**Preprint Title** Genome Mining Reveals the Structure of the Clostrisin and Cellulosin Biosynthetic Gene Clusters, Two Novel Lantibiotics

**Authors** Moisés Alejandro Alejo Hernandez, Katia Pamela Villavicencio Sánchez, Rosendo Sánchez Morales, Karla Georgina Hernández-Magro Gil, David Silverio Moreno-Gutiérrez, Eddie Guillermo Sanchez-Rueda, Yanet Teresa-Cruz, Armando Hernández Garcia, Alba Romero-Rodríguez, Oscar Juárez, Siseth Martínez-Caballero, Mario Figueroa and Corina Diana Ceapa

**Publication Date** 06 Dez. 2023

**Article Type** Full Research Paper

**Supporting Information File 1** Supporting Information.docx; 1.6 MB

**ORCID® iDs** Rosendo Sánchez Morales - https://orcid.org/0009-0002-0687-6591; Armando Hernández Garcia - https://orcid.org/0000-0002-2401-8139; Corina Diana Ceapa - https://orcid.org/0000-0001-8661-4211

# Genome Mining Reveals the Structure of the Clostrisin and Cellulosin Biosynthetic Gene Clusters, Two Novel Lantibiotics

Moisés Alejandro Alejo Hernandez[1], Katia Pamela Villavicencio Sánchez[1], Rosendo Sánchez Morales[1], Karla Georgina Hernández-Magro Gil[1], David Silverio Moreno-Gutiérrez[1,2], Eddie Guillermo Sanchez-Rueda[2], Yanet Teresa-Cruz[3], Armando Hernández Garcia[2], Alba Romero-Rodríguez[3], Oscar Juárez[4], Siseth Martínez-Caballero[5], Mario Figueroa[6], Corina-Diana Ceapă[1]

Addresses:

[1] Laboratory of Microbiology, Institute of Chemistry, National Autonomous University of Mexico (UNAM), Mexico City, Mexico.

[2] Biomolecular Engineering and Bionanotechnology Laboratory, Institute of Chemistry, UNAM, Mexico City, Mexico.

[3] Laboratory A-107, Biomedical Research Institute, UNAM, Mexico City, Mexico.

[4] Department of Biological Sciences, Illinois Institute of Technology, Chicago, Illinois 60616.

[5] Institute of Chemistry, National Autonomous University of Mexico, Mexico City, Mexico.

[6] Facultad de Química, Universidad Nacional Autónoma de México (UNAM), Ciudad de México, México.

To whom correspondence should be addressed: Laboratory of Microbiology, Institute of Chemistry, National Autonomous University of Mexico (UNAM), Mexico City, Mexico.
Email: corina.ceapa@iquimica.unam.mx

## Abstract

Antimicrobial resistance is an issue of enormous proportions that poses a significant threat to global public health. This study focuses on lanthipeptides, ribosomally encoded peptides that display significant structural diversity and hold promising potential as antibiotics. Genome mining was employed to locate biosynthetic gene clusters (BGC) for class II lanthipeptides. A phylogenetic study analyzing homologous sequences of functional LanM sequences revealed a unique evolutionary clade of 39 LanMs associated with 28 Clostridial bacterial genomes. *In silico* exploration identified nine complete BGCs, including one from *Clostridium cellulovorans* 743B, that harbors two new lanthipeptides: Clostrisin and Cellulosin. The lanthipeptide BGCs were heterologously expressed in *Escherichia coli*. Molecular weights associated with the expected post-translational modifications of the purified lanthipeptides were confirmed by MALDI-TOF. Both peptides demonstrated antimicrobial activity against multidrug-resistant bacteria, such as a clinical strain of *Staphylococcus epidermidis* MIQ43 and a reference strain of resistant *Pseudomonas aeruginosa* PA14. This study showcases the immense potential of genome mining in identifying new class II lanthipeptides.

## Keywords

Antimicrobials, multi-drug resistant bacteria, natural products, *Clostridium* sp.

## Introduction

Antimicrobial resistance (AMR) is a significant public health challenge. Only in 2019, there were 4.95 million deaths associated with AMR [1], a number expected to increase exponentially. One fundamental objective of the Global Action Plan on Antimicrobial Resistance by the World Health Organization (WHO) is the investment in developing new drugs, diagnostic tools, vaccines, and other interventions [2]. In this context, many antibiotics have been derived from bacterial natural products (NPs), which have proven to be a valuable source of antimicrobial agents. During the latter part of the 20th century, the discovery of NPs was hindered by the limitation of traditional methods, which often led to the rediscovery of previously identified NPs. Next-generation whole genome sequencing technologies have a newfound ability to explore and identify biosynthetic gene clusters (BGCs) responsible for NP production. This renewed focus offers the advantage of preventing redundant discovery and predicting novelty, resistance, and bioactivities. Furthermore, the increasing availability of genomic data has led to the development of bioinformatics tools, such as antiSMASH[3], Bagel 4[4], and RiPPMinner[5], that have emerged to streamline the process of exploring and discovering BGCs in bacteria, known as genome mining [6 – 8].

Lanthipeptides are a class of ribosomally synthesized and post-translationally modified peptides (RIPPs). Lanthipeptide biosynthetic genes have undergone complex evolutionary processes that have produced chemically diverse active peptides [9]. These genes are predominantly found in bacteria and have evolved through selective pressures driven by competition for resources and defense against predators[10]. The resulting peptides exhibit unique structural features due to the presence of thioether bridges between serine and threonine residues dehydrated (Dha/Dhb) Subsequent cyclization with cysteine residues yields cyclic peptides with the modified amino acid lanthionine (Dha-Cys) and methyllanthionine (Dhb-Cys). Additional post-translational modifications contributing to their distinct biological activities were identified [11 - 13].

The essential biosynthesis of lanthipeptides involves the translation of lanA genes, which encode a precursor peptide with two segments: an N-terminal leader peptide recognized by post-translational enzymes and a C-terminal core peptide where the lanthionine and methyllanthionine are formed. The post-translational modification enzymes involved are often multifunctional, and their activity is tightly controlled at multiple levels to ensure the production of high-quality peptides[14]. To date, five different classes of biosynthetic machinery have been identified [15] and are the basis for classifying lanthipeptide gene clusters. For class II lanthipeptides, discussed in this work, multidomain LanM enzymes catalyze the formation of lanthionine and methyllanthionine[16]. Moreover, lanPt genes encode a membrane protein with two ABC transporter domains and a C39 peptidase domain. This protein cleaves the leader peptide to generate mature lanthipeptides and exports them to the extracellular environment [17].

The diversity of lanthipeptides produced by these biosynthetic pathways has led to the discovery of new compounds with potential therapeutic applications. Class I and II lanthipeptides have exhibited antimicrobial activity against various pathogens, including drug-resistant strains [10]. These molecules are relevant as they have strong antibiotic effects against pathogenic bacteria such as *Staphylococcus aureus*, *Enterococcus*, *Clostridioides difficile*, and *Mycobacterium tuberculosis*. These effects are attributed to their affinity for the lipid II component of Gram-positive bacterial cell walls [18]. Additionally, there have been reports of lantibiotics such as CMB001 displaying activity against resistant Gram-negative bacteria, including *Acinetobacter baumannii* [19]. In addition, some lanthipeptides have been shown to have anticancer and immunomodulatory properties [20, 21].

A low development of resistance to lanthipeptides has been observed, with nisin being the most well-known case in the food industry. While some strains exhibit resistance due to changes in the cell wall, biofilm formation, or the expression of resistance proteins such as ABC transporters or proteases [22], specific mutations in nisin have rendered previously resistant strains susceptible[23]. The structural

diversity of these peptides, coupled with their successful production in *Escherichia coli*[24] has driven research to consider them as a source for new antibiotics, as indicated in the present study.

# Results and discussion

In recent years, there has been a significant increase in the number of lanthipeptides discovered, along with a deeper understanding of their evolution, ecological importance, and gene composition [25]. Many of these novel peptides show antimicrobial properties [26]. Identifying the genetic and biochemical mechanisms underlying the production of these compounds can aid in developing effective antibiotics against drug-resistant pathogens with each discovery.

**Genome Mining of LanM enzymes sequence diversity to localize novel evolutive clades.**

Genome mining has emerged as a crucial research area in discovering novel antimicrobial compounds[27]. Due to its unique position between basic and applied research, it has become an essential tool for identifying compounds with potential therapeutic applications. Genetic and genomic data mining enables the study of the evolution of genes and genomes across diverse species and populations. This provides insights into the origins and evolution of genes that have evolved as "tools" in the ongoing biological battles in the microbial world. The biodiversity of specialized metabolites is strain-specific, meaning that even closely related organisms can have distinct metabolic capabilities. Consequently, researchers can unveil new clusters and metabolites by tracking the conserved genes participating in biosynthetic processes.

Recent years have witnessed notable successes in using genome mining to discover novel antimicrobial compounds. For example, the discovery of teixobactin[28] has paved the way for developing antibiotics with innovative mechanisms of action. It has opened new avenues for developing antibiotics with novel mechanisms of action. Similarly, the discovery of halicin, a molecule identified through a machine learning-based approach, has shown promise as a broad-spectrum antimicrobial agent. Halicin is effective against various pathogens, including some resistant to existing antibiotics [25]. Thus, peptides, including ribosomally synthesized and post-translationally modified peptides (RiPPs), have been regarded as important sources of antibiotics, both historically and through recent discoveries[30]. Structural diversity is a key characteristic of RIPPs, a class of natural products many microorganisms produce. RiPPs[31] have many biological activities, including antimicrobial, antitumor, and immunomodulatory properties. Lanthipeptides are a sub-class of RiPPs characterized by thioether bonds, and lanthionine rings confer stability and rigidity to the peptide backbone. The structural diversity of lanthipeptides is driven by multiple modification enzymes, which can modify the amino acid residues in the peptide backbone in a highly selective and specific manner[16]. Low resistance is another important feature of RiPPs, particularly lanthipeptides, which are highly effective against drug-resistant pathogens. This is due to their unique mode of action, which involves the disruption of the bacterial cell membrane or cell wall[18]. The biological activities of lanthipeptides are also dependent on their structural diversity, with different modifications leading to different biological effects. For example, some lanthipeptides have been found to have potent antibacterial activity, while others are effective against cancer cells. The study of structural diversity, low resistance, and biological activities of RiPPs and lanthipeptides is an active area of research with many potential applications in medicine and biotechnology[16,21].

This study used 28 experimental amino acid sequences of LanM enzymes from MIBiG [26] to identify homologous sequences using BLAST [27] (in BV-BRC [28]). As a result, 315 homologous proteins were considered for a phylogenetic study (the selection criteria were over 30% similarity over 80% of protein sequence length). A phylogenetic tree was constructed (Figure 1A) with homologous and experimental LanM enzyme amino acid sequences. The phylogenetic tree allowed us to identify novel

evolutionary clades based on the taxonomy of the organisms to which the LanMs belong. The Clostridia clade was selected for further study (Figure 1B). Our analysis also indicated the presence of other yet unexplored LanM clades, such as the one from Bacilli, Cyanobacteria, and Actinobacteria, which could be interesting to study further due to their sequence uniqueness.

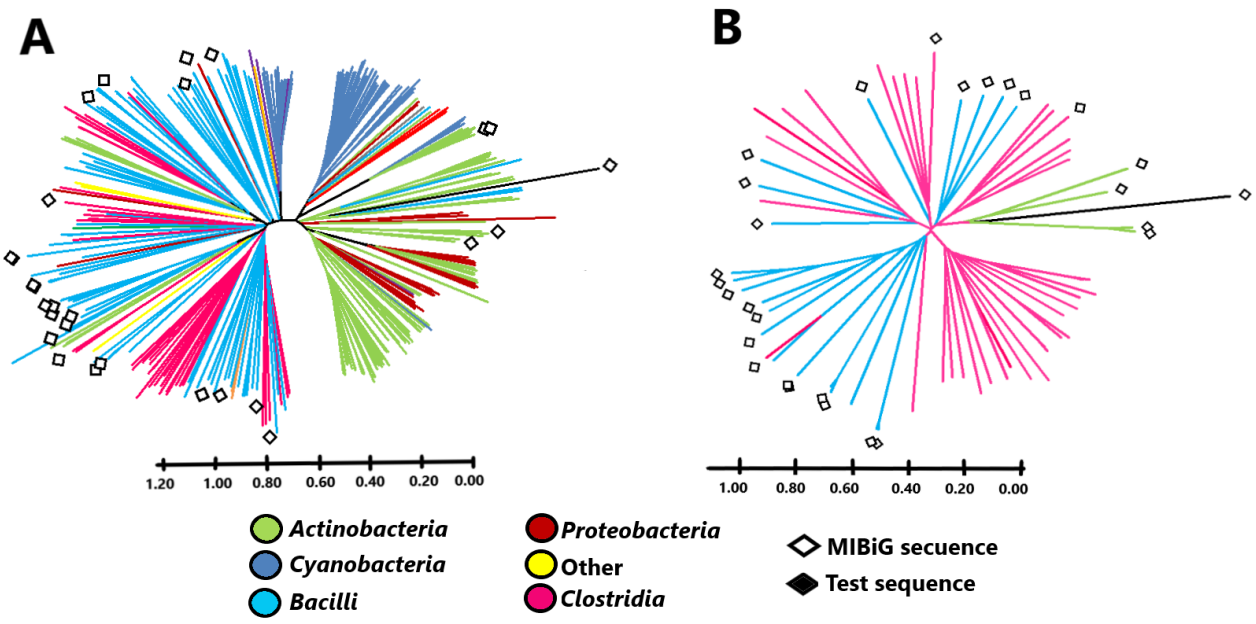


**Figure 1. A.** Unrooted phylogenetic tree of all the experimental LanM sequences identified by BLAST. This analysis involved 343 amino acid sequences; there were a total of 1964 positions in the final dataset. **B.** Unrooted phylogenetic tree of the Clostridia clade. This analysis involved 60 amino acid sequences; there were 1366 positions in the final dataset. Phylogenetic trees were constructed from the alignments using the Neighbor-Joining [29], and the evolutionary distances were calculated using the Poisson correction method. A LanL class IV lantibiotics-modifying enzyme from Venezuelin[36], was used as a test sequence. Square symbols refer to previously characterized LanM sequences from the MIBiG database.

**Identification and selection of Clostrisin and Cellulosin BGCs**

The Clostridia clade of LanM proteins from our dataset consists of 41 sequences associated with 28 genomes. After analyzing the genomes with antiSMASH [3], we identified 232 different BGCs of NPs. The RiPPs accounted for 46% (107 clusters). A total of 37 BGCs were identified to be responsible for Class II lanthipeptides, which in turn helped determine the presence of precursor peptides, gene duplications, paralogues, transporters, and resistance genes. From these BGCs, 23 competent precursor peptides were identified based on their physicochemical characteristics, the presence of serine and threonine residues, and the characteristic cleavage site of the C39 peptidases domain, which is represented by the sequences GG or GA (Supplementary Table 1). We then selected fifteen precursor peptides in 9 BGCs deemed competent due to the presence of essential genes lanA, lanM, and lanPt. These genes were found to be the minimal machinery required for Class II lanthipeptide biosynthesis (Figure 2). Notably, transporter genes and regulatory elements were not considered in the second filter as their presence is not essential for biosynthesis, albeit it is for resistance. As a result, the *C. cellulovorans* 743[37] cluster was chosen because it was thought to be complete, competent, novel, and most interesting, primarily due to the presence of two different, apparently independent clusters,

each with its lanthipeptide precursor (Figure 2). The nucleotide sequence and amino acid sequence of gens ClosA, ClosM, ClosPt, CellA, CellM, and CellPt are reported in the Supporting Information.

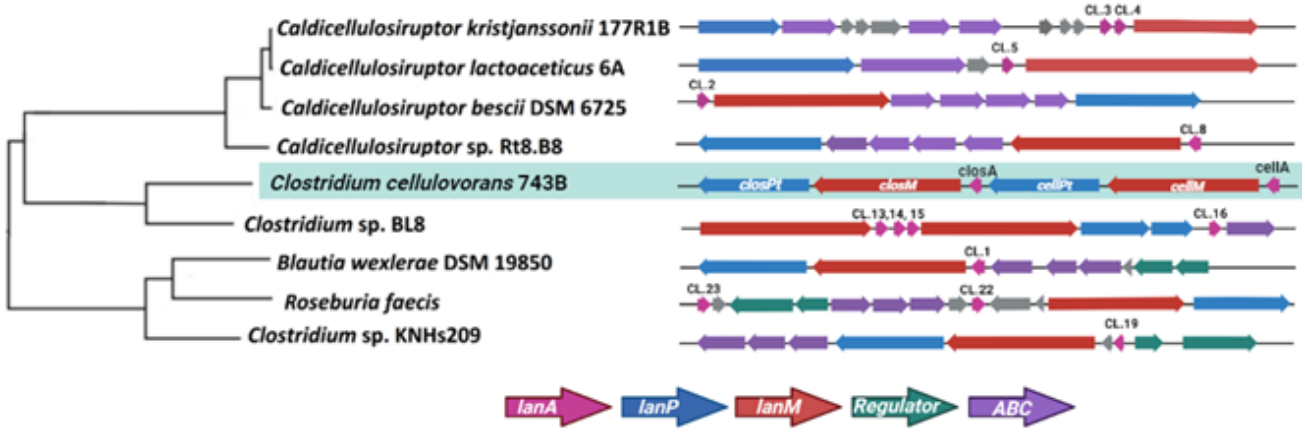


**Figure 2.** A phylogenetic tree built with 1000 shared clostridial genes (left) with the Lanthipeptide II BGC gene structures considered competent (right).

To confirm the novelty of the lanthipeptides from the cluster of *C. cellulovorans* 743, the precursor peptides pre-Clostrisin and pre-Cellulosin were analyzed, generating a similarity network constructed using the percent identity between the precursor peptides from all experimentally characterized lanthipeptide classes to date (Figure 3.A.). The positions of Clostrisin and Cellulosin in the similarity network were notable, as they exhibited connections with groups of characterized lanthipeptides while maintaining sufficient sequence divergence based on an identity percentage of no more than 60%. RiPPMinner [5] predicts 4 and 5 possible lanthionine cycles for the precursors. The amino acid sequence of the precursors shares some similarity with the characterized peptides FlvA.2g [38], Enterocin Wβ [39], Plantirocin Wβ [40], Thusin α[41], and for Cellulosin with Lichenicidin A1[42] FlvA.1[38], FlvA.2[38], BhtA α[43], SmB β[44], Smb α[44], Thusin β[35] (Figure 3.B.).

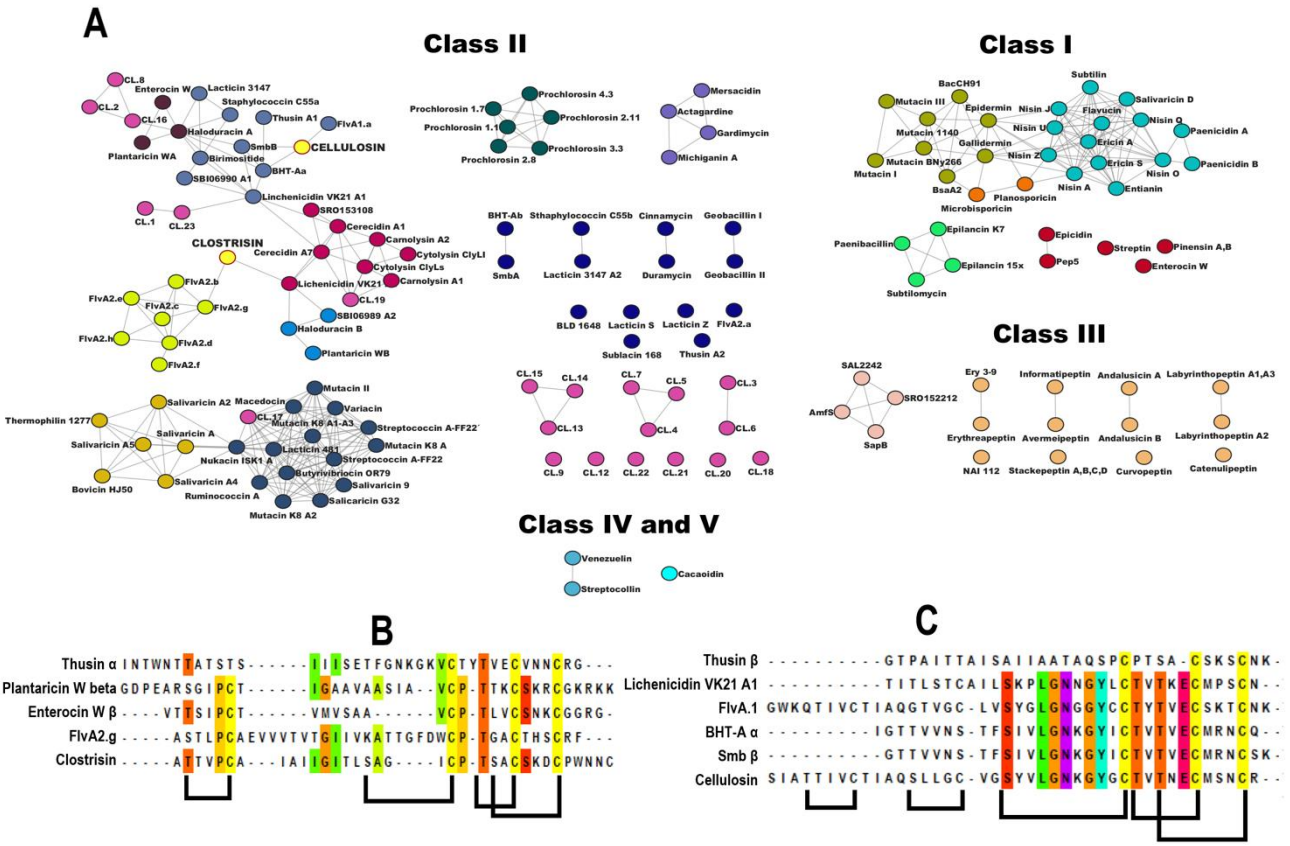


**Figure 3 A.** In the similarity network of precursor peptides, each node represents a characterized lanthipeptide (identified in pink) or one of all experimental lanthipeptides (each color represents a different family), and each line represents a minimum of 27% of sequence identity (EFI - Enzyme Similarity Tool [39]). **B.** Alignment of the central peptide of Clostrisin with FlvA.2g[38], Enterocin Wβ[39], Plantirocin Wβ[40], Thusin α[41] **C.** Alignment of the central peptide of Cellulosin with Lichenicidin A1[42] FlvA.1[38], FlvA.2[38], bhtA α[43], SmB β[44], Smb α[44], Thusin β[41].

### *In silico* characterization of the post-translational modification enzymes of Clostrisin and Cellulosin

The Clostrisin and Cellulosin clusters comprise two transporter protein-peptidases, that we named ClosPT and CellPt, and two LanM enzymes, ClosM and CellM, as well as the precursor peptides (Figure 2). The amino acid sequences of the C39 peptidase domain of ClosPt and CellPt were subjected to BLAST analysis, revealing homologous sequences from other Clostridial species. A phylogenetic tree was constructed based on this sequence showing various organisms within the same class (as detailed in Supplementary Figure 1.A). Additionally, structure models for ClosPt and CellPt were generated using AlphaFold 2.0 [46,47], (Supplementary Figure 1.B and 1.D) and compared to the protein PCAT1 [17] (the closest homologous protein with a structure resolved by crystallography, here presented in complex with its peptide ligand (PDB 6V9Z). The amino acid sequence identity between ClosPt and PCAT1 was 29%, while for CellPt, it was 24%. Structural alignment revealed RMSD values below 3 Å, meeting the minimum criteria for structural conservation (Supplementary Table 2). Furthermore, the catalytic residues within the peptidase domains were confirmed. ClosPt catalytic residues are Cys35 and His111; for CellPt, they are Cys18 and His 92. These residues maintained a distance and structural positions like the catalytic residues in PCAT1 (where the catalytic residues are Cys21 and His99).

A similar procedure was carried out for ClosM and CellM, leading to the identification of the closest sequences within the Clostridial class (Supplementary Figure 2.A). Structure models for ClosM and CellM generated using AlphaFold 2.0 (Supplementary Figure 2.B and 2.C). were compared with the CylM[48] protein (PDB 5DZT), as it remains the sole LanM enzyme characterized through crystallography to date. The amino acid sequence identity between ClosM and CylM was 23%, while for CellM, it was 24%. Structural alignment was performed, and the obtained RMSD values were below 3 Å (Supplementary material, Table 3). Our findings suggested that these peptides and their associated biosynthetic enzymes hold promise as novel entities with untapped bioactive potential.

**Expression and purification of pre-Clostrisin, pre-Cellulosin, and C39 peptidase domain**

As we could not access the producing microorganisms, we synthesized the gene sequences reported in public databases (Supplementary List 1) *de novo*. To express the biosynthetic machinery for the two lanthipeptides, we transformed the *E. coli* strain NiCo21(DE3) with the following vectors: pRSF-Duet_ClosA_ClosM, pRSF-Duet_CellA_CellM, and pRSF-Duet_C39 peptidase domain. The latter contains the peptidase domain of ClosPt, as the complete ABC transporter was deemed too large to synthesize and express in this system. This construct was functional in previous studies and retained catalytic activity without the rest of the transporter domains [43]. Electrophoresis and Western Blot confirmed the expression of all the protein and peptide molecules. After purification, we estimated yields between 0,8 and 1,5 g/L of culture for all products.

**The activity of the C39 peptidase domain cleaving the leader peptide of pre-Clostrisin and pre-Cellulosin**

The proteolytic activity was monitored by mixing the purified samples of C39 peptidase domain with purified pre-Clostrisin and pre-Cellulosin. The SDS-PAGE electrophoretic pattern confirmed the proteolytic reaction (Supplementary Figure 3.D). This was further confirmed through MALDI-TOF analysis.

**Chemical characterization of pre-Clostrisin and pre-Cellulosin through MALDI-TOF MS**

Samples containing pre-Clostrisin, pre-Cellulosin, Clostrisin, and Cellulosin were characterized using MALDI-TOF. Expected masses are presented in Table 4, and the full MALDI-TOF mass spectrum in Figure 4.A-D of the Supporting Information. In Figure 4.A, an ion at *m/z* 3252.91 was associated to [clostrisin-$4H_2O$+H] +, while in Figure 4.B, the pre-clostrisin ion was observed at *m/z* 8609.6 [pre-clostrisin-$4H_2O$+H] +. Four dehydrations from a maximum of seven possible amino acid residues subject to dehydration were observed in both cases. The presence of these masses confirmed the *in-silico* predictions that the leader peptide cleavage is performed at a GA amino acid sequence instead of the regular GG site, which is used in most lanthipeptide precursors.

In the case of pre-cellulosin and cellulosin In Figure 4.D and 4C, the ions at *m/z* 9288.31 [pre- cellulosin $-4H_2O + H]^+$ and *m/z* 4246.06 [cellulosin -$5H_2O$+H] $^+$ confirmed the presence of the predicted peptides (Figure 7). The leader peptide was cleaved at the GG amino acid sequence in this case.

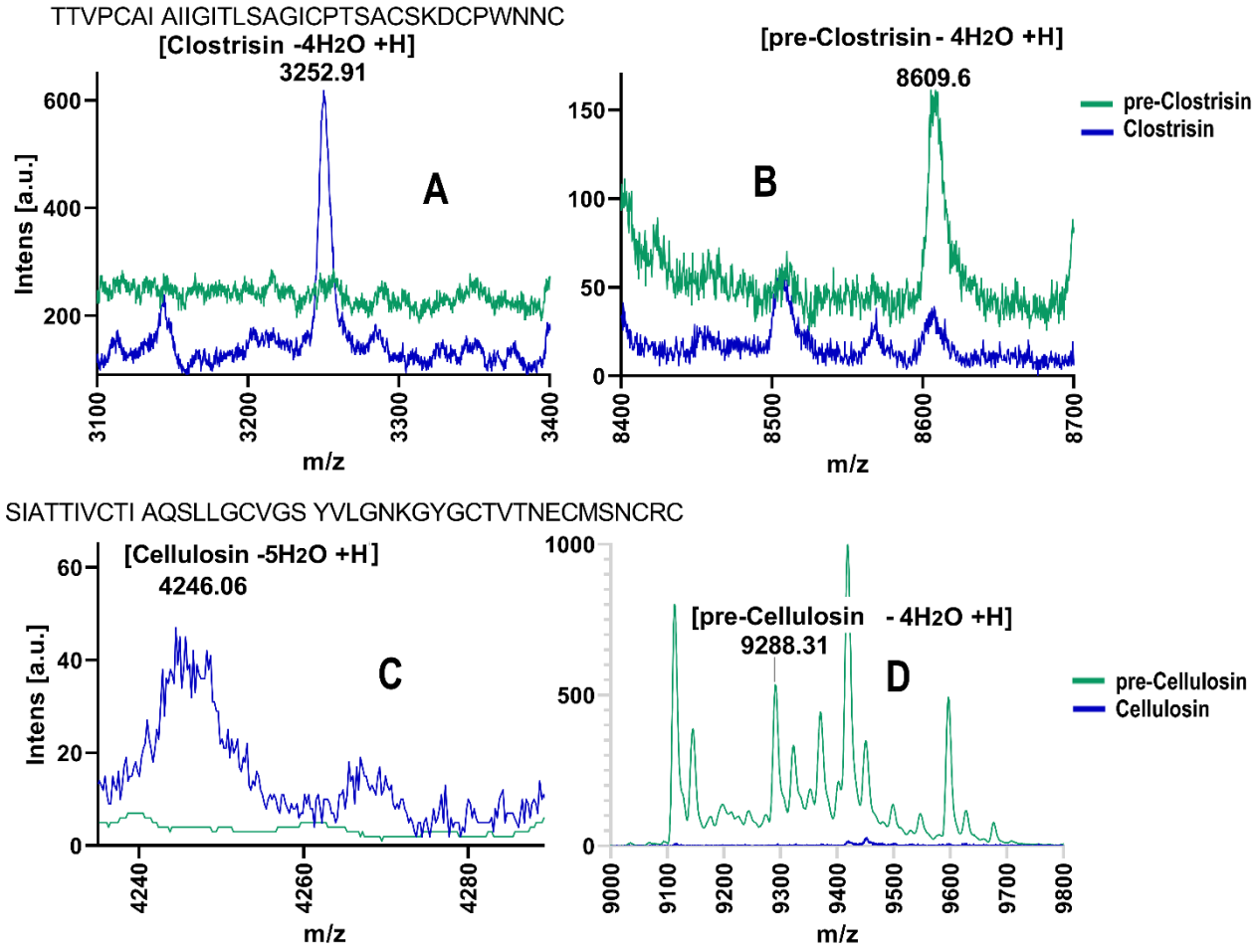


**Fig. 4.** MALDI-TOF mass spectrophotometry of purified fractions of **A.** Clostrisin **B.** Pre-Clostrisin **C.** Cellulosin **D.** Pre-Cellulosin.

**Antimicrobial Activity of Clostrisin and Cellulosin**

Antimicrobial activity assays of Clostrisin and Cellulosin were performed using microplate bioassays with strains from the American Type Culture Collection (ATCC): *Staphylococcus aureus* 43300 (MRSA), *Pseudomonas aeruginosa* ATCC 15442 (PA14), *Escherichia coli* IM08B as well as clinical strains isolated from patients in Mexico: *Staphylococcus epidermidis* MIQ43 (multidrug-resistant clinical sample) (internal code from the MicroIQ laboratory library), and *Pseudomonas aeruginosa* MIQPA25 [50] (multidrug-resistant clinical sample isolated from cystic fibrosis patients), and *Clostridioides difficile* R20291. All strains have a multidrug-resistant profile (Supplementary Table 5). These bacteria were chosen due to their classification in the ESKAPE group and the species of origin (in the case of Clostridioides).

We tested the purified pre-Clostrisin and pre-Cellulosin precursor peptides alongside a mixture with the C39 peptidase domain. The latter mature lanthipeptides (following the cleavage of the leader peptide by the C39 peptidase domain), will be referred to as Clostrisin and Cellulosin. Each assay was performed at the highest obtainable concentrations: 1.4, 2.8, and 5.6 µg/mL of pre-Clostrisin and 1.2, 2.4, and 4.8 µg/mL of pre-Cellulosin. We used the same concentrations of precursor peptides in mixtures of Clostrisin and Cellulosin. In our testing against *Clostridioides difficile*, we evaluated pre-

Clostrisin at 5.6µg/mL and pre-Cellulosin at 4.8 µg/mL. Additionally, to ensure that the observed activity was not due to the C39 peptidase domain protein, we tested them on each bacterium at 10.5 µg/mL, where they showed no inhibition.

The peptides exhibited no antibiotic activity against *Escherichia coli* ATCC IM08B, *Clostridioides difficile* R20291, *Acinetobacter baumannii* ATCC BAA 747, and *S. aureus* ATCC 43300 (MRSA) at the concentrations employed. It was expected that the lack of effect on *E. coli* would occur, as the pre-peptide accumulation did not impact the expression host. This assumption was made due to reports of auto-proteolysis for other lanthipeptides during production in this cell system. This leads us to believe that at least a few mature peptides might have been present. Despite this, the growth curves of the producing strains were not distinct from the control ones, transformed with the empty vector. As a result, it was not surprising that there was no effect on the *E. coli* test strains, which confirmed previous observations.

For *S. epidermidis* MIQ43, *P. aeruginosa* ATCC PA14, and MIQPA25, both Clostrisin and Cellulosin samples displayed robust bacteriostatic activity at the highest estimated concentration of 5.6 µg/mL and 4.8 µg/mL respectively, compared to the bacteria grown in LB media (Control), and lower but statistically significant effects at lower concentrations. As expected, the immature peptides pre-Clostrisin and pre-Cellulosin showed no discernible effect at the same concentration. These findings highlight the activity acquisition after the leader peptide's proteolysis and the presence of mature lanthipeptides (Figure 5).

Finally, it was extremely interesting to test the lanthipeptides on *C. difficile*, due to the origin of these peptides in the Clostridia clade. The strict ecological environments of Clostridia are the human and animal microbiota and soil for both anaerobic and aerotolerant organisms. We interpret a lack of effect on *C. difficile*, an intestinal clinical isolate, due to either the presence of resistance mechanisms or the fact that these peptides evolved to be active against soil microbiota members. The habitat of *C. cellulovurans*, the strain of origin of the lanthipeptide clusters, appears to be plant-associated due to its metabolism and isolation environment (wood).

Our results raise numerous questions regarding the ecological functions of the lantibiotics we discovered. Our bioinformatics studies have not provided any clues about the resistance mechanisms of these peptides. Therefore, we plan to conduct further research to gain insights into these functions. This could help us understand the potential for these peptides to display antimicrobial resistance in the future.

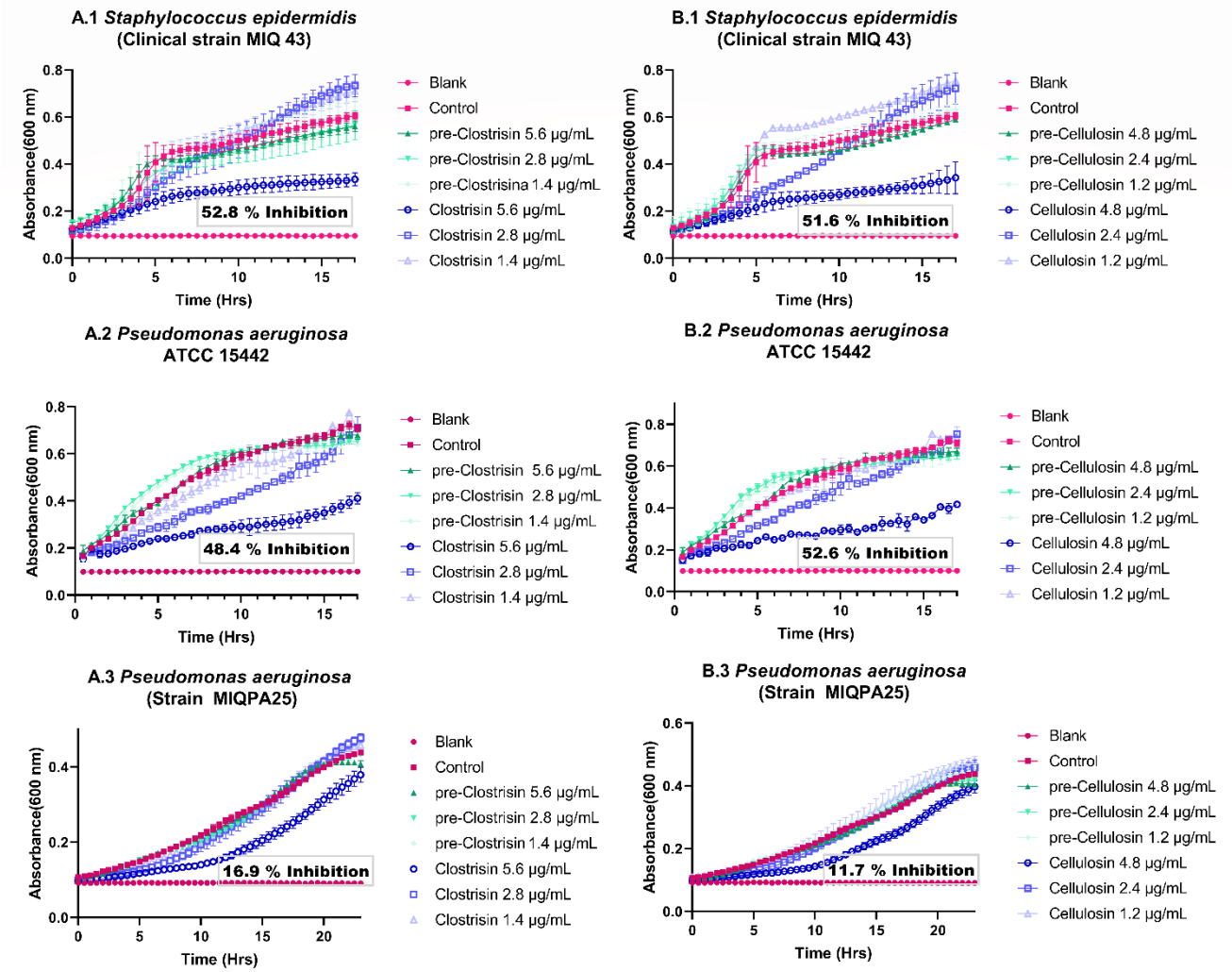


**Figure 5.** Growth curves of the strains with the bacterial activity of the samples **A. Clostrisin** and **B. Cellulosin** with the bacteria with **1.** *S. epidermidis* MIQ43 and the **2.** *P. aeruginosa* ATCC PA14 and **3.** MIQPA25. Tests were performed in sterile polystyrene 96-well flat-bottom-shaped microtiter plates. Inoculum densities were adjusted at an $OD_{600nm}$ of 0.1. The plates were incubated at 36°C for 16h at 220 rpm, and the optical density (OD) was measured at 600 nm.

## Atomic Force Microscopy

The study employed atomic force microscopy in air (AFM) to observe the bacterial morphological changes on the surface of *S. epidermidis* triggered by the impact of lanthipeptides. The samples were treated with an estimated 22.4 ng/µL concentration for 1 and 5 hours. The AFM images of untreated *S. epidermidis* MIQ43 displayed characteristic clustered cocci (see Figures 6.A & 6.F). The treatment with pre-Cellulosin and Cellulosin samples resulted in more pronounced morphological changes after 5 hours of treatment. The changes included the loss of the characteristic cocci structure, with the formation of blebs that increased the membrane's rugosity. Besides, cytosolic leakage and remnants of the outer membrane were identified (see Figures 6. B, C, G, H). The treatment with pre-Clostrisin and Clostrisin samples led to the formation of numerous blebs visible after 1 hour. Interestingly, no cocci structures were detected after 5 hours of treatment (see Figure 6. D, E, I, J). Based on the observations, it can be inferred that Clostrisin and Cellulosin directly interact with the membrane of *S. epidermidis*. However, no significant differences were found in the samples regarding using the peptidase C39 peptidase domain. The findings of this study provide valuable insights into the effects of lanthipeptides

on *S. epidermidis*, and confirm the previous results acquired during the antimicrobial activity assays. Further studies are needed to confirm the mechanism of action of Clostrisin and Cellulosin and their molecular targets in *S. epidermidis*, a Gram-positive bacterium.

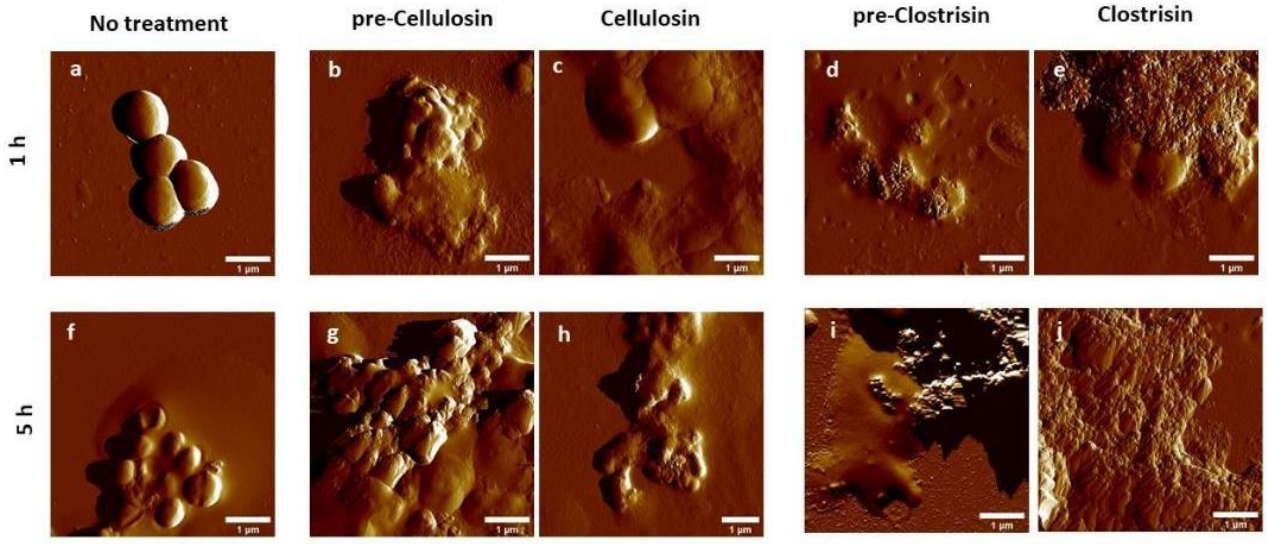


**Figure 6. Atomic Force Microscopy images (peak force mode) of *S. epidermidis* MIQ43 incubated with the different samples for 1 hour (a-e) and 5 hours (f-j).** From left to right: no treatment, pre-Cellulosin, Cellulosin, pre-Clostrisin, Clostrisin. All samples were evaluated at an estimated concentration of 22.4 ng/µL. Scale bar: 1 µm.

# Conclusion

The pressing need for new antibiotics has made the search for them a top priority amongst the scientific community. In this study, we utilized genomic mining to forecast the uniqueness of two lanthipeptides derived from the class Clostridia. These peptides were expressed in an *E. coli* host and produced at substantial concentrations to perform bioactivity assays. The C39-associated domain was purified to facilitate the transformation from pre-peptide to mature configuration. The mature Clostrisin and Cellulosin exhibited antimicrobial properties against *S. epidermidis* and *P. aeruginosa*-resistant strains and failed to exhibit the same effects against *E. coli* and *C. difficile*. As a result, we present the discovery of two new lantibiotics that have demonstrated efficacy in combating Gram-positive and Gram-negative pathogens.

# Author contributions

CDC, MF, and MAAH developed the idea, performed experiments, and wrote and edited the manuscript. PKVS, RSM, KGHMG, DSMG, EGSR, YTC, AHG, ARR, OJ, and SMC performed experiments and wrote and edited the manuscript.

# Acknowledgments and funding

This research was financially supported by PAPIIT (project numbers IA201721, IA205122, IN203923) and CONACHYT (CF-2022/319816, CF-2022/319596). MAAH, EGSR acknowledges financial support from CONAHCYT, Mexico, for postgraduate studies. We thank LESP Aguascalientes, Dr. Rodolfo Garcia Contreras from the Faculty of Medicine, UNAM, Mexico, and Dr. Victor Coria, National Pediatric Institute, Mexico, for sharing the multi-resistant strains. We thank Dr. Daniel

Finkelstein for the grant awarded to KPVS and KGMG for the PAPIIT project (IT20242). We acknowledge using Grammarly AI to improve the manuscript's English writing.

# Methodology

**Gene expression in *E. coli* and purification of ClosA, CellA, and C39 peptidase domain**

*E. coli* NiCO21 (DE3) cells were transformed with the following vectors: pRSF-Duet_C39 peptidase domain, pRSF-Duet_ClosA_ClosM, and pRSF-Duet_CellA_CellM. The transformed bacteria were cultivated in 1L LB medium containing 30 mg/L kanamycin at 23°C. Induction with 0.5 mM IPTG was performed when the $OD_{600nm}$ was between 0.6 and 0.8 Cultures were incubated for 14 h, at 200 rpm, at 18$^{o}$C. Subsequently, cells were centrifuged at 10000 rpm for 30 minutes at 4°C. The resulting cell pellets were resuspended in 20 mL of LanA starting buffer (20 mM Tris pH 7.5, 500 mM KCl, 10% glycerol) and lysed by sonication (4.0 seconds on, 10 seconds off, for a total of 10 minutes at 4.0 potency). Finally, the sample was centrifuged for 30 minutes at 4°C, at 10,000 rpm. Supernatants were stored at –80 C until purification.

The protein products ClosA, CellA, C39 peptidase domain were linked with a 6xHis N-Terminal tag to perform purifications using affinity chromatography with a 1 mL His Trap nickel affinity column. The supernatants were filtered (0.45 µm filter) and loaded into the column. The column was washed with 5 column volumes of LanA wash buffer 1 (20 mM Tris pH 7.5, 500 mM KCl, 10% glycerol, 0.5 mM imidazole), followed by 10 column volumes of LanA wash buffer 2 (20 mM Tris pH 7.5, 500 mM KCl, 10% glycerol, 30 mM imidazole). Finally, the elution was performed with 3 column volumes of LanA elution buffer (20 mM Tris pH 7.5, 500 mM KCl, 10% glycerol, 750 mM imidazole). Purifications were monitored through 20% sodium dodecyl sulfate-polyacrylamide gel electrophoresis (SDS-PAGE)[51].

Expression was confirmed through Western blot analysis using the 6xHis epitope tag on a lysate sample (Supplementary Figure 3.A) for pre-Clostrisin, pre-Cellulosin, and C39 peptidase domain. Affinity chromatography purification (Supplementary Figure 3.B) was performed for the C39 peptidase domain, resulting in a purity of approximately 96%, based on densitometric analysis of the Coomassie-stained gel. Protein yield was estimated using spectrophotometric quantification. Analysis using the NanoDrop device at 280 nm indicated a 1.25 mg/L yield. For pre-Clostrisin, the same analysis revealed a purity of around 41% and a 1 mg/L yield. In the case of pre-Cellulosin, the analysis showed a purity of 60% and a 0.8 mg/L yield.

**Protease activity of the peptidase C39 domain of ClosPt**

To optimize the heterologous expression process of protease ClosPt, we refrained from attempting to purify membrane proteins due to technical complexities. Instead, we exclusively focused on cloning the peptidase C39 domain into the pRSF-Duet_C39 peptidase domain vector. The sample containing the peptidase domain, C39 peptidase domain, was mixed with samples of ClosA and CellA at a 1:5 molar ratio within a reaction buffer composed of 50 mM HEPES at pH 7.0, 150 mM NaCl, and 5 mM DTT. Cofactors were added as indicated, with the following concentrations: 0.5 mM ATP and 1 mM MgCl2[17]. The reactions were incubated at 4 °C for 36 h. Subsequently, the samples were subjected to analysis using 20% SDS-PAGE.

**Antimicrobial activity of Clostrisin and Cellulosin**

Antimicrobial activity assays were conducted using the reference strains, *Acinetobacter baumannii* ATCC BAA 747, *Pseudomonas aeruginosa* PA14 ATCC 15442 and MIQPA25 (Clinical Strain),

*Escherichia coli* ATCC 10799, *Staphylococcus aureus* ATCC 43300 (MRSA) and *Staphylococcus epidermidis* MIQ43. Tests were performed in 96-well polystyrene microtiter plates with bacterial inoculation at an initial OD 600 nm of 0.1. The minimum inhibitory concentration (MIC) for each purified lanthipeptide was determined by testing three different concentrations in three separate wells and repeating the process in three independent experiments. For the highest tested concentration of lanthipeptide (5.6 µg/mL for pre-Clostrisin and 4.8 µg/mL for pre-Cellulosin) the well contained 178 µL of LB broth, 12 µL of purified lanthipeptide and 10 µL of bacterial culture. The plates were incubated at 37°C for 16 h, measuring the optical density (OD) at 600 nm, under shaking.

Three independent experiments were conducted, each consisting of three replicates per concentration. Data collected from each assay were deemed valid based on the indicated values of inhibition and growth observed in the controls, as per our standardized data. The results for MIC and percentage inhibition were averaged. The percentage of inhibition was derived by analyzing the 16 h growth curve data using the Standard Protocol established in our laboratory, comparing the growth curves obtained across the conducted experiments. All data processing and final figures were generated using Prism GraphPad (version 8.0.2).

**Atomic Force Microscopy (AFM)**

AFM characterization was conducted using a MultiMode 8-HR (Bruker). Samples containing *S. epidermis* at an $OD_{600}$>1 were incubated with the lanthipeptides at room temperature for 1 and 5 hours. Subsequently, 9 µL of the samples were combined with 1 µL of a 0.1% w/v poly-L-lysin solution in water (Sigma-Aldrich) and immediately deposited onto freshly cleaved mica to be incubated for 10 min at room temperature to allow adsorption. The surface was then rinsed using 600 µL of ultrapure 0.2 µm filtered water and slowly dried using compressed air. Imaging was performed using a Digital Instruments NanoScope V, acquiring 1024 samples per line with silicon nitride cantilevers possessing a nominal spring constant of 0.32 Nm−1 and a 0.8–1.0 Hz scan rate. Imaging was conducted at room temperature using the ScanAsyst™ air mode. Images were processed using NanoScope Analysis V.1.80.

**MALDI-TOF MS**

Matrix-assisted laser desorption– ionization time of flight mass spectrometry (MALDI-TOF MS) analyses were carried out at the Gas and Liquid Chromatography Laboratory UNAM Institute of Chemistry, on a FLEX-PC mass spectrometer (Bruker Microflex). For MALDI-TOF MS analysis, the samples used were pre-Clostrisin, pre-Cellulosin, Clostrisin, and Cellulosin, previously dialyzed to remove excess imidazole. Samples were spotted on a MALDI target plate with a matrix solution containing sinapinic acid (1:5), with a Linear Acquisition operation mode, a POS voltage polarity, and 659 shots.

# References:

(1) Murray, C. J.; Ikuta, K. S.; Sharara, F.; Swetschinski, L.; Robles Aguilar, G.; Gray, A.; Han, C.; Bisignano, C.; Rao, P.; Wool, E.; Johnson, S. C.; Browne, A. J.; Chipeta, M. G.; Fell, F.; Hackett, S.; Haines-Woodhouse, G.; Kashef Hamadani, B. H.; Kumaran, E. A. P.; McManigal, B.; Agarwal, R.; Akech, S.; Albertson, S.; Amuasi, J.; Andrews, J.; Aravkin, A.; Ashley, E.; Bailey, F.; Baker, S.; Basnyat, B.; Bekker, A.; Bender, R.; Bethou, A.; Bielicki, J.; Boonkasidecha, S.; Bukosia, J.; Carvalheiro, C.; Castañeda-Orjuela, C.; Chansamouth, V.; Chaurasia, S.; Chiurchiù, S.; Chowdhury, F.; Cook, A. J.; Cooper, B.; Cressey, T. R.; Criollo-Mora, E.; Cunningham, M.; Darboe, S.; Day, N. P.

J.; De Luca, M.; Dokova, K.; Dramowski, A.; Dunachie, S. J.; Eckmanns, T.; Eibach, D.; Emami, A.; Feasey, N.; Fisher-Pearson, N.; Forrest, K.; Garrett, D.; Gastmeier, P.; Giref, A. Z.; Greer, R. C.; Gupta, V.; Haller, S.; Haselbeck, A.; Hay, S. I.; Holm, M.; Hopkins, S.; Iregbu, K. C.; Jacobs, J.; Jarovsky, D.; Javanmardi, F.; Khorana, M.; Kissoon, N.; Kobeissi, E.; Kostyanev, T.; Krapp, F.; Krumkamp, R.; Kumar, A.; Kyu, H. H.; Lim, C.; Limmathurotsakul, D.; Loftus, M. J.; Lunn, M.; Ma, J.; Mturi, N.; Munera-Huertas, T.; Musicha, P.; Mussi-Pinhata, M. M.; Nakamura, T.; Nanavati, R.; Nangia, S.; Newton, P.; Ngoun, C.; Novotney, A.; Nwakanma, D.; Obiero, C. W.; Olivas-Martinez, A.; Olliaro, P.; Ooko, E.; Ortiz-Brizuela, E.; Peleg, A. Y.; Perrone, C.; Plakkal, N.; Ponce-de-Leon, A.; Raad, M.; Ramdin, T.; Riddell, A.; Roberts, T.; Robotham, J. V.; Roca, A.; Rudd, K. E.; Russell, N.; Schnall, J.; Scott, J. A. G.; Shivamallappa, M.; Sifuentes-Osornio, J.; Steenkeste, N.; Stewardson, A. J.; Stoeva, T.; Tasak, N.; Thaiprakong, A.; Thwaites, G.; Turner, C.; Turner, P.; van Doorn, H. R.; Velaphi, S.; Vongpradith, A.; Vu, H.; Walsh, T.; Waner, S.; Wangrangsimakul, T.; Wozniak, T.; Zheng, P.; Sartorius, B.; Lopez, A. D.; Stergachis, A.; Moore, C.; Dolecek, C.; Naghavi, M. *The Lancet* **2022**, *399*, 629–655. doi:10.1016/S0140-6736(21)02724-0.

(2) "Global Action Plan on Antimicrobial Resistance**",** World Health Organization. **2015.**

(3) Blin, K.; Shaw, S.; Kloosterman, A. M.; Charlop-Powers, Z.; Van Wezel, G. P.; Medema, M. H.; Weber, T. *Nucleic Acids Res* **2021**, *49*, W29–W35. doi:10.1093/NAR/GKAB335.

(4) Hart, T.; Moffat, J. *BMC Bioinformatics* **2016**, *17*, 1–7. doi:10.1186/S12859-016-1015-8.

(5) Agrawal, P.; Khater, S.; Gupta, M.; Sain, N.; Mohanty, D. *Nucleic Acids Res* **2017**, *45*, W80–W88. doi:10.1093/NAR/GKX408.

(6) Choirunnisa, A. R.; Arima, K.; Abe, Y.; Kagaya, N.; Kudo, K.; Suenaga, H.; Hashimoto, J.; Fujie, M.; Satoh, N.; Shin-Ya, K.; Matsuda, K.; Wakimoto, T. *Beilstein Journal of Organic Chemistry* **2022**, *18*, 1017–1025. doi:10.3762/bjoc.18.102.

(7) Biermann, F.; Wenski, S. L.; Helfrich, E. J. N. *Beilstein Journal of Organic Chemistry* **2022**, *18*, 1656–1671. doi:10.3762/bjoc.18.178.

(8) Rivera-Chávez, J.; Ceapă, C. D.; Figueroa, M. *Planta Med* **2022**, *88*, 702–720. doi:10.1055/a-1795-0562.

(9) Van Kraaij, C.; De Vos, W. M.; Siezen, R. J.; Kuipers, O. P. *Nat Prod Rep* **1999**, *16*, 575–587. doi:10.1039/a804531c.

(10) Arnison, P. G.; Bibb, M. J.; Bierbaum, G.; Bowers, A. A.; Bugni, T. S.; Bulaj, G.; Camarero, J. A.; Campopiano, D. J.; Challis, G. L.; Clardy, J.; Cotter, P. D.; Craik, D. J.; Dawson, M.; Dittmann, E.; Donadio, S.; Dorrestein, P. C.; Entian, K. D.; Fischbach, M. A.; Garavelli, J. S.; Göransson, U.; Gruber, C. W.; Haft, D. H.; Hemscheidt, T. K.; Hertweck, C.; Hill, C.; Horswill, A. R.; Jaspars, M.; Kelly, W. L.; Klinman, J. P.; Kuipers, O. P.; Link, A. J.; Liu, W.; Marahiel, M. A.; Mitchell, D. A.; Moll, G. N.; Moore, B. S.; Müller, R.; Nair, S. K.; Nes, I. F.; Norris, G. E.; Olivera, B. M.; Onaka, H.; Patchett, M. L.; Piel, J.; Reaney, M. J. T.; Rebuffat, S.; Ross, R. P.; Sahl, H. G.; Schmidt, E. W.; Selsted, M. E.; Severinov, K.; Shen, B.; Sivonen, K.; Smith, L.; Stein, T.; Süssmuth, R. D.; Tagg, J. R.; Tang, G. L.; Truman, A. W.; Vederas, J. C.; Walsh, C. T.; Walton, J. D.; Wenzel, S. C.; Willey, J. M.; Van Der Donk, W. A. *Nat Prod Rep* **2013**, *30*, 108–160. doi:10.1039/c2np20085f.

(11) Guder, A.; Wiedemann, I.; Sahl, H. G. *Biopolymers - Peptide Science Section* **2000**, *55*, 62–73. doi:10.1002/1097-0282(2000)55:1.

(12) Acedo, J. Z.; Bothwell, I. R.; An, L.; Trouth, A.; Frazier, C.; Van Der Donk, W. A. *J Am Chem Soc* **2019**, *141*, 16790–16801. doi:10.1021/jacs.9b07396.

(13) Yang, X.; Van Der Donk, W. A. *Chemistry* **2013**, *19*, 7662–7677. doi:10.1002/CHEM.201300401.

(14) Willey, J. M.; Van Der Donk, W. A. *Annu Rev Microbiol* **2007**, *61*, 477–501. doi:10.1146/annurev.micro.61.080706.093501.

(15) Deisinger, J. P.; Arts, M.; Kotsogianni, I.; Puls, J. S.; Grein, F.; Ortiz-López, F. J.; Martin, N. I.; Müller, A.; Genilloud, O.; Schneider, T. *iScience* **2023**, *26*. doi:10.1016/j.isci.2023.106394.

(16) Repka, L. M.; Chekan, J. R.; Nair, S. K.; Van Der Donk, W. A. *Chem Rev* **2017**, *117*, 5457–5520. doi:10.1021/acs.chemrev.6B00591.
(17) Lin, D. Y. W. Huang, S.; Chen, J. *Nature 2015 523:7561* **2015**, *523*, 425–430. doi:10.1038/nature14623.
(18) Bauer, R.; Dicks, L. M. T. *Int J Food Microbiol* **2005**, *101*, 201–216. doi:10.1016/j.ijfoodmicro.2004.11.007.
(19) Karczewski, J.; Krasucki, S. P.; Asare-Okai, P. N.; Diehl, C.; Friedman, A.; Brown, C. M.; Maezato, Y.; Streatfield, S. J. *Front Microbiol* **2020**, *11*, 1–15. doi:10.3389/fmicb.2020.598789.
(20) Yates, K. R.; Welsh, J.; Udegbunam, N. O.; Greenman, J.; Maraveyas, A.; Madden, L. A. *Blood Coagulation and Fibrinolysis* **2012**, *23*, 396–401. doi:10.1097/mbc.0B013E3283538875.
(21) van Staden, A. D. P.; van Zyl, W. F.; Trindade, M.; Dicks, L. M. T.; Smith, C. *Appl Environ Microbiol* **2021**, *87*. doi:10.1128/AEM.00186-21.
(22) Zhou, H.; Fang, J.; Tian, Y.; Lu, X. Y. *Ann Microbiol* **2014**, *64*, 413–420. doi:10.1007/s13213-013-0679-9.
(23) Field, D.; Blake, T.; Mathur, H.; O' Connor, P. M.; Cotter, P. D.; Paul Ross, R.; Hill, C. *Mol Microbiol* **2019**, *111*, 717–731. doi:10.1111/mmi.14183.
(24) Kuthning, A.; Mösker, E.; Süssmuth, R. D. *Appl Microbiol Biotechnol* **2015**, *99*, 6351–6361. doi:10.1007/S00253-015-6557-6.
(25) Baltz, R. H. *J Ind Microbiol Biotechnol* **2019**, *46*, 281–299. doi:10.1007/s10295-018-2115-4.
(26) Xue D, Older EA, Zhong Z, Shang Z, Chen N, Dittenhauser N, Hou L, Cai P, Walla MD, Dong SH, Tang X, Chen H, Nagarkatti P, Nagarkatti M, Li YX, Li J. Correlational networking guides the discovery of unclustered lanthipeptide protease-encoding genes. Nat Commun. 2022 Mar 28;13(1):1647. doi: 10.1038/s41467-022-29325-1.
(27) Baltz, R. H. *J Ind Microbiol Biotechnol* **2021**, *48*. doi:10.1093/jimb/kuab044.
(28) Ling, L. L.; Schneider, T.; Peoples, A. J.; Spoering, A. L.; Engels, I.; Conlon, B. P.; Mueller, A.; Schäberle, T. F.; Hughes, D. E.; Epstein, S.; Jones, M.; Lazarides, L.; Steadman, V. A.; Cohen, D. R.; Felix, C. R.; Fetterman, K. A.; Millett, W. P.; Nitti, A. G.; Zullo, A. M.; Chen, C.; Lewis, K. *Nature* **2015**, *517*, 455–459. doi:10.1038/nature14098.
(29) Stokes, J. M.; Yang, K.; Swanson, K.; Jin, W.; Cubillos-Ruiz, A.; Donghia, N. M.; MacNair, C. R.; French, S.; Carfrae, L. A.; Bloom-Ackerman, Z.; Tran, V. M.; Chiappino-Pepe, A.; Badran, A. H.; Andrews, I. W.; Chory, E. J.; Church, G. M.; Brown, E. D.; Jaakkola, T. S.; Barzilay, R.; Collins, J. J. *Cell* **2020**, *180*, 688-702.e13. doi:10.1016/j.cell.2020.01.021.
(30) Zhong, Z.; He, B.; Li, J.; Li, Y. X. *Synth Syst Biotechnol* **2020**, *5*, 155–172. doi:10.1016/j.synbio.2020.06.002.
(31) Cao, L.; Do, T.; James Link, A. *J Ind Microbiol Biotechnol* **2021**, *48*. doi:10.1093/jimb/kuab005.
(32) Kautsar, S. A.; Blin, K.; Shaw, S.; Navarro-Muñoz, J. C.; Terlouw, B. R.; van der Hooft, J. J. J.; van Santen, J. A.; Tracanna, V.; Suarez Duran, H. G.; Pascal Andreu, V.; Selem-Mojica, N.; Alanjary, M.; Robinson, S. L.; Lund, G.; Epstein, S. C.; Sisto, A. C.; Charkoudian, L. K.; Collemare, J.; Linington, R. G.; Weber, T.; Medema, M. H. *Nucleic Acids Res* **2020**, *48*, D454–D458. doi:10.1093/nar/gkz882.
(33) Boratyn, G. M.; Thierry-Mieg, J.; Thierry-Mieg, D.; Busby, B.; Madden, T. L. *BMC Bioinformatics 2019 20:1* **2019**, *20*, 1–19. doi:10.1186/S12859-019-2996-X.
(34) Olson, R. D.; Assaf, R.; Brettin, T.; Conrad, N.; Cucinell, C.; Davis, J. J.; Dempsey, D. M.; Dickerman, A.; Dietrich, E. M.; Kenyon, R. W.; Kuscuoglu, M.; Lefkowitz, E. J.; Lu, J.; Machi, D.; Macken, C.; Mao, C.; Niewiadomska, A.; Nguyen, M.; Olsen, G. J.; Overbeek, J. C.; Parrello, B.; Parrello, V.; Porter, J. S.; Pusch, G. D.; Shukla, M.; Singh, I.; Stewart, L.; Tan, G.; Thomas, C.; VanOeffelen, M.; Vonstein, V.; Wallace, Z. S.; Warren, A. S.; Wattam, A. R.; Xia, F.; Yoo, H.; Zhang,

Y.; Zmasek, C. M.; Scheuermann, R. H.; Stevens, R. L. *Nucleic Acids Res* **2023**, *51*, D678–D689. doi:10.1093/nar/gkac1003.
(35) Saitou, N.; Nei, M. *Mol Biol Evol* **1987**, *4*, 406–425. doi:10.1093/oxfordjournals.molbev.A040454.
(36) Ren, H.; Shi, C.; Bothwell, I. R.; Van Der Donk, W. A.; Van Der Donk, W. A.; Van Der Donk, W. A.; Zhao, H.; Zhao, H.; Zhao, H.; Zhao, H.; Zhao, H. *ACS Chem Biol* **2020**, *15*, 1642–1649. doi:10.1021/acschembio.0c00267.
(37) Tamaru, Y.; Miyake, H.; Kuroda, K.; Nakanishi, A.; Kawade, Y.; Yamamoto, K.; Uemura, M.; Fujita, Y.; Doi, R. H.; Ueda, M. *J Bacteriol* **2010**, *192*, 901–902. doi:10.1128/JB.01450-09
(38) Zhao, X.; Van Der Donk, W. A. *Cell Chem Biol* **2016**, *23*, 246–256. doi:10.1016/j.chembiol.2015.11.014.
(39) Sawa, N.; Wilaipun, P.; Kinoshita, S.; Zendo, T.; Leelawatcharamas, V.; Nakayama, J.; Sonomoto, K. *Appl Environ Microbiol* **2012**, *78*, 900–903. doi:10.1128/aem.06497-11.
(40) Sand, S. L.; Nissen-Meyer, J.; Sand, O.; Haug, T. M. *Biochim Biophys Acta Biomembr* **2013**, *1828*, 249–259. doi:10.1016/j.bbamem.2012.11.001.
(41) Xin, B.; Zheng, J.; Liu, H.; Li, J.; Ruan, L.; Peng, D.; Sajid, M.; Sun, M. *Front Microbiol* **2016**, *7*, 1–12. doi:10.3389/fmicb.2016.01115.
(42) Begley, M.; Cotter, P. D.; Hill, C.; Ross, R. P. *Appl Environ Microbiol* **2009**, *75*, 5451–5460. doi:10.1128/aem.00730-09.
(43) Hyink, O.; Balakrishnan, M.; Tagg, J. R. *FEMS Microbiol Lett* **2005**, *252*, 235–241. doi:10.1016/j.femsle.2005.09.003.
(44) Yonezawa, H.; Kuramitsu, H. K. *Antimicrob Agents Chemother* **2005**, *49*, 541–548. doi:10.1128/aac.49.2.541-548.2005.
(45) Zallot, R.; Oberg, N.; Gerlt, J. A. The EFI Web Resource for Genomic Enzymology Tools: Leveraging Protein, Genome, and Metagenome Databases to Discover Novel Enzymes and Metabolic Pathways. *Biochemistry*. 2019, pp 4169–4182. doi:10.1021/acs.biochem.9b00735.
(46) Jumper, J.; Evans, R.; Pritzel, A.; Green, T.; Figurnov, M.; Ronneberger, O.; Tunyasuvunakool, K.; Bates, R.; Žídek, A.; Potapenko, A.; Bridgland, A.; Meyer, C.; Kohl, S. A. A.; Ballard, A. J.; Cowie, A.; Romera-Paredes, B.; Nikolov, S.; Jain, R.; Adler, J.; Back, T.; Petersen, S.; Reiman, D.; Clancy, E.; Zielinski, M.; Steinegger, M.; Pacholska, M.; Berghammer, T.; Bodenstein, S.; Silver, D.; Vinyals, O.; Senior, A. W.; Kavukcuoglu, K.; Kohli, P.; Hassabis, D. *Nature* **2021**, *596*, 583–589. doi:10.1038/s41586-021-03819-2.
(47) Mirdita, M.; Schütze, K.; Moriwaki, Y.; Heo, L.; Ovchinnikov, S.; Steinegger, M. *Nat Methods* **2022**, *19*, 679–682. doi:10.1038/s41592-022-01488-1.
(48) Dong, S. H.; Tang, W.; Lukk, T.; Yu, Y.; Nair, S. K.; van der donk, W. A. *Elife* **2015**, *4*. doi:10.7554/eLife.07607.
(49) Bobeica, S. C.; Dong, S. H.; Huo, L.; Mazo, N.; McLaughlin, M. I.; Jiménez-Osés, G.; Nair, S. K.; van der Donk, W. A. *Elife* **2019**, *8*, 1–27. doi:10.7554/eLife.42305
(50) Gutiérrez-Santana, J. C.; Gerónimo-Gallegos, A.; Martínez-Corona, M. B.; López-López, M.; Toscano-Garibay, J. D.; Cuevas-Schacht, F.; Coria-Jiménez, V. R. *Curr Microbiol* **2022**, *79*, 1–10. doi:10.1007/s00284-022-03048-4
(51) Shi, Y.; Yang, X.; Garg, N.; Van Der Donk, W. A. *J Am Chem Soc* **2011**, *133*, 2338–2341. doi:10.1021/JA109044R.