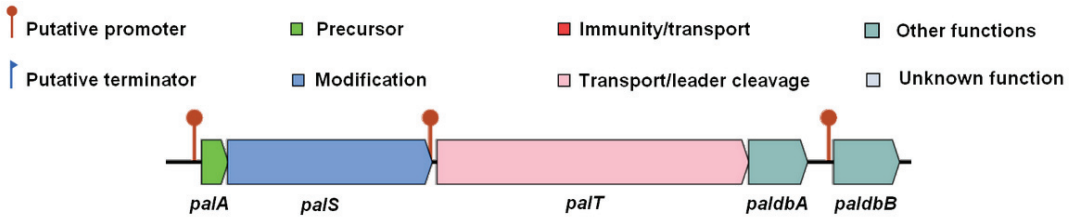


Figure 11. Novel glycocin biosynthesis gene cluster *pal* identified by BAGEL4 in *A. pallidus* 8.

```
PalA          MKDLLKELMYEVDLEEMENLQGSGYSAAQCAWMALSCVNYIPGVGFGCGGYS-ACELYKRYC-    61
SunA          ----MEKLFKEVKLEELENQKGSGLGKAQCAALWLQCA---SGGTIGCGGGAVACQNYRQFCR    56
```

Figure 12. Novel glycocin precursor (PalA) and sublancin 168 precursor (SunA) alignment. Letters in green indicate identcal amino acids at the same position.

Gene *palS* is coding for a protein with 53% sequence similarity to the SunS-family peptide glycosyltransferase (WP_097849814.1), *palT* – a protein with 38% sequence similarity to the SunT superfamily leader cleavage/ABC-type transporter (WP_057576215.1). *paldbA* codes for protein with 50% sequence similarity to the thioredoxin-like enzyme (WP_061859981.1). *paldbB* codes for a protein with 47% sequence similarity to the DsbB disulfide bond formation protein B (WP_098593227.1).

One of the gene clusters coding for sactipeptides encodes a precursor with high similarity to bacteriocin sporulation-killing factor (SKF). This gene cluster was named *skf*. It was observed between base pairs 15,767-22,038 in contig with accession number LVHY01000110.1. The operon contains 7 genes, which were named *skfA, skfB, skfC, skfD, skfE, skfF* and *skfG* (**Fig. 13**). This operon has similar gene composition as the SKF biosynthetic machinery encoded in various *Bacillus* sp. bacteria. BLAST analysis showed that *skfA* is coding for 55 amino acid precursor (**Fig. 14**) with 69% sequence similarity to bacteriocin SKF precursor (CUB15133.1).

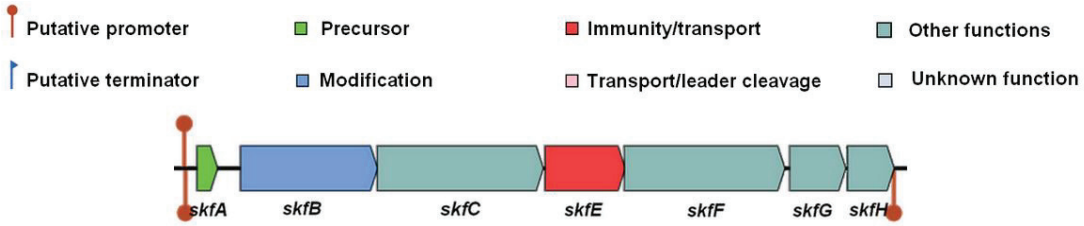


Figure 13. Sactipeptide biosynthesis gene cluster *skf* identified by BAGEL4 in *A. pallidus* 8.

```
SkfA (A)        MEKQKKQWEPTAIKELKKPAGTSIVKAAGCMSCWASKSISMTRVCWLPHPAMRLL     55
SkfA (B)        MTRNQKEWESVSKKNLKRPGGTSIVKAAGCMGCWASKSIAMTRVCALPHPAMRAI     55
```

Figure 14. Alignment of sactipeptide precursors (SkfA) encoded in *A. pallidus* (A) and in *Bacillus* sp. (B). Letters in green indicate identical amino acids at the same position.

Gene *skfB* is coding for a protein with 62% sequence similarity to a rSAM maturase (ASS69691.1). *skfC* is coding for a protein with 62% sequence similarity to the SKF biosynthesis protein with intramembrane metallo-protease domain (SCA84087.1). *skfD* is coding for a protein with 66% sequence similarity to the ABC transporter of SKF peptide (SCA84088.1). *skfE* is coding for a protein with 43% sequence similarity to the SKF system integral membrane protein (WP_076458337.1). *skfF* is coding for a protein with 59% sequence similarity to the HEAT repeat domain-containing protein (KKB75502.1). *skfG* is coding for a protein with 50% sequence similarity to the thioredoxin protein (EXF51921.1).

It can be concluded that *A. pallidus* 8 possibly secretes at least two bacteriocins: a novel glycocin and a sactipeptide with high sequence similarity to the SKF bacteriocin. The other two gene clusters contains genes related to biosynthesis of sactiobiotics: ABC-transporter and rSAM enzyme. The lack of other genes that could be part of a bacteriocin biosynthetic machinery rises doubts if these gene clusters are coding for bacteriocins synthesis.

5.6 Biosynthesis of new glycocin - pallidocin and its characterization

The biosynthetic gene cluster of pallidocin – *pal* (**Fig. 11**), which is 4805 bps in length and encodes genes: *palA*, *palS*, *palT*, *paldbA* and *paldbB,* was amplified by PCR as a single unit and cloned to expression vector pBAD24. The gene cluster *pal* was expressed in heterologous host *E. coli* BL21(DE3). The antibacterial activity of the synthesized peptide was observed in the culture's supernatant, from which bacteriocin was subsequently purified using a hydrophobic polyaromatic resin (Amberlite XAD16N, Sigma-Aldrich), cation exchange chromatography (HiTrap SP HP column, GE Healthcare Life Sciences) and RP-HPLC (Jupiter Proteo column, C-12, Phenomenex). The yield of the peptide was enough for MS analysis and initial antibacterial activity screening, but not for the quantification by measuring the absorbance at 280 nm wavelength.

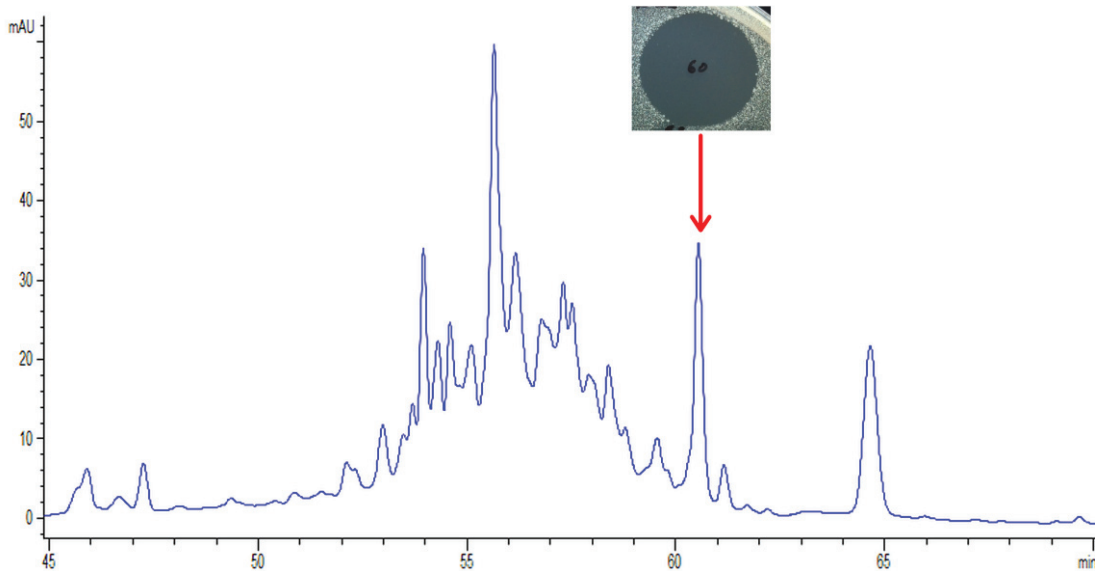


Figure 15. Purification of novel glycocin – pallidocin. Chromatogram represents the last step of purification by RP-HPLC. Absorbance of elution fractions measured at 280 nm. Inhibitory activity of elution fractions measured by spot on a lawn assay. *P. genomospecies* 1 was used as a sensitive strain.

Purified bacteriocin (**Fig. 15**) was submitted for mass spectrometry analysis to determine post-translational modifications. The predicted monoisotopic mass of the unmodified pallidocin core peptide is 4061.76 Da. The monoisotopic mass of purified native pallidocin observed by LC-ESI-MS was 4219.79 Da (**Supplementary Fig. 1a**). The monoisotopic mass of the peptide, after treatment with TCEP, was 4223.82 Da (**Supplementary Fig. 1b**). It suggests that the peptide has a post-translational modification with a mass of +162 Da and two disulfide bonds (-4 Da). To identify the modified residue, pallidocin was fragmented with chymotrypsin and further analyzed by LC-ESI-MS/MS mass spectrometry (**Supplementary Fig. 2**). It was revealed that Cys25 of the core peptide has a modification with a mass of +162 Da, which might be a hexose moiety. To identify the sugar, pallidocin was analyzed by GC-MS (**Supplementary Fig. 3**). Analysis demonstrated that the moiety attached to the Cys25 residue is glucose. The proposed structure of the novel glycocin pallidocin is depicted in **Fig. 16**.

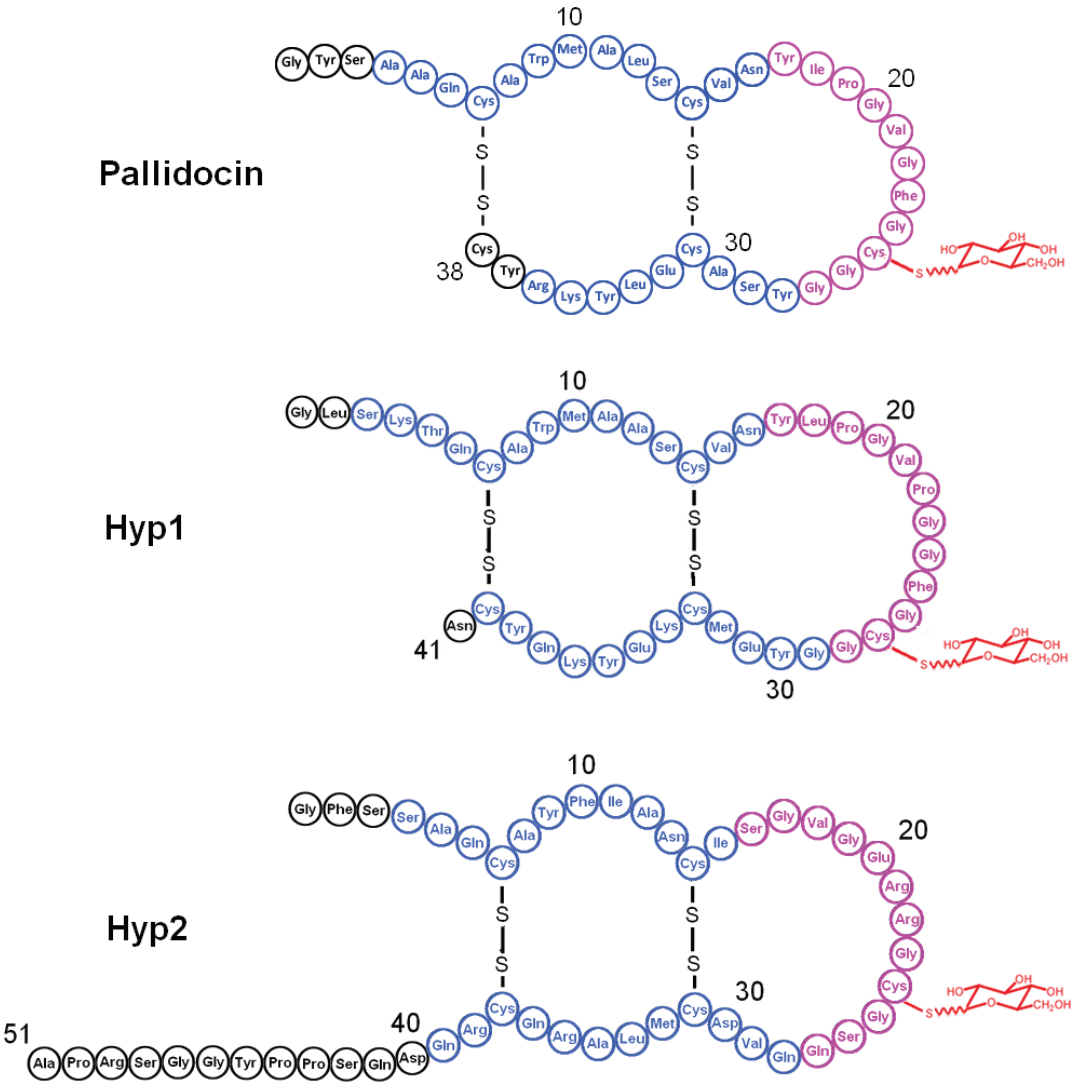


Figure 16. Proposed structures of native pallidocin, Hyp1 and Hyp2 glycocins. α-helical structure highlighted in blue color, coil structure highlighted in purple. The secondary structure predictions were made by PSIPRED program.

A wide range of temperatures and pH's were applied in order to evaluate stability of purified pallidocin (**Table 5**). It was incubated in buffers with diverse pH values and in NB medium at diverse temperatures.

Table 5. Effects of pH and temperature on the activity of pallidocin.

Temperature or pH treatment	Residual activity (%) after treatment
Room temperature, 24 h.	100
Room temperature, 10 days.	100
Room temperature, 30 days	12
60˚C, 3 h.	100
90˚C, 3 h.	100
121˚C, 15 min.	50
pH 2, 3 h.	100
pH 4, 3 h.	100
pH 6, 3 h.	100
pH 8, 3 h.	100
pH 10, 3 h.	100

The residual antibacterial activity after incubation was evaluated by agar well diffusion assay. The peptide was stable at room temperature for 10 days, after 30 days it retained 12% of the activity. 50% antibacterial activity retained after autoclaving pallidocin at 121°C for 15 min. Incubation at the pH range 2-10 did not affect the pallidocin antibacterial activity. This assay demonstrates high thermos-stability of pallidocin. It retains antibacterial activity even in a very broad range of pH values.

5.7 Functional assessment of the genes palS and palT

palS and *palT* were cloned and co-expressed with *palA-his* in *E. coli* BL21(DE3) to determine their functions. Gene *palA* was fused with a His7-tag coding sequence in the C-terminus (**Fig. 17**) to facilitate the purification. The highest yield of synthesized bacteriocin precursor peptides was observed in the insoluble fraction of cell lysate. Peptides were purified from the insoluble fraction, tested for activity and analyzed by MALDI-TOF MS to evaluate the presence of modifications (**Table 6**).

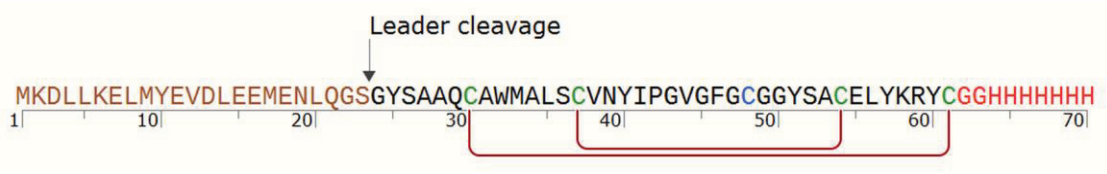


Figure 17. Engineered pallidocin precursor peptide PalA-His. Letters in brown indicates leader sequence, letters in black indicates core peptide, letters in red indicates engineered His7-tag sequence, letters in green and brown lines indicates disulfide bonds, letter in blue indicates glycosylated Cys residue.

Table 6. Results of glycocin precursors co-expression with *palS* and *palT*. Glc refers to peptide's modification with a mass of +162 Da, + refers to antibacterial activity of the purified peptide, NA – no activity, NM – no modifications, NLC – no leader cleavage.

	Glycocin precursor genes					
Co-expressed genes	*palA-his*	*sunA-his*	*hyp1-his*	*hyp2-his*	*enfA4-9-his*	*gccF-his*
No co-expression	NA, NM, NLC	NA, NM, NLC	NA, NM, NLC	NA, NM, NLC	NA, NM, NLC	NA, NM, NLC
Co-expression with *palS*	+, Glc, NLC	+, Glc, NLC	+, Glc, NLC	NA, Glc, NLC	NA, NM, NLC	NA, NM, NLC
Co-expression with *palT*	NA, NM, leader cleaved	NA, NM, NLC	NA, NM, NLC	NA, NM, NLC	NA, NM, NLC	NA, NM, NLC
Co-expression with *palST*	+, Glc, leader cleaved	+, Glc, NLC	+, Glc, NLC	NA, Glc, NLC	NA, NM, NLC	NA, NM, NLC

As expected, expression of only bacteriocin precursor gene *palA-his* resulted in synthesis of pre-PalA-His – the unmodified precursor peptide with leader (**Supplementary Fig. 4a**), and with no antibacterial activity. Co-expression of *palA-his* and *palS* genes resulted in the biosynthesis of pre-PalA-His-Glc – the precursor peptide with leader and a mass increment of 162 Da, which portrays glucosylation (**Supplementary Fig. 4b**). It was also active against the sensitive strain *P. genomospecies* 1 NUB36187. Co-expression of *palA-his* with *palT* resulted in the biosynthesis of the PalA-His core peptide with a mass corresponding to the unmodified leaderless peptide (**Supplementary Fig. 4c**); it also lacked antibacterial activity. Co-expression of the genes *palA-his* with *palS* and *palT* resulted in the biosynthesis of a PalA-His-Glc core peptide (mature pallidocin) with

the mass corresponding to the leaderless peptide with a mass increment of 162 Da, corresponding to glucosylation (**Supplementary Fig. 4d**). The glucosylated core peptide had high antibacterial activity (**Table 6**). The MALDI-TOF MS was not accurate and reproducible enough to unambiguously identify the small mass differences representing disulfide bonds in these peptides.

Tese experiments confirmed that PalS is S-glycosyltransferase which links the glucose to the Cys. Meanwhile PalT is bacteriocin peptidase that cleaves the leader sequence. Glucosylated pallidocin core peptide and glucosylated precursor peptide have antibacterial activity.

5.8 The role of disulfide bonds on antibacterial activity of post-translationally modified pallidocin precursor

To confirm the presence and importance of disulfide bonds in pallidocin, purified glycosylated precursor peptide with leader pre-PalA-His-Glc and leaderless PalA-His-Glc core peptide (mature pallidocin) were treated with TCEP and IAA. After the treatment the masses of pre-PalA-His-Glc and PalA-His-Glc core peptides were increased by +234 Da and +228 Da (**Table 7**), respectively, as all free Cys were alkylated, indicating that all disulfide bonds were reduced in the peptides.

Table 7. The role of disulfide bonds on activity of pallidocin precursors and core peptides. Pre-PalA-His – pallidocin precursor with leader and a His7-tag in the C-terminus. Pre-PalA-Glc – glycosylated pallidocin precursor with leader and the His7-tag. Pal-His-Glc – glycosylated pallidocin core peptide without leader and with the His7-tag (mature pallidocin). Legends: - peptide is not active, + peptide is active.

	pre-PalA-His		**pre-PalA-His-Glc**		**PalA-His-Glc (mature)**	
	Antibacterial activity	Cys alkylation	Antibacterial activity	Cys alkylation	Antibacterial activity	Cys alkylation
Native	-	-	+	-	+	-
Treated with IAA	-	-	+	-	+	-
Treated with TCEP and IAA	-	+	-	+	-	+

Notably, the disruption of disulfide bonds (based on TCEP and/or IAA experiments) results in the loss of antibacterial activity against the indicator *P. genomospecies* 1 NUB36187 (**Table 7**). After treatment only with IAA, no loss of antibacterial activity and mass increment was observed (**Table 7**), confirming that all disulfide bonds are present in both, pre-PalA-His-Glc and PalA-His-Glc core peptide.

These results idicates that synthesized precursor pre-PalA-His-Glc and PalA-His-Glc core peptide have disulfide bonds. The presence of disulfide bonds in these peptides is important for antibacterial activity.

5.9 Co-expression of *his-Xa-palA* and *palS*

To develop a strategy for more efficient pallidocin production, the genes *his-Xa-palA* and *palS* were co-expressed in *E. coli* BL21(DE3). The bacteriocin precursor PalA was engineered by adding a His6-tag in the N-terminus and a Factor Xa peptidase cleavage site (IEGR) in front of the core peptide (**Fig. 18**) for convenient leader removal. After the co-expression, glycosylated precursor peptide pre-His-Xa-PalA-Glc was purified and the leader was cleaved *in vitro* using Factor Xa. The generated leaderless PalA-Glc core peptide (mature pallidocin) was further purified with RP-HPLC followed by MALDI-TOF MS analysis for mass confirmation. The yield of synthesized active mature pallidocin (**Fig. 21**) was ~15 µg per 100 mL of bacterial culture.

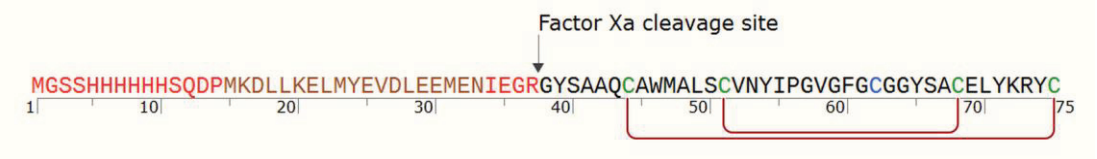


Figure 18. Engineered pallidocin precursor peptide His-Xa-PalA. Letters in red indicates engineered His6-tag and Factor Xa peptidase recognition site, letters in brown indicates leader sequence, letters in black indicates core peptide, letters in green and brown lines indicates Cys forming disulfide bonds, letter in blue indicates glycosylated Cys residue.

Activities of glycosylated precursor peptide with leader and mature pallidocin were compared by agar well diffusion assay. pre-His-Xa-PalA-Glc (10 ng/µL of peptide per 50 µL of sample) resulted in 80 AU/mL and mature pallidocin (0.5 ng/µL of peptide per 50 µL of sample) – 160 AU/mL. It indicates that the leader reduces the activity of the peptide and it is approximately 40 times lower than the activity of the core peptide.

5.10 Pallidocin alignment with other glycocins

BLAST analysis of PalA identified hypothetical peptides which have low sequence similarity to PalA (**Fig. 19**). We named these peptides Hyp1 (WP_061859978.1) and Hyp2 (WP_035395705.1). They were encoded in genomes of *B. megaterium* (Hyp1,) and *Bacillus* sp. JCM 19047 (Hyp2). Hyp1 consists of a 23-residue leader sequence and 41-residue core peptide, Hyp2 – a 20-residue leader sequence and 51-residue core peptide. Each of the peptides Hyp1 and Hyp2 have five Cys residues in the core sequence. These two putative novel glycocin precursors were investigated and examined further for possible posttranslational modifications by the biosynthetic machinery of pallidocin.

```
PalA        MKDLLKELMYEVDLEEMENLQGSGYSAAQCAWMALSCVNYIPGV--GF--GCGGYSACELYKRYC--------------    61
Hyp1        MNKNISKLMEEVSVEEMEQLQGKGLSKTQCAWMAASCVNYLPGVPGGF--GCGGYEMCKEYKQYCN-------------    64
Hyp2        MDNLLREIS-EEDLELYDG--GSGFSSAQCAYFIANCIS---GV--GERRGCGSQQVDCMLARQCRQDQSPPYGGSRPA    71
```

Figure 19. Pallidocin precursor peptide alignment with two putative novel glycocin precursor peptides Hyp1 and Hyp2. Conserved regions are highlighted in green color. **:** indicates predicted or experimentally determined leader cleavage sites.

Alignment of all glycocins (**Fig. 20**) showed that the leader sequences of SunA, ThuA, PalA, Hyp1 and Hyp2 share motifs of characteristic residues. In contrast, the leader sequences of GccF and EnfA49 display no significant similarities. All glycocins have a double-glycine type motif – proteolytic cleavage site at the end of the leader sequence. GccF and EnfA49 precursors have leader cleavage sites after the Gly-Gly motif. SunA, ThuA and PalA precursors have a Gly-Ser-Gly motif

where the leader sequence is cleaved between the Ser and second Gly residue. In case of peptide Hyp2, it has putative leader cleavage site with motif Gly-Ser-Gly, whereas, peptide Hyp1 has a – Gly-Lys-Gly.

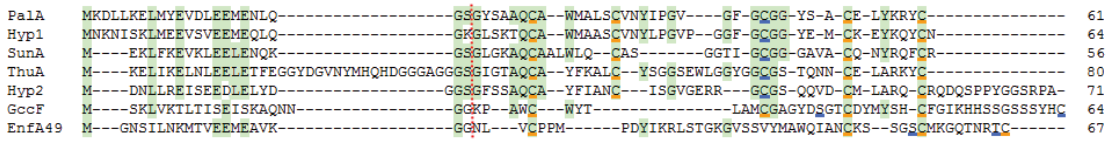


Figure 20. Alignment of glycocin precursor peptides. Conserved regions are highlighted in green color, Cys forming disulfide bonds are underlined in orange, glycosylated amino acids are underlined in blue. **:** indicates predicted or experimentally determined leader cleavage sites.

All glycocins characterized to date have 5 Cys residues in the core peptide and form two disulfide bonds between them. Hyp1 and Hyp2 core peptides have 5 Cys residues also and it is very probable that they form these bonds too. Early studies and the current one showed that the Cys residue in the conserved region Gly-Cys-Gly-Gly/Ser of SunA, ThuA and PalA is glucosylated. This motif is present in Hyp1 and Hyp2 peptides as well and it suggests that Cys in this region is the target for glycosylation.

The secondary structure prediction tool PSIPRED proposes that two α-helices are present in the PalA, Hyp1 and Hyp2 core peptides (**Fig. 16**). Moreover, the prediction that four out of five Cys residues reside in these helical structures is consistent with the NMR structures of sublancin 168 and glycocin F (Venugopal et al., 2011; Garcia De Gonzalo et al., 2014).

5.11 Application of the PalS and PalT enzymes for biosynthesis of other glycocins in *E.coli*

To determine whether PalS and PalT are able to modify and process different heterologous glycocin precursors, they were co-expressed with genes coding for glycocin precursor peptides with leaders (**Fig. 20**): pre-SunA-His, pre-GccF-His, pre-EnfA49-His, pre-Hyp1-His and pre-Hyp2-His. The peptides were engineered

by adding His6-tag sequence in the C-terminus to facilitate the purification. As before, the highest yield of the peptides was observed in the insoluble fraction of cell lysate. After co-expression the peptides were purified from the insoluble fraction and analyzed by MALDI-TOF MS (**Table 6**). The mass spectrometry analysis could not confirm that PalT cleaved off leaders. However, it demonstrated that the PalS glycosyltransferase does modify: pre-SunA-His, pre-Hyp1-His and pre-Hyp2-His, with a mass of +162 Da consistently pointing to monoglycosylation (**Supplementary Fig. 5, 6, 7**). Antibacterial activity analysis of the peptides showed that only pre-SunA-His-Glc and pre-Hyp1-His-Glc (glycosylated precursors with leaders) but not the pre-Hyp2-His-Glc, were active against indicator *P. genomospecies* 1 NUB36187 (**Table 6**).

We designed genes *core_hyp1-his* and *core_hyp2-his* coding for leaderless Hyp1 and Hyp2 core peptides with a His6-tag sequence at the C-terminus. As previously, *palS* was co-expressed with the new genes. The peptides were produced, purified and analyzed by mass spectrometry. LC-ESI-MS analysis detected masses of unmodified leaderless Hyp1-His and Hyp2-His core peptides (**Supplementary Fig. 8**), but glycosylated peptides were not observed. However, both of the peptides showed antibacterial activity against *P. genomospecies* 1 NUB36187 (**Fig. 21**).

Figure 21. Antibacterial activity of glycocins. P – PalA-Glc core peptide (mature pallidocin); 1 – Hyp1-His-Glc core peptide; 2 – Hyp2-His-Glc core peptide. Activity tested by spot on lawn assay. *P. genomospecies* 1 NUB36187 was used as a sensitive strain.

The observed activity suggests that Hyp1-His and Hyp2-His core peptides were partially glycosylated and exhibited antibacterial activity, but their glycosylation rate and their concentration was too low for detection by mass spectrometry. To confirm the importance of glycosylation, *core_hyp1-his* and *core_hyp2-his* genes

were induced in the expression host *E. coli*, while expression of *palS* was not present. The produced Hyp1-His and Hyp2-His core peptides did not show any antibacterial activity.

These results indicates that the leader plays an important role in the glycosylation. Furthermore, the glycosylation is critical factor determining the antimicrobial activity of Hyp1 and Hyp2.

5.12 The antibacterial activity spectrum of the novel glycocins and MIC determination for pallidocin

The antibacterial activity spectrum of purified pallidocin was tested against a number of gram-positive and gram-negative bacteria using a spot on lawn assay (**Table 8**). Purified pallidocin exhibited antibacterial activity against *B. cereus* ATCC 14579, *B. megaterium* DSM 319 and some thermophilic bacteria: *G. stearothermophilus* B4109, B4111, B4112, B4114 strains, *P. genomospecies* 1 NUB36187, *P. toebi* B4110, B4162 strains, *P. caldoxylosilyticus* B4119 and *C. debilis* B4165.

The novel synthesized glycocins Hyp1-His-Glc and Hyp2-His-Glc (glycosylated leaderless core peptides) were active against *Geobacillus* sp. B4113 and *G. stearothermophilus* B4163 strains, but exhibited no activity against *G. stearothermophilus* B4112, *P. toebi* B4110, *P. caldoxylosilyticus* B4119 and *B. cereus* ATCC 14579 strains, in contrast to pallidocin.

The MIC values were determined for some pallidocin susceptible strains. In liquid NB medium, the MICs of pallidocin against *B. megaterium* DSM 319 was 37 nM, *P. genomospecies* 1 NUB36187 – 2 pM (represents ±10 ng/L, the lowest MIC for a bacteriocin ever published), *G. stearothermophilus* B4114 – 246 pM, *G. toebi* B4162 – 493 pM and *P. caldoxylosilyticus* B4119 – 985 pM.

Table 8. Screening of glycocins against gram-positive and gram-negative bacteria strains. No inhibition zone defines -, observed inhibition zone defines +, weak inhibition zone defines ±, not applicable – NA.

Indicator strains	Pallidocin	Hyp1-His	Hyp2-His
Enterococcus faecalis LMG8222	-	-	-
Enterococcus faecalis LMG16216	-	-	-
Enterococcus faecium LMG16003	-	-	-
Enterococcus faecium LMG11423	-	-	-
Escherichia coli LMG15862	-	-	-
Escherichia coli LMG8223	-	-	-
Klebsiella pneumoniae LMG20218	-	-	-
Pseudomonas aeruginosa LMG6395	-	-	-
Staphylococcus aureus LMG15975	-	-	-
Staphylococcus aureus LMG8224	-	-	-
Staphylococcus aureus LMG10147	-	-	-
Listeria monocytogenes LMG10470	-	-	-
Acinetobacter baumannii LMG1041	-	-	-
Enterobacter aerogenes LMG2094	-	-	-
Enterobacter cloaceae subsp. cloaceae LMG2783	-	-	-
Samonella enterica subsp. enterica LMG7233	-	-	-
Geobacillus stearothermophilus B4109	+	±	±
Geobacillus stearothermophilus B4111	+	±	±
Geobacillus stearothermophilus B4112	+	-	-
Geobacillus stearothermophilus B4163	-	+	+
Geobacillus stearothermophilus B4114	+	+	+
Geobacillus stearothermophilus B4161	-	-	-
Geobacillus sp. B4113	-	+	+
Parageobacillus genomospecies 1 NUB36187	+	+	+
Parageobacillus toebi B4162	+	+	+
Parageobacillus toebi B4110	+	-	-
Caldibacillus debilis B4165	+	+	+
Bacillus cereus ATCC14579	+	-	-
Bacillus megaterium DSM319	+	+	+
Bacillus subtilis ATCC6633	-	NA	NA
Bacillus spp. (various isolates of laboratory)	-	NA	NA

Based on the results, it can be concluded that pallidocin, Hyp1-His-Glc and Hyp2-His-Glc acts on closely related *(Para)Geobacillus*, *Caldibacillus* and *Bacillus* spp. strains. The activity spectrum of the novel glycocins Hyp1-His-Glc and Hyp2-His-Glc was partially similar to the pallidocin spectrum. Pallidocin has very low MIC value, but it should be noted that some *Geobacillus* spp. strains used in the MIC assay easily acquired resistance for pallidocin antibacterial activity as observed by growing colonies in the inhibition area or the growth in some 96 plate wells with pallidocin concentration higher than the MIC.

5.13 Purification of bacteriocins produced by *G. stearothermophilus* 15

The *G. stearothermophilus* 15 was grown in NB medium and secreted bacteriocins (**Fig. 22**) were purified from the supernatant. Proteins were extracted by ammonium sulfate precipitation. Obtained crude bacteriocin extract (CBE) was further purified by ion-exchange chromatography (AEC and CEC). Partially purified bacteriocins were analyzed by Glycine-SDS-PAGE. After analysis the gels were stained with PageBlue Protein Staining Solution and overlaid with sensitive strain. It was determined that *G. stearothermophilus* 15 secretes two bacteriocins (**Fig. 23**). The molecular weights of the bacteriocins were approximately 18 kDa and 26 kDa. We named these bacteriocins geobacillin 18 (Geo18) and geobacillin 26 (Geo26), respectively. Geo18 was additionally purified by hydrophobic interaction chromatography (HIC) column and analyzed by Tricine-SDS-PAGE (**Fig. 24**).

Figure 22. Antibacterial activity in the supernatant determines by agar-well diffusion assay. It shows that *G. stearothermophilus* 15 secretes antibacterial compounds to the medium. Sensitive strain *P. genomospecies* 1 NUB36187.

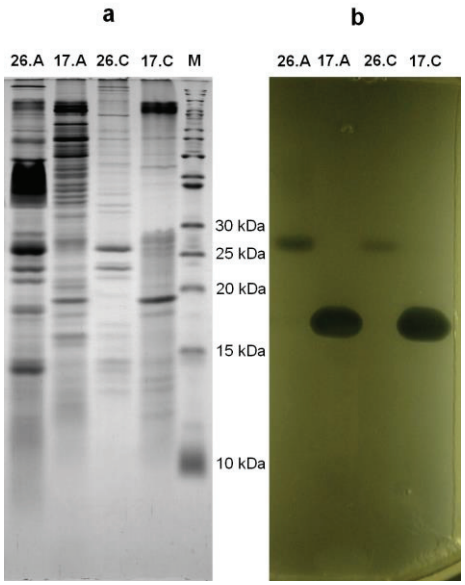


Figure 23. Glycine-SDS-PAGE analysis of bacteriocins produced by *G. stearothermophilus* 15: (**a**) gel stained with PageBlue Protein Staining Solution, (**b**) overlaid gel with indicator *P. genomospecies* 1 NUB36187. Line 26.A indicates flow-throw sample after purification on AEC column. Line 17.A indicates eluate after purification on AEC column. Line 26.C indicates eluate after purification on CEC column. Line 17.C indicates eluate after purification on CEC column. Line M indicates PageRuler Unstained Protein Ladder.

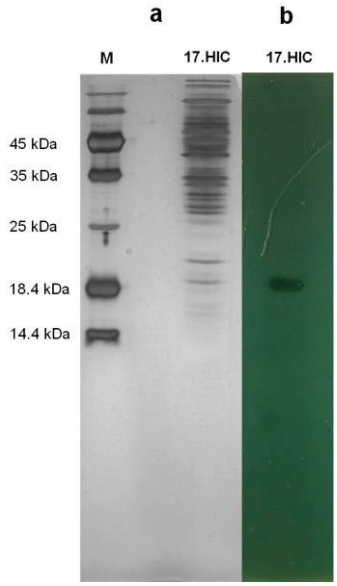


Figure 24. Trycine-SDS-PAGE analysis of Geo18 produced by *G. stearothermophilus* 15: (**a**) silver-stained gel, (**b**) overlaid gel with indicator *P. genomospecies* 1 NUB36187. Line 17.HIC indicates eluate after purification on HIC column. Line M indicates Pierce Unstained Protein MW Marker.

Protein band of Geo18 was detectable in the gel only after silver staining which subsequently was not compatible with MS analysis for protein identification. Unfortunately both bacteriocins were purified not completely, additional protein bands of impurities were visible in the stained gels.

5.14 Effects of temperature and enzymes on Geo18 and Geo26 activities

A wide range of temperatures and enzymes were applied in order to evaluate partially purified Geo18 and Geo26 stability (**Table 9 and 10**). Geo26 was incubated at diverse temperatures and it was revealed that it is stable up to 70°C temperature for 2 h. After incubation at 80°C, only 25% of its initial activity retained

and no activity was observed after incubation at 90°C. Effects of proteinase K, amylase and lipase were evaluated also. No antibacterial activity was observed after Geo26 incubation with proteinase K, meanwhile amylase and lipase did not affect its activity.

Geo18 retained its 50% activity after incubation at 70°C, and after autoclaving at 121°C for 15 min. Incubation at 60°C did not affect its activity. Proteinase K inhibits Geo18 antibacterial activity completely. Incubation with lipase and amylase retained only 12.5% of initial antibacterial activity.

Table 9. Effects of enzymes and temperature on the activity of Geo26.

Temperature or enzyme treatment	Residual activity (%) after treatment
60 ˚C, 2 h.	100
70 ˚C, 2 h.	100
80 ˚C, 2 h.	50
90 ˚C, 2 h.	0
Proteinase K, 2 h.	0
Amylase, 2 h.	100
Lipase, 2 h.	100

Table 10. Effects of enzymes and temperature on the activity of Geo18.

Temperature or enzyme treatment	Residual activity (%) after treatment
60 ˚C, 2 h.	100
70 ˚C, 2 h.	50
80 ˚C, 2 h.	50
90 ˚C, 2 h.	50
100 ˚C, 2 h.	50
121 ˚C, 15 min.	50
Proteinase K, 2 h.	0
Amylase, 2 h.	12.5
Lipase, 2 h.	12.5

These results indicates proteinaceous nature of geobacillin 26 and geobacillin 18. Geobacillin 26 is heat-labile antibacterial protein, meanwhile geobacillin 18 demonstrates stability in high tempertures.

5.15 Determination of pI values of bacteriocins Geo18 and Geo26

Both bacteriocins were analyzed by isoelectric focusing (IEF) to determine their isoelectric points (pI). They were separated by protein pI on IPG strips, which were subsequently overlaid with the sensitive strain and incubated in a thermostat. After the incubation, the overlaid IPG strip was compared to stained IPG strip with protein standards (**Fig. 25**).

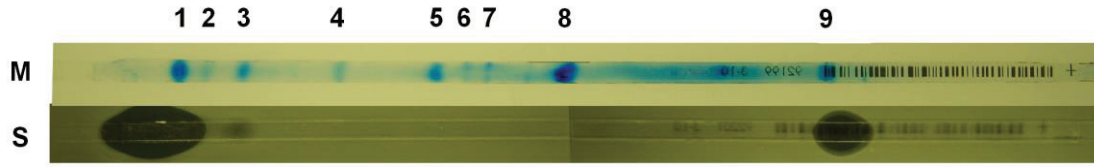


Figure 25. Geo18 and Geo26 analysis on IPG strips. Part M indicates IPG strip with pI standards: 1 – pI 8.2, 2 – pI 8.0, 3 – pI 7.8, 4 – pI 7.5, 5 – pI 7.1, 6 – pI 7.0, 7 – pI 6.8, 8 – pI 6.0, 9 – pI 5.1. Part S indicates IPG strip with samples of purified Geo17 and Geo26 and overlaid with indicator *P. genomospecies* 1 (NUB36187).

Inhibition zones above the strip indicates that Geo26 has pI higher than 8.2, and Geo18 has pI lower than 5.1. Additionally, after the IEF analysis, the strips were analyzed in a second dimension by SDS-PAGE (**Fig. 26**).

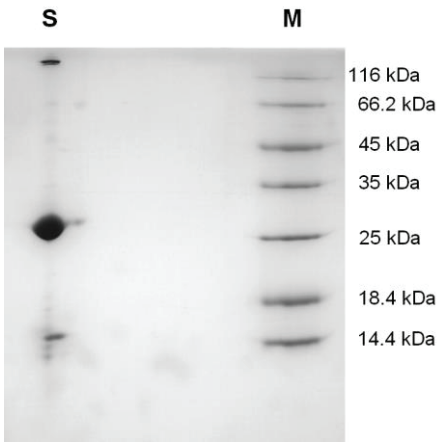


Figure 26. Second dimension SDS-PAGE analysis of IPG strip with protein Geo26. Line S indicates the part of IPG strip corresponding to the inhibition zone and containing the Geo26 protein. The protein band was cut from the gel for MS/MS analysis. Line M indicates PageRuler Unstained Protein Ladder.

Subsequently, the gels were stained by colloidal coomassie staining solution and discrete 26 kDa protein band of bacteriocin was cut and submitted for mass spectrometry analysis. Concentration of 18 kDa bacteriocin was to low and it was not visible in the stained gel.

5.16 MS/MS analysis of Geo26 and its gene identification

Mass spectrometry analysis was used to determine *de novo* amino acid sequence of the Geo26. The cut gel slice with the Geo26 protein (**Fig. 26**) was digested with trypsin and analyzed by MS/MS method. 26 peptides and their sequences were revealed (**Supplementary Table 2**). Each peptide's sequence was analyzed by the BLAST method in the NCBI database. The peptides matched the sequence of a hypothetical proteins encoded in genomes of various *Geobacillus, Bacillus* spp. strains including the *G. stearothermophilus* 15 (**Fig. 27, 28**). Geo26 sequence had no similarities to proteins with known function.

The gene is located in a contig of genome sequence of strain 15 with accession number LVHZ01000156.1, between base pairs 1529-2248. There were no genes flanking the *geo26* and coding for a signal sequence cleavage enzyme, transporter proteins or other post-translational modification enzymes (**Fig. 29**).

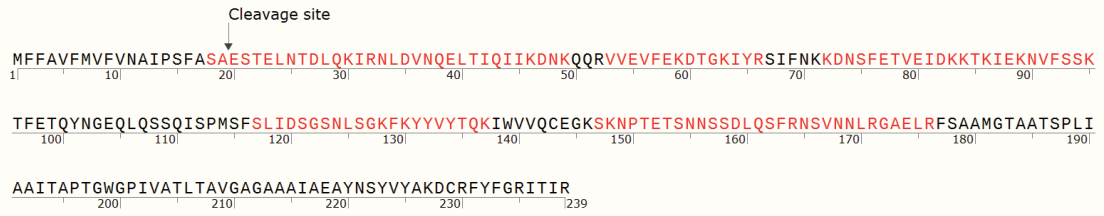


Figure 27. Geo26 protein sequence encoded in the genome of *G. stearothermophilus* 15. Peptide sequences revealed by MS/MS indicated in red. Cleavage site of the signal peptide indicated with an arrow.

```
Geo26      1 -----MKKIFLMFFAVFMVFVNAIPSFASAEST----ELNTDLQKIRNLDV--------NQELTIQIIKD
A          1 -----MKKFFLMFFAVLILFGTNVSLFARAETT----ESSLEMQKIENLDA--------NKDVKIEIIKD
B          1 -----MKRFLLMFFAALVFFGTNASLFANAETT----DSSLEMQKIENLDA--------NKDINIEIIKD
C            ------------------------------------------------------------------------
D          1 MSLNRKLKHAGLGMALTLSATTLIPNTSLAETSVKKYESVSELKSSDNIGKVETTSMPGHQDVHSKIIED

Geo26     54 NKQQRVVEVFEKDTGKIYRSIFNKKDNSFETVEIDKKTKIEKNVFSSKTFETQYNGEQLQ-SSQISPMSF
A         54 NKQQRVVQVVDKQTGKIYKSTFNKKDNSFQTVELDSKSKLEKDVLNSKTIESQNSTAQLQ-SAQVASASF
B         54 NKQQRVVQVVDKQTGKVYKSIFNKKDNSFQTVEVDLKSKVENNVLNSKTLELQSSTMQLQ-SAQMVSSSF
C          1 ---------------------------------MKSKVENNVLNSKALELQSSAMQLQ-SAQMVSSSF
D         71 NKEKRVVETWDNEGN--YRATFYKDRNEVVTETLD-----SNGVVLSKTSTAQTQVKPLKDNTQYNRPSI

Geo26    123 SLIDSGSNLSGKFKYYVYTQKIWVVQCEGKSKNPTETSNNSSDLQSFRNSVNNLRGAELRFSAAMGTAAT
A        123 SLIDSGSNVSGRFKYYVYTQKIWVVQCDGQSKNPVETSNNSSDLQSFRNTVNNLRGSEIRLAAALGTALT
B        123 SLIDSGSNLSGKFKYYAYTQKIWVVQSEGESKNPIETSNNSSDLQSFRNSVNNLRGSEIRFQAAWVTGVS
C         35 SLIDSGSNLSGKFKYYAYTQKIWVVQSEGESKNPIETSNNSSDLQSFRNSVNNLRGSEIRFQAAWVTGVS
D        134 ELVDSGQDWLGRYKYYFYTENVWVIQIPGRTKNPIQSANNKTDLGAFRTAVNNLRVEQIGLTKAMGGVAV

Geo26    193 SPLIAAITAPTGWGPIVATLTAVGAGAAAIAEAYNSYVYAKDCRFYFGRITIR
A        193 STVVAAVTAPTGWGPIVATLTAIGAGATAIAEGYNSYVLAKDCRFYFGRITLR
B        193 STMVAVITSPTGWGPIVATLTAAGATAAAIAEAYNSYMLAKDCRFYFGRITLR
C        105 STMVAVITSPTGWGPIVATLTAAGATVAAIAEAYNSYMLAKDCRFYFGRITLR
D        204 DAAVAVATVPTGWPAVLAGLKTLGAGADALIAGEKVNNLTKDCTYNFQQVSLR
```

Figure 28. Geo26 alignment with proteins, which have sequence similarities. A, B, C ir D – hypothetical proteins identified in *Anoxybacillus* sp. *G. thermocatenulatus*, *G. kaustophilus* and *B. thuringiensis* bacteria genomes, respectively.

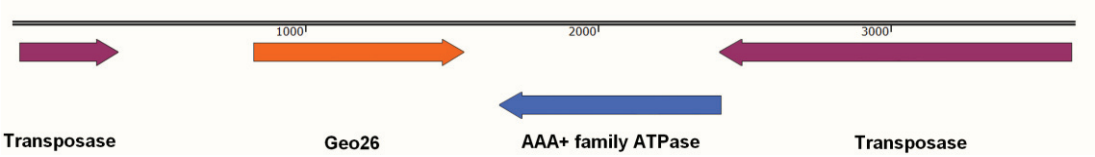


Figure 29. Genomic context of *geo26*.

The protein was 239 amino acid length, calculated molecular weight based on the sequence – 26.6 kDa and pI – 9.18. The protein sequence analysis by SignalP 4.1 Server observed 19 amino acid length signal peptide. Its predicted cleavage site – Ala-Ser-Ala-Glu, before the Glu. Calculated molecular weight of the mature protein, without the signal peptide – 24.5 kDa, with the pI of 9.18.

These values agree with those identified by SDS-PAGE and IEF analyses and it can be concluded that Geo26 is novel high-molecular-weight bacteriocin. Its genomic context indicates no genes that could assist in the biosynthesis of the bacteriocin.

5.17 Cloning and expression of Geo26

Gene *geo26* encoded in *G. stearothermophilus* 15 was cloned to expression vector pBAD24, transformed to *E. coli* BL21(DE3) host and induced for protein expression. After the induction, antibacterial activity was observed only in the cell lysate, but not in the supernatant (**Fig. 30a**).

Additionally, *geo26* was cloned to expression vector pNZ8048 under the control of nisin promoter. The construct pNZ8048-geo26 was transformed to *L. lactis* NZ9000 host and induced for protein biosynthesis by supplementing the culture with nisin. After the induction, antibacterial activity was observed only in the cell lysate, but not in the supernatant (**Fig. 30b**).

Figure 30. Antibacterial activity of cell lysate after geo26 expression in: (**a**) *E. coli* BL21(DE3), (**b**) *L. lactis* NZ9000. 1 – not induced cells (negative control), 2 – induced cells. Indicator strain *P. genomospecies* 1 NUB36187, the activity tested by agar-diffusion assay.

A His6-tag sequence was introduced in N-terminus of Geo26 and its gene *his-geo26* was cloned to pBAD24 vector. After gene expression in *E. coli* BL21(DE3), soluble and insoluble fractions of cell lysate was loaded on IMAC column. Unfortunately, the protein did not attach to the column and it was not purified.

6 Discussion

We identified and characterized a new post-translationally modified bacteriocin – pallidocin. The gene cluster of the pallidocin biosynthetic machinery is encoded in a thermophilic bacterium, i.e. the *A. pallidus* 8, which is unprecedented. The characterized bacteriocin belongs to the relatively new and small subclass of glycocins (Class I – post-translationally modified bacteriocins) and shares genetic

and structural similarities with sublancin 168 (Oman et al., 2011), glycocin F (Stepper et al., 2011), thurandacin (Wang et al., 2013) and enterocin F4-9 (Maky et al., 2015) gene clusters (**Supplementary Fig. 9**).

For the first time it was shown that a glycocin biosynthetic gene cluster can be cloned and functionally expressed in a heterologous host – *E. coli*. Surprisingly, this bacteriocin from gram-positive bacteria was synthesized and secreted in this gram-negative bacterium. Structural characterization of the purified recombinant pallidocin revealed posttranslational modifications – glycosylation and two disulfide bonds within two predicted α-helices. Taking into account other glycocin structures, the letter two features seem to be specific for the class of glycocins (Norris and Patchett, 2016). Pallidocin has an S-linked glucose moiety to a Cys residue, which is very uncommon among bacteria. Only a few cases of S-linked glycopeptides have been described and confirmed to date, i.e. for glycocin F, sublancin 168 and thurandacin (Oman et al., 2011; Stepper et al., 2011; Wang et al., 2013). Pallidocin exhibits stability after exposure to high temperatures and a wide range of pH.

Experiments when only two or three genes were co-expressed revealed that *palS* encodes an S-glycosyltransferase, which introduces glucose to the Cys25 residue in pallidocin. When *palA-his* was co-expressed with *palS* and *palT* then the mature pallidocin (PalA-His-Glc core peptide) was purified also. Additionally, the PalT protein shares sequence similarity to other bacteriocin ABC transporters which leads to the conclusion that PalT has a dual, peptidase and transport, function.

In contrast to previous *in vitro* synthesis of sublancin and thurandacin (Oman et al., 2011; Wang et al., 2013; Wang and van der Donk, 2011), we generated a glycosylated pallidocin precursor and mature pallidocin *in vivo*. It resulted in a peptide with already formed disulfide bonds, which, as we showed, are necessary for antibacterial activity. We demonstrate biosynthesis of novel glycocins *in vivo* in another (gram-negative) host, surpassing enzymatic *in vitro* glycosylation of the peptide and the necessity of chemical formation of disulfide bonds *in vitro*.

We applied PalS for possible modifications of other glycocins and produced two novel glycocins: Hyp1 and Hyp2. PalS specifically monoglycosylated a certain group of glycocin precursors with leaders – pre-PalA-His, pre-SunA-His, pre-Hyp1-His and pre-Hyp2-His, but it was not able to modify pre-GccF-His and pre-EnfA4-9-His. The distinctive feature of the glucosylated peptides is a Cys residue flanked by glycines (Gly-Cys-Gly-Gly/Ser) in the interhelical loop. Keeping in mind that PalS, ThuS and SunS form a sugar S-linkage with a Cys residue in Gly-Cys-Gly-Gly/Ser motif of interhelical loop, and that this motif is present only in SunA, ThuA, PalA, Hyp1 and Hyp2, we can distinguish and designate this group of glycocins as the *sublancin type*.

Previous research on sublancin S-glycosyltransferase, SunS, suggested that this enzyme recognizes an α-helical segment of its substrate and then glycosylates only a Cys residue in a flexible loop following this helix (Wang and van der Donk, 2011). Studies on thurandacin glycosyltransferase, ThuS, showed that it glycosylates its peptide substrate ThuA at Cys28 or both Ser19 and Cys28. Because of that, ThuS represents the first glycosyltransferase that catalyzes both O- and S-glycosylation of proteins. Earlier studies also demonstrated that SunS is not able to modify ThuA, although ThuS is able to modify the SunA peptide generating its bisglucosylated product. Moreover, SunA and ThuA with changed short sequences in their interhelical loops were also glucosylated by ThuS in these regions. On basis of this knowledge, it was suggested that the peptide sequence selectivity of ThuS is much more relaxed than that of SunS (Wang et al., 2013). Here, we demonstrate that PalS has a quite flexible substrate selectivity too and that it may monoglycosylate precursor peptides with leader: pre-PalA-His, pre-SunA-His, pre-Hyp1-His and pre-Hyp2-His.

We indicated that disruption of disulfide bonds in glycosylated precursor peptide pre-PalA-His-Glc or PalA-His-Glc core peptide leads to the loss of antibacterial activity. It supports suggestions of earlier studies on sublancin (Wang and van der Donk, 2011) that Cys modification in the interhelical loop facilitates correct

disulfide bonds formation *in vivo* between the four Cys located in two helices, which are stabilized by these bonds. They might be formed by the host's disulfide bond formation enzymes or spontaneously. Our observations also agree with indications and assumptions in other studies on sublancin (Oman et al., 2011) that the type of sugar attached to Cys22 in sublancin is not important for bioactivity and the glycosylation may constitute an unusual type of self-resistance of the bacteriocin producer.

In earlear studies on sublancin and thurandacin, it has been observed that the leader must be cleaved off to gain antibacterial activity (Oman et al., 2011, Wang et al., 2013). In contrary, our synthesized glycosylated precursors with leaders: pre-PalA-His-Glc, pre-SunA-His-Glc and pre-Hyp1-His-Glc, were active against a sensitive strain. It suggests that the leader sequence cleavage of glycosylated glycocin precursor is not absolutely necessary for activity. In fact, glycocins with leader attached exhibit antibacterial activity. However, the absecnce of the leader increases the activity. It should be noted that in contrary to previous studies on glycocins, we have used the thermophilic indicator strain, which as we show exhibits extreme susceptibility, even to leader containing glycocin.

In addition, we show that leaderless Hyp1-His and Hyp2-His core peptides can be modified by PalS, resulting in active antibacterial peptides. However, the processivity of PalS is lower when compared to leader containing pre-Hyp1-His and pre-Hyp2-His. Glycosylated precursor peptides with the leader were detectable by mass spectrometry, and, additionally one of them – pre-Hyp1-His-Glc, showed antibacterial activity. Glycosylated core peptides were not detectable by the mass spectrometry but their antibacterial activity was observed. These observations are very similar to leaderless SunA core peptide glycosylation by ThuS, also resulting in a very low yield (Wang et al., 2013). The role of pallidocin leader is the enhancement of PalS processivity and the decrement of antibacterial activity of glycocin while it is inside the producer cell.

Pallidocin exhibits antimicrobial activity against gram-positive bacteria. Most of the tested thermophilic bacteria exhibited susceptibility for pallidocin. Mesophilc *B. megaterium* DSM 319 and *B. cereus* ATCC 14579 strains were susceptible also. It indicates a rather narrow activity spectrum restricted to closely related bacteria. Enterocin F4-9 exhibits antimicrobial activity against *E. faecalis* and *E. coli* JM109 (Maky et al., 2015). Glycocin F has narrow activity spectrum also and restricted to the *Lactobacillus* genus (Stepper et al., 2011). Notably, sublancin is active against several gram-positive bacteria, like *Staphylococcus* and, especially *Bacillus* species (Dubois et al., 2009). Pallidocin MIC values for thermophilic bacteria are extremely low when compared to the values for *B. megaterium* DSM 319 and *B. cereus* ATCC 14579. Notably, glycosylated leaderless Hyp1 and Hyp2 core peptides demonstrated a different activity spectrum against thermophilic bacteria as compared to pallidocin. As in case of sublancin, the susceptible strains easily develop resistance to pallidocin.

Studies on sublancin and thurandacin have generally been focused on *in vitro* analysis and on glycosyltransferase as potential tool for antibody generation and other purposes. S-linked glycopeptides are highlighted to be more chemically and biologically stable compounds than O-linked glycopeptides (Oman et al., 2011; Wang et al., 2013). Here, we identified and characterized the new antibacterial peptide pallidocin and expanded the currently small glycocin family. In addition, we demonstrated that glycosyltransferase PalS could be used as a tool for the biosynthesis of novel glycosylated antibacterial peptides *in vivo*. We have developed a system for heterologous expressions and screening of novel *sublancin type* precursors with a minimal number of genes required. Pallidocin or Hyp1 and Hyp2 could be applied in industry processes facing thermophilic bacteria contaminations.

The *G. stearothermophilus* 15 is producing two antibacterial compounds which we named geobacillin 18 (Geo18) and geobacillin 26 (Geo26). Using chromatographic methods they were purified from the culture supernatant. SDS-PAGE analysis

revealed that Geo18 is about 18 kDa and Geo26 is 26 kDa molecular weight. Treatment by proteolytic enzyme confirmed that both compounds are proteinaceous nature and can be referred as bacteriocin like inhibitory compounds (BLIS). Antibacterial activity decrement of Geo18 after the treatment with lipase and amylase may indicate that it contains lipid or carbohydrate moieties. It should be noted that, there are speculations explaining the decrement of BLIS activity after the treatment with lipase or amylase. These enzymes may be not completely pure from the proteolytic enzymes. Small residual amounts of proteolytic enzymes in the powder of lipase or amylase may be present (Norris and Patchett 2014). Geo18 showed high stability in elevated temperatures, meanwhile Geo26 was heat-labile and was completely inactivated by treatment in 80˚C temperature. Geo26 was submitted to mass spectrometry analysis and the amino acid sequence was revealed. The gene encoding Geo26 has no sequence similarity to any known enzymes or characterized bacteriocins. Cloning and heterologous expression of gene *geo26* confirms its antibacterial properties.

Proteinaceous nature, heat sensitivity and determined amino acid sequence of Geo26 indicates that it is novel high-molecular-weight (class III) bacteriocin. There should be done more extensive work and experiments on heterologous expression, purification and characterization of this recombinant bacteriocin.

Amino acid sequence of the Geo18 was not determined, because it was purified not completely and the amount of the compound in the sample was too low for mass spectrometry analysis. The lack of information about the amino acid sequence of Geo18 let us to define it only as a BLIS, which has resistance to the heat treatment. This property is very valuable in biotechnological applications.

7 Conclusions

- Identified novel poist-translationally modified bacteriocin pallidocin, which is encoded in thermophilic *A. pallidus* 8 and belongs to a glycocin subclass.

- After expression of bacteriocin biosynthetic gene cluster derived from the thermophilic bacteria, the first time ever a complete biosynthesis and secretion of glycocin was achieved in heterologous host *E. coli.*

- Pallidocin is active against thermophilic *(Para)Geobacillus*, *Caldibacillus* and mesophilic *Bacillus* genus bacteria. It has extremely low MIC value against *P. genomospecies* 1 NUB36187 – 2 pM. Pallidocin is stable at 121°C temperature and pH range of 2-10.

- Palidocin S-glycosyltransferase PalS glycosylates sublancin precursor and two novel glycocin precursors: Hyp1 and Hyp2. It can also glycosylate Hyp1 and Hyp2 core peptides. Glycosylated Hyp1 and Hyp2 core peptides expose activity against thermophilic *(Para)Geobacillus*, *Caldibacillus* and mesophilic *Bacillus* genus bacteria.

- Identified novel bacteriocin geobacillin 26, produced by thermophilic *G. stearothermophilus* 15. It is 24.5 kDa molecular weight and is fully inactivated after incubation at 90°C temperature. Its amino acid sequence shares no significant similarities to any known function protein sequences.

8 Publications and conferences

The results of this study have been published in papers:

- Pranckutė R., **Kaunietis A.**, Kananavičiutė R., Lebedeva J., Kuisienė N., Šaleikienė J., Čitavičius D. 2015. Differences of antibacterial activity spectra and properties of bacteriocins, produced by *Geobacillus* sp. bacteria isolated from different environments. J. Microbiol. Biotechnol. Food. Sci. 5(2):155-161. doi: 10.15414/jmbfs.2015.5.2.155-161.
- **Kaunietis A.**, de Jong A., Pranckutė R., Buivydas A., Kuipers O.P. 2016. Draft genome sequences of two *Geobacillus* species strains, isolated from oil wells and surface soil above oil pools. Genome Announc. 4(5):e01129-16. doi: 10.1128/genomeA.01129-16.
- **Kaunietis A.**, Buivydas A., Čitavičius D., Kuipers O.P. Heterologous biosynthesis and characterization of a novel glycocin from a thermophilic bacteria. Nature Communications (*manuscript under consideration*).

Other publications:

- Lastauskienė E., Zinkevičienė A., Girkontaitė I., **Kaunietis A.**, Kvedarienė V. 2014. Formic acid and acetic acid induce a programmed cell death in pathogenic *Candida* species. Curr. Microbiol. 69(3):303-10. doi: 10.1007/s00284-014-0585-9.
- Pranckutė R., **Kaunietis A.**, Kuisiene N., Čitavičius D. 2014. Development of synbiotics with inulin, palatinose, α-cyclodextrin and probiotic bacteria. Polish J. Microbiol. 63(1):33-41.
- Pranckutė R., **Kaunietis A.**, Kuisiene N., Čitavičius D.J. 2016. Combining prebiotics with probiotic bacteria can enhance bacterial growth and secretion of bacteriocins. Int. J. Biol. Macromol. 89:669-676. doi: 10.1016/j.ijbiomac.2016.05.041.

The results of this study have been presented as posters in international conferences:

- **Kaunietis A.**, Pranckutė R., Čitavičius D. Optimization of the secretion of bacteriocins and the bacterial cell growth of *Geobacillus* sp. 15 strain. The 6th Congress of European Microbiologists (FEMS), Maastricht, the Netherlands, 7-11 June 2015.
- **Kaunietis A.**, de Jong A., Buivydas A., Kuipers O.P. Identification, biosynthesis and characterization of a novel bacteriocin encoded in the genome of thermophilic bacteria *Aeribacillus pallidus*. The 7th Congress of European Microbiologists (FEMS), Valencia, Spain, 9-13 July 2017.

9 Financial support

- Research Council of Lithuania (RCL) doctoral scholarship (2013 – 2014).
- Erasmus+ EU grant (2015 – 2016).

10 CV

Personal information

Name, surname: ARNOLDAS KAUNIETIS

Date of birth: 01-02-1987

Address: S. Konarskio str. 22-28, LT-03124, Vilnius, Lithuania

E-mail: kavenis1@gmail.com

Phone: +37068531352

Education

2010 achieved bachelor's degree of biology in the field of Molecular Biology at Vilnius University.

2012 achieved master's degree of microbiology at Vilnius University.

Conferences and publications

Presented posters at international conferences:

- **Kaunietis A.**, Pranckutė R., Čitavičius D. Optimization of the secretion of bacteriocins and the bacterial cell growth of *Geobacillus* sp. 15 strain. The 6th Congress of European Microbiologists (FEMS), Maastricht, the Netherlands, 7-11 June 2015.
- **Kaunietis A.**, de Jong A., Buivydas A., Kuipers O.P. Identification, biosynthesis and characterization of a novel bacteriocin encoded in the genome of thermophilic bacteria *Aeribacillus pallidus*. The 7th Congress of European Microbiologists (FEMS), Valencia, Spain, 9-13 July 2017.

List of publications:

- Pranckutė R., **Kaunietis A.**, Kananavičiutė R., Lebedeva J., Kuisienė N., Šaleikienė J., Čitavičius D. 2015. Differences of antibacterial activity spectra and properties of bacteriocins, produced by *Geobacillus* sp. bacteria isolated

from different environments. J. Microbiol. Biotechnol. Food. Sci. 5(2):155-161. doi: 10.15414/jmbfs.2015.5.2.155-161.

- **Kaunietis A.**, de Jong A., Pranckutė R., Buivydas A., Kuipers O.P. 2016. Draft genome sequences of two Geobacillus species strains, isolated from oil wells and surface soil above oil pools. Genome Announc. 4(5):e01129-16. doi: 10.1128/genomeA.01129-16.
- Lastauskienė E., Zinkevičienė A., Girkontaitė I., **Kaunietis A.**, Kvedarienė V. 2014. Formic acid and acetic acid induce a programmed cell death in pathogenic *Candida* species. Curr. Microbiol. 69(3):303-10. doi: 10.1007/s00284-014-0585-9.
- Pranckutė R., **Kaunietis A.**, Kuisiene N., Čitavičius D. 2014. Development of synbiotics with inulin, palatinose, α-cyclodextrin and probiotic bacteria. Polish J. Microbiol. 63(1):33-41.
- Pranckutė R., **Kaunietis A.**, Kuisiene N., Čitavičius D.J. 2016. Combining prebiotics with probiotic bacteria can enhance bacterial growth and secretion of bacteriocins. Int. J. Biol. Macromol. 89:669-676. doi: 10.1016/j.ijbiomac.2016.05.041.
- **Kaunietis A.**, Buivydas A., Čitavičius D., Kuipers O.P. Heterologous biosynthesis and characterization of a novel glycocin from a thermophilic bacteria. Nature Communications (*manuscript under consideration*).

Traineeships

Department of Molecular Genetics, Groningen Biomolecular Sciences and Biotechnology Institute, Faculty of Science and Engineering, University of Groningen (01-11-2015 – 01-05-2017).

Financial support

- Research Council of Lithuania (RCL) doctoral scholarship (2013 – 2014).
- Erasmus+ EU grant (2015 – 2016).

11 Supplementary information

11.1 Supplementary Tables

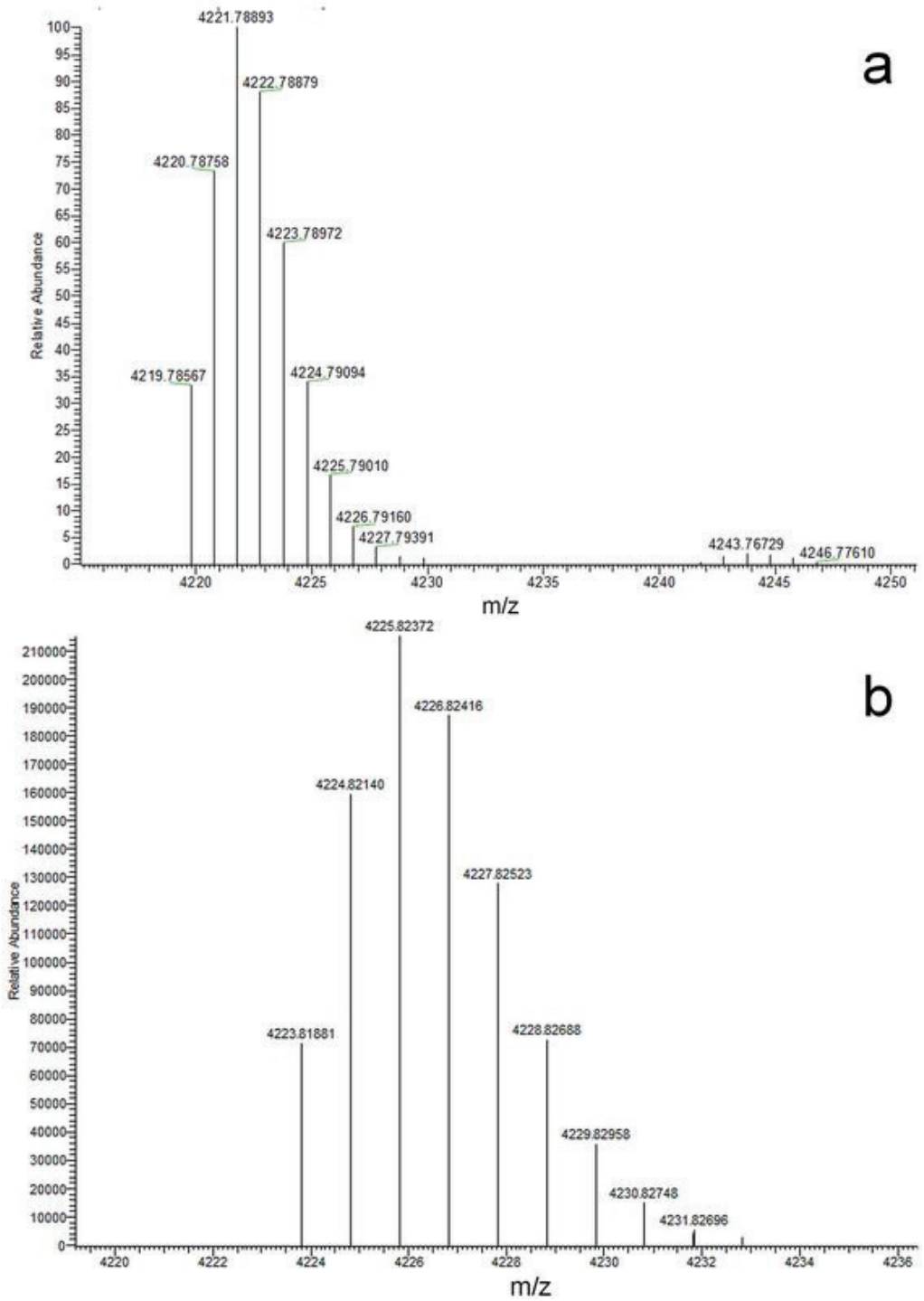
Supplementary Table 1. Primers used for PCR.

Primer name	Primer sequence (5' to 3')
F-PalA-USER	ATTCACCAUGAAAGATTTATTAAAAGAATTAATGTATGAG
R-PalA-USER	AAACAGCCUCGGTATATCGGTTTCTTTTTTTCATG
F-pBAD-USER	AGGCTGTTUTGGCGGATGAG
R-pBAD-USER	ATGGTGAAUTCCTCCTGCTAG
F-PalA-BspHI	ATTCGTCATGAAAGATTTATTAAAAGAATTAATGTATGAG
R-PalA-HindIII	TTTTAAAGCTTAACAATAACGTTTATACAATTCGC
F-PalS-In-Fusion	CAGGAGGAATTCACCATGGGGAACTTAAGAGATTTTTATCAAC
R-PalS-In-Fusion	CAAAACAGCCAAGCTTGTTAAATTTTATTTATACTATCTATAAAA CTTAATAGATTATC
F-PalT-In-Fusion	CAGGAGGAATTCACCATGATTTTAAGGAAATTTGCACATGTAAG AC
R-PalT-In-Fusion	CAAAACAGCCAAGCTTGTCATAGGACCTCTACCTCCAGATTTC
F-PalA-His	CATCATCATCATCATTAAGCTTGCGGCCGCATAATG
R-PalA-His	ATGATGTCCTCCACAATAACGTTTATACAATTCGCAAG
F-Hyp1-NisP	GGTCTGAGCAAAACCCAATGC
R-Core_Hyp-His	CATGGTATATCTCCTTATTAAAGTTAAACAAAATTATTTCTACAG GG
F-Hyp2-NisP	GGTTTCAGCAGCGCGCAGT
F-PalA-BamHI	GAAAAGGATCCGATGAAAGATTTATTAAAAGAATTAATGTATGA GGTAGATTTAGAAGA
R-PalA-HindIII[2]	TAATTTAAGCTTAACAATAACGTTTATACAATTCGCAAGCGC
F-PalA-Xa	GGATATTCAGCTGCCCAATGTGCATG
R-PalA-Xa	ACGACCTTCGATATTCTCCATTTCTTCTAAATCTACCTCATAC
F-Geo26	TTTTACATGTTTTTTGCTGTGTTTATG
R-Geo26	ATTACAAGCTTAACGAATAGTAATTCTACCAAAG
F-His-Geo26	TTTTTTGCTGTGTTTATGGTCTTTGTAAATGCT
R-His-Geo26	GTGGTGATGGTGATGATGCATGGTGAATTCCTCCTGCTAGCC
27F	AGAGTTTGATCMTGGCTCAG
1492R	TACGGYTACCTTGTTACGACTT

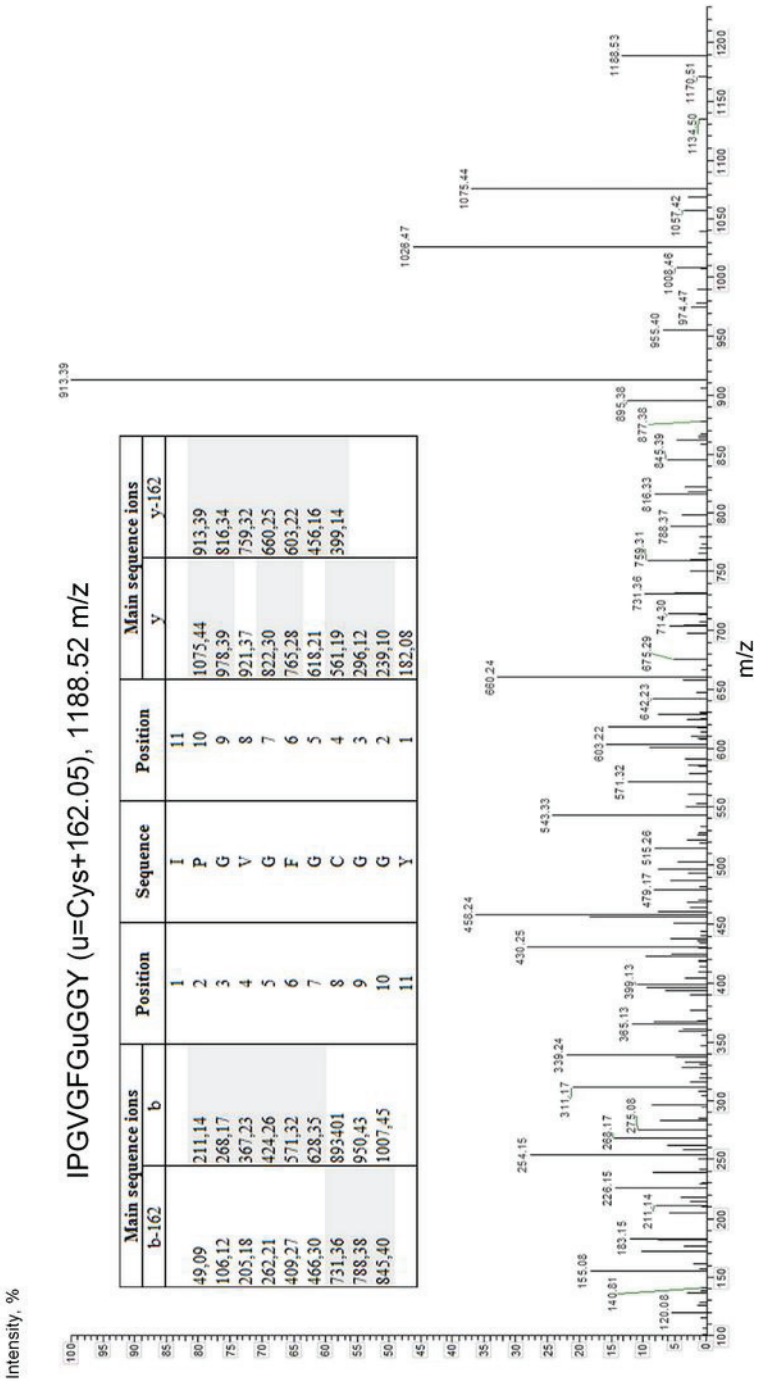
Supplementary Table 2. Peptide sequences of digested Geo26 identified by MS/MS analysis.

SLFDKK	FKYYVYTQK	TKLEKNVFSSK	LWVVQACEGK
SLVHPSY	VVEVFEKDTGK	DNSFETVELDKKTK	TLNNDLMLLK
NSVNNLR	ESTELNTDL(AG/Q)K	NSVNNLR(GA/Q)ELR	SLLDSGSNLSGK
VVEVFEK	VVEVFEKDTGKLYR	(DL/EV)CDETVELDKK	ESTELNTDLQKLR
SSDLQSFR	SAESTELNTDLQKLR	DNSFETVELDKK	KDNSFETVELDKK
YYVYTQK	NPTETSNNSSDLQSFR	SAESTELNTDLQK	
LKSAASLNSR	NLDVNQELTLQLLKDNK	LEKNVFSSK	

11.2 Supplementary Figures

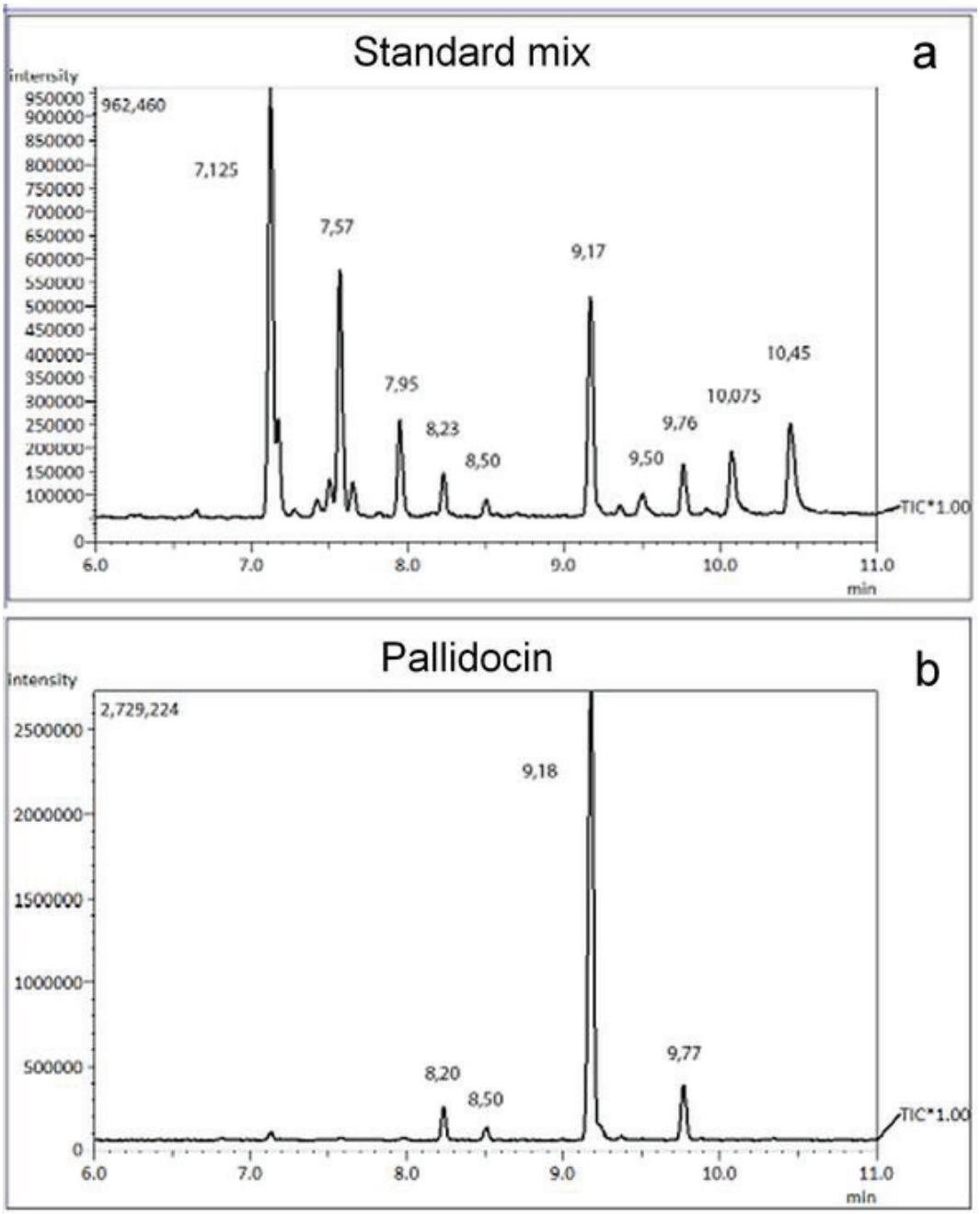


Supplementary Figure 1. Pallidocin analysis by LC-ESI-MS. (**a**) Native pallidocin analysis, (**b**) reduced with TCEP pallidocin analysis.

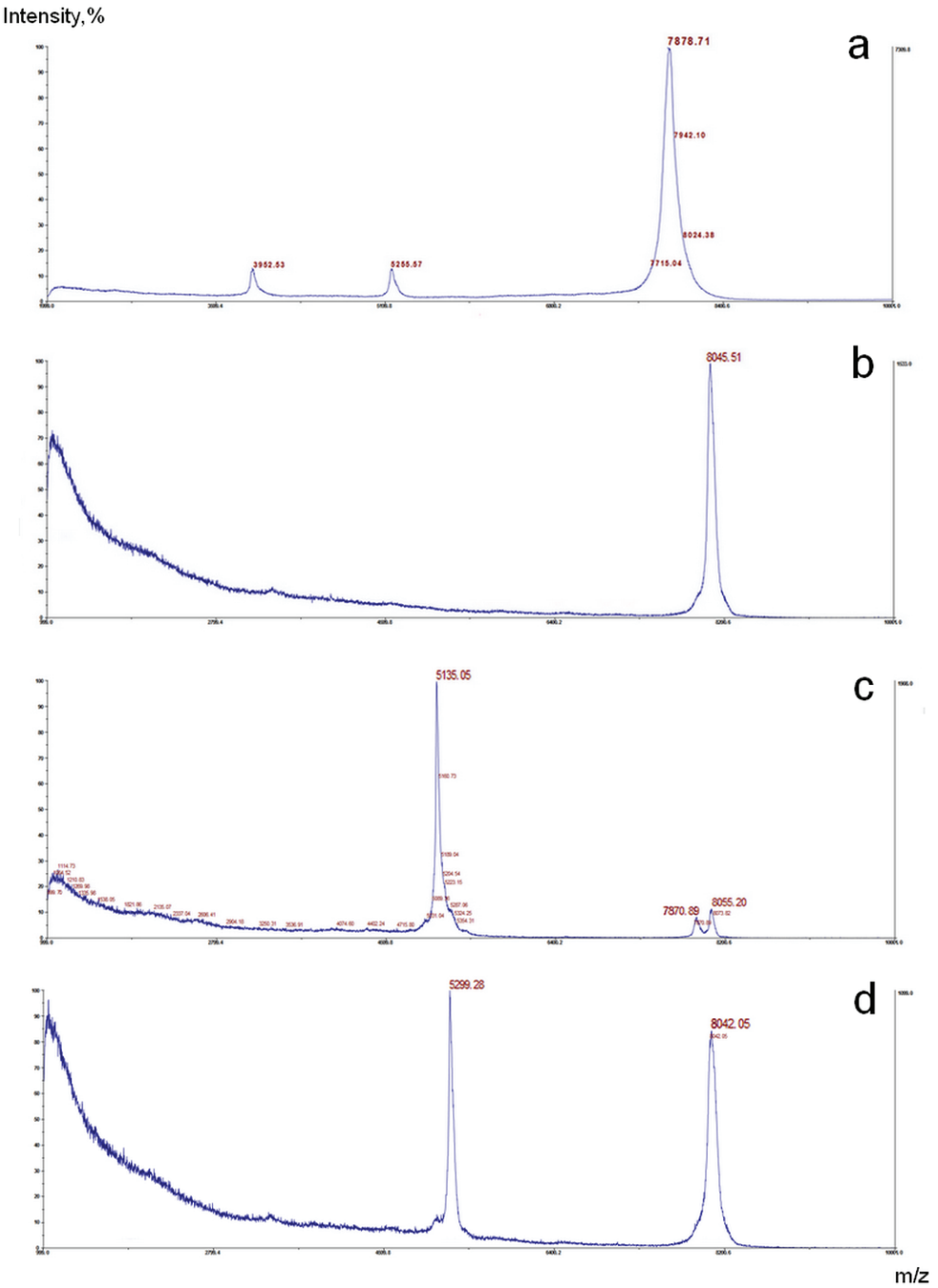


Main sequence ions b-162	Main sequence ions b	Position	Sequence	Position	Main sequence ions y	Main sequence ions y-162
		1	I	11		
49,09	211,14	2	P	10	1075,44	913,39
106,12	268,17	3	G	9	978,39	816,34
205,18	367,23	4	V	8	921,37	759,32
262,21	424,26	5	G	7	822,30	660,25
409,27	571,32	6	F	6	765,28	603,22
466,30	628,35	7	G	5	618,21	456,16
731,36	893401	8	C	4	561,19	399,14
788,38	950,43	9	G	3	296,12	
845,40	1007,45	10	G	2	239,10	
		11	Y	1	182,08	

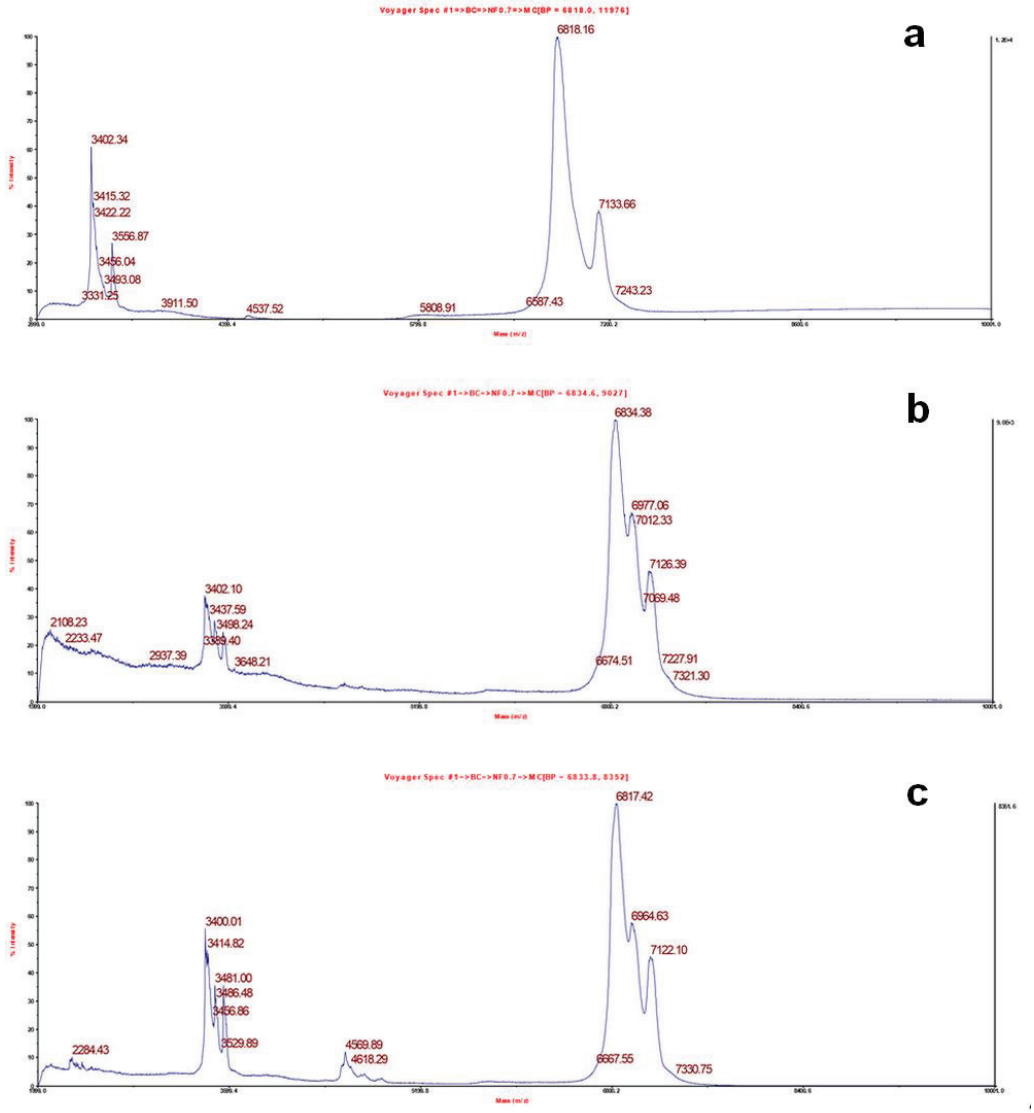
Supplementary Figure 2. ESI-Q-MSMS analysis of fragment IPGVGFGCGGY after pallidocin digestion with chymotrypsin. Theoretical fragmentation patterns of IPGVGFGCGGY are represented in the table. Gray cells indicates observed masses of the fragmented IPGVGFGCGGY sequence. Columns b-162 and y-162 represents ions of fragmented IPGVGFGCGGY sequence without modification. Columns b and y represents ions of fragmented IPGVGFGCGGY sequence with modification.



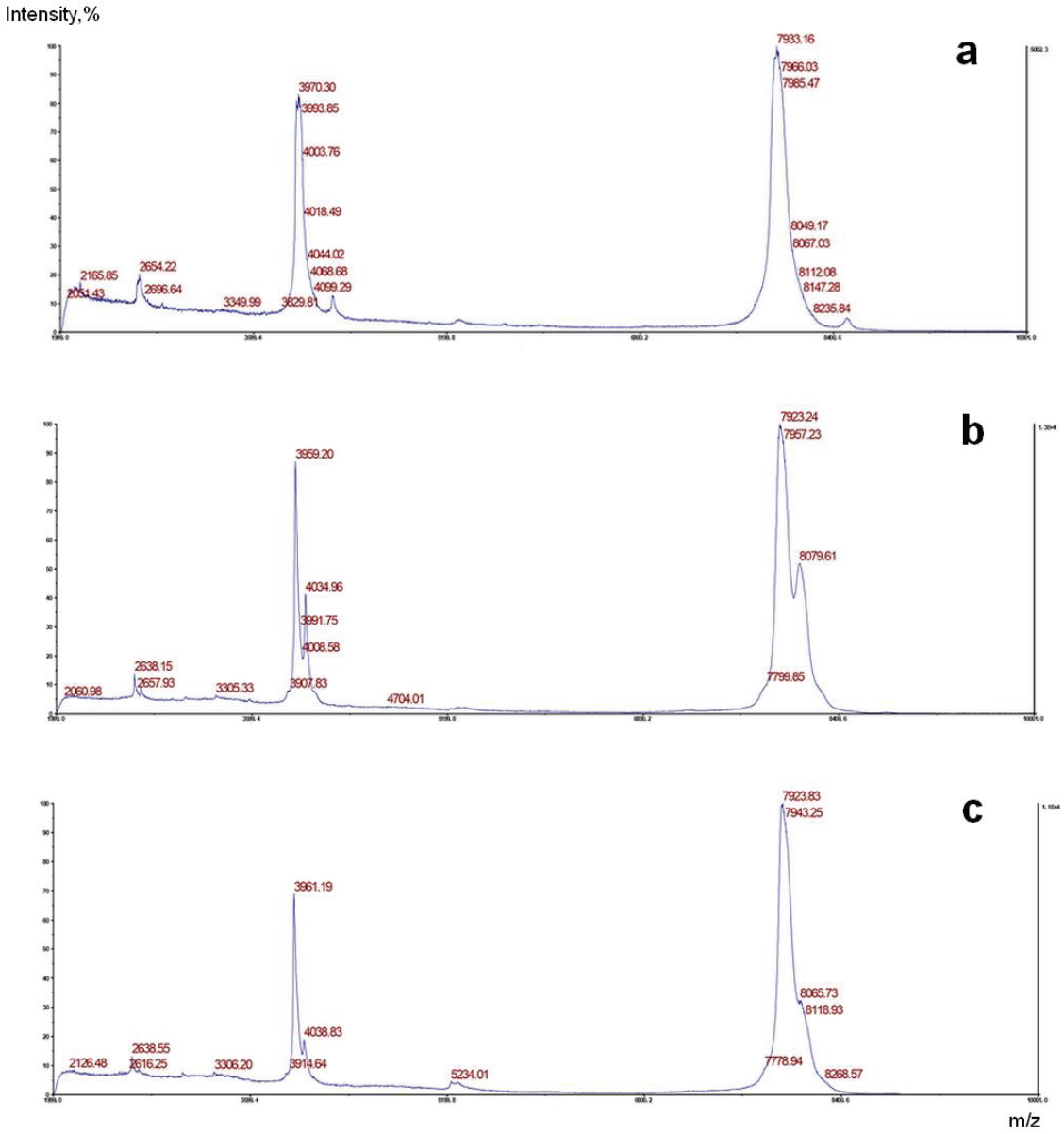
Supplementary Figure 3. GC-MS analysis chromatogram. Pallidocin analysis represented in part **b**. Standard mixture of sugars analysis represented in part **a**, with retention time of mannose – 7.125 min, α and β galactose – 7.57 min and 7.95 min, α and β glucose – 8.23 min and 8.5 min, mannitol (internal standard) – 9.17 min, mono-O-acetylated mannitol (internal standard) – 9.76 min, α and β N-acetylgalactosamine – 9.5 min and 10.075 min, N-acetylglucosamine – 10.45 min.



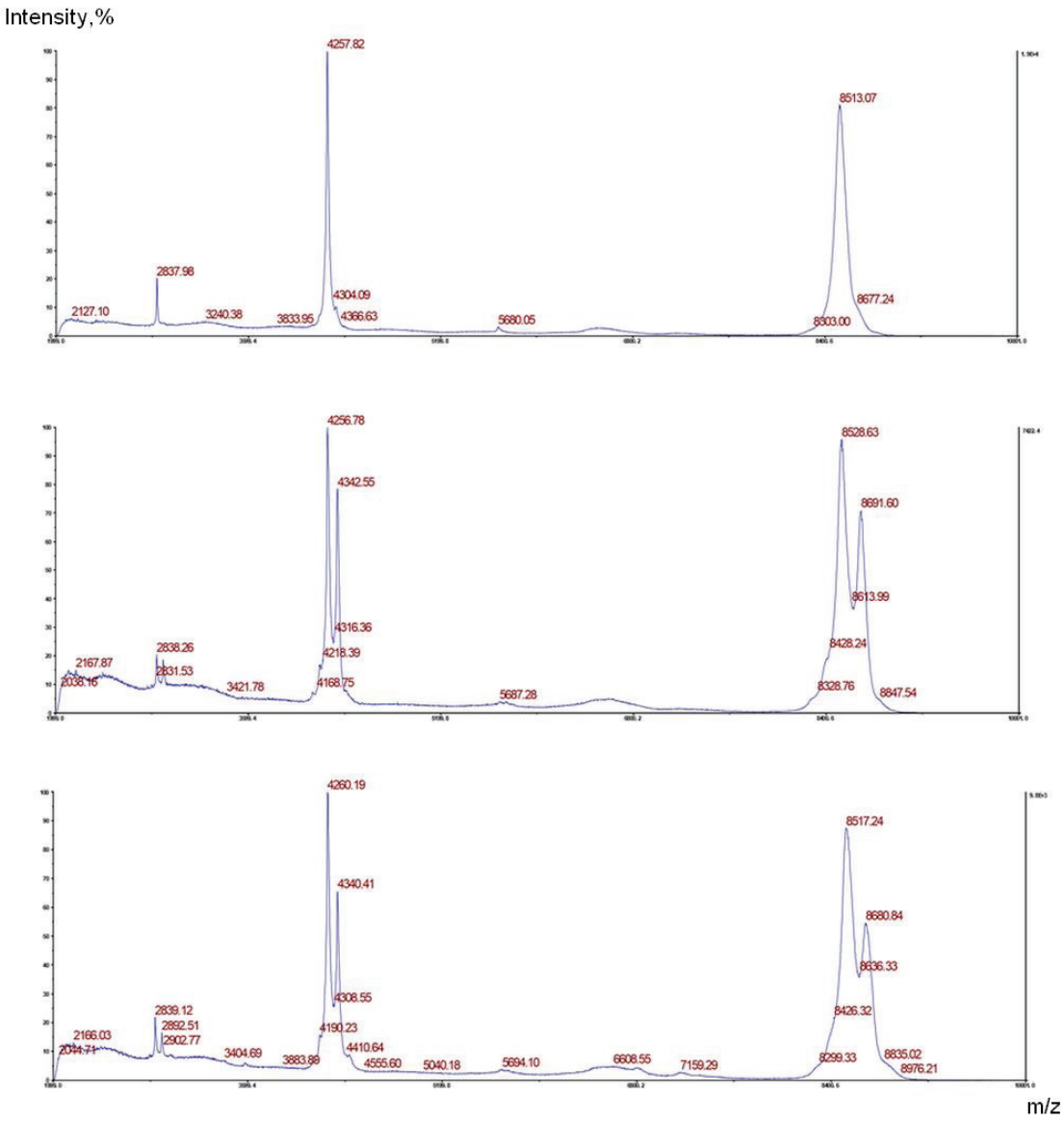
Supplementary Figure 4. MALDI-TOF MS analysis of purified pallidocin precursor peptide after *palA-his* co-expression with *palS, palT* or *palST*. Calculated pre-PalA-His (unmodified precursor with leader) average mass – 7877.93 Da, (**a**) observed mass after only *palA-his* expression – 7878.71 Da, which represents pre-PalA-His. Calculated pre-PalA-His-Glc (glycosylated precursor with leader) with 2 disulfide bonds average mass – 8035.93 Da, (**b**) observed mass after *palA-his* co-expression with *palS* – 8045.51 Da, which represents pre-PalA-His-Glc. Calculated PalA-His core peptide (unmodified peptide without leader) with 2 disulfide bonds average mass – 5133.77 Da, (**c**) observed mass after *palA-his* co-expression with *palT* – 5135.05 Da, which represents PalA-His core peptide. Calculated PalA-His-Glc core peptide (glycosylated peptide without leader) with 2 disulfide bonds average mass – 5295.77 Da, (**d**) observed mass after *palA-his* co-expression with *palST* – 5299.28 Da, which represents PalA-His-Glc core peptide (mature pallidocin).



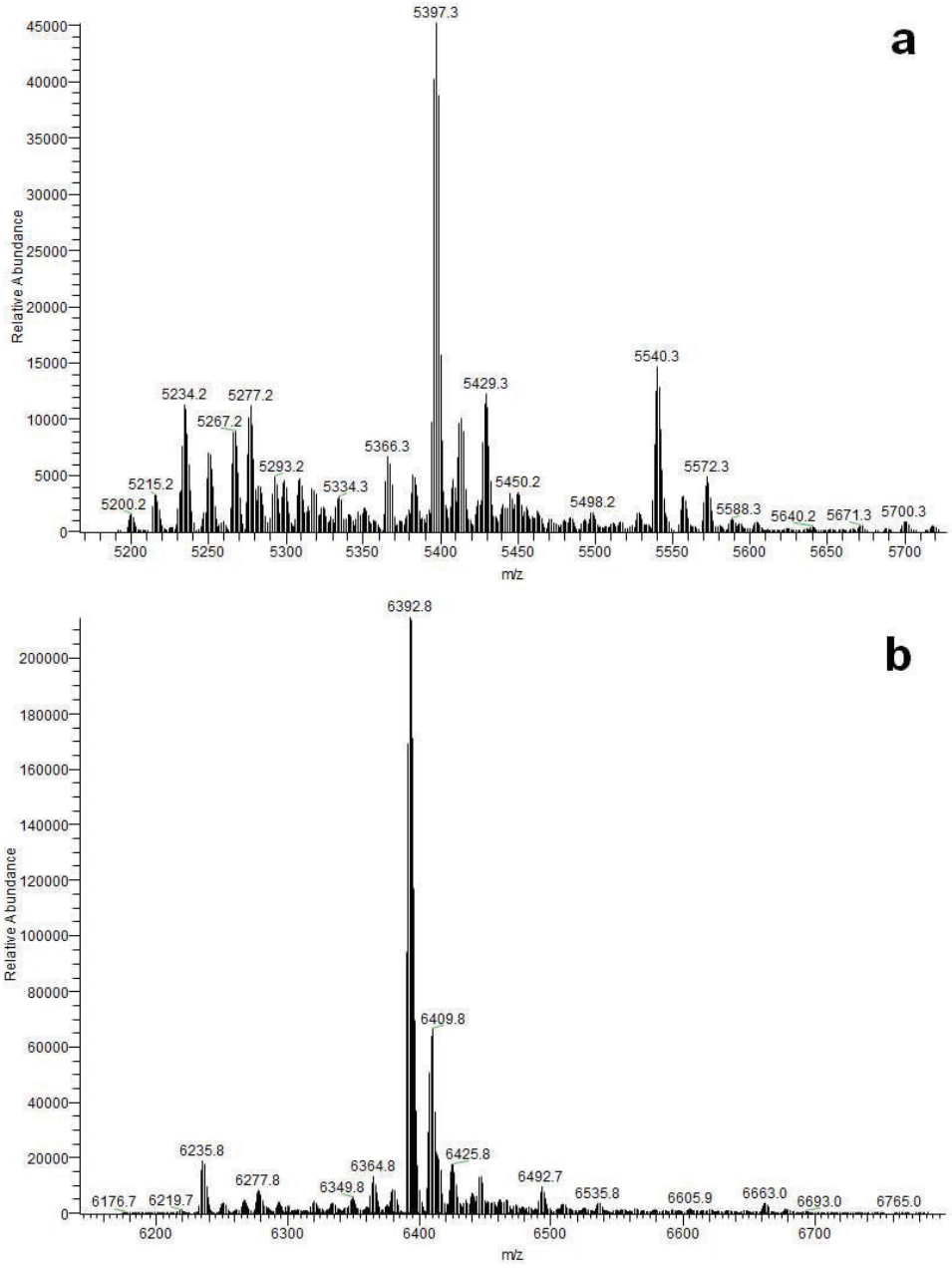
Supplementary Figure 5. MALDI-TOF MS analysis of purified sublancin precursor after *sunA-his* co-expression with *palS* or *palST*. Calculated pre-SunA-His (unmodified precursor with leader) with 2 disulfide bonds average mass is 6800.7 Da, (**a**) observed masses after only *sunA-his* expression – 6818.16 Da, which represents pre-SunA-His with methionine oxidation (+16 Da), and peak of 7133.66 Da – pre-SunA-His with possible unknown modifications. Calculated pre-SunA-His-Glc (glycosylated precursor with leader) with 2 disulfide bonds and glucose average mass is 6962.7 Da, (**b**) observed masses after *sunA-his* co-expression with *palS* – 6834.38 Da, which represents pre-SunA-His with two oxidations of methionines (+32 Da), and 6977.06 Da, which represents pre-SunA-His-Glc with possible methionine oxidation (+16 Da), peak with the mass of 7126.39 – pre-SunA-His with possible unknown modifications. Calculated SunA-His-Glc core peptide (glycosylated peptide without leader) with 2 disulfide bonds and glucose average mass is 4701.1 Da, (**c**) observed masses after *sunA-his* co-expression with *palST* – 6817.42 Da, which represents pre-SunA-His, and 6964.63 Da, which represents pre-SunA-His-Glc, peak with the mass of 7122.10 Da – pre-SunA-His with possible unknown modifications. Other peaks of smaller masses are possible degraded pre-SunA-His or pre-SunA-His-Glc peptides.



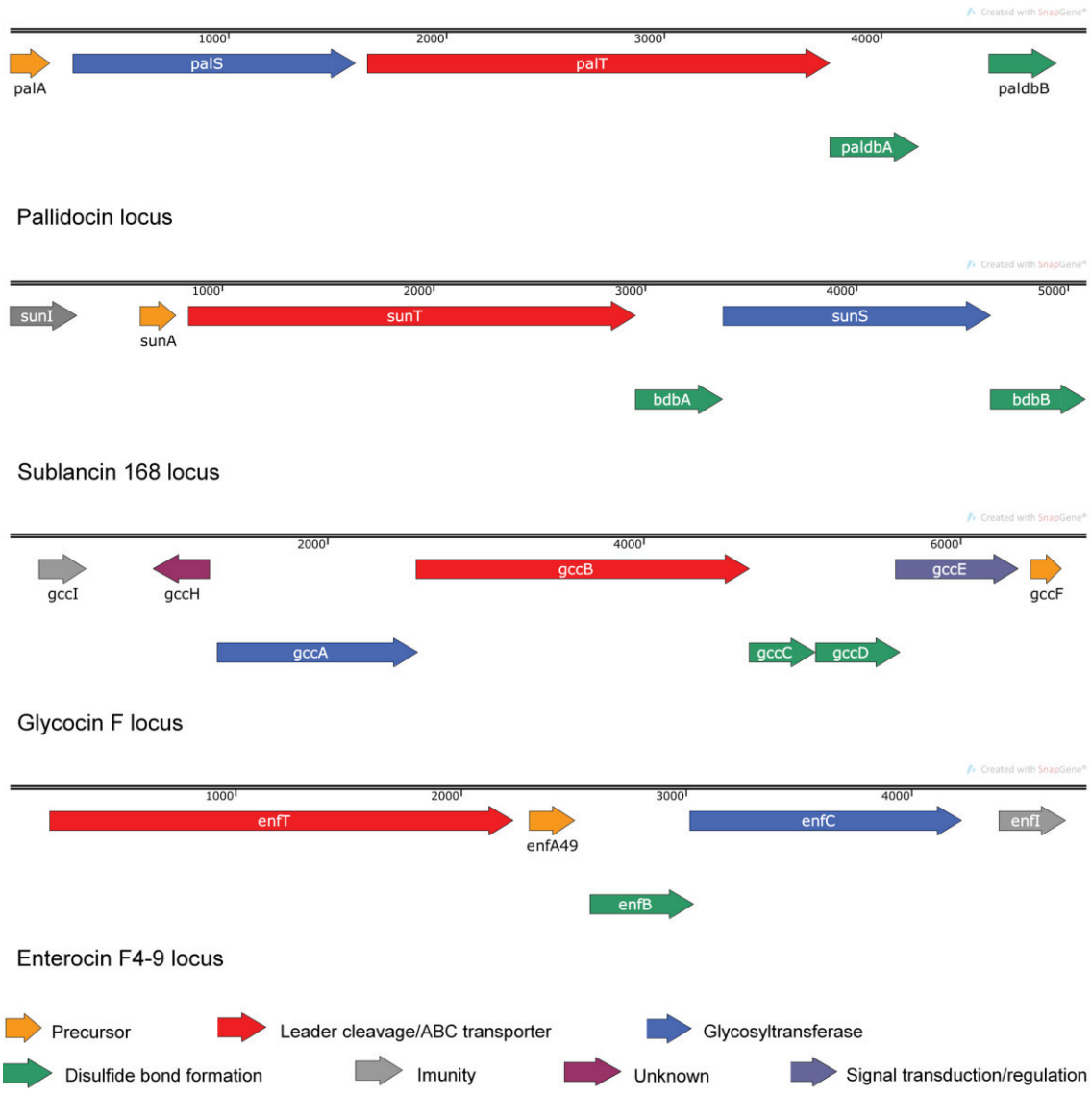
Supplementary Figure 6. MALDI-TOF MS analysis of purified Hyp1 precursor after *hyp1-his* co-expression with *palS* or *palST*. Calculated pre-Hyp1-His (unmodified precursor with leader) with 2 disulfide bonds average mass is 7912.0 Da, (**a**) observed mass after only *hyp1-his* expression – 7933.16, which represents pre-Hyp1-His with methionine oxidation (+16 Da). Calculated pre-Hyp1-His-Glc peptide (glycosylated precursor with leader) with 2 disulfide bonds and glucose average mass is 8074.0 Da, (**b**) observed masses after *hyp1-his* co-expression with *palS* – 7923.24 Da, which represents pre-Hyp1-His, and 8079.61 Da, which represents pre-Hyp1-His-Glc. Calculated Hyp1-His-Glc core peptide (glycosylated peptide without leader) with 2 disulfide bonds and glucose average mass is 5397.0 Da, (**c**) observed masses after *hyp1-his* co-expression with *palST* – 7923.83 Da, which represents pre-Hyp1-His, and 8065.73 Da, which represents pre-Hyp1-His-Glc. Other peaks of smaller masses are possible degraded pre-Hyp1-His or pre-Hyp1-His-Glc peptides.



Supplementary Figure 7. MALDI-TOF MS analysis of purified Hyp2 precursor peptide after *hyp2-his* co-expression with *palS* or *palST*. Calculated pre-Hyp2-His (unmodified precursor with leader) with 2 disulfide bonds average mass is 8514.2 Da, (**a**) observed mass after only *hyp2-his* expression – 8513.07 Da, which represents pre-Hyp2-His. Calculated Hyp2-His-Glc (glycosylated precursor with leader) with 2 disulfide bonds and glucose average mass is 8676.2 Da, (**b**) observed masses after *hyp2-his* co-expression with *palS* – 8528.63 Da, which represents pre-Hyp2-His with methionine oxidation (+16 Da), and 8691.6 Da, which represents pre-Hyp2-His-Glc with methionine oxidation (+16 Da). Calculated Hyp2-His-Glc core peptide (glycosylated peptide without leader) with 2 disulfide bonds and glucose average mass is 6396.0 Da, (**c**) observed masses after *hyp2-his* co-expression with *palST* – 8517.24 Da, which represents pre-Hyp2-His, and 8680.74 Da, which represents pre-Hyp2-His-Glc. Other peaks of smaller masses are possible degraded pre-Hyp2-His or pre-Hyp2-His-Glc peptides.



Supplementary Figure 8. LC-ESI-MS analysis of purified leaderless Hyp1 and Hyp2 peptides after *core_hyp1-his/core_hyp2-his* co-expression with *palS*. Calculated average mass of Hyp1-His-Glc core peptide (glycosylated peptide without leader) with 2 disulfide bonds – 5529.23 Da, Hyp1-His core peptide (unmodified peptide without leader) with 2 disulfide bonds – 5367.17 Da, (**a**) observed mass – 5397.3 Da (possible oxidation of two methionines +32 Da), which represents Hyp1-His core peptide. Calculated average mass of Hyp2-His-Glc core peptide (glycosylated peptide without leader) with 2 disulfide bonds – 6528.19 Da, Hyp2-His core peptide (unmodified peptide without leader) with 2 disulfide bonds – 6366.13 Da, (**b**) observed mass – 6392.8 Da (possible extra formyl or dimethyl +28 Da), which represents Hyp2-His.



Supplemenatry Figure 9. Gene clusters coding for glycocins.

12 Acknowledgements

I am very grateful to my PhD supervisors Prof. Donaldas J. Čitavičius and Prof. Lilija Kalėdienė for advises and possibility to do my PhD. Also I am sincerely grateful to my supervisor at the University of Groningen – Prof. Oscar P. Kuipers for the possibility to do my Erasmus+ traineeship, and consultant Dr. Andrius Buivydas for advises in the lab.

I would like to express my gratitude to Dr. Danutė Labeikytė from the department of Biochemistry and Molecular biology for assistance and kindness help in electrophoretic analysis experiments. I am also grateful to Dr. Marija Ger and Dr. Mindaugas Valius from the Proteomics Center at Vilnius University for their assistance with mass spectrometry analysis.

Special thanks for exciting time and friendly environment to all my lab collages in the department of Molecular Genetics at University of Groningen and the department of Microbiology and Biotechnology at Vilnius University.

13 References

1 Abdel-Mohsein HS, 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. Animal Science Journal 82(2):340-51. doi: 10.1111/j.1740-0929.2010.00835.x.

2 Acedo JZ, van Belkum MJ, Lohans CT, McKay RT, Miskolzie M, Vederas JC. 2015. Solution structure of acidocin B, a circular bacteriocin produced by *Lactobacillus acidophilus* M46. Applied Environmental Microbiology 81: 2910–2918. doi: 10.1128/AEM.04265-14.

3 Ahmad A, Hartman HB, Krishnakumar S, Fell DA, Poolman MG, Srivastava S. 2017. A genome scale model of *Geobacillus thermoglucosidasius* (C56-YS93) reveals its biotechnological potential on rice straw hydrolysate. Journal of Biotechnology 251:30-37. doi: 10.1016/j.jbiotec.2017.03.031.

4 Al Khatib Z, Lagedroste M, Fey I, Kleinschrodt D, Abts A, Smits SH. 2014. Lantibiotic immunity: inhibition of nisin mediated pore formation by NisI. PLoS One 9(7):e102246. doi: 10.1371/journal.pone.0102246.

5 Aliyu H, Lebre P, Blom J, Cowan D, De Maayer P. 2016. Phylogenomic re-assessment of the thermophilic genus *Geobacillus*. Systematic and Applied Microbiology 39(8):527-533. doi: 10.1016/j.syapm.2016.09.004.

6 Alkhalili RN, 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. International Journal of Molecular Sciences 17(8):pii: E1363. doi: 10.3390/ijms17081363.

7 Alvarez-Sieiro P, Montalbán-López M, Mu D, Kuipers OP. Bacteriocins of lactic acid bacteria: extending the family. Applied Microbiology and Biotechnology 100(7):2939-51. doi: 10.1007/s00253-016-7343-9.

8 Arnison PG, Bibb MJ, Bierbaum G, Bowers AA, Bugni TS, Bulaj G, Camarero JA, Campopiano DJ, Challis GL, Clardy J, Cotter PD, Craik DJ, Dawson M, Dittmann E, Donadio S, Dorrestein PC, Entian KD, Fischbach MA, Garavelli JS, Göransson U, Gruber CW, Haft DH, Hemscheidt TK, Hertweck C, Hill C,

Horswill AR, Jaspars M, Kelly WL, Klinman JP, Kuipers OP, Link AJ, Liu W, Marahiel MA, Mitchell DA, Moll GN, Moore BS, Müller R, Nair SK, Nes IF, Norris GE, Olivera BM, Onaka H, Patchett ML, Piel J, Reaney MJ, Rebuffat S, Ross RP, Sahl HG, Schmidt EW, Selsted ME, Severinov K, Shen B, Sivonen K, Smith L, Stein T, Süssmuth RD, Tagg JR, Tang GL, Truman AW, Vederas JC, Walsh CT, Walton JD, Wenzel SC, Willey JM, van der Donk WA. 2013. Ribosomally synthesized and post-translationally modified peptide natural products: overview and recommendations for a universal nomenclature. Natural Product Reports 30(1):108-60. doi: 10.1039/c2np20085f.

9 Asensio C, Perez-Diaz JC, Martinez MC and Baquero F. 1976. A new family of low molecular weight antibiotics from E*nterobacteria*. Biochemical and Biophysical Research Communications 69(1):7-14. doi: 10.1016/s0006-291x(76)80264-1.

10 Barbour A, Tagg J, Abou-Zied OK, Philip K. 2016. New insights into the mode of action of the lantibiotic salivaricin B. Scientific Reports 6:31749. doi: 10.1038/srep31749.

11 Başbülbül Özdemir G and Biyik HH. 2012. Isolation and characterization of a bacteriocin-like substance produced by *Geobacillus toebii* strain HBB-247. Indian Journal of Microbiology 52(1):104-8. doi: 10.1007/s12088-011-0227-x.

12 Başbülbül Özdemir G and Biyik HH. 2012. Isolation and characterization of toebicin 218, a bacteriocin, produced by *Geobacillus toebii* HBB-218. African Journal of Biotechnology 11(30):7711-7719. doi: 10.5897/AJB11.1331.

13 Bayro MJ, Mukhopadhyay J, Swapna GV, Huang JY, Ma LC, Sineva E, Dawson PE, Montelione GT, Ebright RH. 2003. Structure of antibacterial peptide microcin J25: a 21-residue lariat protoknot. Journal of American Chemical Sociaty 125:12382–12383. doi: 10.1021/ja036677e.

14 Beatson SA, Sloan GL, Simmonds RS. 1998. Zoocin A immunity factor: a *femA*-like gene found in a group C streptococcus. FEMS Microbiology Letters 163(1):73-7.

15 Buchan DWA, Minneci F, Nugent TCO, Bryson K, Jones DT. 2013. Scalable web services for the PSIPRED Protein Analysis Workbench. Nucleic Acids Research 41(W1): W340-W348. doi: 10.1093/nar/gkt381.

16 Burgess SA, Flint SH, Lindsay D, Cox MP, Biggs PJ. 2017. Insights into the *Geobacillus stearothermophilus* species based on phylogenomic principles. BMC Microbiology 17(1):140. doi: 10.1186/s12866-017-1047-x.

17 Chandra TJ, Mani PS. 2011. A study of 2 rapid tests to differentiate Gram positive and Gram negative aerobic bacteria. J Med Allied Sci. 1(2):84-85.

18 Chen Y, Simmonds RS, Young JK, Timkovich R. 2013. Solution structure of the recombinant target recognition domain of zoocin A. Proteins 81(4):722-7. doi: 10.1002/prot.24224.

19 Chevallet M, Luche S, Rabilloud T. 2006. Silver staining of proteins in polyacrylamide gels. Nature Protocols 1(4):1852-8. doi: 10.1038/nprot.2006.288.

20 Chikindas ML, Weeks R, Drider D, Chistyakov VA, Dicks LM. 2017. Functions and emerging applications of bacteriocins. Current Opinion in Biotechnology 49:23-28. doi: 10.1016/j.copbio.2017.07.011.

21 Cladera-Olivera F, Caron GR, Brandelli A. 2004. Bacteriocin-like substance production by *Bacillus licheniformis* strain P40. Letters in Applied Microbiology 38(4):251-6. doi: 10.1111/j.1472-765X.2004.01478.x.

22 Cotter PD, Hill C and Ross RP. 2005. Bacteriocins: developing innate immunity for food. Nature Reviews Microbiology 3:777-788. doi: 10.1038/nrmicro1273.

23 Cotter PD, Ross RP and Hill C. 2013. Bacteriocins — a viable alternative to antibiotics? Nature Reviews Microbiology 11(2):95-105. doi: 10.1038/nrmicro2937.

24 Cui Y, Zhang C, Wang Y, Shi J, Zhang L, Ding Z, Qu X, Cui H. 2012. Class IIa bacteriocins: diversity and new developments. International Journal of Molecular Sciences 13(12):16668-707. doi: 10.3390/ijms131216668.

25 Daba H, Pandian S, Gosselin JF, Simard RE, Huang J, Lacroix C. 1991. Detection and activity of a bacteriocin produced by *Leuconostoc mesenteroides*. Applied Environmental Microbiology 57(12):3450-5.

26 Dinsdale AE, Halket G, Coorevits A, Van Landschoot A, Busse HJ, De Vos P, Logan NA. 2011. Emended descriptions of *Geobacillus thermoleovorans* and *Geobacillus thermocatenulatus*. International Journal of Systematic and Evolutionary Microbiology 61:1802-10. doi: 10.1099/ijs.0.025445-0.

27 Dorenbos R, Stein T, Kabel J, Bruand C, Bolhuis A, Bron S, Quax WJ and van Dijl JM. 2002. Thiol-Disulfide Oxidoreductases Are Essential for the Production of the Lantibiotic Sublancin 168. The Journal of Biological Chemistry 277(19):16682-16688. doi: 10.1074/jbc.M201158200.

28 Dubois JYF, Kouwen TRHM, Schurich AKC, Reis CR, Ensing HT, Trip EN, Zweers JC and van Dijl JM. 2009. Immunity to the bacteriocin sublancin 168 is determined by the SunI (YolF) protein of *Bacillus subtilis*. Antimicrobial Agents and Chemotherapy 53(2):651-661. doi: 10.1128/AAC.01189-08.

29 Favret ME, Yousten AA. 1989. Thuricin: the bacteriocin produced by *Bacillus thuringiensis*. Journal of Invertebrate Pathology 53(2):206-16.

30 Fikes JD, Crabtree BL, and Barridge BD. 1983. Studies on the mode of action of a bacteriocin produced by *Bacillus stearothermophilus*. Canadian Journal of Microbiology 29(11):1576-1582.

31 Flühe L, Burghaus O, Wieckowski BM, Giessen TW, Linne U, Marahiel MA. 2013. Two [4Fe-4S] clusters containing radical SAM enzyme SkfB catalyze thioether bond formation during the maturation of the sporulation killing factor. Journal of American Chemical Society 135(3):959-62. doi: 10.1021/ja310542g.

32 Flühe L, Knappe TA, Gattner MJ, Schäfer A, Burghaus O, Linne U, Marahiel MA. 2012. The radical SAM enzyme AlbA catalyzes thioether bond formation in subtilosin A. Nature Chemical Biology 8(4):350-7. doi: 10.1038/nchembio.798.

33 Gabrielsen C, Brede DA, Hernández PE, Nes IF, Diep DB. 2012. Genome sequence of the bacteriocin-producing strain *Lactococcus garvieae* DCC43. Journal of Bacteriology 194(24):6976-7. doi: 10.1128/JB.01864-12.

34 Gabrielsen C, Brede DA, Nes IF, Diep DB. 2014. Circular bacteriocins: biosynthesis and mode of action. Applied and Environmental Microbiology. 80(22):6854-62. doi: 10.1128/AEM.02284-14.

35 Garcia De Gonzalo CV, Denham EL, Mars RAT, Stülke J, van der Donk WA, van Dijl JM. 2015. The phosphoenolpyruvate:sugar phosphotransferase system is involved in sensitivity to the glucosylated bacteriocin sublancin. Antimicrobial Agents and Chemotherapy 59:6844–6854. doi: 10.1128/AAC.01519-15.

36 Garcia De Gonzalo CV, Zhu L, Oman TJ, and van der Donk WA. 2014. NMR structure of the S-linked glycopeptide sublancin 168. ACS Chemical Biology 9(3):796–801. doi: 10.1021/cb4008106.

37 Garg N, Goto Y, Chen T, van der Donk WA. 2016. Characterization of the stereochemical configuration of lanthionines formed by the lanthipeptide synthetase GeoM. Biopolymers 106(6):834-842. doi: 10.1002/bip.22876.

38 Garg N, Oman TJ, Andrew Wang TS, De Gonzalo CV, Walker S, van der Donk WA. 2014. Mode of action and structure-activity relationship studies of geobacillin I. Journal of Antibiotics 67(1):133-6. doi: 10.1038/ja.2013.112.

39 Garg N, Tang W, Goto Y, Nair SK, van der Donk WA. 2012. Lantibiotics from *Geobacillus thermodenitrificans*. PNAS 109(14):5241-6. doi: 10.1073/pnas.1116815109.

40 Gillor O, Kirkup BC, Riley MA. 2004. Colicins and microcins: the next generation antimicrobials. Advances in applied microbiology 54:129-46. doi: 10.1016/S0065-2164(04)54005-4.

41 González-Pastor JE, Hobbs EC, Losick R. 2003. Cannibalism by sporulating bacteria. Science 301(5632):510-3. doi: 10.1126/science.1086462.

42 Gopal N, Hill C, Ross PR, Beresford TP, Fenelon MA, and Cotter PD. 2015. The prevalence and control of *Bacillus* and related spore-forming bacteria in the dairy industry. Frontiers in Microbiology 6:1418. doi: 10.3389/fmicb.2015.01418.

43 Grande-Burgos MJ, Pulido RP, Del Carmen López Aguayo M, Gálvez A, Lucas R. 2014. The cyclic antibacterial peptide enterocin AS-48: isolation, mode of action, and possible food applications. International Journal of Molecular Sciences 15(12):22706-22727. doi: 10.3390/ijms151222706.

44 Gratia A. 1925. Sur un remarquable exemple d'antagonisme entre deux souches de colibacille. Compt. Rend. Soc. Biol. 93:1040–1042.

45 Hammami I, Triki MA, Rebai A. 2011. Purification and characterization of the novel bacteriocin IH7 with antifungal and antibacterial properties. Journal of Plant Pathology 93 443–454. doi: 10.4454/jpp.v93i2.1200.

46 Hammami R, Zouhir A, Le Lay C, Hamida JB, Fliss I. 2010. BACTIBASE second release: a database and tool platform for bacteriocin characterization. BMC Microbiology 10:22. doi: 10.1186/1471-2180-10-22.

47 Hasper HE, Kramer NE, Smith JL, Hillman JD, Zachariah C, Kuipers OP, de Kruijff B, Breukink E. 2006. An alternative bactericidal mechanism of action for lantibiotic peptides that target lipid II. Science 313(5793):1636-7. doi: 10.1126/science.1131167.

48 Hata T, Tanaka R, Ohmomo S. 2009. Isolation and characterization of plantaricin ASM1: A new bacteriocin produced by *Lactobacillus plantarum* A-1. International Journal of Food Microbiology 137(1):94–99. doi: 10.1016/j.ijfoodmicro.2009.10.021.

49 Heddle JG, Blance SJ, Zamble DB, Hollfelder F, Miller DA, Wentzell LM, Walsh CT, Maxwell A. 2001. The antibiotic microcin B17 is a DNA gyrase poison: characterisation of the mode of inhibition. Journal of Molecular Biology 307(5):1223-34. doi: 10.1006/jmbi.2001.4562.

50 Hegarty JW, Guinane CM, Ross RP, Hill C, Cotter PD. 2016. Bacteriocin production: a relatively unharnessed probiotic trait? F1000Research 5:2587. doi: 10.12688/f1000research.9615.1.

51 Hegemann JD, Zimmermann M, Xie X, Marahiel MA. 2015. Lasso peptides: an intriguing class of bacterial natural products. Accountsn of Chemical Research 48(7):1909-19. doi: 10.1021/acs.accounts.5b00156.

52 Heng NC, Ragland NL, Swe PM, Baird HJ, Inglis MA, Tagg JR, Jack RW. 2006. Dysgalacticin: a novel, plasmid-encoded antimicrobial protein (bacteriocin) produced by *Streptococcus dysgalactiae* subsp. *equisimilis*. Microbiology 152(Pt 7):1991-2001. doi: 10.1099/mic.0.28823-0.

53 Himeno K, Rosengren KJ, Inoue T, Perez RH, Colgrave ML, Lee HS, Chan LY, Henriques ST, Fujita K, Ishibashi N, Zendo T, Wilaipun P, Nakayama J, Leelawatcharamas V, Jikuya H, Craik DJ, Sonomoto K. 2015. Identification, characterization, and three-dimensional structure of the novel circular

bacteriocin, enterocin NKR-5-3B, from *Enterococcus faecium*. Biochemistry 54(31):4863-76. doi: 10.1021/acs.biochem.5b00196.

54 Hsu ST, Breukink E, Tischenko E, Lutters MA, de Kruijff B, Kaptein R, Bonvin AM, van Nuland NA. 2004. The nisin-lipid II complex reveals a pyrophosphate cage that provides a blueprint for novel antibiotics. Nature Structural and Molecular Biology 11(10):963-7. doi: 10.1038/nsmb830.

55 Ihnken LAF, Chatterjee C, and van der Donk WA. 2008. *In vitro* reconstitution and substrate specificity of a lantibiotic protease. Biochemistry 47(28):7352-7363. doi: 10.1021/bi800278n.

56 Ito Y, Kawai Y, Arakawa K, Honme Y, Sasaki T, Saito T. 2009. Conjugative plasmid from *Lactobacillus gasseri* LA39 that carries genes for production of and immunity to the circular bacteriocin gassericin A. Applied and Environmental Microbiology 75(19):6340-51. doi: 10.1128/AEM.00195-09.

57 Iwatani S, Horikiri Y, Zendo T, Nakayama J, Sonomoto K. 2013. Bifunctional gene cluster *lnqBCDEF* mediates bacteriocin production and immunity with differential genetic requirements. Applied and Environmental Microbiology 79(7):2446-9. doi: 10.1128/AEM.03783-12.

58 Jack RW, Tagg JR and Ray B. 1995. Bacteriocins of Gram-positive bacteria. Microbiological Reviews 59(2):171–200.

59 Ji S, Li W, Baloch AR, Wang M, Cao B. 2015. Improved production of sublancin via introduction of three characteristic promoters into operon clusters responsible for this novel distinct glycopeptide biosynthesis. Microbial Cell Factories 14:17. doi: 10.1186/s12934-015-0201-0.

60 Jiang Y, Xin F, Lu J, Dong W, Zhang W, Zhang M, Wu H, Ma J, Jiang M. 2017. State of the art review of biofuels production from lignocellulose by thermophilic bacteria. Bioresource Technology 245:1498-1506. doi: 10.1016/j.biortech.2017.05.142.

61 Joerger MC, Klaenhammer TR. 1986. Characterization and purification of helveticin J and evidence for a chromosomally determined bacteriocin produced by *Lactobacillus helveticus* 481. Journal of Bacteriology 167(2):439-46.

62 Joerger MC, Klaenhammer TR. 1990. Cloning, expression, and nucleotide sequence of the *Lactobacillus helveticus* 481 gene encoding the bacteriocin helveticin J. Journal of Bacteriology 172(11):6339-47.

63 Józefiak D, Kierończyk B, Juśkiewicz J, Zduńczyk Z, Rawski M, Długosz J, Sip A, Højberg O. 2013. Dietary nisin modulates the gastrointestinal microbial ecology and enhances growth performance of the broiler chickens. PLoS One 8(12):e85347. doi: 10.1371/journal.pone.0085347.

64 Kalmokoff ML, Cyr TD, Hefford MA, Whitford MF, Teather RM. 2003. Butyrivibriocin AR10, a new cyclic bacteriocin produced by the ruminal anaerobe *Butyrivibrio fibrisolvens* AR10: characterization of the gene and peptide. Canadian Journal of Microbiology 49(12):763-73. doi: 10.1139/w03-101.

65 Kaur S, Kaur S. 2015. Bacteriocins as potential anticancer agents. Frontiers in Pharmacology 6:272. doi: 10.3389/fphar.2015.00272.

66 Kemperman R, Jonker M, Nauta A, Kuipers OP, Kok J. 2013. Functional analysis of the gene cluster involved in production of the bacteriocin circularin A by *Clostridium beijerinckii* ATCC 25752.

67 Khan H, Flint SH, Yu PL. 2013. Determination of the mode of action of enterolysin A, produced by *Enterococcus faecalis* B9510. Journal of Applied Microbiology 115(2):484-94. doi: 10.1111/jam.12240.

68 Khusainov R, Moll GN. and Kuipers OP. 2013. Identification of distinct nisin leader peptide regions that determine interactions with the modification enzymes NisB and NisC. FEBS Open Bio 3(1):237-242. doi: 10.1016/j.fob.2013.05.001.

69 Kiss A, Balikó G, Csorba A, Chuluunbaatar T, Medzihradszky KF, Alföldi L. 2008. Cloning and characterization of the DNA region responsible for Megacin A-216 production in *Bacillus megaterium* 216. Journal of Bacteriology 190(19):6448-57. doi: 10.1128/JB.00557-08.

70 Klaenhammer TR. 1993. Genetics of bacteriocins produced by lactic acid bacteria. FEMS Microbiology Reviews 12(1-3):39-85. doi: 10.1016/0168-6445(93)90057-G.

71 Korenblum E, der Weid I, Santos AL, Rosado AS, Sebastián GV, Coutinho CM, Magalhães FC, Paiva MM, Seldin L. 2005. Production of antimicrobial substances by *Bacillus subtilis* LFE-1, B. *firmus* HO-1 and B. *licheniformis* T6-5 isolated from an oil reservoir in Brazil. Journal of Applied Microbiology 98(3):667-75. doi: 10.1111/j.1365-2672.2004.02518.x.

72 Kotake Y, Ishii S, Yano T, Katsuoka Y, and Substrate HH. 2008. Recognition mechanism of the peptidase domain of the quorum-sensing-signal−producing ABC transporter ComA from *Streptococcus.* Biochemistry 47(8): 2531-2538. doi: 10.1021/bi702253n.

73 Kouwen TR, Trip EN, Denham EL, Sibbald MJ, Dubois JY, van Dijl JM. 2009. The large mechanosensitive channel MscL determines bacterial susceptibility to the bacteriocin sublancin 168. Antimicrobial Agents and Chemotherapy 53(11):4702-11. doi: 10.1128/AAC.00439-09.

74 Kuisienė N, Raugalas J, Chitavichius D. 2004. *Geobacillus lituanicus* sp. nov. International Journal of Systematic and Evolutionary Microbiology 54:1991-5. doi: 10.1099/ijs.0.02976-0.

75 Lagedroste M, Smits SHJ, and Schmitt L. 2017. Substrate specificity of the secreted nisin leader peptidase NisP. Biochemistry 56(30):4005-4014. doi: 10.1021/acs.biochem.7b00524.

76 Lagha AB, Haas B, Gottschalk M, Grenier D. 2017. Antimicrobial potential of bacteriocins in poultry and swine production. Veterinary Research 48:22. https://doi.org/10.1186/s13567-017-0425-6.

77 Laviña M, Pugsley AP, Moreno F. 1986. Identification, mapping, cloning and characterization of a gene (sbmA) required for microcin B17 action on *Escherichia coli* K12. Journal of General Microbiology 132:1685–1693. doi: 10.1099/00221287-132-6-1685.

78 Levengood MR, Patton GC, van der Donk WA. 2007. The leader peptide is not required for posttranslational modification by lacticin 481 synthetase. Journal of American Chemical Society 129:10314–10315. doi: 10.1021/ja072967+.

79 Liu J, Wu Q, Li L, Xi X, Wu D, Zhou M, Chen T, Shaw C and Wang L. 2017. Discovery of phylloseptins that defense against Gram-positive bacteria and inhibit the proliferation of the non-small cell lung cancer cell line, from the skin

secretions of *Phyllomedusa* frogs. 22(9):1428. doi: 10.3390/molecules22091428.

80 Lubelski J, Khusainov R, Kuipers OP. 2009. Directionality and coordination of dehydration and ring formation during biosynthesis of the lantibiotic nisin. Journal of Biological Chemistry 284(38):25962-72. doi: 10.1074/jbc.M109.026690.

81 Lubelski J, Rink R, Khusainov R, Moll GN, Kuipers OP. 2008. Biosynthesis, immunity, regulation, mode of action and engineering of the model lantibiotic nisin. Cellular and Molecular Life Sciences 65(3):455-76. doi: 10.1007/s00018-007-7171-2.

82 Mahanta N, Hudson GA, Mitchell DA. 2017. Radical S-adenosylmethionine enzymes involved in RiPP biosynthesis. Biochemistry doi: 10.1021/acs.biochem.7b00771.

83 Maky MA, Ishibashi N, Zendo T, Perez RH, Doud JR, Karmi M, Sonomoto K. 2015. Enterocin F4-9, a novel O-linked glycosylated bacteriocin. Applied and Environmental Microbiology 81(14):4819-26. doi: 10.1128/AEM.00940-15.

84 Martínez-Bueno M, Valdivia E, Gálvez A, Coyette J, Maqueda M. 1998. Analysis of the gene cluster involved in production and immunity of the peptide antibiotic AS-48 in *Enterococcus faecalis*. Molecular Microbiology 27(2):347-58. doi: 10.1046/j.1365-2958.1998.00682.x.

85 Martirani L, Varcamonti M, Naclerio G, De Felice M. 2012. Purification and partial characterization of bacillocin 490, a novel bacteriocin produced by a thermophilic strain of *Bacillus licheniformis*. Microbial Cell Factories 1(1):1.

86 Masuda Y, Ono H, Kitagawa H, Ito H, Mu F, Sawa N, Zendo T, Sonomoto K. 2011. Identification and characterization of leucocyclicin Q, a novel cyclic bacteriocin produced by *Leuconostoc mesenteroides* TK41401. Applied and Environmental Microbiology 77(22):8164-70. doi: 10.1128/AEM.06348-11.

87 Masuda Y, Zendo T, Sonomoto K. 2012. New type non-lantibiotic bacteriocins: circular and leaderless bacteriocins. Beneficial Microbes 3(1):3-12. doi: 10.3920/BM2011.0047.

88 Mathur H, Fallico V, O'Connor PM, Rea MC, Cotter PD, Hill C, Ross RP. 2017. Insights into the mode of action of the sactibiotic thuricin CD. Frontiers Microbiology 8:696. doi: 10.3389/fmicb.2017.00696.

89 Mathur H, O'Connor PM, Cotter PD, Hill C, Ross RP. 2014. Heterologous expression of thuricin CD immunity genes in *Listeria monocytogenes*. Antimicrobials Agents and Chemotherapy (6):3421-8. doi: 10.1128/AAC.00090-14.

90 Mathur H, Rea MC, Cotter PD, Hill C, Ross RP. 2016. The efficacy of thuricin CD, tigecycline, vancomycin, teicoplanin, rifampicin and nitazoxanide, independently and in paired combinations against *Clostridium difficile* biofilms and planktonic cells. Gut Pathogens 8:20. doi: 10.1186/s13099-016-0102-8.

91 Medema MH, Blin K, Cimermancic P, de Jager V, Zakrzewski P, Fischbach MA, 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 Research 39:W339–W346. doi: 10.1093/nar/gkr466.

92 Meier-Kolthoff JP, Göker M, Klenk HP. 2014. Taxonomic use of DNA G+C content and DNA-DNA hybridization in the genomic age. International Journal of Systematic and Evolutionary Microbiology 64:352-356. doi: 10.1099/ijs.0.056994-0.

93 Molchanova N, Hansen PR and Franzy H. 2017. Advances in development of antimicrobial peptidomimetics as potential drugs. Molecules 22(9): 1430. doi: 10.3390/molecules22091430.

94 Mormak DA and Casida, LE. 1985. Study of *Bacillus subtilis* Endospores in soil by use of a modified endospore stain. Appl. Environ. Microbiol. 49(6): 1356–1360.

95 Müller WM, Ensle P, Krawczyk B, Süssmuth RD. 2011. Leader peptide-directed processing of labyrinthopeptin A2 precursor peptide by the modifying enzyme LabKC. Biochemistry 50(39):8362-73. doi: 10.1021/bi200526q.

96 Nguyen LT, Haney EF, Vogel HJ. 2011. The expanding scope of antimicrobial peptide structures and their modes of action. Trends in Biotechnology 29(9):464-472. doi: 10.1016/j.tibtech.2011.05.001.

97 Nigutová K, Serencová L, Piknová M, Javorský P, Pristas P. 2008. Heterologous expression of functionally active enterolysin A, class III bacteriocin from *Enterococcus faecalis*, in *Escherichia coli*. Protein Expression and Purification 60(1):20-4. doi: 10.1016/j.pep.2008.03.006.

98 Nilsen T, Nes IF, Holo H. 2003. Enterolysin A, a cell wall-degrading bacteriocin from *Enterococcus faecalis* LMG 2333. Applied and Environmental Microbiology 69(5):2975-84.

99 Nissen-Meyer J, Oppegård C, Rogne P, Haugen HS, Kristiansen PE. 2010. Structure and mode-of-action of the two-peptide (class-IIb) bacteriocins. Probiotics and Antimicrobial Proteins 2(1):52-60. doi: 10.1007/s12602-009-9021-z.

100 Norris GE and Patchett ML. 2016. The glycocins: in a class of their own. Current Opinion in Structural Biology 40:112-119. doi: 10.1016/j.sbi.2016.09.003.

101 Norris GE., Patchett ML. 2014. Glycosylated ribosomally synthesized peptide toxins. In book: Natural Products Analysis: Instrumentation, Methods, and Applications. Wiley. Chapter 14:507-543. doi: 10.1002/9781118876015.ch14.

102 Novotny JF Jr, Perry JJ. 1992. Characterization of bacteriocins from two strains of *Bacillus thermoleovorans*, a thermophilic hydrocarbon-utilizing species. Applied and Environmental Microbiology 58(8):2393-6.

103 Oman TJ, Boettcher JM, Wang H, Okalibe XN, & van der Donk WA. 2011. Sublancin is not a lantibiotic but an S-linked glycopeptide. Nature Chemical Biology 7(2):78–80. doi: 10.1038/nchembio.509.

104 Ovchinnikov KV, Chi H, Mehmeti I, Holo H, Nes IF, Diep DB. 2016. Novel group of leaderless multipeptide bacteriocins from Gram-positive bacteria. Applied and Environmental Microbiology 82(17):5216-24. doi: 10.1128/AEM.01094-16.

105 Paik SH, Chakicherla A and Hansen JN. 1998. Identification and characterization of the structural and transporter genes for, and the chemical and biological properties of, sublancin 168, a novel lantibiotic produced by *Bacillus subtilis* 168. Journal of Biological Chemistry 273(36):23134-23142. doi: 10.1074/jbc.273.36.23134.

106 Pan SJ, Rajniak J, Maksimov MO, Link AJ. 2012. The role of a conserved threonine residue in the leader peptide of lasso peptide precursors. Chemical Communications 48(13):1880-2. doi: 10.1039/c2cc17211a.

107 Patton GC, Paul M, Cooper LE, Chatterjee C, van der Donk WA. 2008. The omportance of the leader sequence for directing lanthionine formation in lacticin 481. Biochemistry 47(28):7342-7351. doi: 10.1021/bi800277d.

108 Perez RH, Zendo T and Sonomoto K. 2014. Novel bacteriocins from lactic acid bacteria (LAB): various structures and applications. Microbial Cell Factories 13(Suppl 1):S3. doi: 10.1186/1475-2859-13-S1-S3.

109 Petersen TN, Brunak S, von Heijne G, Nielsen H. 2011. SignalP 4.0: discriminating signal peptides from transmembrane regions. Nature Methods 8:785-786. doi: 10.1038/nmeth.1701.

110 Plat A, Kuipers A, Crabb J, Rink R, Moll GN. 2017. Mutagenesis of nisin's leader peptide proline strongly modulates export of precursor nisin. Antonie Van Leeuwenhoek 110(3):321-330. doi: 10.1007/s10482-016-0802-6.

111 Pokusaeva K, Kuisiene N, Jasinskyte D, Rutiene K, Saleikiene J, Čitavičius D. 2009. Novel bacteriocins produced by *Geobacillus stearothermophilus*. Central European Journal of Biology 4(2):196–203. doi: 10.2478/s11535-009-0009-1.

112 Rea MC, Sit CS, Clayton E, O'Connor PM, Whittal RM, Zheng J, Vederas JC, Ross RP, Hill C. 2010. Thuricin CD, a posttranslationally modified bacteriocin with a narrow spectrum of activity against *Clostridium difficile*. PNAS 107(20):9352-7. doi: 10.1073/pnas.0913554107.

113 Ríos Colombo NS, Chalón MC, Navarro SA, Bellomio A. 2017. Pediocin-like bacteriocins: new perspectives on mechanism of action and immunity. Current Genetics 1-7. doi: 10.1007/s00294-017-0757-9.

114 Roy RS, Kim S, Baleja JD, Walsh CT. 1998. Role of the microcin B17 propeptide in substrate recognition: solution structure and mutational analysis of McbA1-26. Chemical Biology 5(4):217-28. doi: 10.1016/S1074-5521(98)90635-4.

115 Salomón RA, Farías RN. 1995. The peptide antibiotic microcin 25 is imported through the TonB pathway and the SbmA protein. Journal of Bacteriology 177:3323–3325. doi: 10.1128/jb.177.11.3323-3325.1995.

116 Sawa N, Zendo T, Kiyofuji J, Fujita K, Himeno K, Nakayama J, Sonomoto K. 2009. Identification and characterization of lactocyclicin Q, a novel cyclic bacteriocin produced by *Lactococcus* sp. strain QU 12. Applied and Environmental Microbiology 75(6):1552-8. doi: 10.1128/AEM.02299-08.

117 Schägger H. 2006. Tricine-SDS-PAGE. Nature Protocols 1(1):16-22. doi: 10.1038/nprot.2006.4.

118 Scholz R, Vater J, Budiharjo A, Wang Z, He Y, Dietel K, Schwecke T, Herfort S, Lasch P, Borriss R. 2014. Amylocyclicin, a novel circular bacteriocin produced by *Bacillus amyloliquefaciens* FZB42. Journal of Bacteriology. 196(10):1842-52. doi: 10.1128/JB.01474-14.

119 Shafia F. 1966. Thermocins of *Bacillus stearothermophilus*. Journal of Bacteriology 92(2):524-525.

120 Sharp RJ, Bingham AHA, Comer MJ and Atkinson A. 1979. Partial characterization of a bacteriocin (Thermocin) from *Bacillus stearothermophilus* RS93. Journal of General Microbiology 111:449-451.

121 Shimafuji C, Noguchi M, Nishie M, Nagao J, Shioya K, Zendo T, Nakayama J, Sonomoto K. 2015. In vitro catalytic activity of N-terminal and C-terminal domains in NukM, the post-translational modification enzyme of nukacin ISK-1. Journal of Bioscience and Bioengineering 120(6):624-9. doi: 10.1016/j.jbiosc.2015.03.020.

122 Shkundina I, Serebryakova M, Severinov K. 2014. The C-terminal part of microcin B is crucial for DNA gyrase inhibition and antibiotic uptake by sensitive cells. Journal of Bacteriology 196(9):1759-67. doi: 10.1128/JB.00015-14.

123 Simmonds RS, Pearson L, Kennedy RC, Tagg JR. 1996. Mode of action of a lysostaphin-like bacteriolytic agent produced by *Streptococcus zooepidemicus* 4881. Applied and Environental Microbiology 62(12):4536-41.

124 Simmonds RS, Simpson WJ, Tagg JR. 1997. Cloning and sequence analysis of *zooA*, a *Streptococcus zooepidemicus* gene encoding a bacteriocin-like inhibitory substance having a domain structure similar to that of lysostaphin. Gene 189(2):255-61.

125 Sit CS, McKay RT, Hill C, Ross RP, Vederas JC. 2011. The 3D structure of thuricin CD, a two-component bacteriocin with cysteine sulfur to α-carbon cross-links. Journal of American Chemical Society 133(20):7680-3. doi: 10.1021/ja201802f.

126 Sit CS, van Belkum MJ, McKay RT, Worobo RW, Vederas JC. 2011. The 3D solution structure of thurincin H, a bacteriocin with four sulfur to α-carbon crosslinks. Angewandte Chemie 50(37):8718-21. doi: 10.1002/anie.201102527.

127 Smith MK, Draper LA, Hazelhoff PJ, Cotter PD, Ross RP, Hill C. 2016. A bioengineered nisin derivative, M21A, in combination with food grade additives eradicates biofilms of *Listeria monocytogenes*. Frontiers Microbiology 7:1939. doi: 10.3389/fmicb.2016.01939.

128 Solbiati JO, Ciaccio M, Farías RN, González-Pastor JE, Moreno F, Salomón RA. 1998. Sequence analysis of the four plasmid genes required to produce the circular peptide antibiotic microcin J25. Journal of Bacteriology 181(8):2659-62.

129 Sprules T, Kawulka KE, Gibbs AC, Wishart DS and Vederas JC. 2004. NMR solution structure of the precursor for carnobacteriocin B2, an antimicrobial peptide from *Carnobacterium piscicola*. European Journal of Biochemistry 271:1748–1756. doi: 10.1111/j.1432-1033.2004.04085.x.

130 Stepper J, Shastri S, Loo TS, Preston JC, Novak P, Man P, Moore CH, Havlíček V, Patchett ML, Norris GE. 2011. Cysteine S-glycosylation, a new post-translational modification found in glycopeptide bacteriocins. FEBS Letters 585(4):645-50. doi: 10.1016/j.febslet.2011.01.023.

131 Strieker M, Tanović A, Marahiel MA. 2010. Nonribosomal peptide synthetases: structures and dynamics. Current Opinion in Structural Biology 20(2):234-40. doi: 10.1016/j.sbi.2010.01.009.

132 Subramanian S, Smith DL. 2015. Bacteriocins from the rhizosphere microbiome - from an agriculture perspective. Frontiers in Plant Science 6:909. doi: 10.3389/fpls.2015.00909.

133 Swe PM, Cook GM, Tagg JR, Jack RW. 2009. Mode of action of dysgalacticin: a large heat-labile bacteriocin. Journal of Antimicrobial Chemotherapy. 63(4):679–686. doi: 10.1093/jac/dkn552.

134 Tietz JI, Schwalen CJ, Patel PS, Maxson T, Blair PM, Tai HC, Zakai UI, Mitchell DA. 2017. A new genome-mining tool redefines the lasso peptide biosynthetic landscape. Nature Chemical Biology 13(5):470-478. doi: 10.1038/nchembio.2319.

135 Tosukhowong A, Zendo T, Visessanguan W, Roytrakul S, Pumpuang L, Jaresitthikunchai J, Sonomoto K. 2012. Garvieacin Q, a novel class II bacteriocin from *Lactococcus garvieae* BCC 43578. Applied and Environmental Microbiology 78(5):1619-23. doi: 10.1128/AEM.06891-11.

136 Tymoszewska A, Diep DB, Wirtek P, Aleksandrzak-Piekarczyk T. 2017. The non-lantibiotic bacteriocin garvicin Q targets Man-PTS in a broad spectrum of sensitive bacterial genera. Scientific Reports 7(1):8359. doi: 10.1038/s41598-017-09102-7.

137 van Belkum MJ, Martin-Visscher LA, Vederas JC. 2010. Cloning and characterization of the gene cluster involved in the production of the circular bacteriocin carnocyclin A. Probiotics and Antimicrobial Proteins. 2(4):218-25. doi: 10.1007/s12602-010-9056-1.

138 van Belkum MJ, Martin-Visscher LA, Vederas JC. 2011. Structure and genetics of circular bacteriocins. Trends in Microbiology 19(8):411-8. doi: 10.1016/j.tim.2011.04.004.

139 van Heel AJ, de Jong A, Montalbán-López M, Kok J, and Kuipers OP. 2013. BAGEL3: automated identification of genes encoding bacteriocins and (non-)bactericidal posttranslationally modified peptides. Nucleic Acids Research 41(W1):W448–W453. doi: 10.1093/nar/gkt391.

140 Venugopal H, Edwards PJ, Schwalbe M, Claridge JK, Libich DS, Stepper J, Loo T, Patchett ML, Norris GE, Pascal SM. 2011. Structural, dynamic, and

chemical characterization of a novel S-glycosylated bacteriocin. Biochemistry 50(14):2748–2755. doi: 10.1021/bi200217u.

141 Von Tersch MA, Carlton BC. 1983. Bacteriocin from *Bacillus megaterium* ATCC 19213: comparative studies with megacin A-216. Journal of Bacteriology 155(2):866-71.

142 Wachsman MB, Castilla V, de Ruiz Holgado AP, de Torres RA, Sesma F, Coto CE. 2003. Enterocin CRL35 inhibits late stages of HSV-1 and HSV-2 replication in vitro. Antiviral Research 58(1):17-24. doi: 10.1016/S0166-3542(02)00099-2.

143 Walsh CT, Wencewicz TA. 2014. Prospects for new antibiotics: a molecule-centered perspective. Journal of Antibiotics. 67(1):7-22. doi: 10.1038/ja.2013.49.

144 Wang H, Oman TJ, Zhang R, Garcia De Gonzalo CV, Zhang Q, and van der Donk WA. 2013. The glycosyltransferase involved in thurandacin biosynthesis catalyzes both O- and S-glycosylation. Journal of the American Chemical Society 136(1):84-7. doi: 10.1021/ja411159k.

145 Wang S, Wang Q, Zeng X, Ye Q, Huang S, Yu H, Yang T, Qiao S. 2017. Use of the antimicrobial peptide sublancin with combined antibacterial and immunomodulatory activities to protect against methicillin-resistant *Staphylococcus aureus* infection in mice. Journal of Agricultural and Food Chemistry 65(39):8595-8605. doi: 10.1021/acs.jafc.7b02592.

146 Wieckowski BM, Hegemann JD, Mielcarek A, Boss L, Burghaus O, Marahiel MA. 2015. The PqqD homologous domain of the radical SAM enzyme ThnB is required for thioether bond formation during thurincin H maturation. FEBS Letters 589(15):1802-6. doi: 10.1016/j.febslet.2015.05.032.

147 Wirawan RE, Swanson KM, Kleffmann T, Jack RW, Tagg JR. 2007. Uberolysin: a novel cyclic bacteriocin produced by *Streptococcus uberis*. Microbiology 153(Pt 5):1619-30. doi: 10.1099/mic.0.2006/005967-0.

148 World Health Organization. 2015. Global Action Plan on Antimicrobial Resistance. WHO Document Production Services, Geneva, Switzerland. http://www.who.int/iris/bitstream/10665/193736/1/9789241509763_eng.pdf?ua=1.

149 Yule R and Barridge BD. 1976. Isolation and characterization of a bacteriocin produced by *Bacillus* s*tearothermophilus* strain NU-10. Canadian Journal of Microbiology 22(12):1743-1750. https://doi.org/10.1139/m76-257.