# Semi-*in vitro* Reconstitution of Roseocin, a Two-Component Lantibiotic from an Actinomycete

Mangal Singh and Dipti Sareen*

Department of Biochemistry, Basic Medical Sciences Block-II, Panjab University, Sector-25, Chandigarh-160014, INDIA
*Correspondence: diptsare@pu.ac.in

## ABSTRACT

Lantibiotics are lanthionine containing peptide natural products that belong to the class of ribosomally synthesized and post-translationally modified peptides (RiPPs). Recent expansion in the availability of microbial genomic data and *in silico* analysis tools have accelerated the discovery of these promising alternatives to antibiotics. Following the genome-mining approach, a biosynthetic gene cluster for a putative two-component lantibiotic roseocin was identified in the genome of an actinomycete, *Streptomyces roseosporus* NRRL 11379. Post-translationally modified lanthipeptides of this cluster were obtained by heterologous expression of the genes in *E. coli*, and *in vitro* reconstituted to their bioactive form. The two lanthipeptides displayed antimicrobial activity against Gram-positive bacteria only in synergy, reminiscent of a two-component lantibiotic**.** Structural analysis of the α-component identified a disulfide bridge flanking two of its four thioether bridges and the β-component having six thioether bridges with its C-terminus extended than the previously known two-component lantibiotics.

**Keywords:** Lantibiotics, two-component, Lanthionine, RiPPs, genome mining, lanthipeptides, roseocin, synergistic, thioether

## INTRODUCTION

Many of the clinically used antibiotics are derived from the genus *Streptomyces.* It is the largest genus of the Actinobacteria, Gram-positive and has a high GC content genome. The arsenal of antimicrobial compounds that this genus produces includes polyketides, non-ribosomal, and ribosomal peptides. Recent studies have focussed on the discovery of ribosomally synthesized and post-translationally modified peptides (RiPPs) in light of the need for novel antimicrobials to combat the rising antimicrobial drug resistance [1]. A particular focus has been on the discovery of novel lantibiotics (a class of RiPPs) that display their activity against the drug-resistant bacteria. Here, we have heterologously produced and characterized a lantibiotic roseocin in *E. coli*, from *Streptomyces roseosporus* NRRL 11379, by employing a semi-*in vitro* reconstitution approach. The lantibiotic comprises two peptides, one of which likely binds the

lipid II and the other acts synergistically with the lipid II binding peptide to exert the antimicrobial effect.

As with the polyketides and non-ribosomal peptide antibiotics, lantibiotics are also synthesized from a biosynthetic gene cluster (BGC). A genetically encoded lantibiotic is initially synthesized as a linear peptide (precursor peptide) with two important regions, an N-terminal leader peptide region separated by a proteolytic cleavage site from its C-terminal core peptide region that is rich in Cys/Ser/Thr residues. While the leader peptide guides the precursor peptide to different enzymes, the core peptide forms the final lantibiotic structure. Lantibiotics are synthesized in two steps by (1) dehydration of Ser/Thr residues in the core peptide followed by (2) intra-peptide Michael addition of cysteine residues to form lanthionine or thioether bridges (cyclization). These lanthionine rings are installed by various enzymes which define the four classes of lantibiotics. In class I lantibiotics two separate enzymes for dehydration (LanB) and cyclization (LanC) are encoded in the BGC along with the gene for the precursor peptide (LanA). Besides this, a lantibiotic BGC encodes a protease (LanP) for the proteolytic removal of the leader peptide, a transporter (LanT) for extracellular transport of the modified peptide, immunity proteins (LanF, LanE, LanG and sometimes LanI) and a two-component response regulation system (LanR and LanK). Class II, III and IV lantibiotic BGCs encode a single lanthionine synthetase instead of separate dehydratase and cyclase enzymes for the lanthionine ring formation. The class III and IV lantibiotics have bioactivities majorly other than antimicrobial action like morphogenetic[2], antiviral, antiallodynic[3], antinociceptive [4] and antidiabetic activities[5]. The class II lantibiotics with a single lanthionine synthetase are antimicrobial and also the most heterologously characterized lantibiotics.

In class II lantibiotics, lanthionine ring installation in the core peptide is done by a single lanthionine synthetase LanM, followed by the concomitant proteolytic removal of the leader and extracellular transport by a bifunctional transporter ($LanT_p$). The promiscuity of LanM in lanthionine ring installation on its precursor peptides can be low (dedicated LanM for a single peptide, as in haloduracin) to very high (single LanM acting on >17 precursor peptides in prochlorosins). Class II lantibiotics are further of two types depending upon the constituent peptides and their bioactivity. Single component lantibiotics contain peptides which are all antimicrobial [6] and the two-component lantibiotics contain two different types of peptides that show synergistic antimicrobial activity. Most of the two-component lantibiotics have a lipid II binding α-peptide along with a β-peptide, which acts in synergy with the α-peptide, thereby enhancing its activity manifolds [7]. Two-component lantibiotics are synthesized either by a single LanM which installs lanthionine rings on these two different types of peptide i.e. both

on the α and β precursor peptide [8–10], or by two separate LanM enzymes, each of which is specific to either the α or the β precursor peptide(s) [11–13]. While there are many examples of characterized and putative two-component lantibiotics [14,15] synthesized by two LanMs, only three are known to be processed by a single LanM i.e. carnolysin and cytolysin which are homologs [8,9] and bicereucin with D-amino acids as the major post-translational modification instead of lanthionine [10].

Understanding of genetic organization of lantibiotic encoding genes has led to the discovery of novel lantibiotics by combinatorial approach of genome mining, activity-based screening of thus identified potential producers and/or *in vitro* reconstitution of BGCs from the native producers [11,16]. Till date, two-component lantibiotics have been isolated and are confined to only firmicutes [14]. In our earlier genome mining study, we identified three putative two-component lantibiotic clusters in actinomycetes having two precursor peptides and two LanM, following a microbial genome database mining strategy for novel bifunctional transporter $LanT_p$ [15]. In addition, a cluster identified in *S. roseosporus* NRRL 11379, encoded for a single LanM to process two precursor peptides. The presence of six cysteine residues in each of the two peptides led us to speculate that it might form a highly constrained structure with a different bioactivity than the existing lantibiotics. Another point of exploration was to analyze whether these peptides show synergistic, as in two-component lantibiotics, or additive bioactivity. Hence, this cluster was undertaken for characterization by a semi-*in vitro* reconstitution approach, involving heterologous production of the lanthipeptides in *E. coli* [17].

## RESULTS

### Bioinformatic analysis of Roseocin as a two-component lantibiotic

The roseocin cluster was identified following a genome mining study for novel HalT homologs [15]. The roseocin genetic cluster encoded for two precursor peptides and a single lanthionine synthetase RosM (**Fig. 1**). Also, a putative bifunctional transporter $RosT_p$, immunity proteins RosEFG and a transcriptional regulator were encoded by the cluster (**Table S1**). The two RosA peptides display an overall 68% similarity (40% identity) in 75 amino acid overlap (Fig. 1B). Although the leader peptide is highly similar, the core-peptide region is variable in the composition of lanthionine forming moieties. RosA2α and RosA1β both contain six Cys with five- and twelve- Ser/Thr residues, respectively. Alignment of RosA peptides using MUSCLE with previously characterized two-component lantibiotics could not conclusively identify the α- and β-component of roseocin (Fig. 1C and 1D). Putative α-peptide which displayed limited sequence similarity on the N-terminus was designated as RosA2α, while β-peptide which is

comparatively extended on the C-terminus was designated as RosA1β. In general, α- and β-core peptides of roseocin are extended than those are previously characterized.

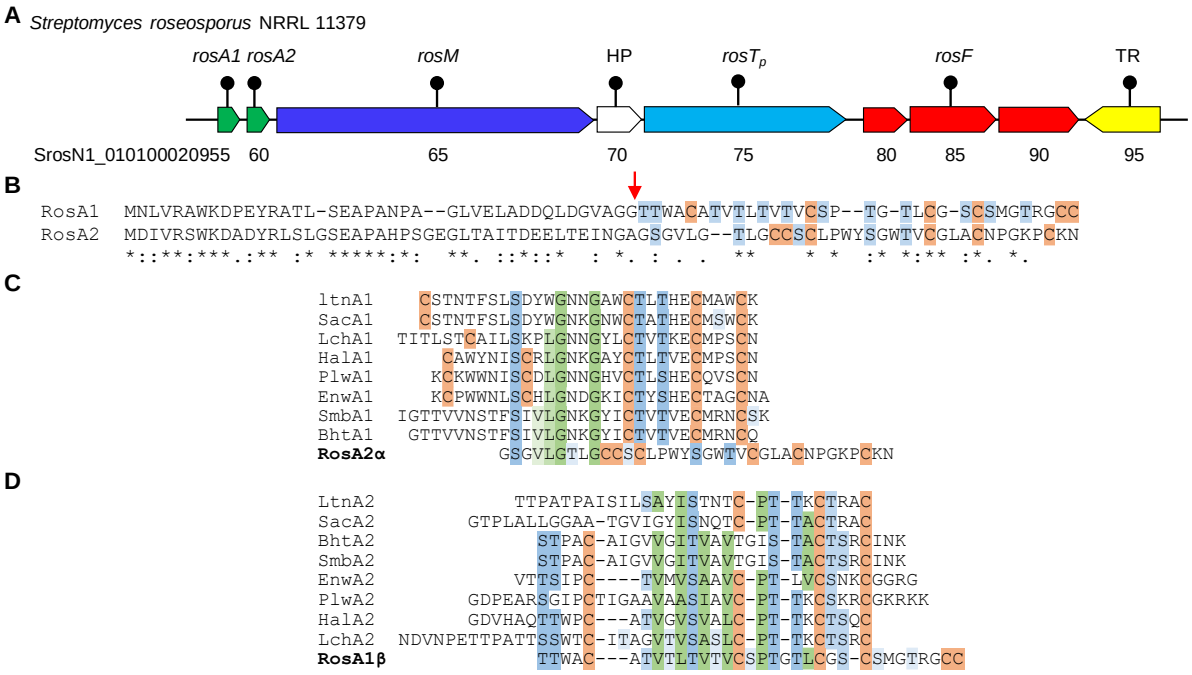


**Fig. 1. Roseocin is a homolog of two-component lantibiotics.**
(A) The roseocin biosynthetic gene cluster encoded in *S. roseosporus* NRRL 11379. Locus tags are mentioned below the genes. (B) Alignment of the full-length RosA1 and RosA2. The cleavage site for the CPD of $RosT_p$ is present next to the double glycine motif (marked with an arrow). (C) & (D) Multiple sequence alignment of RosA2α (ZP_04710450.1) and RosA1β (ZP_04710449.1) core peptides with α- and β- peptides of two-component lantibiotics lacticin 3147 (LtnA1, LtnA2); staphylococcin C55 (SacA1, SacA2); plantaricin W (PlwA1, PlwA2); BhtA (BhtA1, BhtA2); lichenicidin (BliA1, BliA2); haloduracin (HalA1, HalA2) and enterocin W (EnwA1, EnwA2). HP-hypothetical protein, accession numbers are given in *Supplemental information.*

## RosA peptides are *in vivo* dehydrated by RosM

To obtain the post-translationally modified RosA peptides, two recombinant plasmids were constructed, each having lanthionine synthetase *rosM*, together with either *rosA1* (pRSFDuet-*rosA1*-*rosM*) or *rosA2* (pRSFDuet-*rosA2*-*rosM*). Two *E. coli* BL21(DE3) hosts, harboring each of these constructs were induced with IPTG for expression and the peptides were found in soluble fraction. The hexahistidine tagged lanthipeptides were purified by Ni-NTA affinity chromatography followed by RP-HPLC. A yield of approximately 4 mg/liter culture of HPLC purified product was obtained for each of the RosA peptides. As was expected, the two RosA peptides were found to be modified by RosM, as MALDI-TOF MS analysis identified a reduction of mass in multiple of 18 Da from the calculated mass of the unmodified peptides (**Fig. 2**). $His_6$-mRosA1 was nine-fold dehydrated at 8803 *m/z* (calc. mono. 8803 *m/z* $[M-9H_2O+Na]^+$) , with the N-terminal methionine excised by the inherent activity of *E. coli* N-

aminopeptidase [18]. Similarly, $His_6$-mRosA2 (m-indicates modified) is four-fold dehydrated at 9558 *m/z* (calc. mono. 9558 *m/z* $[M-4H_2O+Na]^+$) along with the methionine excision. In both the peptides, number of dehydration are less than the total number of Ser/Thr residues, which is not uncommon in lanthipeptide biosynthesis [19], and often not all these dehydrated residues are linked into lanthionine bridges [11]. An additional peak for a gluconoylated derivative was also observed for both the hexahistidine tagged RosA peptides with a mass shift of +178 Da and +258 Da. As per earlier reports, proteins that are expressed in *E. coli* with an N-terminus Gly-Ser-Ser-$[His]_6$- undergo spontaneous α-N-6-Phosphogluconoylation (+258 Da) and upon phosphate removal by the host cell phosphatase, the mass shift becomes 178 Da [20]. However, these extra peaks for the gluconoylated product do not affect the final bioactivity, as the N-terminal region gets removed upon the *in vitro* proteolytic removal of the leader peptide. In conclusion, MALDI-TOF MS analysis confirmed the successful post-translational modification of both $His_6$-mRosA1 and $His_6$-mRosA2 by single RosM, heterologously. Although mass analysis of purified peptide can detect the number of dehydration, the extent of cyclization cannot be determined as there is no change in mass upon the intra-peptide Michael addition of cysteine residues for lanthionine formation.

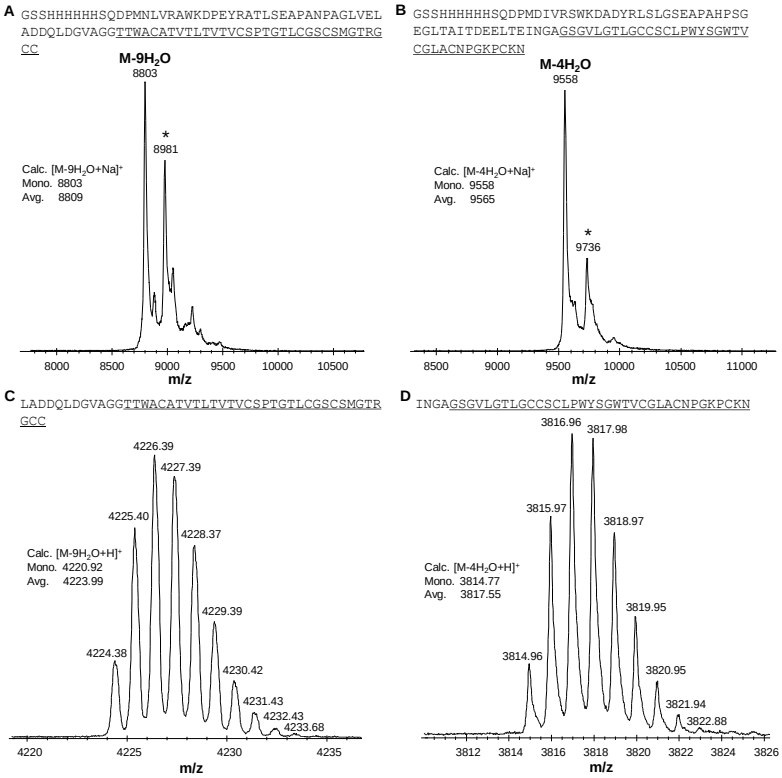


**Fig. 2. $His_6$-mRosA1 and $His_6$-mRosA2 are dehydrated by RosM in *E. coli.*** MALDI-TOF MS spectra of purified $His_6$-mRosA1 and $His_6$-mRosA2. (A) Nine-fold dehydrated $His_6$-mRosA1 with an observed mass of 8803 *m/z* (calc. mono. 8803 *m/z* $[M-9H_2O+Na]^+$). (B) The major peak for $His_6$-mRosA2, indicating four-fold dehydration with an observed mass of 9558 *m/z* (calc. mono. 9558 *m/z* $[M+-4H_2O\ Na]^+$). Peaks for alpha-N-gluconoylated products were also observed for both the heterologously produced RosA peptides (marked with an asterisk). Both $His_6$-mRosA1 and $His_6$-mRosA2 were treated with GluC for high-resolution mass spectrometric analysis. (C) Mass spectrum of isotopically separated RosA1β' with obs. 4223.38 *m/z* (calc. mono. 4220.92 *m/z*) and (D) RosA2α' with obs. 3814.96 *m/z* (calc. mono. 3814.77 *m/z*).

**RosA2α' has four while RosA1β' has six thioether rings**

Cysteine residue not involved in lanthionine ring can be detected by alkylation with iodoacetamide (IAA) and observed with mass spectrometry. A mass shift of +57.05 Da is observed when a carbamidomethyl adduct with free cysteine is formed [11]. So, alkylation of the modified peptides was carried out in reducing conditions followed by MALDI-TOF MS analysis. Alkylation of $His_6$-mRosA2 indicated an increase of mass corresponding to two carbamidomethylation in the presence of tris(2-carboxyethyl) phosphine (TCEP) at 9679 *m/z* (calc. avg. 9679 *m/z* $[M-4H_2O+2IAA+Na]^+$; **Fig. 3**). For $His_6$-mRosA1, alkylation did not lead to any change in its mass indicating that none of the six cysteine residues are free and that all are cyclized (Fig. 1B). Since both the RosA peptides contain six cysteine residues with nine and four dehydrated Ser/Thr residues, it indicated the presence of six and four thioether rings in $His_6$-mRosA1 and $His_6$-mRosA2, respectively. The observation of two alkylations in $His_6$-mRosA2 under reduced conditions, directed towards the possibility of the two cysteines to be involved in a disulfide linkage. This was further investigated following the leader peptide removal.

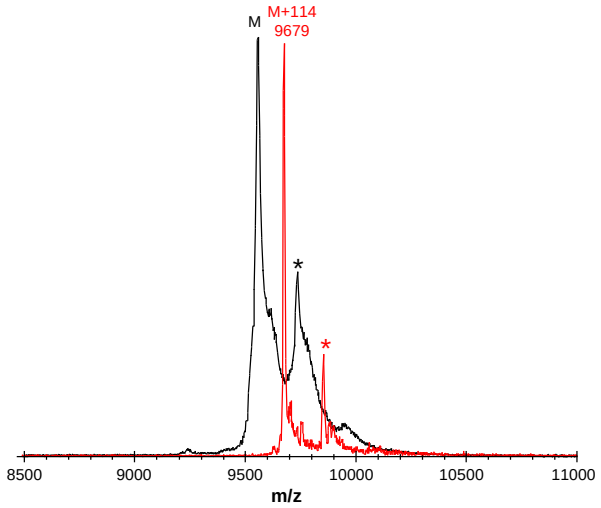


**Fig. 3. $His_6$-mRosA1 has six while $His_6$-mRosA2 has four thioether rings.** Both $His_6$-mRosA1 and $His_6$-mRosA2 were subjected to alkylation with IAA and analyzed with MALDI TOF MS. For $His_6$-mRosA2, a mass shift corresponding to two carbamidomethylation (+114 Da) was observed at 9679 *m/z* (calc. avg. 9679 *m/z* $[M+Na+2IAA]^+$) indicating that two of the six cysteines residues are not cyclized. In the case of $His_6$-mRosA1, no new peaks were observed upon alkylation (hence not shown) indicating all of the six cysteine residues are cyclized.

**RosA peptides contain overlapping thioether rings with a disulfide bond in RosA2α'**

Both the lanthipeptides were subjected to MALDI-TOF MS/MS for structural analysis. By using multiple sequence alignment (MSA) of RosA peptides, the conserved GA/GG motif was identified which demarcates the boundary between the leader and the core peptide (Fig. 1B). Leader peptides ending in GA/GG motif are proteolytically removed by cysteine protease domain (CPD) of the bifunctional transporter $LanT_p$ and indeed such a CPD domain is present in $RosT_p$ of the roseocin cluster (Fig. 1A). For *in vitro* studies, removal of the leader peptide has been achieved using commercial proteases such as LysC, trypsin or GluC to obtain

bioactive lanthipeptides [17]. Since, both the RosA peptides have multiple GluC cleavage sites in the leader region (at the C-terminal of glutamate residues) and none in the core-peptide, we utilized GluC for *in-vitro* removal of the leader peptide along with the hexahistidine tag. Treatment of RosA peptides with GluC removed the leader peptide thereby reducing the size to <5kDa and thus allowed high-resolution MALDI-TOF MS and sequence analysis with tandem MS. The GluC treated fragments were identified to be proteolytically cleaved at 3$^{rd}$ and 5$^{th}$ Glu residues for RosA1 and RosA2, respectively (Fig. 2C and 2D). These leader peptide removed RosA peptides with the trace of leader peptide were termed as RosA2α' and RosA1β'.

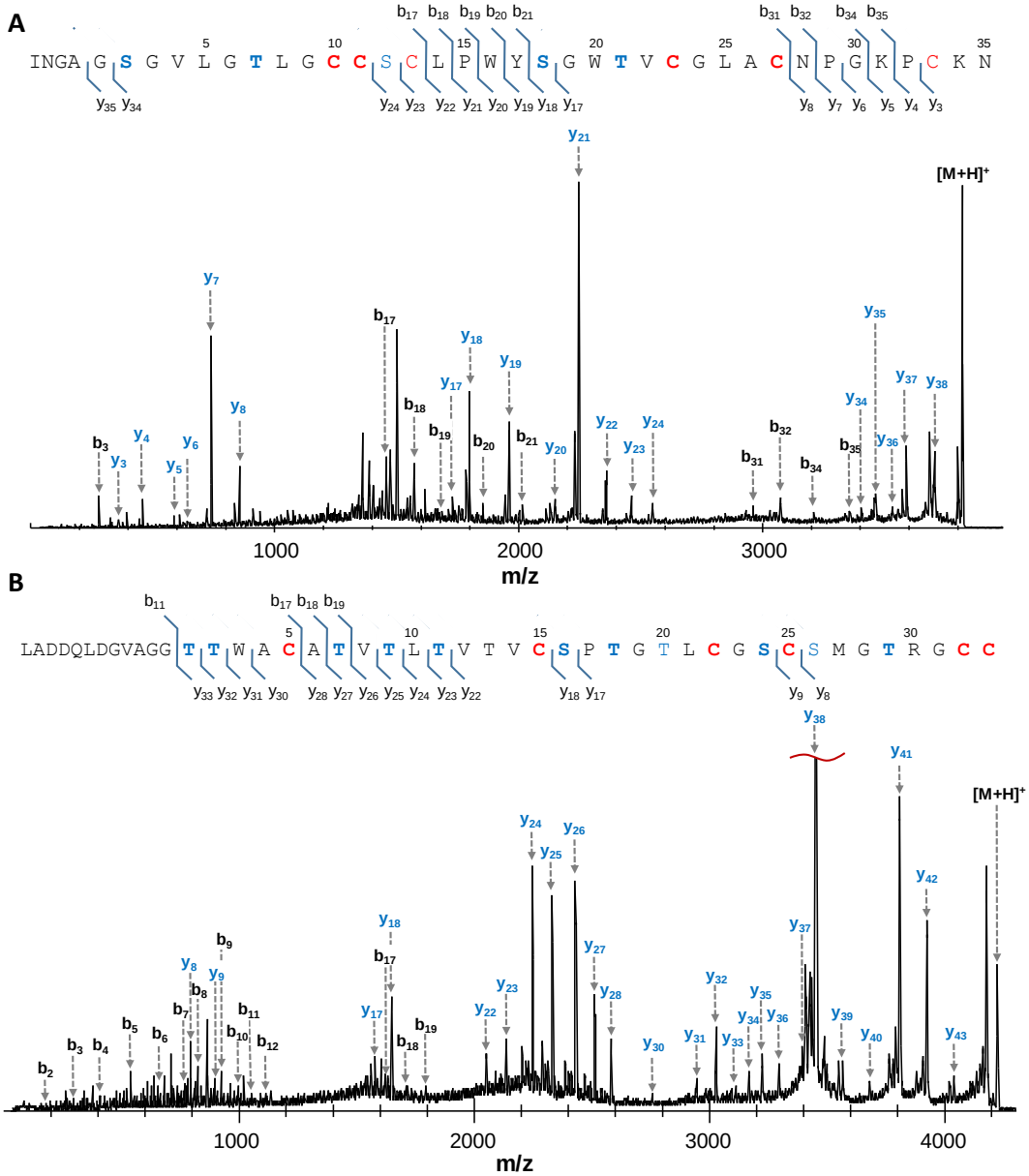


**Fig. 4. RosA2α' and RosA1β' core peptides contain multiple overlapping thioether bridges.** Fragment ions were not observed in the region protected with lanthionine rings and the mass shifts corresponding to dehydrated Ser/Thr are observed. b and y ions observed in MALDI TOF MS/MS spectrum for (A) RosA2α' and (B) RosA1β'. Residues in bold are modified by RosM. $[M+H]^+$ indicates precursor ion.

In general, as reported for other lantibiotics, fragmentation was not observed in the region protected with lanthionine rings. For RosA2α', where a probable disulfide bond presence was speculated by alkylation assay, MS/MS analysis was carried out with both reduced and non-reduced forms of the peptide. A ~2 Da increase in mass was observed upon TCEP treatment of RosA2α', indicating an addition of two protons to the disulfide bonded Cys residues (**Fig. S2**). Reduction of RosA2α' also led to an enhancement in fragmentation (**Fig. 4A and S3**) and hence identification of b and y ions with mass shifts of corresponding dehydrations of Ser/Thr residues (**Table S2**). $MS^n$ analysis of reduced RosA2α' identified fragment ions on either side of Cys13 and Cys33, which indicated that these regions are not protected or involved in lanthionine rings. Furthermore, it was noted that RosA2α' underwent alkylation only upon its reduction with TCEP (Fig. S2), and MS/MS analysis of the alkylated RosA2α' also identified Cys13 and Cys33 residues to be carbamidomethylated (**Fig. S4**). Identification of distant cysteines as disulfide partners indicated RosA2α' to be having a conformationally constrained structure. Also, no fragment ions were observed in two long stretches indicating the presence of thioether bridges (Fig. 4A). For RosA1β', MS/MS data along with the absence of alkylation, suggested cyclization of all the six cysteines into lanthionine rings (Fig. 4B). Fragment ions were majorly observed for the N-terminus and a few ions in the C-terminus indicating lanthionine protection (**Table S3**).

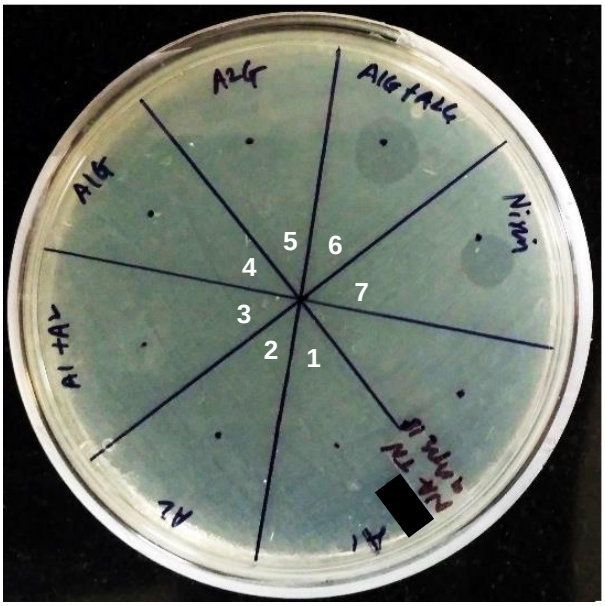


**Fig. 5. RosA2α' and RosA1β' displayed synergistic antimicrobial activity against *L. monocytogenes* 839.** The $His_6$-RosA1 and $His_6$-RosA2 peptides each were assayed at a concentration of 30 uM. A 40 µl volume was used for the spot of a single peptide, while 20 µl of each was used for spots containing both the peptides; nisin was assayed using 3 µl of a 300 µM solution. Sector-1, $His_6$-mRosA1; Sector-2, $His_6$-mRosA2; Sector-3, combined $His_6$-mRosA1 and $His_6$-mRosA2; Sector-4, GluC digest of $His_6$-mRosA2 (to generate RosA2α'); Sector-5, GluC digest of $His_6$-RosA1 (to generate RosA1β'); Sector-5, combined digests (RosA2α' and RosA1β'); Sector-7, nisin.

**RosA2α' and RosA1β' display synergistic antimicrobial activity**

RosA peptides were tested for their antimicrobial activity against *L. monocytogenes* MTCC 839, *Bacillus subtilis* MTCC 121, *Staphylococcus aureus* MTCC 1430. *Pseudomonas aeruginosa* MTCC 1934, *Escherichia coli* MTCC 1610 and *Vibrio cholerae* MTCC 3904.

RosA2α' and RosA1β' did not display any antimicrobial activity when tested individually, but a zone of inhibition was observed against Gram-positive bacteria when both the peptides were spotted together (**Table 1**). Uncleaved $His_6$-mRosA1 and $His_6$-mRosA2 peptides alone and in combination did not display any antimicrobial activity, as was expected from a leader peptide attached lantibiotic (**Fig. 5**). Such a synergistic antimicrobial activity of two separate post-translationally modified peptides which display little to no activity alone is a characteristic of two-component lantibiotics (Navaratna et al., 1998). In general, synergistic activity was observed against Gram-positive bacteria except for *S. aureus* MTCC 1430 where no activity was observed. Another exception was the observation of a zone of inhibition with RosA2α' alone against *B. subtilis* MTCC 121, which enhanced to a distinct zone with the addition of RosA1β' (not shown).

**Table 1. Bioactivity spectrum of roseocin**. Roseocin displayed synergistic antimicrobial activity against most of the Gram-positive bacteria and displayed weak to no antimicrobial activity against Gram-negative bacteria.

| **Indicator strain** | **RosA2α'** | **RosA1α'** | **Combined** |
|---|---|---|---|
| *Listeria monocytogenes* MTCC 839 | - | - | +++ |
| *Bacillus subtilis* MTCC 121 | ++ | - | +++ |
| *Staphylococcus aureus* MTCC 1430 | - | - | - |
| *Pseudomonas aeruginosa* MTCC 1934 | - | - | - |
| *Escherichia coli* MTCC 1610 | - | - | + |
| *Vibrio cholerae* MTCC 3904 | - | - | - |

Antimicrobial activity (+) weak, (++) moderate, (+++) high

## DISCUSSION

Lantibiotics are believed to be the solution for the current problem of antimicrobial resistance and hence researchers are discovering novel lantibiotics through various approaches. The earlier approach was dependent upon culture-based activity screening with the limitation of re-discovery of known compounds, being time-consuming and is subject to the selective conditions in which the culture can be induced for the secondary metabolite production. However, the currently followed *in silico* approach promises the identification of potential producers which can then be confidently taken up for wet lab experiments. In our initial attempt to identify class II lantibiotic biosynthetic gene clusters encoded in the sequenced bacteria, we reported several novel putative lantibiotic clusters. Two of these clusters reported in our study have already been characterized to encode two-component lantibiotics, bicereucin and flavecins [10,12]. The lantibiotic cluster characterized in the current study attracted our attention for its special attributes like a single RosM for processing of two RosA precursor peptides

which are homologs of two-LanM processed two-component lantibiotics (Fig. 1). These peptides have altogether different core region, extended C-terminus and excessive Ser/Thr and Cys residues. The putative two-component lantibiotic from an actinomycete was expected to be having more lanthionine rings than already reported lantibiotics which are isolated from *firmicutes* and hence would probably be more efficacious.

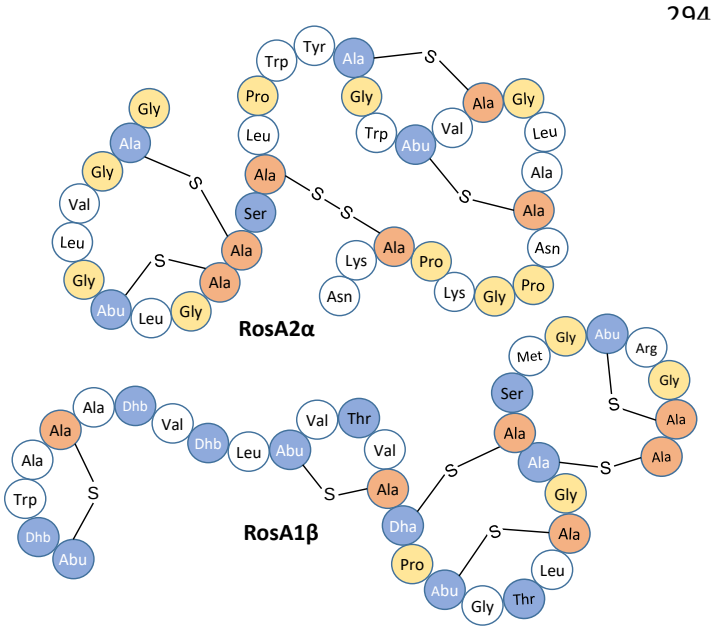


**Fig. 6. Structure of RosA2α and RosA1β.** Predicted structures of RosA2α and RosA1β as per the fragmentation and mass shifts obtained in the tandem MS analysis. RosA2α contains a disulfide bond between Cys13 and Cys33 that flank two of the four lanthionine rings. RosA1β is installed with six lanthionine rings and its structure was predicted using Halβ as the template, tandem MS analysis, and alkylation assay.

In a previous study to interrogate the biosynthetic capacity of the potential producers, Kersten *et al*. (2011) executed a MS-guided genome mining approach for natural product peptidogenomics in eight different *Streptomyces* strains, which are the known producers of various antimicrobial natural products [21]. The enormous MS data of *Streptomyces roseosporus* NRRL 15998 provided by them mention the presence of an identical cluster with two precursor peptide (73 aa and 80 aa, respectively) encoding genes, but this previous study didn't focus on the bioactivity analysis of the lanthipeptides. Also, the product of only one of the two LanA genes could be detected by them from the n-butanolic extract of *S. roseosporus*, and in the absence of the second gene product, it could not be studied either. However, our successful heterologous production of the two modified RosA peptides by *in vivo* promiscuous activity of RosM, followed by *in-vitro* leader peptide removal for generation of bioactive lanthipeptides led to the structure prediction and bioactivity evaluation of the encoded two-component lantibiotic roseocin.

It was initially speculated that RosA2α' with six cysteines would form five lanthionine rings due to the availability of only five Ser/Thr residues in the precursor peptide. Mass spectrometry analysis of RosA2α' confirmed only four dehydration of the total five Ser/Thr residues further limiting the number of lanthionine rings that could form to only four, leaving two Cys

uncyclized (Fig. 2A and 2C). The structure of RosA2α, as suggested by sequence homology and tandem MS analysis, was unlike any other previously known α-component. The Cys13 and Cys33 residues were not found cyclized in the thioether bridges as these were alkylated by Iodoacetamide treatment (Fig. 3 and 4A). Moreover, alkylation was observed only in the presence of a reducing agent which indicated that these residues are involved in a disulfide bridge and hence are in the oxidized state naturally (Fig. S2). While, other two-component lantibiotics like Halα, Plwα, and Enwα have a disulfide bond between the two nearby cysteine residues, in the case of RosA2α' these cysteine residues are distant and found to be flanking two thioether bridges (**Fig. 6**). The presence of this disulfide bond in RosA2α' contributed to a globular structure and prevented fragmentation for sequence analysis. Interestingly, the conserved lipid II binding CTxTxEC motif found in the α-component of all the haloduracin related two-component lantibiotics is not present in RosA2α' peptide. It is perplexing in light of its importance for interaction with lipid II target and its conservation in all the α-peptides characterized till date.

RosA1β' with twelve Ser/Thr was speculated to undergo sufficient dehydration to form lanthionine rings with six cysteines. The mass analysis confirmed nine dehydrations of the total twelve residues (Fig. 2B and 2D), which were still sufficient for six cysteines to undergo Michael addition to these residues. Indeed, the presence of six lanthionine rings was confirmed by the absence of alkylation upon iodoacetamide treatment. The structure of RosA1β' as suggested by sequence homology and tandem MS data is consistent with those of the other β-peptides for at least first three rings on the N-terminus (Fig. 1D and 4B). Besides being an elongated peptide, RosA1β' contains two additional lanthionine rings than the usual four in other β-peptides of the two-component lantibiotics. The presence of extra lanthionine rings could endow the peptide with a more constrained and stable structure, and the length extension might allow better spanning of the bacterial membrane. As with the presence of multiple sites for Michael addition (nine dehydrated residues for six cysteines), RosA1β' could undertake myriad of configurations. Lack of fragmentation from the middle of the peptide till the C-terminus, probably due to thioether bridge overlaps made the assignment of the ring structure in this region difficult (Fig. 4B). Additionally, as the C-terminal is extended and does not show homology with known lantibiotics a conclusive rings pattern could only be predicted. Also, the methyllanthionine encoded by the Dhx-Dhx-Xxx-Xxx-Cys motif possibly has LL stereochemistry, instead of the usual DL stereochemistry. This motif has previously been associated with the formation of LL stereochemistry in flavecins [12] and carnolysin[9].

As with the other lantibiotics, roseocin displayed synergistic antimicrobial activity essentially against the Gram-positive bacteria (Table 1), with the four- and twelve-residue long traces of the leader peptides still attached to RosA2α' and RosA1β', respectively (Fig. 5). The extent of proteolytic processing can be a determining factor for the bioactivity and hence, it is speculated that an enhancement would be observed upon removal of the remaining trace residues from roseocin peptides. In case of the class I lantibiotic gallidermin a similar length 12 aa long leader trace prevented any antimicrobial activity and removal of this trace was necessary for it to be antimicrobial [22]. In some of the type II lantibiotic clusters, two proteases are responsible for proteolytic processing of the modified lanthipeptide. For example, in case of two-component lantibiotic lichenicidin, the Bliβ peptide is processed by the bifunctional $LicT_p$ followed by subtilisin family protease LicP, for removal of a six-residue stretch in order for it to get activated [13]. Previous studies where commercial protease was utilized for leader peptide cleavage, relied on N-aminopeptidase for removal of the trace left behind by GluC treatment [12,23], which was hindered in case of RosA peptides, as N-aminopeptidase activity is obstructed by the aspartate residues present in the trace of RosA1β'. In future, we aim to co-express $RosT_p$ in a compatible plasmid for *in vivo* cleavage of the leader peptide, thus leading to extracellular transport of the fully processed roseocin peptides, RosA2α and RosA1β [24]. However, the purification strategy will have to be worked-out accordingly because of the absence of the hexahistidine tag.

A recent study described large-scale analysis of lanthipeptide biosynthetic gene clusters from actinobacteria and grouped these clusters into gene cluster families (GCF) [14]. Putative two-component lantibiotics with twelve members were grouped in Lant_GCF.30 (encoding two LanM and two LanA substrates), none of which have been characterized experimentally. Additionally, the currently known few lantibiotics reported from actinobacteria are all single peptide lantibiotics. Here, we characterized a Lant_GCF.73 group member as a two-component lantibiotic from *S. roseosporus* NRRL 11379, which comprises of a single RosM for the synthesis of RosA2α and RosA1β. The alpha peptide is installed with four lanthionines, and a disulfide bridge, while the beta-component comprised of six thioether bridges. Looking for the homologous clusters, we identified a biosynthetic gene cluster encoding a natural variant of roseocin in *Glycomyces harbinensis* [25]. A study on such variants is important to understand the role of modifications that affect bioactivity, for designing semisynthetic derivatives with improved characteristics [26], which we plan to carry out in future. Overall, roseocin is the first example of a two-component lantibiotic from a non-*firmicute*.

# METHODS

## CONTACT FOR REAGENT AND RESOURCE SHARING

Further information and requests for resources and reagents should be directed to and will be fulfilled by the Lead Contact, Dipti Sareen (diptsare@pu.ac.in).

## EXPERIMENTAL MODEL AND SUBJECT DETAILS

*Streptomyces roseosporus* NRRL 11379 was grown in ISP2 medium (1 liter media with 4 g yeast extract, 10 g malt extract and 4 g dextrose at pH 7) at 200 rpm, 28$^{o}$C for genomic DNA isolation. All the *E. coli* strains were grown in LB at 37$^{o}$C, 200 rpm with or without kanamycin, as required. Indicator strains were cultured in nutrient broth at 37$^{o}$C, 200 rpm, except for *B. subtilis* which was cultured at 30$^{o}$C. Ready-to-use and individual media components were purchased from Himedia.

## METHOD DETAILS

### Cloning of *rosA1* and *rosA2* with *rosM* in pRSFDuet1 vector

Genomic DNA was isolated using a previously described method [27] from *S. roseosporus* NRRL 11379. The three genes were PCR amplified using respective primers with restriction sites, on a *BIO-RAD* MyCycler$^{TM}$ Thermal Cycler using Q5 DNA polymerase. Gene for RosA1 and RosA2 was inserted in MCS-1 of pRSFDuet-1 using BamHI and HindIII sites, and RosM in MCS-2 using NdeI and XhoI restriction sites. Chemically competent *E. coli* DH10B was prepared, transformed by heat shock and selected on a kanamycin plate. Colonies carrying recombinant plasmid (pRSFDuet-*rosA1-rosM* and pRSFDuet-*rosA2-rosM*) were screened by colony PCR, which was followed by gene sequencing of the complete ORF from SciGenome labs using appropriate vector-specific primers (sequencing primers in **Table S4**). The 3315 bp long *rosM* was additionally sequenced using gene-specific primers (M1, M2, and M3). Sequencing results were analyzed on FinchTV and gene sequence was found 100% identical to the sequence on NCBI. These constructs were then transformed into *E. coli* BL21(DE3) for production of modified RosA peptides.

### Production and purification of *in vivo* modified RosA peptides

A single colony was inoculated in 10 mL LB for overnight. The culture was used to inoculate 2-L of LB added with 50 μg/mL of kanamycin. The culture was incubated at 37$^{o}$C, 200 rpm until the $A_{600}$ reached 0.6-0.8, and the temperature was lowered at this point to 18$^{o}$C and induced with 0.1 mM IPTG for an additional 24 hrs. The cells were harvested by centrifugation at 5000 g for 15

min. The cell pellet was resuspended in 50 mL start buffer (50 mM Tris-HCl, pH 8.0, 500 mM NaCl, 1 mM Imidazole, 1 mM PMSF), and cell lysis was carried out using Sonics VC 505 sonicator. The lysate was cleared by centrifugation at 25000 g for 30 min at 4$^{o}$C (on *SIGMA 3K30* using 12158 rotor). The cleared lysate was loaded on a manually packed 2 mL Ni Sepharose$^{TM}$ High Performance (GE Healthcare) metal affinity resin. The resin was washed with wash buffer (start buffer with 30 mM imidazole), and the peptide was eluted in elution buffer (wash buffer containing 500 mM imidazole). Eluted fractions were analyzed on 16%/6M urea tricine SDS-PAGE [28]. Further purification and desalting of the eluted fractions were done with an Agilent 300 SB-C18 semi-preparative column on an Agilent 1260 infinity series HPLC system. Sample loading and desalting was done with mobile phase A (5:95 ACN:$H_2O$, with 0.1% TFA) and gradient of 0-60% of solvent B (95:5 ACN:$H_2O$, with 0.085% TFA) at 4 mL/min was used to elute the peptides. Elution was monitored at 220 and 280 nm. Modified $His_6$-mRosA1and $His_6$-mRosA2 started eluting out at 43% and 40% of mobile phase B, respectively. Collected fractions were lyophilized and re-dissolved in MilliQ water for further analysis.

**Molecular weight and sequence analysis of RosA peptides**

Matrix Assisted Laser Desorption/Ionization time-of-flight mass spectrometry (MALDI-TOF MS) was carried for determining accurate masses and post-translational modifications in the peptides, on a Bruker ultrafleXtreme$^{TM}$ MALDI-TOF/TOF system maintained at CSIC, PGIMER, Chandigarh. Intact peptides with mass of >5 kDa were analyzed in linear mode and proteolytic digests (peptides of <5 kDa) were analyzed in the reflectron mode. HPLC purified peptides (in 50:50 ACN:$H_2O$ 0.1% TFA) were mixed with sinapinic acid and alpha-Cyano-4-hydroxycinnamic acid (5 mg/mL in 50:50 ACN:$H_2O$, 0.1% TFA) for analysis in linear and reflectron mode, respectively. Mass spectra were recorded in positive ion mode. The Bruker flexControl was used for data acquisition, and SeeMS and mMass program were used for data analysis.

**IAA modification of the RosA peptides**

The purified peptides dissolved in 50 mM tris HCl pH 8.0 were incubated with 1 mM TCEP at 37$^{o}$C for 30 min followed by addition of 5 mM IAA for 60 min (kept in dark). The sample was processed to remove unbound IAA and DTT with Pierce$^{TM}$ C18 spin column and analyzed with MALDI-TOF MS.

**Preparation of bioactive roseocin and growth inhibition assay**

Purified peptides were dissolved in sterile MilliQ grade water and treated with 0.1 µg/µl endoproteinase GluC (P8100, NEB) in 20:1 ratio at a final concentration of 30 µM for 16 h at

37°C. The RosA peptide digests were directly used for activity testing against the indicator strains. Indicator strains were grown to 0.1 OD before spreading Nisin was prepared at 300 µM concentration by dissolving 40 mg of 2.5% nisin powder (Sigma - N5764) in 0.05% acetic acid.

## QUANTIFICATION AND STATISTICAL ANALYSIS

Protein quantitation was done with densitometry using lysozyme as a standard with ImageJ 1.48v software.

## KEY RESOURCES

| REAGENT or RESOURCE | SOURCE | IDENTIFIER |
|---|---|---|
| **Antibodies** | | |
| | | |
| **Bacterial and Virus Strains** | | |
| *Streptomyces roseosporus* | NRRL | NRRL 11379 |
| *Bacillus subtilis* | MTCC | MTCC 121 |
| *Staphylococcus aureus* subsp. *aureus* | MTCC | MTCC 1430 |
| *Escherichia coli* | MTCC | MTCC 1610 |
| *Pseudomonas aeruginosa* | MTCC | MTCC 1934 |
| *Listeria monocytogenes* | MTCC | MTCC 839 |
| *Vibrio cholerae* (classical O1) | MTCC | MTCC 3904 |
| **Biological Samples** | | |
| | | |
| **Chemicals, Peptides, and Recombinant Proteins** | | |
| Endoproteinase, GluC | NEB | Cat#P8100S |
| Q5® High-Fidelity DNA Polymerase | NEB | Cat#M0491S |
| PageRuler Low Range Unstained Protein Ladder | Thermo Scientific | Cat#26632 |
| Nisin | *Sigma-Aldrich* | Cat# N5764-1G |
| $His_6$-mRosA1 | This paper | N/A |
| $His_6$-mRosA2 | This paper | N/A |
| **Critical Commercial Assays** | | |
| | | |
| **Deposited Data** | | |
| | | |
| **Experimental Models: Cell Lines** | | |
| | | |
| **Experimental Models: Organisms/Strains** | | |
| *E. coli* BL21 (DE3) pRSFDuet-*rosA1-rosM* | This paper | N/A |
| *E. coli* BL21 (DE3) pRSFDuet-*rosA2-rosM* | This paper | N/A |
| **Oligonucleotides** | | |
| Primer: *rosA1* forward with BamHI site: cagtg**ggatcc**GATGAACCTTGTTCGCGCATGGA | This paper | N/A |
| Primer: *rosA1* reverse with HindIII site: cagtg**aagctt**TCAGCAACAGCCGCGCGT | This paper | N/A |
| Primer: *rosA2* forward with BamHI site: cagtg**ggatcc**GATGGACATAGTCCGGTCCTGGA | This paper | N/A |

| | | |
|---|---|---|
| Primer: *rosA2* reverse with HindIII site: cagtg**aagctt**CTAGTTCTTGCAGGGCTTACCGG | This paper | N/A |
| Primer: *rosM* forward with NdeI site: tgcag**catATG**CCTGACGACGCAAGC | This paper | N/A |
| Primer: *rosM* forward with NdeI site: acgtg**ctcgag**TCAGCACGTCGTCCTCCC | This paper | N/A |
| See Table S4 for primers used in sequencing | | |
| **Recombinant DNA** | | |
| pRSFDuet-1 | Merck Millipore | Cat#71341-3 |
| pRSFDuet-*rosA1-rosM* | This paper | N/A |
| pRSFDuet-*rosA2-rosM* | This paper | N/A |
| **Software and Algorithms** | | |
| Proteomics Toolkit: MS/MS Fragment Ion calculator | Institute for Systems Biology, Seattle | http://db.systemsbiology.net/ |
| FinchTV | Geospiza, Inc. | https://digitalworldbiology.com/FinchTV |
| ImageJ 1.48v | NIH | https://imagej.nih.gov/ij/ |
| SeeMS 3.0.18194.0 | 29 | http://proteowizard.sourceforge.net/ |
| mMass version 5.5.0 | 30 | http://www.mmass.org/ |
| Cello v.2.5 | 31 | http://cello.life.nctu.edu.tw/ |
| TMHMM server v.2.0 | 32 | http://www.cbs.dtu.dk/services/TMHMM/ |
| **Other** | | |
| | | |

## SUPPLEMENTAL INFORMATION TITLES AND LEGENDS

Supplemental information includes 4 figures, and 5 tables.

## AUTHOR CONTRIBUTIONS

Supervision D.S.; Resources D.S. Conceptualization D.S. and M.S.; Methodology D.S. and M.S.; Investigation M.S.; Formal analysis M.S.; Writing-original draft, M.S.; Writing-Review & Editing, M.S and D.S. Funding acquisition M.S. and D.S.

## ACKNOWLEDGMENTS

MS acknowledges the independent Senior Research Fellowship (SRF) No. 3/1/3/JRF-2011/HRD-99(11005), awarded by the Indian Council of Medical Research, Government of India, New Delhi. The financial assistance received to the lab from DST-PURSE and UGC-SAP grant, Government of India, New Delhi, is also acknowledged. MALDI-TOF MS facility maintained at CSIC, PGIMER, Chandigarh is acknowledged.

## REFERENCES


1. Basi-chipalu, S. a Review : Lantibiotics , a Promising Antimicrobial Agent. **21,** 119–128 (2016).
2. Ueda, K. *et al.* AmfS, an extracellular peptidic morphogen in *Streptomyces griseus*. *J. Bacteriol.* **184,** 1488–1492 (2002).
3. Férir, G. *et al.* The Lantibiotic Peptide Labyrinthopeptin A1 Demonstrates Broad Anti-HIV and Anti-HSV Activity with Potential for Microbicidal Applications. *PLoS One* **8,** e64010 (2013).
4. Iorio, M. *et al.* A glycosylated, labionin-containing lanthipeptide with marked antinociceptive activity. *ACS Chem. Biol.* **9,** 398–404 (2014).
5. Iftime, D. *et al.* Streptocollin, a Type IV Lanthipeptide Produced by *Streptomyces collinus* Tü 365. *ChemBioChem* **16,** 2615–2623 (2015).
6. Wang, J. *et al.* Cerecidins, novel lantibiotics from *Bacillus cereus* with potent antimicrobial activity. *Appl. Environ. Microbiol.* **80,** 2633–2643 (2014).
7. Bakhtiary, A. *et al.* Insights into the Mechanism of Action of the Two-Peptide Lantibiotic Lacticin 3147. *J. Am. Chem. Soc.* **139,** 17803–17810 (2017).
8. Booth, M. C. *et al.* Structural analysis and proteolytic activation of *Enterococcus faecalis* cytolysin, a novel lantibiotic. *Mol. Microbiol.* **21,** 1175–1184 (1996).
9. Lohans, C. T., Li, J. L. & Vederas, J. C. Structure and biosynthesis of carnolysin, a homologue of enterococcal cytolysin with D-amino acids. *J. Am. Chem. Soc.* **136,** 13150–13153 (2014).
10. Huo, L. & van der Donk, W. A. Discovery and characterization of bicereucin, an unusual D-amino acid-containing mixed two-component lantibiotic. *J. Am. Chem. Soc.* **138,** 5254–7 (2016).
11. McClerren, A. L. *et al.* Discovery and *in vitro* biosynthesis of haloduracin, a two-component lantibiotic. *Proc. Natl. Acad. Sci. U. S. A.* **103,** 17243–8 (2006).
12. Zhao, X. & van der Donk, W. A. Structural Characterization and Bioactivity Analysis of the Two-Component Lantibiotic *Flv* System from a Ruminant Bacterium. *Cell Chem. Biol.* **23,** 1–11 (2016).
13. Caetano, T., Krawczyk, J. M., Mösker, E., Süssmuth, R. D. & Mendo, S. Heterologous expression, biosynthesis, and mutagenesis of type II lantibiotics from *Bacillus licheniformis* in *Escherichia coli*. *Chem. Biol.* **18,** 90–100 (2011).
14. Zhang, Q., Doroghazi, J. R., Zhao, X., Walker, M. C. & van der Donka, W. A.

Expanded natural product diversity revealed by analysis of lanthipeptide-like gene clusters in Actinobacteria. *Appl. Environ. Microbiol.* **81,** 4339–4350 (2015).

15. Singh, M. & Sareen, D. Novel LanT associated lantibiotic clusters identified by genome database mining. *PLoS One* **9,** e91352 (2014).
16. Begley, M., Cotter, P. D., Hill, C. & Ross, R. P. Identification of a novel two-peptide lantibiotic, lichenicidin, following rational genome mining for LanM proteins. *Appl. Environ. Microbiol.* **75,** 5451–5460 (2009).
17. Shi, Y., Yang, X., Garg, N. & van der Donk, W. a. Production of lantipeptides in *Escherichia coli*. *J. Am. Chem. Soc.* **133,** 2338–41 (2011).
18. Lowther, W. T. & Matthews, B. W. Structure and function of the methionine aminopeptidases. *Biochim. Biophys. Acta - Protein Struct. Mol. Enzymol.* **1477,** 157–167 (2000).
19. Xie, L. *et al.* Lacticin 481: in vitro reconstitution of lantibiotic synthetase activity. *Science* **303,** 679–681 (2004).
20. Geoghegan, K. F. *et al.* Spontaneous α-N-6-phosphogluconoylation of a 'His tag' in *Escherichia coli*: The cause of extra mass of 258 or 178 Da in fusion proteins. *Anal. Biochem.* **267,** 169–184 (1999).
21. Kersten, R. D. *et al.* A mass spectrometry-guided genome mining approach for natural product peptidogenomics. *Nat. Chem. Biol.* **7,** 794–802 (2011).
22. Valsesia, G., Medaglia, G., Held, M., Minas, W. & Panke, S. Circumventing the effect of product toxicity: Development of a novel two-stage production process for the lantibiotic gallidermin. *Appl. Environ. Microbiol.* **73,** 1635–1645 (2007).
23. Garg, N., Tang, W., Goto, Y., Nair, S. K. & Donk, W. A. Van Der. Lantibiotics from *Geobacillus thermodenitrificans*. *Proc. Natl. Acad. Sci. U S A* **109,** 5241–6 (2012).
24. Wang, J., Ge, X., Zhang, L., Teng, K. & Zhong, J. One-pot synthesis of class II lanthipeptide bovicin HJ50 via an engineered lanthipeptide synthetase. *Sci. Rep.* **6,** 38630 (2016).
25. Labeda, D. P., Testa, R. T., Lechevalier, M. P. & Lechevalier, H. A. Glycomyces, a New Genus of the Actinomycetales. *Int. J. Syst. Bacteriol.* **35,** 417–421 (1985).
26. Gomes, K. M., Duarte, R. S. & Bastos, M. do C. de F. Lantibiotics produced by Actinobacteria and their potential applications (A review). *Microbiol. (United Kingdom)* **163,** 109–121 (2017).
27. Kumar, R., Wani, S. I., Chauhan, N. S., Sharma, R. & Sareen, D. Cloning and characterization of an epoxide hydrolase from *Cupriavidus metallidurans*-CH34.

*Protein Expr. Purif.* **79,** 49–59 (2011).

28. Schägger H. Tricine – SDS-PAGE. *Nat. Protoc.* **1,** 16–23 (2006).

29. Chambers, M. C. *et al.* A cross-platform toolkit for mass spectrometry and proteomics. *Nat. Biotechnol.* **30,** 918–920 (2012).

30. Niedermeyer, T. H. J. & Strohalm, M. mMass as a Software Tool for the Annotation of Cyclic Peptide Tandem Mass Spectra. *PLoS One* **7,** e44913 (2012).

31. Yu, C., Chen, Y., Lu, C. & Hwang, J. Prediction of protein subcellular localization. *Proteins.* **64,** 643–651 (2006).

32. Krogh, A., Larsson, B., von Heijne, G. & Sonnhammer, E. L. Predicting transmembrane protein topology with a hidden Markov model: application to complete genomes. *J. Mol. Biol.* **305,** 567–580 (2001).