EFFICIENT *IN VIVO* SYNTHESIS OF LASSO PEPTIDE PSEUDOMYCOIDIN PROCEEDS IN THE ABSENCE OF LEADER AND LEADER PEPTIDASE

Doctoral Thesis

by

Tatyana Zyubko

Skoltech

Skolkovo Institute of Science and Technology

EFFICIENT *IN VIVO* SYNTHESIS OF LASSO PEPTIDE PSEUDOMYCOIDIN PROCEEDS IN THE ABSENCE OF LEADER AND LEADER PEPTIDASE

Doctoral Thesis

by

Tatyana Zyubko

DOCTORAL PROGRAM IN LIFE SCIENCES

Supervisor

Professor Konstantin Severinov

Moscow - 2019

Abstract

The continuous spread of pathogenic bacteria resistant to antibiotics is one of the most important threats facing medicine today. One of the possible solutions to antibiotic crisis is intensified search for new antibacterial agents. Nature remains the major source both of new active compounds and of molecular scaffolds to develop new synthetic molecules.

Ribosomally synthesized and post-translationally modified peptides (RiPPs) are a major class of natural products. These peptides possess a high degree of structural diversity and a wide range of bioactivities. One intriguing family of RiPPs are the Lasso Peptides. These compact peptides (up to 40 residues) demonstrate an unusual topology resembling a lasso and typically consist of a 7–9-residue N-terminal macrolactam ring pierced by a trapped C-terminal tail.

This knot-like fold of lasso peptides is responsible for high compactness, extraordinary stability, and diverse bioactivities, such as antimicrobial, antiviral, receptor antagonistic or enzyme inhibitory activities. Moreover, this topology represents a promising scaffold for peptide epitope grafting.

Unfortunately, synthesizing these peptides through a set of chemical reactions is ineffective so far, and the biotechnological approach is the only way to obtain them nowadays. Biosynthesis of lasso peptide starts with the ribosomal translation of a precursor peptide consisting of a cleavable leader part and a modifiable core that subsequently undergoes post-translational modifications. The lasso peptide maturation is catalyzed by two processing enzymes: the leader peptidase and a lasso cyclase. This process is directed by RiPP Recognition Element (RRE), which recognizes the leader part of the precursor peptide and recruits the maturation enzymes to the substrate. However, the structure of the enzymatic maturation complex is not established.

We have bioinformatically identified a lasso peptide biosynthetic gene cluster in the genome of *Bacillus pseudomycoides* DSM 12442. Its heterologous expression in *Escherichia coli* results in the production of lasso peptide called pseudomycoidin. Uniquely, the pseudomycoidin cluster

deleted for both genes coding for essential processing proteins, leader peptidase and RRE, still was able to produce pseudomycoidin. Furthermore, the subsequent exclusion of the leader part from the precursor peptide did not lead to a decrease in pseudomycoidin yield. Thus, the lasso-cyclase from *B. pseudomycoides* can function in the absence of other processing enzymes and the leader sequence. Hence, the minimized pseudomycoidin synthesis system is a promising candidate for the development of a universal tool for large-scale *in vivo* synthesis of various pseudomycoidin derivatives and, possibly, other lasso-compounds.

Publications

The results of the projects were presented at the following scientific conferences and published in international peer-reviewed journals:

1. Efficient *in vivo* Synthesis of Lasso Peptide Pseudomycoidin Proceeds in the Absence of Leader and Leader Peptidase. Zyubko, Tatyana; Serebryakova, Marina; Andreeva, Julia; Metelev, Mikhail; Lippens, Guy; Dubiley, Svetlana; Severinov, Konstantin, Chem. Sci. 2019 Aug 30; DOI:10.1039/C9SC02370D.

2. Reiterative Synthesis by the Ribosome and Recognition of the N-Terminal Formyl Group by Biosynthetic Machinery Contribute to Evolutionary Conservation of the Length of Antibiotic Microcin C Peptide Precursor. Zukher I, Pavlov M, Tsibulskaya D, Kulikovsky A, Zyubko T, Bikmetov D, Serebryakova M, Nair SK, Ehrenberg M, Dubiley S, Severinov K., MBio. 2019 Apr 30; 10(2): e00768-19. DOI: 10.1128/mBio.00768-19.

Conferences:

1. 35th European Peptide Symposium, “The novel lasso peptide pseudomycoidin can be matured *in vivo* with only one specific processing enzyme”, Tatyana Zyubko, Yulia Piskunova, Svetlana Dubiley, Marina Serebryakova, Leah B. Bushin, Tatyana Artamonova, Mikhail Metelev, Mohammad R. Seyedsayamdost, and Konstantin Severinov, 26th – 31st August, 2018, Dublin City University Dublin, Ireland.

2. **Second prize** on 8th International Meeting on Antimicrobial Peptides, “The novel lasso peptide pseudomycoidin from Bacillus sp. synthesized in branched-cyclic form”, Tatyana Zyubko, Yulia Piskunova, Svetlana Dubiley, Marina Serebryakova, Leah B. Bushin, Tatyana Artamonova, Mikhail Metelev, Mohammad R. Seyedsayamdost, and Konstantin Severinov, 2nd – 4th September, 2018, Royal College of Surgeons of Edinburgh, Edinburgh, EH8 9DW, UK

Acknowledgements

The author would like to express her thanks to Konstantin Severinov, Svetlana Dubiley, Mikhail Metelev, Marina Serebryakova, Tatyana Artamonova, Guy Lippens and special thanks to Anastasia Sharapkova and Mikhail Khodorkovskii.

Table of contents

Abbreviations

ABC	–	ATP-binding cassette transporter
ADP	–	adenosine diphosphate
AMP	–	antimicrobial peptides
AS-B	–	asparagine synthetase B
ATP	–	adenosine triphosphate
BGC	–	biosynthetic gene clusters
BLAST	–	basic local alignment search tool
HPLC	–	high-performance liquid chromatography
IPTG	–	isopropyl-β-D-thiogalactopyranoside
IsoP	–	isopeptidase
LAP	–	linear azol(in)e-containing peptide
MALDI	–	matrix-assisted laser desorption/ionization
MBP	–	maltose binding protein
MccB17	–	microcin B17
McC	–	microcin C
MccJ25	–	microcin J25
MS	–	mass spectrometry
NMR	–	nuclear magnetic resonance
ORF	–	open reading frame
PCR	–	polymerase chain reaction
PSI-BLAST	–	position-specific iterated BLAST
RiPP	–	ribosomally synthesized and post-translationally modified peptides
RNAP	–	ribonucleic acid polymerase
RP-HPLC	–	reversed-phase high-performance liquid chromatography
RRE	–	RiPP precursor peptide recognition element
TFA	–	trifluoroacetic acid
tRNA	–	transfer ribonucleic acid
UV	–	ultra violet

1. Literature Review

1.1. RiPPs, the major class of Natural Products

Ribosomally synthesized and post-translationally modified peptides (RiPPs) make up a broad class of gene-encoded natural products. These peptides demonstrate a high degree of structural diversity as well as a wide range of bioactivities, including antimicrobial [1]–[9].

RiPP biosynthetic gene clusters (BGCs) exhibit similar organization and code for short precursor peptides as well as varying numbers of maturation enzymes involved in post-translational processing of precursor [6], [10]. Additionally, the clusters often encode an export system for the mature product and in some cases of self-toxic antimicrobial peptides (AMP), a self-immunity system [2], [4], [6], [11].

Typically, the export system consists of the ABC-exporters belonging to one of the two major types: transporters that exclusively export AMPs or bifunctional transporters that ensure both the maturation of AMP by cleaving the leader region of precursor and secreting the mature peptide (Peptidase-Containing ATP-binding Transporters) [12].

Thus, the structure of RiPP biosynthetic gene clusters can be schematically summarized as represented in Figure 1.

RiPP biosynthesis relies on the ribosomal translation of the precursor peptide that subsequently undergoes enzymatic maturation. Typically, a precursor peptide consists of a C-terminal modifiable core part and a cleavable leader sequence located at the N-terminus. Far less frequently (i.e., in bottromycins) the follower sequence is present at the C-terminus of the precursor peptide [13], [14]. The leader peptide guides biosynthesis by interacting with specific processing enzymes [15]–[19], while the core part is subjected to various enzymatic post-translational modifications [2]–[4], [6], [15], [20].

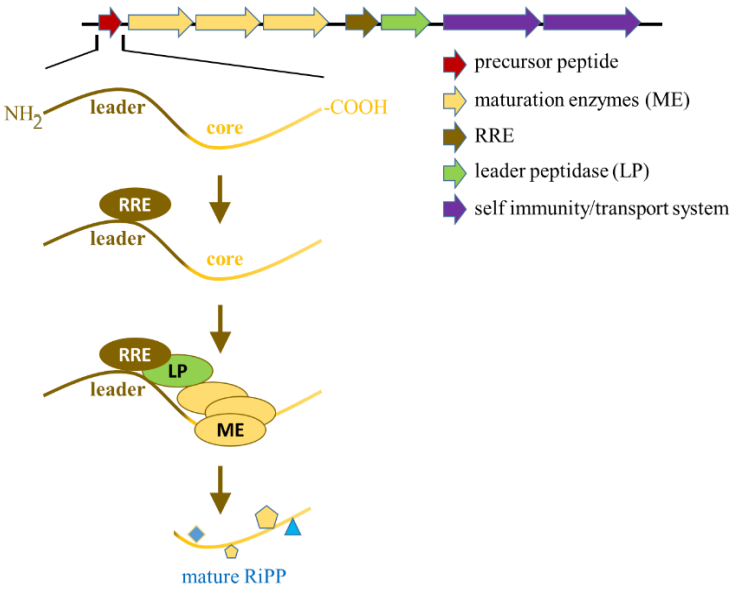


***Figure 1*. The genetic organization of a typical RiPP biosynthetic gene cluster**. Genes are marked with arrows and coloured as indicated. A typical RiPP biosynthetic gene cluster codes for a precursor peptide containing a cleavable leader part as well as a modifiable core part; a recognition element which binds the leader peptide and guides subsequent modifications; a leader peptidase; and a varying number of maturation enzymes and transporters or/and a self-immunity system. A general scheme of RiPP biosynthetic pathway is also presented. The translated precursor peptide and processing enzymes are shown as a line and circles, respectively.

Elaborate post-translational modifications give rise to structurally different peptides demonstrating a wide range of biological activities. Many RiPPs exhibit increased stability compared to unmodified peptides [15]. For instance, modifications of the N- and C-termini restrict degradation by exoproteases, while macrocyclization tends to improve metabolic stability and reduce structural flexibility.

RiPPs can be classified into several distinct subclasses based on structural commonalities. The major subfamilies of RiPPs include lantipeptides, thiopeptides, cyanobactins, linear azol(in)e-containing peptides (LAPs), and lasso peptides (Figures 2 and 3).

The first members of three of these subfamilies – microcin C, lasso peptide microcin J25, and LAP microcin B17 (Figure 3) were isolated from *Enterobacteriaceae* [21] and assigned to the same class of antimicrobial compounds – microcins, although their structures are unrelated and modes of antibacterial action differ dramatically (Figure 4) [6], [15], [22].

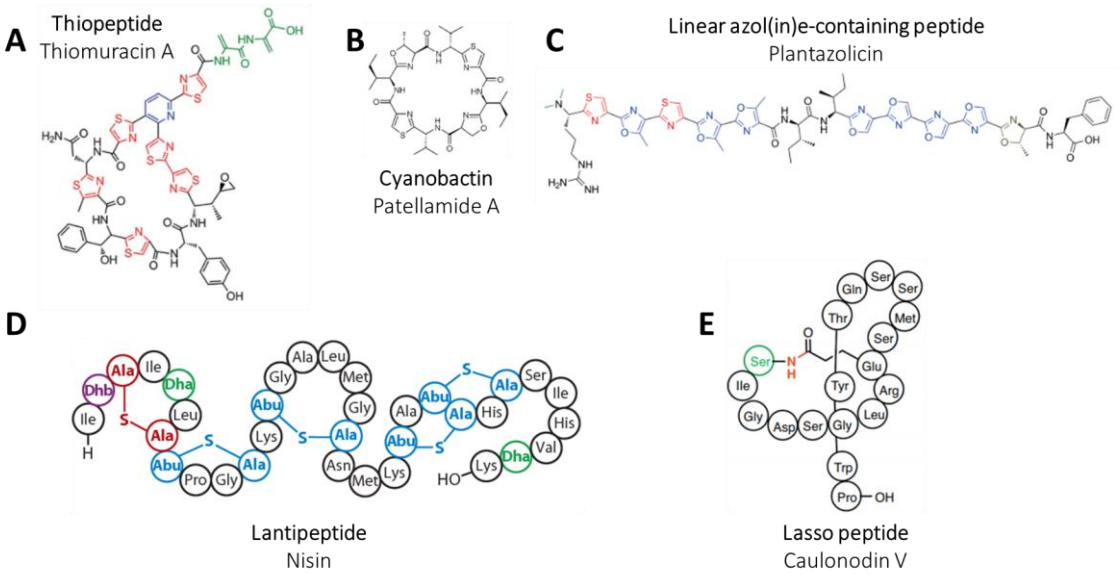


***Figure 2*. Some representatives of RiPPs**. (A) Thiopeptides: Thiomuracin A. The central dehydropiperidine or pyridine rings are in blue, the thiazole and thiazoline rings are in red, and the dehydroalanine and dehydrobutyrine residues are in green [6]. (B) Cyanobactins: Patellamide A [23]. (C) Linear azol(in)e-containing peptides: Plantazolicin. Cys resulting in thiazoles are in red; Ser/Thr resulting in oxazoles and methyloxazoles are in blue; Thr leading to methyloxazoline are in brown; Ser converted to dehydroalanine are in light blue; N-terminal acetylated/ methylated residues are in green [6]. (D) Lantipeptides: Nisin. Highlighting the canonical nisin- and mersacidin-lipid II–binding motifs (blue and red dashed circles, respectively) and tailoring post-translational modifications (brown) [24]. (E) lasso peptides: Caulonodin V. Bond formed during macrocyclization is in red [25]

Microcin B17 (MccB17) is produced by various *E. coli* strains [26] and inhibits DNA gyrase [27]–[30]. Mature MccB17 is a 43-amino-acid peptide containing four thiazole and four oxazole heterocycles [31]; it is an archetypical example of Thiazole-Oxazole Modified Microcins, a subclass of the Linear azol(in)e-containing peptides [6]. MccB17 is synthesized from ribosomally translated 69-amino-acid precursor peptide McbA by the McbBCD enzymatic maturation complex [20], [32].

The entire BGC of MccB17 consists of seven genes *mcbABCDEFG* [26], encoding the precursor peptide (McbA), three processing enzymes (McbBCD) [20], self-immunity protein (McbG) [33], and a dedicated two-component ABC-transporter (McbEF) ensuring active export of MccB17 out of producing cells [26], [33]. McbG protects DNA gyrase from the antibiotic action. Thus, *mcbABCD* subcluster is responsible for the biosynthesis of MccB17, while *mcbEFG* confers self-immunity to the producer as well as export of mature MccB17 out of cells [33], [34].

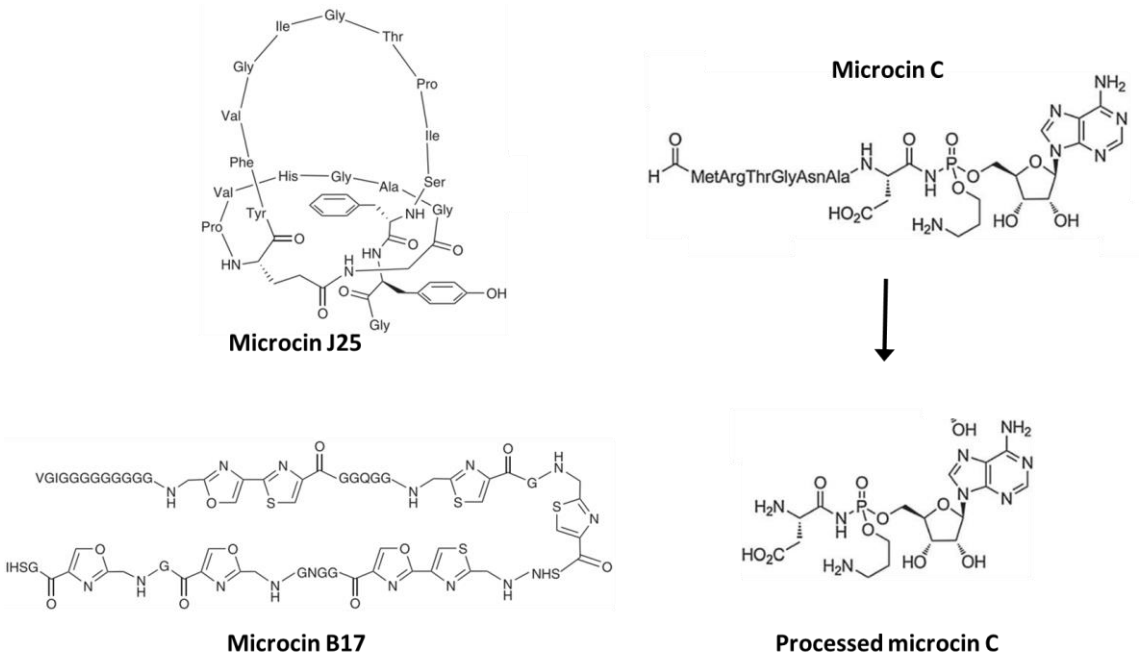


***Figure 3*. Chemical structures of (A) microcin J25; (B) microcin B17 and (C) microcin C** [6], [15]**.**

Microcin C (microcin C7/C51, McC) is a peptide nucleotide antibiotic that inhibits translation in sensitive cells by interacting with aspartyl-tRNA synthetase [35], [36]. The ribosomally synthesized MccA precursor peptide is matured by the MccB, MccD, and MccE proteins [36]. This maturation results in a N-formylated heptapeptide (fMRTGNAD) with a modified adenosine monophosphate attached to the γ-carboxyl group of the C-terminal aspartate (Figure 3) [37], [38]. The peptide part is processed into toxic payload in target cells.

McC is a Trojan-Horse antibiotic [36], [39]: it is converted into toxic form after penetrating the target cells. The peptide moiety ensures penetration into the target cells through the YejABEF ABC-transporter [40] and is subsequently degraded. The remaining part of McC, nonhydrolyzable aspartyl-adenylate [40], [41], inhibits translation. This type of processing can be regarded as an after-maturation leader-peptide cleavage, that takes place outside the producing cells as opposed to the situation with other RiPPs, which are fully matured inside the producers

[6]. The protection of producing cells against McC is ensured by a joint action of three proteins – MccE and MccF detoxifying enzymes and the MccC efflux pump [11], [42].

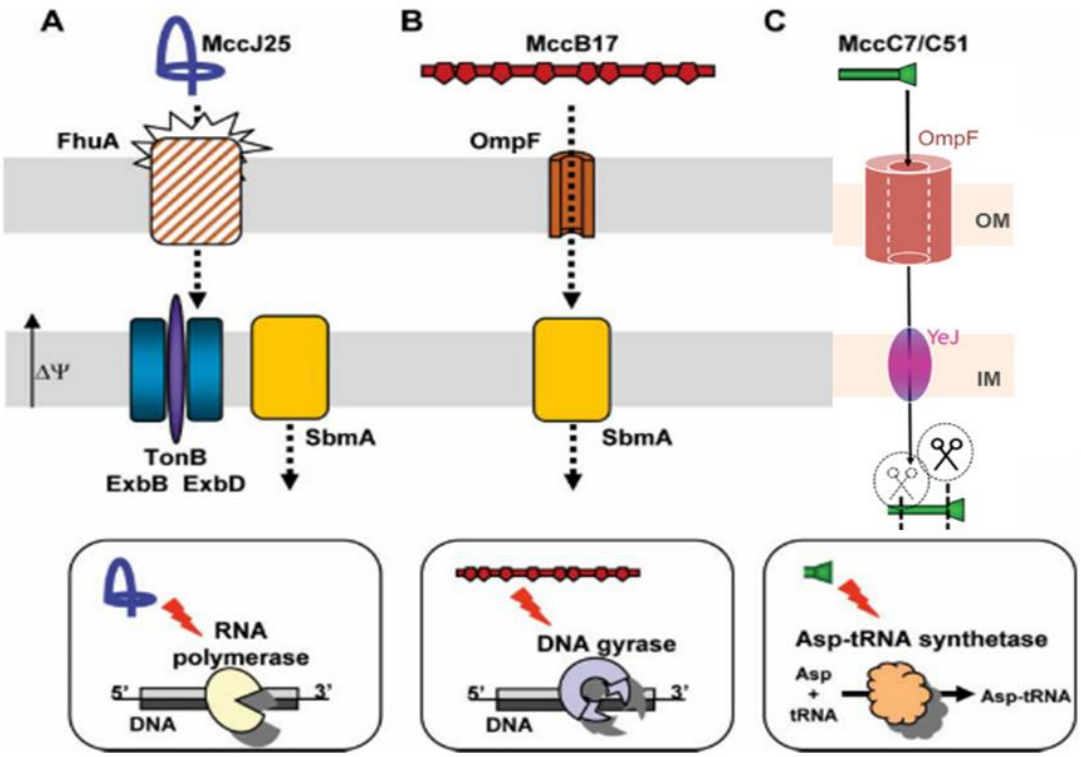


***Figure 4*. Schematics of the uptake and antibacterial action of lasso peptide microcin J25, thiazole/oxazole modified microcin B17, and nucleotide-peptide microcin C**. (A) MccJ25 penetrates into target cells through high affinity receptor FhuA in the outer membrane (OM) and TonB/ExbB/ExbD complex as well as the SbmA protein in the inner membrane (IM) and clogs the RNA polymerase secondary channel. (B) MccB17 enters through the OM via the OmpF porin and through IM via SbmA. Inside the target cells, MccB17 arrests DNA replication by binding the DNA gyrase and fixing it on DNA. (C) The uptake of full-sized (inactive) McC is mediated by OmpF and the YejABEF ABC-transporter in the outer and inner membranes, respectively. To become an active compound, McC undergoes proteolytic cleavage inside target cells. The resulting modified aspartyl-adenylate inhibits translation by targeting the aspartyl-tRNA synthetase. *Adapted from [2], [7]*.

Microcin J25 (MccJ25) is a 21 aminoacid peptide with an unusual structure resembling a lasso; it inhibits bacterial DNA-dependent RNA-polymerase (RNAP). To date, MccJ25 is the most studied representative of the Lasso peptide subfamily of RiPPs. It consists of an 8-residue N-terminal macrolactam ring through which the C-terminal tail (13-residue) is threaded [43]. The tail is immobilized by the two aromatic bulky side chains Phe19 and Tyr20 located, respectively, above and below the plane of the ring [44] (Figure 3).

MccJ25 precursor peptide is translated by a ribosome and further processed with two specific maturation enzymes – McjB and McjC [45]. McjB is a double-function protein that recognizes the precursor peptide and cleaves its leader part [7], [46], while McjC installs, enzyme at a cost of one ATP, an isopeptide bond between the N-terminal amino group of Gly1 formed after cleavage and the side-chain carboxylate of Glu8 forming the ring [7], [43]. Mature MccJ25 is exported out of producing cells by the ABC transporter McjD.

MccJ25 reaches the target Gram-negative bacteria cytoplasm through the FhuA receptor and TonB and SbmA transporters located in the outer and inner membranes, respectively. In sensitive cells MccJ25 inhibits DNA transcription by blocking the secondary channel of RNAP [43], [47].

The growing number of the sequenced genomes of microorganisms accompanied by recent advances in bioinformatics [48]–[50] reveals that microcins C, B17 and J25 are not unique to *Enterobacteriaceae* but each are representatives of larger families of structurally related peptides, whose BGCs were discovered in the genomes of various microorganisms [49], [48].

1.2. Lasso peptides: structure, classification

Lasso peptides are naturally produced small molecules with an unusual topology resembling a lasso. Biosynthesis clusters for these peptides are widespread throughout bacterial phyla [49] and are also found in archaea. To date, more than 3000 lasso peptide BGCs have been predicted in sequenced prokaryotic genomes [49], [51], [52]. About 70 various lasso peptides have been isolated so far (Table 1), mainly from *Actinobacteria* and *Proteobacteria*. Only one lasso peptide, paeninodin, was isolated from *Firmicutes* (Table 1) [49], [53]–[56].

Lasso peptides contain a 7–9-residue N-terminal macrolactam ring followed by a C-terminal peptide tail, which is normally threaded through the ring forming a “lasso”- like topology [44] (Figure 5). The lasso peptide tails are prevented from escaping with a steric lock made either of disulfide bonds formed by side chains of cysteines or by the bulky side chains located on both

sides of the ring, thus forming a molecular rotaxane [57], [58]. Thus, the threaded lasso peptides have three distinct regions: the ring, the loop above it and the threaded tail below it. Such molecules possess an inherent chirality with almost the same free energy for both right- and left-handed topoisomers [59], [60], but only the right-handed lasso peptides were observed in nature. The unthreading of the tail leads to a "branched-cyclic" lasso isomer (Figure 5).

The distinctive lariat knot structure contributes to lasso peptides' high compactness and typically extraordinary stability against protease degradation and denaturation [2], [6], [61]–[65]. In some cases lasso peptides also possess high thermal stability [51], [63], [64], [66]–[69]. In particular, microcin J25 remains active after autoclaving [66], xanthomonins maintain the lasso structures for up to 8 h at 95°C in water [63] and caulosegnin I – for up to 4 h at 95°C [64]. In contrast, a number of lasso peptides undergo a conversion into a branched cyclic form upon heating [62], [67], [69]–[71] or in the presence of organic solvents [72]: caulosegnin I, caulosegnin III.

The heat stability of lasso peptides depends on many factors such as the size of the ring, the plug residues and the amino acid sequence forming the macrolactam ring [69]. In fact, the xanthomonins contain the smallest so far known macrolactam rings consisting of 7 residues only, which makes the WT xanthomonins and variants thereof extremely stable. Only variants of xanthomonin II with the plug smaller than Ser residue are thermally unstable fold [63]. For the heat-sensitive lasso peptides having 8aa rings, such as caulosegnin I, the thermal stability can be accomplished with a larger lower plug residue, e.g. with the E16W substitution [71]. The heat-sensitive astexin-1 involving 9aa macrolactam ring the same as caulosegnin I can be made heat-stable by the F15W substitution of the lower plug residue [57]. At the same time, the caulosegnin III, caulonodin VI and caulondin V containing the largest known 9aa-macrocycles do not possess thermal stability, even when the biggest amino acid, Trp, is used as a lower plug [65], [71].

Interestingly, the antimycobacterial peptide lassomycin, confirmed by nuclear magnetic resonance (NMR) spectroscopy in dimethyl sulfoxide-d6 to be an unthreaded peptide [73],

probably unwinds in this organic solvent since the artificially prepared branched cyclic form lacks bioactivity [74], [75].

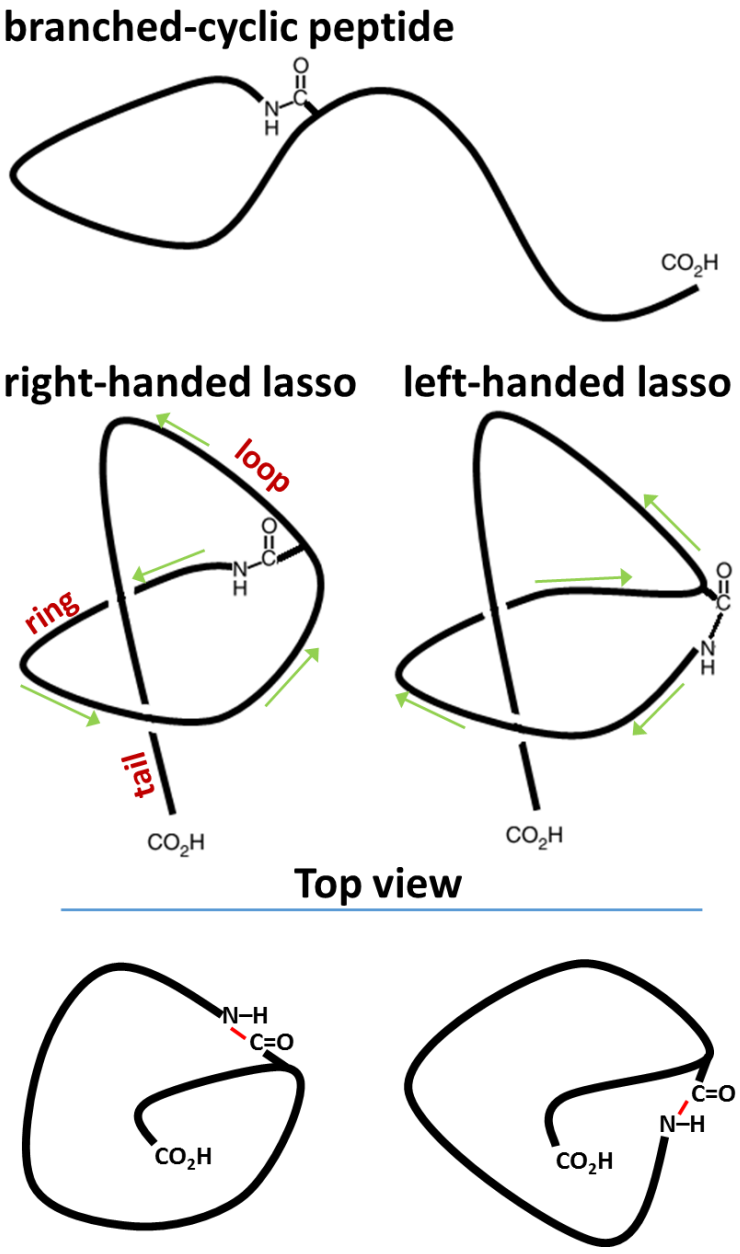


***Figure 5*. A scheme of an unthreaded branched-cyclic peptide in comparison with a threaded one**. Lasso peptides are chiral; only right-handed ones have been found.

The chromatographic behavior of the branched cyclic and threaded lasso peptides is usually distinct (different retention time), while their masses are identical (Figure 6) [62], [69], [76].

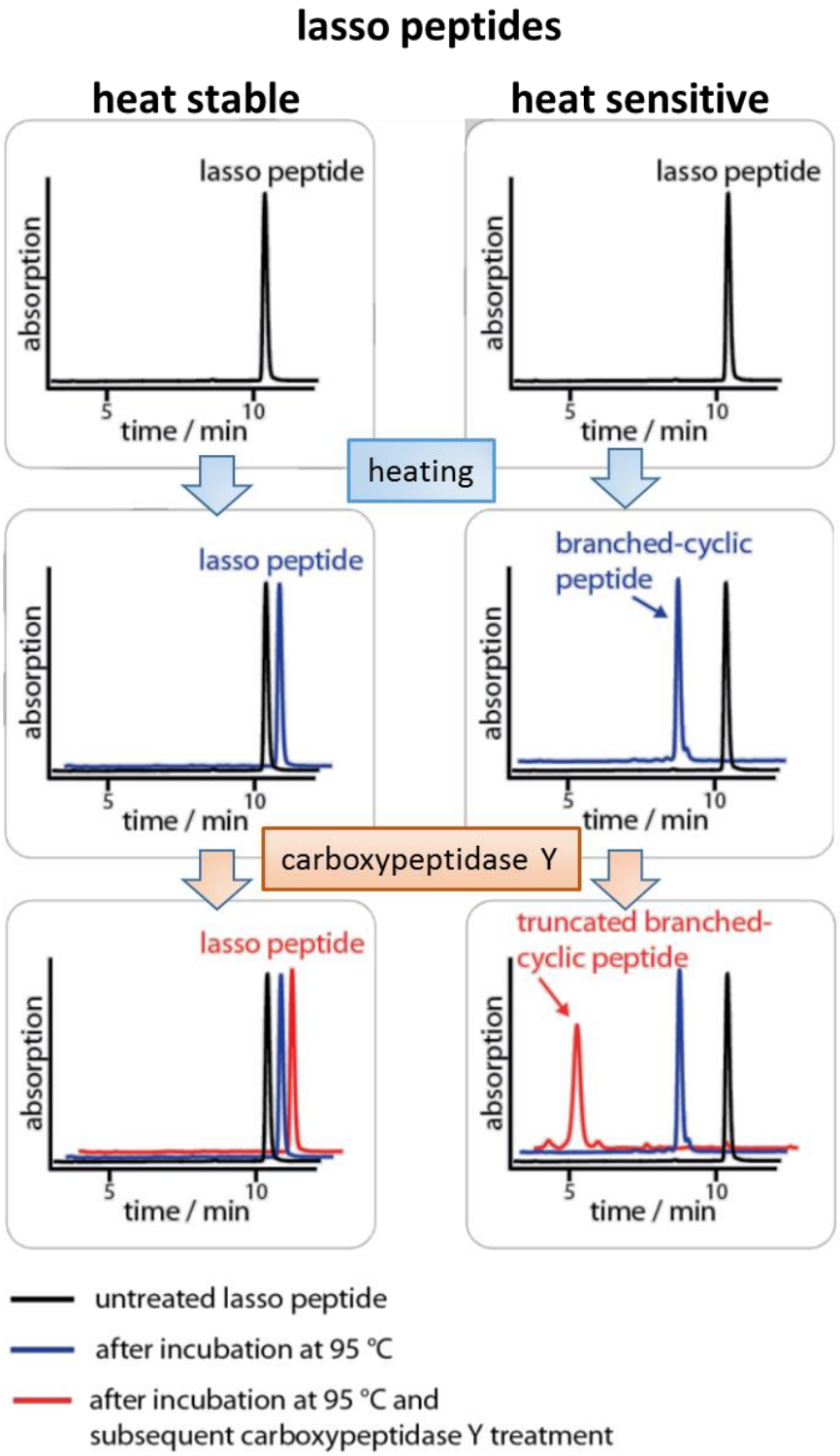


***Figure 6*. Comparison of the LC traces of a heat stable and heat sensitive lasso peptide:** after thermal treatment (middle panel) and subsequent carboxypeptidase Y digestion (lower panel). *Adapted from* [69].

To identify the threaded lasso-structure, carboxypeptidase Y (CPDY) digestion is commonly used [64], [69]. CPDY is capable of nonspecific hydrolyzation of a peptide bond at the C-terminal end of the peptide. In the case of lasso peptides, the action is directed to the tail region.

The tail trapped through the ring is assumed to be protected from degradation through the steric hindrance formed by the surrounding macrolactam ring: the enzyme would stop cleaving the tail close to the ring resulting in a ring-loop containing product. On the contrary, an unthreaded peptide will undergo further degradation to a stable macrocycle product (Figure 6, lower panel).

The unique lasso structure arises in the course of unusual post-translational modifications of a linear precursor peptide. Macrolactam ring is formed by an amide bond between the backbone amine of the N-terminal residue and the side chain of an internal glutamic or aspartic residue at positions 7, 8 or 9, while the tail becomes trapped within.

The lasso peptides can be categorized into four classes [49] (Figure 7). Class I includes six peptides (sviceucin; siamycin I, II, III; specialicin; humidimycin) having four cysteines in the aminoacid sequence. Cysteines allow the structure to be maintained by two disulfide bonds formed between the ring and 1) the loop as well as 2) the tail sections. The majority of class I peptides are additionally stabilized with a plug amino acid located below the ring (see Table 1).

Lasso peptides with no disulfide bonds fall into class II. To date, it is the largest class including about 60 lasso peptides (Table 1). Typically, the tails of class II lasso peptides are trapped in the ring due to steric hindrance mediated by large amino acids (Figure 7). Microcin J25 belongs to class II. Three lasso peptides BI-32169 [77], Pandonodin [78] and Lp2006 [49] contain only a single disulfide bond. BI-32169 belongs to class III as its S-S bond links the ring and the tail region, while Lp2006 and Pandonodin fall into class IV with a tail-to-tail disulfide bond.

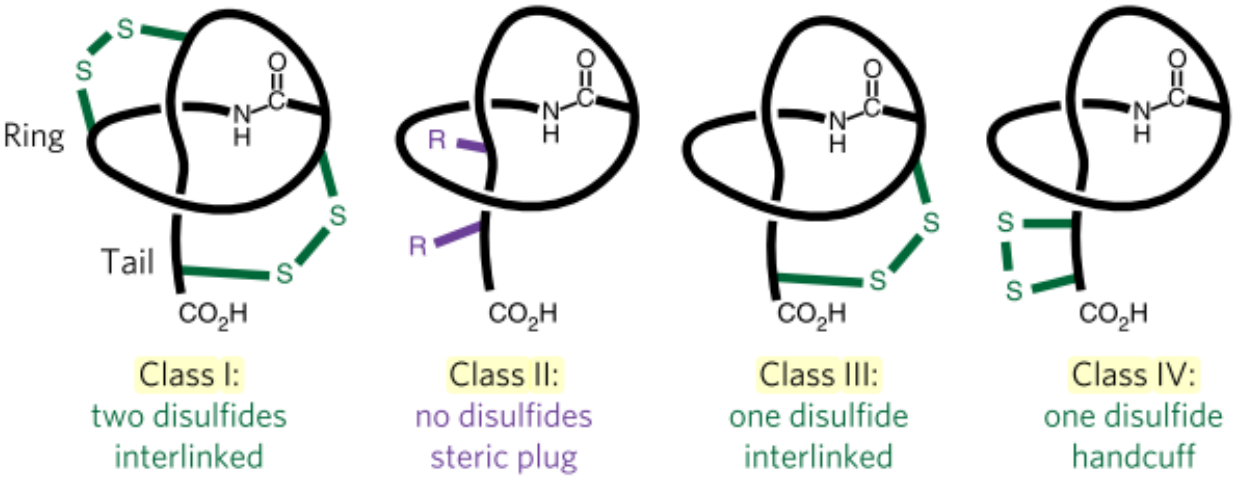


***Figure 7*. Four classes of lasso peptides categorized by the number of disulfide bonds in the structure and their location**. Disulfide bridges are coloured in green, the bulky side chains are indicated by violet. *Adapted from [49]*.

While the size of the ring ranges from 7 to 9 residues, the loop and tail regions are more varied in length (Table 1). For example, the loop section of ubonodin [79] is 18 residues (Tyr9-Tyr26) or 11 residues (Tyr9-Phe19) in MccJ25 compared with only two amino acids in capistruin (Ala10-Arg11). The tail length normally ranges from 5to 10 amino acids, but recently described ubonodin [79] has only 2, while pandonodin [78] has 17 residues below the ring (Table 1). Since the tail is typically linear and unprotected, it often loses some residues after the synthesis is completed through proteolytic degradation. Albusnodin and anantin lose one tail residue; anantin and astexin1 lose two; astexin 2,-3 – three; benenodin-1, burhizin, capistruin and caulonodin I-VI lose four; caulonodin VII and caulosegnin I – five; caulosegnin II – six; and, finally caulosegnin III loses seven tail residues after synthesis.

Table 1. An overview of lasso peptides known to date. Cysteines involved in disulfide bond formation (classes I, III, IV) are underlined; residues involved in ring formation are coloured dark red; steric-lock residues are coloured green. Residues typically cleaved off from the main product are coloured in light grey. The actual number of amino acids before cleavage is indicated in brackets. The names of lasso peptides from *Proteobacteria*, *Actinobacteria* and *Firmicutes* are coloured in violet, black and orange, respectively.

Lasso Peptide	**Core sequence**	**Ring formed by**	**Plug (if known)**	**Length, aa**	
				Ring	**Core**
Class I					
Humidimycin [80]	**CLGIGSCDDFAGCGYAIVCFW**	Cys1/Asp9	Phe20, S-S: 1-13; 7-19	9	21

RP-71955/siamycin III/aborycin [81], [82]	CLGIGSCNDFAGCGYAVVCFW	Cys1/Asp9	Phe20, S-S: 1-13; 7-19	9	21
Siamycin I / MS-271 / NP-06 [58], [83]	CLGVGSCNDFAGCGYAIVCFW	Cys1/Asp9	Phe20, S-S: 1-13; 7-19	9	21
Siamycin II [84]	CLGIGSCNDFAGCGYAIVCFW	Cys1/Asp9	Phe20, S-S: 1-13; 7-19	9	21
Ssv-2083 / sviceucin [85]	CVWGGDCTDFLGCGTAWICV	Cys1/Asp9	S-S: 1-13; 7-19	9	20
Specialicin [86]	CLGVGSCVDFAGCGYAVVCFW	Cys1/Asp9	Phe20, S-S: 1-13; 7-19	9	21
Class II					
Achromosin [87]	GIGSQTWDTIWLWD	Gly1/Asp8		8	14
Acinetodin [88]	GGKGPIFETWVTEGNYYG	Gly1/Glu8	Tyr16/Tyr17	8	18
Actinokineosin [89]	GYPWWDYRDLFGGHTFISP	Gly1/Asp9		9	19
Albusnodin [90]	GQGGGQSEDKRRAYNC	Gly1/Glu8		8	16 (15)
Anantin B1 (B2) [49], [91]	GFIGWGNDIFGHYSGGF	Gly1/Asp8		8	16 (17)
Astexin-1 [62]	GLSQGVEPDIGQTYFEESRINQD	Gly1/Asp9	Tyr14/Phe15	9	19 (23)
Astexin-2 [70]	GLTQIQALDSVSGQFRDQLGLSAD	Gly1/Asp9	Phe15/Arg16	9	21 (24)
Astexin-3 [70]	GPTPMVGLDSVSGQYWDQHAPLAD	Gly1/Asp9	Tyr15/Trp16	9	22 (24)
Benenodin-1 [76]	GVGFGRPDSILTQEQAKPM	Gly1/Asp8	Glu14/Gln15/ Lys17	8	19
Brevunsin [92]	GDMGEEVIEGLVRDSLYPPAG	Gly1/Glu9	Tyr17	9	21
Burhizin/Burhizin-23 [61], [93]	GGAGQYKEVEAGRWSDRIDSDDE	Gly1/Glu8		8	17 (23)
Capistruin [67]	GTPGFQTPDARVISRFGFN	Gly1/Asp9	Arg11/Arg15	9	19
Cattlecin [94]	SYHWGDYHDWHHGWYGWWDD	Ser1/Asp9		9	20
Caulonodin I [65]	GDVLNAPEPGIGREPTGLSRD	Gly1/Glu8		8	17 (21)
Caulonodin II [65]	GDVLFAPEPGVGRPPMGLSED	Gly1/Glu8		8	17 (21)
Caulonodin III [65]	GQIYDHPEVGIGAYGCEGLQR	Gly1/Glu8		8	17 (21)
Caulonodin IV [61]	SFDVGTIKEGLVSQYYFA	Ser1/Glu9	Tyr16/Phe17	9	18
Caulonodin V [61]	SIGDSGLRESMSSQTYWP	Ser1/Glu9	Tyr16/Trp17	9	18
Caulonodin VI [61]	AGTGVLLPETNQIKRYDPA	Ala1/Glu9	Arg15/Tyr16	9	19
Caulonodin VII [61]	SGIGDVFPEPNMVRRWD	Ser1/Glu9	Arg15/Trp16	9	17
Caulosegnin I [64]	GAFVGQPEAVNPLGREIQG	Gly1/Glu8	Arg15/Glu16	8	19
Caulosegnin II [64]	GTLTPGLPEDFLPGHYMPG	Gly1/Glu9	His15/Tyr16	9	19

Caulosegnin III [64]	GALVGLLLEDITVARYDPM	Gly1/Glu9	Arg15/Tyr16	9	19
Chaxapeptin [95]	GFGSKPLDSFGLNFF	Gly1/Asp8	Leu12/Asn13	8	15
Citrocin [96]	GGVGKIIEYFIGGGVGRYG	Gly1/Glu8	Arg17/Tyr18	8	19
Citrulassin A [49]	LLGLAGNDRLVLSKN	Leu1/Asp8		8	15
Fuscanodin/ Fusilassin [51], [72]	WYTAEWGLELIFVFPRFI	Tyr1/Glu9	Arg16/Phe17	9	18
Klebsidin [88]	GSDGPIIEFFNPNGVMHYG	Gly1/Glu8	Tyr18	8	19
Lagmysin [49]	LAGQGSPDLLGGHSLL	Leu1/Asp8		8	16
Lariatin A (B) [97]	GSQLVYREWVGHSNVIKPGP	Gly1/Glu8	His12/Asn14	8	18 (20)
Lassomycin [73], [98]	GLRRLFADQLVGRRNI	Gly1/Asp8		8	16
Leepeptin [99]	LYGVRNDEEINWHFDYWT	Leu1/Glu8	Phe14/Tyr16+ Trp17	8	18
Microcin J25 [66]	GGAGHVPEYFVGIGTPISFYG	Gly1/Glu8	Phe19/Tyr20	8	21
Moomysin [49]	SYHWGDYHDWHHGWYGWWDD	Ser1/Asp9		9	20
Mycetohabin-15 [93]	GGSGQYREAGVGRFL	Gly1/Glu8	Arg13/Phe14	8	15
Mycetohabin-16 [93]	GGSGKYREAGVGRFLD	Gly1/Glu8	Arg13/Phe14	8	16
Paeninodin [100]	AGPGTSTPDAFQPDPDEDVHYDS	Ala1/Asp9	Tyr21	9	23
Propeptin [101]	GYPWWDYRDLFGGHTFISP	Gly1/Asp9		9	19
Pseudomycoidin [102]	AGPGKRLVDQVFEDEDEQGALHHS	Ala1/Asp9		9	24
Res-701-1 [103]	GNWHGTAPDWFFNYYW	Gly1/Asp9	Phe12/Tyr14	9	16
Res-701-3 [103]	GNWHGTSPDWFFNYYW	Gly1/Asp9	Phe12/Tyr14	9	16
Rhodanodin [61]	GVLPIGNEFMGHAATPGITE	Gly1/Glu8		8	17 (20)
Rubrivinodin [61]	GAPSLINSEDNPAFPQRV	Gly1/Glu9	Arg17	9	18
Sphaericin [61]	GLPIGWWIERPSGWYFPI	Gly1/Glu9	Phe16/Pro17	9	18
Sphingonodin I [61]	GPGGITGDVGLGENNFGLSDD	Gly1/Asp8		8	17 (21)
Sphingonodin II [61]	GMGSGSTDQNGQPKNLIGGISDD	Gly1/Asp8		8	19 (23)
Sphingopyxin I [61]	GIEPLGPVDEDQGEHYLFAGGITADD	Gly1/Asp9	His15/Tyr16	9	21 (26)
Sphingopyxin II [61]	GEALIDQDVGGGRQQFLTGIAQD	Gly1/Asp8		8	19 (23)
Sro15-2005 [104]	GYFVGSYKEYWSRRII	Gly1/Glu9		9	16
Streptomonomicin [105]	SLGSSPYNDILGYPALIVIYP	Ser1/Asp9	Tyr13/Leu16	9	21
Subterisin [106]	GPPGDRIEFGVLAQLPG	Gly1/Glu9	Leu12/Gln14	9	17
Sungsanpin [107]	GFGSKPIDSFGLSWL	Gly1/Asp8	Leu12/Trp14	8	15

Syanodin I [61]	GISGGTVDAPAGQGLAGILDD	Gly1/Asp8		8	17 (21)
Ubonodin [79]	GGDGSIAEYFNRPMHIHDWQIMDSGYYG	Gly1/Glu8	Tyr26/Tyr27	8	28
Ulleungdin [108]	GFIGWGKDIFGHYGG	Gly1/Asp8		8	15
Xanthomonin I [63]	GGPLAGEEIGGFNVPGISEE	Gly1/Glu7	Phe12	7	16 (20)
Xanthomonin II [63]	GGPLAGEEMGGITTLGISQD	Gly1/Glu7	Ile12	7	14 (20)
Xanthomonin III [63]	GGAGAGEVNGMSPIAGISEE	Gly1/Glu7		7	13 (20)
Zucinodin [61]	GGIGGDFEDLNKPFDV	Gly1/Glu8		8	16
Class III					
Bi-32169 [77]	GLPWGCPSDIPGWNTPWAC	Gly1/Asp9	Trp13/Trp17 S-S: 6-19	9	19
Class IV					
Lp2006 [49]	GRPNWGFENDWSCVRVC	Gly1/Glu8	S-S: 13-17	8	17
Pandonodin [78]	GVLGNDAEGITLLPLCFKPICIPTLPPLTG GHA	Gly1/Glu8	Leu15/Phe17 S-S: 16-21	8	33

The vast majority of the lasso peptides characterized to date do not have any extra tailoring modifications and the existing classification ignores them as well. Actually, only six additionally modified lasso peptides have been described so far: **lassomycin** with methylated C-terminal carboxyl [73], [98], **RES**-701-2/4 having Trp side chain hydroxylation [103], **siamycin** I carrying epimerized C-terminal Trp residue in D-isomeric form [58], [83], **albusnodin**, an acetylated lasso peptide [90], and **paeninodin** from *Firmicutes*, subjected to C-terminal phosphorylation [100], [109] (Figure 8).

The newly reported lasso peptide **citrulassin** A, a representative of a previously unknown subfamily of citrulassins, in its mature form carries Arg9 modified to citrulline [49]. Such modification is very rare for bacteria and had never been found in RiPP before [110].

The group of additionally modified lasso peptides is expanding [49], [61], [100], [111], and existing classification most likely will have to be updated soon.

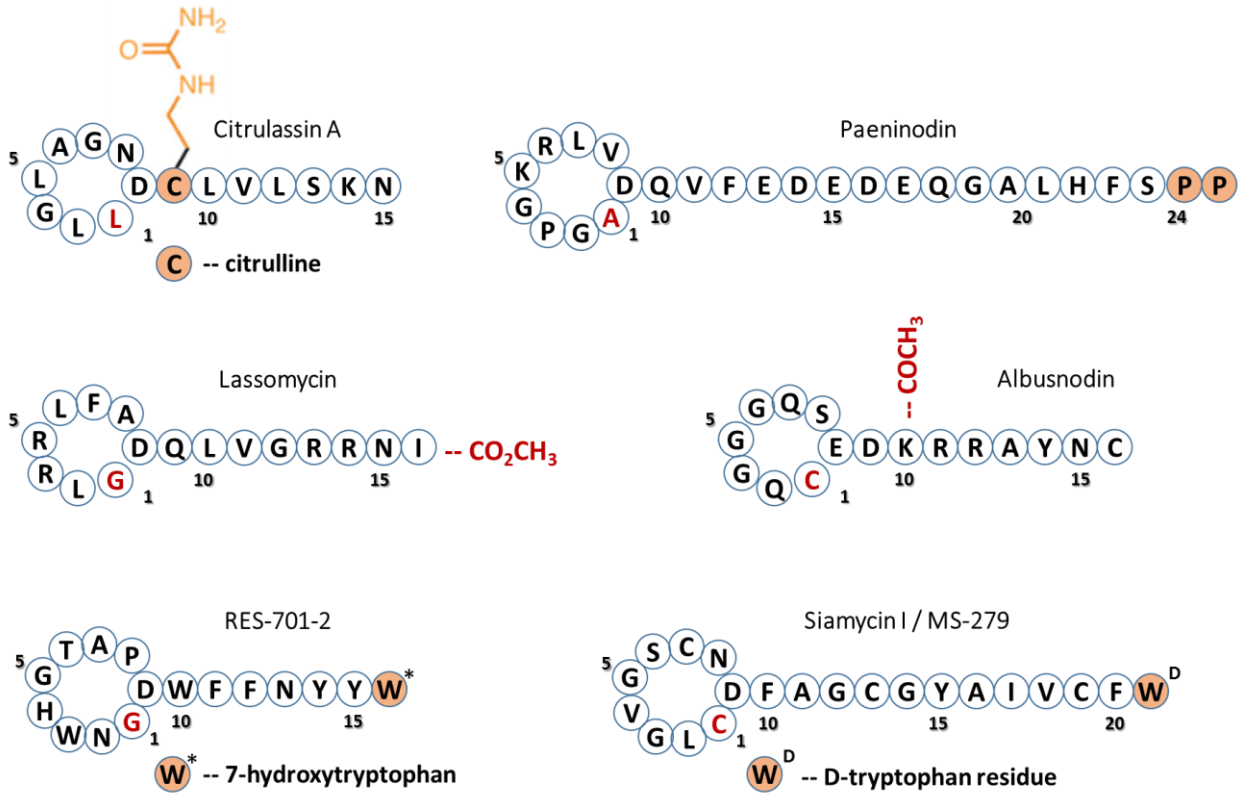


***Figure 8. Lasso* peptides with additional modifications**. Peptides carrying C-terminal tailoring modifications: lassomycin, [73] with a methyl group at C-terminal carboxyl; RES-701-2 hydroxylated at the side chain of C-terminal Trp [103]; siamycin I / MS-279 [112] with D-Trp residue at the C-terminus, albusnodin with acetyllysine post-translational modification, and paeninodin phosphorylated at the side chain of C-terminal serine [109]. A recently reported lasso peptide citrulassin A possesses Arg9 modified to citrulline [49], [25].

1.3. Chemical synthesis of lasso peptides

Lasso peptides belonging to class II are mechanically interlocked molecules resembling [1]rotaxanes [54]. The epitope molecular grafting approach has been effectively applied to create a microcin J25 derivative that was able to specifically bind and inhibit integrin [113] as well as D-enantiomer of BI-32169, which shows a strong glucagon receptor antagonist activity and much higher enzymatic stability compared to the L-lasso peptide [114], [115].

An approach to creating rotaxanes and catenanes from the lasso peptide precursor MccJ25 was developed [116]. However, the only BI-32169 lasso peptide has been totally synthesized *in vitro* [114], while branched cyclic MccJ25 can be obtained via amide or disulfide bond formation using solid-phase peptide synthesis approaches [117], [118] or through cyclization with copper-

catalyzed azide-alkyne cycloaddition [60]. The rotaxane-like topology is the main obstacle to the chemical synthesis of lasso peptides [119], [120].

Several reports about successful total chemical synthesis of lasso peptides were published. Recently peptide-based rotaxanes reminiscent of natural MccJ25 (Figure 9) were developed [121] and the cryptand-imidazolium supported method enabled the total synthesis of lasso peptide BI-32169 and its D-enantiomer [114]. This synthetic strategy is anticipated to facilitate the construction of diverse lasso peptide libraries in the future.

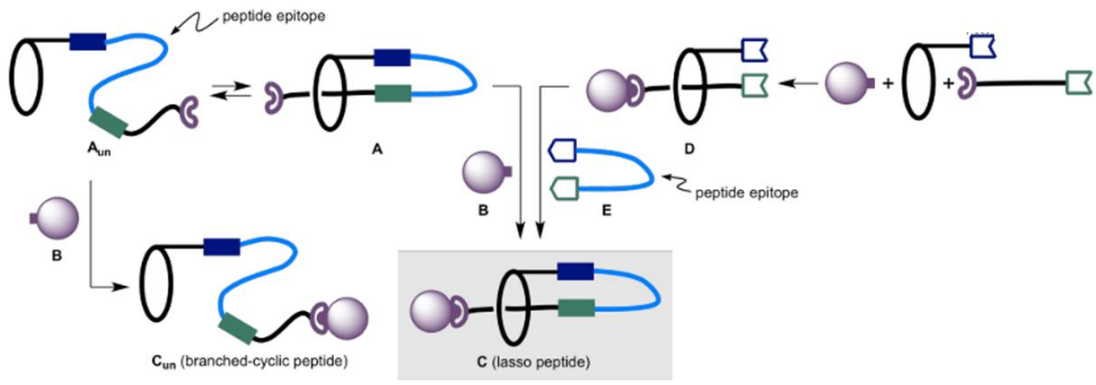


***Figure 9*. Two proposed strategies to construct lasso peptides using epitope grafting approach**. 1) Generating branch-cyclic form with covalent link between a macrocycle and a peptide-containing segment. Resulting molecules are in a thermodynamic equilibrium with pseudo [1]rotaxane. Attaching a capping agent B would lead to catching and stabilizing a trapped variant of lasso peptide C; 2) Starting with [2]rotaxane D formation and subsequent attachment of a linear peptide E by two sequential chemical ligations. *Reprinted from* [121].

1.4. Enzymatic biosynthesis of lasso peptides

The lasso peptide biosynthesis starts with ribosome-driven translation of a small precursor peptide consisting of the leader and the core parts. Typical RiPP biosynthesis is directed by RiPPs Recognition Element (RRE) that binds to the leader peptide [6], [122]–[126], while the core part undergoes further maturation by specific enzymes. Similar to most RiPPs, the biosynthetic machinery of lasso peptides incorporates the core maturation enzymes as well as auxiliary proteins typically co-encoded in the same genetic cluster [49], [58], [73], [83], [100], [103], [109].

Apart from one well-known example, the recently reported lasso peptide albusnodin [90], only three processing proteins are necessary and sufficient for the lasso peptide formation (Figure 10) [49], [61], [127]–[129]. The first one is a RiPPs Recognition Element (B1-protein) that binds the precursor peptide and directs further enzymatic modifications; the second one is the leader peptidase (B2-protein) that specifically hydrolyzes the peptide bond at the leader-core junction. The third one is the lasso cyclase (C-protein) catalyzing an isopeptide bond formation in the prefolded core peptide resulting in N-terminal macrocycle creation (Figure 10), while C-terminal tail is trapped within it.

Interestingly, the BGCs found in *Proteobacteria* typically encode a fused B1+B2 protein as a single polypeptide chain. Though the function of each of these enzymes was revealed, the general structure of the biosynthetic maturation complex is still obscure. The lasso peptide biosynthesis likely requires folding of precursor peptide with processing enzymes in addition to the chemical transformations [72].

The vast majority of lasso peptides described to date (about 70 compounds, Table 1) are not subjected to any extra post-translational modifications, and three processing enzymes listed above have been demonstrated to be sufficient to complete lasso peptide formation. In case of albusnodin biosynthesis, a gene coding for acetyltransferase was proved to be indispensable for lasso peptide formation, the first example of a lasso peptide biosynthesis with an obligate tailoring post-translational modification.

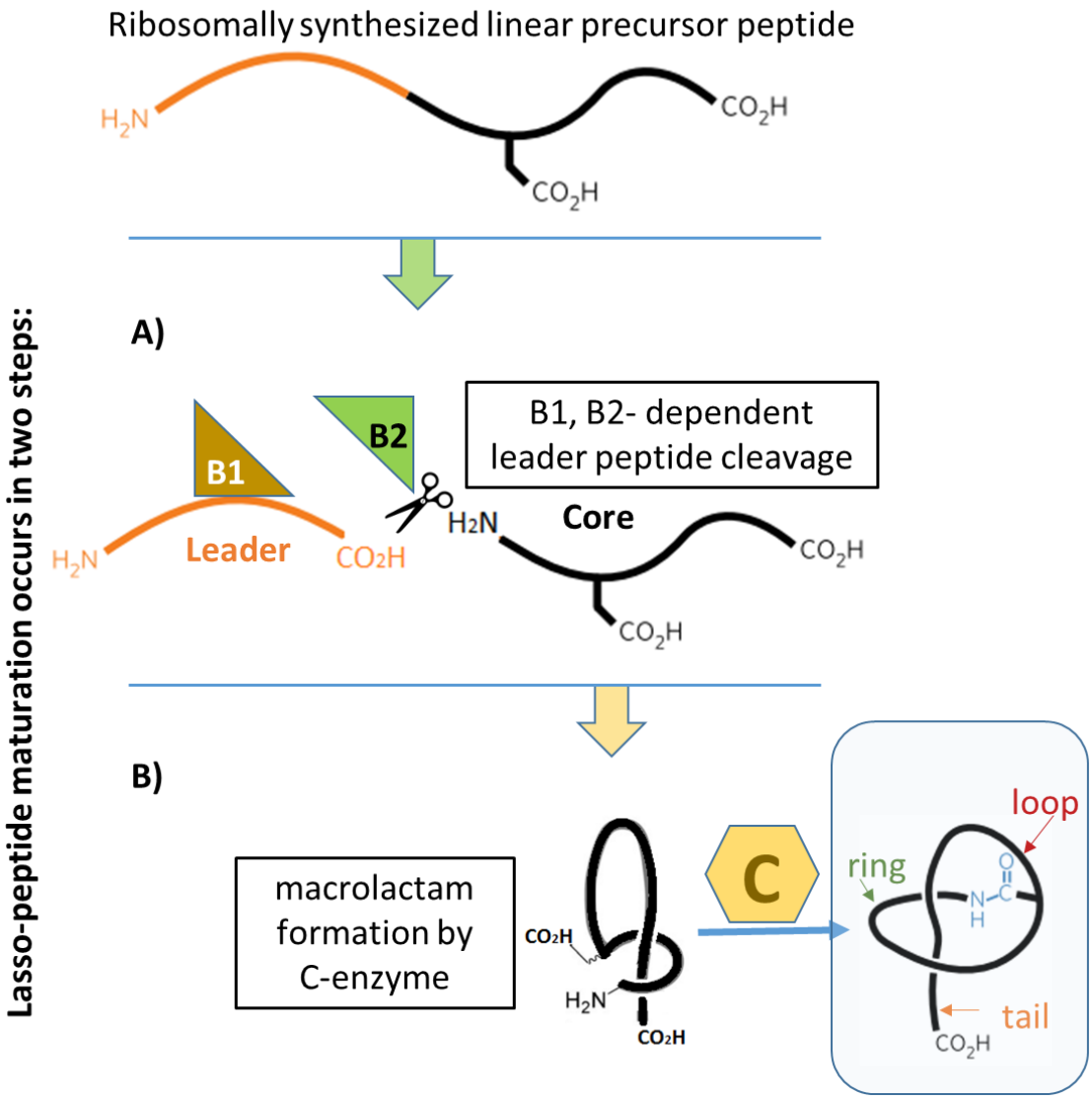


Figure 10. **A typical lasso peptide biosynthetic pathway by essential processing enzymes occurs in two steps:** A) leader peptide cleavage; B) catalysis of the isopeptide bond formation in the prefolded core peptide resulting in the N-terminal macrolactam ring trapped with the C-terminal tail.

1.4.1. Lasso precursor peptide structure

Synthesized by the ribosome, the precursor peptides typically consists of a cleavable leader part as well as a modifiable core. Both leader and core peptides demonstrate diversity in lengths and, especially, aminoacid sequences [42], [49], [57].

Leader peptide and its role in biosynthesis

One the background of wide variability of leader peptides sequence composition there is conserved threonine residue (or isoleucine in cases of lasso-precursors from *Acidobacteriaceae* [49]) at the penultimate position [16], [49], [62]–[64], [67], [130]. The leader sequence length varies from 18 (in caulosegnin II [64]) to 37 residues (in MccJ25 [128]) in isolated peptides and extends to 88 residues in predicted ones [49].

RiPP leader peptides can be considered as multifunctional due to their ability to contribute to the maturation [17], [46], [131] and the secretion as well as to protect the precursor peptide from degradation [6], [15], [123]. In case of lasso peptide biosynthesis the leader sequences guide the maturation by interacting with the B1 protein [15]–[17], [124], [126], [132]–[134].

Most of predicted leader sequences (63%) contain the L1 motif ($Y^{-15}xxP^{-12}xLx_3Gx_5T^{-2}x$) whose first portion ($Y^{-15}xxP^{-12}$) is found in all lasso peptide-containing phyla [49], [85], [90], [95], [105], [135]. There is a strong correlation of the L1 motifs with three RRE motifs in case of separately encoded B1/B2 subunits or with a single motif of fused B proteins [49]. *In vivo* studies of paeninodin [136] and chaxapeptin [137] biosynthesis confirmed that Y^{-15} and P^{-12} residues are essential for interaction with RRE, while in case of the fusilassin precursor peptide Y^{-17} and P^{-14} residues were confirmed to be essential [51]. Replacing conserved T^{-2} with Ala does not alter recognition by B1 protein but affects the enzymatic proteolytic cleavage by B2 [136].

A rarer conserved motif $LI^{-11}xLG^{-8}xA^{-6}xxxT^{-2}x$ was bioinformatically observed in a number of predicted lasso peptides from *Proteobacteria* and *Actinobacteria*. The importance of Thr^{-2}, Gly^{-8} as well as of four hydrophobic residues at positions -12, -11, -9 and -6 was confirmed *in vivo* for the caulonodin V biosynthesis [65]. Interestingly, no conserved leader motifs described above were found in microcin J25 precursor peptide.

Apart from interacting with the processing proteins, the leader sequences were additionally proposed to function as intramolecular chaperones assisting in the folding of the precursor peptides

during maturation [60], [138]. Indeed, maturation of leaderless precursor peptide of MccJ25 results in formation of branched-cyclic peptides [60], [130]. The full-sized leader peptide was shown to be a crucial element for the lasso peptide caulonodin V [65] and for paeninodin production [136]. The *in vitro* synthesis of the MccJ25 dramatically decreased in the leaderless system [130], while the fusilassin synthesis dropped to 10% compared with the full-length one [72].

Core peptides and their variety

The core peptides convert into the mature lasso form and demonstrate a high variety in both length and sequence composition. Validated core peptides range from 14 (achromosin [87]) to 28 residues (sphingopyxin I [61]), while the predicted ones are 12-49 aa in length [49]. The sequence composition appears to be restricted by only two requirements for residues involved in macrolactam formation. Firstly, the Asp or Glu residues carrying carboxyl group at the end of the side chain are to be present at positions 7-9. Secondly, proline is to be absent at the first position with a clear trend towards over-representation (62%) of glycines in this position. Gly residues are actually over-represented in the macrocycle as well, 22.8% versus 7.1% expected [49].

1.4.2. RiPP Recognition Element – B1 protein

RiPP Recognition Element is a small protein (of ca. 90 amino acids [62]) binding the precursor peptide and recruiting it for future maturation [124], [126], [139], [140]. It has been found in over a half of all known prokaryotic RiPP classes [6], [49], [124], [134], [141]. RRE subunits do not have any sequence homology to other proteins of known functions, but exhibit some structural similarity of the conserved peptide domain forming Winged Helix-Turn-Helix (wHTH) motifs [124], [126], [140]. Most lasso-RRE include a LDxxxxRYFxL sequence [42], [61].

Typically, RREs are fused within one of RiPP biosynthetic enzymes [15], [124] and far less frequently they represent a stand-alone subunit, such as PqqD-protein that was shown to be

essential for pyrroloquinoline quinone (PQQ) biosynthesis [142]–[144]. The wHTH domains of fused LynD (residues 2–81) cyclodehydratase involved in cyanobactin biosynthesis [145], [146] and NisB (residues 153–223) lanthipeptide dehydratase involved in nisin biosynthesis [133] are closely related to the PqqD-protein [124]. PqqD-homologs have also been demonstrated to be involved in the maturation of sactipeptide thurincin H [125], lantibiotic NAI-107 [147], and lasso peptides [49], [57], [105], [131].

RREs consist of three N-terminal β-strands and three C-terminal α-helices. They typically bind the leader peptide in a cleft between the β3-strand and α3-helix [69]. A number of co-crystallization and NMR experiments have clearly demonstrated that a wHTH motifs are strongly associated with their respective leader precursor peptides [18], [126], [132], [133], [148]. During this interaction the leader sequence can form a β-strand with the β-sheets of wHTH motifs of enzymes LynD [18], NisB [133], TfuB1 [148], PqqD [132], etc. Surface plasmon resonance spectroscopy and isothermal calorimetry techniques have demonstrated micromolar affinity of PqqD-protein to the precursor peptide (KD ~ 200 nM) [149].

In case of lasso peptide biosynthesis, the RRE is commonly referred to as the "B1" protein. The gene expression and complementation experiments *in vivo* have proved the B1 protein to be indispensable in lasso peptides maturation [72], [83], [85], [104], [105], [135]. B1 subunits were shown to tightly bind cognate leader peptides and orchestrate subsequent enzymatic modification [42], [49], [124], [127], [131], [135]. They presumably recruit the B2 leader peptidase to the substrate [131]. In addition to some specific residue pairs important for B1/B2 interaction as revealed by signs of coevolutionary pressure in cognate pairs, opposite charges probably provide for primary electrostatic attraction: RRE proteins are notably acidic pI (4.7 ± 1.2) while leader peptidases are strongly basic pI (10.5 ± 1.1). Lasso cyclases exhibit neutral pI (7.4 ± 1.1) [51]. The mechanism of leader peptide recognition by the B1 proteins may involve a fit of conserved residues of the leader peptide onto a hydrophobic cleft B1 protein [148].

In case of the lasso-clusters coding for the separate B1/B2 proteins, the conserved YxxP motif in leader sequence has been confirmed as indispensable for RRE recognition [136], [148]. Typically, the lasso RREs do not correlate with any core peptide motifs (and vice versa) [49], [136], [150], [151]. But in the case of lariatin biosynthesis, the PqqD homolog LarB1 interacts with the core sequence and probably enables substrate to be further matured by LarB2 and, perhaps, LarC enzymes [17].

The B1 subunit was also proposed to be a novel type of peptide chaperones [132], [149]. However, the exact role of B1 as well as the site of its interaction with B2 remains elusive, though the C-terminal region of B1 is suspected to be involved [131]. Presumably, the hydrophobic patch on the surface of the intermolecular β-sheet formed by B1 and leader may be used in the interaction with the core or the B2 subunit [148].

Given that the PqqD-protein contains a bound phosphate group [152], the B1 protein was presumed to be an ATP-binding subunit catalyzing hydrolysis of ATP. In fact, McjB, a fused RRE-protease enzyme, was shown to be an ATP-dependent cysteine protease. *In vitro* experiments confirmed that processing of precursor peptide McjA occurs in the presence of ATP [130]. On the contrary, the precursor peptide of paeninodin, PadeA, can be successfully processed in the absence of ATP by separated RRE (PadeB1) and leader peptidase (PadeB2) *in vitro* [131].

1.4.3. Leader peptidase – B2 protein

Another essential lasso-synthesis enzyme is the leader peptidase commonly referred to as the B2 enzyme. These proteins exhibit homology to eukaryotic transglutaminases belonging to a large family of cysteine proteases [43]. The B2 enzymes have been postulated to catalyze the proteolytic cleavage of the leader sequence from precursor peptide by establishing the γ-glutamyl-ε-lysine cross-links [153]. The B2 proteins contain conserved Cys-His-Asp/Cys-His residues forming the catalytic triad/dyad typical for transglutaminase peptide family [154], [155]. In fact,

the single substitution of Cys150, His182 or Asp194 to Ala in McjB completely annuls the production of mature MccJ25 *in vivo* [156].

The interaction of B2 enzymes with their peptide substrate is presumably mediated by the B1 subunits [134]. Indeed, *in vivo* proteolysis of paeninodin precursor peptide was demonstrated to be RRE-dependent [131] and only carried by the cognate leader peptidase PadeB2. Mutational analysis of paeninodin leader sequences revealed that only Thr^{-2} residue is crucial for the proteolysis by PadeB2 *in vitro*. This residue is highly conserved in lasso precursors [136].

Typically, the lasso peptide BGCs originating from *Proteobacteria* code for fused B1/B2 bifunctional B-proteins [61] consisting of N-terminal recognition and C-terminal peptidase parts, corresponding to B1 and B2 subunits, respectively [42], [49], [61], [104], [105], [157]. The fused B-enzymes seem to be evolutionary descendants of separated ones [49]. Artificially fused B1–B2 proteins support lasso peptide biosynthesis [100]. Conversely, artificially split B protein from rubrivinodin biosynthetic cluster remains active as well [131].

1.4.4. Macrolactam synthetase – C-protein

The final essential step in the lasso peptide maturation is catalyzed by the macrolactam synthetase called protein C (of ca. 500-700 amino acids). These adenylate-forming enzymes catalyze the formation of the isopeptide bond between the amino group of the first amino acid of the core and the side chain carboxyl group of aspartic or glutamic residues at the position 7, 8 or 9. The reaction results in the macrolactam ring being formed.

C-proteins are homologous to β-lactam synthetases [156], [158] as well as to the C-terminal domain of asparagine synthetase B (AS-B) [154], [158], [159]. AS-B is involved in coordination of Mg^{2+} for ATP binding in ATP-dependent amidation of aspartate to form asparagine [154]. *In vitro* experiments with McjC enzyme confirmed the Mg^{2+}/ATP-dependent catalysis of the isopeptide bond [158]. Single aminoacid substitutions of residues forming the

ATP-Mg^{2+} binding site with Ala (S199A, D203A and D302A) significantly decreased MccJ25 production [130], [156] (Figure 11).

The reaction of cyclisation mediated by protein C is suggested to proceed in two separate steps: (1) side chain carboxylate adenylation and (2) lactam formation [130]. At the first step, the McjC enzyme and its homologs have been proposed to activate the γ -carboxyl group of Asp/Glu residues via adenylation in $Mg^{2+/}$ATP-dependent manner [130], [155]. At the second step, the newly formed N-terminal amine displaces adenosine diphosphate (ADP) from the modified carboxylate by a nucleophilic attack via a tetrahedral transition state or an oxyanion intermediate, resulting in isopeptide linkage [130]. Given the knotted structure of mature lasso peptides stabilized with lock residues around the ring, the reaction of cyclisation is assumed to occur on a prefolded substrate (Figure 11).

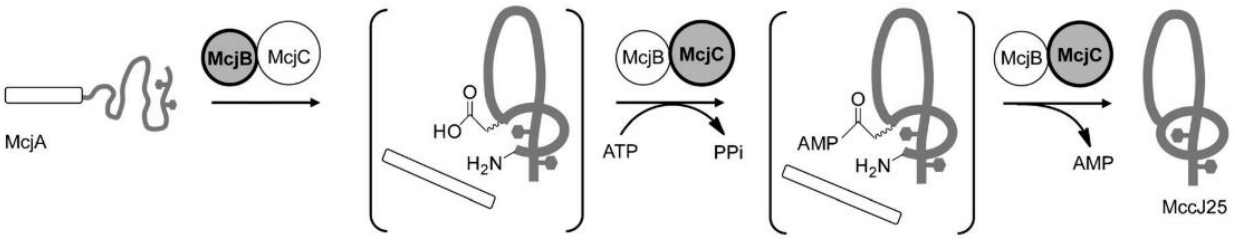


Figure 11. **Proposed two-step biosynthetic scheme of microcin J25 maturation in the presence of ATP and Mg^{2+} ions**. McjB and McjC presumably interact to form the synthetase. The functional enzyme at each step is indicated with bold outlines, the grey empty bar represents the leader peptide. *Adapted from* [130].

Interestingly, the C enzyme has demonstrated a high degree of substrate promiscuity. A broad mutation analysis of various core peptides revealed the incredibly relaxed specificity of the lasso-cyclase towards single amino acid substitutions [16], [62]–[65], [67], [136], [150], [151], [160] as well as towards the ring size [51], [64]. The MccJ25 biosynthesis was only abolished with substitutions in three core peptide positions: Gly1, Gly2, and Glu8 [150]. Dramatic changes in core peptide McjA, including insertions (insA5–6, insG22, insD22) or deletions (ΔF10-V11, ΔG12-G14), did not prevent maturation by the McjC enzyme [156]. Notably, mature variants of

mutated MccJ25 possess either a lasso or a branched-cyclic topology, suggesting that either linear or folded precursor peptides can be recognized and cyclized by McjC or the two final forms can interconvert [156]. The fusilassin lasso-cyclase was also readily able to tolerate substitutions of the first core position, and a range (±1 residue) of ring sizes [51].

1.4.5. Possible biosynthetic mechanism

Akin to some RiPPs matured with the assembled large synthase involving several processing enzymes [6], [161], lasso peptide maturation is assumed to require co-dependent efforts of several interacting processing enzymes. Unfortunately, the exact steps in the biosynthesis of lasso peptides as well as its kinetics remain poorly explored since most recombinant lasso maturation enzymes are poorly soluble and can be obtained in low yields. While the B1 proteins can be successfully isolated [17], [124], [131], B2 and C are less stable or co-purify with chaperones [130], [131], except for two examples: the MccJ25 and fusilassin biosynthetic enzymes. Copurification of fusilassin BGC proteins TfuB2 and TfuC with N-terminal 6xHis tags led to good yield, and the complex between the two proteins was confirmed to be active *in vitro* [72].

In the case of MccJ25 synthetic machinery, McjB and McjC also probably form a synthetase complex [130]. Inactive McjC (e.g., D302A mutant) was sufficient to allow the cleavage of the leader peptide by McjB that did not occur in the absence of McjC. Conversely, inactive variants of McjB (e.g., C150S, C150A/H182A mutants) were sufficient to lead to cyclisation of the core peptide by McjC, but McjC alone was not able to do so [130]. The leader peptide was also shown to be essential for biosynthesis, and a leaderless core in *in vitro* system resulted in a dramatically low yield of mature MccJ25 [130].

Similarly, *in vitro* biosynthesis of fusilassin with separated B1 and B2 proteins probably requires a preformed synthetase complex. The presence of all three processing proteins and ATP together was shown to be essential for fusilassin formation [51], [72]. A leaderless system was still

functional, but only at ~10% level of complete system. Interestingly, the TfuB1 protein was required even in the leaderless system which suggests its role not only in the binding of the leader peptide, but also, for example, in interacting with the TfuB2/TfuC complex and ensuring its stability [72].

Unlike MccJ25, fusilassin and the paeninodin *in vitro* biosynthesis has been demonstrated to proceed independently of C enzyme with the B1 and B2 subunits complex only [131] (Figure 12). The lariatin biosynthesis occurs in the same way [135].

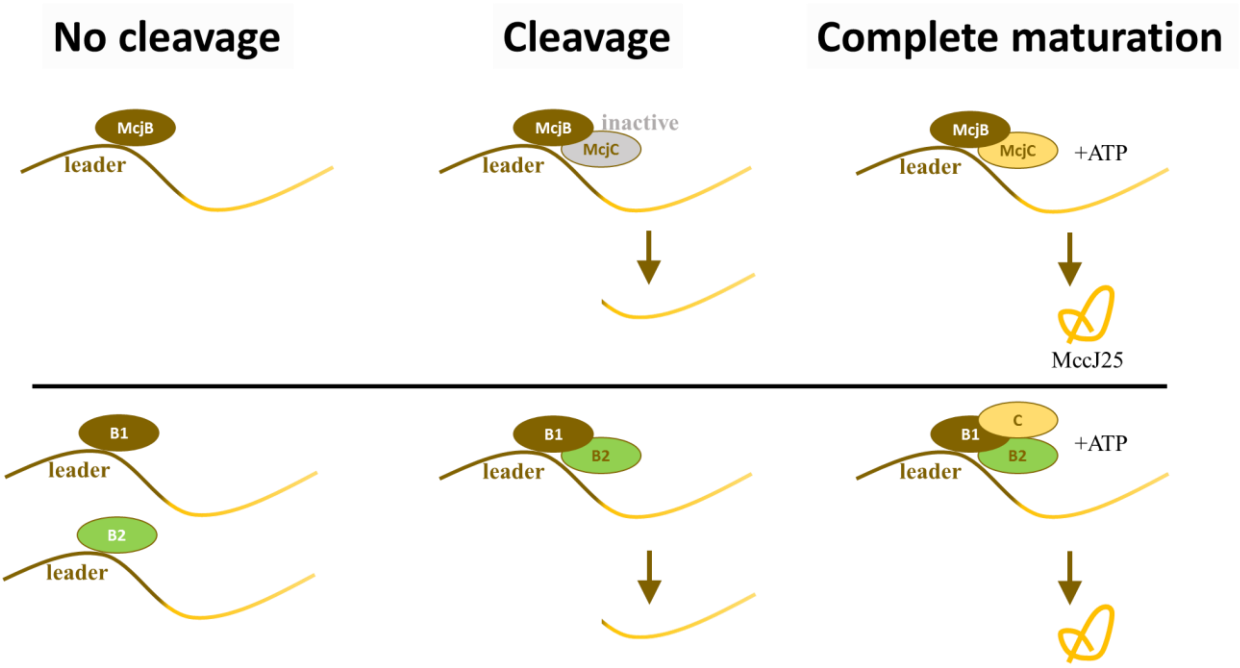


***Figure 12*. A scheme of *in vitro* processing of lasso-precursor peptides with various combinations of precursor peptides, maturation enzymes and ATP**. The top panel represents data relevant to MccJ25 biosynthetic system with fused B-protein. The bottom panel shows the scheme for separated B1 and B2. *Adapted from refs*. [130], [131].

In brief, the lasso peptide maturation occurs in equivalent concentrations of precursor peptide and their maturation enzymes indicating the formation of a stable synthetase [72]. Complete maturation of MccJ25 or leader peptide cleavage require the presence of both processing enzymes (McjB, McjC), ATP and Mg^{2+} ions, while paeninodin precursor peptide can be cleaved

with combination of B1 and B2 proteins in the absence of C-enzyme or ATP. However, no cleavage of paeninodin precursor peptide was observed in the absence of either B1 or B2 subunits.

In summary, it is still impossible to outline the detailed mechanism of lasso peptide biosynthetic machinery and they have to be clarified in the future. Only two moderately efficient *in vitro* lasso peptide synthesis systems were reported, for MccJ25 [130] and fusilassin [51], [72], and *in vivo* studies may be a promising way for further investigations.

1.5. Cluster architecture

The proteins involved in the lasso peptide biosynthesis are encoded by small clusters widespread throughout all bacterial domains as well as in archaea. Lasso peptide BGCs possess relatively similar architecture with varying number of auxiliary genes present in the same cluster. Mandatorily, the clusters contain one (or, in rare cases, several) genes coding for lasso precursor peptides (A-peptide) and three genes coding for essential processing proteins responsible for the lasso peptide formation: B1, B2 and C [49], [157]. These four essential genes form the so-called "core cluster".

Most of lasso peptide BGCs (ca. 62%) include additional genes coding for transport systems [49], [85], typically represented by ATP-binding cassette (ABC) transporters, which export mature peptides out of the producing cells [46], [49], [52], [128], [162].

Thus, the most studied to date BGC producing MccJ25 consists of three core genes *mcjABC* as well as a single transporter gene *mcjD* [128]. Since MccJ25 exhibits antibacterial activity [163], efficient export of mature peptide out of producing cells necessary for self-immunity [2], [4], [6], [11].

Various auxiliary genes found in lasso-clusters were predicted to code for extra modification enzymes, especially, various transferases such as: kinases (in 32% of BGCs), nucleotidyltransferases (26%), glycosyltransferases (14%), and acetyltransferases (7%) [49], [61],

[157], [164]. To date we know only one example of an extra modification enzyme encoded at a significant distance from the main cluster. Citrulline located in the middle of citrulassin A is converted from Arg9 by an unknown enzyme, which is present neither in the biosynthetic cluster nor in the surrounding 20 kb of DNA [49]. A general scheme of the lasso peptide BGC architecture can be represented in the RiPP-specific canonical form (Figure 13).

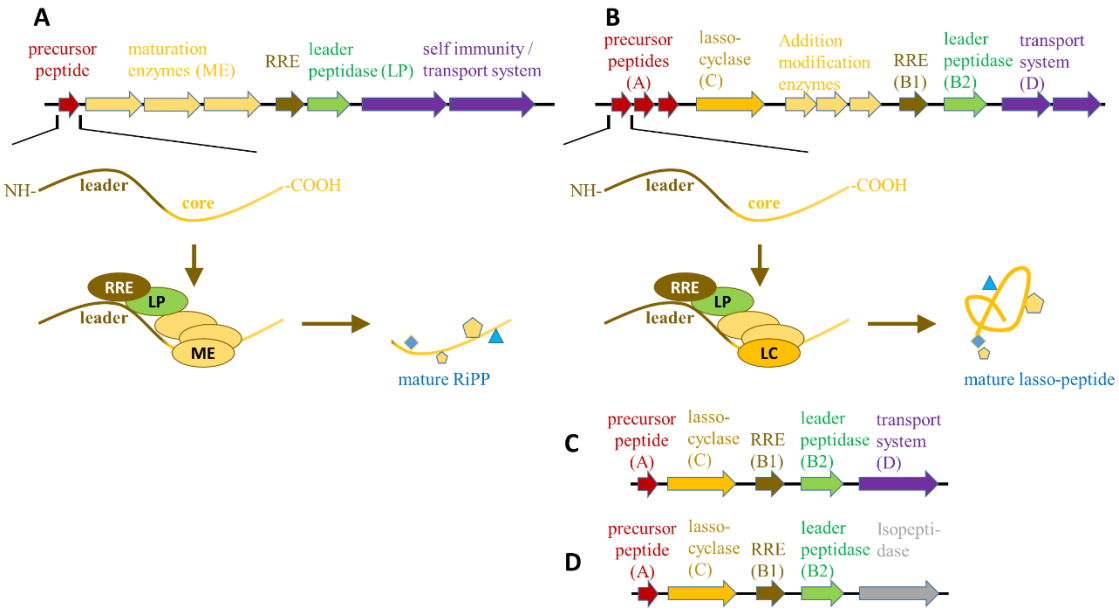


***Figure 13*. Comparison of cluster architectures and biosynthetic pathway of (A) RiPPs and (B) lasso peptides**. (C) and (D) represent schemes of typical lasso peptide "core clusters", and a cluster containing isopeptidase, respectively.

The presence of genes encoding a serine kinase strongly correlates with C-terminal Ser in the precursor peptide [49], [100], [109]. These clusters possess the *CAKB1B2D* organization and often contain additional genes. Its first reported representative was shown to be responsible for the production of lasso peptide paeninodin, which is phosphorylated with multiple phosphate groups at its C-terminus [100], [109]. The product of *padeK* gene located between the genes of precursor peptide (*padeA*) and RRE (*padeB1*) is the tailoring kinase. *In vitro* studies have clearly demonstrated that ThcoK, a homolog of PadeK is capable of phosphorylating paeninodin precursor peptide on the side chain of the C-terminal Ser [100], [109]. The same study also provided

evidence for multiple phosphorylation (up to three phosphate groups) by ThcoK of precursor peptide ThcoA *in vitro*.

Interestingly, a special type of cluster was found in the genomes of a subset *Proteobacteria*. It is devoid of any transporter genes or additional post-translational modifications enzymes genes, but carries a highly conserved gene encoding an isopeptidase (IsoP), a unique member of the cysteine protease family [42], [49], [157]. IsoP has been shown to specifically hydrolyze the isopeptide bond of the macrolactam ring converting lasso peptide to a linear form (Figure 14) [111], [165], [166]. The astexin IsoP was proved to be highly selective acting only against cognate lasso peptide [111]. Such type of activity contributes to self-immunity of producing cells [33].

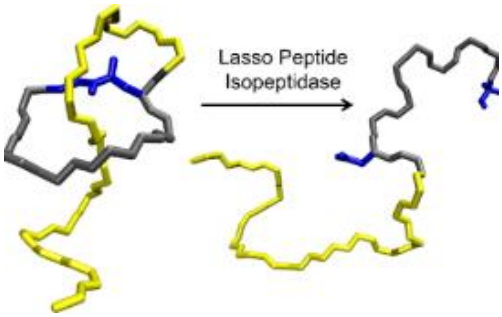


Figure 14. **Isopeptidases encoded by lasso-biosynthetic clusters catalyze proteolytic cleavage of isopeptide bond in lasso-macrocycles resulting in lasso peptide linearization.** *Reprinted from* [111].

Curiously, a number of recently isolated and characterized lasso peptides produced by clusters possessing an IsoP instead of an ABC transporter (including astexin-1 [62], sphingopyxin I [61], and caulosegnin I [64]), apparently lack antimicrobial activity, yet they might possibly have other unknown functions.

1.6. Genome mining approaches

A number of genome mining approaches to identify lasso peptide BGCs have been proposed. Until recently, the mainstream methods were either precursor-centric [42], [104], [157] or protein homology-based [49], [52], [61]–[64], [167], [168]. The precursor-centric approach

starts with searching for short open reading frames (ORF) encoding putative precursor peptides based on their conserved residues. The next step is screening for the rest of essential genes B1/B2 / B and C [157] nearby. This method allowed to identify a number of lasso peptide BGCs at the dawn of lasso peptide studies. However, searching for small ORFs fitting the soft criteria for the lasso peptide precursors usually results in a large number of putative candidates and low hit rate [42]. Moreover, this method may overlook non-canonical peptide precursors.

The protein homology-based methods demonstrate much greater selectivity and allow one to manual check each precursor peptide candidate. Mostly B protein-centric [61], [88], [100], [157], but also C-centric [49] and even kinase- [100], [109] and isopeptidase-centric [169] approaches were developed. The "B protein-centric" search of homologs of ATP-dependent cysteine protease (B-protein) is based on the Position-Specific Iterated (PSI)-BLAST iterative approach followed by case-by case search for a gene coding for a suitable putative precursor. Due to the uniqueness of B-proteins for the lasso peptide family, this method is widely used for seeking new lasso BGCs.

To date the most powerful bioinformatics tool to identify new lasso peptide BGCs is RODEO (Rapid ORF Description and Evaluation Online). It has been developed and implemented by the research group of Douglas Mitchell. Using the C-centric approach they found more than 1300 new clusters, while more recent paper reported about 3000 novel lasso BGCs [49], [51]. Such a large number reveals the general patterns of lasso clusters distribution throughout the bacterial domains, showing their presence outside *Proteo-* or *Actinobacteria* from which most previously characterized clusters came.

1.7. Targets, mechanisms of action

Approximately one third of all known lasso peptides possess antimicrobial activity (astexin-1 [62], capistruin [67], chaxapeptin [95], lariatin A, B [97], lassomycin [73], microcin J25 [170], propeptin [101], siamycin I, II, III, [81], [171], [172] sviceucin [168], streptomonomicin

[105], klebsidin [87], acinetodin [87], actinokineosin [89], achromosin [87], sphaericin [62], ubonodin [79]). Some lasso peptides have diverse other bioactivities, such as potentiating antifungal activity (humidimycin strongly increases caspofungin efficacy [80], antiviral (siamycins I, II, III), receptor antagonistic (BI-32169 [77], RES-701-1 [173] and RES-701-3 [103]), or enzyme inhibitory activities (propeptin, siamycin I, anantin [91], sungsanpin [107]).

MccJ25, capistruin, citrocin, ubonodin, acinetodin, and klebsidin, antimicrobial peptides from *Proteobacteria*, inhibit Gram-negative bacterial RNAP. Despite the lack of strong sequence similarity of these peptides, they bind to the same target, the secondary channel of RNA polymerase blocking the access of NTPs to the catalytic center [44], [88], [174], [175] Interestingly, *in vitro* citrocin is ~100-fold more potent RNAP inhibitor than microcin J25, but has a higher minimal inhibition concentration in bioactivity tests [96]. It may indicate that citrocin, which crosses the outer membrane through Ton and Tol-Pal, penetrates target cells much less efficiently than MccJ25, which uses the FhuA receptor and TonB and SbmA transporters. Lariatin A displays anti-mycobacterial activity [176], [177]. Lassomycin targets the ATP-dependent protease ClpC1P1P2 [73], propeptin inhibits prolyl endopeptidase [101], and siamycin-I targets cell wall biosynthesis of Gram-positive bacteria by binding to lipid II [178], [179].

Surprisingly, some lasso peptides have eukaryotic cell-associated activity. Siamycin I and II exhibit HIV inhibition activity preventing viral subunits from merging with eukaryotic cells [84], [180]. Sungsanpin [107] and ulleungdin [108] demonstrate the lung cancer cell (line A549) inhibitory activities. BI-32169 possesses human glucagon receptor inhibitory activity [77]. Anantin was shown to be an antagonist of the atrial natriuretic factor by competitive binding to its receptor in bovine adrenal cortex [91]. RES-701-1 [173] and RES-701-3 [103] are antagonists of human endothelin type B receptor.

As most lasso peptide clusters encode export pumps, they must be functioning outside the producer cells. Presumably, some lasso peptide can signal to other bacteria of the same or different

species. Most effects associated with the non-bacterial targets were found via a systematic screening and might merely be coincidental, whereas the antibacterial activities are likely to be native functions.

1.8. Lasso peptides as an efficient scaffold for molecular grafting

Unmodified linear peptides are generally unstable *in vivo*, and this restricts their harnessing as drugs despite their high selectivity and potency. Lasso peptides possess a unique three-dimensional structure providing them with extraordinary stability and making them an attractive scaffold for the development of non-natural bioactive compounds [16], [54], [113], [118], [181], [182].

Presently, biotechnological approach is the easiest and fastest way to obtain lasso peptides. Molecular genetic engineering proved to be an effective method for modifying the core part of the precursor peptide without preventing recognition and processing by the biosynthesis machinery *in vivo* [130], [150]. Validated and predicted lasso peptides demonstrate remarkable sequence diversity which makes them promising candidates for creating novel unnatural lasso peptide analogs enriched with desirable features [49].

2. Project objectives

Despite significant advances in the field of molecular synthesis, numerous compounds of natural origin are still out of grasp for the total chemical synthesis – especially the compounds with complex topology. Being one of the great examples of such scaffolds lasso peptides are obtainable via biosynthetic means only. However, biosynthetic pathways are not understood well enough, which is also complicated by the fact that the full set of lasso maturation enzymes is difficult to isolate due to solubility and low yield problems. State-of-the-art robust approaches to investigating biochemical machinery of lasso peptides thus presently mainly rely on *in vivo* studies. We have bioinformatically found a novel *Firmicutes* lasso peptide biosynthetic gene cluster and confirmed that it produces a lasso peptide we called pseudomycoidin. It is only the second example of a lasso peptide isolated from this phylum so far. The cluster possesses an unusual architecture due to the presence of two additional genes presumably responsible for extra post-translational modifications of the lasso peptide. The main goal of this project was to investigate the biosynthetic pathways of lasso peptides using pseudomycoidin as a model.

The objectives of this project were as follows:

- To develop a robust production system *in vivo* for pseudomycoidin and its mutational variants.
- To determine the structure of pseudomycoidin.
- To characterize pseudomycoidin variants produced by incomplete clusters.
- To isolate processing enzymes and assess the mode of their action.
- To check whether processing enzymes are able to function *in vitro*.
- To investigate the tolerance of processing enzymes to substrate alterations.
- To obtain different variants of pseudomycoidin to characterize the tolerance of the biosynthetic machinery to aminoacid substitutions.

3. Materials and methods

3.1. Bioinformatics analysis

The bioinformatic search for the new lasso-cluster in the *Firmicutes* (taxid:1239) genomes was carried out against non-redundant GenBank CDS translations (excluding models (XM/XP) and uncultured/environmental sample sequences). The web-based Position-Specific Iterated Basic Local Alignment Search Tool (PSI-BLAST) was applied using C protein from paeninodin cluster (PadeC) as a bite [100]. The number of target sequences (hitlist size) was set to 500; the matrix for scoring parameters was BLOSUM62; expected value threshold set up to 10. The number of identified C proteins derived from other than paeninodin origin Paenibacillaceae family were then checked manually to search the nearby presence of genes coding for B2 protein homologue as well as a leader peptide. As a result, the cluster from *Bacillus pseudomycoides* was chosen for future research.

3.2. Bacterial strains and growth conditions

Bacillus pseudomycoides DSM 12442 strain was purchased from the German Collection of Microorganisms and Cell Cultures (DSMZ). *Escherichia coli* DH5a was used as a general host for cloning and BL21 (DE3) cells were applied for the protein production as well as a heterologous host for the *psm* cluster expression. For the MALDI MS analysis of the lasso peptide production *E. coli* BL21(DE3) cells harboring corresponding sets of plasmids were grown at 30°C on M9 minimal agar (48 mM Na_2HPO_4, 22 mM KH_2PO_4, 9 mM NaCl, 19 mM NH_4Cl, 0.1 mM $CaCl_2$, 2 mM $MgSO_4$, 1.5% agar) supplemented with 0.2% lactose and 0.2% arabinose as a sole carbohydrate source, 0.01% thiamine and appropriate antibiotics (50 μg/ml for ampicillin, 50 μg/l kanamycin, 34 μg/ml chloramphenicol). For the MALDI-MS analysis of the pseudomycoidin production in the *B. subtilis* 168 cells harboring the *psm* genes under inducible promoters were grown at 37°C on M9 minimal agar supplemented with 0.4% glycerol, 50 μg/ml L-Tryptophan, 5 μg/ml for erythromycin, and 10 μg/l chloramphenicol. For protein purification, *E. coli* BL21 (DE3)

cells were grown at 37°C in 2YT medium containing 50 μg/ml kanamycin to OD_{600} of ~0.6. Protein production was induced with 1 mM isopropyl-β-D-thiogalactopyranoside (IPTG), followed by growth at 18°C for 16 h.

For the MALDI-MS analysis of production of the pseudomycoidin derivatives a single obtained colony carrying corresponding plasmids was transferred to 10 ml of LB medium (suppl. with 100 μg/ml ampicillin and 50 μg/ml kanamycin) and grown at 37°C overnight. Then 100 μl of the overnight cultures were used to inoculate 10ml of fresh M9 minimal medium (48 mM Na_2HPO_4, 22 mM KH_2PO_4, 9 mM NaCl, 19 mM NH_4Cl, 0.1 mM $CaCl_2$, 2 mM $MgSO_4$, 0.4% glucose or glycerol,) supplemented with 0.4% glycerol as a carbon source, 10 μg/ml thiamine, 100 μg/ml ampicillin, 50 μg/ml kanamycin and 0.2 μM and cultivated in the presence of non-hydrolyzable inductor of translation 0.2 μM IPTG for 6 hours at 30°C. The culture supernatants were subjected to the MALDI-MS analysis.

To produce the lasso peptide for the NMR analysis a single colony with corresponding sets of plasmids was transferred to 40 ml of LB medium (suppl. with 100 μg/ml ampicillin and 50 μg/ml kanamycin) and grown at 37°C overnight. The overnight cultures were used to inoculate 4 liters of fresh LB medium (suppl. with 0.4% glucose, 100 μg/ml ampicillin and 50 μg/ml kanamycin) for growth at 37°C aerobically up to OD_{600} = 0.8. The cells were then harvested by centrifugation and transferred into 4 liters of fresh M9 minimal medium supplemented with 0.4% glycerol, 10 μg/ml thiamine, 100 μg/ml ampicillin, 50 μg/ml kanamycin and 0.2 μM IPTG. The induced cells were incubated with shaking for 10 hours at 30°C followed by centrifugation for 20 minutes at 10,000 g at room temperature. The obtained conditioned supernatant was used for the pseudomycoidins extraction.

For the His_6-PsmB1 and PsmB2 fused with maltose binding protein (MBP) purification, *E. coli* BL21 (DE3) cells harboring pET28_*His-psmB1* or pET28_*MBP-psmB2* vectors,

respectively, were grown at 37°C in 2YT medium containing 50 μg/ml kanamycin to OD_{600} of ~0.6. Protein production was induced with 1 mM IPTG followed by growth at 18°C for 16 h.

All reagents, media and enzymes used in this study were obtained from standard commercial sources.

3.3. Molecular Cloning and Protein production

All oligonucleotides used in this work are listed in the Table S2, Supporting Information. The maps of created plasmids are also present in the Supporting Information section. Oligonucleotides and DNA sequencing verification of all the constructed plasmids were made by Evrogen (Moscow, Russia). The *psm* cluster genes for molecular cloning were produced by polymerase chain reaction (PCR) amplification with Phusion High-Fidelity DNA-Polymerase (Thermo Fisher Scientific).

The fragment of the *psm* cluster spanning the *psmC* and *psmA* genes was PCR-amplified and inserted into pCA24 plasmid [183] harboring T5*lac* promoter using the Gibson assembly kit (NEB, USA). PCR fragments containing either *psmKB1B2ND* or *psmB1B2ND* gene sets were inserted into pCA24 compatible pBAD30 plasmid between the SacI and SalI sites. The *psmN* and *psmB1B2* deletions were introduced into pBAD-*psmKB1B2ND* plasmid containing arabinose-inducible promoter using the Gibson assembly kit and the corresponding primer set. For the ectopic complementation experiment, the *psmN* and *psmK* genes were cloned into the pCA24 and pBAD30 compatible pCOLADuet-1 plasmid between BamHI and XhoI to get T7*lac* promoter driven expression of the His_6-tagged proteins. The overlap–extension PCR [184] was carried out to create the D134A and D136A double mutant of the PsmN in the pCOLA-*psmN* plasmid. The D158A and D159A double mutant of PsmK was obtained using the same technique and pCOLA-*psmK* plasmid as a template.

The combinations of the compatible Duet plasmids (Novagen) containing T7*lac* inducible promoters were used to produce the lasso peptide for the lasso topology study. The

pseudomycoidin production system was designed based on the combination of two commercially available co-expression replicative vectors: pCOLADuet-1 and pETDuet-1 (Novagen) each of them harboring two multiple cloning sites (MCS), both under the control of the individual T7*lac* promoter. Each of the genes A, B1, C and D was cloned into a distinct MCS, and B2 was introduced just downstream of B1 with its own ribosome binding site resulting in pETDuet_*psmCA* and pCOLADuet_*psmB1B2D* vectors carrying the *psmCA* and *psmB1B2D* genes respectively.

Gene *psmA* and its 14 substituted variants, *psmB1*, *psmB2*, *psmC* and *psmD* (sequences are listed in Table S1, Supporting Information) were PCR-amplified from the *B. pseudomycoides* genome with Phusion High-Fidelity DNA-Polymerase using oligonucleotide primer pairs listed in Table S2, Supporting Information. The final PCR products were flanked by the endonuclease restriction sites suitable for further cloning procedure. Gene $psmA^C$ was obtained by hybridization of $psmA^C$_NcoI_F & $psmA^C$_SacI_R primers to each other forming two single-stranded overhangs related to the NcoI & SacI endonuclease restriction sites. The *psmB1*-containing PCR product additionally carried one more ribosome binding site following the psmB1 gene.

PCR product containing genes *psmA* and $psmA^C$ was cloned into MCS(I) of pETDuet-1 using NcoI & SacI restriction enzymes resulting in the pETDuet_*psmA* and the pETDuet_$psmA^C$ plasmids respectively. Next, psmC gene was cloned with *NdeI*&*XhoI* endonucleases into MCS(II) of pETDuet_*psmA* and pETDuet_$psmA^C$ forming pETDuet_*psmAC* and pETDuet_$psmA^C C$ expression vectors. *PsmD* gene was cloned into MCS(II) of pCOLADuet-1 with digestion by *NdeI*&*XhoI* endonucleases thus producing pCOLADuet_*psmD*. Next, the *psmB1* containing PCR product was introduced into MCS(I) of pCOLADuet_*psmD* with *NcoI* &*BamHI* restriction enzymes yielding the pCOLADuet_*psmB1D* plasmid. Finally, *psmB2* was similarly put into MCS(I) of pCOLADuet_*psmB1D* with endonucleases *BamHI* &*SalI* resulting in pCOLADuet_*psmB1B2D*. Apart from two original ribosome binding sites this co-expression system contained an additional one positioned 7 nucleotides before the start codon of *psmB2* gene.

Thus, the obtained vectors maintained the independent translation of every psm biosynthetic protein under its own RBS.

To get T7*lac* promoter-driven expression of the His_6-tagged PsmB1 and MBP-PsmB2 proteins the *psmB1* gene was subcloned into pET28 plasmid (Novagen) between NdeI and XhoI sites and *psmB2* gene was introduced into pET28*_MBP* plasmid [185] with BamHI and XhoI endonucleases resulting in pET28*_His-psmB1* and pET28*_MBP-psmB2* vectors, respectively.

To create *psmCAD* expressing *B. subtilis* strain the *psmCA* genes the PCR fragment which contains *psmC* and *psmA* genes was inserted into pHT-01 shuttle vector (MoBiTec, GmbH) harboring inducible P*grac-lac* promoter between BamHI and SpeI. The expression cassette containing *psmCA* operon flanked by P*grac-lac* promoter and the transcription terminator was PCR-amplified using pHT-*psmCA* vector template and inserted into EcoRI site of the pBS2E integrating plasmid [186]. The resulting *pBS2E-psmCA* was transformed into *B. subtilis* 168 strain as described early [187]. Integration-positive clones were selected on erythromycin-containing LB plates and checked by PCR for the presence of insert. The *psmD* gene was cloned into the pHT-01 vector between BamHI and SpeI. The pHT-*psmD* vector was transformed into *B. subtilis* 168 *amy::psmCA* strain and selected on LB agar supplemented with erythromycin and chloramphenicol.

Proteins fused to His and MBP N-terminal affinity tags were purified using Talon CellThru Co2+-chelating resin (Takara-Clontech) and Amylose (NEB, USA) resin, respectively, as described [188]. Purity was checked by SDS PAGE and the concentration was measured using Bradford assay (BioRad, USA).

3.4. Pseudomycoidin and its derivatives production

Various combinations of *psm* genes on compatible pETDuet-1/pCOLADuet-1 expression vectors were transformed into *E. coli* BL21DE3 electrocompetent cells. For production, a single transformed colony was selected on LB-agar plates with 100 μg/ml ampicillin and 50 μg/ml

kanamycin. 1 ml of induced culture supernatants of *E. coli* cells harboring corresponding sets of *psm* genes were analyzed with analytical C18 column and eluted with a linear gradient of aqueous acetonitrile from 19 to 25% supplemented with 0.1% trifluoroacetic acid (TFA), using ultra violet (UV) detection at 210 nm. Elution fractions were analyzed by MALDI-MS.

3.5. Chromatographic purification

Cells harboring one of three sets of genes were induced in 8 liters of fresh M9 liquid medium as described above. Pseudomycoidins were extracted from the culture supernatant in two steps. Firstly, the solid-phase extraction on the C18 gravity column (Sep-Pak C18, 55-105 μm particle size) was applied. The column was washed with 7% acetonitrile/0.1% TFA. Pseudomycoidin-containing fraction was eluted with 30% acetonitrile/0.1% TFA, dried in a rotary vacuum evaporator and dissolved in MQ water. Secondly, reversed-phase high-performance liquid chromatography (RP-HPLC) on the C18 column (Jupiter, 5 μm particle size C18 300 Å, LC Column 250 x 4.6 mm) applying gradient elution from 19 to 25% aqueous ACN/ 0.1% TFA in 40 min using UV detection at 210 nm was performed. Collected fractions were confirmed by MALDI MS. Fractions containing lasso peptides were lyophilized to dryness and stored at -20°C until further studies. These cultivation conditions followed by two-step purification resulted in about 0.5 mg of pure peptide per liter of culture supernatant.

3.6. Biochemical reactions

For the proteolytic activity test of PsmB1 and PsmB2 enzymes, 25 μM synthetic PsmA peptide (Genscript, USA) was incubated in the reaction buffer A (20 mM Tris HCl, pH 7.5, 5 mM $MgSO_4$, 100 mM NaCl, 1 mM DTT, 1 mM ATP) in the presence of either 1 μM His_6-PsmB1, or 1 μM MBP-PsmB2 proteins, or the combination of thereof at 30°C for 16 h.

Endoproteolytic digestion of the pseudomycoidin was performed with 0.1μg/μl GluC from *Staphylococcus aureus* V8 (Sigma-Aldrich) in 50 mM ammonium bicarbonate, pH 7.8 at 37°C for 3h. Carboxypeptidase Y treatment was carried out in 50 mM MES and 1 mM $CaCl_2$

buffer with 0.5 U carboxypeptidase Y at pH 6.75 for 4 h at 25°C [62], [88]. Chymotrypsin assays were performed in 100 mM Tris-HCl and 10 mM $CaCl_2$ with 0.5 μg of chymotrypsin at pH 8.0 for 12 h at 25°C [64].

3.7. Mass spectrometry analysis

Ion spectra were recorded on an AB Sciex TOF/TOF 5800 MALDI mass spectrometer on Reflectron mode. The instrument was calibrated with the mass standard kit for calibration of AB Sciex TOF/TOF instruments (AB Sciex). Samples (0.5 μl) were spotted on a steel plate air-dried at room temperature and covered with 0.5 μl of α-cyano-4-hydroxycinnamic acid matrix solution (AB Sciex) and again air-dried at room temperature. The accuracy of the mass peak measurements was 30 ppm.

MALDI-TOF MS analysis was performed on UltrafleXetreme MALDI-TOF/TOF mass spectrometer (Bruker Daltonik, Germany) equipped with Nd laser. Samples for analysis were prepared as described earlier [189] Due to high instability of the phosphorylated forms of pseudomycoidin the MH+ molecular ions were measured in linear mode; the accuracy of average mass peak measurement was within 1 Da. Spectra of fragmentation were obtained in LIFT mode, the accuracy of daughter ions measurement was within 1 Da range. Mass-spectra were processed with the use of FlexAnalysis 3.2 software (Bruker Daltonik, Germany) and analyzed manually.

High-resolution mass spectra were recorded on a Fourier transform ion cyclotron resonance mass spectrometer (Varian 902-MS) equipped with a 9.4-T magnet in positive MALDI mode. The instrument was calibrated with the ProteoMass peptide MALDI-MS calibration kit (Sigma- Aldrich). The accuracy of the mass peak measurements was reached 2.5 ppm. Samples (0.5 μl) were spotted on a steel plate with 0.5 μl of a 2, 5-dihydroxybenzoic acid matrix (Sigma-Aldrich) and air-dried at room temperature.

4. Results and discussion

4.1. Bioinformatics studies

Since most lasso peptides characterized to date are derived from *Proteo-* or *Actinobacteria* [49], [57], lasso peptides from other bacterial phyla were of great interest to us. The non-redundant protein database was used for the PSI-BLAST search for the homologs of B2 protein from the lariatin cluster in the *Firmicutes* phylum. As a result, a likely lasso peptide biosynthesis cluster in the *Bacillus pseudomycoides* DSM 12442 genome (Figure 15.A) was identified. The cluster, which we will refer to as *psm*, is similar to the previously characterized *Paenibacillus dendritiformis* C454 lasso peptide paeninodin *pade* cluster.

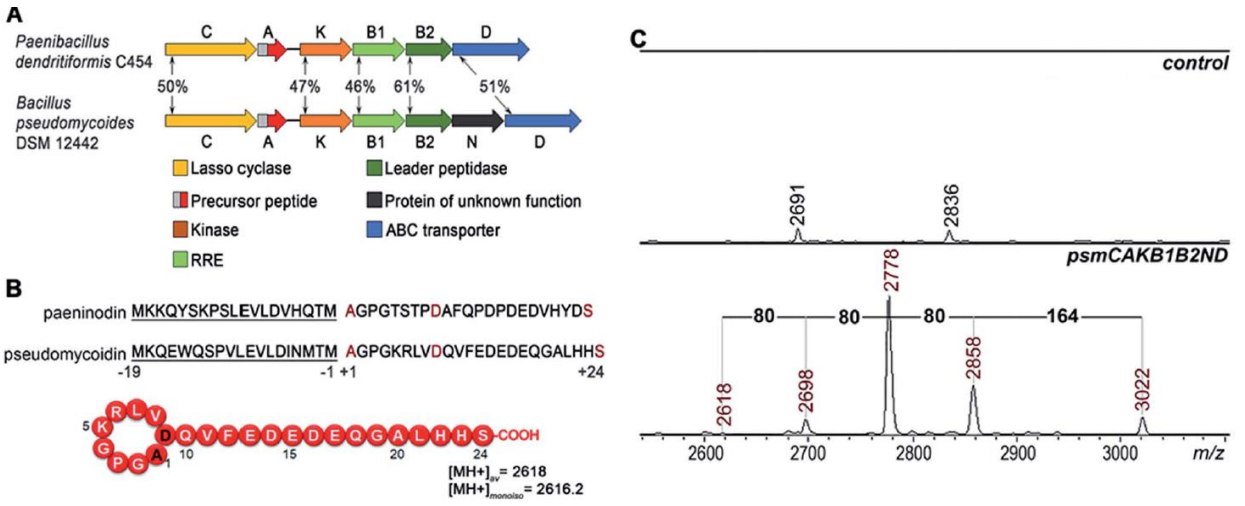


***Figure 15*. Heterologous expression of the *B. pseudomycoides* DSM 12442 lasso peptide *psm* gene cluster in *E. coli*.** (A) Comparison of lasso peptide gene clusters from *P. dendritiformis* C454 and *B. pseudomycoides* DSM 12442. Genes are indicated by arrows (letters indicate gene names, i.e., "B" stands for either "padeB" or "psmB"), and homologous genes are indicated by same colors. Numbers indicate the percentage of identity between homologous protein sequences. Known or predicted gene product functions are listed at the bottom. RRE – RiPP precursor peptide recognition element. (B) At the top, the precursor of *P. dendritiformis* C454 paeninodin is aligned with the corrected *B. pseudomycoides* DSM 12442 pseudomycoidin precursor sequence. Leader parts are underlined. In the core parts, the N-terminal alanine and the inner aspartate that form the ring and C-terminal serine that becomes phosphorylated in paeninodin are indicated with red-color font. The predicted structure of pseudomycoidin is shown at the bottom with average and monoisotopic [MH+] of a molecule with the structure shown as indicated. (C) MALDI-MS spectra of *E. coli* cells harboring control (upper panel) and *psm* cluster (lower panel) expression plasmids. Average m/z values of different mass-peaks are highlighted with red-color font.

The *B. pseudomycoides* DSM 12442 cluster, contains a *psmC* gene coding for a homolog of known lasso peptide macrolactam synthetase, a putative precursor peptide gene (*psmA*), a *psmK* gene, encoding a histidine-containing protein kinase family protein (HprK), two genes coding for split B enzyme homologs (*psmB1* and *psmB2*), and the ABC-type transporter gene *psmD* (gene sequences are provided in Table S1). The *psm* cluster also contains an additional gene, *psmN*, located between *psmB2* and *psmD*. Such a gene is absent in the *P. dendritiformis* C454 cluster. The product of *psmN* is annotated as a putative uncharacterized NTP transferase, but a PSI-BLAST search against non-redundant protein sequence database revealed the presence of a PsmN homolog (66.5% identity) encoded in *P. dendritiformis* C454 genome outside of paeninodin BGC. Similar genes have been found at the same location in several other putative lasso peptide clusters [49], [100], [109].

The *psmB1* gene (GenBank accession: WP_000095761.1) codes for a 99 amino acid chain (ClustalW alignment gives 66.7% identity in amino acid sequence with PadeB1, E-value 2.90e-16) containing a PqqD-like domain between amino acids 29-93 that should interact with the precursor peptide and direct subsequent enzymatic modifications. The *psmB2* gene (GenBank accession: WP_018764971.1) codes for predicted 136 aminoacid leader peptidase PsmB2 (73.1% identical to PadeB2, E-value 4.90e-21) with a transglutaminase core covering the interval of amino acids from 19 to 128. It should cleave the leader part of precursor peptide. The *psmC* gene (GenBank accession: WP_006096609.1) codes for 645 aminoacid long predicted lasso-cyclase (71.4% identity with PadeC, E-value 1.19e-21) and is homologous to ATP-dependent asparagine synthetase B in the region from amino acids 246 to 571. It should catalyze the isopeptide bond formation leading to a macrolactam ring. The *psmD* gene (GenBank accession: EEM14408) encodes a 575 aminoacid efflux pump (72.8% identity with PadeD, E-value 4.47e-109) with similarities to ABC-type between amino acids 44-571. The *psmK* gene (GenBank accession: WP_006096607.1) encodes a 309 aminoacid protein (71.4% identity with PadeK, E-value 9.01e-07) containing a serine kinase domain between amino acids 84-181. Finally, the *psmN* gene

(GenBank accession: WP_006096606) located between *psmB2* and *psmD*, encodes a 395 aminoacid protein with similarity to an uncharacterized nucleotidyltransferase (E-value 9.91e-58) between residues 41 to 288. The *psmA* gene (GenBank accession: EEM14413. 1) is located downstream of *psmC* and encodes a 43 aminoacid predicted precursor peptide consisting of a 19-residue leader part and a 24-residue core. The leader peptide $PsmA_L$ MKQEWQSPVLEVLDINMTM has a conserved threonine at the penultimate position and is similar with the leader peptide of paeninodin (78.9%).

Based on similarity to the *pade* cluster, one could expect that the *psm* cluster products may include lasso peptides with or without terminal phosphorylation and, possibly, with some additional unknown modifications introduced by PsmN. The predicted sequence of PsmA ends with an asparagine residue as opposed to C-terminal serine of paeninodin precursor PadeA, which is subject to phosphorylation [90], [100].

We resequenced the *psmA* gene and found a single nucleotide difference with the published sequence at the last codon, which is reported to be AAC coding for asparagine. The sequence obtained by us has an A**G**C coding for a serine residue at the same site. The correct sequence thus codes for a PsmA peptide with a terminal serine. Thus, the core part of PsmA consists of 24 residues AGPGKRLVDQVFEDEDEQGALHHS and likely undergoes post-translational modifications forming a macrocycle due to the isopeptide bond between the amino group of Ala^1 and side chain carboxyl group of Asp^9 and may also be phosphorylated at the C-terminus.

The expected average [MH+] of the lasso peptide formed from PsmA, if one assumes processing at the leader-core junction based on similarity to paeninodin, is 2618 (Figure 15.B).

4.2. Heterologous expression of various configurations of *psm* cluster

When *B. pseudomycoides* DSM 12442 cells and the cultured medium were analyzed by MALDI-MS, no peaks matching expected lasso peptide cluster products – with or without terminal

phosphates – were observed. When natural producer strains do not lead to lasso peptides sought under standard culturing conditions, heterologous expression is typically used to isolate them [42], [61], [63], [64]. The small size of the *psm* cluster makes it convenient for heterologous expression in *E. coli,* a model organisms successfully used for production of a number of other lasso peptides [167]. We therefore introduced the *psm* cluster in a heterologous *E. coli* host.

The *psmCA* genes were cloned on an expression plasmid under the control of a T5-*lac* promoter, while the remaining genes were cloned on a compatible plasmid under the control of an arabinose-inducible *Para* promoter. As can be seen in Figure 15.C five difference mass peaks were present in the potential area of interest of the mass spectrum of cells carrying plasmids with *psm* genes compared to control (cells carrying empty plasmid vectors). One peak, with average m/z = 2618, matched the expected mass of a lasso peptide formed from the PsmA precursor.

Three peaks, with the average m/z = 2698, 2778, and 2858 differed from the m/z = 2618 peak and from each other by 80 atomic unit increments, that is a value matching a phosphate group. The fifth difference peak had an m/z value of 3022 and could not be matched to any known lasso peptide modification.

The obtained data led us to conclude that the lasso peptide from *B. pseudomycoides* can be successfully synthesized, cyclized and exported out of heterologous producing cells *E. coli*. Having established a robust system of novel lasso peptide production in *E. coli*, we proceeded to study the biosynthetic capabilities of the cluster and constructed several of its truncated variants.

4.2.1. Mutational analysis of the *psm* cluster: deletion of *psmK*

A set of expression plasmids carrying deletions of *psm* genes was prepared and introduced in *E. coli* and the products were analyzed. Deletion of *psmK* resulted in accumulation of the m/z = 2618 peak only (Figure 16). The amplitude of its signal was strongly increased compared to the mass spectrum of cells with full clusters.

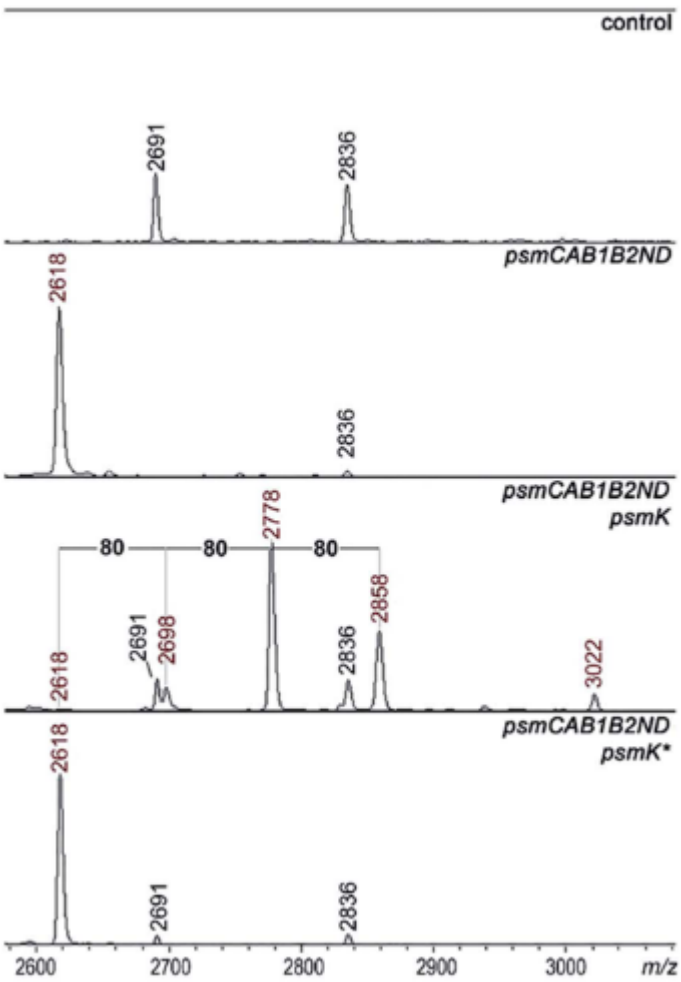


Figure 16 **Identification of *psm* cluster products produced in the presence of the PsmK kinase.** MALDI-MS spectra of *E. coli* cells harboring control empty vector plasmids (upper panel), plasmids expressing the *psm* cluster without the psmK gene (*psmCAB1B2ND*) or cells expressing the mutated cluster and harboring an additional plasmid expressing wild-type or D158A D159A double mutant psmK (psmK*). Mass-peaks that are common to both control and experimental cells are indicated by black-color font. The average m/z values of mass-peaks of *psm* cluster products are highlighted with red-color font.

The MS-MS fragmentation spectrum of this peak allowed us to validate the predicted amino acid sequence of the C-terminal part of PsmA core peptide as well as the N-terminal macrocycle formed by Ala^{1} - Asp^{9} residues with m/z = 877 (AGPGKRLVD) that is stable against MS fragmentation. (Figure 17, upper panel). We will refer to this compound as pseudomycoidin.

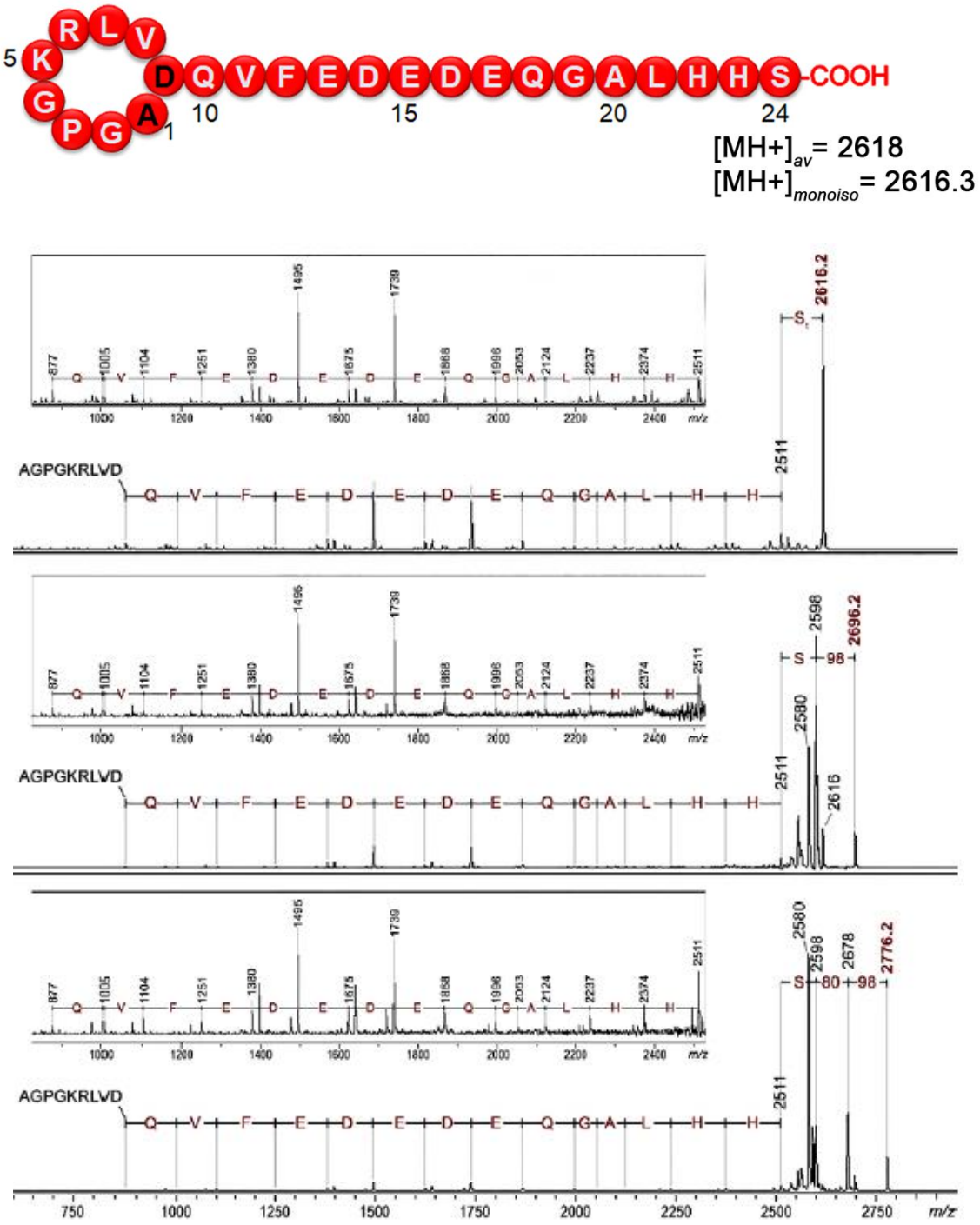


Figure 17. **MALDI MS-MS analysis of unphosphorylated and phosphorylated forms of pseudomycoidin**. Upper panel – the fragmentation spectrum of unphosphorylated pseudomycoidin (average m/z = 2618, monoisotopic m/z = 2616.2). The peaks corresponding to the C-terminal fragments of the lasso peptide are marked. Low-intensity daughter peaks are magnified in the insert. The daughter peak with m/z = 877 matches the macrocycle formed from the N-terminal AGPGKRLVD peptide. Middle and lower panels – fragmentation spectra of monophosphorylated (average m/z = 2698, monoisotopic m/z = 2696.2) and diphosphorylated (average m/z = 2778, monoisotopic m/z = 2776.2) pseudomycoidin. Fragment analysis shows the presence of phosphate group(s) attached to the C-terminal serine residue. Mass differences of 98 and 80 Da match H_3PO_4 and -HPO_3 groups, respectively.

The disappearance of m/z = 2698, 2778, and 2858 Da peaks in cells lacking *psmK* is consistent with them being, respectively, the products of mono-, di-, and triphosphorylation of

pseudomycoidin (Figure 16). When these cells were supplemented with a compatible *psmK* expression plasmid, the 2778 and 2858 peaks reappeared. MS-MS analysis of the m/z = 2698 and m/z = 2778 peaks indicates that the +80 Da orthophosphate and +160 Da pyrophosphate are located at the C-terminal serine residue of pseudomycoidin, consistent with the PadeK-dependent phosphorylation of paeninodin, for which the side chain of the terminal Ser was shown to be a substrate [100]. As expected, production of PsmK with D158A and D159A substitutions in the predicted metal binding site did not restore the phosphorylated pseudomycoidin production (Figure 16).

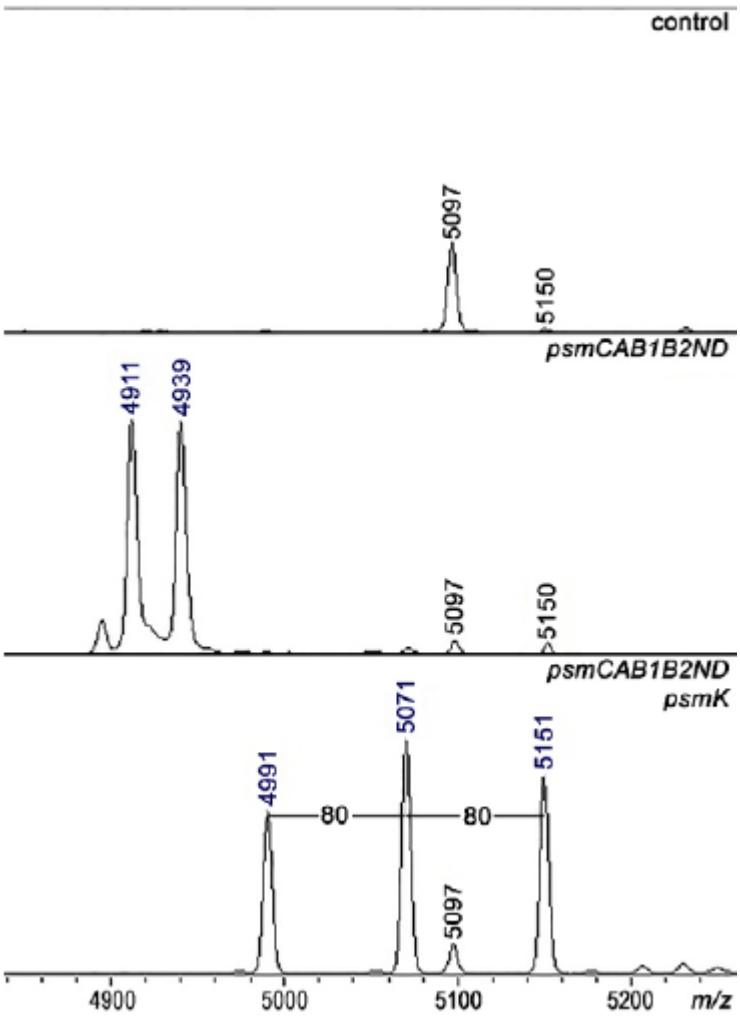


***Figure 18*. Production of phosphorylated linear precursor peptide PsmA.** MALDI-MS analysis of the *E. coli* BL21 (DE3) cells containing control plasmids (upper panel), plasmids expressing *psmCAB1B1ND* (middle panel), or plasmids expressing complete *psm* cluster (lower panel). Average *m/z* = 4911 and 4939 match the PsmA precursor peptide and its formylated form, respectively. Average *m/z* = 4991, 5071, and 5151 match mono-, di-, and triphosphorylated PsmA precursor peptide, respectively. The *m/z* values of mass-peaks specific for cells carrying the *psm* genes are marked with blue color font.

Inspection of mass-spectra of producing cells in the area corresponding to full-sized PsmA revealed multiple phosphorylated PsmA in cells carrying intact psmK. The result thus shows that PsmA precursors can be phosphorylated and phosphorylation may occur prior to cyclisation (Figure 18).

4.2.2. Mutational analysis of the *psm* cluster: deletion of *psmN*

Deletion of *psmN* gene led to disappearance of the m/z = 3022 peak while the other four pseudomycoidin peaks remained unchanged. When the *psmN* gene was introduced into a separate plasmid, the 3022 peak reappeared and became more prominent. In addition, new peaks with m/z = 2940 and m/z = 2860 were observed. The mass difference between monophosphorylated pseudomycoidin (m/z = 2698) and the m/z = 2860 peak is 162 Da. The difference between m/z = 2860 and m/z = 3022 peaks is also 162 Da. In addition, the difference between diphosphorylated pseudomycoidin (m/z = 2778) and the m/z = 2940 peak is also 162 Da. It thus appears that PsmN is required for installation of a 162 Da modification onto phosphorylated forms of pseudomycoidin.

Cellular extracts prepared from cultures expressing the *psm* cluster without *psmN* or the same cultures harboring the additional *psmN* expression plasmid were prepared, treated with calf intestinal alkaline phosphatase (CIAP) and analyzed by MS (Figure 19).

In extracts lacking PsmN, CIAP treatment resulted, as expected, in accumulation of unphosphorylated pseudomycoidin (m/z = 2618) and decreased signals from phosphorylated forms. In cells expressing *psmN*, the diphosphorylated pseudomycoidin peak disappeared, and the m/z = 2618 signal increased, but the m/z = 2860 and 3022 peaks remained. The results thus suggest that modification(s) in m/z = 2860 and 3022 compounds prevents the removal of terminal phosphate groups.

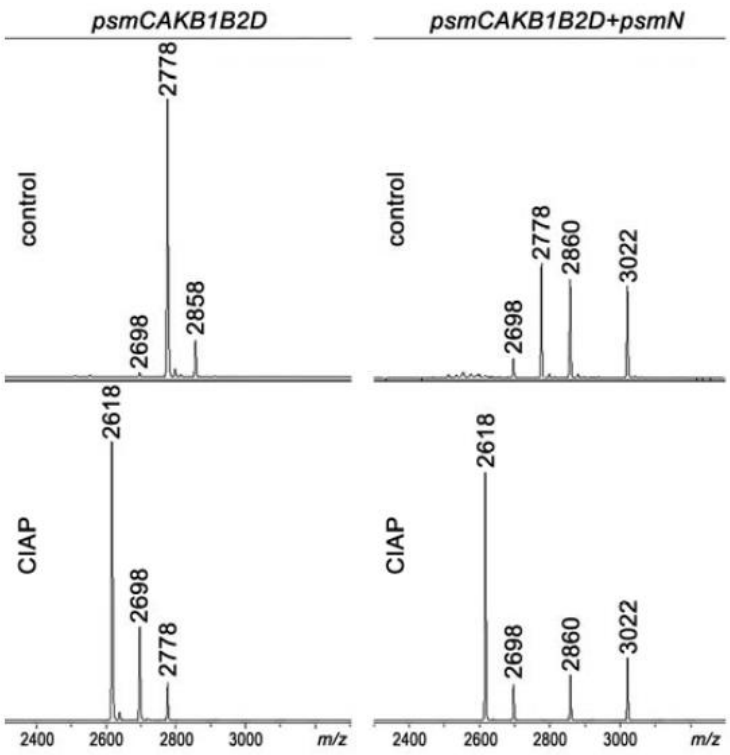


***Figure 19*. The +162 Da PsmN-dependent modification protects phosphorylated pseudomycoidin from enzymatic dephosphorylation**. MALDI-MS analysis of extracts prepared from *E. coli* cells carrying the *psm* cluster with deleted *psmN* deletion and an additional control (left panels) or *psmN*-overexpressing plasmid (right panels). Where indicated, extracts were treated with CIAP prior to mass-spectrometric analysis.

MS-MS fragmentation spectra of compounds from monoisotopic m/z = 2858.3 and m/z 3020.3 peaks were obtained (Figure 20). As a control, fragmentation of monoisotopic m/z = 2856.3 triphosphorylated pseudomycoidin was also performed. This resulted in an expected fragmentation pattern with peaks matching the successive loss of the phosphate groups from the C-terminal serine residue. The fragmentation of the m/z = 2858.3 compound revealed a loss of the 162 Da unit, followed by the loss of phosphate. The fragmentation of the m/z = 3020.3 compound proceeded similarly – through the sequential loss of 324 (2*162) and 98 Da from the C-terminal serine residue of the lasso peptide.

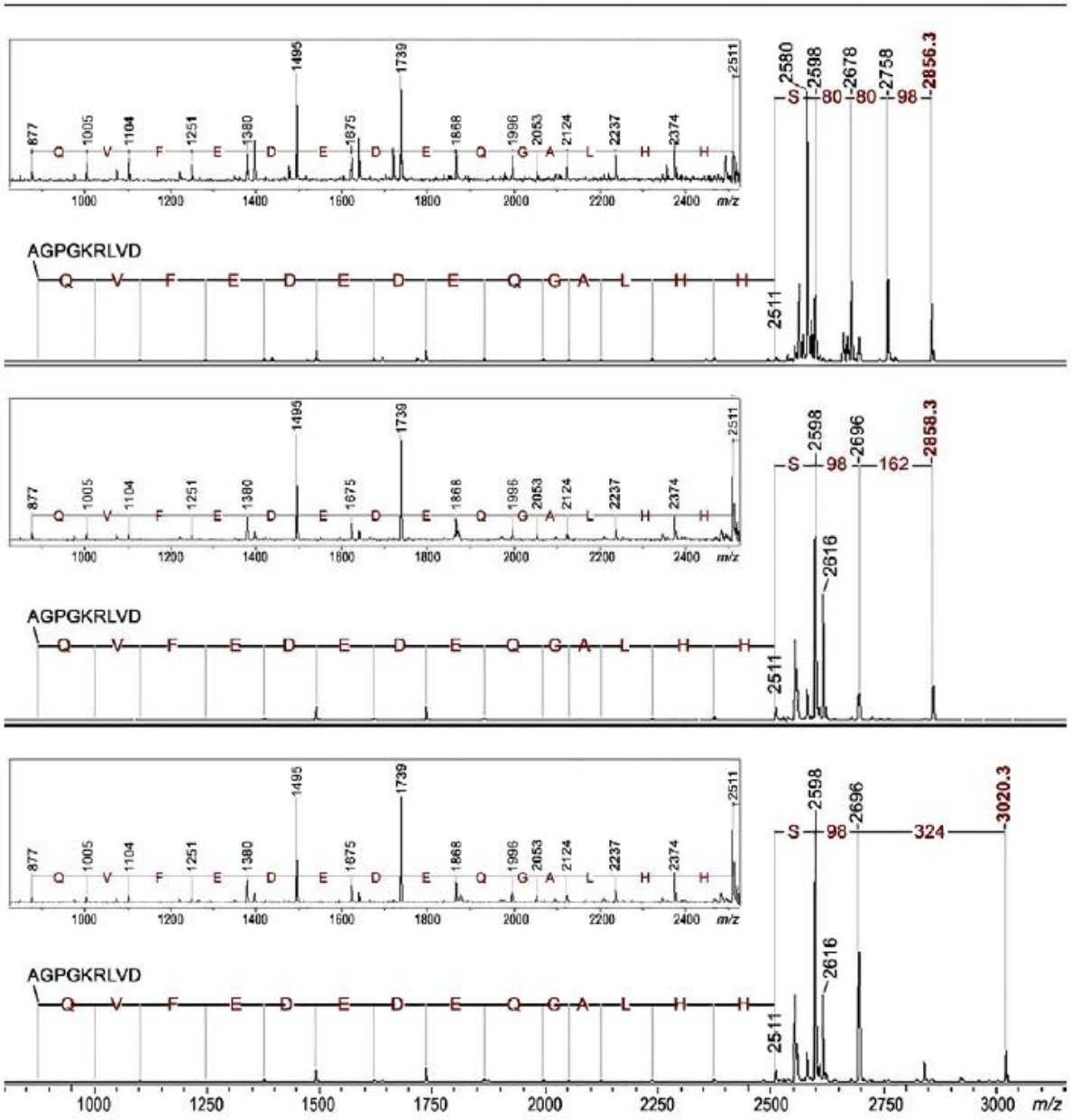


***Figure 20*. MALDI MS-MS analysis of different forms of pseudomycoidin.** Upper panel – fragmentation spectrum of triphosphorylated lasso pseudomycoidin (monoisotopic m/z = 2856.3, average m/z = 2858). Peaks corresponding to C-terminal lasso peptide fragments are marked. Low-intensity daughter peaks corresponding to fragments that arise due to breakage of peptide bonds inside pseudomycoidin are magnified in the insert. Mass differences of 98 and 80 Da match the H_3PO_4 and HPO_3 groups, respectively. Middle panel – fragmentation spectrum of monoisotopic *m/z* = 2858.3 (average m/z = 2860) pseudomycoidin form. A mass difference of 162 Da between the mother ion and the first daughter ion (m/z = 2696) corresponds to the addition of a $C_6H_{10}O_5$ group, which matches an unmodified hexose residue. The mass difference between the first daughter ion and the dehydrated C-terminal Ser pseudomycoidin (*m/z* = 2598) equals 98 Da, which matches the H_3PO_4 group, and indicates that the hexose moiety is attached to the C-terminal phosphate. Lower panel –fragmentation spectrum of monoisotopic m/z = 3020.3 (average m/z = 3022) pseudomycoidin form. The mass difference of 324 Da between the daughter peak of phosphorylated peptide (*m/z* = 2696) and the mother peak matches the dihexose moiety.

The 162 and 324 Da mass differences match the addition of one ($C_6H_{10}O_5$) or two hexose residues ($C_{12}H_{20}O_{10}$), respectively. High-resolution MS of unphosphorylated and

monophosphorylated pseudomycoidin resulted in mono-isotopic m/z values of 2616.2431 and 2696.2118, respectively, which are in 0.5 ppm accordance with the calculated values for brutto formulae of expected compounds (2616.2439 and 2696.2104, respectively). An m/z value of 3020.3175 was obtained for the major peak observed in cells expressing *psmN*. This is in 0.5 ppm accordance with calculated values for a brutto formula of monophosphorylated pseudomycoidin with two hexose residues added (3020.3162). Indeed, the mass difference between m/z = 3020.3162 and m/z = 2696.2104 is 324.1057, which matches within 0.5 ppm a dihexose residue $C_{12}H_{20}O_{10}$ (324.1056).

PsmN belongs to a family of uncharacterized nucleotidyl-transferases and has no closely related proteins with a known function. Protein modeling using Phyre2 server (http:// www.sbg.bio.ic.ac.uk/phyre2/html/) reveals limited structural homology between PsmN and 200-aminoglycoside nucleotidyl-transferases (PDB 5cfs; PDB 4xje). This allowed us to identify possible catalytic metal binding amino acids in PsmN. A PsmN D134A D136A double mutant was inactive when overexpressed in cells bearing the *psm* cluster with the *psmN* deleted (Figure 21.A, lower panel), i.e., no hexose-modified pseudomycoidin peaks were observed.

Overall, the data above show that in addition to variable numbers of phosphate groups attached to the C-terminal serine residue, pseudomycoidin can be decorated with an unidentified hexose(s) installed on the phosphate(s). PsmN is required for this modification.

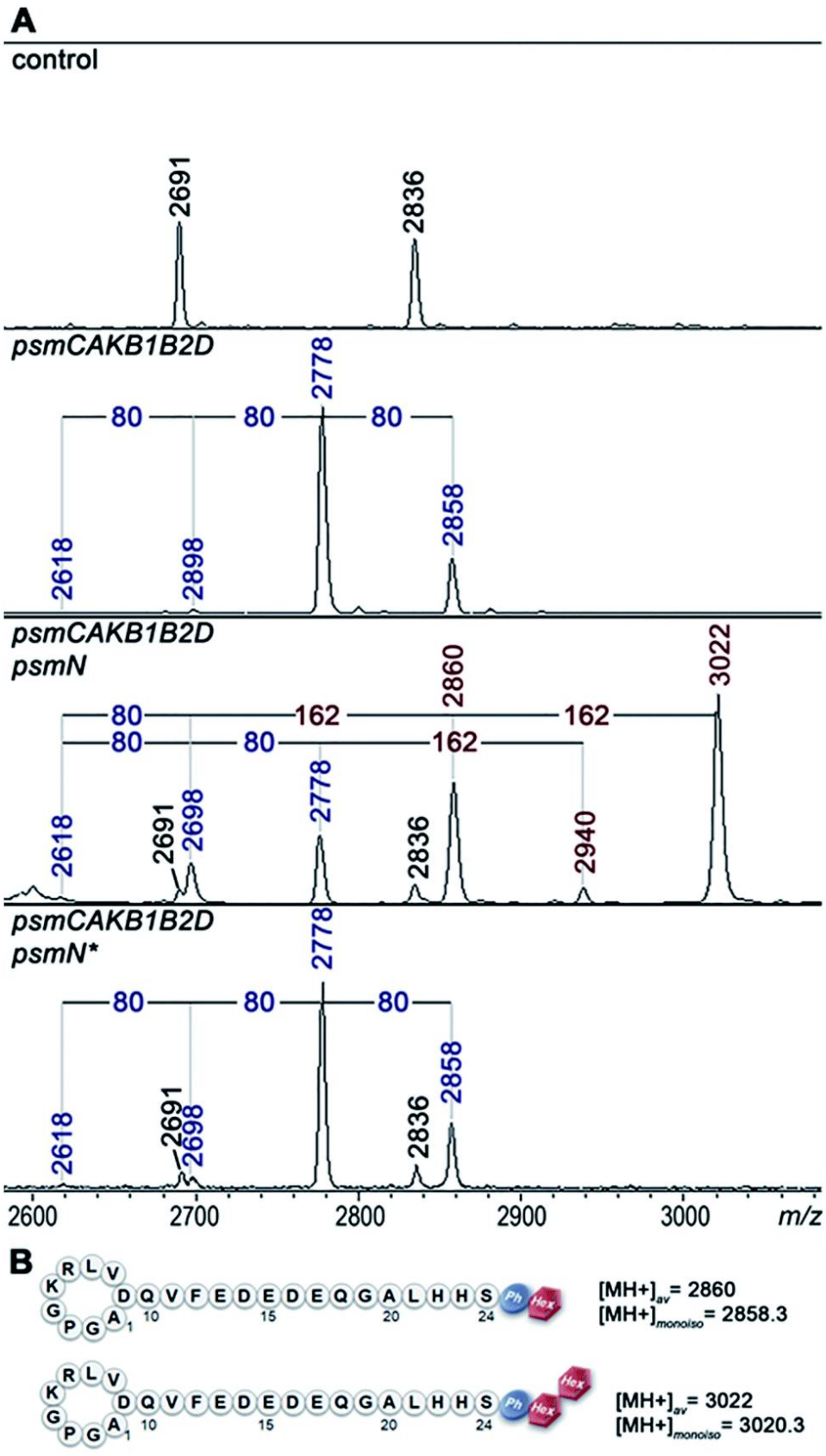


***Figure 21*. PsmN is required for an additional post-translational modification of phosphorylated pseudomycoidin.** (A) MALDI-MS analysis of *E. coli* BL21(DE3) cells containing control plasmids, or plasmids expressing the *psm* cluster with *psmN* deleted, with or without additional plasmid expressing wild-type, or mutant (*psmN**, encodes the D134A D136A double mutant) *psmN*. The average m/z values of mass-peaks specific for cells carrying the *psm* genes are marked with blue font and for *psmN*-dependent peaks, brown color font. (B) Schematic structures of fully modified pseudomycoidins.

4.2.3. Mutational analysis of the *psm* cluster: deletion of *psmB1, psmB2* and $psmA^L$.

Plasmid lacking *psmB1* and *psmB2* genes were also created. Surprisingly, the cells carrying *psm* cluster without *psmB1* and *psmB2* genes, still were able to synthesize pseudomycoidin and its phosphorylated variants (Figure 22). We this conclude that the PsmC enzyme does not require the presence of PsmB2 protein to cyclize the core peptide, in stark difference to other lasso peptide biosynthetic systems known [51], [68], [95]. Although both PsmB1 and PsmB2 are functional *in vivo* [148] the unexpected result suggested that in the absence of these two proteins their role is performed by some host proteins.

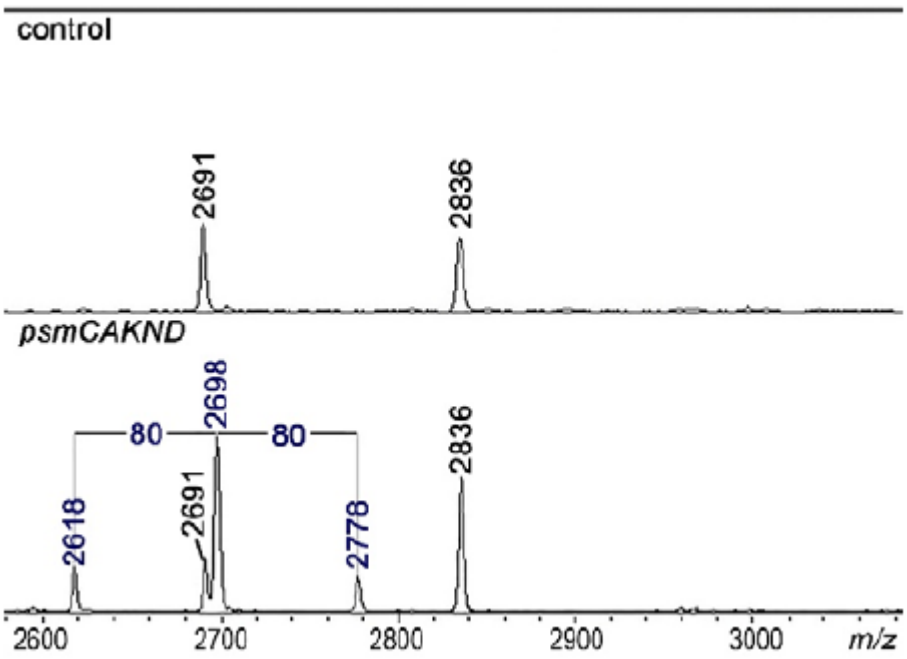


***Figure 22*. PsmB1 and PsmB2 are dispensable for pseudomycoidin maturation.** MALDI MS analysis of the *E. coli* BL21 (DE3) cells harboring the *psm* cluster without the *psmB1B2* genes. Mass-peaks corresponding to unmodified lasso peptide (average *m/z* = 2616) and its mono- and diphosphorylated forms (average *m/z* = 2698 and 2778, respectively), are labeled with the blue color font.

To confirm further functionality of PsmB1 and PsmB2, an *in vitro* experiment was performed using the purified recombinant proteins and synthetic full-sized PsmA precursor peptide of 4910 Da. As can be seen from Figure 23, the precursor peptide was specifically cleaved at the core–leader junction in the presence of both PsmB1 and PsmB2. No cleavage was observed

when PsmB1 or PsmB2 were used separately. These data completely agree with the reported *in vitro* assay of B1 and B2 processing enzymes from paeninodin clusters [131].

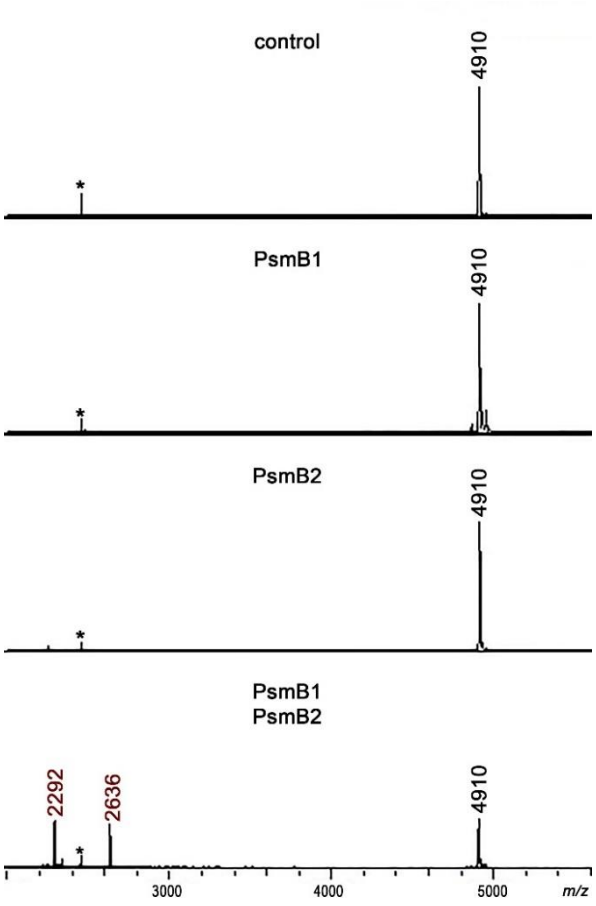


***Figure 23*. PsmB1 and PsmB2 cleave the leader peptide of the pseudomycoidin precursor peptide *in vitro*.** MALDI MS spectra of the resulting reaction products of incubation of the PsmA propeptide with proteins are indicated. The mass peak with m/z = 4910 corresponds to PsmA; the mass peaks with *m/z* = 2292 and *m/z* = 2636 correspond to the leader and the core peptide parts, respectively. The peak marked with an asterisk is the MH+ ion of the full-sized PsmA.

The PsmB1/PsmB2 cleavage was specific. In fact, the synthetic peptide corresponding to the precursor peptide encoded by a similar cluster from *Paenibacillus polymyxa* ATCC 842 (Figure 24.A) was not cut despite having a leader sequence that is highly similar (14 out of 19 identical residues) to the PsmA leader sequence and an identical leader–core junction sequence (Figure 24.B).

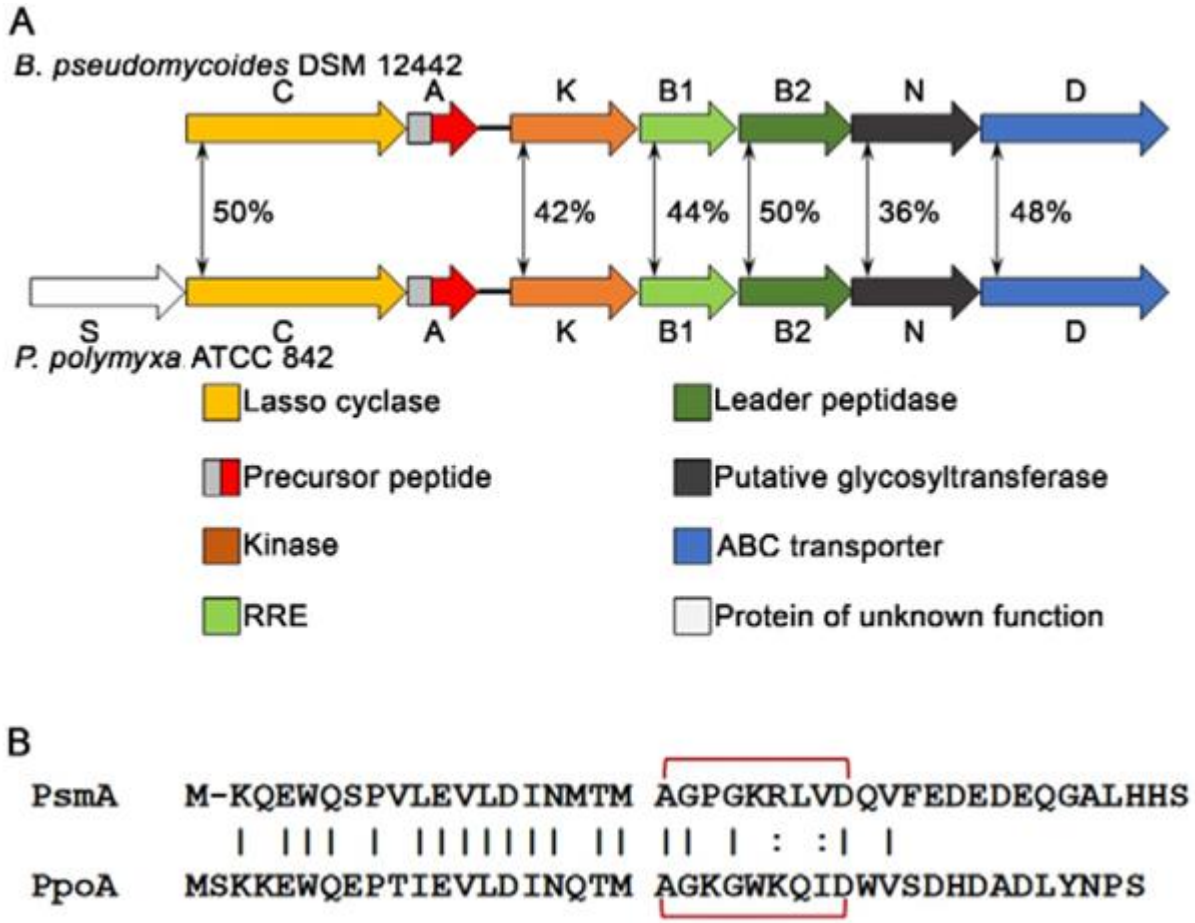


***Figure 24*. Comparison of lasso peptide gene clusters from *B. pseudomycoides* DSM 12442 and *P. polymyxa* ATCC 842.** (A) A schematic representation of the gene arrangements in the clusters. Genes are indicated by arrows, and homologous genes are indicated by identical colors. Numbers indicate the percentage of identity between homologous protein sequences. Known or putative gene product functions are listed at the bottom. PpoS has homology to the sulfotransferase family of proteins. (B) An alignment of precursor peptides from *B. pseudomycoides* DSM 12442 and *P. polymyxa* ATCC 842. Amino acids forming the macrolactam ring are linked by red brackets.

To further investigate pseudomycoidin synthesis in the absence of PsmB1 and PsmB2 enzymes, we concentrated on analysis of a simplified *psmCAD* cluster lacking the psmK and psmN genes. *E. coli* cells harboring such a cluster produced pseudomycoidin without the phosphate or hexose modifications (Figure 25).

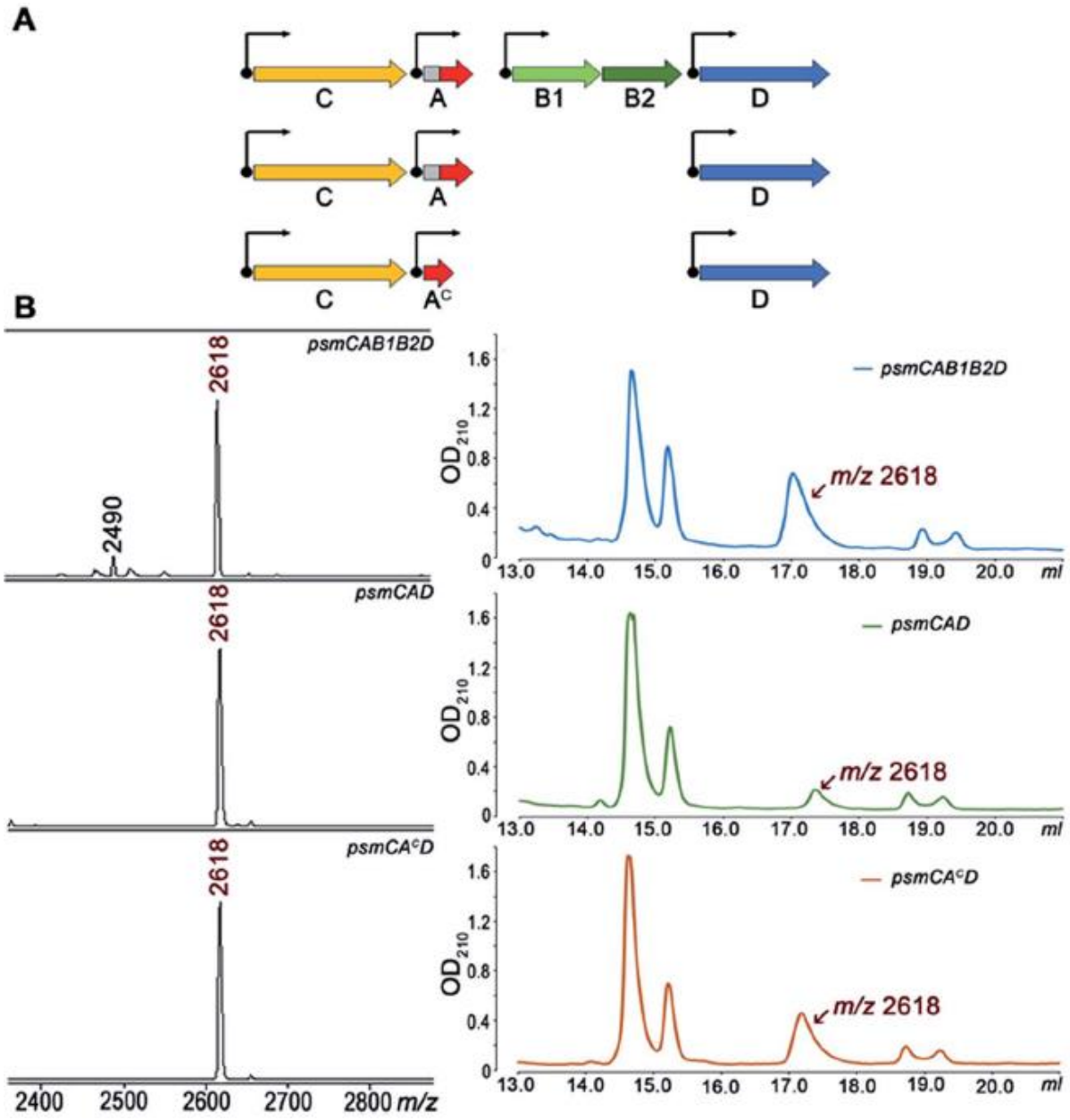


***Figure 25*. PsmB1/PsmB2 proteins are dispensable for pseudomycoidin lasso structure formation.** (A) Three-gene arrangements used for pseudomycoidin production in *E. coli*. (B) MALDI MS spectra of the cells harboring the indicated sets of the *psm* genes (left panels) and HPLC traces of fractionation of corresponding conditioned media (right panels). HPLC fractions containing the lasso peptide (average m/z = 2618) were identified by MALDI MS.

Pseudomycoidins obtained from expression systems carrying *psmCAB1B2D* and *psmCAD* genes were compared by TOF-TOF high-resolution mass spectrometry. As Figures 26-2 demonstrate both peptides have the same monoisotopic masses as well as coinciding MS-MS fragmentation spectra. MS-MS analysis confirmed the identity of these compounds in terms of the amino acid sequence and the macrolactam ring location.

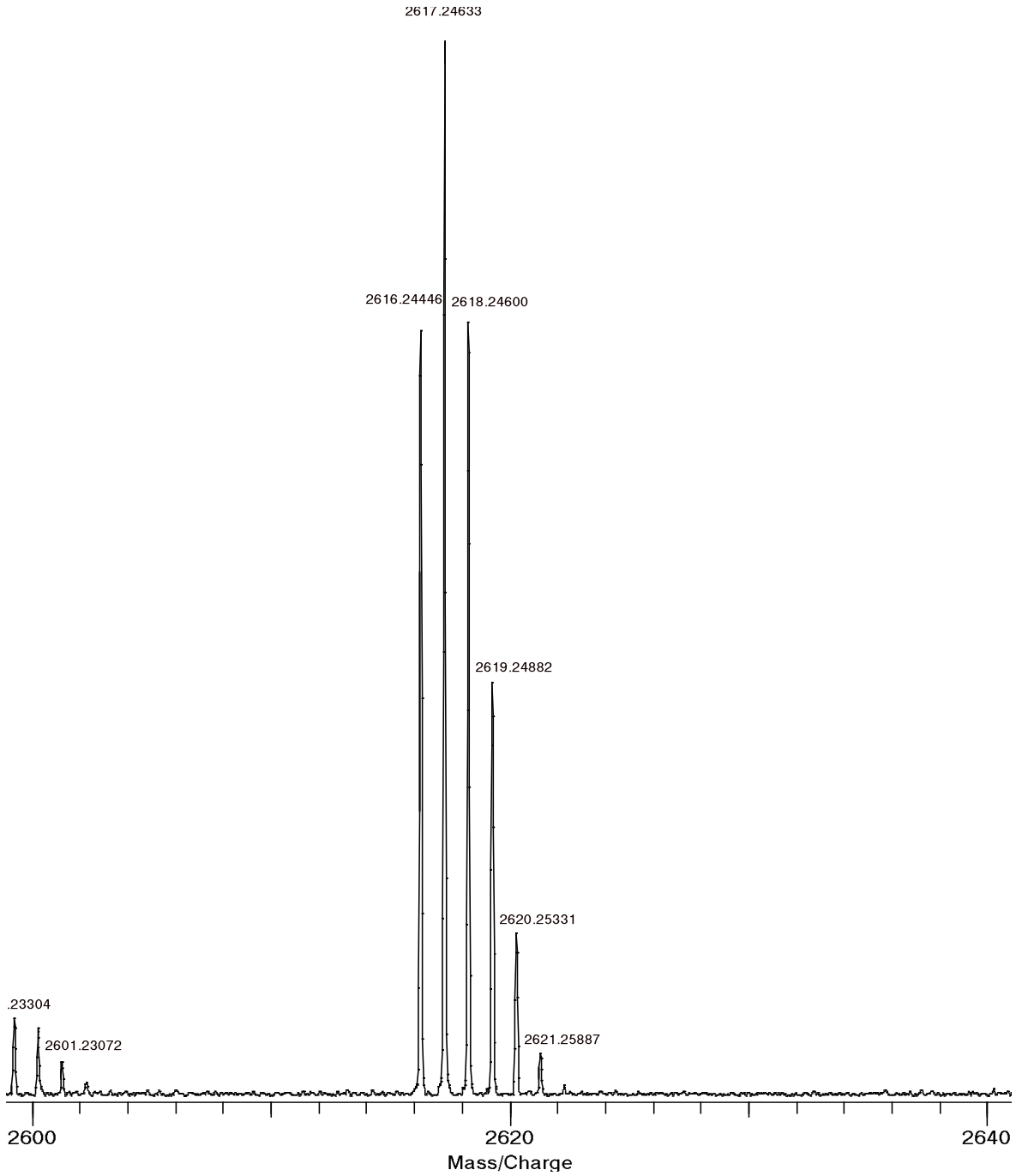


***Figure 26*. High-resolution FT-ICR-MS spectrum of pseudomycoidins obtained from cells expressing *psmCAB1B2D* genes.** The m/z values of major peaks M+H+ are indicated. Calculated value for brutto formulae: 2616.2439; error: 0.03 ppm. Internal calibration (ACTH fragment18-39 (2465.1989Da) and Insulin oxidized B chain (3494.6513Da)**.**

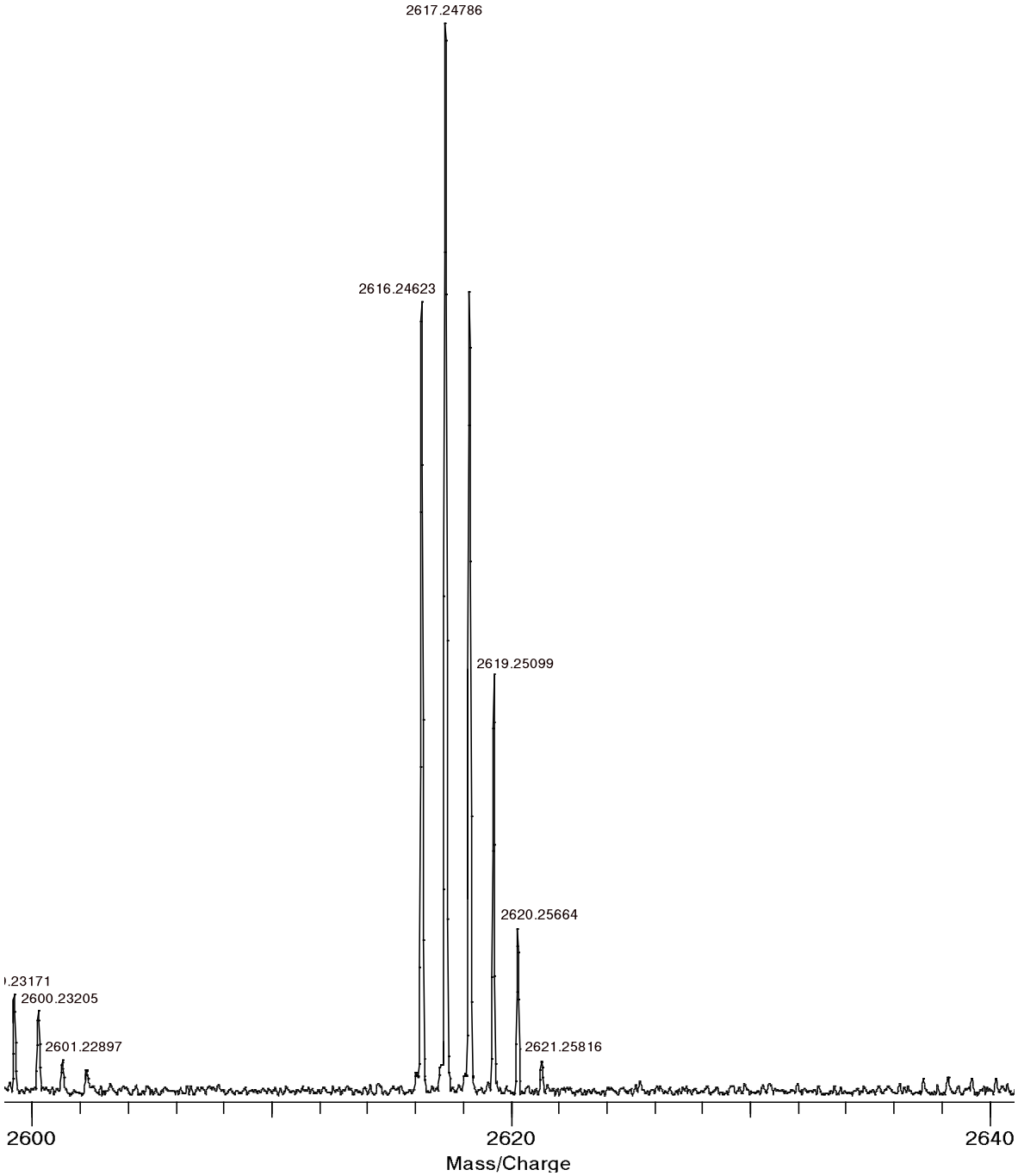


***Figure 27*. High-resolution FT-ICR-MS spectrum of pseudomycoidins obtained from cells expressing *psmCAD* genes.** The m/z values of major peaks M+H+ are indicated. Calculated value for brutto formulae: 2616.2439; error: 0.04 ppm. Internal calibration (ACTH fragment18-39 (2465.1989Da) and Insulin oxidized B chain (3494.6513Da)**.**

To test whether the ability to produce a lasso peptide without the universally conserved B1 and B2 enzymes is a specific feature of the heterologous production host, we introduced the *psmCAD* cluster into *B. subtilis*, an organism that is phylogenetically closer to the native producer

than *E. coli*. As can be seen in Figure 28 the cells harboring *psm* expression plasmid produced correctly processed linear core peptide of PsmA as well as its cyclized form, i.e., pseudomycoidin. We therefore conclude that PsmB1 and PsmB2 though functional and specific to the pseudomycoidin propeptide PsmA are not required for pseudomycoidin production either in *E. coli* or in *B. subtilis*.

To determine if the PsmA leader sequence is needed for the cyclisation reaction catalyzed by PsmC, we next created a pseudomycoidin expression system that encoded the core peptide of PsmA and $PsmA^{C}$, without the leader part.

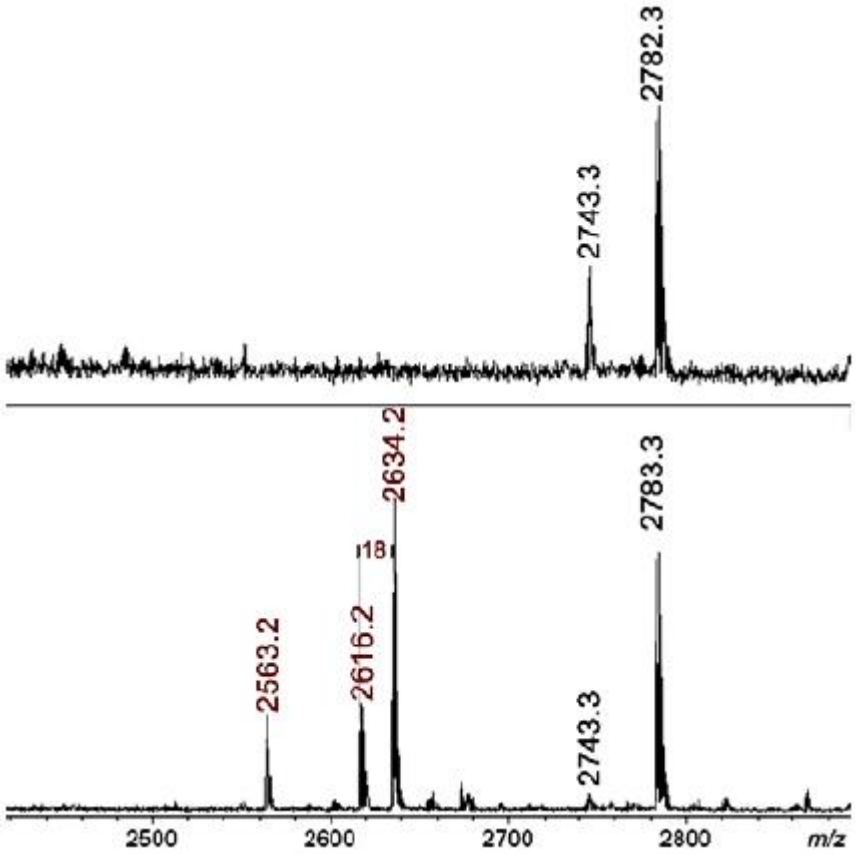


***Figure 28*. Pseudomycoidin production in *B. subtilis*.** MALDI MS spectra of *B. subtilis* cells harboring the control plasmid (upper panel) or a derivative plasmid expressing the *psmCAD* gene set (lower panel). The monoisotopic *m/z* = 2634.2 peak corresponds to the core part of PsmA propeptide (amino acids 1-24); the monoisotopic *m/z* = 2563.2 peak matches incorrectly processed PsmA (amino acids 2-24); the monoisotopic *m/z* = 2616.2 corresponds to pseudomycoidin lasso peptide.

While the recombinant core peptide was produced with N-terminal formyl-methionine incorporated during translation initiation, we expected this residue to be efficiently removed by the *E. coli* methionine aminopeptidase [190], [191] This cleavage results in the free NH-group of the first amino acid of the core peptide (Ala^1) that can be subjected to cyclization by the C-enzyme. Analysis of *E. coli* cells expressing *psmCA*^{C}*D* genes revealed robust production of pseudomycoidin. The yield was higher than that in cells expressing *psmCAD*, suggesting that removal of the leader in the absence of PsmB1 and PsmB2 is a limiting step in production.

Overall, we conclude that production of lasso peptide pseudomycoidin requires neither the leader part of the precursor peptide nor the universally conserved B1 and B2 enzymes.

To check whether the core part of PsmA can form a lasso in the absence of PsmC, plasmids expressing *psmAB1B2D* and *psmAD* were tested. Cells lacking PsmC produced only linear PsmA and fragments thereof. Thus, PsmC is strictly required for pseudomycoidin lasso structure formation.

We also checked pseudomycoidin production in the absence of the putative PsmD transporter by cells expressing *psmCA* and *psmCA*^{C}. Pseudomycoidin was detected inside the cells in both cases, indicating that PsmD is not required for its synthesis. We conclude that a gene encoding the core part of the PsmA peptide and the lasso peptide macrolactam synthetase PsmC gene are sufficient for production of pseudomycoidin. We also attempted *in vitro* synthesis of pseudomycoidin from a chemically synthesized PsmA core peptide using recombinant PsmC, alone or in the presence of PsmB1/B2 proteins. No cyclisation was observed; however, the PsmC enzyme was very poorly produced and likely misfolded.

4.3. Elucidating the 3-dimensional structure of pseudomycoidins

We were next interested to determine whether pseudomycoidin is a threaded lasso peptide or a branched cycle (Figure 5) that cannot be distinguished by MS analysis.

4.3.1. The thermal stability assay

High temperature treatment has been shown to stimulate the release of the tail in some threaded lasso compounds [62], [67], [69]–[71]. We compared chromatographic behavior of pseudomycoidin produced from full-sized PsmA or from its core part after 4 hours incubation at 95°C with untreated controls. While heat treatment caused partial degradation of the peptide, no change in HPLC retention time for full-sized pseudomycoidin peak was observed (Figure 29).

Since in studied systems chromatographic behavior of branched cycles and threaded lasso isomers is usually distinct (Figure 6) [69], the result suggests that pseudomycoidin is either a branched cycle or a stable threaded lasso, such as MccJ25 or capistruin [67], [174].

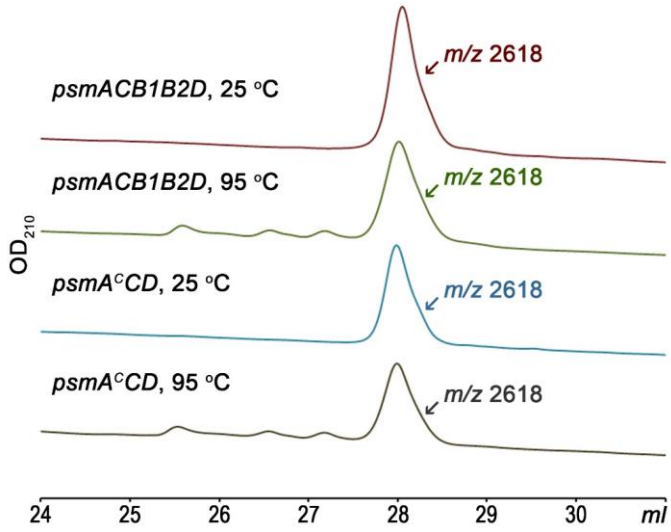


***Figure 29*. Thermostability of pseudomycoidin**. UV-HPLC traces of pseudomycoidin purified from cells harboring *psmCAB1B2D* or *psmCA^{C}D* gene sets before and after 2-hour incubation at 95°C. Subsequent MS/MS analysis confirmed the identity of heat-treated compounds with the untreated ones. The arrow indicates the peak containing the lasso peptide with the average m/z = 2618.

4.3.2. Protease degradation assays

We performed several additional tests to distinguish branched cycles from threaded lassos [192] and to compare pseudomycoidin produced from full-sized or the core part of PsmA and in the presence or the absence of PsmB1/B2 proteins.

The carboxypeptidase Y digestion assay was performed. Despite the broad amino acid specificity of CPDY, the efficiency of cleavage is different and depends on the amino acid sequence of the substrate. Therefore, to have a reference substrate with a known topology, we treated pseudomycoidin produced by cells harboring *psmCAB1B2D* with trypsin, which was found to cut inside the ring between Lys^5 and Arg^6 residues (Figure 30, upper panel). The resulting molecule should be Y-shaped and its tail shall be free from any possible topological constrains. Equal amounts of trypsin-digested pseudomycoidin and intact compounds produced by cells with various combinations of *psm* genes were combined and treated with CPDY. The 18 Da difference between trypsin-digested and intact substrates allowed us to follow protease digestion of both substrates in the same reaction. CPDY was expected to cleave from C-termini of both intact and trypsin digested peptides. It was also expected to cleave amino acids from a C-terminus that was formed after the opening of the ring in the trypsin digested substrate; however, mass peaks corresponding to this activity were minor and did not interfere with analysis (Figure 30). Cleavage of threaded lasso substrates but not of trypsin-treated ones was expected to be limited by topology. As can be seen from Figure 30, digestion with carboxypeptidase led to accumulation of products ending at His^{23}, Ala^{20} , Gly^{19} , Gln^{18} and Gln^{10}. Peaks corresponding to each of these intermediates were split, with a characteristic 18 Da difference indicating that they were produced from both intact and the Y-shaped pseudomycoidin.

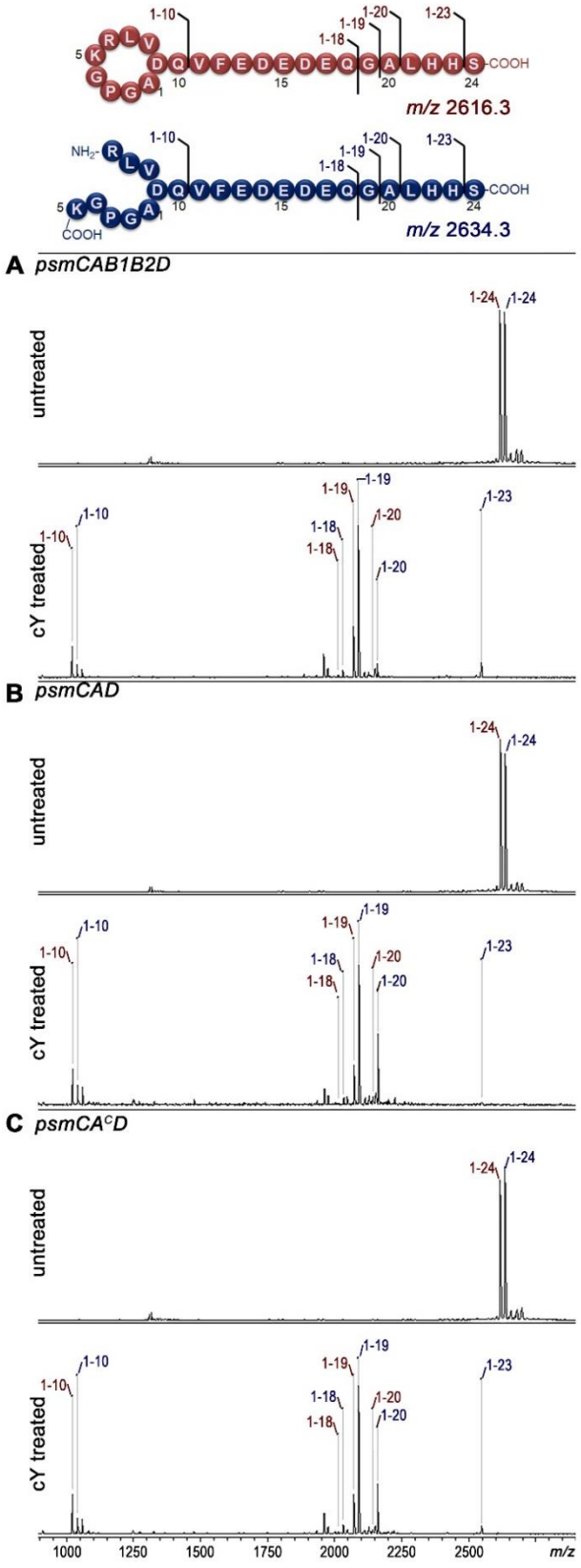


***Figure 30*. Digestion of intact and trypsin-treated pseudomycoidin with carboxypeptidase Y.** Lasso peptides produced under expression of (A) *psmCAB1B2D*, (B) *psmCAD* and (C) *psmCA^{C}D* were partially digested with trypsin to open the macrocycle ring resulting in two C-termini branched peptides, as schematically shown at the top. The trypsin digestion product is distinguished from intact pseudomycoidin by an 18-Da mass shift. Mixtures of trypsin-treated and intact pseudomycoidin were treated with carboxypeptidase Y and the products of the reactions were analyzed with MALDI MS. The smallest proteolytic fragments identified (1-10) correspond to N-terminal macrocycle extended by one amino acid. The peaks corresponding to branched peptides originating from trypsin-digested pseudomycoidin and intact lasso peptide are marked in red and blue, respectively.

The absence of differences in carboxypeptidase digestion product accumulation is consistent with the branched cycle topology of pseudomycoidin. The pseudomycoidins obtained from *psmCAB1B2D*, *psmCAD* and *psmCA^{C}D* clusters appeared to be identical.

Typically, the stable threaded lasso peptides of the class II (such as microcin J25), after cleavage in the loop region form an intermediate molecular [2]rotaxane with a tail entrapped in the ring and unable to escape because of the lock residues (Figure 31). The molecular mass of this treated peptide is equivalent to that of the untreated one with the 18 Da mass shift due to hydrolysis of the peptide bond.

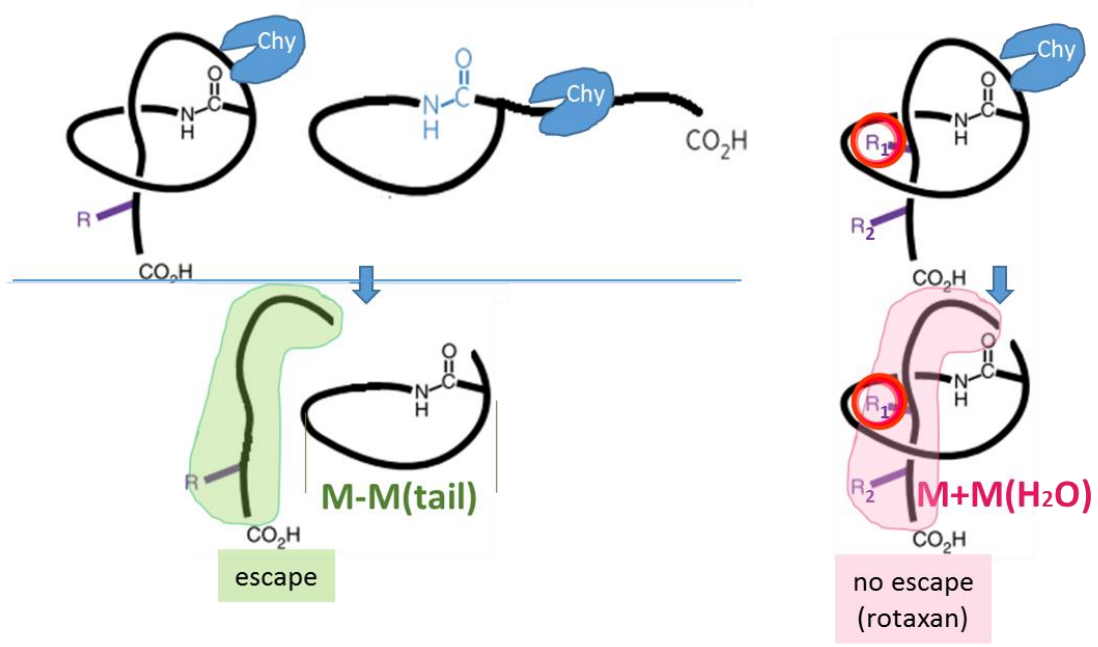


***Figure 31*. Possible molecular fragment formation after loop-region cleavage of the threaded lasso peptide interlocked via bulky residues.** A lasso peptide carrying two lock amino acids ought to form a molecular [2]rotaxane with an interlocked tail, while the branch-cyclic peptide as well as the lasso peptide locked via only one residue should lose the tail from the ring.

Chymotrypsin digestions of pseudomycoidins followed by MS analysis revealed that the main mass peaks corresponded to the fragments 1-12, 1-21 and 1-22. No peak matching the hydrolyzed full-size pseudomycoidin was observed (Figure 32). For all three tested pseudomycoidins, the same degradation spectra were obtained that also verifies the equivalency of the peptides.

The emergence of 1-21 and 1-22 fragments did not contradict the hypothesis of a threaded lasso, because the nearly full-size peptide was present with only some lost C-terminal residues that may possibly be located below the ring. The appearance of 1-12 ion indicated that the C-terminal

tail was completely separated from the ring part and no [2]rotaxane was formed. The data are consistent with either the branched cycle topology or threaded lasso with only one lock residue or containing an extremely short loop region such that F12 is located below the ring.

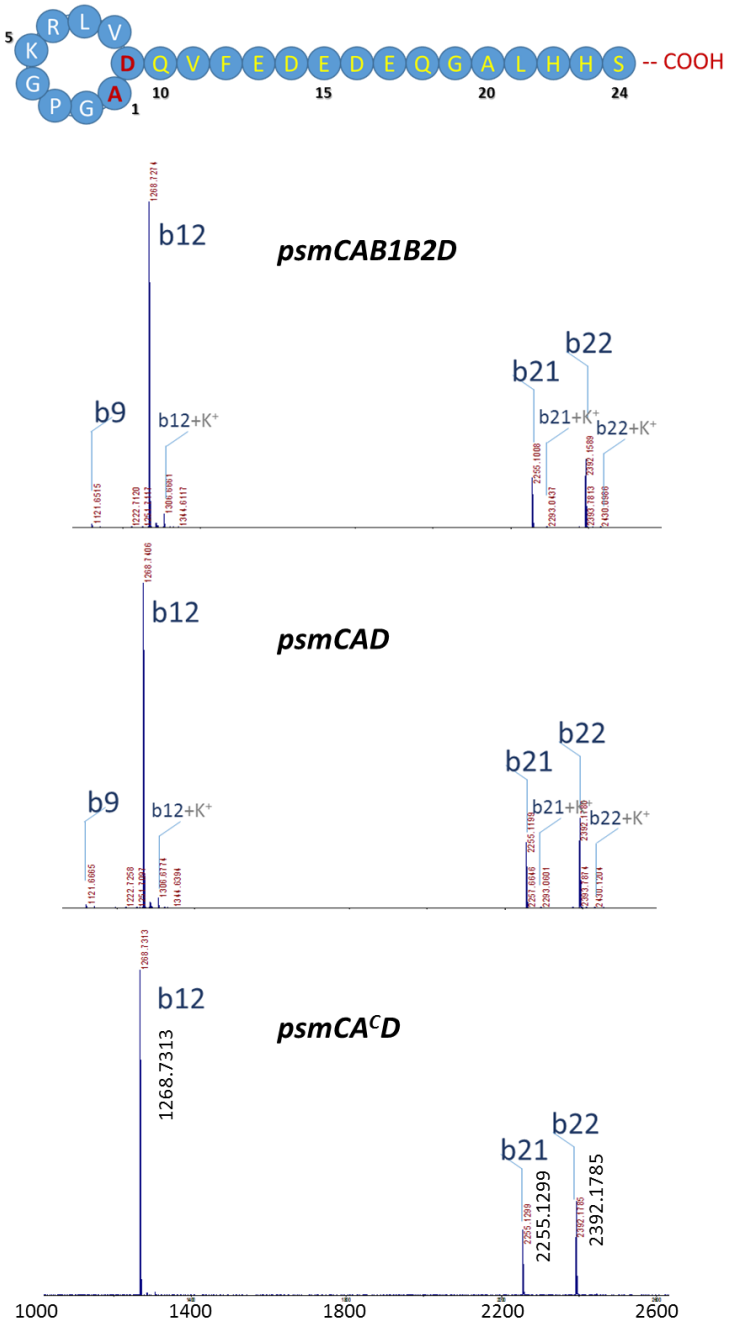


***Figure 32*. MALDI-MS spectra of chymotrypsin-digested pseudomycoidins**. The b-ions (labeled in the spectrum in blue font) are much more prominent.

Another test included digestions with Glu-C protease. Similarly to chymotrypsin digestions, GluC treatment should result in loop/tail section cleavage of pseudomycoidins. Under our conditions, this protease cleaves pseudomycoidin in three positions between E^{13}-D^{14}, E^{15}-D^{16} and E^{17}-Q^{18} in the tail. For a stable threaded lasso such as microcin J25, cleavage in the tail results in the formation of a [2]rotaxane intermediate consisting of a ring and a linear tail trapped within it.

HPLC analysis of Glu-C treated pseudomycoidin revealed the disappearance of intact compounds and the appearance of additional peaks. According to MALDI-MS, these peaks contained separate rings or tail products of Glu-C digestion (Figure 33) with no peaks containing both fragments, as would be expected for a [2]rotaxane topology.

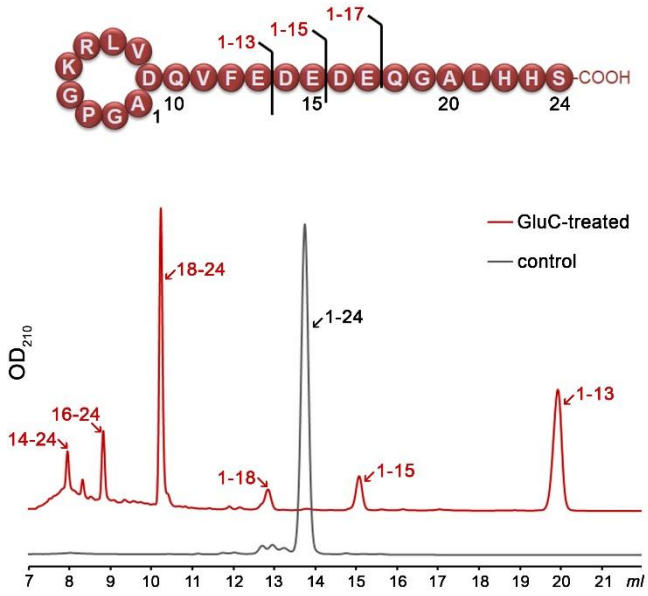


***Figure 33*. HPLC traces of pseudomycoidin cleaved with Glu-C protease.** The Glu-C cleavage sites are schematically shown at the top of the figure. The products of the Glu-C proteolysis were separated by reverse phase HPLC along with untreated control (red and grey traces, respectively). HPLC peaks containing proteolytic fragments were identified using MALDI MS. Pseudomycoidin was produced in *E. coli* cells harboring the minimal *psmCA^{C}D* set of genes.

Taken together, the results suggest that pseudomycoidin, once purified from cells harboring *psmCAB1B2D* or smaller sets of genes, is a branched-cyclic peptide

4.4. Mutational analysis of the core peptide

Given the fact that pseudomycoidin synthesis machinery is homologous to that of validated lasso peptides, the branched lasso topology may arise when the tail is withdrawn from the macrolactam cycle during purification. In fact, the tailoring modifications (large hexose residues on the C-terminus) could serve as a lock of the threaded structure. Unfortunately, the hexose-modified peptide could not be obtained in sufficient amount due to the extremely low yield under laboratory conditions.

At that point, two general questions have appeared: (1) could we achieve the lasso form pseudomycoidin by introducing a large residue instead of the smaller one into the tail region of the core peptide; (2) would the C-enzyme tolerate such substitutions and still cyclase the peptide?

A number of mutant expression clusters were designed based on our *psm* expression system. We have replaced the *psmA* gene with the mutated ones coding for PsmA variants, which differ from the wild-type in the tail/loop region at positions 15-18, 21-23. Each of these amino acids was substituted either with a small Ala residue or with a large Phe residue resulting in 14 pseudomycoidin variants. The calculated masses of the expected mutants are listed in the Table 2. The created plasmid vectors were co-expressed with *psmCB1B2D* in *E. coli* cells under conditions described above. Pseudomycoidin variants productions in the culture medium were monitored by MS-analysis. In all cases the expected products were detected with their masses corresponding to cyclized peptides (Figure 34). We conclude that this limited *in vivo* analysis for substrate substitution tolerance of the biosynthetic machinery reveals the capability of PsmC processing enzyme to recognize a broad range of substrates and to cyclize them.

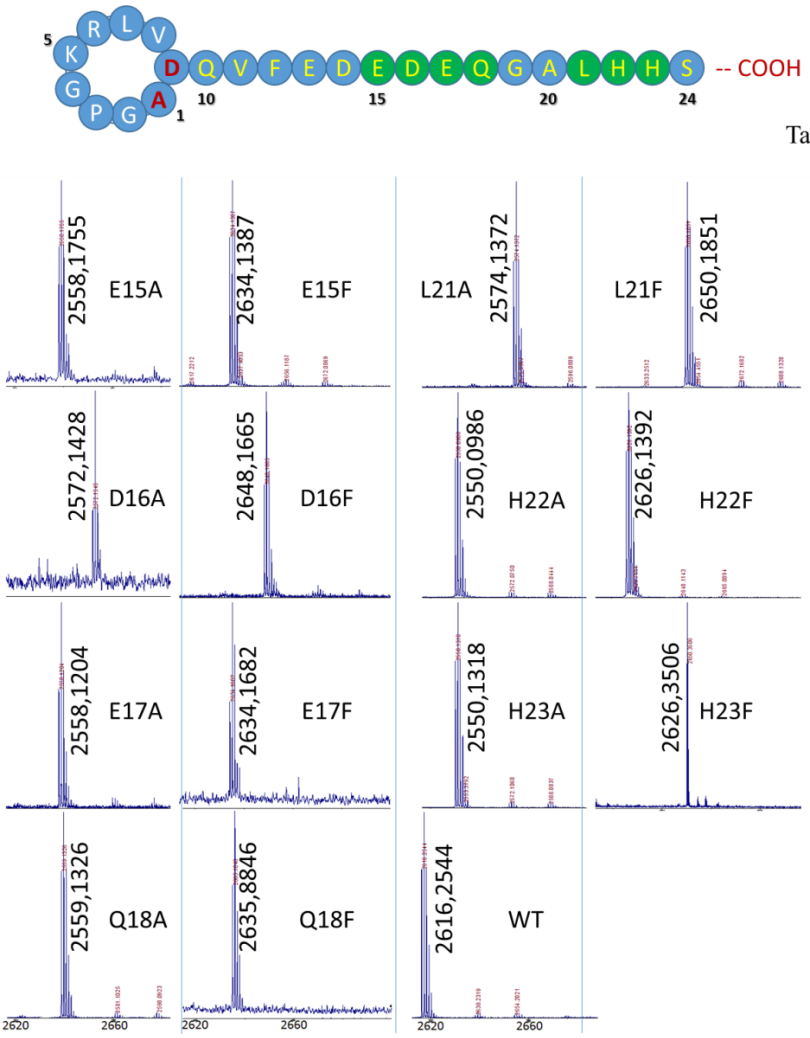


Table 2. The list of calculated masses of pseudomycoides variants

Core peptide	calculated monosiotopic m/z
WT	2616,75705
E15A	2558,720887
E15F	2634,817065
D16A	2572,747504
D16F	2648,843683
E17A	2558,720887
E17F	2634,817065
Q18A	2559,705648
Q18F	2635,801826
L21A	2574,677198
L21F	2650,773376
H22A	2550,695554
H22F	2626,791733
H23A	2550,695554
H23F	2626,791733

***Figure 34*. MALDI-MS spectra of culture media of induced cells harboring the *psmCAB1B2D* genes and *psmA* carrying mutations encoding single aminoacid substitutions listed in Table 2.** Wild-type Pseudomycoidin is schematically shown at the top of the figure. Amino acids substituted with Ala or Phe are coloured in green. Mass-peaks matching the expected pseudomycoidin variants are labeled.

We next examined the topology of peptides carrying the substituted C-terminal amino acids at positions 21-23. Six pseudomycoidin variants were produced as previously described and further subjected to HPLC analysis and confirmed by MS. Figure 35 shows comparison of UV-HPLC traces of WT pseudomycoidin and six mutants with indicated peaks of interest. All three Ala substituted pseudomycoidin variants L21A, H22A and H23A demonstrated the same HPLC-behavior and retention time. Amino acid substitutions with phenylalanine resulted in L21F, H22F and H23F pseudomycoidins with the former demonstrating different HPLC-behavior than that of wild-type pseudomycoidin.

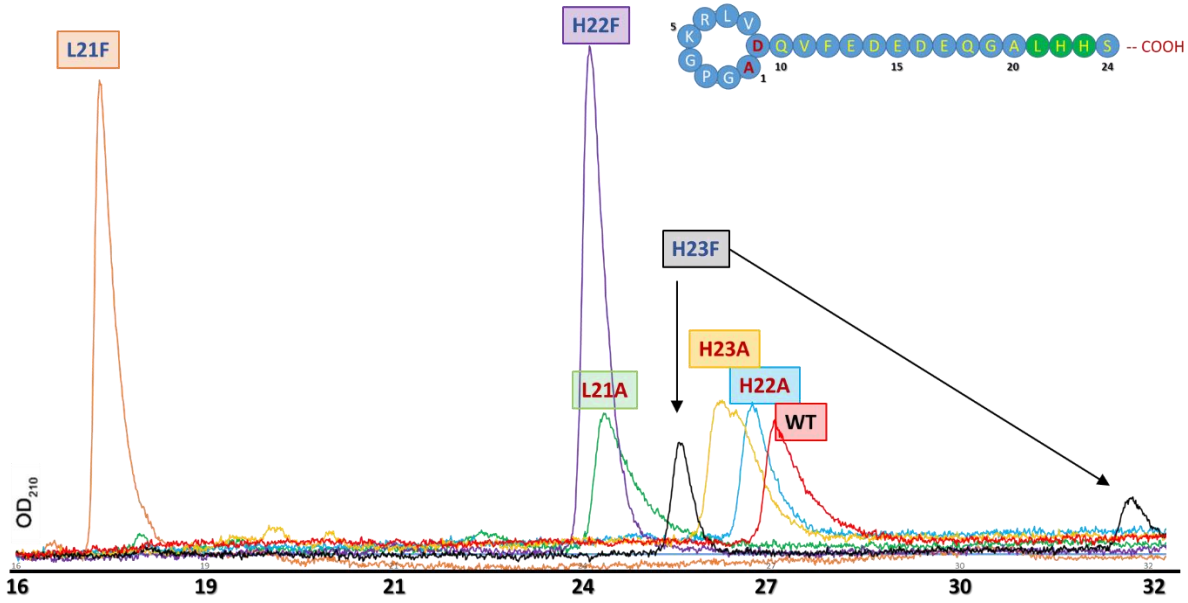


Figure 35. **UV-HPLC (210 nm) traces of six pseudomycoidin variants and of wild-type pseudomycoidin.** Wild-type pseudomycoidin is schematically shown at the top of the figure. Amino acids subjected to single substitution with Ala or Phe are coloured in green. HPLC traces corresponding to pseudomycoidin variants are highlighted with different color. HPLC fractions confirmed by MALDI-MS analysis to contain the expected peptides are labeled.

4.4.1. Pseudomycoidin L21F derivate have lasso topology

The mutant L21F peptide produced by cells harboring either *psmCAB1B2D* or *psmCA^{C}D* clusters eluted much earlier than wild-type pseudomycoidin (Figure 36).

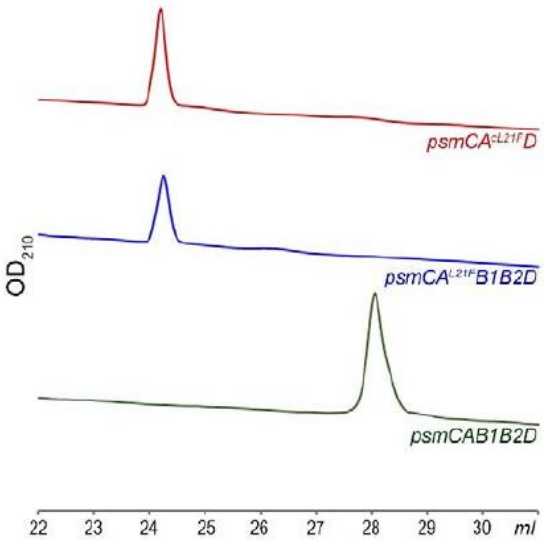


Figure 36. **HPLC traces of the wild-type and L21F mutant forms of pseudomycoidin.** Lasso peptides were produced using either complete (*psmCAB1B2D*) or minimal (*psmCA^{C}D*) set of pseudomycoidin biosynthesis genes as indicated. Wild-type pseudomycoidin (average m/z = 2618) and the L21F mutant (average m/z = 2652) were identified in HPLC peaks by MALDI MS.

After incubation at 95°C and rechromatography, the pseudomycoidin L21F peak became split, with a new peak eluting later, at a position similar to where the wild-type pseudomycoidin eluted (Figure 37.A). Mass-spectrometric analysis revealed that this later peak still had an m/z = 2652 and thus corresponded to chemically intact pseudomycoidin L21F. Material from each peak was next treated with CPDY. For peak B that appeared after thermal treatment (Figure 37.B), a pattern that was previously observed for wild-type pseudomycoidin and consistent with branched lasso topology was observed. In contrast, for the earlier eluting peak A, the cleavage rate was considerably slower, as only cleavage products corresponding to removal of C-terminal 4 amino acids were observed (Figure 37.B).

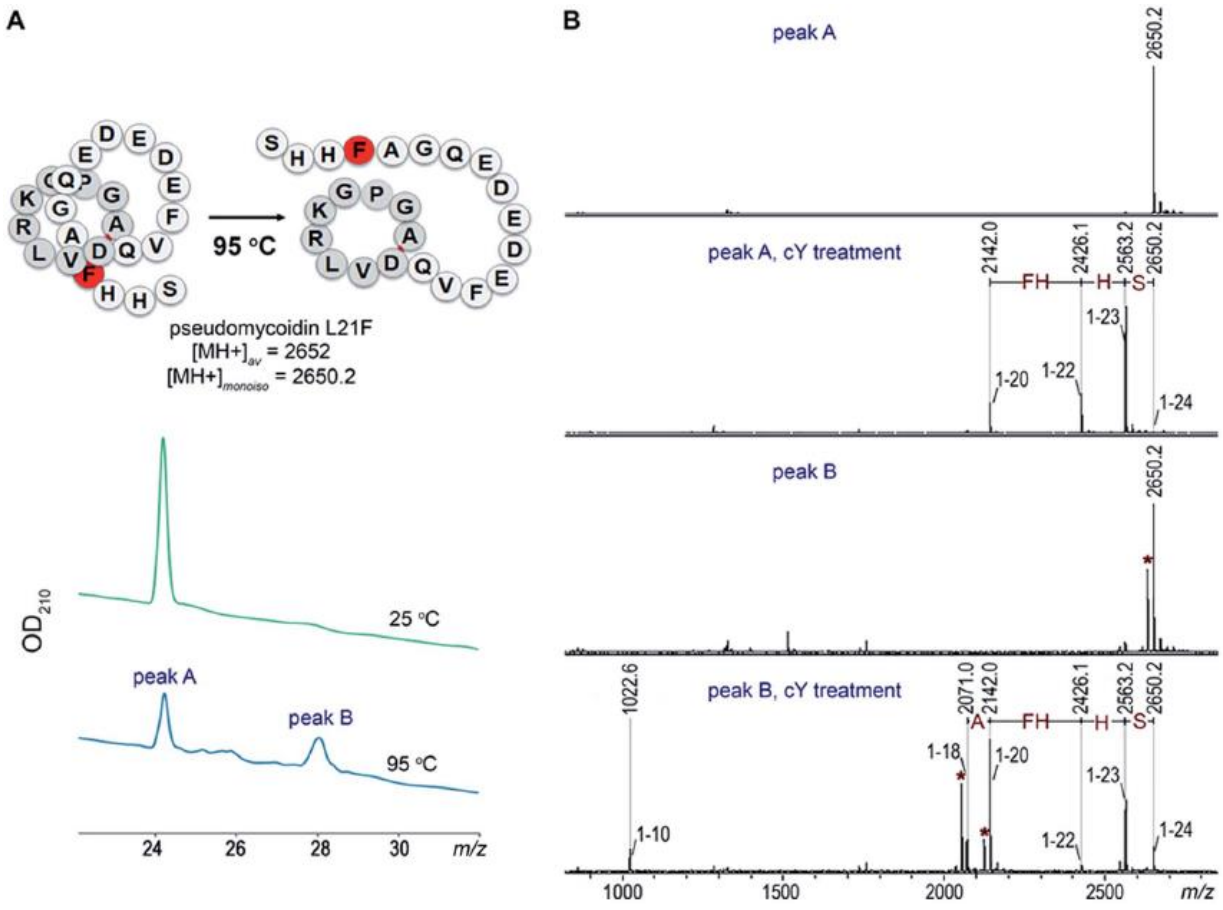


Figure 37. **L21F mutant of pseudomycoidin is a threaded lasso peptide.** An L21F mutant of pseudomycoidin was produced in the truncated expression system psmCA^{C}D involving the mutant core peptide L21F. PsmL21F is schematically shown at the top of the figure. The leucine substituted with phenylalanine is coloured in red. (A) Comparison of UV-HPLC (210 nm) traces of the pseudomycoidin L21F before (green) and after (blue) incubation at 95°C for 2h. One more peak appearing under heating is labeled as "B". (B) Subsequent MALDI MS analysis confirmed the identity of the "B" compound (two upper panels) with the compound collected from HPLC "A" peak (two lower panels) before and after 80 min treatment with carboxypeptidase Y. The smallest proteolytic fragments identified (1–10) correspond to N-terminal macrocycle extended by one amino acid. Dehydrated fragments are marked with asterisk.

In order to confirm our findings by an independent approach, we recorded NMR spectra of wild-type and L21F pseudomycoidins purified from cells harboring *psmCAB1B2D* and *psmCA^{C}D* gene sets, correspondingly, in an aqueous phosphate buffer.

4.5. NMR-analysis

With the help of the team of Dr. Guy Lippens from Toulouse Biotechnology Institute, France, we recorded NMR spectra of wild-type and L21F pseudomycoidins purified from cells harboring *psmCAB1B2D* and *psmCA^{C}D* gene sets, correspondingly, in an aqueous phosphate buffer.

Homonuclear TOCSY and NOESY spectra were recorded at 800 MHz and led to a full assignment of both peptides. Importantly, we observed the characteristic NOE peak between the amide proton of Ala^{1} and the Hb protons of Asp^{9} in both spectra (Figure 38), confirming the cyclic nature of both peptides [193].

However, upon closer comparison of the spectra, both the amide and methyl proton spectral regions of the wild-type pseudomycoidin appear more compact than those of the L21F mutant. As an example, we show the methyl groups of V11 that resonate at 0.84 ppm in the spectrum regions of the wild-type pseudomycoidin, whereas they shift to 0.3 ppm in the L21F mutant (Figure 38).

Another salient difference between both spectra is that while the amide protons of the wild-type peptide are all sharp and of comparable intensity, several amide protons of the L21F mutant pseudomycoidin are severely broadened. This is the case for A1 and G2 also for Q10 and F12.

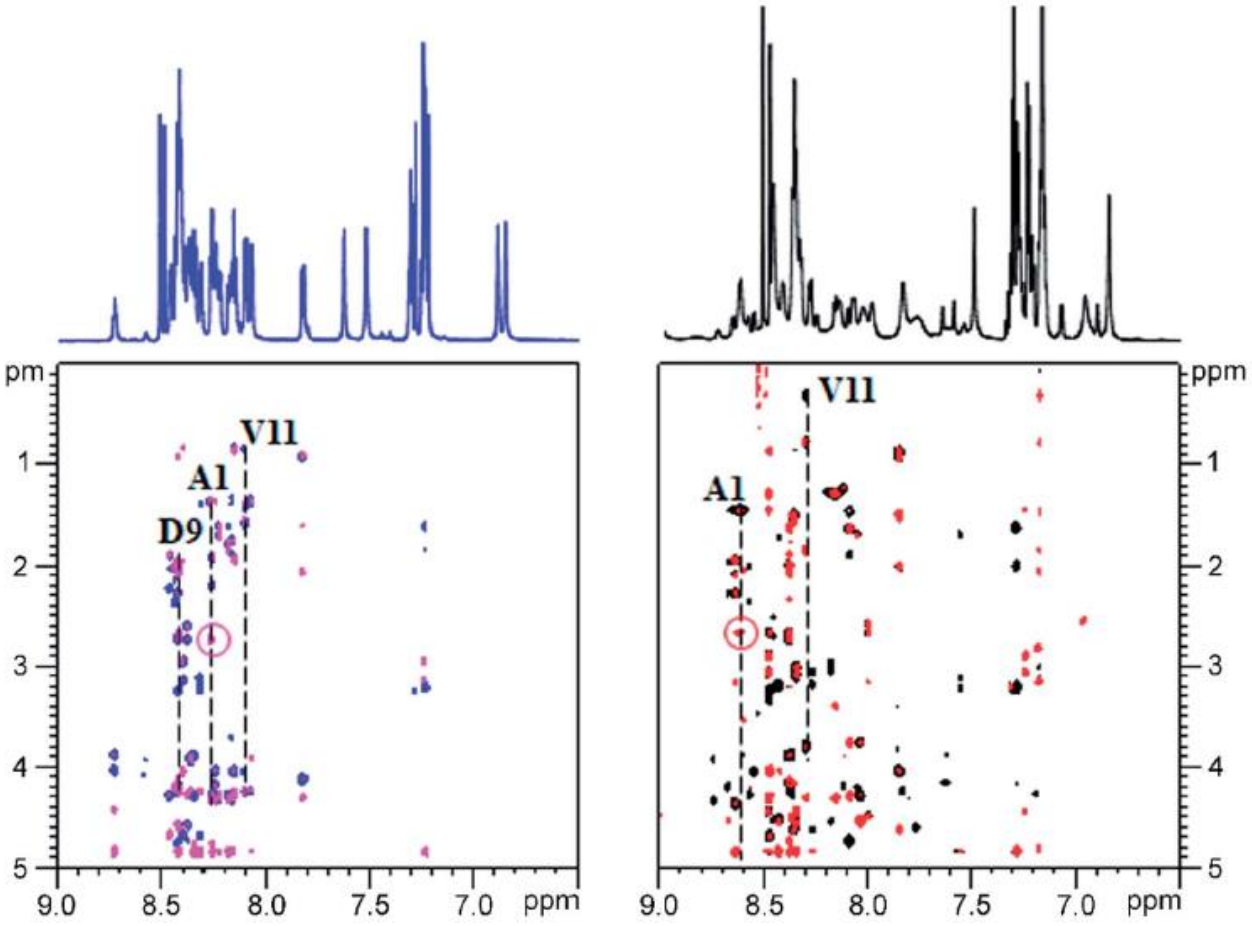


Figure 38. **1D spectra (top) and homonuclear TOCSY/NOESY spectra (bottom) of wild-type (left) and mutant L21F (right) pseudomycoidin.** The amide and methyl proton regions of the wild-type peptide spectrum are reduced when compared to those of the L21F mutant, and peaks are sharper in the former. The characteristic NOE contact between A1 HN and D9 Hb is circled in both NOESY spectra.

Although a full conformational analysis by NMR is beyond the scope of this work, these findings confirm the hypothesis of the C-terminus threading through the macrolactam ring when a bulky Phe occupies position 21. The tail is however not static, but slides back and forth through the ring, thereby leading to signal broadening. Overall, these data are consistent with threaded lasso topology of pseudomycoidin L21F and suggest that the branched lasso topology of wild-type pseudomycoidin is derived from threaded lasso after the tail is withdrawn from the ring during purification.

4.6. Antibacterial activity assay

E. coli cells that produced pseudomycoidin L21F and cells producing wild-type pseudomycoidin grew well and did not inhibit the growth of several Gram-positive and Gram-negative bacteria tested (laboratory strains of *E. coli B, B. subtilis* 168, or *Arthrobacter sp. FB24)* similarly to paeninodin, which was also devoid of any antibacterial activity [81]. The result suggests that the lasso topology in itself is not sufficient for antibacterial activity.

Conclusions

The principal result of this work is the demonstration that a lasso peptide can be formed *in vivo* in the absence of B1/B2 enzymes that heretofore were thought to be a universal component of lasso peptide biosynthesis machinery [51], [72], [130]. In systems that have been studied, the B1 protein recognizes the lasso peptide precursor leader sequence and then recruits the B2 protease to cleave at the leader–core junction and generate an N-terminus, which is used by the C macrolactam synthetase to form an amide bond with a side chain of a Glu or Asp residue in the core.

In the case of pseudomycoidin, this requirement must be bypassed through involvement of a cellular protease(s) that cleaves at the leader–core junction. The necessary enzymes are present in either *E. coli* or *B. subtilis* since both surrogate hosts produce pseudomycoidin in the absence of B1/B2 proteins. Given the phylogenetic distance between *E. coli* and *B. subtilis*, it is likely that the PsmA precursor has a specific structure, which exposes the leader–core junction to proteolytic attack by divergent proteases.

In systems that have been investigated, the B and C enzymes are thought to interact with the core part, once generated, being passed over to the macrolactam synthetase [17], [130]. In the case of pseudomycoidin, the situation must be different since the C protein must be able to convert the peptide part to a lasso in the absence of any additional factors, as exemplified by the efficient synthesis of pseudomycoidin in the presence of the PsmA core peptide and PsmC only. This work also reveals the remarkable tolerance of PsmC lasso-cyclase to the amino acid substitution, at least in the tail region.

The idea of universality of requirement of the B1/B2 enzymes (sometimes fused into a single B protein) may have originated due to a bias during bioinformatics searches used to reveal lasso peptide clusters. In fact, we identified a putative lasso peptide cluster from a cyanobacterium *Stanieria cyanosphaera* PCC 7437 (core genes locus tags STA7437_RS03725 and

STA7437_RS03720) that does not encode B proteins, while coding for an apparently functional macrolactam synthetase, a precursor peptide, and a PsmK kinase homolog. It is thus possible that a number of lasso peptide clusters were overlooked during prior searches.

We here show that *psmN* is required for installation of unique mono and/or dihexose modification on phosphorylated pseudomycoidin. The exact nature of hexose modification of pseudomycoidin is beyond the scope of the present study but the results suggest that PsmN is either a glycosyl transferase of a novel kind, since it does not contain recognizable motifs characteristic of known enzymes of this class or, alternatively, is an enzyme that produces a substrate used by one of glucosyl transferases present in the *E. coli* host. Interestingly, the cyanobacterial cluster mentioned above encodes a clearly recognizable (unlike PsmN) glycosyltransferase that may perform a similar function. Be as it may, the psmN-like genes are common in many lasso peptide clusters and must play an evolutionary conserved essential function. It is possible that the bulky hexose residues appended to phosphorylated C termini of pseudomycoidin and related peptides may serve as a lock to maintain the threaded lasso conformation. If true, this will be an additional, third strategy for stabilizing this topology.

The unique fold and exceptional stability make lasso peptides an attractive scaffold for development of non-natural bioactive compounds. Genetic engineering reveals that the core part of the precursor peptide can tolerate significant alterations without preventing recognition and processing by the biosynthesis machinery *in vivo* [130], [150].

The availability of the minimal heterologous pseudomycoidin *in vivo* production system should make it straightforward to generate libraries of various pseudomycoidin derivatives and understand the determinants of PsmC specificity.

Supporting information

Table S1. The nucleic acid sequences of the genes, consisting of the *psm* cluster.

psmA	atgaaacaagaatggcaatcaccagtattagaagtgcttgatatcaacatgacaatggctggacctggaaagagacttgttgaccaagtttttgaag atgaggatgagcaaggcgcacttcaccacaactaa
psmB1	atgagtagtaaacaaacaatttcattacacagttttgttgtacaaggtcaggaaaatgtagtgagcgatatggatggggagaaagtgatgatgagcat ccataatgggaaatactataacttgggaggtataggtggtgaaatttggagtttaattaatgaattaatatctgtgaatcaagtgattgatatattattgtct cgatatatgattgaagaggcagaatgtaaggaacaggtactctctttcttaaatcatttgtatgcagaagggttaatctcggttgatgaaaaactttga
psmB2	atgggaacaaaattattacttatagaagctttcttttatttaggctgggctcgtattcttaaaagtataccgttttctaaagctgctcctatattaggaattca tatggcagaaacgtctttaagccatgatgaatcggaaaaggttacattaaggaaaatttctcaagcagtgcatatgatgagccgttatacattttggga aagccagtgtctcgttaaagctatcgctgcaatgaaaatgttggagaagcgtcaaattgaaagtacccctttacttagggaccgcgaaagatgagac aggaaatatgattgcccatgcgtggttacgtagcgggccctttatattactggcgtagaagaaatgaagagatttactgttgttggtaaatttgcaaaa aggatagggtaa
psmC	atgagtgcaataaccggaatttatcatttaaatggagaaccgataccgcctgaacataggattggaatcatgaaggatttatctcgttatccaggtgac aataccgatatatggcacaaggaaaatatttttttgggctcccatgcaaaatggataacatctgaatctattaatgaacaattaccttactatgattatga aacacagttggttataactgctgatgctattattgataatcgcaatgaattattcgatagattgcaagtagaaaaggtacatagaagagagatgcctga tagtaagttgattttactcgcttatcaaaagtggggagagcatgtaccaaagtatctaattggtgattttgcttttatgatttgggatgagagaaagcaatt actctttggtgcacgtgatttttcaggaagtcggacttttttacttttaccgcaatcaaaatcaatttgctttttgtactagcatacaaccgctgtttttctttacc atacgtcgggaaggaattgaatgagcaatggttagctgaatatttagcgatacctaatatgcttgatacaatagatgcattttcaacagtgcataaaaa tatagagcaagtccccccttctcatactatctccgtagttaagggaagagttagtattgcgaaatacgacacgttatttgtagaggagcagttgcaact caaatcaaatgaagattatgaagaggcgttcagggacgtatttgaaagcgcagtaacatcaaggttacgtactcatcgtaatgttggttctcatttaag tgggggattagattcaggtgctgttgctagctttgctgcgagagctctgcaaaaagagaacaaaaaattatatacattcagctcagtacctgcagaa ggatttactgattggactcgtggccatacgattgctgatgaaaagccatttatccaaactacagtacaacatgtagggaatattaaagggaattatttg aatttcccgggagaaaattctttctcagaaattgatgattggcttgagattatggagatgccatataaatttttcgaaaattccgtctgggtgaaaggaat ttatgaaaaagcccaacaacaagggattggggtgctcttaaatggagcgagaggaaatttcacaatctcttgggggccagctttagattattatgcg actttacttaaaaaattcagctggttgcagctttatcgtgaaattgatttgtatggtaagaacgttggggtcaaaaaatcgagaatactgaaagttgtag gagagaaagcatttccatcgataaaaaacatgttttcctcggaacagcaatacgaattgccaacgttagttagtccagaacttgcaaaacgtactgat gtatttagaaagcttgaagaacatggtgtgaaagtaacaggggcttctatcccaaatatatatgaaataagacagaaacattttgaacagctaaactt gtggaatacaacgggtacaagtggaacaaagctctccttgcgctatggtttacaagggcatgatccaacgaatgacttacgggtaattcgattctgc ttatcgttaccactcgaacaatttgttcaaaatggcttaaatcgtgcacttgtccggaggtccacaaaaggatttcttcctgacaaagtaagattaaatc aacgtgtgcggggaatacaggcagctgatgtggttcatcgcatgattccttcatggaaaatgtttatcgaggaacttcagcagcttagtaaagaccc

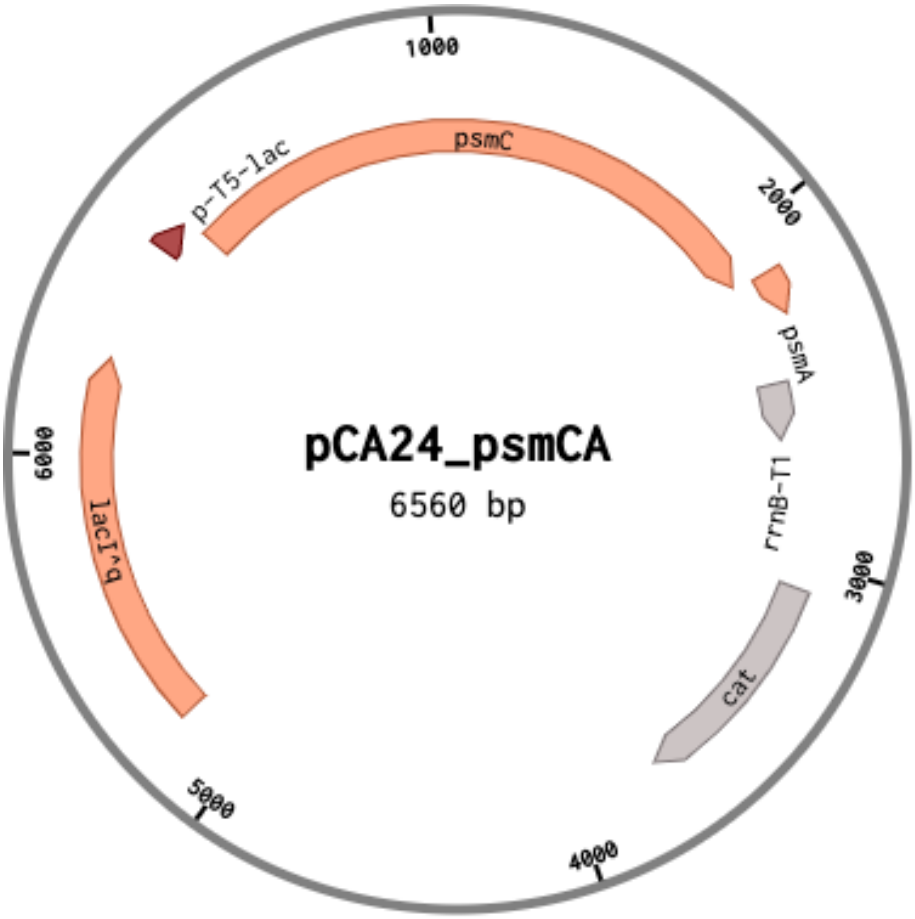
	agtggcctctaacttttttaatatagagattattaagagtgcgattttgaaaattcaaggagaaccacaagtagaatacgtgtttgatccggattttagaa ttttaatgcgtagtcttattgtctatcgatttatcaaaaaatatatttga
psmD	atgagaaatattttatattttacaaagcagttatactcttttactggaaacattctcattgttaatctcttgggaatggtattggtcagtcttgtagaagggat cggtattttactattaatcccaatgataagtattagtggtattgcgaagatgaatgaacaggcaagccctctttcaggggttcttggattcctaggtgag tttccaactactttaggattaccacttgtattaggaatatatattatattggtcgtgggacaagcgcttttgcaaaggaatataacagttcgtaatgcgaa aatccagatacgttttattaatcatttgagattggaaacttatcgtggattgttgcaggccaactggagttttttgtgaaaaagagaaaatcagatcttat caattctttgacaacggaattagggcgtgttagcgcaggcactaatttatttttacagttgctcacttctttaatgtttacacttatacagattggtttcgcttt ttggttagcggcagatataacaatttttgtactatgttgcgctgtggtacttgcctttctttcacgaagatttattaaaaaatcaaaaagattaggtggtcg aacctcagagaattcaaaagcgtatcttgcaggtataactgatcattttaatggaattaaggatattaagagcaatacactagaacgatctcgttatatg tggttacgtaattgggtggaagatgtggaacgtgaacagtttgcatatattaaggtgagaaataattctcaactcttttataaaattgcatcagctgttttc attgctattcttatctttctatctgttaagttgtttcagacgaaaacggagcaattattattaattattcttatattttctcgtttgtggccacgatttgcaggtat acaatcaaacatggaacaaattgcagcgaatcttcccgcttttaaagatttgctgcaattacaggaagagtgtaagaaaagtatagaattgcaagata tagagcaggattatgattgtataaatcctatgcagatagaacaatctatcgagtgtcagaatatctctttccgctacaataaagaaaatgctgtgtactc acttaaagatattaatttgcgtattcctgcaaattgtatgacagctatctcagggcgttctggtgctggaaagagtacattaattgatatcataatgggac ttatgcagccagagagcggtcagctattaatagataacaaaccgttgataagtaaggacctattatcattgagaaaatctattagctatgttcctcaag atccatttctattcaatgccagtataagagacaatttgagaatggttgaaccaaatgcaagtgaggaacagctttgggaagctttggaattttctgcttc cgctgagtttgttagtaggctctcacaaggattggatacaatgattggcgataggggagtcaaactttcaggggagaacgacaacgtcttgttctg gctagagcgatacttcgaaagccttctattttagtattagatgaggcaacaagtgcgttggatgctgaaaatgaagcgaaaatacaagaagcattgg aaagattgaaaggaagcatgacaattattgtcattgcacatagactatctactatacgtaaagctgaccaagtaattgtaatggatcaaggacaaatta ttcaacatggtagtttcacacaactagcgaaagatgaaagtggtttgtttagtagcctacttggacaacaggccaaagtaaacatttaa
psmA^{C}	atggctggacctggaaagagacttgttgaccaagtttttgaagatgaggatgagcaaggcgcacttcaccacagctaa

Table S2. The list of oligonucleotides used in the study.

psmA_NcoI_F	AATTAACCATGGCCATGAAACAAGAATGGCAATCACC
psmA_SacI_R	AATTAAGAGCTCTTAGCTGTGGTGAAGTGCGC
psmB1_NcoI_F	AATTAACCATGGCCATGAGTAGTAAACAAACAATTTCATTACAC
psmB1_BglII_RBS2_BamHI_R	AATTAAGGATCCTTCTCCTTAGATCTTCAAAGTTTTTCATCAACCGAGATT
psmB2_BamHI_F	AATTAAGGATCCATGAATATCATAAATAAAGCCAAAGC
psmB2_SalI_R	AATTAAGTCGACTTACCCTATCCTTTTTGCAAATTTACC
psmC_NdeI_F	AATTAACATATGAGTGCAATAACCGGAATTTATC
psmC_XhoI_R	AATTAACTCGAGTCAAATATATTTTTTGATAAATCGATAGAC
psmD_NdeI_F	AATTAACATATGAGAAATATTTTATATTTTACAAAGCAG
psmD_XhoI_R	AATTAACTCGAGTTAAATGTTTACTTTGGCCTGTTG
psmAc_NcoI_F	CATGGCTGGACCTGGAAAGAGACTTGTTGACCAAGTTTTTGAAGATGAGGATGAG CAAGGCGCACTTCACCACAGCTAAGAGCT
psmAc_SacI_R	CTTAGCTGTGGTGAAGTGCGCCTTGCTCATCCTCATCTTCAAAAACTTGGTCAACA AGTCTCTTTCCAGGTCCAGC
psmB1_F_NdeI	TTATATCATATGAGTAGTAAACAAACAA
psmB1_R_XhoI	TTATATCTCGAGTCAAAGTTTTTCATCAACCGA
psmB2_F_BamHI	TTATATGGATCCGATGAATATCATAAATAAAGC
psmB2_R_XhoI	TTATATCTCGAGTTACCCTATCCTTTTTGC
Psm_CA_FBgl	TTATAAGATCTTGGGAGGAAAAAAATGAGTGCAA
Psm_CA_RSal	TTATAGTCGACATAAATCTCATTTAAATATTAGTGCATT
PsmK_FSac	TTATAGAGCTCGGAGGAAAAAAATGATACATACTGTAAAGAAAGAGAT
PsmB1_FSAC	GAGGAGGAAAAATAATGAGTAGTAAACAA
PsmD_RSAL	TTATAGTCGACTTAAATGTTTACTTTGGCCTGT
deletion_N_F	ATGAGAAATATTTTATATTTTACAAAGCAGTTATAC
deletion_N_R	ATAACTGCTTTGTAAAATATAAAATATTTCTCATCTCATTCTCCTTCAAGATTTTT
deletion_B_F	AAAAATCTTGAAGGAGAATGAGATG
deletion_B_R	GTATACATCTCATTCTCCTTCAAGATTTTTTTATTTTTCCTCCTCTTTTAAAGACTCT AG
psmN_FB(BamhI)	TTATAGGATCCGATGTATACCGGTTGTAATCTTGATT
psmN_RX (XhoI)	TTATATCTCGAGTTACTTCTCCTTACGGTACCTCATTCC
psmp_F_bamHi	ATGGATCCATGATACATACTGTAAAGAAAGAG
psmp_R_xhoi	ATAATTCTCGAGTTATTTTTCCTCCTCTTTTAAAGACTC

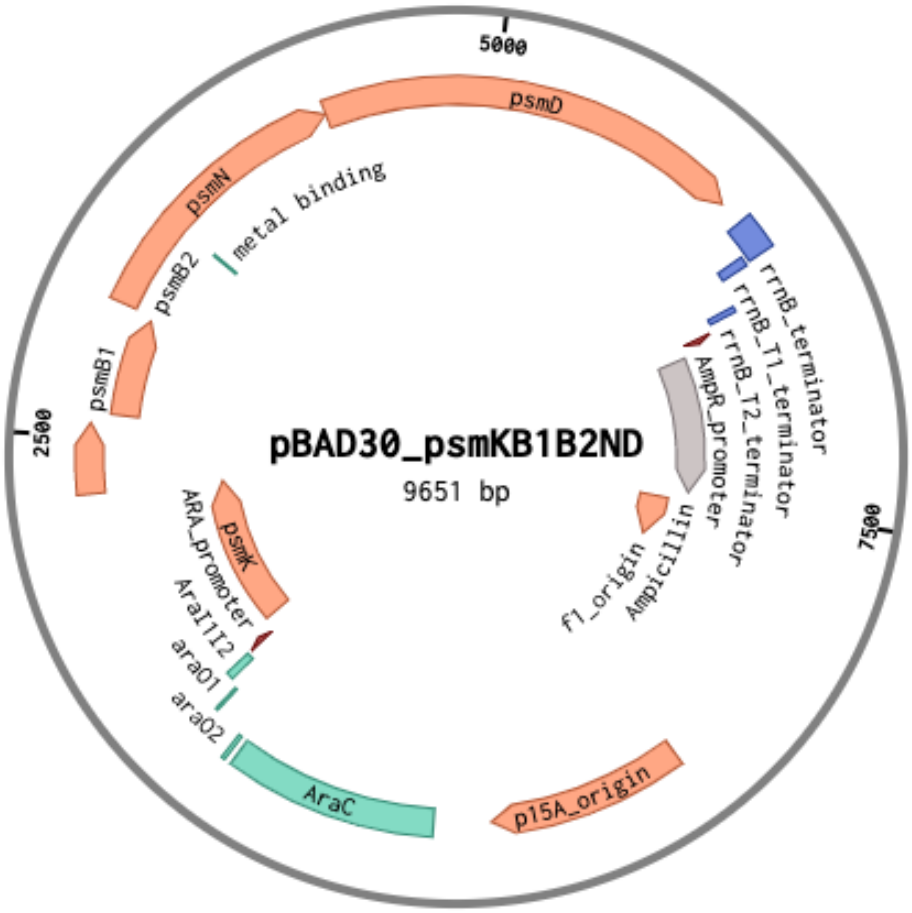
psmN_D134-6_F	TCTTCGAACATCAGGGGCGCTGGCGGCCCTTGTTCCTATAGA
psmN_D134-6_R	TCTATAGGAACAAGGGCCGCCAGCGCCCCTGATGTTCGAAGA
psmP_D158-9_F	GCTATCAGCTTTTAAGCGCGGCGGTAATTGCTGTGTCTCTT
psmP_D158-9_R	AAGAGACACAGCAATTACCGCCGCGCTTAAAAGCTGATAGC
PsmA_RMut	AATTAAGAGCTCTTAGCTGTGGTGAAATGCGCCTTGCTCAT
psmB2_R_BamHI	TTATATGGATCCTTACCCTATCCTTTTTGC
pHT_CA_F_seq	ATTAAAGGATCCTATGAGTGCAATAACCGGAATT
pHT_CA_R_seq	AATATTACTAGTATTAGCTGTGGTGAAGTGCGCC
pHT_Pgrac_CA_F	ATTATATGAATTCGGTACCAGCTATTGTAACATAATCG
pHT_Pgrac_CA_R	TTATTATGAATTCTGAACATCAAATCGCTTTATTCTTT
pHT_psmD_BamHI	AATATTGGATCCATGAGAAATATTTTATATTTTACAAAGCAGTTATACT
pHT_psmD_SpeI_R	AATATTACTAGTGTTAAATGTTTACTTTGGCCTGTTG

The maps of constructed plasmids



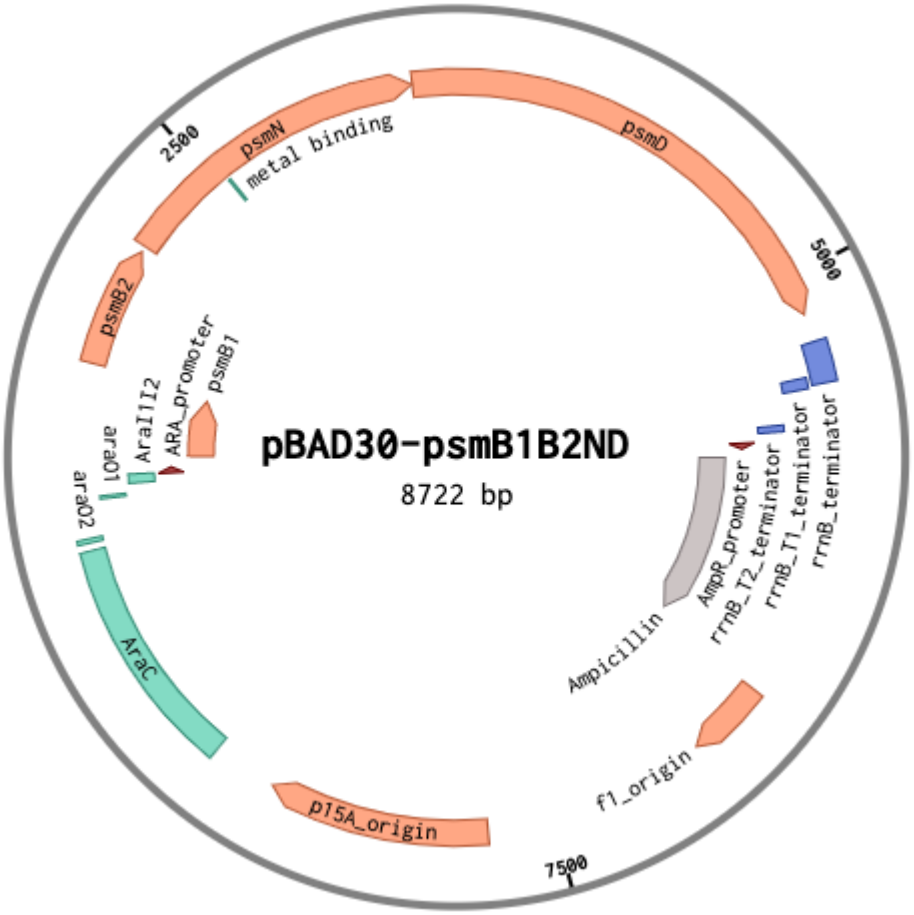
The *psmCA* genes were PCR-amplified with **Psm_CA_FBgl** & **Psm_CA_RSal** oligoes and introduced using the Gibson assembly kit (NEB, USA) into the

pCA24 commercial plasmid vector restricted with Bgl and Sal endonuclease.



The *psmKB1B2ND* genes were PCR-amplified with **PsmK_FSac** & **PsmD_RSAL** oligoes and introduced into the

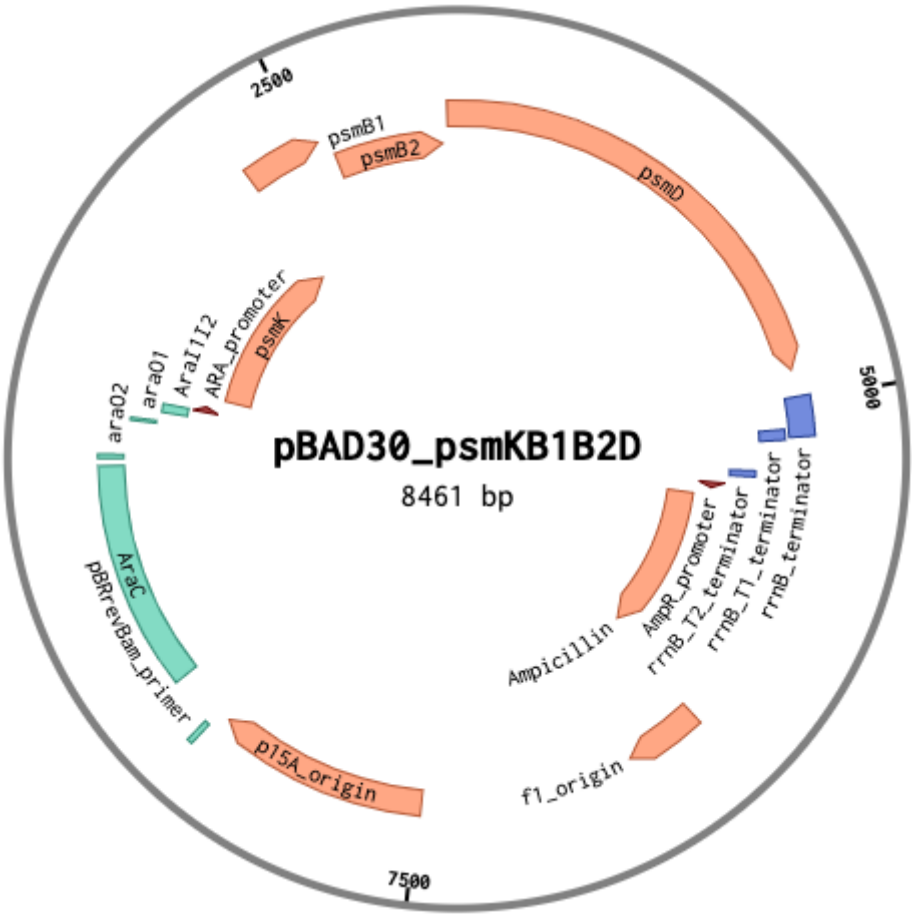
pBAD30 commercial plasmid vector with SacI and SalI endonucleases.



The deletion of *psmK*

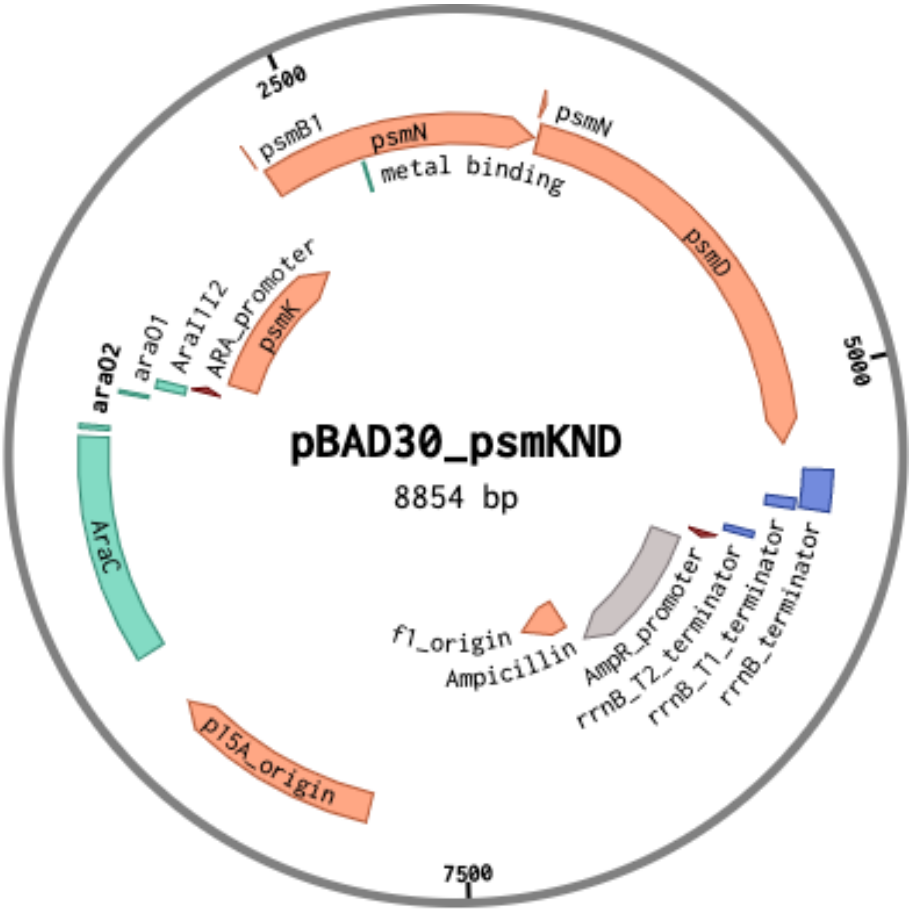
The *psmB1B2ND* genes were PCR-amplified with **PsmB1_FSAC** & **PsmD_RSAL** oligoes and introduced into the

pBAD30 commercial plasmid vector with SacI and SalI endonucleases.



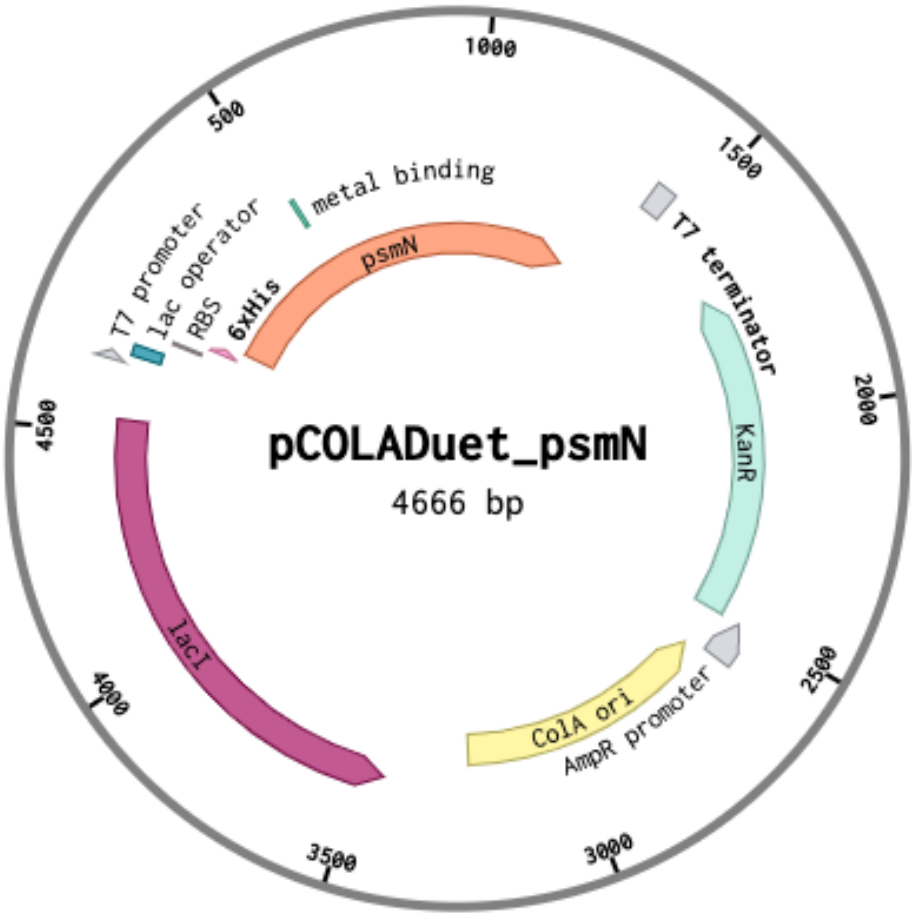
The deletion of *psmN*

The pBAD-*psmKB1B2ND* plasmid vector was PCR-amplified with **deletion_N_F** & **deletion_N_R** oligoes and ligated with Gibson assembly kit (NEB, USA).



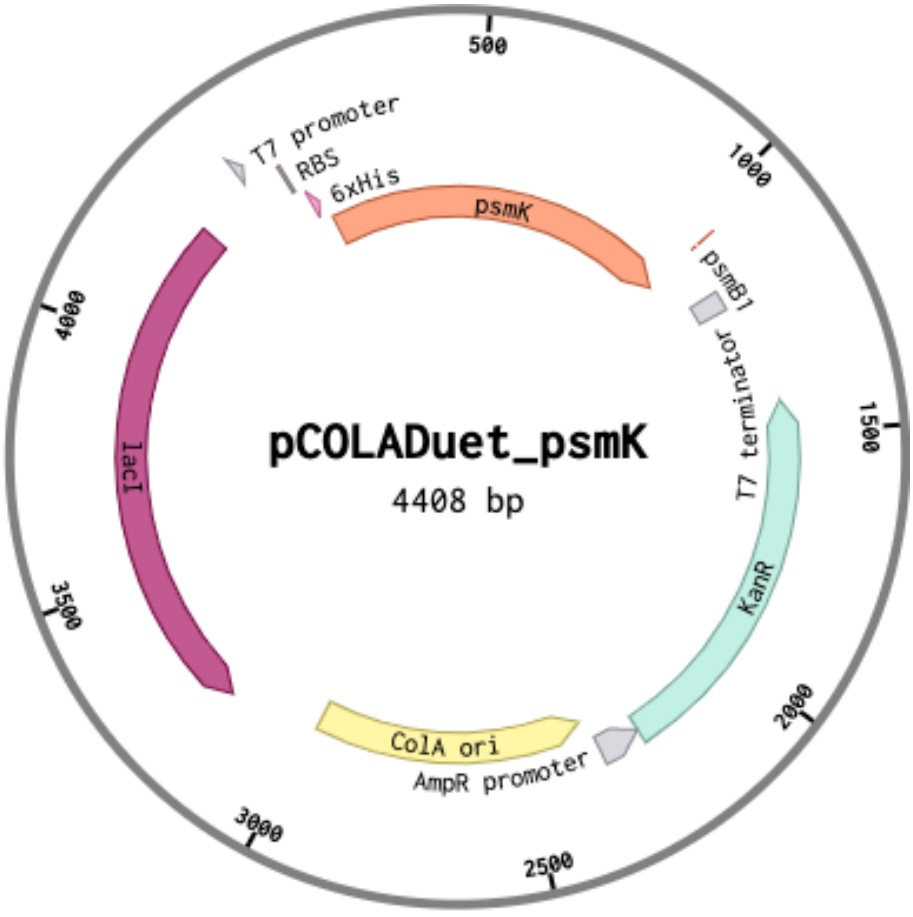
The deletion of *psmB1B2*

The pBAD-*psmKB1B2ND* plasmid vector was PCR-amplified with **deletion_B_F** & **deletion_B_R** oligoes and ligated with Gibson assembly kit (NEB, USA).



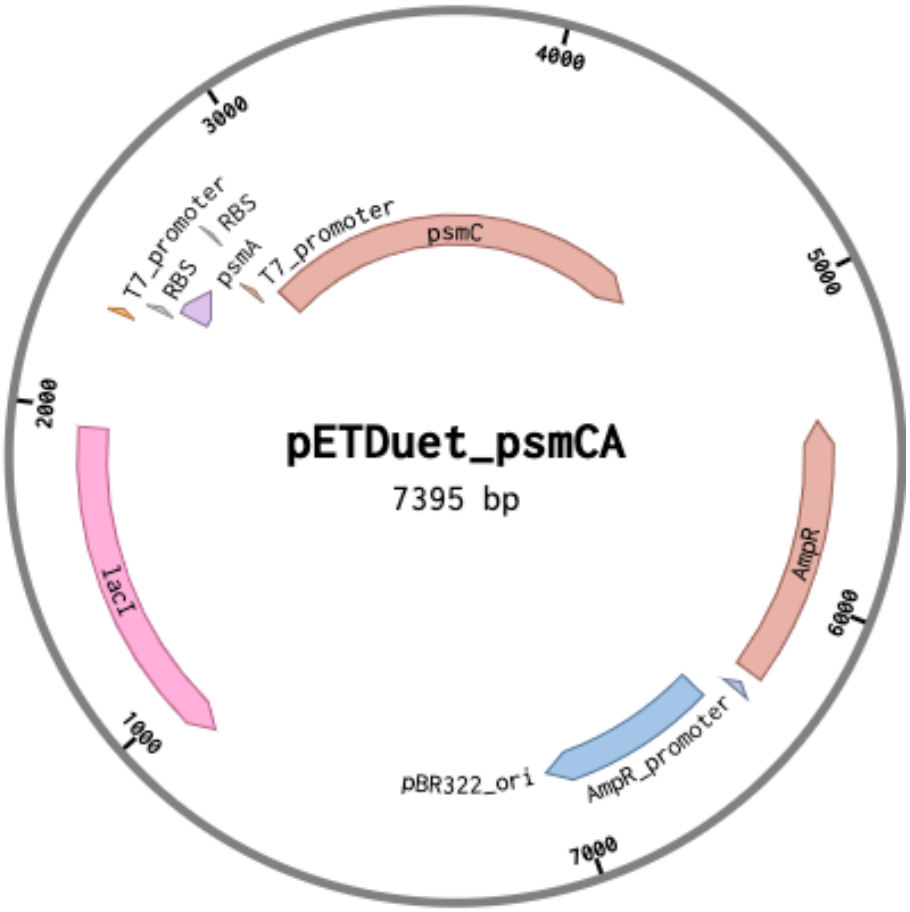
The *psmN* gene were PCR-amplified with **psmN_FB(BamhI)** & **psmN_RX (XhoI)**oligoes and introduced into the

pCOLADuet-1 commercial plasmid vector with BamHI and XhoI endonucleases.



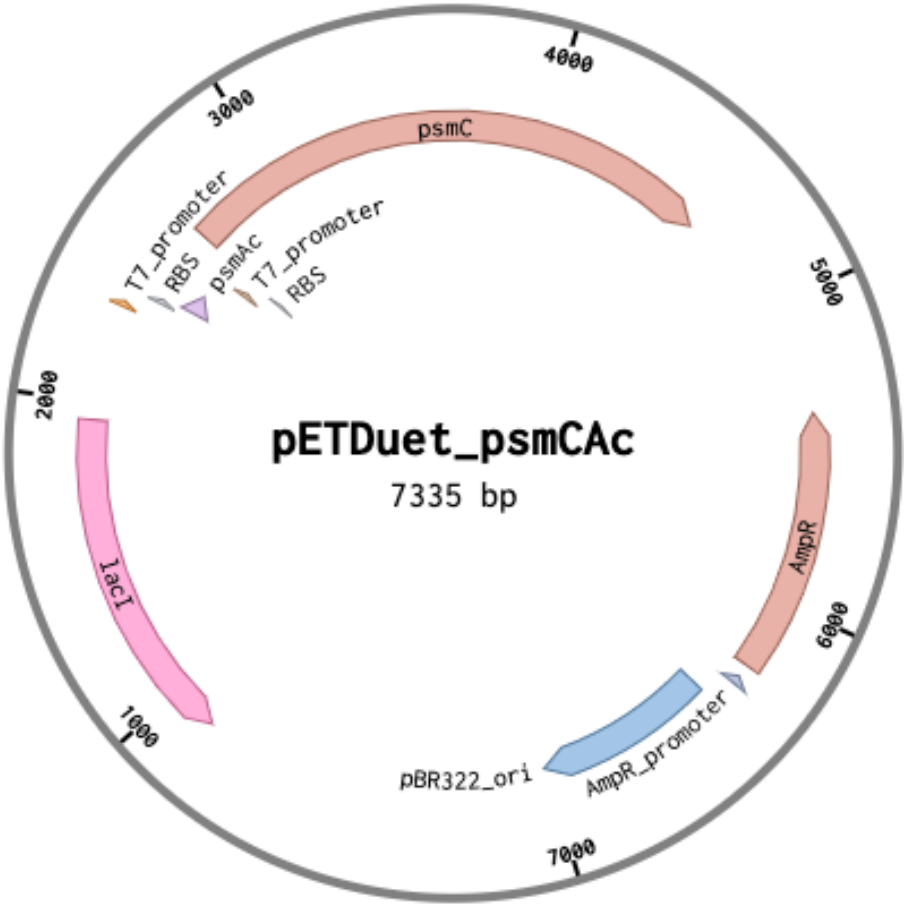
The *psmK* gene were PCR-amplified with **psmp_F_bamHi** & **psmp_R_xhoi** oligoes and introduced into the

pCOLADuet-1 commercial plasmid vector with BamHI and XhoI endonucleases.



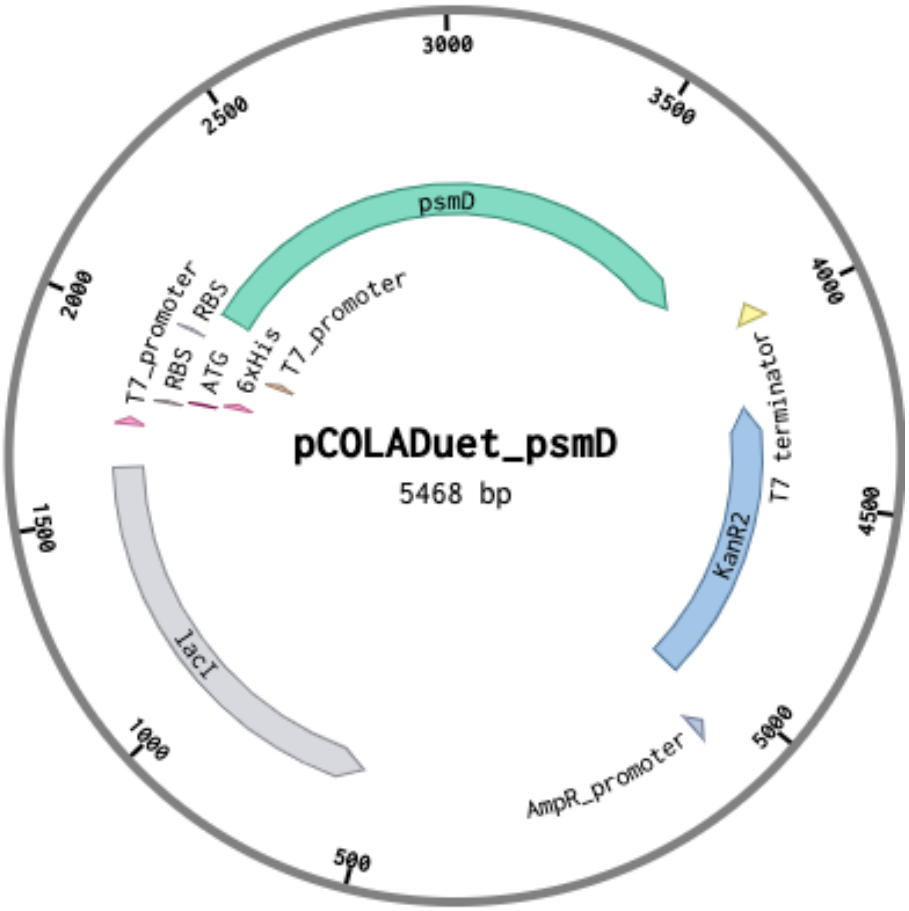
The *psmC* gene was PCR-amplified with **psmC_NdeI_F & psmC_XhoI_R** oligoes and introduced into the second multiple cloning site of the pETDuet-1 vector with NdeI and XhoI endonucleases.

The *psmA* gene was PCR-amplified with oligoes, digest with BspHI and SacI and introduced into the first multiple cloning site of the pETDuet_ *psmC* with NcoI and SacI endonucleases.

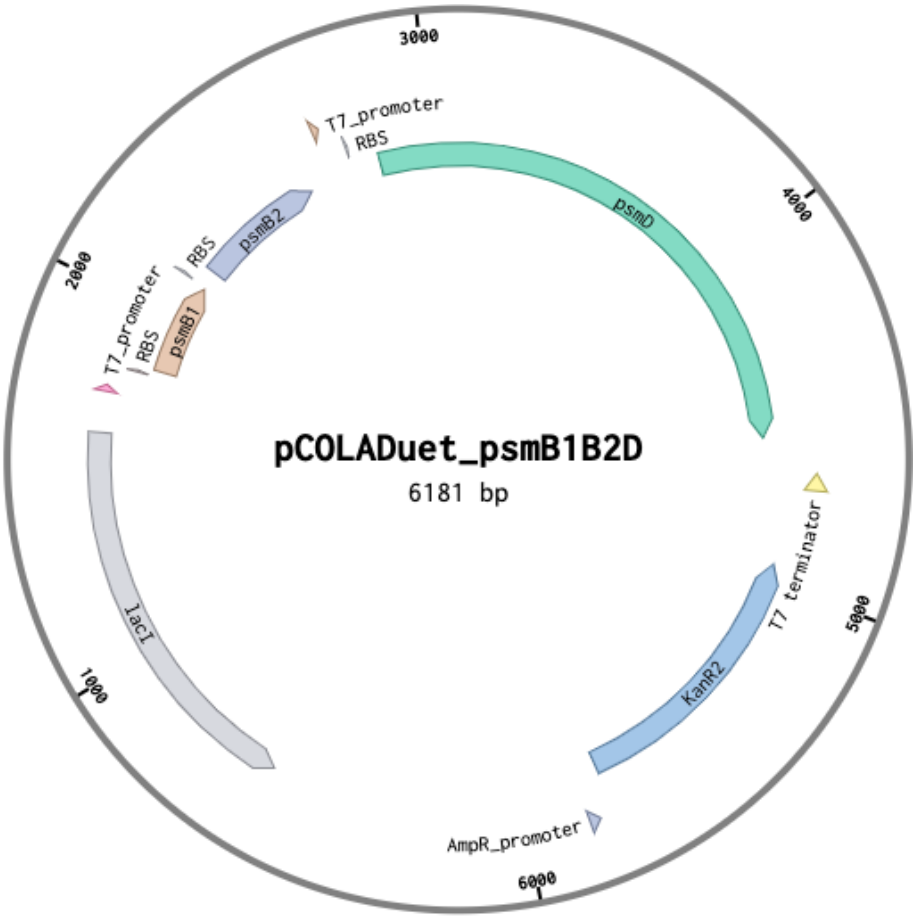


The *psmC* gene was PCR-amplified with **psmC_NdeI_F & psmC_XhoI_R** oligoes and introduced into the second multiple cloning site of the pETDuet-1 vector with NdeI and XhoI endonucleases.

The *psmA*^{C} gene was obtained by annealing of **psmAc_NcoI_F** & **psmAc_SacI_R** oligoes and introduced into the first multiple cloning site of the pETDuet_ *psmC* with NcoI and SacI endonucleases



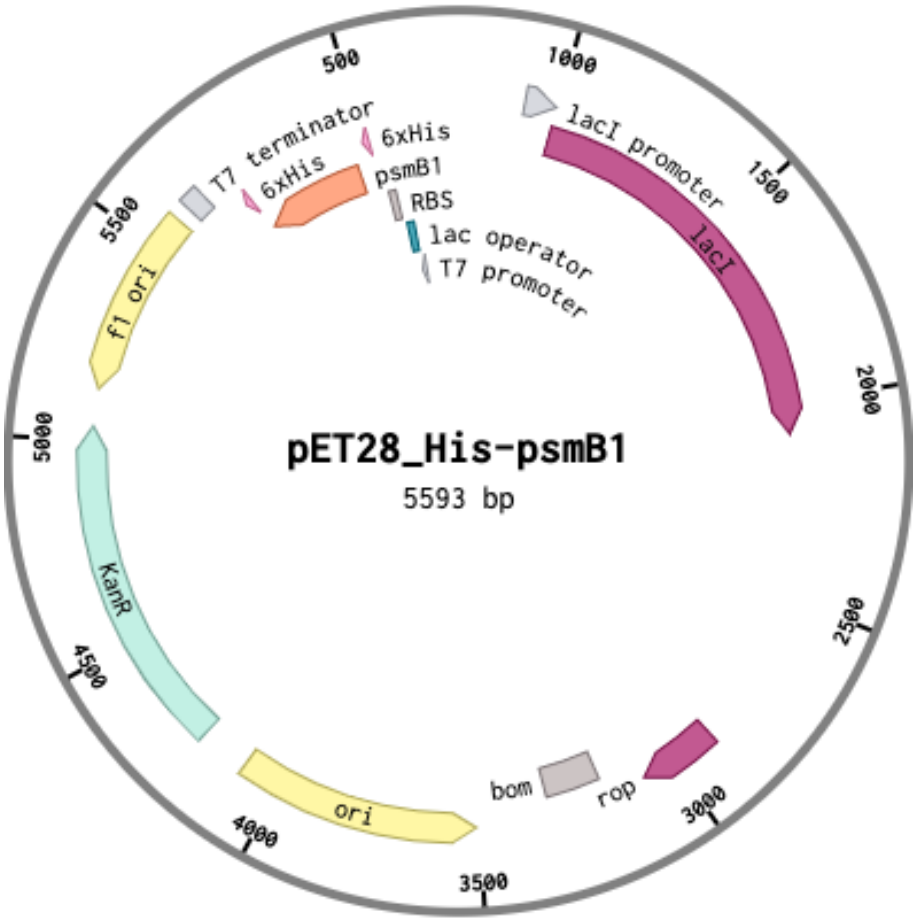
The *psmD* gene was PCR-amplified with **psmD_NdeI_F** & **psmD_XhoI_R** oligoes and introduced into the second multiple cloning site of the pCOLADuet-1 vector with NdeI and XhoI endonucleases.



The *psmD* gene was PCR-amplified with **psmD_NdeI_F** & **psmD_XhoI_R** oligoes and introduced into the second multiple cloning site of the pCOLADuet-1 vector with NdeI and XhoI endonucleases.

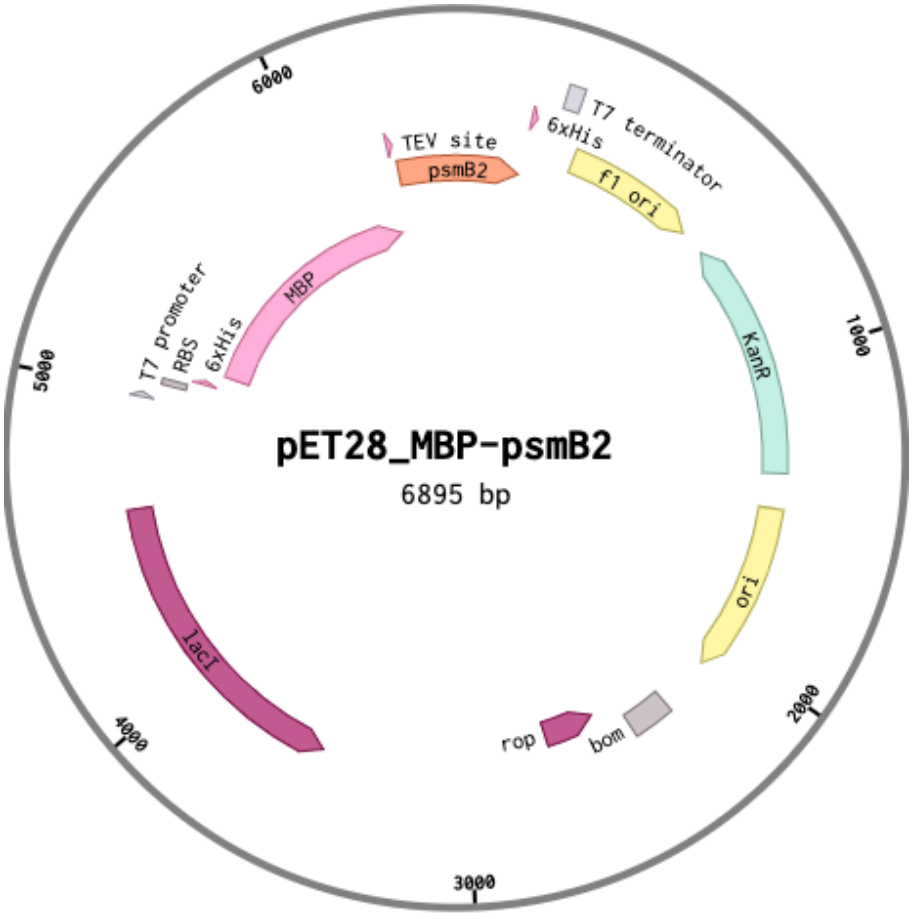
The *psmB1* gene was PCR-amplified with **psmB1_NcoI_F** & **psmB1_BglII_RBS2_BamHI_R** oligoes and introduced into the first multiple cloning site of the pCOLADuet_ *psmD* with NcoI and BamHI endonucleases.

The *psmB2* gene was PCR-amplified with **psmB2_BamHI_F & psmB2_SalI_R** oligoes and introduced into the pCOLADuet_ *psmB1D* with BamHI and SalI endonucleases.



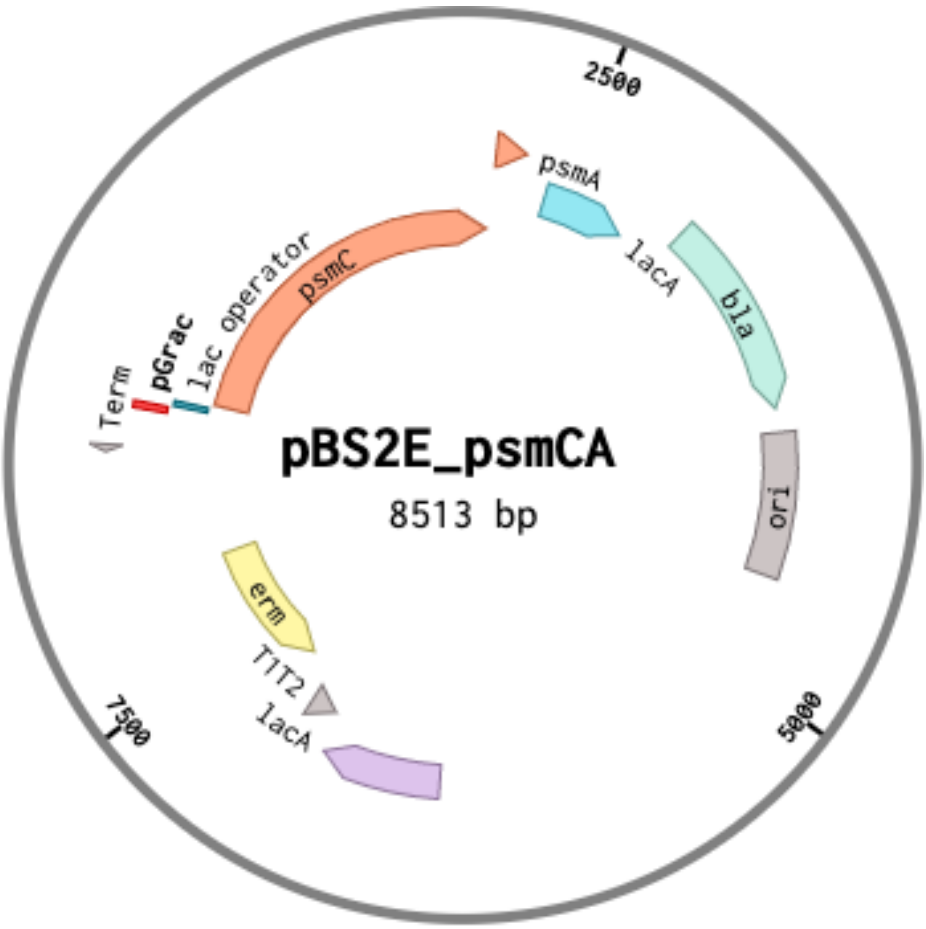
The *psmB1* gene were PCR-amplified with **psmB1_F_NdeI** & **psmB1_R_XhoI** oligoes and introduced into the

pET28 commercial plasmid vector with NdeI and XhoI endonucleases.



The *psmB2* gene was PCR-amplified with **psmB2_F_BamHI** & **psmB2_R_XhoI** oligoes and introduced into the

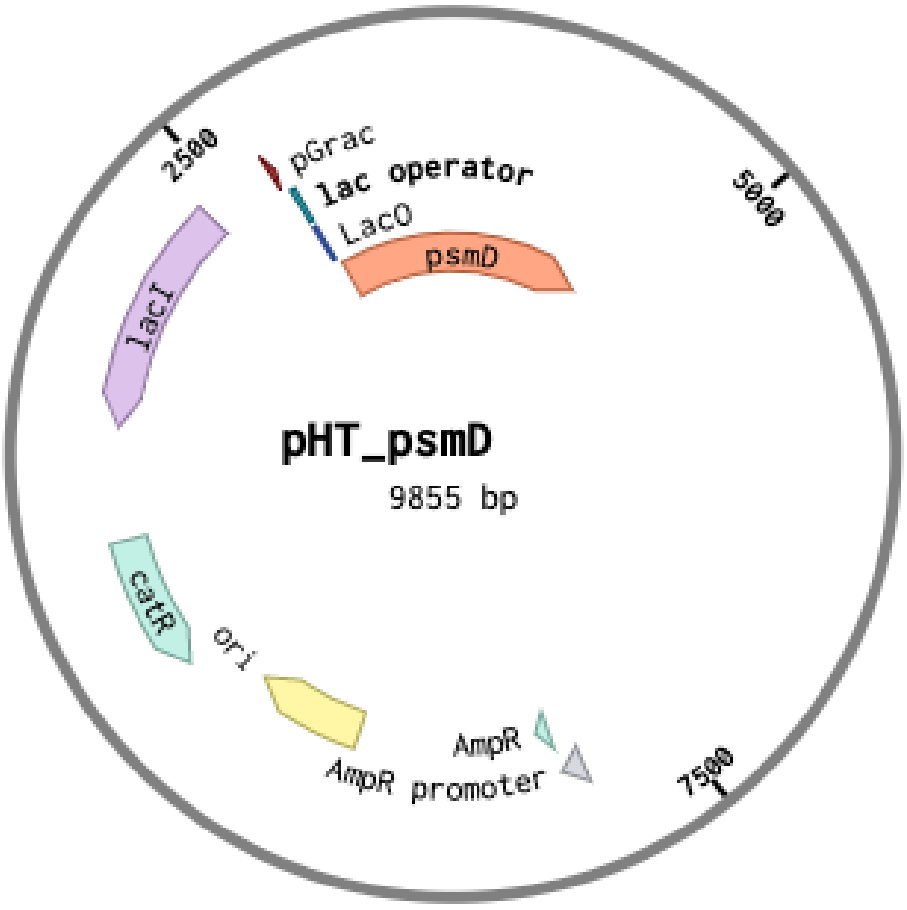
pET28_MBP commercial plasmid vector with BamHI and XhoI endonucleases



The pBS2E was restricted using EcoRI endonuclease and ligated with

psmCA operon flanked by P*grac-lac* promoter and the terminator amplified by PCR with **pHT_Pgrac_CA_F** & **pHT_Pgrac_CA_R** oligoes and the

pHT-*psmCA* vector template (obtained by ligation of *psmCA* genes amplified with **pHT_CA_F_seq** & **pHT_CA_R_seq** oligoes and pHT-01 shuttle vector with BamHI and SpeI endonucleases)



The *psmD* gene was PCR-amplified with **pHT_psmD_BamHI** & **pHT_psmD_SpeI_R** oligoes and introduced into the

pHT-01 commercial plasmid vector with BamHI and SpeI endonucleases.

5. References

[1] Y. Xue *et al.*, “Gram-negative bacilli-derived peptide antibiotics developed since 2000,” *Biotechnol. Lett.*, vol. 40, no. 9–10, pp. 1271–1287, 2018.

[2] S. Duquesne, D. Destoumieux-Garzón, J. Peduzzi, and S. Rebuffat, “Microcins, gene-encoded antibacterial peptides from enterobacteria,” *Nat. Prod. Rep.*, vol. 24, no. 4, pp. 708–734, 2007.

[3] J. O. Melby, N. J. Nard, and D. A. Mitchell, “Thiazole/oxazole-modified microcins: Complex natural products from ribosomal templates,” *Curr. Opin. Chem. Biol.*, vol. 15, no. 3, pp. 369–378, 2011.

[4] M. Körner and J. C. Reubi, *Handbook of Biologically Active Peptides*. 2006.

[5] M. Metelev *et al.*, “Klebsazolicin inhibits 70S ribosome by obstructing the peptide exit tunnel,” *Nat. Chem. Biol.*, vol. 13, no. 10, pp. 1129–1136, 2017.

[6] P. G. Arnison *et al.*, “Ribosomally synthesized and post-translationally modified peptide natural products: overview and recommendations for a universal nomenclature,” *Nat. Prod. Rep.*, vol. 30, no. 1, pp. 108–160, 2013.

[7] S. Rebuffat, “Microcins in action: Amazing defence strategies of Enterobacteria,” *Biochem. Soc. Trans.*, vol. 40, no. 6, pp. 1456–1462, 2012.

[8] X. Just-Baringo, F. Albericio, and M. Álvarez, “Thiopeptide antibiotics: Retrospective and recent advances,” *Mar. Drugs*, vol. 12, no. 1, pp. 317–351, 2014.

[9] P. Zhao *et al.*, “Actinobacteria–Derived peptide antibiotics since 2000,” *Peptides*, vol. 103, pp. 48–59, 2018.

[10] K. L. Dunbar and D. A. Mitchell, “Revealing nature’s synthetic potential through the study of ribosomal natural product biosynthesis,” *ACS Chem. Biol.*, vol. 8, no. 3, pp. 473–487, 2013.

[11] K. Severinov, E. Semenova, A. Kazakov, T. Kazakov, and M. S. Gelfand, “Low-molecular-weight post-translationally modified microcins,” *Mol. Microbiol.*, vol. 65, no. 6, pp. 1380–1394, 2007.

[12] K. Beis and S. Rebuffat, “Multifaceted ABC transporters associated to microcin and bacteriocin export,” *Res. Microbiol.*, no. xxxx, 2019.

[13] M. A. Skinnider *et al.*, “Genomic charting of ribosomally synthesized natural product chemical space facilitates targeted mining,” *Proc. Natl. Acad. Sci. U. S. A.*, vol. 113, no. 42, pp. E6343–E6351, 2016.

[14] J. P. Gomez-Escribano, L. Song, M. J. Bibb, and G. L. Challis, “Posttranslational β-methylation and macrolactamidination in the biosynthesis of the bottromycin complex of ribosomal peptide antibiotics,” *Chem. Sci.*, vol. 3, no. 12, pp. 3522–3525, 2012.

[15] T. J. Oman and W. A. Van Der Donk, “Follow the leader: The use of leader peptides to guide natural product biosynthesis,” *Nat. Chem. Biol.*, vol. 6, no. 1, pp. 9–18, 2010.

[16] S. J. Pan, J. Rajniak, M. O. Maksimov, and A. J. Link, “The role of a conserved threonine residue in the leader peptide of lasso peptide precursors,” *Chem. Commun.*, vol. 48, no. 13, pp. 1880–1882, 2012.

[17] W. L. Cheung, M. Y. Chen, M. O. Maksimov, and A. J. Link, “Lasso peptide biosynthetic protein LarB1 binds both leader and core peptide regions of the precursor protein LarA,” *ACS Cent. Sci.*, vol. 2, no. 10, pp. 702–709, 2016.

[18] J. Koehnke *et al.*, “Structural analysis of leader peptide binding enables leader-free cyanobactin processing,” *Nat. Chem. Biol.*, vol. 11, no. 8, pp. 558–563, 2015.

[19] M. A. Ortega *et al.*, “Substrate specificity of the lanthipeptide peptidase ElxP and the oxidoreductase ElxO,” *ACS Chem. Biol.*, vol. 9, no. 8, pp. 1718–1725, 2014.

[20] Y. M. Li, J. C. Milne, L. L. Madison, R. Kolter, and C. T. Walsh, “From peptide precursors to oxazole and thiazole-containing peptide antibiotics: Microcin B17 synthase,” *Science (80-.).*, vol. 274, no. 5290, pp. 1188–1193, 1996.

[21] J. Camadro, *Antimicrobial peptides from enterobacteria*. 2007.

[22] A. Chakraborty *et al.*, *Trapping effect analysis of*

AlGaN/InGaN/GaN Heterostructure by conductance frequency measurement, vol. XXXIII, no. 2. 2014.

[23] K. Sivonen, N. Leikoski, D. P. Fewer, and J. Jokela, "Cyanobactins-ribosomal cyclic peptides produced by cyanobacteria," *Appl. Microbiol. Biotechnol.*, vol. 86, no. 5, pp. 1213–1225, 2010.

[24] P. J. Knerr and W. A. van der Donk, "Discovery, Biosynthesis, and Engineering of Lantipeptides," *Annu. Rev. Biochem.*, vol. 81, no. 1, pp. 479–505, 2012.

[25] K. J. Hetrick and W. A. van der Donk, "Ribosomally synthesized and post-translationally modified peptide natural product discovery in the genomic era," *Curr. Opin. Chem. Biol.*, vol. 38, pp. 36–44, 2017.

[26] J. L. San Millan, R. Kolter, and F. Moreno, "Plasmid genes required for microcin B17 production," *J. Bacteriol.*, vol. 163, no. 3, pp. 1016–1020, 1985.

[27] M. Lavina, A. P. Pugsley, and F. Moreno, "Identification, mapping, cloning and characterization of a gene (sbm A) required for microcin B17 action on Escherichia coli K12," *J. Gen. Microbiol.*, vol. 132, no. 6, pp. 1685–1693, 1986.

[28] M. R. Baquero, M. Bouzon, J. Varea, and F. Moreno, "sbmC, a stationary-phase induced SOS Escherichia coli gene, whose product protects cells from the DNA replication inhibitor microcin B17," *Mol. Microbiol.*, vol. 18, no. 2, pp. 301–311, 1995.

[29] J. L. Vizán, C. Hernández-Chico, I. del Castillo, and F. Moreno, "The peptide antibiotic microcin B17 induces double-strand cleavage of DNA mediated by E. coli DNA gyrase.," *EMBO J.*, vol. 10, no. 2, pp. 467–476, 1991.

[30] W. M. Parks, A. R. Bottrill, O. A. Pierrat, M. C. Durrant, and A. Maxwell, "The action of the bacterial toxin, microcin B17, on DNA gyrase," *Biochimie*, vol. 89, no. 4, pp. 500–507, 2007.

[31] A. Bayer, S. Freund, G. Nicholson, and G. Jung, "Posttranslational Backbone Modifications in the Ribosomal Biosynthesis of the Glycine-Rich Antibiotic Microcin B17," *Angew. Chemie Int. Ed. English*, vol. 32, no. 9, pp. 1336–1339, 1993.

[32] O. Genilloud, F. Moreno, and R. Kolter, “DNA sequence, products, and transcriptional pattern of the genes involved in production of the DNA replication inhibitor microcin B17,” *J. Bacteriol.*, vol. 171, no. 2, pp. 1126–1135, 1989.

[33] M. C. Garrido, M. Herrero, R. Kolter, and F. Moreno, “The export of the DNA replication inhibitor Microcin B17 provides immunity for the host cell.,” *EMBO J.*, vol. 7, no. 6, pp. 1853–1862, 1988.

[34] O. Lomovskaya, F. Kawai, and A. Matin, “Differential regulation of the mcb and emr operons of Escherichia coli: Role of mcb in multidrug resistance,” *Antimicrob. Agents Chemother.*, vol. 40, no. 4, pp. 1050–1052, 1996.

[35] A. Metlitskaya *et al.*, “Aspartyl-tRNA synthetase is the target of peptide nucleotide antibiotic microcin C,” *J. Biol. Chem.*, vol. 281, no. 26, pp. 18033–18042, 2006.

[36] K. Severinov and S. K. Nair, “Microcin C: Biosynthesis and mechanisms of bacterial resistance,” *Future Microbiol.*, vol. 7, no. 2, pp. 281–289, 2012.

[37] J. E. Gonzalez-Pastor, J. L. San Millan, M. A. Castilla, and F. Moreno, “Structure and organization of plasmid genes required to produce the translation inhibitor microcin C7,” *J. Bacteriol.*, vol. 177, no. 24, pp. 7131–7140, 1995.

[38] A. Z. Metlitskaya, G. S. Katrukha, A. S. Shashkov, D. A. Zaitsev, T. A. Egorov, and I. A. Khmel, “Structure of microcin C51, a new antibiotic with a broad spectrum of activity,” *FEBS Lett.*, vol. 357, no. 3, pp. 235–238, 1995.

[39] M. Serebryakova *et al.*, “A Trojan-Horse Peptide-Carboxymethyl-Cytidine Antibiotic from Bacillus amyloliquefaciens,” *J. Am. Chem. Soc.*, vol. 138, no. 48, pp. 15690–15698, 2016.

[40] M. Novikova *et al.*, “The Escherichia coli Yej transporter is required for the uptake of translation inhibitor microcin C,” *J. Bacteriol.*, vol. 189, no. 22, pp. 8361–8365, 2007.

[41] A. Metlitskaya *et al.*, “Maturation of the translation inhibitor microcin,” *J. Bacteriol.*, vol. 191, no. 7, pp. 2380–2387, 2009.

[42] M. O. Maksimov, I. Pelczer, and A. J. Link, “Precursor-centric

genome-mining approach for lasso peptide discovery," *Proc. Natl. Acad. Sci. U. S. A.*, vol. 109, no. 38, pp. 15223–15228, 2012.

[43] S. Rebuffat, A. Blond, D. Destoumieux-Garzon, C. Goulard, and J. Peduzzi, "Microcin J25, from the Macrocyclic to the Lasso Structure: Implications for Biosynthetic, Evolutionary and Biotechnological Perspectives," *Curr. Protein Pept. Sci.*, vol. 5, no. 5, pp. 383–391, 2005.

[44] K. A. Wilson *et al.*, "Structure of microcin J25, a peptide inhibitor of bacterial RNA polymerase, is a lassoed tail," *J. Am. Chem. Soc.*, vol. 125, no. 41, pp. 12475–12483, 2003.

[45] D. J. Clarke and D. J. Campopiano, "Maturation of McjA precursor peptide into active microcin MccJ25," *Org. Biomol. Chem.*, vol. 5, no. 16, pp. 2564–2566, 2007.

[46] W. L. Cheung, S. J. Pan, and A. J. Link, "Much of the microcin J25 leader peptide is dispensable," *J. Am. Chem. Soc.*, vol. 132, no. 8, pp. 2514–2515, 2010.

[47] P. Vincent and R. Morero, "The Structure and Biological Aspects of Peptide Antibiotic Microcin J25," *Curr. Med. Chem.*, vol. 16, no. 5, pp. 538–549, 2009.

[48] C. L. Cox, J. R. Doroghazi, and D. A. Mitchell, "The genomic landscape of ribosomal peptides containing thiazole and oxazole heterocycles," *BMC Genomics*, vol. 16, no. 1, pp. 1–16, 2015.

[49] J. I. Tietz *et al.*, "A new genome-mining tool redefines the lasso peptide biosynthetic landscape," *Nat. Chem. Biol.*, vol. 13, no. 5, pp. 470–478, 2017.

[50] S. W. Lee *et al.*, "Discovery of a widely distributed toxin biosynthetic gene cluster," *Proc. Natl. Acad. Sci. U. S. A.*, vol. 105, no. 15, pp. 5879–5884, 2008.

[51] A. J. Dicaprio, A. Firouzbakht, G. A. Hudson, and D. A. Mitchell, "Enzymatic Reconstitution and Biosynthetic Investigation of the Lasso Peptide Fusilassin," *J. Am. Chem. Soc.*, vol. 141, no. 1, pp. 290–297, 2019.

[52] W. L. Cheung-Lee and A. J. Link, "Genome mining for lasso peptides: past, present, and future," *J. Ind. Microbiol. Biotechnol.*,

vol. 46, no. 9–10, pp. 1371–1379, 2019.

[53] R. J. Young, “Physical Propreties in Drug Desin,” *Top Med Chem*, vol. 9, no. July, pp. 1–68, 2014.

[54] X. W. Wang and W. Bin Zhang, “Chemical Topology and Complexity of Protein Architectures,” *Trends Biochem. Sci.*, vol. 43, no. 10, pp. 806–817, 2018.

[55] M. Metelev *et al.*, “Acinetodin and Klebsidin, RNA Polymerase Targeting Lasso Peptides Produced by Human Isolates of Acinetobacter gyllenbergii and Klebsiella pneumoniae,” *ACS Chem. Biol.*, vol. 12, no. 3, pp. 814–824, 2017.

[56] S. Kodani *et al.*, “Sphaericin, a Lasso Peptide from the Rare Actinomycete Planomonospora sphaerica,” *European J. Org. Chem.*, vol. 2017, no. 8, pp. 1177–1183, 2017.

[57] J. D. Hegemann, M. Zimmermann, X. Xie, and M. A. Marahiel, “Lasso Peptides: An Intriguing Class of Bacterial Natural Products,” *Acc. Chem. Res.*, vol. 48, no. 7, pp. 1909–1919, 2015.

[58] S. Microbiology, *Lasso Peptides Bacterial Strategies to Make and Maintain Bioactive Entangled Scaffolds*. Microbiology, Springerbriefs, 2014.

[59] K. Jeanne Dit Fouque, J. Moreno, J. D. Hegemann, S. Zirah, S. Rebuffat, and F. Fernandez-Lima, “Metal ions induced secondary structure rearrangements: Mechanically interlocked lasso: Vs. unthreaded branched-cyclic topoisomers,” *Analyst*, vol. 143, no. 10, pp. 2323–2333, 2018.

[60] A. L. Ferguson, S. Zhang, I. Dikiy, A. Z. Panagiotopoulos, P. G. Debenedetti, and A. J. Link, “An experimental and computational investigation of spontaneous lasso formation in microcin J25,” *Biophys. J.*, vol. 99, no. 9, pp. 3056–3065, 2010.

[61] J. D. Hegemann, M. Zimmermann, S. Zhu, D. Klug, and M. A. Marahiel, “Lasso peptides from proteobacteria: Genome mining employing heterologous expression and mass spectrometry,” *Biopolymers*, vol. 100, no. 5, pp. 527–542, 2013.

[62] M. Zimmermann, J. D. Hegemann, X. Xie, and M. A. Marahiel, “The astexin-1 lasso peptides: Biosynthesis, stability, and structural

studies," *Chem. Biol.*, vol. 20, no. 4, pp. 558–569, 2013.

[63] J. D. Hegemann *et al.*, "XanthomoninsI-III: A new class of lasso peptides with a seven-residue macrolactam ring," *Angew. Chemie - Int. Ed.*, vol. 53, no. 8, pp. 2230–2234, 2014.

[64] J. D. Hegemann, M. Zimmermann, X. Xie, and M. A. Marahiel, "Caulosegnins I-III: A highly diverse group of lasso peptides derived from a single biosynthetic gene cluster," *J. Am. Chem. Soc.*, vol. 135, no. 1, pp. 210–222, 2013.

[65] M. Zimmermann, J. D. Hegemann, X. Xie, and M. A. Marahiel, "Characterization of caulonodin lasso peptides revealed unprecedented N-terminal residues and a precursor motif essential for peptide maturation," *Chem. Sci.*, vol. 5, no. 10, pp. 4032–4043, 2014.

[66] R. A. Salomon and R. N. Farias, "Microcin 25, a novel antimicrobial peptide produced by Escherichia coli," *J. Bacteriol.*, vol. 174, no. 22, pp. 7428–7435, 1992.

[67] T. A. Knappe, U. Linne, L. Robbel, and M. A. Marahiel, "Insights into the Biosynthesis and Stability of the Lasso Peptide Capistruin," *Chem. Biol.*, vol. 16, no. 12, pp. 1290–1298, 2009.

[68] K. Jeanne Dit Fouque, V. Bisram, J. D. Hegemann, S. Zirah, S. Rebuffat, and F. Fernandez-Lima, "Structural signatures of the class III lasso peptide BI-32169 and the branched-cyclic topoisomers using trapped ion mobility spectrometry–mass spectrometry and tandem mass spectrometry," *Anal. Bioanal. Chem.*, vol. 411, no. 24, pp. 6287–6296, 2019.

[69] J. D. Hegemann, "Factors Governing the Thermal Stability of Lasso Peptides," *ChemBioChem*, vol. Early View, 2019.

[70] C. D. Allen, M. Y. Chen, A. Y. Trick, D. T. Le, A. L. Ferguson, and A. J. Link, "Thermal unthreading of the lasso peptides astexin-2 and astexin-3," *ACS Chem. Biol.*, vol. 11, no. 11, pp. 3043–3051, 2016.

[71] J. D. Hegemann *et al.*, "The ring residue proline 8 is crucial for the thermal stability of the lasso peptide caulosegnin II," *Mol. Biosyst.*, vol. 12, no. 4, pp. 1106–1109, 2016.

[72] J. D. Koos and A. J. Link, “Heterologous and in Vitro Reconstitution of Fuscanodin, a Lasso Peptide from Thermobifida fusca,” *J. Am. Chem. Soc.*, vol. 141, no. 2, pp. 928–935, 2019.

[73] E. Gavrish *et al.*, “Lassomycin, a ribosomally synthesized cyclic peptide, kills mycobacterium tuberculosis by targeting the ATP-dependent protease ClpC1P1P2,” *Chem. Biol.*, vol. 21, no. 4, pp. 509–518, 2014.

[74] P. W. R. Harris, G. M. Cook, I. K. H. Leung, and M. A. Brimble, “An Efficient Chemical Synthesis of Lassomycin Enabled by an On-Resin Lactamisation-Off-Resin Methanolysis Strategy and Preparation of Chemical Variants,” *Aust. J. Chem.*, vol. 70, no. 2, pp. 172–183, 2017.

[75] S. Lear *et al.*, “Total chemical synthesis of lassomycin and lassomycin-amide,” *Org. Biomol. Chem.*, vol. 14, no. 19, pp. 4534–4541, 2016.

[76] C. Zong, M. J. Wu, J. Z. Qin, and A. J. Link, “Lasso Peptide Benenodin-1 Is a Thermally Actuated [1]Rotaxane Switch,” *J. Am. Chem. Soc.*, vol. 139, no. 30, pp. 10403–10409, 2017.

[77] T. A. Knappe, U. Linne, X. Xie, and M. A. Marahiel, “The glucagon receptor antagonist BI-32169 constitutes a new class of lasso peptides,” *FEBS Lett.*, vol. 584, no. 4, pp. 785–789, 2010.

[78] W. L. Cheung-Lee, L. Cao, and A. J. Link, “Pandonodin: a proteobacterial lasso peptide with an exceptionally long C-terminal tail,” *ACS Chem. Biol.*, 2019.

[79] W. L. Cheung-Lee *et al.*, “Discovery of ubonodin, an antimicrobial lasso peptide active against members of the Burkholderia cepacia complex,” *ChemBioChem*, 2019.

[80] V. Valiante *et al.*, “Hitting the caspofungin salvage pathway of human-pathogenic fungi with the novel lasso peptide humidimycin (MDN-0010),” *Antimicrob. Agents Chemother.*, vol. 59, no. 9, pp. 5145–5153, 2015.

[81] O. Potterat, H. Stephan, J. W. Metzger, V. Gnau, H. Zähner, and G. Jung, “Aborycin – A Tricyclic 21-Peptide Antibiotic Isolated from Streptomyces griseoflavus,” *Liebigs Ann. der Chemie*, vol. 1994, no. 7, pp. 741–743, 1994.

[82] G. Helynck, C. Dubertret, J. F. Mayaux, and J. Leboul, “Isolation of RP 71955, a new Anti-Hiv-1 peptide secondary metabolite,” *J. Antibiot. (Tokyo).*, vol. 46, no. 11, pp. 1756–1757, 1993.

[83] R. Katahira, M. Yamasaki, Y. Matsuda, and M. Yoshida, “MS-271, A novel inhibitor of calmodulin-activated myosin light chain kinase from Streptomyces sp. - II. Solution structure of MS-271: Characteristic features of the ‘lasso’ structure,” *Bioorganic Med. Chem.*, vol. 4, no. 1, pp. 121–129, 1996.

[84] K. L. Constantine *et al.*, “High-resolution solution structure of siamycin II: Novel amphipathic character of a 21-residue peptide that inhibits HIV fusion,” *J. Biomol. NMR*, vol. 5, no. 3, pp. 271–286, 1995.

[85] Y. Li *et al.*, “Characterization of Sviceucin from Streptomyces Provides Insight into Enzyme Exchangeability and Disulfide Bond Formation in Lasso Peptides,” *ACS Chem. Biol.*, vol. 10, no. 11, pp. 2641–2649, 2015.

[86] I. Kaweewan, H. Hemmi, H. Komaki, S. Harada, and S. Kodani, “Isolation and structure determination of a new lasso peptide specialicin based on genome mining,” *Bioorganic Med. Chem.*, vol. 26, no. 23–24, pp. 6050–6055, 2018.

[87] I. Kaweewan, M. Ohnishi-Kameyama, and S. Kodani, “Isolation of a new antibacterial peptide achromosin from Streptomyces achromogenes subsp. achromogenes based on genome mining,” *J. Antibiot. (Tokyo).*, vol. 70, no. 2, pp. 208–211, 2017.

[88] M. Metelev *et al.*, “Acinetodin and Klebsidin, RNA Polymerase Targeting Lasso Peptides Produced by Human Isolates of Acinetobacter gyllenbergii and Klebsiella pneumoniae,” *ACS Chem. Biol.*, vol. 12, no. 3, pp. 814–824, 2017.

[89] N. Takasaka, I. Kaweewan, M. Ohnishi-Kameyama, and S. Kodani, “Isolation of a new antibacterial peptide actinokineosin from Actinokineospora spheciospongiae based on genome mining,” *Lett. Appl. Microbiol.*, vol. 64, no. 2, pp. 150–157, 2017.

[90] C. Zong, W. L. Cheung-Lee, H. E. Elashal, M. Raj, and A. J. Link, “Albusnodin: An acetylated lasso peptide from: Streptomyces albus,” *Chem. Commun.*, vol. 54, no. 11, pp. 1339–1342, 2018.

[91] D. F. Wyss, H. W. Lahm, M. Manneberg, and A. M. Labhardt, “Anantin — a peptide antagonist of the atrial natriuretic factor (anf): II. determination of the primary sequence by NMR on the basis of proton assignments,” *J. Antibiot. (Tokyo).*, vol. 44, no. 2, pp. 172–180, 1991.

[92] S. Kodani, H. Hemmi, Y. Miyake, I. Kaweewan, and H. Nakagawa, “Heterologous production of a new lasso peptide brevunsin in Sphingomonas subterranea,” *J. Ind. Microbiol. Biotechnol.*, vol. 45, no. 11, pp. 983–992, 2018.

[93] E. V Bratovanov *et al.*, “Genome Mining and Heterologous Expression Reveal Two Distinct Families of Lasso Peptides Highly Conserved in Endofungal Bacteria,” *ACS Chem. Biol.*, 2019.

[94] S. Sugai, M. Ohnishi-Kameyama, and S. Kodani, “Isolation and identification of a new lasso peptide cattlecin from Streptomyces cattleya based on genome mining,” *Appl. Biol. Chem.*, vol. 60, no. 2, pp. 163–167, 2017.

[95] S. S. Elsayed *et al.*, “Chaxapeptin, a Lasso Peptide from Extremotolerant Streptomyces leeuwenhoekii Strain C58 from the Hyperarid Atacama Desert,” *J. Org. Chem.*, vol. 80, no. 20, pp. 10252–10260, 2015.

[96] W. L. Cheung-Lee, M. E. Parry, A. J. Cartagena, S. A. Darst, and A. James Link, “Discovery and structure of the antimicrobial lasso peptide citrocin,” *J. Biol. Chem.*, vol. 294, no. 17, pp. 6822–6830, 2019.

[97] M. Iwatsuki, H. Tomoda, R. Uchida, H. Gouda, S. Hirono, and S. Omura, “Lariatins, antimycobacterial peptides produced by Rhodococcus sp. K01-B0171, have a lasso structure,” *J. Am. Chem. Soc.*, vol. 128, no. 23, pp. 7486–7491, 2006.

[98] Y. Su *et al.*, “Discovery and characterization of a novel C-terminal peptide carboxyl methyltransferase in a lassomycin-like lasso peptide biosynthetic pathway,” *Appl. Microbiol. Biotechnol.*, vol. 103, no. 6, pp. 2649–2664, 2019.

[99] J. P. Gomez-Escribano *et al.*, “Heterologous expression of a cryptic gene cluster from Streptomyces leeuwenhoekii C34T yields a novel lasso peptide, leepeptin,” *Appl. Environ. Microbiol.*, vol. 85, no. 23,

2019.

[100] S. Zhu *et al.*, “Insights into the unique phosphorylation of the lasso peptide paeninodin,” *J. Biol. Chem.*, vol. 291, no. 26, pp. 13662–13678, 2016.

[101] K. I. Kimura, F. Kanou, H. Takahashi, Y. Esumi, M. Uramoto, and M. Yoshihama, “Propeptin, a new inhibitor of prolyl endopeptidase produced by Microbispora. I. Fermentation, isolation and biological properties,” *J. Antibiot. (Tokyo).*, vol. 50, no. 5, pp. 373–378, 1997.

[102] T. Zyubko *et al.*, “Efficient In vivo synthesis of lasso peptide pseudomycoidin proceeds in the absence of both the leader and the leader peptidase,” *Chem. Sci.*, vol. 10, no. 42, pp. 9699–9707, 2019.

[103] T. Ogawa *et al.*, “RES-701-2, -3 and -4, Novel and Selective Endothelin Type B Receptor Antagonists Produced by Streptomyces sp.: I. Taxonomy of Producing Strains, Fermentation, Isolation, and Biochemical Properties,” *J. Antibiot. (Tokyo).*, vol. 48, no. 11, pp. 1213–1220, 1995.

[104] R. D. Kersten *et al.*, “A mass spectrometry-guided genome mining approach for natural product peptidogenomics,” *Nat. Chem. Biol.*, vol. 7, no. 11, pp. 794–802, 2011.

[105] M. Metelev *et al.*, “Structure, bioactivity, and resistance mechanism of streptomonomicin, an unusual lasso peptide from an understudied halophilic actinomycete,” *Chem. Biol.*, vol. 22, no. 2, pp. 241–250, 2015.

[106] M. Kuroha, H. Hemmi, M. Ohnishi-Kameyama, and S. Kodani, “Isolation and structure determination of a new lasso peptide subterisin from Sphingomonas subterranea,” *Tetrahedron Lett.*, vol. 58, no. 35, pp. 3429–3432, 2017.

[107] S. Um *et al.*, “Sungsanpin, a lasso peptide from a deep-sea streptomycete,” *J. Nat. Prod.*, vol. 76, no. 5, pp. 873–879, 2013.

[108] S. Son *et al.*, “Ulleungdin, a Lasso Peptide with Cancer Cell Migration Inhibitory Activity Discovered by the Genome Mining Approach,” *J. Nat. Prod.*, vol. 81, no. 10, pp. 2205–2211, 2018.

[109] S. Zhu, C. D. Fage, J. D. Hegemann, D. Yan, and M. A. Marahiel,

"Dual substrate-controlled kinase activity leads to polyphosphorylated lasso peptides," *FEBS Lett.*, vol. 590, pp. 3323–3334, 2016.

[110] T. Goulas *et al.*, "Structure and mechanism of a bacterial host-protein citrullinating virulence factor, Porphyromonas gingivalis peptidylarginine deiminase," *Sci. Rep.*, vol. 5, no. June, pp. 1–17, 2015.

[111] M. O. Maksimov and A. J. Link, "Discovery and characterization of an isopeptidase that linearizes lasso peptides," *J. Am. Chem. Soc.*, vol. 135, no. 32, pp. 12038–12047, 2013.

[112] Z. Feng, Y. Ogasawara, S. Nomura, and T. Dairi, "Biosynthetic Gene Cluster of a d-Tryptophan-Containing Lasso Peptide, MS-271," *ChemBioChem*, vol. 19, no. 19, pp. 2045–2048, 2018.

[113] T. A. Knappe *et al.*, "Introducing lasso peptides as molecular scaffolds for drug design: Engineering of an integrin antagonist," *Angew. Chemie - Int. Ed.*, vol. 50, no. 37, pp. 8714–8717, 2011.

[114] M. Chen, S. Wang, and X. Yu, "Cryptand-imidazolium supported total synthesis of the lasso peptide BI-32169 and its d-enantiomer," *Chem. Commun.*, vol. 55, no. 23, pp. 3323–3326, 2019.

[115] J. D. Hegemann *et al.*, "Rational improvement of the affinity and selectivity of integrin binding of grafted lasso peptides," *J. Med. Chem.*, vol. 57, no. 13, pp. 5829–5834, 2014.

[116] C. D. Allen and A. J. Link, "Self-Assembly of Catenanes from Lasso Peptides," *J. Am. Chem. Soc.*, vol. 138, no. 43, pp. 14214–14217, 2016.

[117] R. Soudy, L. Wang, and K. Kaur, "Synthetic peptides derived from the sequence of a lasso peptide microcin J25 show antibacterial activity," *Bioorganic Med. Chem.*, vol. 20, no. 5, pp. 1794–1800, 2012.

[118] R. Hammami, F. Bédard, A. Gomaa, M. Subirade, E. Biron, and I. Fliss, "Lasso-inspired peptides with distinct antibacterial mechanisms," *Amino Acids*, vol. 47, no. 2, pp. 417–428, 2015.

[119] F. Coutrot and E. Busseron, "A new glycorotaxane molecular machine based on an anilinium and a triazolium station," *Chem. - A*

Eur. J., vol. 14, no. 16, pp. 4784–4787, 2008.

[120] C. Clavel, K. Fournel-Marotte, and F. Coutrot, "A pH-sensitive peptide-containing lasso molecular switch," *Molecules*, vol. 18, no. 9, pp. 11553–11575, 2013.

[121] F. Saito and J. W. Bode, "Synthesis and stabilities of peptide-based [1]rotaxanes: molecular grafting onto lasso peptide scaffolds," *Chem. Sci.*, vol. 8, no. 4, pp. 2878–2884, 2017.

[122] S. Hayashi *et al.*, "Genome mining reveals a minimum gene set for the biosynthesis of 32-membered macrocyclic thiopeptides lactazoles," *Chem. Biol.*, vol. 21, no. 5, pp. 679–688, 2014.

[123] X. Yang and W. A. Van Der Donk, "Ribosomally synthesized and post-translationally modified peptide natural products: New insights into the role of leader and core peptides during biosynthesis," *Chem. - A Eur. J.*, vol. 19, no. 24, pp. 7662–7677, 2013.

[124] B. J. Burkhart, G. A. Hudson, K. L. Dunbar, and D. A. Mitchell, "A prevalent peptide-binding domain guides ribosomal natural product biosynthesis," *Nat. Chem. Biol.*, vol. 11, no. 8, pp. 564–570, 2015.

[125] B. M. Wieckowski, J. D. Hegemann, A. Mielcarek, L. Boss, O. Burghaus, and M. A. Marahiel, "The PqqD homologous domain of the radical SAM enzyme ThnB is required for thioether bond formation during thurincin H maturation," *FEBS Lett.*, vol. 589, no. 15, pp. 1802–1806, 2015.

[126] C. A. Regni, R. F. Roush, D. J. Miller, A. Nourse, C. T. Walsh, and B. A. Schulman, "How the MccB bacterial ancestor of ubiquitin E1 initiates biosynthesis of the microcin C7 antibiotic," *EMBO J.*, vol. 28, no. 13, pp. 1953–1964, 2009.

[127] A. Appel, "Functional and structural properties of A novel LASSO biosynthetic machinery (ALASSO)," pp. 1–24, 2015.

[128] J. O. Solbiati, M. Ciaccio, R. N. Farías, J. E. González-Pastor, F. Moreno, and R. A. Salomón, "Sequence analysis of the four plasmid genes required to produce the circular peptide antibiotic microcin J25," *J. Bacteriol.*, vol. 181, no. 8, pp. 2659–2662, 1999.

[129]J. O. Solbiati, M. Ciaccio, R. N. Parías, and R. A. Salomón, "Genetic analysis of plasmid determinants for microcin J25 production and immunity," *J. Bacteriol.*, vol. 178, no. 12, pp. 3661–3663, 1996.

[130]K. P. Yan *et al.*, "Dissecting the Maturation Steps of the Lasso Peptide Microcin J25 in vitro," *ChemBioChem*, vol. 13, no. 7, pp. 1046–1052, 2012.

[131]S. Zhu *et al.*, "The B1 Protein Guides the Biosynthesis of a Lasso Peptide," *Sci. Rep.*, vol. 6, no. August, p. 35604, 2016.

[132]R. L. Evans, J. A. Latham, Y. Xia, J. P. Klinman, and C. M. Wilmot, "Nuclear Magnetic Resonance structure and binding studies of PqqD, a chaperone required in the biosynthesis of the bacterial dehydrogenase cofactor pyrroloquinoline quinone," *Biochemistry*, vol. 56, no. 21, pp. 2735–2746, 2017.

[133]M. A. Ortega, Y. Hao, Q. Zhang, M. C. Walker, W. A. Van Der Donk, and S. K. Nair, "Structure and mechanism of the tRNA-dependent lantibiotic dehydratase NisB," *Nature*, vol. 517, no. 7535, pp. 509–512, 2015.

[134]K. M. Davis *et al.*, "Structures of the peptide-modifying radical SAM enzyme SuiB elucidate the basis of substrate recognition," *Proc. Natl. Acad. Sci. U. S. A.*, vol. 114, no. 39, pp. 10420–10425, 2017.

[135]J. Inokoshi, M. Matsuhama, M. Miyake, H. Ikeda, and H. Tomoda, "Molecular cloning of the gene cluster for lariatin biosynthesis of Rhodococcus jostii K01-B0171," *Appl. Microbiol. Biotechnol.*, vol. 95, no. 2, pp. 451–460, 2012.

[136]J. D. Hegemann, C. J. Schwalen, D. A. Mitchell, and W. A. Van der Donk, "Elucidation of the roles of conserved residues in the biosynthesis of the lasso peptide paeninodin," *Chem. Commun.*, vol. 54, no. 65, pp. 9007–9010, 2018.

[137]H. Martin-Gómez, U. Linne, F. Albericio, J. Tulla-Puche, and J. D. Hegemann, "Investigation of the biosynthesis of the lasso peptide chaxapeptin using an E. coli-based production system," *J. Nat. Prod.*, vol. 81, no. 9, pp. 2050–2056, 2018.

[138]C. J. Tsai, B. Ma, and R. Nussinov, "Intra-molecular chaperone:

The role of the N-terminal in conformational selection and kinetic control," *Phys. Biol.*, vol. 6, no. 1, pp. 1–8, 2009.

[139] K. L. Dunbar, J. R. Chekan, C. L. Cox, B. J. Burkhart, S. K. Nair, and D. A. Mitchell, "Discovery of a new ATP-binding motif involved in peptidic azoline biosynthesis," *Nat. Chem. Biol.*, vol. 10, no. 10, pp. 823–829, 2014.

[140] A. M. Burroughs, L. M. Iyer, and L. Aravind, "Natural history of the E1-like superfamily: Implication for adenylation, sulfur transfer, and ubiquitin conjugation," *Proteins Struct. Funct. Bioinforma.*, vol. 75, no. 4, pp. 895–910, 2009.

[141] J. A. Latham, I. Barr, and J. P. Klinman, "At the confluence of ribosomally synthesized peptide modification and radical S-adenosylmethionine (SAM) enzymology," *J. Biol. Chem.*, vol. 292, no. 40, pp. 16397–16405, 2017.

[142] Y. Q. Shen, F. Bonnot, E. M. Imsand, J. M. Rosefigura, K. Sjölander, and J. P. Klinman, "Distribution and properties of the genes encoding the biosynthesis of the bacterial cofactor, pyrroloquinoline quinone," *Biochemistry*, vol. 51, no. 11, pp. 2265–2275, 2012.

[143] S. R. Wecksler *et al.*, "Interaction of PqqE and PqqD in the pyrroloquinoline quinone (PQQ) biosynthetic pathway links PqqD to the radical SAM superfamily," *Chem. Commun.*, vol. 46, no. 37, pp. 7031–7033, 2010.

[144] H. Toyama, L. Chistoserdova, and M. E. Lidstrom, "Sequence analysis of pqq genes required for biosynthesis of pyrroloquinoline quinone in Methylobacterium extorquens AM1 and the purification of a biosynthetic intermediate," *Microbiology*, vol. 143, no. 2, pp. 595–602, 1997.

[145] E. W. Schmidt *et al.*, "Schmidt EW 2005 Patellamides A und C from Prochloron didemni.pdf," vol. 102, no. 20, pp. 7315–7320, 2005.

[146] J. A. McIntosh, Z. Lin, M. D. B. Tianero, and E. W. Schmidt, "Aestuaramides, a natural library of cyanobactin cyclic peptides resulting from isoprene-derived Claisen rearrangements," *ACS Chem. Biol.*, vol. 8, no. 5, pp. 877–883, 2013.

[147] M. A. Ortega *et al.*, “Structure and tRNA Specificity of MibB, a Lantibiotic Dehydratase from Actinobacteria Involved in NAI-107 Biosynthesis,” *Cell Chem. Biol.*, vol. 23, no. 3, pp. 370–380, 2016.

[148] T. Sumida, S. Dubiley, B. Wilcox, K. Severinov, and S. Tagami, “Structural Basis of Leader Peptide Recognition in Lasso Peptide Biosynthesis Pathway,” *ACS Chem. Biol.*, vol. 14, no. 7, pp. 1619–1627, 2019.

[149] J. A. Latham, A. T. Iavarone, I. Barr, P. V. Juthani, and J. P. Klinman, “PqqD is a novel peptide chaperone that forms a ternary complex with the radical S-adenosylmethionine protein PqqE in the pyrroloquinoline quinone biosynthetic pathway,” *J. Biol. Chem.*, vol. 290, no. 20, pp. 12908–12918, 2015.

[150] O. Pavlova, J. Mukhopadhyay, E. Sineva, R. H. Ebright, and K. Severinov, “Systematic structure-activity analysis of microcin J25,” *J. Biol. Chem.*, vol. 283, no. 37, pp. 25589–25595, 2008.

[151] S. J. Pan and A. J. Link, “Sequence diversity in the lasso peptide framework: Discovery of functional microcin J25 variants with multiple amino acid substitutions,” *J. Am. Chem. Soc.*, vol. 133, no. 13, pp. 5016–5023, 2011.

[152] T. Y. Tsai, C. Y. Yang, H. L. Shih, A. H. J. Wang, and S. H. Chou, “Xanthomonas campestris PqqD in the pyrroloquinoline quinone biosynthesis operon adopts a novel saddle-like fold that possibly serves as a PQQ carrier,” *Proteins Struct. Funct. Bioinforma.*, vol. 76, no. 4, pp. 1042–1048, 2009.

[153] E. Kvamme, “Transglutaminases,” *Glutamine Glutamate Mamm. Vol. I*, vol. 35, pp. 69–80, 2018.

[154] V. C. Yee, L. C. Pedersen, I. Le Trong, P. D. Bishop, R. E. Stenkamp, and D. C. Teller, “Three-dimensional structure of a transglutaminase: Human blood coagulation factor XIII,” *Proc. Natl. Acad. Sci. U. S. A.*, vol. 91, no. 15, pp. 7296–7300, 1994.

[155] K. S. Makarova, L. Aravind, and E. V. Koonin, “A superfamily of archaeal, bacterial, and eukaryotic proteins homologous to animal transglutaminases,” *Protein Sci.*, vol. 8, no. 8, pp. 1714–1719, 1999.

[156] S. J. Pan, J. Rajniak, W. L. Cheung, and A. J. Link, “Construction

of a single polypeptide that matures and exports the lasso peptide microcin J25," *ChemBioChem*, vol. 13, no. 3, pp. 367–370, 2012.

[157] M. O. Maksimov and A. J. Link, "Prospecting genomes for lasso peptides," *J. Ind. Microbiol. Biotechnol.*, vol. 41, no. 2, pp. 333–344, 2014.

[158] S. Duquesne, D. Destoumieux-Garzón, S. Zirah, C. Goulard, J. Peduzzi, and S. Rebuffat, "Two Enzymes Catalyze the Maturation of a Lasso Peptide in Escherichia coli," *Chem. Biol.*, vol. 14, no. 7, pp. 793–803, 2007.

[159] M. A. Scofield, W. S. Lewis, and S. M. Schuster, "Nucleotide sequence of Escherichia coli asnB and deduced amino acid sequence of asparagine synthetase B," *J. Biol. Chem.*, vol. 265, no. 22, pp. 12895–12902, 1990.

[160] R. Ducasse *et al.*, "Sequence determinants governing the topology and biological activity of a lasso peptide, microcin J25," *ChemBioChem*, vol. 13, no. 3, pp. 371–380, 2012.

[161] M. A. Ortega and W. A. Van Der Donk, "New Insights into the Biosynthetic Logic of Ribosomally Synthesized and Post-translationally Modified Peptide Natural Products," *Cell Chem. Biol.*, vol. 23, no. 1, pp. 31–44, 2016.

[162] H. G. Choudhury *et al.*, "Structure of an antibacterial peptide ATP-binding cassette transporter in a novel outward occluded state," *Proc. Natl. Acad. Sci. U. S. A.*, vol. 111, no. 25, pp. 9145–9150, 2014.

[163] F. E. Lopez, P. A. Vincent, A. M. Zenoff, R. A. Salomón, and R. N. Farías, "Efficacy of microcin J25 in biomatrices and in a mouse model of Salmonella infection," *J. Antimicrob. Chemother.*, vol. 59, no. 4, pp. 676–680, 2007.

[164] M. O. Maksimov, S. J. Pan, and A. James Link, "Lasso peptides: Structure, function, biosynthesis, and engineering," *Nat. Prod. Rep.*, vol. 29, no. 9, pp. 996–1006, 2012.

[165] M. O. Maksimov, J. D. Koos, C. Zong, B. Lisko, and A. James Link, "Elucidating the specificity determinants of the AtxE2 lasso peptide isopeptidase," *J. Biol. Chem.*, vol. 290, no. 52, pp. 30806–30812, 2015.

[166] C. D. Fage *et al.*, “Structure and Mechanism of the Sphingopyxin I Lasso Peptide Isopeptidase,” *Angew. Chemie - Int. Ed.*, vol. 55, no. 41, pp. 12717–12721, 2016.

[167] T. A. Knappe, U. Linne, S. Zirah, S. Rebuffat, X. Xie, and M. A. Marahiel, “Isolation and structural characterization of capistruin, a lasso peptide predicted from the genome sequence of Burkholderia thailandensis E264,” *J. Am. Chem. Soc.*, vol. 130, no. 34, pp. 11446–11454, 2008.

[168] A. R. Ducasse*, E. L. , Y. Lia, A. Blonda, S. Ziraha, E. G. C. Goularda, J.-L. Pernodetc, and S. Rebuffata, “Sviceucin, a lasso peptide from Streptomyces sviceus: isolation and structure analysis,” *J. Pept. Sci.*, pp. 67–68, 2012.

[169] J. R. Chekan, J. D. Koos, C. Zong, M. O. Maksimov, A. J. Link, and S. K. Nair, “Structure of the Lasso Peptide Isopeptidase Identifies a Topology for Processing Threaded Substrates,” *J. Am. Chem. Soc.*, vol. 138, no. 50, pp. 16452–16458, 2016.

[170] P. A. Vincent, M. A. Delgado, R. N. Farías, and R. A. Salomón, “Inhibition of Salmonella enterica serovars by microcin J25,” *FEMS Microbiol. Lett.*, vol. 236, no. 1, pp. 103–107, 2004.

[171] M. Tsunakawa *et al.*, “Siamycins I and II, new anti-HIV peptides: I. Fermentation, isolation, biological activity and initial characterization.,” *J. Antibiot. (Tokyo).*, vol. 48, no. 5, pp. 433–434, 1995.

[172] M. Shao, J. Ma, Q. Li, and J. Ju, “Identification of the anti-infective aborycin biosynthetic gene cluster from deep-sea-derived Streptomyces sp. SCSIO ZS0098 enables production in a heterologous host,” *Mar. Drugs*, vol. 17, no. 2, pp. 1–9, 2019.

[173] Y. Morishita *et al.*, “RES-701-1,a novel and selective endothelin type B receptor antagonist produced by streptomyces sp. RE-701. I. Characterization of producing strain, fermentation, isolation, physico-chemical and biological properties,” *J. Antibiot. (Tokyo).*, vol. 47, no. 3, pp. 269–275, 1994.

[174] K. Kuznedelov *et al.*, “The antibacterial threaded-lasso peptide capistruin inhibits bacterial RNA polymerase,” *J. Mol. Biol.*, vol. 412, no. 5, pp. 842–848, 2011.

[175] N. R. Braffman, F. J. Piscotta, J. Hauver, E. A. Campbell, A. James Link, and S. A. Darst, "Structural mechanism of transcription inhibition by lasso peptides microcin J25 and capistruin," *Proc. Natl. Acad. Sci. U. S. A.*, vol. 116, no. 4, pp. 1273–1278, 2019.

[176] J. Inokoshi, N. Koyama, M. Miyake, Y. Shimizu, and H. Tomoda, "Structure-Activity Analysis of Gram-positive Bacterium-producing Lasso Peptides with Anti-mycobacterial Activity," *Sci. Rep.*, vol. 6, no. March, pp. 5–9, 2016.

[177] S. Zhu, Y. Su, S. Shams, Y. Feng, Y. Tong, and G. Zheng, "Lassomycin and lariatin lasso peptides as suitable antibiotics for combating mycobacterial infections: current state of biosynthesis and perspectives for production," *Appl. Microbiol. Biotechnol.*, vol. 103, pp. 3931–3940, 2019.

[178] M. Daniel-Ivad *et al.*, "An Engineered Allele of afsQ1 Facilitates the Discovery and Investigation of Cryptic Natural Products," *ACS Chem. Biol.*, vol. 12, no. 3, pp. 628–634, 2017.

[179] S. Tan, K. C. Ludwig, A. Müller, T. Schneider, and J. R. Nodwell, "The Lasso Peptide Siamycin-I Targets Lipid II at the Gram-Positive Cell Surface," *ACS Chem. Biol.*, vol. 14, no. 5, pp. 966–974, 2019.

[180] P. F. Lin *et al.*, "Characterization of siamycin I, a human immunodeficiency virus fusion inhibitor," *Antimicrob. Agents Chemother.*, vol. 40, no. 1, pp. 133–138, 1996.

[181] R. S. Al Toma *et al.*, "Site-directed and global incorporation of orthogonal and isostructural noncanonical amino acids into the ribosomal lasso peptide capistruin," *ChemBioChem*, vol. 16, no. 3, pp. 503–509, 2015.

[182] C. Zong, M. O. Maksimov, and A. J. Link, "Construction of Lasso Peptide Fusion Proteins," *ACS Chem. Biol.*, vol. 11, no. 1, pp. 61–68, 2016.

[183] T. Baba *et al.*, "Construction of Escherichia coli K-12 in-frame, single-gene knockout mutants: The Keio collection," *Mol. Syst. Biol.*, vol. 2, 2006.

[184] S. N. Ho, H. D. Hunt, R. M. Horton, J. K. Pullen, and L. R. Pease,

"Site-directed mutagenesis by overlap extension using the polymerase chain reaction," *Gene*, vol. 77, no. 1, pp. 51–59, 1989.

[185] O. Bantysh, M. Serebryakova, K. S. Makarova, S. Dubiley, K. A. Datsenko, and K. Severinov, "Enzymatic synthesis of bioinformatically predicted microcin C-like compounds encoded by diverse bacteria," *MBio*, vol. 5, no. 3, pp. 1–11, 2014.

[186] J. Radeck *et al.*, "The Bacillus BioBrick Box: Generation and evaluation of essential genetic building blocks for standardized work with Bacillus subtilis," *J. Biol. Eng.*, vol. 7, no. 1, p. 29, Dec. 2013.

[187] A. P. Pomerantsev, A. Camp, and S. H. Leppla, "A new minimal replicon of Bacillus anthracis plasmid pXO1," *J. Bacteriol.*, vol. 191, no. 16, pp. 5134–5146, 2009.

[188] O. Bantysh, M. Serebryakova, K. S. Makarova, S. Dubiley, and K. A. Datsenko, "Enzymatic Synthesis of Bioinformatically Predicted Microcin C-Like," *MBio*, vol. 5, no. 3, pp. e01059-14, 2014.

[189] D. Tsibulskaya *et al.*, "The Product of Yersinia pseudotuberculosis mcc Operon Is a Peptide-Cytidine Antibiotic Activated Inside Producing Cells by the TldD/E Protease," *J. Am. Chem. Soc.*, vol. 139, no. 45, pp. 16178–16187, 2017.

[190] F. Sherman, J. W. Stewart, and S. Tsunasawa, "Methionine or not methionine at the beginning of a protein," *BioEssays*, vol. 3, no. 1, pp. 27–31, 1985.

[191] Q. Xiao, F. Zhang, B. A. Nacev, J. O. Liu, and D. Pei, "Protein N-terminal processing: Substrate specificity of escherichia coli and human methionine aminopeptidases," *Biochemistry*, vol. 49, no. 26, pp. 5588–5599, 2010.

[192] H. Martin-Gómez and J. Tulla-Puche, "Lasso peptides: chemical approaches and structural elucidation," *Org. Biomol. Chem.*, vol. 16, no. 28, pp. 5065–5080, 2018.

[193] X. Xie and M. A. Marahiel, "NMR as an Effective Tool for the Structure Determination of Lasso Peptides," *ChemBioChem*, vol. 13, no. 5, pp. 621–625, 2012.