

# Reconstruction of the lantibiotic ruminococcin-A biosynthesis machinery in *Escherichia coli* and structural characterization

Elvis Legala Ongey - Dissertation

# **Reconstruction of the lantibiotic ruminococcin-A biosynthesis machinery in *Escherichia coli* and structural characterization**

vorgelegt von  
M. Sc. Elvis Legala Ongey  
geb. in Bali-Nyonga, Kamerun

von der Fakultät III-Prozesswissenschaften  
der Technischen Universität Berlin  
zur Erlangung des akademischen Grades

Doktor der Naturwissenschaften  
- Dr. rer. nat. -  
genehmigte Dissertation

Promotionsausschuss:

Vorsitzender: Prof. Dr. Prof. Dr. Jens Kurreck-TU Berlin  
Erstgutachter: Prof. Dr. Peter Neubauer, Institut für Biotechnologie - TU Berlin  
Zweitgutachter: Prof. Dr. Nediljko Budisa, Institut für Chemie - TU Berlin  
Drittgutachter: Prof. Dr. Oscar P. Kuipers, Groningen, The Netherlands

Tag der wissenschaftlichen Aussprache: 02. August 2018

Berlin, 2018







## Abstract

Lanthipeptides are a group of therapeutically relevant proteinaceous compounds predominantly produced by Gram-positive bacteria at the ribosomal level, containing thioether cross-linked amino acids called lanthionine as a characteristic structural element. Lanthipeptides are attracting growing interests stemming from their wide range of bioactivities including antimicrobial, antiviral and anticancer. However, purifying them from native sources is challenging due to low production yields and cultivation difficulties that accompany the natural producers. Ruminococcin-A (RumA) is an example which is naturally produced by the obligate anaerobe *Ruminococcus gnavus* E1. Culturing *R. gnavus* E1 is obviously challenging and would certainly negatively affect the development of a stable high-quality production process, limiting biotechnological development and therapeutic exploitation of RumA. In this study, we amplified parts of the RumA biosynthesis operon from *R. gnavus* E1, encoding the linear precursor peptide (preRumA) and the lanthionine synthetase (RumM), and coexpressed them in the heterologous bacterial host *E. coli*.

Our results show that RumM catalyzed the introduction of dehydroamino acids into the core peptide of preRumA and subsequently conjugated the dehydrated residues with cysteine to generate thioether bridges. This was achieved with preRumA expressed as a chimeric fusion protein together with GFP. Results here demonstrate that fusing a larger protein carrier to the N-terminus of the leader peptide does not necessarily obstruct *in vivo* processivity of the lanthionine-generating enzyme in modifying its substrate. A strong interaction was observed between the chimeric fusion product and RumM, supplying some interesting insights into the catalytic mechanisms of class II lanthionine-generating enzymes. Characterizing the structure of preRumA further revealed the presence of three thioether rings, contradicting previous report which concluded that the formation of a third thioether bridge was not possible. Modified preRumA was activated *in vitro* by removing the leader peptide using trypsin to yield the active product and biological activity was achieved. A production yield of 6 mg of pure modified preRumA per litre of *E. coli* culture was attained. Considering the size ratio of the leader-to-core segments of preRumA, this amount would produce a final yield of approximately 1-2 mg of RumA when the leader peptide is removed. This yield exceeds the amount of RumA purified from the native host in the order of  $10^4$ .

This study supplies a system that may be applied as a useful tool for studying peptide engineering to generate analogues of RumA or completely new peptides with superior anti-infective properties. The current system also provides an alternative strategy to derive mechanistic insights on the complex modification machinery of lanthipeptides.

## Zusammenfassung

Lanthipeptide sind eine Gruppe von therapeutisch bedeutsamen proteinogenen Naturstoffen, die hauptsächlich von Gram-positiven Bakterien ribosomal produziert werden. Als kennzeichnendes Strukturelement enthalten sie über Thioether verknüpfte Aminosäuren, sogenannte Lanthionine. Auf Grund ihres breiten Wirkspektrums, welches Bioaktivität sowohl gegen Bakterien, Viren als auch Krebszellen einschließt, rücken Lanthipeptide zunehmend in den Fokus der Forschung. Allerdings ist die Gewinnung von Lanthipeptiden aus deren natürlichen Produzenten äußerst schwierig, da diese nur in geringem Maße produziert werden und die Kultivierung der entsprechenden Bakterien auf Grund ihrer Lebensweise schwierig ist. Ein Beispiel hierfür ist Ruminococcin-A (RumA), das vom Bakterium *Ruminococcus gnavus* E1 produziert wird. Eine Kultivierung von *R. gnavus* ist auf Grund seiner obligat anaeroben Lebensweise schwierig, wodurch die Entwicklung eines stabilen, hoch effizienten Produktionsprozesses erschwert und die biotechnologische Entwicklung bzw. die therapeutische Erforschung limitiert wird. Im Rahmen dieser Arbeit wurden Teile des RumA Biosyntheseoperons aus *R. gnavus* E1 amplifiziert und im heterologen Wirtsorganismus *Escherichia coli* exprimiert. Hierzu zählen die Gene für das lineare „Precursor“-Peptid (preRumA) sowie die Lanthionine-Synthetase (RumM).

Das Enzym RumM katalysiert die Dehydratisierung einzelner Aminosäuren im Kernpeptide von preRumA und anschließend die Bildung der typischen Thioether Brücken durch Konjugation der dehydratisierten Reste mit einem Cystein. Hierfür wurde ein Fusionsprotein aus preRumA und dem grünfluoreszierenden Protein (GFP) produziert. Die Resultate zeigen, dass die Fusion eines größeren Trägerproteins an den N-Terminus des RumA „Leaderpeptides“ dessen Prozessierung *in vivo* durch RumM nicht negativ beeinflusst. Die beobachtete starke Interaktion zwischen dem chimären Fusionsprotein und RumM ermöglicht einen interessanten Einblick in den katalytischen Mechanismus der Klasse II Lanthionin-Synthetasen. Im Gegensatz zur früheren Annahme, dass die Bildung eines dritten Thioetherrings nicht möglich ist, konnte durch die Aufklärung der Struktur von preRumA das Vorhandensein von drei Thioetherringen nachgewiesen werden. Die Aktivierung des modifizierten preRumA erfolgte *in vitro*. Das Leaderpeptid wurde hierfür durch Spaltung mit Trypsin entfernt und so das biologisch aktive Produkt gewonnen. Aus einem Liter *E. coli* Kultur konnten hierbei bis zu 6 mg modifiziertes preRumA gereinigt werden. Bedenkt man das Größenverhältnis zwischen Kernpeptid und Leaderpeptid entspricht diese Menge einer finalen Ausbeute von ca. 1-2 mg aktivem RumA nachdem das Leaderpeptid entfernt wurde. Dies übersteigt die Menge an RumA, die aus dem natürlichen Produzenten gewonnen werden kann, um das 10.000 fache.

Das im Rahmen dieser Arbeit entwickelte Expressionssystem stellt ein nützliches Werkzeug für die Untersuchung von Peptiden dar. Es könnte als Grundlage für die Expression von RumA Analoga oder anderer neuer bioaktiver Peptide mit unter Umständen besseren antiinfektiven Eigenschaften dienen. Darüber hinaus bietet unser System eine alternative Möglichkeit zur Untersuchung der komplexen Modifikationssysteme von Lanthipeptiden.

The present work was performed from January 2015 – March 2018 in the research group of Prof. Dr. Peter Neubauer (Chair of Bioprocess Engineering) at the Department of Biotechnology, Technische Universität Berlin.

## List of Publications Contributing to this Work

- I. **Ongey, E. L.**; Giessmann, R.; Fons, M.; Rappsilber, J.; Adrian, L. and Neubauer, P. (2018) Heterologous biosynthesis, modifications and structural characterization of ruminococcin-A, a lanthipeptide from a gut bacterium, in *E. coli*. *Front Microbiol* 9:1688. <https://doi.org/10.3389/fmicb.2018.01688>
- II. **Ongey E. L.**, Pflugmacher S., Neubauer P. (2018) Bioinspired designs, molecular premise and tools for evaluating ecological importance of antimicrobial peptides. *Pharmaceuticals*, 11(3), 68; <https://doi.org/10.3390/ph11030068>
- III. **Ongey, E. L.**; Yassi, H.; Pflugmacher, S.; Neubauer, P. (2017) Pharmacological and pharmacokinetic properties of lanthipeptides undergoing clinical studies. *Biotechnol Lett* 39: 473. <https://doi.org/10.1007/s10529-016-2279-9>
- IV. **Ongey, E. L.**; Neubauer, P. (2016) Lanthipeptides: chemical synthesis versus *in vivo* biosynthesis as tools for pharmaceutical production. *Microb Cell Fact* 15: 97. <https://doi.org/10.1186/s12934-016-0502-y>

## List Abbreviations

[ppm]	Parts per million
ABG	Active beam guide
Abu	2-aminobutyric acid
AMP	Antimicrobial peptide
CAI	codon adaptation index
CAP	catabolite activator protein
CID	Collision-induced dissociation
DC	Direct current
Dha	Dehydroalanine or dehydrated serine
Dhb	Dehydrobutyrine or dehydrated threonine
DOT	Dissolved oxygen tension
DTT	Dithiothreitol
EDTA	Ethylenediaminetetraacetic acid
ESI	Electrospray ionization
ETD	Electron transfer dissociation
FGA	Fibrinogen alpha chain
FPP	Farnesyl diphosphate
GRAS	Generally regarded as safe
GST	Glutathione S-transferase
HCD	Higher-energy collisional dissociation
HPLC	High-performance liquid chromatography
IMAC	Immobilized metal ion affinity chromatography
IP	inducer peptide
IPTG	Isopropyl $\beta$ -D-1-thiogalactopyranoside
LAB	Lactic acid bacteria
Lan	Lanthionine ring/product of <i>lan</i> gene
<i>lan</i>	Notation for genes constituting the lanthipeptide biosynthesis cluster
LanB	General notation for class I lanthipeptide dehydratase
LanC	General notation for class I lanthipeptide cyclase
LanKC	General notation for class III lanthipeptide synthetase
LanL	General notation for class IV lanthipeptide synthetase
LanM	General notation for class II lanthipeptide synthetase
LanP	General notation for lanthipeptide activation protease
LanT	General notation for lanthipeptide transport protein
LAP	Linear azol(in)e-containing peptide
LTQ	Linear trap quadrupole
MALDI	Matrix-assisted laser desorption/ionization
ManPTS	Mannose phosphotransferase system
MBP	Maltose binding protein
MeLan	Methyllanthionine ring
MRSA	Methicillin-resistant <i>Staphylococcus aureus</i>

MS	Mass spectrometry
MS <sup>2</sup>	Tandem MS
MTP	Microtiter plate
MUB	Membrane-anchored ubiquitin-fold
NAG	N-acetylglucosamine
nLC	Nano-liquid chromatography
ORF	Open reading frames
<i>ori</i>	Origin of replication
PDB	Protein data bank
preRumA	Ruminococcin-A precursor peptide
PTM	Posttranslational modifications
PVDF	Polyvinylidene fluoride
QS	Quorum sensing
RBS	Ribosomal binding site
RF	Radio frequency
RFU	Relative fluorescent unit
RiPPs	Ribosomally synthesized and post-translationally modified peptides
RT	Room temperature
RumA	Ruminococcin-A
RumM	Ruminococcin-A lanthionine synthetase
SAM	S-adenosylmethionine
SDS-PAGE	Sodium dodecyl sulfate polyacrylamide gel electrophoresis
SRM	Selected Reaction Monitoring
SUMO	Small ubiquitin-like modifier
TEV	Tobacco etch virus
Trx	Thioredoxin
UniProt	Universal protein resource
WHO	World Health Organization

# Table of Contents

<b>Abstract.....</b>	<b>I</b>
<b>Zusammenfassung.....</b>	<b>II</b>
<b>List of Publications Contributing to this Work .....</b>	<b>V</b>
<b>List Abbreviations .....</b>	<b>VI</b>
<b>Table of Contents .....</b>	<b>VIII</b>
<b>1. Introduction .....</b>	<b>1</b>
<b>2. Literature Review .....</b>	<b>4</b>
2.1 The Bacteriocins .....	4
2.2 Classification of bacteriocins .....	5
2.3 The classification of lanthipeptides .....	7
2.3.1 Class I lanthipeptides .....	9
2.3.2 Class II lanthipeptides .....	12
2.3.3 Class III lanthipeptides .....	15
2.3.4 Class IV lanthipeptides .....	17
2.4 Biotechnological production versus chemical synthesis of lanthipeptides .....	19
2.5 Pharmacological properties and therapeutic use of lanthipeptides .....	19
2.6 Recent advances in engineering and heterologous production of lanthipeptides .....	19
2.7 Native biosynthesis and regulation of lantibiotic ruminococcin-A .....	21
2.8 Isolation of ruminococcin-A from <i>R. gnavus</i> E1 .....	23
2.9 Carriers for heterologous expression of AMPs .....	24
2.10 Objective of this work .....	24
<b>3. Materials and Methods .....</b>	<b>27</b>
3.1 Materials .....	27
3.2 Software .....	27
3.3 Bacterial strains, growth and cultivation conditions .....	27
3.3.1 Cultivation media .....	27
3.3.2 Monitoring bacterial growth with a spectrophotometer .....	28
3.3.3 Automated fluorescence measurements .....	28
3.4 Vectors .....	28
3.5 Molecular Biology techniques .....	28
3.5.1 Isolation of <i>R. gnavus</i> E1 Genomic DNA isolation .....	29



3.5.2	Preparations of plasmid DNA .....	29
3.5.3	Oligonucleotides design .....	29
3.5.4	PCR techniques .....	29
3.5.4.1	Analytical PCR .....	29
3.5.4.2	Preparative PCR .....	29
3.5.4.3	Colony PCR.....	30
3.5.4.4	Splicing by overlap extension (SOE) PCR .....	30
3.5.5	Purification of DNA fragments and spectrophotometric quantitation .....	30
3.5.6	Agarose gel electrophoresis .....	31
3.5.7	Conventional cloning using restriction enzyme digestion and ligation.....	31
3.5.8	Modification of expression vectors .....	32
3.5.9	Site-directed mutagenesis.....	33
3.5.10	Construction of expression vectors .....	33
3.5.10.1	Amplification of gene fragment from <i>R. gnavus</i> E1 chromosome .....	33
3.5.10.2	Construction of His6-preRumA expression vectors.....	33
3.5.10.3	His6-preRumA <i>vectors</i> with alternative cleavage site at position -1.....	33
3.5.10.4	Construction of His6-GFP-TEV-preRumA expression vectors.....	34
3.5.10.5	Construction of His6-SUMO-preRumA expression vectors .....	34
3.5.10.6	Construction of His6-RumM expression vectors .....	34
3.5.10.7	Construction of bicistronic expression vectors .....	34
3.5.10.8	Construction of RumT125 expression vectors.....	35
3.5.11	Transformation of competent <i>E. coli</i> cells.....	35
3.6	Protein production and analysis.....	35
3.6.1	Expression optimization in EnPresso B growth system.....	35
3.6.2	Protein Expression in TB medium.....	36
3.6.3	Strain screening using GFP fluorescence .....	36
3.6.4	Cell disruption.....	37
3.6.5	SDS-PAGE analysis .....	38
3.6.6	Native PAGE analysis .....	38
3.6.7	Western blot .....	38
3.7	Protein purification and peptide extraction.....	39
3.7.1	IMAC .....	39
3.7.2	Size exclusion chromatography .....	40
3.7.3	TEV Cleavage .....	40
3.7.4	Extraction of preRumA from TEV-digested product.....	40
3.8	Measurement of protein concentrations using Bradford assay.....	41

3.9	Iodoacetamide derivatization and trypsin digestion.....	41
3.10	Mass spectrometric analyses .....	42
3.10.1	MS sample preparation with ZipTip .....	42
3.10.2	LC-ESI-MS analyses of preRumA.....	42
3.10.3	Tandem mass spectrometric analyses of preRumA .....	42
3.11	Biological Assay .....	43
3.12	Identification of optimal production conditions using online DOT & pH.....	43
3.13	Computational analyses .....	44
<b>4.</b>	<b>Results and Discussion .....</b>	<b>46</b>
4.1	Sequence analysis of ruminococcin-A processing enzymes.....	46
4.1.1	Sequence similarity between RumM and other lanthionine synthetases .....	46
4.1.2	Sequence and structural similarities between RumT and other AMS proteins .....	47
4.1.3	Codon adaptability calculations and analysis.....	48
4.2	Vector construction and <i>E. coli</i> expression .....	51
4.2.1	Features of expression vectors.....	51
4.2.2	Separate expression of His6-preRumA and His6-RumM.....	53
4.2.3	Two-plasmid coexpression system for His6-preRumA and His6-RumM .....	54
4.2.4	MS identification of products from the two-plasmid system .....	56
4.2.5	PreRumA Fused to GFP .....	59
4.2.6	LC-ESI-MS analyses of preRumA from WLEOgrA and WLEOgrA/M .....	62
4.3	Quality enhancement, characterization & activation of preRumA .....	65
4.3.1	Optimized vector for His6-RumM expression .....	65
4.3.2	Expression of RumT peptidase domain .....	66
4.3.3	Plasmid-encoded bicistronic operon .....	68
4.3.4	Purification of His6-GFP-TEV-preRumA* and TEV digestion .....	69
4.3.5	Interactions between His6-RumM and His6-GFP-TEV-preRumA* .....	71
4.3.6	Extraction and nLC-ESI-MS analyses of preRumA* .....	73
4.3.7	Mass Spectrometric Fragmentation and MS <sup>2</sup> Analysis of PreRumA* .....	77
4.3.8	Alkyl derivatization and trypsin digestion of preRumA* .....	83
4.4	Bioassay analysis of trypsin activated RumA.....	85
4.5	Microtiter plate cultivations.....	87
4.5.1	Colony screening in microtiter plate using GFP as a reporter.....	88
4.5.2	Cultivation of WLEOgrA*M1 in 24-well plate using TB medium.....	89
4.6	Strain optimization using multicultivation & screening strategies .....	90
4.6.1	Characteristics of the bicistronic and two-vector systems.....	90

4.6.2	Optimization of WLEOgrA*M1 using process-relevant parameters .....	93
4.7	Overall discussion of results .....	97
<b>5.</b>	<b>Conclusions and outlook .....</b>	<b>102</b>
5.1	Conclusions .....	102
5.2	Outlook .....	103
<b>6.</b>	<b>Appendix .....</b>	<b>105</b>
6.1	Structure-function analyses, description of the biosynthesis operons, classification and physicochemical characteristics of randomly selected bacteriocins .....	105
6.1.1	Class I: Posttranslationally modified peptides (<10 kDa) .....	110
6.1.1.1	<i>Ia—The lanthipeptides</i> .....	110
6.1.1.2	<i>Ib—Circular bacteriocins</i> .....	112
6.1.1.3	<i>Ic—Linear azol(in)e-containing peptides (LAPs)</i> .....	114
6.1.1.4	<i>Id—Sactipeptides</i> .....	115
6.1.1.5	<i>Ie—Glycocins</i> .....	117
6.1.1.6	<i>If—Lasso peptides</i> .....	118
6.1.2	Class II: Unmodified bacteriocins (<10 kDa) .....	119
6.1.2.1	<i>Ila—pediocin-like</i> .....	119
6.1.2.2	<i>Ilb—Two-peptide bacteriocins</i> .....	121
6.1.2.3	<i>Ilc—Leaderless bacteriocins</i> .....	122
6.1.2.4	<i>Ild—Non-pediocin-like single-peptide bacteriocins</i> .....	124
6.1.3	Class III: Unmodified high molecular weight bacteriocins .....	125
6.2	Vector descriptions and vector maps .....	126
6.3	Structure prediction and sequence similarity .....	129
6.4	Codon-optimized gene sequence of the peptidase-encoding domain of RumT (UniProt: Q93JP5).....	132
6.5	MS <sup>2</sup> assignments of His6-preRumA fragments .....	132
6.6	MS <sup>2</sup> characterization of preRumA* .....	134
6.7	Multi-parallel cultivations.....	137
6.8	The concept of mass spectrometry (MS) .....	139
6.8.1	Mass spectrometry and peptide sequencing .....	139
6.8.2	Electrospray ionization .....	140
6.8.3	The Orbitrap Fusion Tribrid mass spectrometer .....	141
6.8.4	The quadrupole mass filter.....	143
6.8.5	Ion trap mass analyzer.....	143
6.8.6	Tandem mass spectrometry (MS <sup>2</sup> ) .....	144
6.9	Tables and Figures for materials and methods section.....	146

6.9.1	Instrumentation used .....	147
6.9.2	Expendables.....	149
6.9.3	Software .....	149
6.10	Purification of TEV .....	156
<b>References .....</b>		<b>157</b>
<b>Acknowledgements.....</b>		<b>179</b>
<b>Short Resume .....</b>		<b>182</b>

# 1. Introduction

Economic sensibilities and healthcare awareness are two major concerns in the world today. Clinical problems that range from minor nosocomial infections to substantially critical diseases such as sepsis, invasive endocarditis, bone and soft tissue intra-surgical infections are examples of complications caused by staphylococci and enterococci forms of bacteria in hospital settings. Methicillin-resistant *Staphylococcus aureus* (MRSA) and vancomycin-resistant enterococci are amongst the most disreputable forms of infectious microorganisms. According to World Health Organization's report in 2014, without potent anti-infective therapy, a host of standard medical procedures will fail or become compromised, resulting in very complex and high risk situations (WHO, 2014). The development of novel potent antimicrobial compounds therefore remains an ensuing challenge for humankind. The situations are becoming more unpleasant as many of the available anti-infective agents are persistently losing their effectiveness against these pathogenic microbes. Some scientists have briefly described the present condition as a "new pre-antibiotic era" (Rios et al., 2016), which is perceived to be largely facilitated by the excessive use of conventional antibiotics and the huge growth of bacterial resistances. This has consequently stimulated considerable interests in developing and applying natural compounds especially antimicrobial peptides (AMPs) as the next generation drug molecules (Zhang and Gallo, 2016).

The historical background of antibiotic research and the evolution of bacterial resistances highlight the necessity to evolve with novel strategies to control pathogenic bacteria. Key events recorded during the past years to bring us to the economic realities we face today and the need for new antibiotics in medicine have been analyzed elsewhere (Kirst, 2013). Interestingly, Kirst holds the opinion that using non-lethal means to control bacteria may be the ultimate antidote to the antibiotics dilemma. Zhang and Gallo on the other hand believe that controlling endogenous expression of AMPs may be an ideal approach for treating a wide range of diseases in humans and animals (Zhang and Gallo, 2016). Either way, both approaches require huge efforts to come to fruition. Attempts to explore uncultured microbes, screening diverse resources and targets have proven that the bacteriocins family of AMPs may be used as alternatives to conventional antibiotics based on their remarkable potencies against drug-resistant bacteria (Cotter et al., 2013).

Nevertheless, a critical understanding of how target macromolecular structures in the cell are recognized by a peptide and how they govern the microbial killing process is necessary to facilitate the design of peptides with superior properties that can hamper the development of resistances which perhaps may grossly restrict their use in medicine (Cotter et al., 2013). We recently summarized molecular targets of some AMPs, identifying their evolutionary

characteristics and indicating how the macromolecules can be useful in designing target-specific peptides (Ongey et al., 2018). The desired application of a compound may provide basis for rational design and production of AMPs with higher tenacity strong enough to escape specific resistance machineries. In this regard, gene-encoded peptide engineering is one advantage that natural peptides have over classical antibiotics. Although theoretical investigation of the pharmacodynamics of AMPs and conventional antibiotic with respect to their ability to cause resistance evolution indicates less likelihood for AMPs (Yu et al., 2017), extensive investigations are desired to obtain more insights into the mechanisms of interaction between complex functional components involved in biosynthesis, translocation and self-immunization in order to prevent these situations from occurring. Furthermore, understanding the concept of biosynthesis, and macromolecular targets and their responses to AMPs' interaction may create awareness of how to apply bioengineering to rationally tune bioactive peptides to enhance properties like biostability, potency and specificity as reported in a few examples of class I bacteriocins, also called lanthipeptides (Field et al., 2015, Field et al., 2015).

Pharmacological studies including but not limited to mode of administration, dosing regimen, half-life, distribution and clearance rates, minimal inhibitory concentration (MIC) and potential drug-serum protein interactions are important to assess therapeutic relevance of bioactive peptides. This is partly because in many cases their intrinsic properties do not favour direct use in medicine as we recently discussed in a review of selected examples of lanthipeptides (Ongey et al., 2017). However, these studies require a good amount of the substance under investigation which appear to be challenging to say the least. The absence of well-characterized strategies for producing these peptides hampers the implementation of diverse experimental protocols to evaluate them. We recently concluded from studying the literature that chemical synthesis is not a viable tool to achieve economically feasible production of ribosomally synthesized peptides with complex modifications, but emphasized on biological system as a realistic source to construct a production process (Ongey and Neubauer, 2016). Nevertheless, yields obtained from natural isolates are low, and most native producers require prolonged periods of cultivation, coupled with the high cost of cultivation strategies and growth media supplements. The inability to resolve these challenges have encouraged the use of heterologous production hosts as the only realistic alternative.

A good example of a lanthipeptide that poses significant expression challenges is ruminococcin-A naturally produced by the gut bacterium *Ruminococcus gnavus* E1. The organism, besides being an obligate anaerobe, large-scale production and peptide engineering possibilities of the peptide may be limited by the relatively large size (12.8 kb) of the

ruminococcin-A biosynthesis gene cluster (Gomez et al., 2002). As reported for the class II lanthipeptides nukacin ISK-1 (Nagao et al., 2005) and lichenicidin (Caetano et al., 2011b), and many other lanthipeptides (Shi et al., 2011), active expression in *E. coli* does not require all the genes present in the biosynthesis cluster. In the present study, we utilized dual-vector and single-vector approaches to engineer the pathway for the biosynthesis and modifications of ruminococcin-A in *E. coli*. The genes encoding the lanthionine synthetase (RumM) and the structural peptide (preRumA) were amplified from the genome of *R. gnavus* E1 and coexpressed simultaneously in *E. coli*. We show that biosynthesis and modification of the peptide can be successfully achieved by expressing the structural peptide as a fusion partner to GFP. Further investigations in conjunction with the formation of complexes between the modifying enzyme and the chimeric fusion product supplied some interesting insights into the catalytic mechanisms of class II lanthionine-generating enzymes. We also identified a simple strategy to isolate the product and the yield overwhelmingly exceeded that attainable in the natural producer by several thousand-fold. The structure of modified preRumA was fully characterized, showing interesting new features. The systems developed here provide useful tools for optimizing production of the peptide, as well as performing *in vivo* peptide engineering to improve its physicochemical and pharmacological features.

## 2. Literature Review

### 2.1 The Bacteriocins

Bacteriocins are ribosomally synthesized proteinaceous toxins secreted by bacteria and archaea to prevent the growth of other microorganisms competing with them in their ecological niches, and they have potential clinical, aquaculture, livestock and food applications (Bali et al., 2016). They are produced predominantly by Gram-positive bacteria including gut bacteria (Hegarty et al., 2016) and especially lactic acid bacteria (LAB) (Cascales et al., 2007, Alvarez-Sieiro et al., 2016). Several conventional methods like the ammonium sulfate precipitation, pH-mediated cell adsorption/desorption, membrane filtration and a series of different chromatographic techniques are currently applied to purify bacteriocins from complex culture broth (Jamaluddin et al., 2017). However, applying these techniques on the natural host to supply a sustainable production process is not trivial since most natural producer organisms secrete the compounds in very minute quantities believed to play a role in signaling and repelling other microbes in their ecological environment (Chikindas et al., 2018).

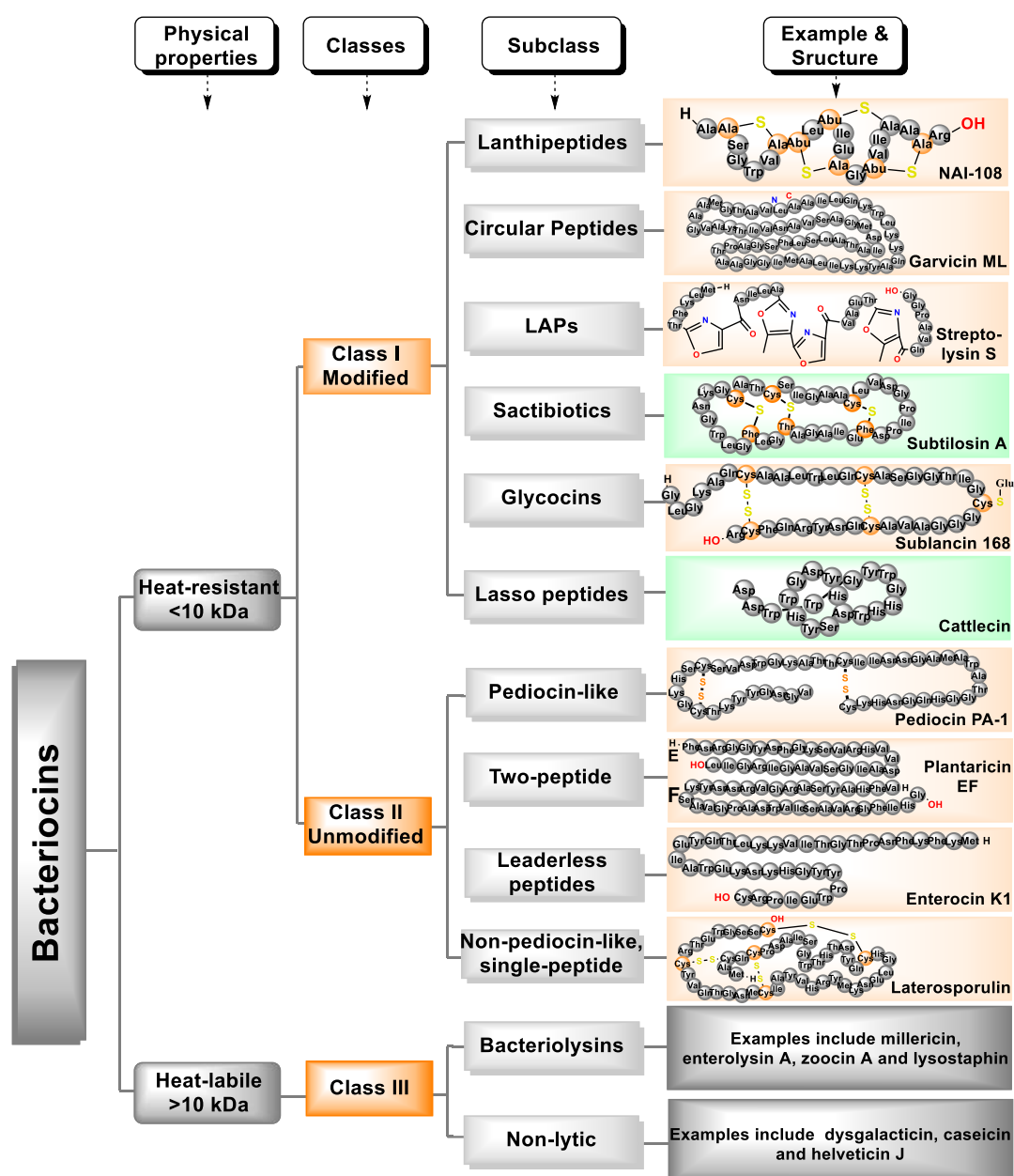
Although data on expression levels in their ecological niches are rare in the literature, studies reveal that when bacteriocins are applied at concentrations higher than those in the natural habitat of the host microorganisms, several different functions are observed amongst which the microbial inhibitory activities of the compound have been well characterized (Dridet et al., 2016). Bacteriocins usually inactivate their targets via membrane disruption or by forming pores (Etayash et al., 2016), or even terminate cell division by targeting and dissociating lipid II (Hasper et al., 2006) which serves as specific anchor molecule for a host of class I bacteriocins including nisin, epidermin and gallidermin. Bacteriocins have high potencies against multi-drug-resistant pathogens, possess relatively negligible toxicity to host cells and demonstrate significant stabilities under physicochemical conditions (de Oliveira Junior et al., 2015). These characteristics position them on high demand as alternative anti-infective agents in food processing and as therapeutic drugs, which may be applied as narrow or broad spectra antimicrobials. Furthermore, they are secreted naturally by gut microbes and LAB frequently found in many commercially useful products, and generally regarded as safe (GRAS) for human consumption (Nes et al., 2007). Additionally, the functions of bacteriocins exceed beyond the margin of antimicrobial agents. For example, subtilisin A produced by *Bacillus subtilis* show lethal activities against viruses (Quintana et al., 2014) and sperms (Sutyak et al., 2008), and others may be used in the treatment of cancers (Kaur and Kaur, 2015).



## 2.2 Classification of bacteriocins

The growing number of bacteriocins isolated from different bacteria sources makes it a challenging task to perfectly divide them into separate classes which explains why a plethora of controversies exist about their classification. For example, carnobacteriocin X was recently studied and from its structural information, it was reclassified from class IId to class IIb (Tulini et al., 2014, Acedo et al., 2017). Nevertheless, genome mining tools with a database like Bagel3 (van Heel et al., 2013) and BACTIBASE (Hammami et al., 2010); and genome mining tools like antiSMASH 2.0 (Blin et al., 2013) have been created to enable information collection and to automatically screen for gene clusters that encode ribosomally synthesized and post-translationally modified peptides (RiPPs). Other genome mining software that are commonly used to identify unique biosynthetic machineries include BLAST, ClusterFinder and RiPPPRISM (Hetrick and van der Donk, 2017). An empirical comparison of these tools shows that Bagel3 is handier and provides more insightful information (Gabere and Noble, 2017). For example, an analysis of complete genome data extracted from public databases of 238 LAB resulted in 785 gene clusters putatively belonging to bacteriocins (Alvarez-Sieiro et al., 2016). Analyses of this nature have greatly facilitated the identification of valuable peptides with enhanced functions and features that may ease their groupings into different classes.

Klaenhammer proposed the first classification of bacteriocins which grouped them into four distinct classes based on biochemical and genetic information existing at the time that made it possible to predict structural features and mechanisms of action (Klaenhammer, 1993). Many other classifications ensued taking into considerations criteria such as molecular sizes, physical properties and chemical structures (Nes et al., 1996, Kemperman et al., 2003, Cotter et al., 2005). Cotter and colleagues revised previous classifications (that essentially distinguished bacteriocins into five classes) and postulated a classification scheme that divided them into two main categories namely; class I (constituting the lanthionine-containing lantibiotics) and class II (the non-lanthionine-containing bacteriocins), while the large murein hydrolases were separately grouped under bacteriolysins (Cotter et al., 2005). In an attempt to solve the problem of contradictory classification that allows one bacteriocin to belong to multiple classes, Zouhir and colleagues utilized a host of bioinformatics tools to develop a sequence-based fingerprints that enabled more than 70% known bacteriocins to be divided into 12 groups (Zouhir et al., 2010). However, bacteriocin sequences with large evolutionary distances would be difficult to fit in this classification because the sequence divergence would be huge. This therefore poses a significant limitation to the sequence homology-based classification and as such, it may be difficult to directly assign uncharacterized bacteriocins under a particular group.



**Figure 1.1** Classification of bacteriocins. Examples depicted in light-green are from non-LAB sources. Non-lytic class III members are yet to be structurally characterized. LAPs stands for Linear azol(in)e-containing peptides

Due to increasing complexity of newly identified compounds, many other classifications have subsequently been proposed based on biosynthesis route and activities of the peptide (Arnison et al., 2013, Cotter et al., 2013). In 2013 Cotter and coworkers suggested a classification scheme that depended on whether the peptide is modified (class I) or unmodified (class II), and emphasized the necessity to eliminate ribosomally synthesized antimicrobial proteins from the list of bacteriocins. This proposal was slightly modified to include the bacteriolysin by

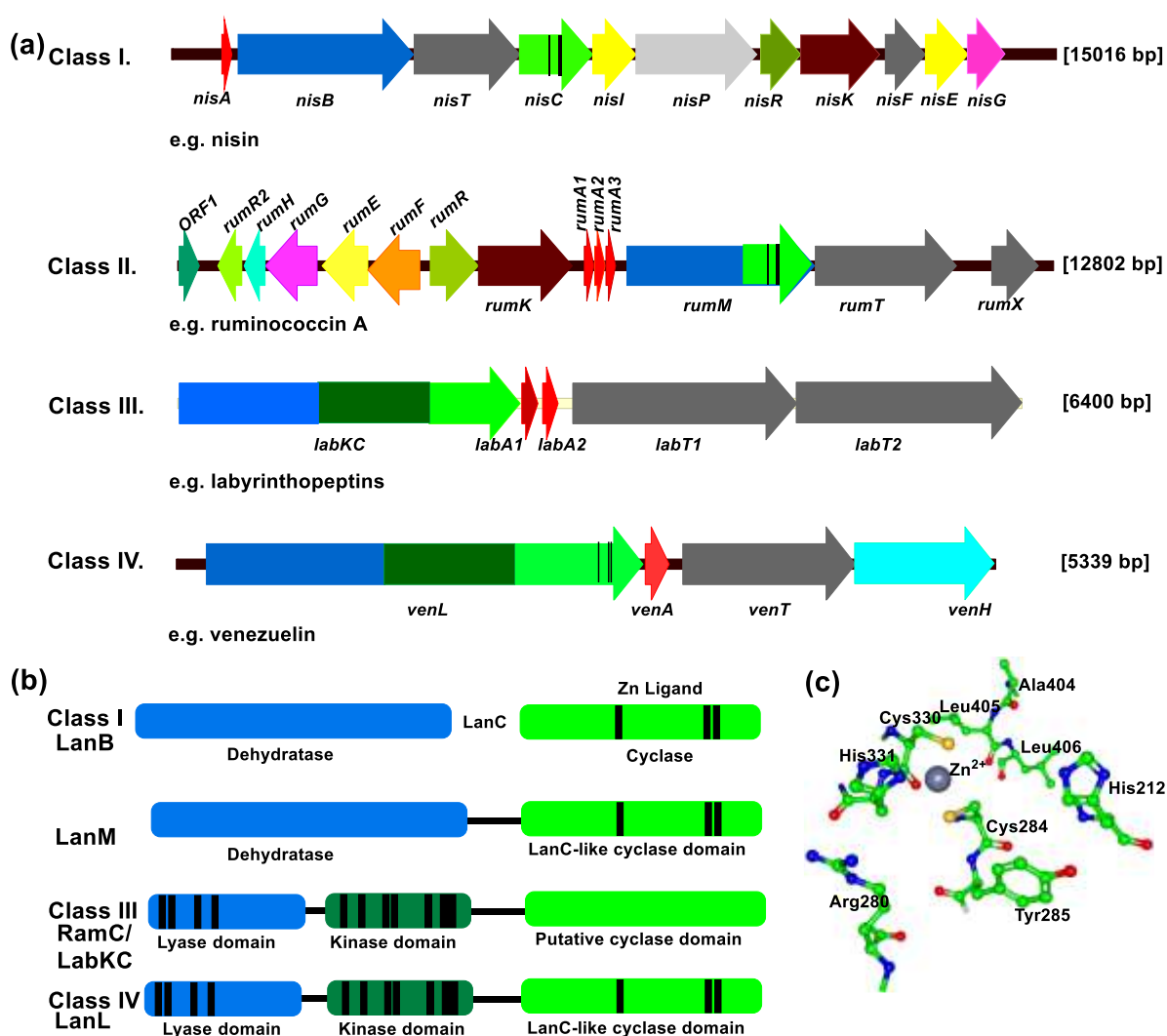
designating a third class (class III) to accommodate unmodified bacteriocins (molecular weights >10 kDa) that elicit a bacteriolytic or non-lytic action (Alvarez-Sieiro et al., 2016).

The classification scheme displayed in Figure 1.1 was suggested as an effective classification model for LAB bacteriocins, but it may also be expanded to include bacteriocins from other organisms as indicated by the examples of sacitibiotics and lasso peptides which are produced by the non-LAB *Bacillus subtilis* 168 and *Streptomyces cattleya*, respectively. It was also suggested that a more general definition for bacteriocins may be considered as follows; “proteinaceous substances produced at ribosomal level, having multifunctional properties whose antimicrobial activities are concentration-dependent” (Chikindas et al., 2018). This largely eliminate the general consideration that they are AMPs, amphipathic in nature, possessing an overall positive charge and since anionic bacteriocins such as subtilisin A do exist (Mathur et al., 2015). Structural description of the biosynthesis operons of the various classes of bacteriocins, the physicochemical characteristics of randomly selected examples, and brief summaries of the main features (including biosynthesis and killing mechanisms) are described in Appendix 6.1.

## 2.3 The classification of lanthipeptides

Lanthipeptides are RiPPs comprising of non-proteinogenic amino acids and thioether cross-linkages formed between dehydrated side chains of threonine/serine and the sulfhydryl group of cysteine. The maturation of lanthipeptides involve a distinct group of enzymes that catalyze three unique reactions namely; dehydration, cyclisation and proteolysis. Other customized enzymes may catalyze the introduction of additional PTMs which are necessary to ensure their activities and/or stability (Repka et al., 2017). The dehydration reaction usually occurs via glutamyl-tRNA-dependent glutamylation, followed by elimination (Garg et al., 2013, Ortega et al., 2015, Repka et al., 2017). Michael addition cyclization reactions involving cysteine and dehydroamino acids may introduce a lanthionine (Lan) ring in the case of dehydrated serine or dehydroalanine (Dha), and methyllanthionine (MeLan) ring when dehydrated threonine or dehydrobutyrine (Dhb) is involved (Schnell et al., 1988, Arnison et al., 2013, Yang and van der Donk, 2015, Repka et al., 2017). Their polycyclic structure is made up of the MeLan/Lan rings from which they derive their name. Additional features such as Dhb, Dha, N-terminal pyruvate, S-aminovinyl-cysteine and D-amino acids may also be present in the molecule (Freund et al., 1991, Knerr and van der Donk, 2012, Lohans et al., 2014, Ortega et al., 2014). For decades, they were globally referred to as lantibiotics because most members possessed antibacterial activities (Schnell et al., 1988), but this has changed over the years as many new members

show diverse functions by producing different categories of biological effects (Knerr and van der Donk, 2012).



**Figure 2.1** Summary of the major features in the lanthipeptides biosynthesis pathway. (a) Gene clusters (drawn to scale) representing examples of each class, showing highly conserved motifs and/or residues (indicated by colour and vertical lines on designated genes) in the lanthionine synthetase encoding genes. The precursor peptides encoding genes is *lanA* is cloured red. (b) Putative and characterized domains of lanthionine-generating enzymes: Class I (*LanB*, *LanC*), II (*LanM*), III (*LanKC*) and IV (*LanL*). (c) Catalytic core structure of NisC edited using X-ray crystallographic data for NisC (PDB: 2G02) and Protein Workshop. The  $Zn^{2+}$  cofactor is shown to coordinates residues Cys284, Cys330 and His331 in the active site. The residues His212 and Tyr285 which are much closer to the  $Zn^{2+}$  catalytic center are also shown.

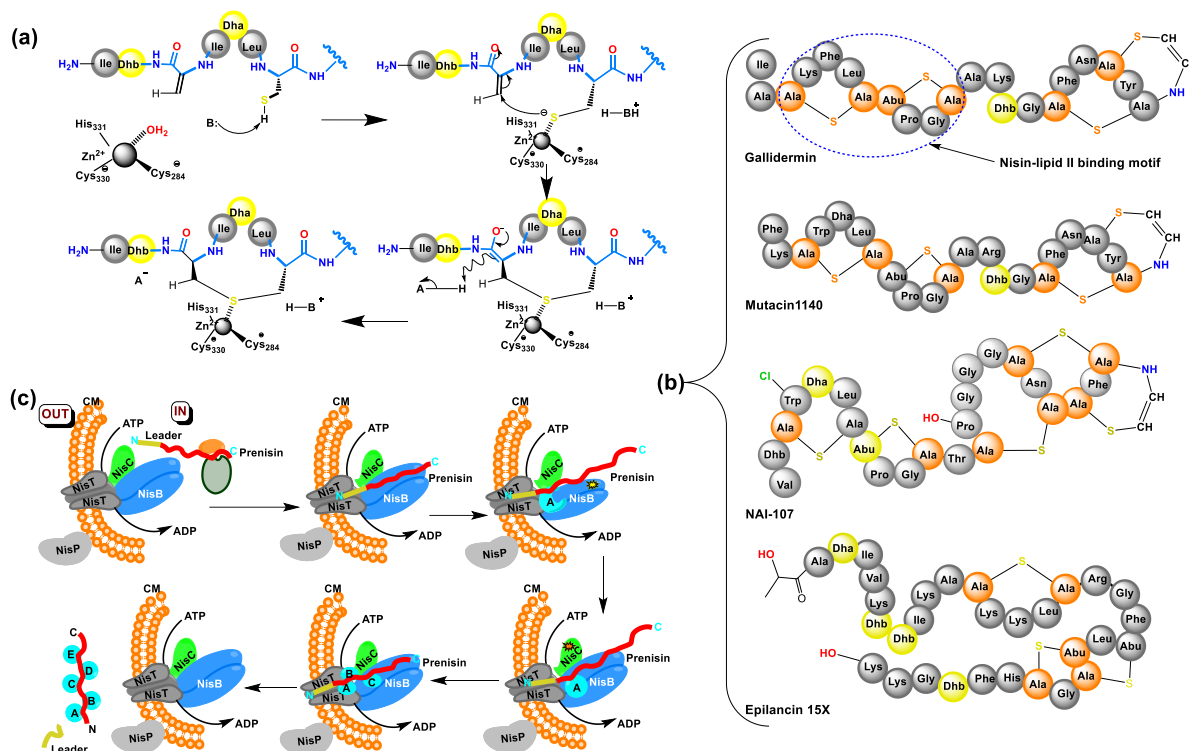
The posttranslational modifications (PTMs) in lantipeptides make them more resistant to protease degradation (Rink et al., 2010) and accord them limited conformational freedom which confers target specificity (Goto et al., 2011). On the basis of their structure, they display

rigid conformational flexibilities that render high biological activities with very low minimal inhibitory concentrations (MICs) against wide variety of infections ranging from antimicrobial to antiallodynic effects (Repka et al., 2017). For these reasons they are seriously considered as future alternative clinical agents against bacterial virulence (Dischinger et al., 2014).

The biosynthesis machinery of lanthipeptides is regulated by a large cluster of genes present in the *lan* operon ("*lan*" is a general notation given to genes that are hosted in the lanthipeptide biosynthesis cluster and their corresponding products are referred to as "Lan"). Note that "Lan" may be used contextually to refer to the lanthionine ring (as mentioned above) or to name proteins expressed by the *lan* operon (e.g. LanM is a lanthionine synthetase encoded by *lanM*). Based on the structural arrangement of the maturation enzymes on the *lan* operon, this class of bacteriocins may be further grouped into four distinct classes namely; class I where a dehydratase (LanB) and a cyclase (LanC) are involved in their maturation; class II modified by a single bifunctional enzyme LanM (which performs both the dehydratase and the cyclase roles); class III processed by the LanKC enzyme; and class IV processed by LanL (Knerr and van der Donk, 2012, Repka et al., 2017). The arrangement of genes encoding the lanthionine-generating enzymes and other accessory proteins in the biosynthetic operon is unique for different groups of these bioactive molecules (Figure 2.1a). The various domains of the modifying enzymes that distinguish the four classes of lanthipeptides are displayed in Figure 2.1b.

### 2.3.1 Class I lanthipeptides

The prototype member of lanthipeptides nisin belongs to class I where two enzymes (LanB dehydratase and LanC cyclase) are involved in their Lan/MeLan rings formation. LanB proteins usually have sizes around 1,000 amino acid residues and although they catalyze the same reaction, they show only limited ( $\approx 30\%$ ) sequence identity across the family and share negligible homology with other known proteins (Knerr and van der Donk, 2012). However, they particularly share highly conserved structures in most of their biosynthesis genes (Götz et al., 2014). Previous reports indicate that reactions catalyzed by LanB enzymes may involve phosphorylation of serine/threonine residues (Chatterjee et al., 2005, Repka et al., 2017). Previous attempts to reconstitute LanB enzyme *in vitro* failed due to lack of evidence that supported dehydration of the peptide substrate (Xie et al., 2002, Mavaro et al., 2011). Nevertheless, very recently Garg and coworkers reported that some bacterial cellular components might be involved in the dehydration process since incubating the peptide substrate with, ATP,  $Mg^{2+}$  and *E. coli* cell extract supplemented with glutamate and spermidine resulted in the desired bioactivity of nisin (Garg et al., 2013).



**Figure 2.2** Maturation of nisin and other class I lanthipeptides. (a) Suggested mechanism for NisC-catalyzed reactions, illustrating the formation of the first N-terminal thioether cross-bridge of nisin. The scheme begins with a water molecule being displaced from Zn<sup>2+</sup> ion by the thiol group of Cys7 residue targeted for conjugation with Dha3. Further deprotonation of this group is believed to be achieved by an active site base or water (Li et al., 2006). The β-carbon of Dha then launches an attack on the nucleophilic thiolate to produce an enolate intermediate which undergoes further protonation to generate a D-configuration at the α-carbon. (b) Representative examples of class I lanthipeptide structures, highlighting the canonical nisin-lipid II-binding motif and additional PTMs. (c) Proposed model for nisin biosynthesis and secretion indicating how the ribosomally synthesized precursor peptide is channeled to the lanthionine synthetase NisB/NisT/NisC complex where the leader sequence is recognized and bound. Nisin precursor peptide alternate (indicated by a mark) between the active site of NisB and NisC until all the five rings are formed. Fully modified nisin is then exported via NisT and subsequent activation by the dedicated membrane-anchored NisP.

Nisin cyclase NisC of the LanC protein family shares about 20–30 % overall sequence identity across the family (Knerr and van der Donk, 2012). The *in vitro* reconstitution of NisC and its X-ray crystal structure have been reported (Li et al., 2006). These achievements uncovered interesting insights on lanthionine cyclase catalytic mechanism. Figure 2.1c clearly shows how the catalytic core of NisC coordinates Zn<sup>2+</sup> ion via a Cys284-Cys330-His331 catalytic triad (Li et al., 2006). The proposed catalytic mechanism for NisC (Figure 2.2a) illustrates that the triad activates the thiol group of cysteine to enable intramolecular nucleophilic attack by Dha or Dhb resulting from the dehydratase reactions (Li et al., 2006). The C-terminal residues Ala404,

Leu405, and Leu406 are most likely involved in stabilizing the catalytic core (Helfrich et al., 2007). Recent studies on NisC and the class II lanthipeptide synthetase HalM2 showed that the Michael-type addition reaction that produces the Lan/MeLan rings is reversible and that these enzymes can actually open up all thioether rings in their products (Yang and van der Donk, 2015). LanC cyclases are usually found in Bacteroidetes, Proteobacteria, Actinobacteria and Firmicutes; meanwhile sequences of LanB dehydratases from Bacteroidetes and Proteobacteria show ancestral links to Firmicutes (Zhang et al., 2012). There is no direct correlation in the phylogenetic distribution of LanBs and LanCs, although enzymes present in the same biosynthesis operon generally have similar evolutionary origin (Repka et al., 2017). Examples of class I lanthipeptides other than nisin are shown in Figure 2.2b.

Other proteins in the class I cluster like LanT serves to export the modified peptide while the membrane-anchored leader peptide cleavage protein LanP removes the leader peptide and consequently activates the peptide (Chatterjee et al., 2005, Xu et al., 2014, Repka et al., 2017). In the case of nisin, the functions of NisT and NisP do not depend on one another as in other classes (Kuipers et al., 2004). There exists evidence of steric hindrance in the active site of NisP (Repka et al., 2017) indicating possible reasons why the protease prefers modified NisA (Kuipers et al., 2004, Lagedroste et al., 2017). Nevertheless, a soluble variant of NisP displayed substrate promiscuity and additional findings indicated that the presence of Lan/MeLan ring in the substrate as well as some specific residues nearer to the cleavage site are necessary for specific cleavage to occur (Montalbán-López et al., 2018).

NisA, NisB, NisC and NisT do interact to form a complex during biosynthesis of nisin as illustrated in Figure 2.2c, although their individual activities are conserved even in the absence of each other (Bakkes et al., 2008, Lubelski et al., 2008). Additionally, dysfunctional LanC may also greatly affects active LanA synthesis as observed for nisin where the absence of NisC resulted in increased levels of dehydrated NisA (Lubelski et al., 2009). Little is known about the order in which the reactions proceed and how much input is achieved by each enzyme.

Although the dehydratase reactions are independent of each other, mutagenesis studies on NisA have revealed that conjugation of the cysteinyl thiol to Dha/Dhb in the cyclization step can prevent further dehydration of serine and threonine residues (Kuipers et al., 2008, Lubelski et al., 2009). This may explain why partially modified products have been observed in some recombinant lanthipeptides expression systems. Biophysical characterization of purified NisA and NisB indicated that the leader peptide creates a docking motif that allows NisB to bind and interact with NisA (Mavaro et al., 2011). Furthermore, the interaction between dehydrated NisA and NisB is more stronger than either the unmodified or mature peptide (Mavaro et al., 2011). These results corroborated findings of Khusainov and coworkers where they showed that NisB

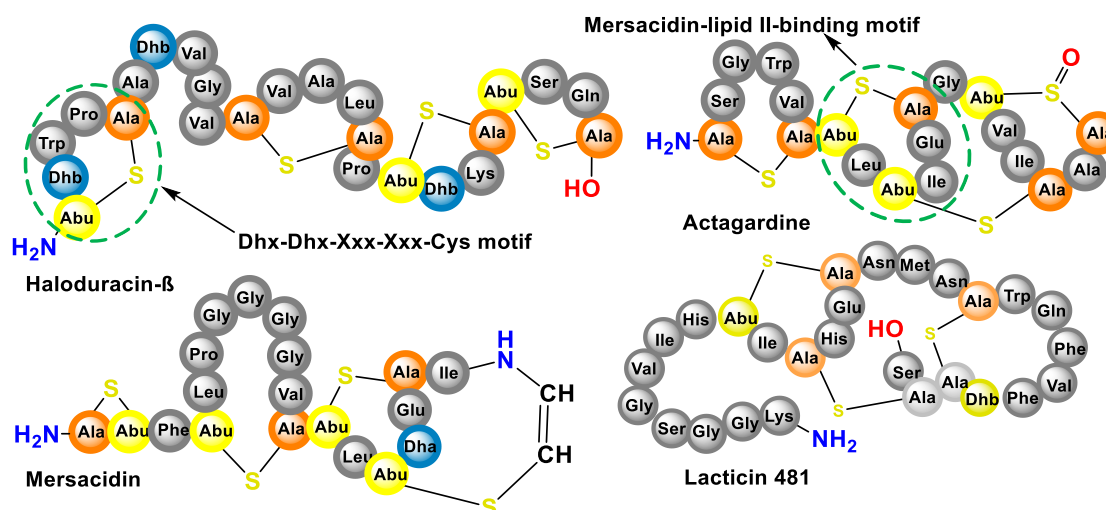
and NisC were co-purified from *L. lactis* together with C-terminal hexahistidine-tagged NisA (Khusainov et al., 2011). Since deleting the core peptide resulted in no co-purification of NisB and NisC, this implies that cooperative interactions involving both the core peptide and the leader sequence exist between the modifying enzymes and NisA. Thus, a model describing a unidirectional N→C terminal modification of NisA has been proposed (Lubelski et al., 2009), wherein the cooperative action of NisB, NisC and NisT proteins are highlighted (Figure 2.2c).

### 2.3.2 Class II lanthipeptides

Lacticin 481 and nukacin ISK-1 are typical examples of class II lanthipeptides whose PTMs are installed by LanM protein which is a bifunctional lanthionine synthetase. The size range of LanM is usually between 900 and 1,200 amino acid residues, possessing an N-terminal dehydratase domain and a C-terminal cyclase domain (Knerr and van der Donk, 2012). Although the dehydratase domain catalyzes the same reaction as does LanB enzymes, both proteins share no sequence homology. However, their C-terminal regions contains the conserved zinc-binding residues and share >25 % homology to LanC proteins but negligible similarities with each other (Knerr and van der Donk, 2012). Removal of the conserved zinc-binding residues proportionately eliminated the cross-linking ability of lacticin 481 synthetase (LctM) (Paul et al., 2007).

*In vitro* and/or *in vivo* characterization of LctM (Xie et al., 2004), bovicin HJ50 synthetase BovM (Ma et al., 2014) and nukacin ISK-1 synthetase NukM (Shimafuji et al., 2015) revealed exciting insights about how LanM enzymes generate the Lan/MeLan rings. Whereas they essentially lack the characteristic ATP-binding motifs, their catalysis requires the hydrolysis of ATP to supply phosphate for phosphorylation (Bierbaum and Sahl, 2009). They also use ADP and Mg<sup>2+</sup> as cofactor in the elimination step involving the phosphate esters of serine/threonine residues in the core peptide to yield Dha/Dhb (Nagao et al., 2006). Furthermore, residues that are critical for catalyzing phosphorylation are more inclined to those of typical serine/threonine kinases (You and Van Der Donk, 2007). Tandem mass spectrometry with a mutant whose elimination properties is removed revealed that phosphorylation occurs in a unidirectional N- to C-terminal catalytic mode (Lee et al., 2009). The order of modification is greatly influenced by the leader peptide which additionally improves catalytic efficiency and product quality (Thibodeaux et al., 2016).





**Figure 2.3** Representative examples of class II lanthipeptide structures, highlighting the canonical mersacidin-lipid II-binding motifs, additional PTMs and the Dhx-Dhx-Xxx-Xxx-Cys motif (where Dhx = Dha or Dhb).

The N-terminal A-ring of class II lanthipeptides like geobacillin II, cytolyisin, haloduracin, and carnolysin share a Dhx-Dhx-Xxx-Xxx-Cys motif as illustrated for haloduracin  $\beta$  in Figure 2.3 (Garg et al., 2012). Geobacillin II was isolated from the thermophilic gram-positive bacterium *Geobacillus thermodenitrificans* NG80-2 (Garg et al., 2012). Geobacillin synthetase GeoM is thermostable and able to catalyze PTMs formation in geobacillin II precursor peptide GeoAll at temperatures ranging between 37 and 80 °C (Garg et al., 2016).

The LanM enzymes are also known to exhibit wide catalytic promiscuity as observed in the case of ProcM and its ProcA peptide substrates (Li et al., 2010, Tang and van der Donk, 2012, Zhang et al., 2014). However, this may also constitute a cost on the catalytic efficiencies of the enzymes as a result of the burden exerted by the heterogeneous Lan/MeLan ring architectures required in the different substrates as report indicates (Thibodeaux et al., 2014). Although the mechanisms involved in this promiscuous catalysis remain to be established, it is suggested that the peptide sequence, and not the modifying enzyme, determines both the directionality of modification of nascent peptide and the final conformational architecture of the product (Yang and van der Donk, 2013). This may explain why it is possible for the enzyme to catalyze the modifications of several unrelated peptides. Nevertheless, recent findings indicate that the peptide sequence alone does not determine the stereochemical outcomes of Lan/MeLan rings since it was revealed that the N-terminal A-ring in geobacillin II has LL-configuration while the rest of the rings possess the normal DL-configurations (Garg et al., 2016). This was the very first report of a single-component lanthipeptide possessing the LL-stereochemistry.

Class II lanthipeptides are exported from the cell by integral membrane proteins that belong to the ATP-binding cassette (ABC)-transporter maturation and secretion (AMS) family of proteins generally referred to as class II LanTs. This class of LanT enzymes perform a dual function in the biosynthesis of class II lanthipeptides; (i) they cleave off the N-terminal leader sequences of the lanthipeptide precursor and (ii) transport the active peptide to the extracellular space where they are required. The N-terminal domain of class II LanT proteins possess residues that are typical for papain-like cysteine proteases (Nishie et al., 2011). This domain acts as a peptidase to cleave off the leader peptide at a conserved Gly-Gly motif located at the C-terminus of the leader sequence preceding the core structure (Chatterjee et al., 2005).

The protease domains of other bifunctional transporters for non-lanthipeptide bacteriocins have been characterized (Havarstein et al., 1995, Wu and Tai, 2004, Ishii et al., 2006). N-terminal 1-150 amino acid residues of lactacin 481 transporter LctT was the first class II protease to be characterized. It was shown that this domain indiscriminately cleaves modified and unmodified LctA (Furgerson Ihnken et al., 2008), indicating that the ABC transporter domain does not affect the protease activity. However, nukacin ISK-1 transporter NukT requires the cooperative action of both peptidase and ABC transporter domains (Nishie et al., 2011), and its protease domain is located in the cytoplasmic side of the lipid bilayer membrane (Nishie et al., 2009). Contrary to the protease domain of NukT, other lanthipeptides like the class I prototype nisin is processed post-excretory by a membrane anchored protease NisP (Lubelski et al., 2008). Interestingly, the gene cluster of the class II two-component lantibiotic lichenicidin produced by *Bacillus licheniformis* constitutes an additional protease LicP that further activates the  $\beta$ -component Bli $\beta$ ', produced by lichenicidin transporter LicT in a processing event that initially cleaves off the leader peptide of Bli $\beta$  at the Gly-Gly site (Caetano et al., 2011).

The non-lanthipeptide bacteriocin-associated ABC transporter of *Streptococcus* ComA peptidase domain induce the formation of an amphiphilic  $\alpha$ -helical structure in its substrate ComC upon binding to it. Moreover, mutagenesis studies revealed conserved hydrophobic residues in the leader peptide of ComC that facilitate this interaction (Kotake et al., 2008). Existing theories suggest that substrate recognition by LanT proteins may also occur via the amphipathic  $\alpha$ -helical conformation in the precursor peptide (Furgerson Ihnken et al., 2008, Nagao et al., 2009).

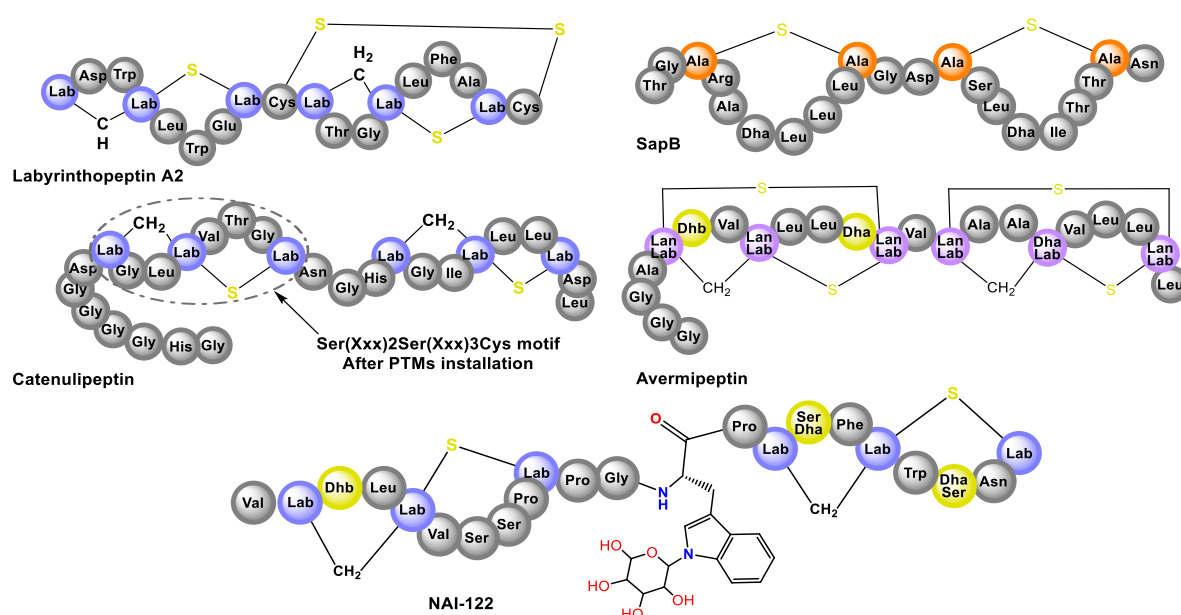
### 2.3.3 Class III lantipeptides

Class III lantipeptides are synthesized by the RamC/LabKC family of lanthionine/labionin synthetases consisting of three catalytic domains—an N-terminal phosphoserine/phosphothreonine lyase domain which has sequence homology with members of the outer surface protein F (OspF) family of effector proteins [OspF catalyzes the removal of the phosphate group from phosphothreonine (Li et al., 2007)], a central Ser/Thr kinase domain, and a C-terminal lanthionine cyclase domain (Goto et al., 2011). The first member of this class, SapB, was isolated slightly over a decade ago from *Streptomyces coelicolor* and shown to possess structural characteristics of regular lantibiotics. It exhibited no antimicrobial activities, but instead enhanced sporulation in streptomycete (Kodani et al., 2004).

The C-terminal domain of the SpaB synthetase (RamC) believed to be involved in the synthesis of the Lan/MeLan rings showed very little sequence homology to the cyclase domain of LanM and lacks the conserved residues that coordinate the  $Zn^{2+}$  ion in the catalytic core of a lanthionine cyclase (Willey and van der Donk, 2007). Mutagenetic investigations of the N-terminal lyase domain displayed structure-functional similarities to the OspF proteins (Goto et al., 2010). The catalytic pathway that employs the kinase-like domain to produce Dha/Dhb follows a two-step process involving phosphorylation and  $\beta$ -elimination (Goto et al., 2011). Other lantipeptides with similar characteristics as SapB are SapT isolated from *S. tendae* (Willey and Gaskell, 2010), labyrinthopeptins from *Actinomadura namibiensis* (Meindl et al., 2010), catenulipeptin from *Catenulispora acidiphila* (Wang and Van Der Donk, 2012), erythreapeptin, avermipeptin and griseopeptin isolated from *Saccharopolyspora erythraea*, *Streptomyces avermitilis* and *Streptomyces griseus* respectively (Völler et al., 2012). Structures of selected examples are shown in Figure 2.4.

The discovery of labyrinthopeptins enabled mechanistic studies on the synthetic pathway of class III lantipeptides (Meindl et al., 2010). The labyrinthopeptins synthetase (LabKC) responsible for processing three similar precursor peptides demonstrated significant sequence homology to RamC (Knerr and van der Donk, 2012). The biosynthetic gene cluster for labyrinthopeptins does not encode a thiol-disulfide oxidoreductase to account for the disulfide bridge in the molecule (Figure 2.4) and no dedicated protease to remove the leader sequence. The lack of conserved  $Zn^{2+}$ -binding residues in LabKC and the presence of two LanT-like transporters also raise series of unanswered questions. *In vitro* reconstitution studies on LabKC indicated GTP is involved in the kinase reaction and not ATP (Müller et al., 2010). However, a different scenario is observed for AciKC which is able to catalyze 5-fold dehydration of AciA in the presence of all nucleoside triphosphates (Wang and Van Der Donk, 2012).

The directionality of LabKC catalysis was established to follow a C→N-terminal mode when processing labyrinthopeptin (Krawczyk et al., 2012). Subsequent investigation with curvopeptin using deuterium-labelled CurA also indicated a C→N-terminal modification order. The data obtained from this study enabled the construction of a comprehensive biosynthesis model that included phosphorylation, elimination, and cyclization (Jungmann et al., 2014). The latter findings supply an insightful understanding of the post-translational modification machinery involved in the biosynthesis of RiPPs.



**Figure 2.4** Representative examples of class III lanthipeptide structures, showing the characteristic carbocyclic rings, the  $\text{Ser(Xxx)}_2\text{Ser(Xxx)}_3\text{Cys}$  motif and additional PTMs like the N-glycosylation of NAI-112

The two products of LabKC labyrinthopeptin A1 and labyrinthopeptin A2 have been studied in details (Meindl et al., 2010, Müller et al., 2010, Sambeth and Süßmuth, 2011, Féir et al., 2013, Krawczyk et al., 2013). The precursor peptides of most class III lanthipeptides share a  $\text{Ser(Xxx)}_2\text{Ser(Xxx)}_3\text{Cys}$  motif as shown in Figure 2.4 (Meindl et al., 2010, Wang and Van Der Donk, 2012). The X-ray crystallographic structure of labyrinthopeptin A2 revealed the presence of a disulfide bond, as well as an unprecedented carbocyclic labionin residue (Meindl et al., 2010), that have also been identified in other new members of the class. Other structural motifs have been observed in the course of characterizing other members such as erythraepectin, flavipeptin, and curvopeptin (Krawczyk et al., 2012, Völler et al., 2012, Völler et al., 2013), indicating structural heterogeneity within class III peptides. Furthermore, the lanthipeptide NAI-

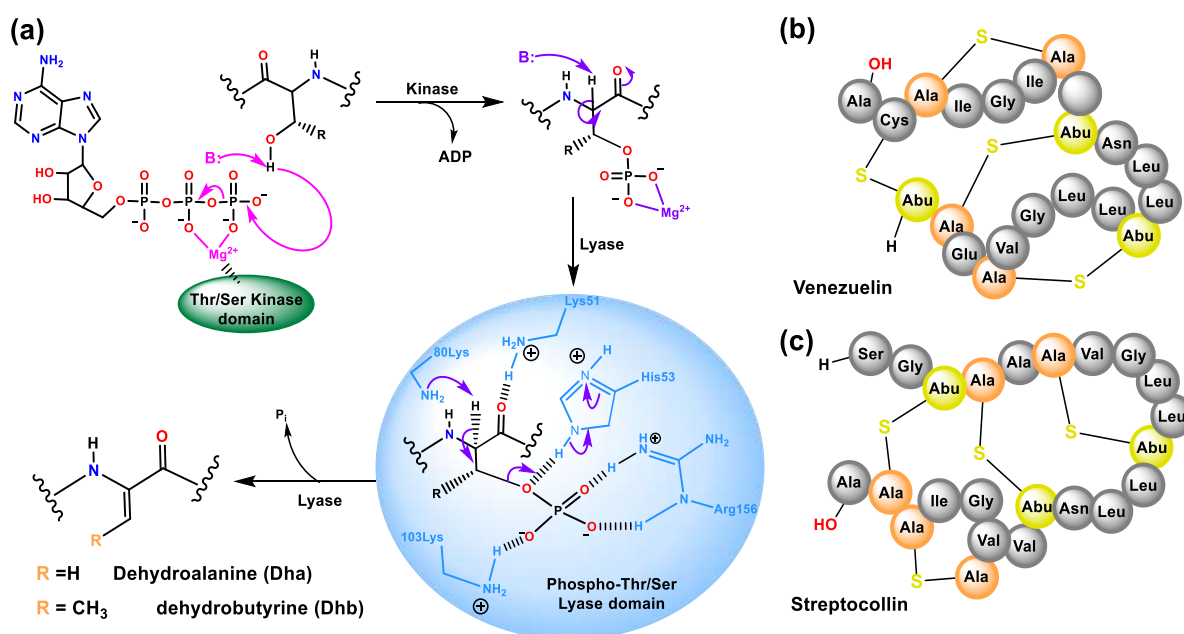
112 from *Actinoplanes* sp. DSM24059 has an N-glycosylated moiety (Iorio et al., 2014), which presents diverse pharmaceutical possibilities for this molecule.

No dedicated protease has been identified in the gene cluster for class III peptides. Preliminary studies on erythraepectin, avermipeptin and griseopeptin suggested that a putative aminopeptidase-like protease(s) may be involved in a stepwise leader peptide removal, producing a mixture of peptides whose N-terminals are processed differently (Völler et al., 2012). Nevertheless, four gene clusters responsible for class III peptides biosynthesis were found to encode prolyl oligopeptidase and a constituent member of this family of peptidases FlaP was shown to remove the leader sequence of the class III lanthipeptide flavipeptin (Völler et al., 2013). Unlike the redundant substrate specificity demonstrated by LctT (Furgerson Ihnken et al., 2008) and NisP preference for modified NisA (Kuipers et al., 2004, Lagedroste et al., 2017), FlaP is specific to modified flavipeptin and capable of distinguishing between N- and C-terminal rings (Völler et al., 2013), indicating that other factors besides size and amino acid sequence may prompt sequence recognition by this activator protein.

### 2.3.4 Class IV lanthipeptides

Class IV lanthipeptides are modified by the lanthipeptide synthetase LanL. It is a trifunctional enzyme comprising of three distinct catalytic domains namely; lyase, kinase, and cyclase domains that install PTMs in the lanthipeptide venezuelin from *Streptomyces venezuelae*, (Goto et al., 2010). LanL enzymes dehydrate threonine/serine residues by employing the kinase domain to phosphorylate side chain hydroxyl groups which are then subsequently eliminated by the lyase domain in a manner that was proposed to follow the scheme displayed in Figure 2.5a (Knerr and van der Donk, 2012). Unlike the class III LanKC, the cyclase domain of LanL possess the conserved Zn<sup>2+</sup>-binding residues but share similar structural features in the kinase and lyase domains (Goto et al., 2010). They also have one LanT-like transporter that possess no protease domain, and no evidence of a dedicated protease was found in the biosynthesis gene cluster (Knerr and van der Donk, 2012).

The similarities between their synthetases suggest that both class III and IV lanthipeptides share a common ancestral lineage, completely different from those of class I and II where the synthetases lack apparent serine/threonine kinases and phosphoserine/threonine lyases. Phylogenetic comparison shows a broad range of phyla including firmicutes, actinomycetes, Bacteroidetes, Cyanobacteria and Proteobacteria, harboring putative LanB, LanM and LanL (Goto et al., 2010).



**Figure 2.5** Features of class IV lantipeptide (a) Suggested mechanism of serine/threonine dehydration by LanL phosphothreonine/serine kinase and lyase domains. Structures of related class IV lantipeptides venezuelin (b) and streptocollin (c).

So far two members of this class have been produced and characterized; venezuelin (see above) and streptocollin from *Streptomyces collinus* Tü 365 (Iftime et al., 2015). The structures of venezuelin and streptocollin (Figures 2.5b & 2.5c, respectively) were elucidated based on mutagenesis and ESI tandem mass spectrometry (Goto et al., 2010, Iftime et al., 2015). Both have globular structures similar to duramycin and cinnamycin but do not inhibit human phospholipase A2. Rather, streptocollin was shown to inhibit protein tyrosine phosphatase 1B by 33% at a concentration of 50  $\mu$ M (Iftime et al., 2015), demonstrating that the molecule may have therapeutic relevance in related applications.

Four genes are implicated in class IV lanthipeptides biosynthesis, LanL, LanA, LanT, and LanH, which are organized in a single operon. Analyzing sequences adjacent to *lanLATH* operon did not reveal any evidence of genes associated with the biosynthesis and/or regulation of lanthipeptide, nor an immunity factor (Iftime et al., 2015). Hypothetically, the putative LanT-like ABC transporter may form a heterodimeric translocation complex with LanH (Méndez and Salas, 2001) to actively ensure efflux of the peptide and prevent the product from harming the producer host cells (Davidson and Chen, 2004). Like class III, neither a dedicated N-terminal leader peptide processing protein nor an enzyme containing a protease domain has been detected in the gene cluster. Due to the inadequacy of empirically verified sequences, it is challenging to predict the cleavage sites for the leader peptide of class III and IV lanthipeptides

*in silico*, something which can be routinely done for class I and II lanthipeptides (Blin et al., 2014).

## **2.4 Biotechnological production versus chemical synthesis of lanthipeptides**

These topics were extensively reviewed in publication I where the suitability of biological systems and chemical approaches were evaluated, and significant evidences pointed to biotechnological procedures as the only realistic alternative for lanthipeptide production at the moment (Ongey and Neubauer, 2016).

## **2.5 Pharmacological properties and therapeutic use of lanthipeptides**

This subject was summarized in publication II where the therapeutic relevant properties of lanthipeptides under consideration for clinical use were analyzed while trying to gain understanding of how extrinsic modifications affect the overall behaviour of the peptides and how such information may be used to engineer other peptides to improve their clinical values (Ongey et al., 2017).

## **2.6 Recent advances in engineering and heterologous production of lanthipeptides**

To fully exploit and understand how lanthipeptides can be engineered to overcome challenges like instability, insolubility and protease degradation, which remain therapeutically unfavourable under physiological conditions, several tools have been developed (Field et al., 2015) and many others are still in the pipeline. Production levels of lanthipeptides from their natural sources are very low and efforts are also being made to improve this. In addition to what has already been discussed in publication I on these subjects, some of the most recent findings are reviewed herein.

The intrinsic modularity of the Lan/MeLan rings and the promiscuous nature of lanthipeptide synthetases can be exploited to generate a wide range of different peptide architectures (Montalbán-López et al., 2016). These perspectives can actually constitute a new generational approach to identifying and characterizing novel anti-infective agents with enhanced therapeutic potential. For instance the promiscuity of nisin biosynthetic machinery (NisBC) was

exploited to heterologously express and install dehydrations in 27 different peptides, with 5 of them demonstrating good antimicrobial activity, while 2 others were fully modified and possessed enhanced antimicrobial properties (van Heel et al., 2016). This study demonstrated an efficient approach to produce lanthipeptides and to easily identify molecules that may be interesting to investigate further.

Tailoring modifications in RiPPs that usually occur on the side chain of amino acids, and sometimes within the peptide backbone (e.g. epimerization), or at the N- or C-terminus, can increase stability of the peptides. Their activities may also be improved by the diverse chemical functionalities supplied by these modifications, which are necessary for optimal interactions with biological targets (Funk and Van Der Donk, 2017). Setting the platforms that may facilitate these modifications is also important. One of such platforms was reported recently where the precursor peptide of Prochlorosin coupled to the C-terminus of bacteriophage M13 minor coat protein pIII was heterologously expressed and subsequently modified by modifying enzyme ProcM in the cytoplasm of *E. coli*, and then displayed on the phage surface (Urban et al., 2017). This technique may take advantage of the promiscuity of lanthionine synthetases to generate additional PTMs on a variety of peptides such as the N-glycosylation of NAI-112 (Iorio et al., 2014) or the D-amino acids of bicereucin from *Bacillus cereus* SJ1 (Huo and van der Donk, 2016). For example, by co-expressing the bicereucin dehydratase BsjM and the zinc-dependent dehydrogenase NpnJ<sub>A</sub> from *Nostoc punctiforme* PCC 73102 (Zhang et al., 2014), it was possible to produce an analog of the D-amino acid-containing opioid dermorphin in *E. coli* (Huo and van der Donk, 2016). *In vivo* and *in vitro* introduction of D- amino acid into the peptide are also possible using NpnJ<sub>A</sub> that reduces dehydroalanine to D-Ala in the presence of NADPH (Yang and van der Donk, 2015).

Furthermore, by engineering the genetic code, it was possible to reassign the natural codon that enable the substitution of a tryptophan residue in lichenicidin with an unusual amino acid L-β-(thieno[3,2-b]pyrrolyl)alanine, and heterologously expressed this bioactive congener of the compound in an evolutionarily adapted *E. coli* strain (Kuthning et al., 2016). This study supported the concept of biocontainment by showing that although the non-canonical amino acid was toxic to the heterologous host, evolutionary adaptation may help to alleviate the situation. Moreover, co-expressing a dedicated suppressor aminoacyl-tRNA synthetases and the PTM enzymes for nisin NisBC in *E. coli*, nisin variants with α-chloroacetamide-containing unnatural amino acids were generated (Zambaldo et al., 2017). These peptide engineering approaches allow for possible expansion of the chemical reactivity space of the peptides, thereby broadening their activity spectrum.

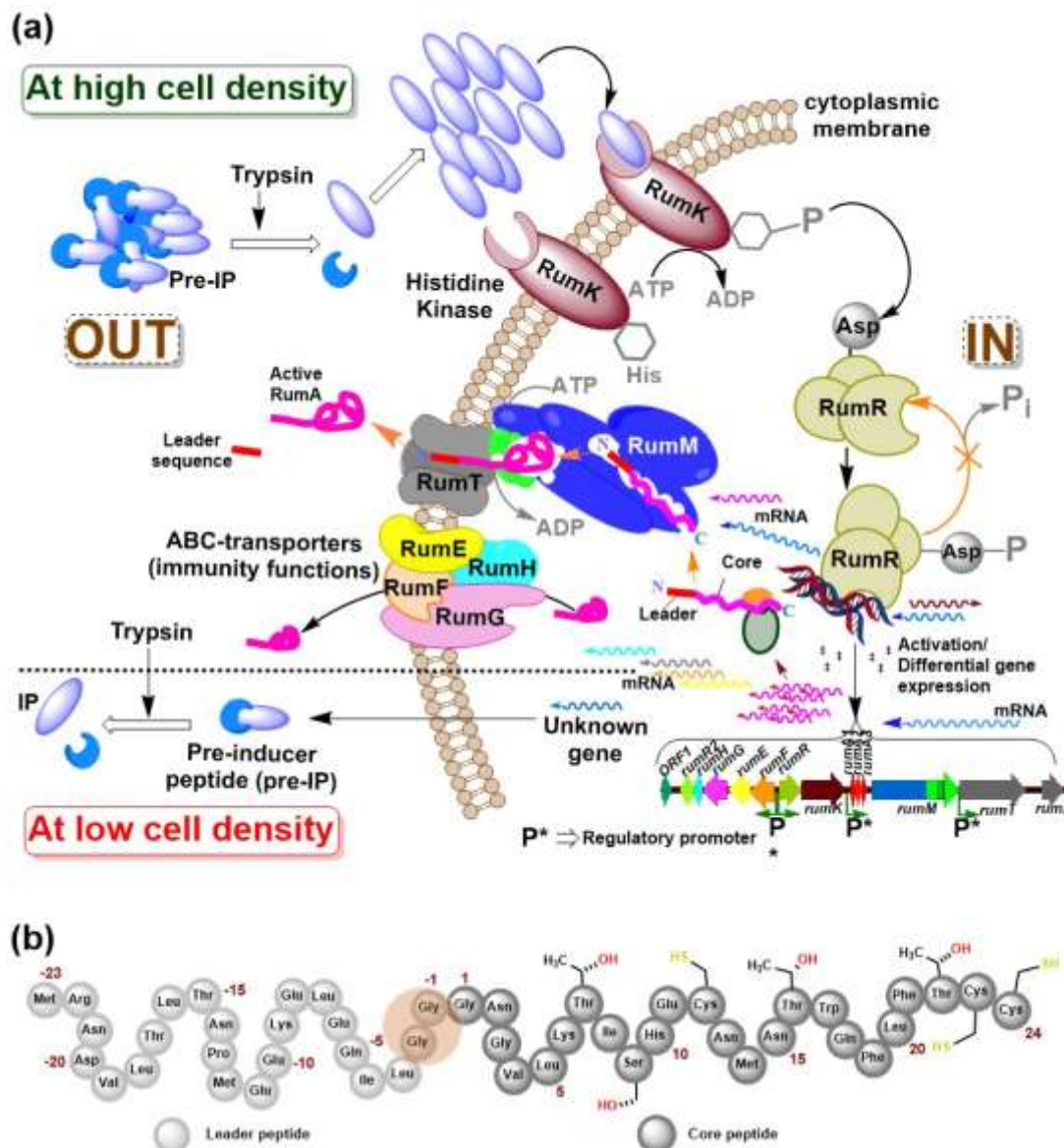


The outer membrane of Gram-negative bacteria prevents lantibiotics from gaining access to the inner phospholipid membrane where lipid II is found and hence their inactivity against this group of bacteria. Efforts to improve the efficiency at which the peptides can overcome this barrier have recorded some progress very recently. For example, part of the anti-Gram-negative peptide apidaecin 1b was fused to the C-terminus of nisin, or truncated versions thereof, producing a two-fold increase in nisin activity against *E. coli* CECT101 (Zhou et al., 2016). Additionally, a nano-engineering approach involving the coupling of nisin molecules to gold surfaces was used to broaden the activity spectrum of nisin to include Gram negative bacteria like *E. coli* and *Pseudomonas aeruginosa* (Vukomanović et al., 2017).

The biosynthetic gene cluster of mathermycin from *Marinactinospora thermotolerans* SCSIO 00652 was isolated and reconstituted in *Streptomyces lividans* and *E. coli* to produce the cinnamycin-like lantibiotic that demonstrated antimicrobial activity against a *Bacillus* strain (Chen et al., 2017). Although purification from *S. lividans* was not successful due to low expression, several attempts to produce in the native strain proved futile. Nevertheless, overexpression was achieved in *E. coli* (Chen et al., 2017). The two-component lanthipeptide *Flv* System from *Ruminococcus flavefaciens* FD-1 was also recombinantly expressed in *E. coli* and found to have one highly conserved  $\alpha$ -peptide and a set of eight structurally diverse  $\beta$ -peptides (Zhao and van der Donk, 2016). This work revealed new features in lanthipeptides which have not been reported previously.

## 2.7 Native biosynthesis and regulation of lantibiotic ruminococcin-A

The ability of bacteria to adapt and survive in an ecological community greatly depends on their potential to sense local environmental changes and devise means to sustain these external influences usually by upregulating or downregulating the expression of specific genes (Dunny and Leonard, 1997). The two-component QS system is implicated in the biosynthesis of most bacteriocins including the prototype lanthipeptide nisin from *L. lactis* (Kuipers et al., 1995, Lubelski et al., 2008) where nisin acts as an inducer in its biosynthetic pathway. Other bacteriocins like plantaricin A from *Lactobacillus plantarum* C11 (Diep et al., 2001) is induced by a bacteriocin-like peptide which is co-transcriptionally synthesized simultaneously in a three-component system. Similar to the latter QS system, a three-component regulatory mechanism is involved in the biosynthesis of the class II lantibiotic ruminococcin-A originally isolated from *Ruminococcus gnavus* E1, a constituent member of the human intestinal microbiota (Gomez et al., 2002).



**Figure 2.6** A hypothetical model describing the trypsin-dependent regulation of ruminococcin-A biosynthesis in *R. gnavus* E1. (a) At lower cell density in the gut, *R. gnavus* E1 synthesizes an inactive inducer peptide (pre-IP). The pre-IP is secreted to the extracellular space. Then at higher cell density, trypsin activates pre-IP to active IP. The histidine kinase RumK monitors environmental changes and senses the accumulation of IP as it reaches a threshold concentration. This leads to the activation of the regulatory protein RumR which in turn triggers differential gene synthesis which also enables RumA to be expressed and secreted concomitantly (Gomez et al., 2002). (b) The peptide sequence of preRumA contains a leader peptide (position -23 to -1), the core peptide (position 1 to 24) and the Gly-Gly motif (position -2/-1) where proteolytic cleavage occurs. Threonine, serine and cysteine side chains involved in thioether crosslinks formation are also indicated.

RiPP-genes are ubiquitously present in anaerobic bacteria where more than 25% host biosynthetic pathways for the natural products (Letzel et al., 2014). *R. gnavus* E1 is a strictly anaerobic Gram-positive bacterium which was first identified in human feces (Ramare et al.,

1993). Ruminococcin C was also reported in *R. gnavus* E1, possessing similar biological activity as ruminococcin-A but different structural and biochemical characteristics (Croston et al., 2011). The biosynthesis machinery of ruminococcin-A consists of 14 ORFs, two, whose functions are yet to be determined. In the schematic drawing of the genetic cluster presented in Figure 2.6a (bottom right), following the ORF1 is a series of five genes *rumFEGHR2*, which have characteristics homologous to the ABC-transporter and other accessory components believed to participate in keeping the producer *R. gnavus* E1 safe from its own toxic product. Multiple copies of the structural peptide-encoding gene appears to be a special feature in the cluster of lanthipeptides from the genus *Ruminococcus* as observed with the *rumA1A2A3* of ruminococcin-A (Gomez et al., 2002) and the two-component lantibiotic *Flv* System from *R. flavefaciens* FD-1 (Zhao and van der Donk, 2016). This physiological arrangement may be necessary to ensure production efficiency and strengthen antimicrobial activity.

Essential components of the gene cluster that are directly implicated in the biosynthesis of RumA are the genes that encode the precursor peptide preRumA, the dual-functional (dehydratase/cyclase) lanthionine synthetase RumM and the dual-functional (protease/transporter) AMS protein RumT (Gomez et al., 2002). Native RumA production is regulated by the abundant trypsin in the gut which activates an inducer peptide (IP) to trigger a cascade of events. The biosynthesis model proposed in Figure 2.6a is an expansion of previous studies (Dabard et al., 2001, Gomez et al., 2002, Gomez et al., 2002), with additional information culled from extended literature. Although the amino acid sequence encoded by the last ORF *rumX* displayed negligible sequence homology with known proteins, its N-terminal 130 residues showed similarity with DNA-associated proteins (Gomez et al., 2002). Studies on bacteriocin-negative *R. gnavus* mutant strain revealed the presence of *rumA<sub>1A2A3</sub>*, *rumTX* genes and 3'-truncated *rumM*, whereas the hypothetical regulatory operon (*rumRK*) and the immunity operon (*rumFEGHR2*) were both absent (Gomez et al., 2002). Although these findings supported the predicted functions of these genes, the characterization of the various products remained open for investigation. Figure 2.6b displays the schematic diagram of the primary structure of preRumA, showing the Gly-Gly motif which is engaged by the N-terminal domain of RumT to activate the peptide, as well as threonine, serine and cysteine side chains supposedly targeted for PTMs formation by RumM.

## 2.8 Isolation of ruminococcin-A from *R. gnavus* E1

Previous investigation in the early 90's on gnotobiotic rats harboring *R. gnavus* E1 in its digestive tract demonstrated the presence of a trypsin-dependent anti-*Clostridium perfringens* (Ramare et al., 1993). Although the strain at that time was characterized as belonging to the

genus *Peptostreptococcus*, subsequent studies verified that it was actually a member of the genus *Ruminococcus* (Dabard et al., 2001). *In vitro* culturing identified RumA (Dabard et al., 2001) and ruminococcin-C (Crosth et al., 2011) as the components responsible for the observed anti-*Clostridium perfringens* activities.

The strict anaerobic nature of *R. gnavus* E1 only allowed successful growth in an anaerobic cabinet, applying brain heart infusion broth as the cultivation medium (Dabard et al., 2001). The culture medium was further supplemented with hemin and yeast extract, as well as trypsin to trigger ruminococcin-A production. Upon all these efforts, Dabard and colleagues succeeded to achieve a product yield of 0.665 microgram RumA per litre of *R. gnavus* E1 culture (Dabard et al., 2001). Such outcome underscores the need to investigate alternative production systems since the current procedure is obviously not optimal for obtaining quantities that would allow the peptide to be evaluated for possible applications in therapeutic treatments, food preservation and livestock farming.

## 2.9 Carriers for heterologous expression of AMPs

Heterologous expression challenges like aggregation, protease cleavage and toxicity to host cells, are prevented by fusing AMPs to tags like glutathione S-transferase (GST), small ubiquitin-like modifier (SUMO), maltose binding protein (MBP) and thioredoxin (Trx) (Li et al., 2009, Li, 2011, Bell et al., 2013, Pane et al., 2016). These tags are then subsequently cleaved to generate the functional peptides. Recombinant production of AMPs in *E. coli* that employ these tags as fusion partners sometimes generate mixed outcomes (Li, 2011). For instance, the performance of GST in various attempts have been inconsistent since some of the system even failed to produce the desired results (Skosyrev et al., 2003, Si et al., 2007, Chen et al., 2008). MBP which is even larger than GST was fused to Human  $\alpha$ -defensins and expressed in *E. coli* as aggregates which were then refolded to obtain the functional peptides (Li and Leong, 2011, Tay et al., 2011). Other tags that have also been used successfully in producing AMPs are thioredoxin and SUMO, which perhaps due to their smaller sizes (11.8 kDa and 11.2 kDa respectively) have yielded higher titres of target precursor peptides expressed in the cytoplasm (Gibbs et al., 2004, Li et al., 2009, Bommarius et al., 2010, Cao et al., 2010, Li et al., 2010).

## 2.10 Objective of this work

Observations in several examples indicate that the natural producers of RiPPs are usually challenging to culture and even when they are successfully grown, the production levels of the desired products are usually very low. Bioinformatics and recombinant technology have allowed the characterization of new sources of natural compounds via metabolic sampling and genome mining approaches. These technologies have enabled the use of ecologically close relations to the native producer strains to produce lanthipeptides in situations where optimizing the natural host failed to produce the desired outcomes. Nevertheless, characteristics like short generation time, high cell density growth, high product yield, and ease of manipulation, which are typical of a surrogate host like *E. coli* may have positive influences on production titres and facilitate certain aspects of peptide bioengineering thus fostering potential therapeutic, agricultural and food processing applications.

Since the production pathways of these peptides include complex modifications machineries which are usually not present in the new expression hosts, they require the transfer of the minimal essential components of the pathway into their own system to be able to produce the active biomolecule. The aim of this work was to recombinantly reconstruct the RumA biosynthesis pathway in an autonomous host and test its amenability to genetic manipulation and to improve the production level of the peptide. Specific biosynthesis genes from the RumA biosynthesis cluster, hosted by the obligate anaerobic bacterium *R. gnavus* E1, had to be isolated and reconstituted in *E. coli* which was the host of choice.

It was desired to first achieve soluble expression of the different genes involved by separately expressing them in *E. coli*. Most importantly, the ruminococcin-A lanthionine synthetase RumM had to be active in order to catalyze formation of PTMs in the precursor peptide preRumA. To address solubility and stability challenges of RumA, a highly expressible and soluble fusion partner was desired. Different cloning and expression strategies were to be applied to design the strains, meanwhile parallelized microtiter-plate screenings of cultivation and production-relevant parameters were also necessary. Active *in vivo* expression of RumM capable of installing all the desired modifications in preRumA was necessary for *E. coli* production to be successful. Product identification and characterization should constitute an important phase of the study. Furthermore, selecting parameters to optimize in both upstream and downstream procedures should be done rationally.

Secondly, the activation of modified preRumA which is the final step in the maturation process performed by the AMS protein RumT was to be tackled. Envisaging the challenges of expressing membrane proteins and the fact that RumT does not engage directly in any PTM formation, alternative activation methods were also to be exploited. This approach involved engineering the precursor peptide to contain alternative cleavage possibilities at position -1 in the leader sequence of preRumA. These included the tobacco etch virus (TEV) protease, GluC,

Trypsin and Factor Xa sites. Moreover, attempts to express, purify and characterize the N-terminal peptidase domain of RumT was also anticipated.

Finally, with a fully functional system in hand, engineering the peptide to expand and/or increasing its antimicrobial activity by incorporating non-canonical amino acids into the peptide would be possible. The system would also offer irrefutable advantages over the native producer in that scale-up studies to achieve economic feasibility would be attainable since complicated and expensive experimental set-ups would not be necessary.

## 3. Materials and Methods

### 3.1 Materials

The various categories of materials used in this study are described in Appendix 6.9 (tables A5.1-A5.3)

### 3.2 Software

Software used in this work are listed in Appendix 6.9 (Table A5.4)

### 3.3 Bacterial strains, growth and cultivation conditions

An overview of bacterial strains used in this work are listed in Appendix 6.9 (Table A5.5). Various *E. coli* strains were used for different purposes including cloning, plasmid maintenance and protein expression. Ruminococcin-A native biosynthesis strain *R. gnavus* E1 was used to isolate the genome from which specific genes in the biosynthesis machinery of RumA were amplified. For storage purposes, *E. coli* strains were grown overnight at 37 °C on an LB agar plate supplemented with the appropriate antibiotic concentrations. Fresh LB medium containing 20 % glycerol was added to the plates containing colonies and then washed off and aliquoted into Roti-Store cryo vials (Carl Roth GmbH, Germany) and stored at -80 °C. *R. gnavus* E1 was cultivated in Brain Heart Infusion (BHI) broth as described elsewhere (Dabard et al., 2001). *B. subtilis* ATCC 6633 cultivations were performed at 30°C with an agitation speed of 200 rpm.

#### 3.3.1 Cultivation media

All media recipes (composed as listed in Table A5.6, Appendix 6.9) were dissolved in ddH<sub>2</sub>O and pH was adjusted when necessary. After dissolution, the media were then autoclaved at 121 °C for 20 min. For solid media cultivation, 1.5 % agar was directly added to the mixture prior to autoclavation, and appropriate amounts of required sterile-filtered antibiotics (as described in Table A5.7, Appendix 6.9) were added to the broth agar medium precooled to around 50 °C. The antibiotic-supplemented broth agar was then poured into sterile petri dishes and allowed to solidify at room temperature (RT). Liquid media were stored at RT while the agar plates were stored at 4 °C. For cloning purposes, SOC medium was used to grow freshly transformed cells. Optimization and expression screening were performed in TB medium and

the EnPresso B growth system (Enpresso GmbH, Berlin, Germany) according to the manufacturer's protocol.

### **3.3.2 Monitoring bacterial growth with a spectrophotometer**

*E. coli* growth was monitored by estimating the concentration of cells using an Ultrospec 3300 spectrophotometer (Amersham). Optical densities (ODs) at 600nm were measured while considering the accuracy limit of the spectrophotometer ( $OD < 0.5$ ). Samples were diluted in 0.9 % NaCl solution for measurements. For online ODs, measurements were performed every 10 min using the Synergy Mx plate reader (Biotek). In this work, online ODs only provided a monitoring tool to ascertain cell growth during fluorescence measurements.

### **3.3.3 Automated fluorescence measurements**

Qualitatively, target product expressions in strains expressing GFP-fused constructs were estimated by measuring their fluorescence. The strains were cultivated in 96-well or 24-well flat-bottom microtiter plates and the relative fluorescence emission of the cells were measured using a Synergy™ Mx monochromator-based multimode microplate reader (Biotek). Time-course GFP fluorescence signal intensities were recorded as relative fluorescent unit (RFU) every 10 min at 528 nm for 15-30 h. As a control, the auto-fluorescence of a strain which does not express GFP was also measured.

## **3.4 Vectors**

An overview of cloning and expression vectors modified or used in this study are described in Appendix 6.2 (Table A1).

## **3.5 Molecular Biology techniques**

Molecular biology procedures were performed following established protocols. T4 DNA ligase, restriction enzymes, nucleotide triphosphates, as well as other DNA-modifying enzymes were purchased from Fermentas (Vilnius, Lithuania), Thermo Fisher Scientific (Waltham) and New England BioLabs (NEB) (Frankfurt am Main, Germany), and used according to the manufacturers' recommended procedures. Only slight modifications were made when it was desired to improve reproducibility.



### **3.5.1 Isolation of *R. gnavus* E1 Genomic DNA isolation**

*R. gnavus* E1 culture was grown at 37°C overnight in a pre-reduced BHI (Difco Laboratories, Detroit, MI) supplemented with 5 g of yeast extract and 5 mg of hemin (Sigma-Aldrich) per litre. The cultivation was performed in an anaerobic cabinet. Cells were harvested by centrifuging 2 ml of the overnight culture for 2 min at 12,000 ×g at 4 °C. Genomic DNA was extracted using GenElute Bacterial Genomic DNA Kit (Sigma-Aldrich), following the manufacturer's instructions. The resulting chromosomal DNA was analyzed with 0.5 % agarose Gel, pre-stained with GelRed (Biotium).

### **3.5.2 Preparations of plasmid DNA**

*E. coli* cells hosting the desired plasmids were cultivated overnight at 37 °C in LB medium. 2-4 ml of the culture was harvested by centrifugation for 1 min at 16,000 ×g. Plasmids were purified from the cell pellets using the Invisorb Spin Plasmid Mini Two Kit (Stratec) according to the manufacturer's protocol. The concentrations of isolated plasmid were estimated using the Nanodrop ND-1000 spectrophotometer (PeQlab) and stored at 4 °C until required for use.

### **3.5.3 Oligonucleotides design**

Vector NTI Advance 11.5 (Invitrogen) was employed in designing oligonucleotide primers for polymerase chain reactions (PCRs) and DNA sequencing. Chemically synthesized primers used for PCRs were purchased from TIB MOLBIOL (Berlin, Germany). An overview of the primers is presented in Appendix 6.9 (Table A5.8).

### **3.5.4 PCR techniques**

#### **3.5.4.1 Analytical PCR**

For analytical purposes, PCR were performed using the DreamTaq DNA polymerase (Thermo Scientific) which has no proof-reading properties. The reaction compositions and conditions were set up following to the manufacturer's recommendation.

#### **3.5.4.2 Preparative PCR**

High-quality PCR products were desired for subsequent cloning purposes. To achieve this, we utilized the Phusion and Q5 high fidelity DNA polymerases (NEB), both possessing the 3'→5' exonuclease activity. All reaction mixtures containing template DNA, deoxynucleotide

triphosphate (dNTP) mix, primer pair and DNA polymerases were assembled and operated in a gradient thermocycler under recommended conditions. Gradient PCR was used to optimize the annealing temperature of each set of primers. Once the optimal amplification condition was determined, it was then used to run a preparative PCR.

### 3.5.4.3 Colony PCR

In this technique, the DreamTaq Green DNA Polymerase (Thermo Scientific) was used in the reaction mix constituted according to the manufacturer's instructions. Two Eppendorf tubes were prepared for each colony, one containing 30 µl ddH<sub>2</sub>O and the other containing 100 µl LB medium supplemented with appropriate antibiotics concentration. Colonies were randomly selected from agar plates using sterile toothpicks. Each colony was first washed in the ddH<sub>2</sub>O and then rinsed in the medium. The inoculated cultures were stored at +4 °C. The Eppendorf tubes containing cells in ddH<sub>2</sub>O were incubated for 5 min at 95 °C to ensure complete disruption of the cell and subsequently placed on ice. 5 µl was obtained from each tube and used as template for the DreamTaq Green reaction mix. Once positive clones were identified, they were matched with the inoculated LB cultures, the plasmids were purified and sent for sequencing.

### 3.5.4.4 Splicing by overlap extension (SOE) PCR

This technique was applied to fuse *gfpmut2* to the gene that encodes RumA precursor peptide (*rumA*). To do this, two special primers *rumA<sup>s</sup>\_f* and *gfp<sup>s</sup>\_r* (Table A5.8, Appendix 6.9) were designed such that their 5' overhangs had partially complementary regions. In a first PCR step, *gfpmut2* and *rumA* were separately amplified from their respective templates using the set of overlap primers respectively designed for each of the target genes. This resulted in two double-stranded DNA fragments with overlapping regions in each of the sequences. In the second stage, quasi-equimolar amounts of the two fragments were mixed together in a mega-primer PCR where the two fragments preliminary served as primers and as template at the same time. They annealed to each other and were joined by amplification using the external primers (see Figure A5.4a and A5.4b, Appendix 6.9, for illustrations). The resulting chimeric product contained an *NheI* restriction site at the 5'-end and a *PstI* site at the 3'-end.

### 3.5.5 Purification of DNA fragments and spectrophotometric quantitation

The quality and quantity of preparative PCR amplicons were preliminarily evaluated via agarose gel electrophoresis. Gene fragments that were successfully amplified were purified

using the HiYield PCR Clean-up/Gel Extraction Kit (Süd-Laborbedarf) according to the manufacturer's protocol. This was done in two ways: either a piece of the agarose gel containing a clear band at the desired product length was excised or the PCR products were applied directly. The purified products were quantified using NanoDrop and directly used for subsequent cloning steps or stored at -20 °C for further use.

### 3.5.6 Agarose gel electrophoresis

With regards to the size range of gene fragments used in this study, 0.5 % or 1 % gels were prepared by dissolving electrophoresis grade agarose in 1X TAE buffer (40 mM Tris-HCl pH 8.5, 20 mM Acetic acid and 1 mM EDTA). The solute was completely dissolved by heating in a microwave. Ethidium bromide (Carl Roth) or GelRed (Biotium) was added (in the ratio of 1:10,000) to the molten gel solution precooled to about 50 °C and then poured into a preassembled casting plate. Analysis of DNA was performed by diluting samples in required amount of ddH<sub>2</sub>O and mixing with 6x DNA loading buffer (Fermentas) in the ratio of 1:6. The sample mixtures were loaded into precasted wells in the gel submerged in 1X TAE buffer in a DNA electrophoresis unit (Biozym), and an electric field was applied. DNA bands were visualized by UV light excitation.

### 3.5.7 Conventional cloning using restriction enzyme digestion and ligation

For analytical purposes, 10 µl of 400-500 ng concentrated DNA was assembled by adding target DNA, sterile ddH<sub>2</sub>O, FastDigest restriction enzymes and FastDigest buffer (Thermo Scientific) according to the manufacturer's manual. The sample mixtures were incubated for 20-30 min at 37 °C. For preparative samples, DNA up to 4 µg was digested for 1-2 h. In cases where the pair of restriction enzymes used were both sensitive to higher temperatures, the digested product was inactivated by heating for 15-20 min at 80 °C. Otherwise the digested DNA was analyzed by agarose gel electrophoresis and corresponding fragments were excised and purified as described earlier (section 3.5.5). Ligation reactions were performed with T4 DNA ligase (NEB) by assembling a 10 µl reaction according to the manufacturer's protocol. The amount of insert added to the mixture was calculated according to the following equation:

$$m_{\text{insert}} [\text{ng}] = X \cdot \frac{\text{length}_{\text{insert}} [\text{bp}]}{\text{length}_{\text{vector}} [\text{bp}]} \cdot m_{\text{vector}} [\text{ng}]$$

Ligation reactions were performed for 10-30 min at RT or overnight at 16 °C. T4 DNA ligase was inactivated by incubating for 20 min at 65 °C. The resulting products were used to transform competent *E. coli* cells.

### 3.5.8 Modification of expression vectors

The main vector backbone used in this study was the pCTUT7 which possessed a pBR322 origin of replication (*ori*) and an Isopropyl  $\beta$ -D-1-thiogalactopyranoside (IPTG)-inducible *CTU* promoter (Kraft et al., 2007), as well as the pJL10 which possessed an RSF *ori* and an IPTG-inducible *CU* promoter (Li et al., 2015). pCTUT7 was derived from pKA100 (Krebber et al., 1996) and pDest15, and possessed mutations in the *lacUV5* sequence of the original pAK100, -35 region from *tac* promoter, ribosomal binding site (RBS) of gene 10 of bacteriophage T7 (T7 RBS) and a recombination cassette amplified from pDest15 (Kraft et al., 2007). The plasmid that hosted the *CU* promoter (pJL10) was a modification of pRSF-1b where the strong T7 promoter was substituted for the *CU* promoter (Li et al., 2015). The difference between the *CTU* (stronger) and *CU* (relatively weaker) promoters were specific mutations at their respective -35 regions (Kraft et al., 2007). It was desired to express the target genes under control of different promoters. To achieve this, some features of the pCTUT7 and pJL07 (Li, 2013), as well as pSEVA1810 needed to be modified. pSEVA1810 is a vector that was designed according to the Standard European Vector Architecture (SEVA) (Martínez-García et al., 2014). It harbours a pUC *ori* and a *pBAD* promoter.

To change the pUC *ori* of pSEVA1810 to p15A *ori*, FseI/AsclI-flanked p15A *ori* amplified from pJL02 (Li, 2013) was digested and inserted into FseI/AsclI-digested pSEVA1810. The resulting vector was named pSEVA15A. The Chloramphenicol resistance marker in pJL02 was replaced with ampicillin resistance by inserting HindIII/Bsu36I-flanked fragment amplified from pJL07 into HindIII/Bsu36I-digested pJL02 to yield pLEOp15A.

The *CTU* promoter in pCTUT7 was replaced with *CU* promoter while maintaining its T7 RBS, by using *SpeI* and *NdeI* to remove an entire cassette comprising of the *lacI* gene and the *CTU* promoter from pCTUT7. *SpeI*/*NdeI*-flanked fragment containing the *lacI* gene together with the *CU* promoter was amplified from pJL10 and subsequently inserted into the backbone of the *SpeI*/*NdeI*-digested pCTUT7 to yield pCUT7. Vector description and plasmid maps are found in Appendix 6.2. Plasmids were purified and sequenced via LGC Genomic (Berlin, Germany).

### 3.5.9 Site-directed mutagenesis

Q5 Site-Directed Mutagenesis Kit (NEB) was used to generate mutants of preRumA that allowed alternative means to remove its leader peptide and activate the compound. Trypsin, GluC, TEV and factor Xa cleavage sites were introduced between the leader and core segments of preRumA. Q5 high-fidelity DNA polymerase produced linear chimeric products which were directly used in the subsequent steps. After degrading methylated DNA followed by ligation, the reaction mix was used to transform chemically competent *E. coli* cells.

#### 3.5.10 Construction of expression vectors

The general characteristics of expression vectors are described in Appendix 6.2 (Table A1).

##### 3.5.10.1 Amplification of gene fragment from *R. gnavus* E1 chromosome

*R. gnavus* E1 genome was used as a template together with primer pair *RumClus\_f/rumClus\_r* (Table A5.8, Appendix 6.9) to amplify a segment of the *rumA* gene cluster (GenBank accession no. AF320327), containing the three genes encoding preRumA and the gene encoding the lanthionine synthetase RumM. The blunt-ended amplicons were directly inserted into a *Sma*I-restricted pCTUT7 vector to yield pLEOrC2.

##### 3.5.10.2 Construction of His6-preRumA expression vectors

For preRumA expression, *Nhe*I/*Pst*I-flanked *rumA1* PCR product was amplified from pLEOrC2. The resulting amplicons were digested and ligated into *Nhe*I/*Pst*I-digested pCTUT7 vector yielding pLEOrA.

##### 3.5.10.3 His6-preRumA vectors with alternative cleavage site at position -1

To introduce Trypsin, GluC, Factor Xa and TEV cleavage sites at position -1 of preRumA, site directed primers (Table A5.8) were designed for the respective purposes. PCR, using pLEOrA as template together with the Q5 Site-Directed Mutagenesis Kit and the respective primer pairs generated pLEOrA\*, pLEOrA<sup>GluC</sup>, pLEOrA<sup>Xa</sup>, pLEOrA<sup>TEV</sup>, each possessing trypsin, GluC, factor Xa and TEV cleavage sites, respectively, at position -1 in the final preRumA expression product.

#### 3.5.10.4 Construction of His6-GFP-TEV-preRumA expression vectors

The final *NheI/PstI*-flanked chimeric fragment from SOE PCR was digested and ligated into corresponding *NheI/PstI*-digested pCTUT7, pJL07 and pLEOp15A to yield pLEOgrA, pLEOp15A\_grA and pLEOt0grA respectively. To eliminate multiple products expression from a single gene construct, additional features like a  $\lambda$ t0 terminator downstream and a tandem stop codon at the 3'-end of the chimeric gene fragment were included in the pLEOp15A\_grA and pLEOt0grA vectors. To substitute the Gly-1 in preRumA fused to GFP with Arg, pLEOrA\* was used as template in the SOE PCR and the resulting products were digested and inserted into pCTUT7 to yield pLEOgrA\*.

#### 3.5.10.5 Construction of His6-SUMO-preRumA expression vectors

To fuse preRumA to SUMO, *NheI/PstI*-flanked gene fragments encoding the preRumA were amplified from pLEOrA\*, pLEOrA<sup>Xa</sup> and pLEOrA<sup>TEV</sup> respectively. The resulting products were digested and inserted into *NheI/PstI*-digested pCTUT7-SUMO plasmid (Kraft et al., 2007) to yield corresponding pLEOsrA\*, pLEOsrA<sup>Xa</sup> and pLEOsrA<sup>TEV</sup>.

#### 3.5.10.6 Construction of His6-RumM expression vectors

For the expression of the ruminococcin-A lanthionine synthetase RumM, PCRs with appropriate primer pairs (Table A5.8) using pLEOrC2 as template were performed, and the resulting amplicons were digested with either *NheI/PstI*, *KpnI/PstI* and *BamHI/PstI*, and correspondingly ligated into the pCTUT7, pRSF-1b and pJL10, generating pLEOrM', pRSF-rM and pLEOrM respectively. Similarly, *NheI/PstI*-digested *rumM* amplicons were inserted into *NheI/PstI*-restricted pCUT7 vector to yield pLEOrM1.

#### 3.5.10.7 Construction of bicistronic expression vectors

Appropriate primer pairs (Table A5.8) were used to amplify a *PstI/HindIII*-flanked gene fragment from either pLEOrM' or pLEOrM1 containing the *CTU* or the *CU* promoter respectively, together with a T7 RBS and *rumM*. The resulting amplicons were digested with *PstI* & *HindIII* and subsequently inserted into *PstI/HindIII*-digested pLEOrA<sup>TEV</sup>, pLEOgrA, pLEOgrA\*, pLEOsrA\*, pLEOsrA<sup>Xa</sup> and pLEOsrA<sup>TEV</sup> to generate pLEOrA<sup>TEV</sup>M1, pLEOgrAM, pLEOgrA\*M1, pLEOsrA\*M, pLEOsrA<sup>Xa</sup>M and pLEOsrA<sup>TEV</sup>M respectively.

### 3.5.10.8 Construction of RumT125 expression vectors

The nucleotide sequence encoding the first 125 N-terminal amino acid residues of the bifunctional AMS protein RumT was amplified from a cDNA optimized for *E. coli* expression and purchased from Thermo Fisher Scientific (Waltham). The resulting *NheI/PstI*-flanked PCR amplicons were digested and inserted into *NheI/PstI*-digested pCTUT7 and pCTUT7-SUMO to yield pLEOrT125 and pLEOsrT125 respectively.

### 3.5.11 Transformation of competent *E. coli* cells

Chemically competent and electrocompetent *E. coli* cells were prepared according to a standard protocols (Sambrook and Russell, 2001) and stored at -80 °C. For chemical transformation, 100 µl of competent cells were thawed on ice and 1 µl of the desired plasmid or ligation product was added and mixed gently. The mixture was incubated on ice for further 30 min and subsequently placed at 42 °C for 45 s. The heat-shocked samples were incubated on ice for 5 min. 900 µl of SOC medium was transferred into the transfected cell culture and grown at 37 °C while shaking for 1 h. The transformed cells were spread on LB agar plates prepared with appropriate antibiotic concentrations and cultivated overnight at 37 °C. Single colonies were obtained from the plate and screened for positive clones via colony PCR. Major strains used in this study are described in Appendix 6.9 (Table A5.5).

## 3.6 Protein production and analysis

### 3.6.1 Expression optimization in EnPresso B growth system

Expression optimization were performed in 24-well deep-well plates using the read-to-use enzyme-based automated glucose-delivery EnPresso B growth system (Enpresso GmbH, Berlin, Germany), following the manufacturer's recommended procedure. *E. coli* W3110 transformed with the desired expression vectors were cultivated in an orbital shaker (Infors HT, Switzerland) at 30 °C, 250 rpm overnight. Prior to induction (after about 12-15 h of cultivation), 3 ml of the main culture was distributed into the 24-well plate and induced with different concentrations of IPTG. Variable amounts of reagent A (the glucose-releasing biocatalyst that slowly degrades a complex polymer component of the EnPresso B medium tablet to deliver glucose to the growing cells) were also supplied to the different wells. The plate was sealed by Breathable Film (Starlab, Hamburg) and placed back into the orbital shaker under the same conditions as before. The optical densities of the cultures at 600 nm (OD<sub>600</sub>) were measured at different time points during a period of cultivation ranging from 18 to 24 h post induction. For

GFP fusion construct, protein expression was followed by measuring GFP fluorescence signals online (section 3.3.3). Cultivation was stopped by obtaining 2 ml of the cultures into 2-ml Eppendorf tubes and harvesting the cell pellets by centrifuging in a Himac /CT15RE centrifuge (VWR, Leuven) for 1 min at 16,000  $\times g$  at 4 °C. The cell pellets were stored at -20 °C for further preparations and analysis.

### 3.6.2 Protein Expression in TB medium

Precultures of *E. coli* W3110 transformed with the desired expression vectors were prepared by using glycerol stocks from -80 °C to inoculate 20 ml of LB medium supplemented with appropriate antibiotics. The preculture was grown overnight at 37 °C, 200 rpm. For expression screening purposes, 100 ml of fresh TB medium containing adequate antibiotic concentrations in a 500-ml baffled Ultra Yield Flask (Thomson Instrument Company, USA) was inoculated with the overnight culture to an initial OD<sub>600</sub> of 0.1. Larger preparative cultures were inoculated in a 2-litre Ultra Yield Flask containing 400 ml of fresh TB medium supplied with suitable antibiotics. The flask was sealed with air-permeable AirOtop Enhanced Seals (Thomson Instrument Company, USA) and incubated in the orbital shaker at 30 °C, 200 rpm. At induction OD<sub>600</sub> between 0.8 and 1, 3 ml each from the 100 ml culture were distributed into a 24-well microtiter plate and induced with different concentrations of IPTG, while the preparative culture was induced with 100  $\mu M$  IPTG (optimized from micro-scale cultivations) and further cultivated for 18-30 h. The OD<sub>600</sub> of the cultures were measured at different time points while online GFP fluorescence signals was used to monitor the expression of GFP fusion constructs. Small scale cultures were harvested by collecting 2 ml and centrifuging for 1 min at 16,000  $\times g$  at 4 °C. Cells from the preparative cultures were collected by centrifugation using the Eppendorf centrifuge 5810R (Eppendorf, Hamburg) for 10 min at 6000  $\times g$  at 4 °C. Cell pellets were stored at -20 °C for further preparations and analysis.

### 3.6.3 Strain screening using GFP fluorescence

The efficiency and reproducibility of strains producing GFP fusion products were evaluated based on their GFP fluorescence over a longer cultivation period. These experiments were performed using the Hamilton Microlab STAR Liquid-handling station (Hamilton, Martinsried) coupled to the Synergy Mx plate reader (Figure A5.5, Appendix 6.9). At the beginning, 96 colonies were selected manually from an LB agar plate and inoculated 150  $\mu l$  of LB or TB medium supplied with appropriate antibiotics in each of the wells of a 96-well flat-bottom microtiter plate. The cultures were induced automatically with 100  $\mu M$  IPTG 8 h after the



cultivation started. Prior to induction, the robot was programmed to sample 10 µl each from all 96 wells and transfer into a new plate (containing 10 µl of 30 % glycerol prepared in LB medium) placed in one of the 96-well microtiter plate (MTP) positions on the deck (Figure A5.5, Appendix 6.9). The latter plate was collected and stored at -80 °C. IPTG solution used for induction was also stationed in one of the 96-well MTP positions. Measurements were performed every 20 mins. The robot transported the culture plate from the incubator to the plate reader and transferred it back into the incubator once measurements were finished. Online OD<sub>600</sub> were also measured to determine if the cells were growing.

### 3.6.4 Cell disruption

Prior to purification, *E. coli* cells were lysed by a combined enzymatic and ultrasonication procedure. Frozen cells from -20 °C were thawed at RT and resolubilized in the IMAC binding buffer (50 mM NaH<sub>2</sub>PO<sub>4</sub>·H<sub>2</sub>O, pH 8, 300 mM, 10 mM Imidazole). Fresh lysozyme and benzonase were added to the cell suspension to a final concentration of 1 mg ml<sup>-1</sup> and 1 U ml<sup>-1</sup> respectively. Ultrasonication was performed on ice for 5-12 min using 30 sec on/off intervals using the UP200S sonicator (Hielscher Ultrasonics, Teltow). Depending on the sample size, different probes were used. Samples from screening experiments were processed using the detergent-based cell disruption technique, BugBuster (Merck, Darmstadt, Germany) protocol, following the manufacturer's recommendations. Smaller lysates were clarified by centrifugation for 15 min at 16,000 ×g, 4 °C using a in a Himac centrifuge, while larger sample volumes were clarified by centrifuging for 20-30 min at 10,000 ×g, 4 °C using an Eppendorf centrifuge 5810R. The samples were filtered through a 0.45-µm pore-sized syringe filter (Carl Roth) and used for subsequent purification steps. When whole cell extracts were desired for SDS-PAGE comparative analyses, BugBuster was added to the collected cell pellets according to the following equation:

$$\text{Volume of BugBuster } [\mu\text{l}] = 67 \cdot \frac{\text{OD}_{600}}{2} \cdot \frac{\text{Original culture volume } [\mu\text{l}]}{1000};$$

where original culture volume is the volume of culture collected; OD<sub>600</sub> is the final optical density of the culture; 2 is the dilution factor included because of the 2X loading buffer used for SDS sample preparation; 67 is the estimation factor; and 1000 converts volume units from milliliter to microliter.

### 3.6.5 SDS-PAGE analysis

Insoluble fractions from the BugBuster experiments were resuspended in resolubilization buffer (100 mM Tris-HCl pH 8.0, 8 M Urea, 100 mM DTT, 1 mM EDTA) with the same volume as used in resuspending the cell pellet. The suspension was vortexed and centrifuged for 5 min at 16,000 ×g at 4 °C. The supernatant, as well as soluble protein samples were mixed in a one-to-one ratio with 2X SDS sample buffer (100 mM Tris-HCl pH 6.8, 200 mM DTT, 4 % SDS, 20 % Glycerol 0.20 % bromphenol blue). The samples were boiled for 5 min at 95 °C and then cooled to RT. 10 µl of each sample were loaded into the wells of the SDS gel. Roti-Mark TRICOLOR (Carl Roth) was used as molecular weight standard.

Larger proteins were processed using 10 % resolving, with 4 % stacking polyacrylamide gel prepared following the protocol described by Sambrook and Russel (2001). The SDS running buffer (25 mM Tris-HCl pH 8.3, 192 mM Glycine, 500 mM Urea, 0.10 % SDS) was used in the electrophoresis chamber. A voltage of 70 V was applied across the two electrodes (anode/cathode) for ~20 min and subsequently increased to 120 V until the bromophenol blue dye front was almost out of the gel. Smaller peptides (<10 kDa) were analyzed using 16 % Tricine PAGE as described elsewhere (Schägger, 2006). Gels were rinsed with ddH<sub>2</sub>O and then washed twice for 2 min with warm ddH<sub>2</sub>O. All gels were visualized by staining with colloidal blue silver Coomassie G-250 as reported elsewhere (Candiano et al., 2004). Background staining were removed by rinsing the gels severally with distilled and when the bands were sufficiently clear, the gels were documented using a scanner.

### 3.6.6 Native PAGE analysis

Sample preparations and running of the gel were performed in the same manner as described for SDS-PAGE except for the fact that detergents (SDS and urea) and DTT were not included in the native PAGE assembling, sample and running buffers. Also, the sample heating step (95 °C, 5 min) was not performed. Additionally, 8 % resolving, with 4 % stacking polyacrylamide gels were used for the analysis.

### 3.6.7 Western blot

Semi-dry electroblotting was applied using a horizontal apparatus (Cleaver Scientific). Gels from SDS-PAGE was rinsed with ddH<sub>2</sub>O and incubated in the transfer buffer (25 mM Tris-HCl pH 8.3, 192 mM Glycine, 10 % MeOH, 0.1 % SDS) for 15 min. Six pieces of equally sized (the

same size as the gel containing proteins to be transferred) Whatman Filter paper (Whatman, UK) previously soaked in the transfer buffer were placed on the anode plate of the electroblotting unit. The blot sandwich was completed by placing a same-sized nitrocellulose membrane, followed by the gel and another set of six wet filter papers. The upper plate (cathode) was fitted onto the blot sandwich and a voltage of 20 V was applied across the electrodes for 1 h at RT. To detect the poly-histidine tag on the protein of interest, the blotted membrane was first incubated in Blocking buffer (5 % skimmed milk powder in TBST) while shaking for 1 h at RT. The membrane was subsequently rinsed with TBST buffer (25 mM Tris-HCl pH 7.4, 150 mM NaCl, 0.1 % Tween 20) and incubated in 1:20,000 dilution of anti-penta-His mouse IgG1 (Quiagen) in blocking buffer (primary antibody). The membrane was incubated for 1 h at RT, washed and further incubated in 1:10,000 dilution of alkaline phosphatase (AP)-conjugated anti-Mouse IgG antibody (Sigma) in blocking buffer. The membrane was rinsed with AP buffer (100 mM Tris-HCl pH 9.5, 100 mM NaCl, 5 mM MgCl<sub>2</sub>) and visualization was achieved via the reaction on the chromogenic BCIP/NBT substrates catalyzed by alkaline phosphatase [80 µl BCIP (20 mg ml<sup>-1</sup> BCIP in 100 % DMF) plus 60 µl NBT (50 mg ml<sup>-1</sup> NBT in 70 % DMF) in 10 ml AP Buffer].

## 3.7 Protein purification and peptide extraction

### 3.7.1 IMAC

For analytical purification of His6-preRumA, His6-GFP-TEV-preRumA and His6-RumM, the clarified lysates obtained after cell disruption were applied to His SpinTrap columns (GE Healthcare) and processed according to the manufacturer's instructions. For comparative analysis, all samples were processed following the same procedure and eluted in the same amount of elution buffer. Equal volumes were then used for SDS-PAGE analysis. Preparative purifications were performed with the Äkta Avant 25 instrument by loading the clarified bacterial lysates onto a 1-ml HisTrap FF Crude column (all from GE Healthcare) at a flow rate of 1 ml min<sup>-1</sup>. UV signals at 280 nm, 260 nm and 484 nm (GFP fluorescence) were used to monitor purification. The column was initially flushed with five column volumes of binding buffer (50 mM NaH<sub>2</sub>PO<sub>4</sub>·H<sub>2</sub>O, pH 8, 300 mM, 10 mM Imidazole) and eluted by applying a gradient range of 0–100 % of elution buffer (50 mM NaH<sub>2</sub>PO<sub>4</sub>·H<sub>2</sub>O pH 8, 300 mM NaCl, 500 mM imidazole). The purified constructs were concentrated using Amicon Ultra centrifugation tubes (Merk Millipore, Darmstadt) with the appropriate molecular weight cut-off limits (10 kDa for SUMO fused proteins, 30 kDa for GFP fused constructs and 100 kDa for His6-RumM). Protein concentrations were measured using the Bradford assay.

### 3.7.2 Size exclusion chromatography

The IMAC concentrated sample was injected onto a HiLoad 16/60 Superdex pg 200 column (GE Healthcare) and eluted with the gel filtration buffer (20 mM  $\text{NaH}_2\text{PO}_4 \cdot \text{H}_2\text{O}$  pH 8, 150 mM NaCl, 10 % glycerol) at a constant flow rate of  $1.5 \text{ ml min}^{-1}$ . Samples were obtained from each of the 2 ml elution fraction and analyzed via SDS-PAGE. Fractions were pooled and concentrated to approx. 4-5 ml. Concentrations were measured using Bradford assay and aliquots were stored at  $-20^\circ\text{C}$ .

### 3.7.3 TEV Cleavage

TEV was applied to cleave off fusion partners from the desired product by adding it to the target sample according to the following equation:

$$V_{\text{TEV}} = \frac{m_{\text{target}}}{C_{\text{TEV}} \times 100};$$

where V represents the volume; m, mass (mg) of target protein to be digested; and C, concentration ( $\text{mg ml}^{-1}$ ) of TEV. All reactions were performed overnight at  $4^\circ\text{C}$ . The procedure followed one of three ways: (i) by adding fresh homemade TEV (see Appendix 6.10) to the IMAC concentrated samples and directly loaded into an appropriate dialysis membrane and then dialyzed in  $1-2 \times 10^3$  volumes of dialysis buffer (20 mM  $\text{NaH}_2\text{PO}_4 \cdot \text{H}_2\text{O}$ , pH 8, 150 mM NaCl, 1 mM DTT, 0.5 mM EDTA, 1 mM PMSF); (ii) by buffer-exchanging the protein samples in dialysis buffer and TEV was subsequently added; and (iii) by thawing gel filtration samples from  $-20^\circ\text{C}$  on ice, and 1 mM DTT, 0.5 mM EDTA and TEV were added.

### 3.7.4 Extraction of preRumA from TEV-digested product

To extract the desired preRumA peptide from TEV-digested samples, the samples were run through a  $\text{Ni}^{2+}$ -NTA column and the flow through containing the desired product was collected, dialyzed in water and dried using Concentrator Plus vacuum concentrator (Eppendorf, Hamburg, Germany). The dried samples were stored at  $-20^\circ\text{C}$  for further use. Alternatively, the TEV-digested samples were extracted with 1-butanol by adding 1 volume of the solvent and stirring the mixture at RT for 1-2 h. Extracts were separated by centrifugation for 5 min at  $3000 \times g$ . The upper organic layer (containing the peptide) was obtained and the concentrations were

measured using the NanoDrop technique. Recovered peptide extracts were dried and stored at -20 for subsequent applications.

### 3.8 Measurement of protein concentrations using Bradford assay

Bradford Coomassie Brilliant Blue method (Bradford, 1976) was used to measure protein concentrations. Samples were assembled by adding 10  $\mu$ l protein solution to 90  $\mu$ l of Roti-Quant Bradford solution (Carl Roth). The resulting solution was mixed and incubated in 96-microwell plate for 5 min at RT. Measurement was performed using the Infinite 200 plate reader (Tecan). The absorption value at 595 nm was compared with a standard curve prepared using known concentrations of BSA.

### 3.9 Iodoacetamide derivatization and trypsin digestion

Iodoacetamide derivatization was used to determine the availability of side chain cysteinyl thiols in the core of preRumA following presumed lanthionine modifications. Meanwhile trypsin was applied to cleave the leader peptide from preRumA and consequently activating the compound. Dried extracts from -20 °C were dissolved in 100 mM ambic (prepared by adding 0.79g ammonium bicarbonate to 100 ml ddH<sub>2</sub>O) to obtain ~1 mg ml<sup>-1</sup> of soluble peptide. 30  $\mu$ l (~30  $\mu$ g of peptide) was obtained and 2  $\mu$ l of 160 mM DTT (final conc. = 10 mM) was added to the peptide solution and incubated at 55°C for 1 h. 400 mM fresh stock solution of iodoacetamide (Sigma) was prepared by weighing 36 mg of the solute crystals and dissolving in 100 mM ambic. 2  $\mu$ l of the alkylating agent was added to the reduced sample (final conc. = 25 mM) and then incubated at room temperature for 45 min in the dark. After the derivatization was completed, 6.3  $\mu$ l of 0.1  $\mu$ g  $\mu$ l<sup>-1</sup> mass spectrometry grade trypsin (Promega, Lyon, France) stock solution was added and incubated on a shaking platform at 37 °C overnight. For peptide identification purposes, gel pieces were excised from SDS-PAGE and in-gel tryptic digestion was performed as described elsewhere (Shevchenko et al., 2006). Prior to mass spectrometric analyses, the tryptic digests were purified using ZipTip.

For modified and non-modified His6-SUMO-RumA\*, the derivatization step was not performed. The IMAC-purified samples in elution buffer 2 (20 mM TRIS-HCl, pH 8, 250 mM NaCl, 300 mM imidazole) were dialyzed in water and then in 100 mM ambic. 15  $\mu$ l of trypsin stock solution (1.5  $\mu$ g) was added to 100  $\mu$ l of protein samples obtained from each of the modified and non-modified His6-SUMO-RumA\* and then incubated on a shaking platform at 37 °C overnight. SDS-PAGE samples were obtained from the tryptic digests and the rest was used for bioassay analyses.

### 3.10 Mass spectrometric analyses

#### 3.10.1 MS sample preparation with ZipTip

ZipTip (Merck) is a 10  $\mu$ l pipette tip possessing a non-polar (C18) chromatographic matrix immobilized inside the end of the tip. This technique is ideal for purifying and concentrating peptides for MS analyses. To do this, two tubes were prepared with 10  $\mu$ l each of 30 or 80 % acetonitrile (ACN) in HPLC-grade H<sub>2</sub>O, containing 0.1 % formic acid (FA). Dried samples from -20 °C were dissolved in 0.1 % FA and placed in ultrasonic bath for 30 s to ensure complete dissolution. The dissolve sample was aspirated and dispensed severally through the column matrix of a pre-equilibrated ZipTip. The loaded ZipTip column was washed thrice with 0.1 % FA by aspirating and dispensing the solvent to remove contaminants, and then eluted in the tube containing 10  $\mu$ l of 30 % ACN tube, followed by the 80 % ACN tube. The column was equilibrated again, and the ZipTip procedure was performed three times. The two elution fractions were combined and dried using a vacuum concentrator. Furthermore, when desired, the ZipTip procedure was also applied directly to the purified TEV-digested samples or tryptic digests.

#### 3.10.2 LC-ESI-MS analyses of preRumA

For the liquid chromatography-electrospray ionization-mass spectrometry (LC-ESI-MS) analyses, dried ZipTip extracted samples were suspended in 15-20  $\mu$ l of 0.1 % FA and dissolved by placing in ultrasonic bath for 30 s. 5  $\mu$ l was injected into an Agilent 1290 Infinity HPLC system (Agilent Technologies, Waldbronn, Germany), followed by an ESI-Triple-Quadrupole LC-MS 6460 electron spray ionization mass spectrometry analysis using multiple reaction monitoring. The material used for the column and pre-column was Poroshell 120 EC-C8 (2.1  $\times$  50 mm, 2.1  $\mu$ m). Solvents allocation was designated as: H<sub>2</sub>O for A and ACN as solvent B. The introduced sample was eluted with an ACN gradient from 5 % to 20 % solvent B within 0.5 min, followed by 70 % B in 4 min and then to 100 % B in 0.2 min. An isocratic elution with 100 % B was finally applied for 1.3 min. The flow rate was set at 0.7 ml min<sup>-1</sup>.

#### 3.10.3 Tandem mass spectrometric analyses of preRumA

Whereas LC-MS analyzes only the precursor ion, LC-MS<sup>2</sup> filters out the precursor ion, fragment it and analyze the product ions with a second mass analyzer (see Appendix 6.8.5). Nano-LC-ESI-MS<sup>2</sup> is a widely used method to identify peptides in a complex mixture. Dried samples from

-20 °C were resolubilized as described in the previous section. For LC-MS<sup>2</sup> analysis, the samples were directly injected into an HPLC-coupled LTQ Orbitrap XL™ Hybrid Ion Trap-Orbitrap mass spectrometer with a nanoelectrospray ion source. Highly intense ions identified in the full MS were directed to the orbitrap analyzer where they were fragmented and processed in the linear ion trap (see Appendix 6.8.5). For the nLC-ESI-MS<sup>2</sup>, the dried samples were redissolved in 50 µl of 0.1 % FA. The processed samples were directly injected into the Orbitrap Fusion mass spectrometer (Thermo Scientific). Specific ions were selected and fragmented with HCD, CID or ETD (Appendix 6.8.3). Different fragmentation energies (in volt) were applied.

### 3.11 Biological Assay

*B. subtilis* ATCC 6633 was reported by Dabard et al. (2001) as one of the indicator strains for RumA biological activity. The strain was cultivated in M<sup>+</sup> medium (Table A5.6, Appendix 6.9) as described earlier and then spread on M<sup>+</sup>-agar plates containing punched wells. The plates were air-dried and 50 µl of each sample was pipetted into separate wells and incubated at 30 °C overnight.

### 3.12 Identification of optimal production conditions using online DOT & pH

To determine certain parameters that may influence a production process, 24 different representative experiments were designed by altering the concentrations of IPTG and reagent A at the point of induction as shown in Figure 5.6. These experiments were aimed at optimizing the expression of His6-GFP-TEV-preRumA and His6-RumM during cultivation in EnPresso B growth system. Three different plates were set up in parallel using the same conditions. One of the plates was a normal 24-well shallow plate which was used to measure online OD<sub>600</sub> and GFP fluorescence. The other two plates were; a 24-well deep-well plate possessing a pH sensor at the bottom of each well (HydroDish HD24) and a 24-well shallow plate with oxygen sensors at the bottom of each well (OxoDish OD24), all from PreSens. The culture volume in the wells of the shallow plates was maintained at 1 ml per well while the deep-well plate was set at 2.5 ml per well. To prevent unprecedented evaporation of liquid between the wells, Duetz System sandwich covers and cover clamps (EnzyScreen, Heemstede, Netherlands) were used to close the microtiter plates (Duetz et al., 2000).

		1	2	3	4	5	6
A	Reagent A [U]	0.8	1.2	1.6	2	2.4	2.8
	IPTG [ $\mu$ M]	20	20	20	20	20	20
B	Reagent A [U]	0.8	1.2	1.6	2	2.4	2.8
	IPTG [ $\mu$ M]	50	50	50	50	50	50
C	Reagent A [U]	0.8	1.2	1.6	2	2.4	2.8
	IPTG [ $\mu$ M]	100	100	100	100	100	100
D	Reagent A [U]	0.8	1.2	1.6	2	2.4	2.8
	IPTG [ $\mu$ M]	500	500	500	500	500	500

**Figure A5.6** 24-well plate experimental design for screening optimal growth and production conditions in EnPresso B medium. The concentrations of reagent A and IPTG are indicated for each well.

The cultivations were performed for 20 h after induction at 30 °C and 200 rpm while measuring dissolved oxygen tension (DOT) and pH online using the SensorDish Reader (PreSens). GFP fluorescence and OD<sub>600</sub> were measured online at 10 min intervals with the Synergy Mx plate reader (Biotek). At line OD measurements of cultures in the first row (A1-6) of the deep-well plate (the pH measurement plate) were performed at different time points by diluting 10  $\mu$ l of the culture in 990  $\mu$ l of 0.9 % NaCl solution in 1.5 ml Eppendorf tubes and measured using the Ultrospec 3300 spectrophotometer (Amersham). Expression of His6-GFP-TEV-preRumA was determined by the GFP fluorescence while both His6-GFP-TEV-preRumA and His6-RumM were determined by SDS-PAGE analysis of samples obtained at the end of the cultivation.

### 3.13 Computational analyses

For illustration purposes, protein structure data were obtained from the protein data bank (PDB) and edited using protein workshop (Moreland et al., 2005). Nucleotide sequence data were obtained from the National Center for Biotechnology Information (NCBI) GenBank database (Benson et al., 2014) and the European Molecular Biology Laboratory (EMBL) nucleotide sequence database (Kulikova et al., 2006). Information about lanthipeptides and other AMPs were acquired from Bagel3 (van Heel et al., 2013), BACTIBASE (Hammami et al., 2010) and the AMP database, APD (Wang et al., 2016). Polypeptide sequence information was acquired from the UniProt database (Consortium, 2016).

The NCBI's Basic Local Alignment Search Tool (BLAST) (Altschul et al., 1990) was used to investigate similarities between nucleotide or protein sequences. Vector NTI Advance 11.5 (Invitrogen) was used for *in silico* molecular biology experimental evaluation including primer



design, cloning, analysis of sequencing data and manipulation of DNA and protein sequences. The Vector NTI was also used to graphically view DNA sequences and to generate plasmid maps and linear graphical presentation of DNA sequences. The codon usage between *E. coli* and *R. gnavus* E1 was compared using the Graphical Codon Usage Analyzer 2.0 (<http://gcua.schoedl.de>) and the Rare Codon Analysis Tool (GenScript) was used to analyze rare codons and calculate the Codon Adaptation Index (CAI). Raptor X server (Källberg et al., 2012) was used to predict protein structures and the prediction data was processed using PyMOL 2.1 (Schrödinger, New York, US).

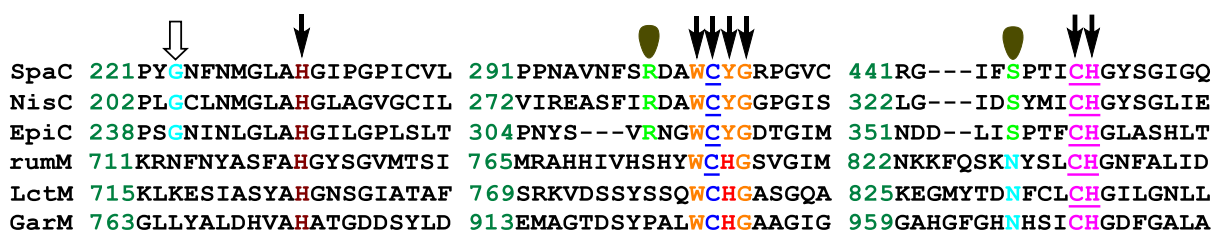
Peptide masses and isotopic distribution were calculated using mMass 5.5.0 (Strohalm et al., 2010). Mass spectra data acquisition and analyses were performed with either the Analyst 1.4.2 software (AB SCIEX) or the MassHunter Qualitative Analysis Software (Agilent) or Xcalibur 4.1 (Thermo Scientific). MaxQuant (Cox and Mann, 2008) and Xi Spectrum Viewer (<http://spectrumviewer.org/>) were also used for MS data analyses. MASCOT search engine (Matrix Science, UK) was used for sequence identification by searching the MS and MS<sup>2</sup> data. Chromatographic data obtained during protein purification with the Äkta avant 25 system were acquired using the UNICORN 6.1 software (GE Healthcare). Chemical structures and schematic representation of biomolecular structures were created with ChemBioDraw Ultra 14.0 (PerkinElmer). ChemBioDraw Ultra 14.0 was also used to determine the chemical formulae of peptides using their amino acid sequences. MicrolabSTAR VENUS one software (Hamilton) was used to control the MicrolabSTAR liquid handling station. Graphs were generated with SigmaPlot (Systat Software).

## 4. Results and Discussion

### 4.1 Sequence analysis of ruminococcin-A processing enzymes

#### 4.1.1 Sequence similarity between RumM and other lanthionine synthetases

The amino acid sequences of three dehydratases (LanB enzymes) of class I lanthipeptides and their corresponding cyclases (LanC enzymes) were aligned together with three class II lanthionine synthetases (LanM enzymes), including RumM. Although the dehydratase domains of the LanM enzymes showed negligible sequence homology with each other and none with their class I equivalents, their C-terminal region was found to contain the conserved zinc-binding residues present in lanthionine cyclases (Li et al., 2006, Helfrich et al., 2007). This domain also shared significant homology with each other as well as other LanC enzymes. In fact, the C-terminus of RumM displayed >30 % identity and >60% similarity with nisin cyclase (NisC), indicating possible structure-function similarities. This type of structure-function similarity has been described in a variety of the lanthionine-generating enzymes (Knerr and van der Donk, 2012, Repka et al., 2017). The conserved active site residues are indicated in Figure 4.1. The reaction catalyzed by cyclases involves deprotonation of cysteinyl thiol and protonation of the resulting enolate intermediates, formed via a conjugate addition reaction between the  $\alpha$ -carbon of Dha/Dhb substrates as described earlier for nisin (see Figure 2.2a, section 2.3.1).



**Figure 4.1** The highly conserved domains of selected class I cyclases including subtilin [SpaC, Swiss-Prot accession number P33115 (Klein et al., 1992)], nisin [NisC, CAA48383 (Engelke et al., 1992)] and epidermin [EpiC, CAA44254 (Schnell et al., 1992)]; as well as class II LanM cyclase domains of ruminococcin-A [RumM, Q9L3F1 (Gomez et al., 2002)], lactacin 481 [LctM: P37609 (Rince et al., 1994)] and actagardine [GarM, C4NFI1 (Boakes et al., 2009)]. The underlined residues are responsible for coordinating the central  $Zn^{2+}$  ion as shown in the crystal structure of NisC (Li et al., 2006). The residues indicated by arrows are essential for an active modification of subtilin, whereas those indicated by single orbitals appear to have no effect on SpaC catalytic activity (Helfrich et al., 2007). The conserved Gly residue denoted by an open arrow was shown to be highly implicated in the maturation of epidermin (Agustin et al., 1992). [see section 3.13 for analysis methods]

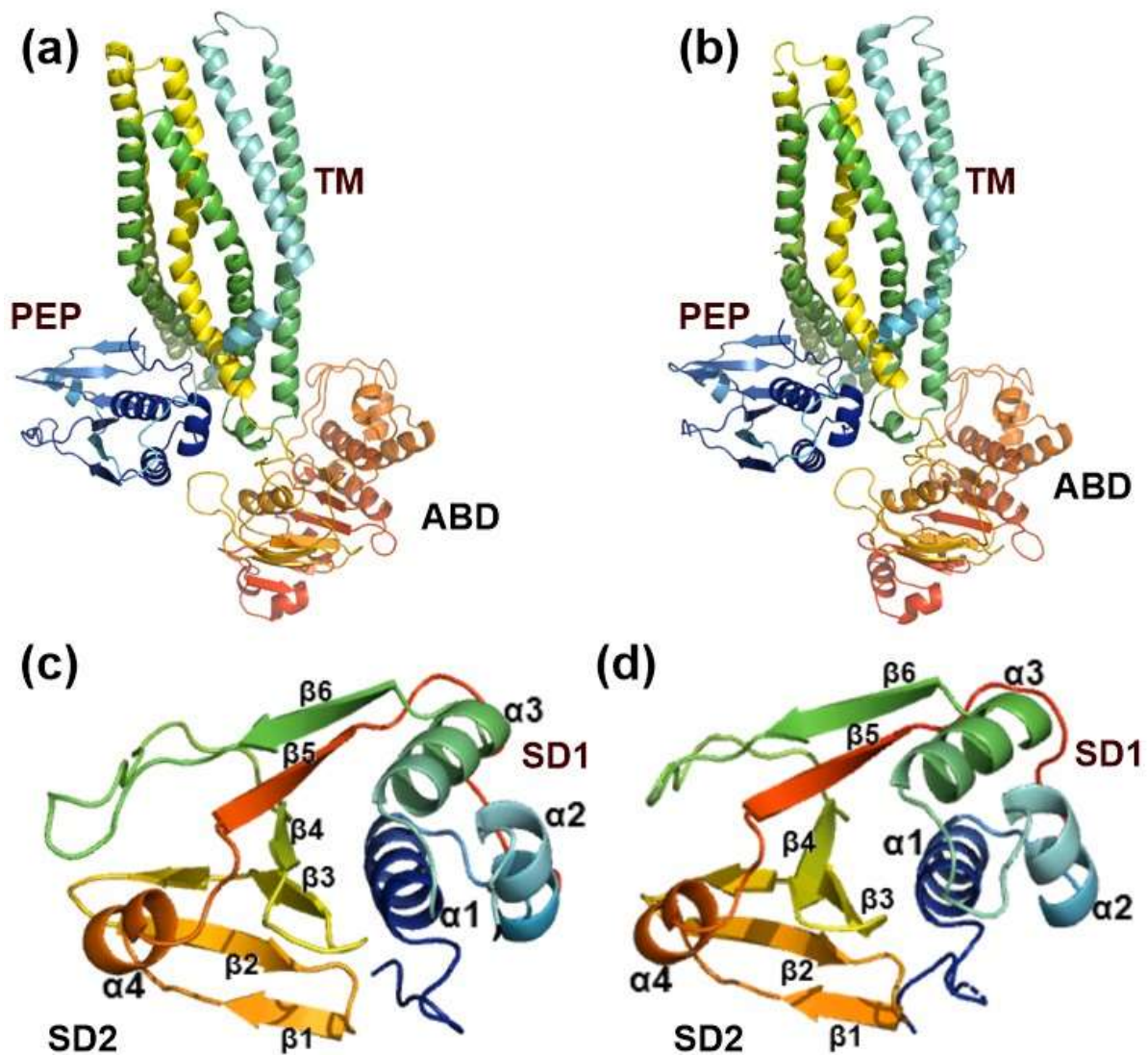
The conserved tyrosine residue in class I cyclases (Figure 4.1) has been shown to be obligatorily implicated in subtilin biosynthesis, serving as an acid/base catalyst (Helfrich et al., 2007). Interestingly, histidine seems to play this role in class II cyclases. Additionally, the conserved non-essential Ser residue in class I cyclases seems to be substituted for conserved Asn in the cyclase domains of LanM enzymes. The mere fact that these substitutions do not alter the putative functions at these positions (i.e. proton donor/acceptor), reflect more on the similarities between the enzymes.

#### **4.1.2 Sequence and structural similarities between RumT and other AMS proteins**

The structure of a related member of a protein family may provide vital information with regards to its function and how it may be expressed. This was the case with RumT where the crystal structure of related family members like PCAT1 from *C. thermocellum* (Lin et al., 2015) and the peptidase domain of ComA from *Streptococcus pneumoniae* (Ishii et al., 2010) have been determined. Although the peptidase domain of the class II lanthipeptide AMS protein LctT has been characterized *in vitro* (Ferguson Ihnken et al., 2008), no structures exist for this class so far. To enable a comparability assessment, RaptorX protein structure modeling tool was used to predict the monomeric structures of five class II lanthipeptide transporters namely; LctT, NukT, MrsT, RumT and BovT.

Although these proteins displayed very dissimilar amino acid sequences, the final predicted structures all six proteins showed strong structural homology. They possess features similar to the crystal structure of the AMS protein PCAT1 (Figure A2, Appendix 6.3). They contain a transmembrane domain with of six membrane-spanning  $\alpha$ -helices, an ATP-binding domain and a peptidase domain (Figures 4.2a & 4.2b). The crystal structures of the peptidase domains of ComA (PDB: 3K8U) and PCAT1 (PDB: 4RY2) show that the peptidase domain has two subdomains namely: N-terminal subdomain and the C-terminal subdomain.

These subdomains are made up of a series of  $\alpha$ -helices and antiparallel  $\beta$ -strands arranged such that some of the strands are connected via a short  $\alpha$ -helix (Ishii et al., 2010, Lin et al., 2015), as well as a short  $\beta$ -strand. All these features were also visible in the predicted LanT peptidase domains (Figures 4.2c & 4.2d). The active site of the predicted peptidase domain of RumT was analyzed in comparison with the crystal structure data of the peptidase domain of ComA (PDB: 3K8U). These results indicate a tight correlation between hydrogen bond distances measured for putative active site residues in RumT and those performed with the actual residues in ComA (Figure A3, Appendix 6.3).

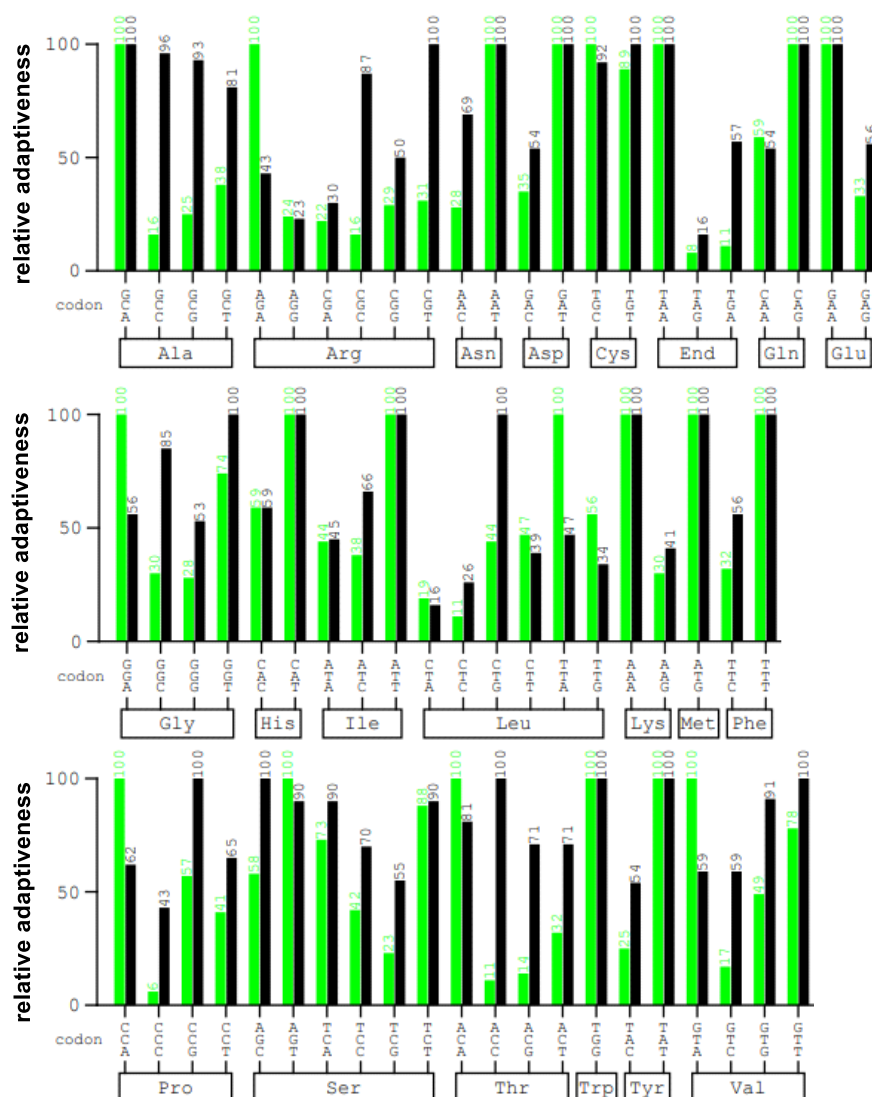


**Figure 4.2** Structure prediction of two LanT proteins: RumT and LctT using RaptorX. The predicted structures of LctT (a) and RumT (b) show strong homology to the crystal structure of the polypeptide processing and secretion transporter PCAT1 (Lin et al., 2015), possessing a transmembrane domain (TM), an ATP-binding domain (ABD) and a peptidase domain (PEP). The peptidase domains of LctT (c) and RumT (d) are also homologous to the crystal structure of the peptidase domain of ComA (Ishii et al., 2010). They have two subdomains: SD1 and SD2. SD1 consists of three alpha-helices ( $\alpha1$  -  $\alpha3$ ) while SD2 consists of 6 beta-strands ( $\beta1$  -  $\beta6$ ) and an alpha helix ( $\alpha4$ ) connecting  $\beta1$  and  $\beta5$ .  $\beta2$  and  $\beta3$  are linked by another short  $\beta$ -strand. The catalytic site is formed by residues in  $\alpha1$  and  $\beta3$ . [see section 3.13 for computational analysis methods used]

#### 4.1.3 Codon adaptability calculations and analysis

Organisms differ in the way they select and apply the triplet codon during the synthesis of a nascent polypeptide chain. How this codon usage bias comes about is a subject under strong debate within the scientific community. However, some suggestions with regards to their role

in temporal protein regulation have been proposed (Shin et al., 2015). Codon bias may have a multiplicative effect on the efficiency of protein synthesis as well as misincorporation which may occur during heterologous protein expression (Kurland and Gallant, 1996). Codon usage analysis comparing *E. coli* and *R. gnavus* was performed using the graphical codon usage analyzer (Fuhrmann et al., 2004). A graphical presentation is shown in Figure 4.3.



**Figure 4.3** Comparison of codon usage between *E. coli* and *R. gnavus*. Green bars represent the relative adaptiveness or the usage of each codon in *R. gnavus* against the black bars which represent codon usage in *E. coli*. For each amino acid, the codon that is used most frequently by the designated organism has the highest relative adaptiveness (100 %) and used as a reference point to scale the other codons for the same amino acid.

The mean difference in codon usage between *R. gnavus* with that of *E. coli* is 25 %. All codons used by *R. gnavus* (green bars) with high frequencies have moderate frequencies in *E. coli* (black bars), and so genes from *R. gnavus* source can be heterologous expressed in *E. coli*.

However, it would be extremely challenging to express genes from *E. coli* source in *R. gnavus*, as many codons like ACC (Thr), ACG (Thr), CGC (Arg) and GCC (Ala) that are frequently utilized by *E. coli* are rare in *R. gnavus*.

Using *E. coli* as a surrogate host may pose some problems if codons like AGA (Arg), AGG (Arg), CGA (Arg), ATA (Ile), CTA (Leu), GGA (Gly), CGG (Arg) and CCC (Pro) are present in the target gene (Nakamura et al., 2000). However, except for AGA (Arg) which has low frequency of usage in *E. coli* compared to *R. gnavus* (100 % → 43 %), the usage preferences for all the other codons which are rarely used in *E. coli* do not vary very much in *R. gnavus*, indicating that expression of those genes may be relatively tolerable in *E. coli*.

**Table 3.1** Properties of the RumM and RumT encoding gene sequences

Rare codon	Amino acid	Frequency of Occurrence	
		<i>rumM</i>	<i>rumT</i>
<b>AGG</b>	Arg	2	6
<b>AGA</b>	Arg	18	10
<b>CGG</b>	Arg	3	3
<b>CGA</b>	Arg	6	2
<b>GGA</b>	Gly	26	23
<b>AUA</b>	Ile	29	19
<b>CUA</b>	Leu	5	13
<b>CCC</b>	Pro	0	0
<b>Total #codons</b>	—	915	689
<b>CAI</b>	—	<b>0.66</b>	<b>0.61</b>
<b>GC content</b>	—	29.9	32.29

We analyzed RumM and RumT coding sequences using GenScript rare codon analysis tool and found out that there are a variety of these rare codons which may cause expression problems in *E. coli* (Table 3.1). We also calculated the codon adaptation index (CAI) of *rumM* and *rumT*. CAI is the relative adaptiveness of a foreign gene to the biosynthesis machinery of the host organism usually reported as a value between 0 and 1. Higher values within this range indicate that there is similar codon usage pattern in the gene of interest as there are in the reference gene (Sharp and Li, 1987). The lower the CAI, the higher the chances for poor expression (Wu et al., 2005). Although the CAIs calculated for *rumM* and *rumT* (Table 3.1) were lower than the expected value to guarantee good expression in *E. coli* (>0.8), studies indicate that the CAI is not a strong determinant of heterologous gene expression (Welch et

al., 2009). Furthermore, few tandem rare codons were found in *rumM* (Figure A4, Appendix 6.3) that may negatively influence heterologous expression of the gene (Kim and Lee, 2006). Additionally, several tandem and even a quadruple rare codon clusters were found in *rumT* that can significantly slow down translation and cause expression problems (Clarke IV and Clark, 2008). For this reason, a codon-optimized gene sequence encoding the first 125 amino acid residues of RumT (Appendix 6.4) was purchased for *E. coli* expression.

## 4.2 Vector construction and *E. coli* expression

### 4.2.1 Features of expression vectors

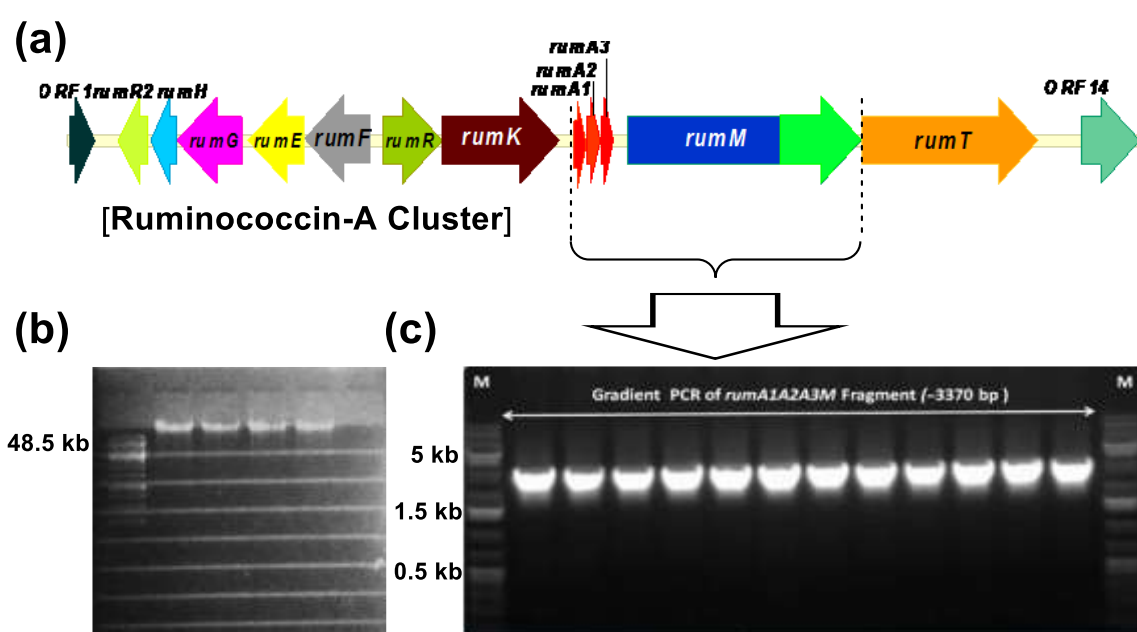
The biosynthesis gene cluster of RumA is huge (12.8 kb) and this may limit overexpression if the whole cluster is isolated and introduced into the heterologous host. In order to facilitate product optimization, we constructed single-vector and two-vector systems harbouring genes that encode the structural peptide and the PTM enzyme, as well as regulatory features that are required for expression in *E. coli*. Studies on lichenicidin which is also a class II lanthipeptide identified genes in its biosynthetic cluster which are not directly involved in the biosynthesis and modification of the component peptides (Caetano et al., 2011). Following this example, and considering data reported by Gomez et al. (2002), we proposed in this study that the hypothetical genes for immunity (*rumFEGHR2*), transport/processing (*rumTX*) and regulatory operon (*rumRK*) are accessory components and not directly involved in the biosynthesis and PTMs formation in preRumA, concluding that *rumA* and *rumM* genes are the two essential elements required.

To maintain two plasmids in the same expression host, the co-transformants were constructed to possess compatible origins of replication and distinct antibiotic resistance genes. Conventional restriction and ligation cloning was employed to facilitate insertion of the gene encoding preRumA and the lanthionine-generating enzyme RumM under control of variable promoters. Two of the most prominent promoters used here were the Isopropyl  $\beta$ -D-1-thiogalactopyranoside (IPTG)-inducible *CTU* and *CU* promoters (Kraft et al., 2007). The promoters are described in detail in section 3.5.8. Others like the *pBAD* and the *T7*-RNA polymerase promoter were also applied. Compatible pair of *oris* used for the two-plasmid systems were either pBR322/RSF or pBR322/p15A. At first, our strategy involved placing *rumM* under control of the relatively weak *CU* promoter and the *rumA* construct under the relatively stronger promoter *CTU*. The plasmid maps of all resultant expression vectors are



displayed in Appendix 6.2 and details on the construction procedures can be found in section 3.8.10.

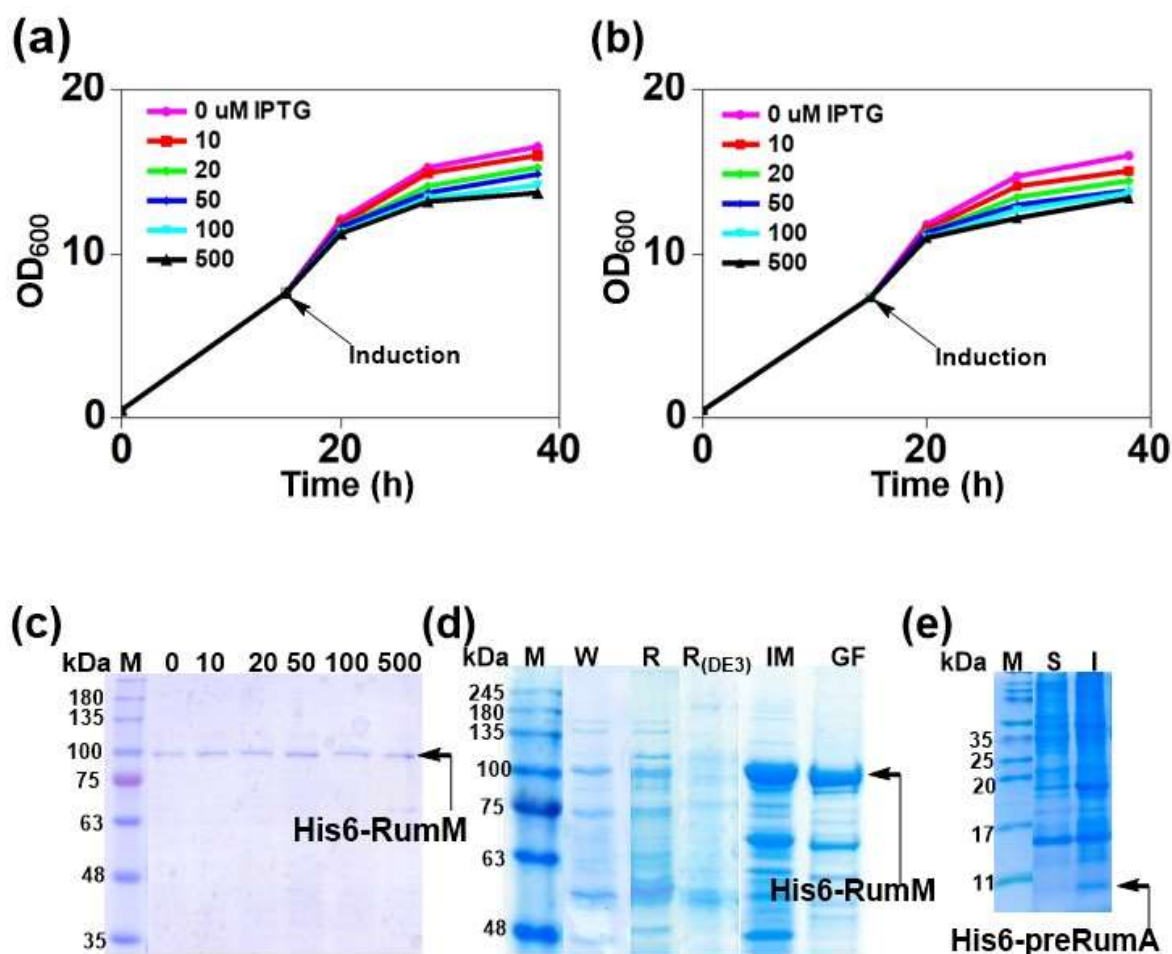
The upstream region of the backbone vectors hosting the *CU* and *CTU* promoters contained the transcriptional terminator tHP and mutations in the catabolite activator protein (CAP) site that weaken CAP binding and reduce transcriptional activation prior to induction (Krebber et al., 1996, Šiurkus et al., 2010). Even with this tight arrangement, leaky expression was still visible when traces of lactose were present in the cultivation medium, as will be seen later for complex media which have yeast extract as one of their major constituents.



**Figure 4.4 Isolation of rA1A2A3M fragment from Ruma biosynthesis gene cluster.** (a) Ruma gene cluster illustrating gradient PCR amplification of Ruma and RumM encoding genes from purified genomic DNA of *R. gnavus* E1. E1 genomic DNA was analyzed on a 0.5 % agarose gel (b) while gradient PCR amplified rA1A2A3M fragment was analyzed on a 1 % agarose gel (c). [see method section 3.5 for molecular biology techniques]

Genomic DNA purified from *R. gnavus* E1 was used as a template to amplify a segment of the *rumA* gene cluster via gradient PCR using the annealing temperature settings of  $55^{\circ}\text{C} \pm 7^{\circ}\text{C}$ . Results showed that the chromosomal DNA purification from *R. gnavus* E1 was successful (Figure 4.4b). The target rA1A2A3M fragment (3370 bp) was also successfully amplified and the correct size was identified on agarose gel (Figure 4.4c). The resulting PCR amplicon was inserted into pCTUT7 to yield pLEOrC2 (section 3.5.10.1).





**Figure 4.5** Growth and protein expression. The quasi-linear growth profiles of WLEOrA (a) WLEOrM' (b) in EnPresso B cultivation medium. (c) SDS-PAGE analysis of HisTrap spin column-purified extracts from WLEOrM' strains induced with different IPTG concentrations (indicated at the top of each well). (d) Comparing His6-RumM expression in extracts from *E. coli* W3110 (W), Rosetta 2 (R) and Rosetta DE3 (R<sub>DE3</sub>). IMAC (IM) and gel filtration (GF) purification of His6-RumM from WLEOrM'. (e) Soluble (S) and insoluble (I) fractions of His6-preRumA extracted from WLEOrA. His6-RumM was analyzed using 10 % standard SDS-PAGE while His6-preRumA was analyzed with 16 % tricine SDS-PAGE gel. Protein bands were visualized using the colloidal blue silver Coomassie G-250 as reported elsewhere (Candiano et al., 2004) [see section 3.6 for protein analysis methods].

#### 4.2.2 Separate expression of His6-preRumA and His6-RumM

Following procedures described in section 3.5.11, *E. coli* W3110 was separately transformed with pLEOrA (expressing His6-preRumA under control of *CTU* promoter) and pLEOrM' (expressing His6-RumM under *CTU* promoter), yielding strains WLEOrA and WLEOrM'. In a similar procedure, *E. coli* Rosetta 2 (DE3) was separately transformed with pLEOrM (expressing His6-RumM under *CU* promoter) and pRSF-*rM* (expressing His6-RumM under T7 promoter), respectively. The strains were cultivated in 24 deep-well plates, using the automated *in situ* glucose delivery EnPresso B growth medium system (section 3.6.1). Varying

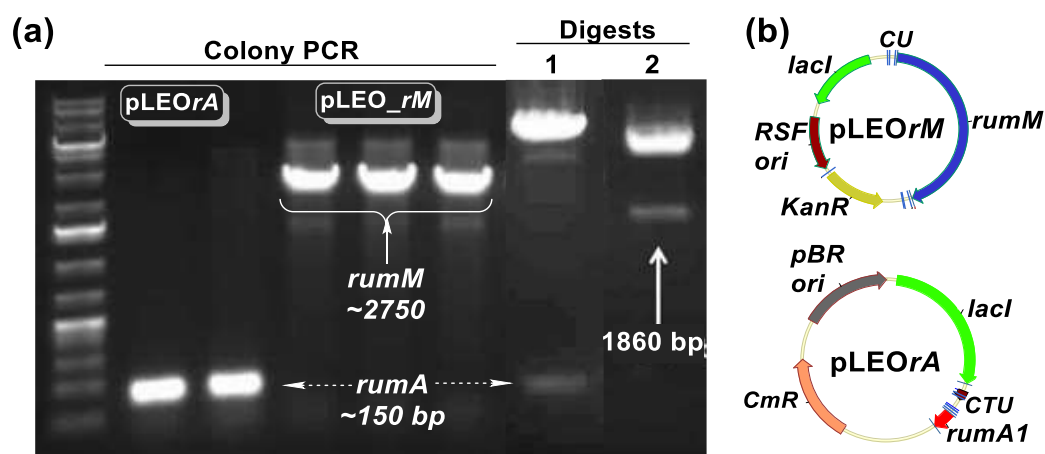
concentrations of IPTG ranging from 10 to 500  $\mu\text{M}$  were supplied to the cultures. Cells were further cultivated for 18 to 24 h.  $\text{OD}_{600}$  of WLEOrA and WLEOrM' strains were measured and plotted as shown in Figures 4.5a and 4.5b respectively. Endpoint samples were obtained and processed using immobilized metal ion affinity chromatography (IMAC) as described in the experimental part (section 3.7.1).

Although the cultures all followed a quasi-linear growth profile as expected, IPTG concentration only slightly decreased the growth rate but did not produce any significant effect on the expression levels of His6-RumM (Figure 4.5c). We investigated if producing RumM in a system that supplies the tRNAs that encode the rare codons (e.g Rosetta) could improve its yield. Results in Figure 4.5d demonstrates an even poorer expression in the Rosetta (DE3) system. However, in Rosetta 2, the expression level was comparable to W3110 that was originally used. We therefore decided to continue with the wild-type *E. coli* W3110 strain throughout the study and only switched when the need was critical to establish some facts. We further cultivated the WLEOrM' strain and purified His6-RumM via IMAC and size exclusion chromatography (section 3.7.2). Approximately  $0.75 \text{ mg l}^{-1}$  of purified protein was obtained after gel filtration (Figure 4.5d). His6-preRumA expressed in WLEOrA was largely present in the insoluble fraction (Figure 4.5e), suggesting that the his-tagged precursor peptide must have been expressed in inclusion bodies as reported for other lantipeptide precursors expressed in the absence of the PTMs enzyme (Li et al., 2009, Li et al., 2010).

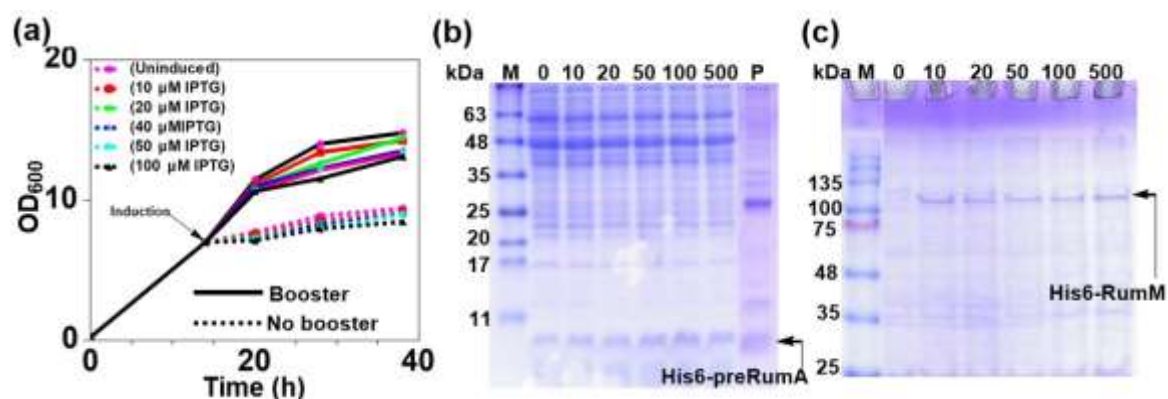
#### 4.2.3 Two-plasmid coexpression system for His6-preRumA and His6-RumM

The RumM encoding gene was cloned in a vector with a weaker promoter compared to the vector that hosted the preRumA encoding gene (see sections 3.5.10.2 and 3.5.10.6 for details on how individual His6-preRumA and His6-RumM vectors were constructed). Results from colony PCR and control digests (Figure 4.6a), together with sequencing indicated that the genes of interest were successfully cloned in-frame into their respective vectors. The arrangement of various expression-relevant features on a plasmid are shown in Figure 4.6b. The concepts that guided the expression vectors design was adopted to limit physiological and metabolic stress on the host organism and enable high-quality expression yield. On the one hand, RumM is relatively large (914 aa) and thus require more cellular resources for its biosynthesis and a slower process from transcription to translation for the protein to be properly folded and soluble. Strong overexpression of RumM may result in truncated products and/or aggregation which would be wasteful. On the other hand, strong overexpression of the structural peptide may have little physiological effect on the host since it is relatively small (47 aa). As such, this strategy may help to conceal the effects of cellular proteases on the small

peptide thereby increasing cellular availability of the unmodified substrate and the chances for RumM to install PTMs on preRumA.



**Figure 4.6** Cloning of preRumA and RumM for coexpression in *E. coli*. (a) Colony PCR and control digest analyses of pLEOrA (1) and pLEOrM (2). (b) A sketch of the plasmid maps indicating the main features.



**Figure 4.7** Growth and protein expression test. Varying concentrations IPTG (indicated in the graph and at the top of each well of the SDS-PAGE) were used to induce protein expression. The graphs in (a) illustrates the growth of WLEOrA/M strain in EnPresso B medium system under the different IPTG concentrations. (b) Soluble WLEOrA/M lysate extracted from cultures grown under varied IPTG concentrations, showing expression of His6-preRumA. The letter P, IMAC-purified His6-preRumA from WLEOrA/M. (c) Soluble His6-RumM purified from the same cultures. His6-RumM was analyzed using 10 % standard SDS-PAGE while His6-preRumA was analyzed with 16 % tricine SDS-PAGE gel. Protein bands were visualized using the blue silver staining protocol. [see section 3.6 for protein analysis methods]

*E. coli* W3110 was cotransformed with the expression vectors pLEOrA and pLEOrM, resulting in the expression strain named WLEOrA/M. To optimize the performance of strain WLEOrA/M in producing both His6-preRumA and His6-RumM, the strain was cultivated in EnPresso B and processed as described in the previous section. Additionally, prior to induction, the booster

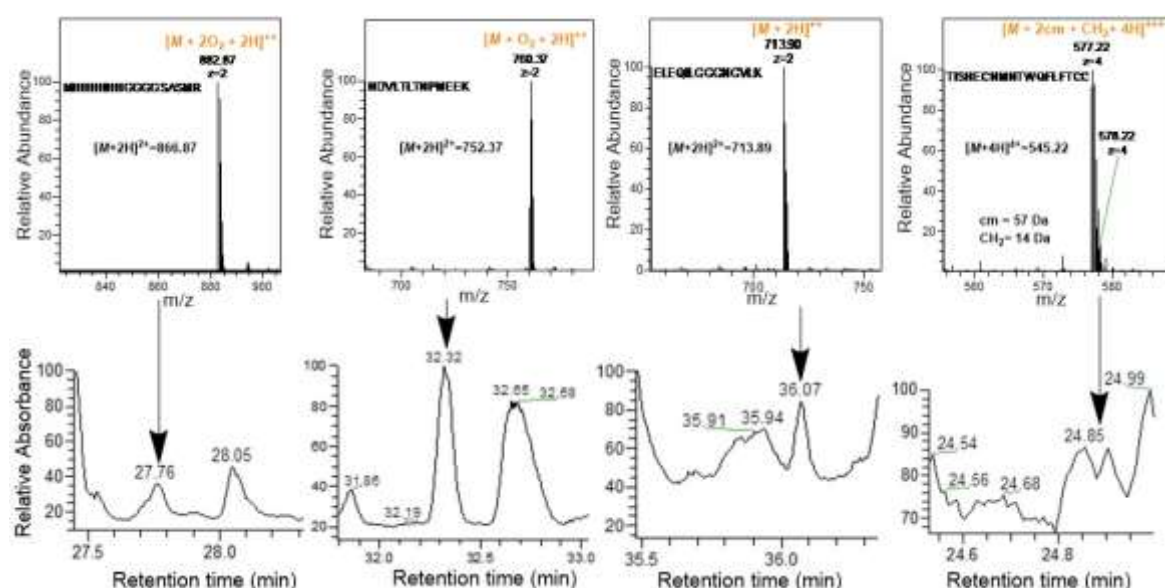
tablet (a complementary component of the growth system) was not added to one half of the overnight EnPresso culture. The other half was boosted, and the cultivation was continued as described in the experimental part (section 3.6.1). Growth of cultures proceeded as expected. Due to limited growth resources, the non-boosted cultures grew very slow and produced end-point ODs that were approximately 1.5 times less than their boosted counterparts (Figure 4.7a). Consequently, the booster tablets were used for all subsequent studies performed with the EnPresso B growth system.

IPTG influenced growth and His6-preRumA expression (Figure 4.7b) but did not show any significant effect on the expression levels of His6-RumM (Figure 4.7c). Results indicate that PTMs improve the solubility of the precursor peptides (Shi et al., 2011). This may explain why soluble His6-preRumA was present in extracts from WLEOrA/M and not in WLEOrA. Although we were able to purify His6-preRumA from WLEOrA/M via IMAC (Figure 4.7b), further purification steps proved to be challenging as the IMAC samples precipitated shortly after elution. Several different buffer conditions were tested in an effort to prevent the formation of the precipitate or to resolubilize them without success. Gels in Figures 4.7b and 4.7c show that soluble His6-RumM and His6-preRumA were indeed expressed in *E. coli* W3110. Furthermore, the SDS-PAGE bands of IMAC-purified His6-preRumA (lane P, Figure 4.7b) and gel filtration-purified His6-RumM (lane GF, Figure 4.5d, section 4.2.2) were excised and subjected to in-gel tryptic digestion preceded by an iodoacetic acid derivatization step for identification via MS. Note should be taken here that the set-up procedures and reaction conditions of the iodoacetic acid reaction were the same as those of the iodoacetamide (section 3.9) since both chemicals share similar physicochemical properties.

#### 4.2.4 MS identification of products from the two-plasmid system

The extracted gel samples from the previous section were desalted and concentrated using ZipTip (section 3.10.1). 5 µl was applied to an Orbitrap-coupled HPLC system for LC-MS/MS<sup>2</sup> analysis (section 3.10.3). Figure 4.8 shows MS peaks of the four tryptic peptides that were produced following complete digestion of the His6-preRumA. Modifications like S-carboxymethyl-cysteine and oxidation of methionine were apparent. The measured masses of all the tryptic peptides were consistent with the calculated exact masses. Worthy to note is the fact that if preRumA from WLEOrA/M was modified by RumM, carboxymethylation of cysteines would be blocked due to the presence of the thioether cross-bridges formed between cysteine and Dha/Dhb residues in the core peptide. Additionally, dehydration of threonine and serine residues are accomplished as a prelude to the cyclization reaction (Arnison et al., 2013, Yang and van der Donk, 2015, Repka et al., 2017). We did not find any evidence of dehydration in

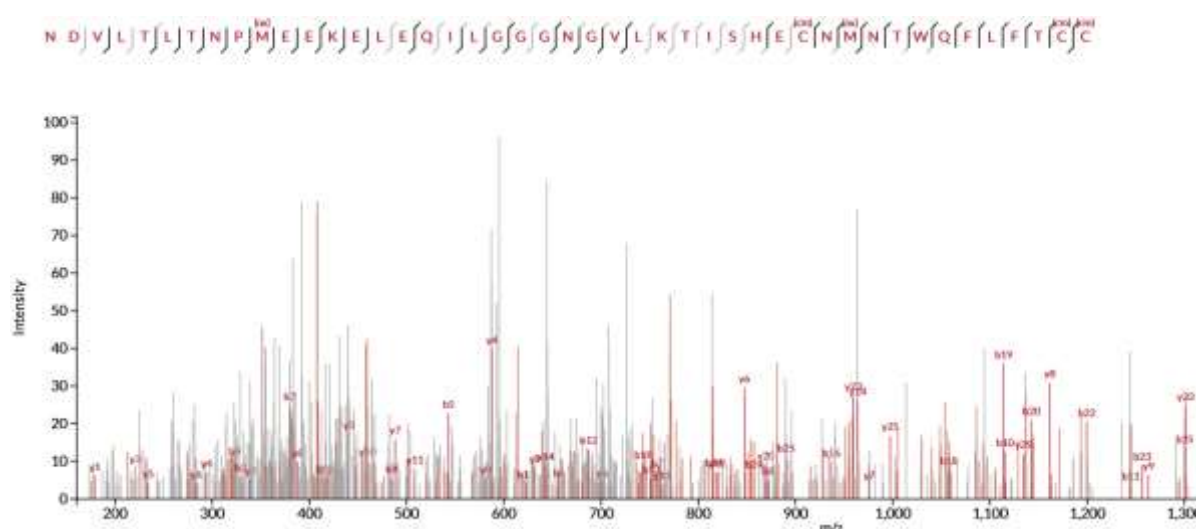
the tryptic peptides, especially in the core fragment, compelling us to conclude that there was no modification in the isolated preRumA.



**Figure 4.8 Sequence verification of His6-preRumA.** Mass spectra and corresponding retention times (indicated by arrows) of trypsin-digested His6-preRumA obtained from LC-MS analysis. The mass peaks represent the masses of tryptic His6-preRumA fragments. Modifications such as oxidation, methylation ( $CH_2$ ) and carbamidomethyl (cm)-cysteine, corresponding to mass shifts of +16, +14 and +57 Da respectively are also indicated. The calculated exact masses of each peptide sequences are labelled in their respective spectra. [see section 3.10 for mass spectrometric analysis methods used]

Although our results in the previous section suggested that RumM may have facilitated solubility of His6-preRumA in the WLEOrA/M strain, MS<sup>2</sup> data obtained by fragmenting the precursor ions (Figure 4.8) did not proffer any evidence to suggest that the precursor peptide was modified in the two-vector system, even though expression of RumM was apparent. Assignments of product ions produced as a result of fragmenting the tryptic fragments of His6-preRumA are presented in Appendix 6.5. The fragmentation pattern was characteristic of a linear peptide. The peptide sequence exhibited b- and y-type ions with high intensities which nicely fitted to all amino acids from N- to C-terminus (Figure 4.9). This is unlike lanthionine-modified peptides whose ring structures cause them to resist fragmentation (Müller et al., 2010). Consequently, there was no ring structure in the preRumA measured. These results also supported the outcome of LC-MS scans. Furthermore, subsequent MASCOT search of the MS and MS<sup>2</sup> data from the orbitrap analyzer identified the preRumA peptide sequence (UniProt ID: P83674) and RumM sequence (UniProt ID: Q9L3F1).

It is challenging to say exactly why preRumA was not modified when coexpressed simultaneously with RumM although the coexpression fostered solubility of the precursor peptide. We may just be tempted to suggest that there was a simple binding interaction between His6-preRumA and His6-RumM that facilitated the solubility of the former. Our suggestion is supported by the fact that interactions between nonmodified peptide precursor and its modifying enzyme have been reported in previous studies (Lubelski et al., 2009, Khusainov et al., 2011, Mavaro et al., 2011, Repka et al., 2017). Meanwhile Nagao et al. (2005) had demonstrated that applying a two-vector coexpression system to produce modified His-tagged lantibiotic in *E. coli* is a rather laborious venture, Basi-Chipalu and coworkers applied the same procedure to modify pseudomycoicidin in *E. coli* (Basi-Chipalu et al., 2015). Their data indicate that functional expression of lanthipeptides in *E. coli* may vary from peptide to peptide, which obviously also involve several different factors in-between. Nevertheless, other methods have been used to successfully express lanthipeptides in *E. coli* (Shi et al., 2011, Shi et al., 2012, Tang and van der Donk, 2012, Kuthning et al., 2015).



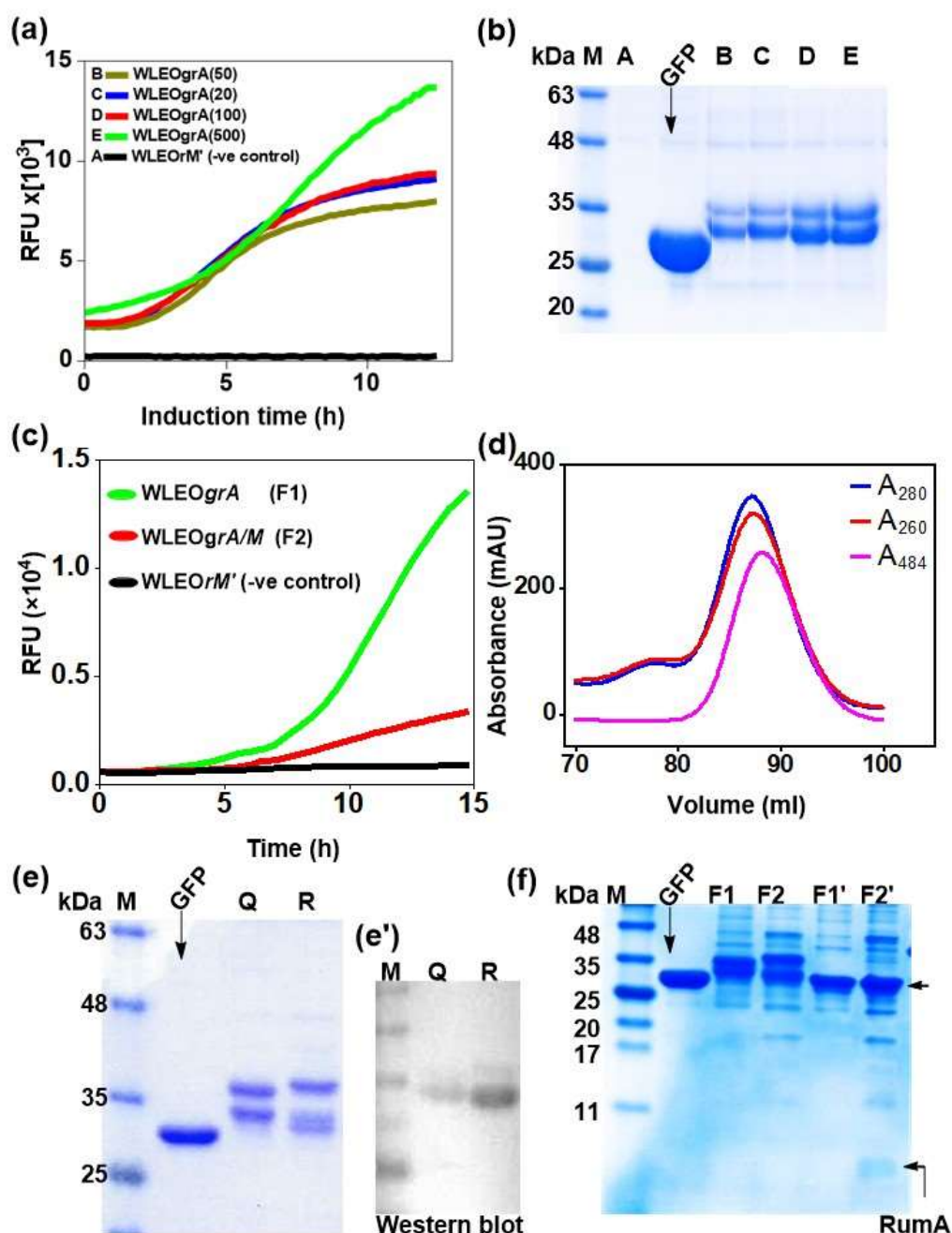
expressed precursor peptide may face intracellular biochemical challenges that may lead to aggregation and/or digestion by host proteases.

#### 4.2.5 PreRumA Fused to GFP

To overcome regular expression challenges, fusion partners like self-cleavable carriers, aggregation-promoting and solubility-enhancing conjugates are usually employed to prevent proteolytic inactivation, conceal lethal effects towards heterologous hosts and promote production of AMPs in *E. coli* (Li, 2011, Bell et al., 2013, Pane et al., 2016). Our results from the previous section, supported by earlier studies, compelled us to posit that the timing between production of the precursor peptide preRumA and subsequent modification by RumM may be crucial in determining the fate of the small peptide, since it may encounter diverse physical and biochemical challenges before the PTMs are installed. We therefore decided to provide a physical support to preRumA by fusing GFP to its N-terminus as described earlier (section 3.5.4.4). Transformation of *E. coli* W3110 with the single plasmid pLEOgrA, and cotransformation with pLEOgrA and pLEOrM (see section 3.5.10 for vectors construction) resulted in strain WLEOgrA (expressing His6-GFP-TEV-preRumA) and strain WLEOgrA/M (expressing both His6-GFP-TEV-preRumA and His6-RumM on separate plasmids). Expression of the chimeric fusion protein was monitored online in a 24-well flat-bottom plate via GFP fluorescence signal intensities (see section 3.3.3). Different concentrations of IPTG were tested and results showed an inducer concentration-dependent expression (Figures 4.10a & 4.10b).

Cultures of the strains WLEOgrA and WLEOgrA/M were induced with 100  $\mu$ M IPTG and time-course fluorescence signal intensities were measured (Figure 4.10c). In all cases, the relative fluorescent unit (RFU) of the control strain remained constant throughout the entire cultivation period. A dramatic drop in fluorescence signals of strain WLEOgrA/M was observed. This was also reflected in the amount of purified His6-GFP-TEV-preRumA obtained from this strain in comparison with WLEOgrA (data not shown). A possible reason for this may be the metabolic burden imposed by the presence of two distinct plasmids. The large modifying enzyme RumM may also utilize a substantial portion of the available resources in the combinatorial system and hence reduced the expression levels of His6-GFP-TEV-preRumA.





**Figure 4.10** Expression of GFP-RumA fusion constructs. (a) GFP fluorescence signals of WLEOgrA cultures induced with varying IPTG concentrations (b) SDS-PAGE analyses of His6-GFP-TEV-preRumA purified from respective WLEOgrA cultures. (c) Comparison between the GFP fluorescence signals of WLEOgrA and WLEOgrA/M strains, cultivated under the same growth conditions. WLEOrM' was used as a control. (d) UV signals showing Gel filtration purification of His6-GFP-TEV-preRumA from WLEOgrA/M lysate. (e) SDS-PAGE analyses of gel filtration-purified His6-GFP-TEV-preRumA from WLEOgrA lysate (lane Q) and WLEOgrA/M lysate (lane R). (e') Western blot analyses of gel filtration-purified His6-GFP-TEV-preRumA from WLEOgrA (lane Q) and WLEOgrA/M (lane R) using anti-Histag antibody and biochemical chromogenic detection method. (f) SDS-PAGE analyses of IMAC-purified GFP, His6-GFP-TEV-preRumA from WLEOgrA lysate (lane F1), WLEOgrA/M lysate (lane F2), TEV-digested His6-GFP-TEV-preRumA from WLEOgrA (lane F1') & TEV-digested His6-GFP-TEV-preRumA from WLEOgrA/M (lane F2') [see section 3.6 for protein analysis methods].



For larger expression experiments, *WLEOgrA* and *WLEOgrA/M* were cultivated in TB medium as described in the experimental part (section 3.6.2) and purified using IMAC and gel filtration. Figure 4.10d shows a gel filtration chromatogram representing the elution of His6-GFP-TEV-preRumA purified from *WLEOgrA/M* lysate. We obtained approx. 100 mg of total protein per litre of *WLEOgrA* culture and approx. 15 mg l<sup>-1</sup> from *WLEOgrA/M* after IMAC and gel filtration purifications. Results demonstrated dissimilar migration properties of products from the two systems on SDS-PAGE (Figure 4.10e), clearly indicating that His6-RumM has an effect on the His6-GFP-TEV-preRumA.

Purified His6-GFP-TEV-RumA was dialyzed in TEV digestion buffer and fresh homemade TEV from -20 °C (see Appendix 6.10 for purification results) was added to the sample as described in section 3.7.3. The digested products together with the non-digested samples were analyzed on SDS-PAGE (Figure 4.10f). Interestingly, the expected cleaved preRumA product band was not observed in digested His6-GFP-TEV-preRumA purified from *WLEOgrA* as opposed to that obtained from *WLEOgrA/M* (Figure 4.10f). One reason for this may be degradation by host proteases carried over from the IMAC purification since cleavage may have exposed the unmodified peptide to such challenging environment. If the peptide is modified, protease degradation would be unlikely since the presence of PTMs would confer stability onto it (Rink et al., 2010). However, additional experiments with protease inhibitors did not show any noticeable improvements (data not shown).

The unequal mobility of His6-GFP-TEV-preRumA obtained from the two systems on SDS-PAGE and the absence of a preRumA band in the TEV-digested sample purified from the system expressing His6-GFP-TEV-preRumA alone supplied the first evidence for the activity of His6-RumM on the GFP fusion construct. It was however not possible at that time to determine if this was just a simple binding activity or if RumM was actually introducing the expected PTMs on preRumA. Additionally, multiple bands were visible in the SDS-PAGE which both were further confirmed via Western blot analysis (section 3.6.7) to belong to the His6-GFP-TEV-preRumA (Figure 4.10e'). This may have been produced via host protease degradation or incomplete translation during expression.

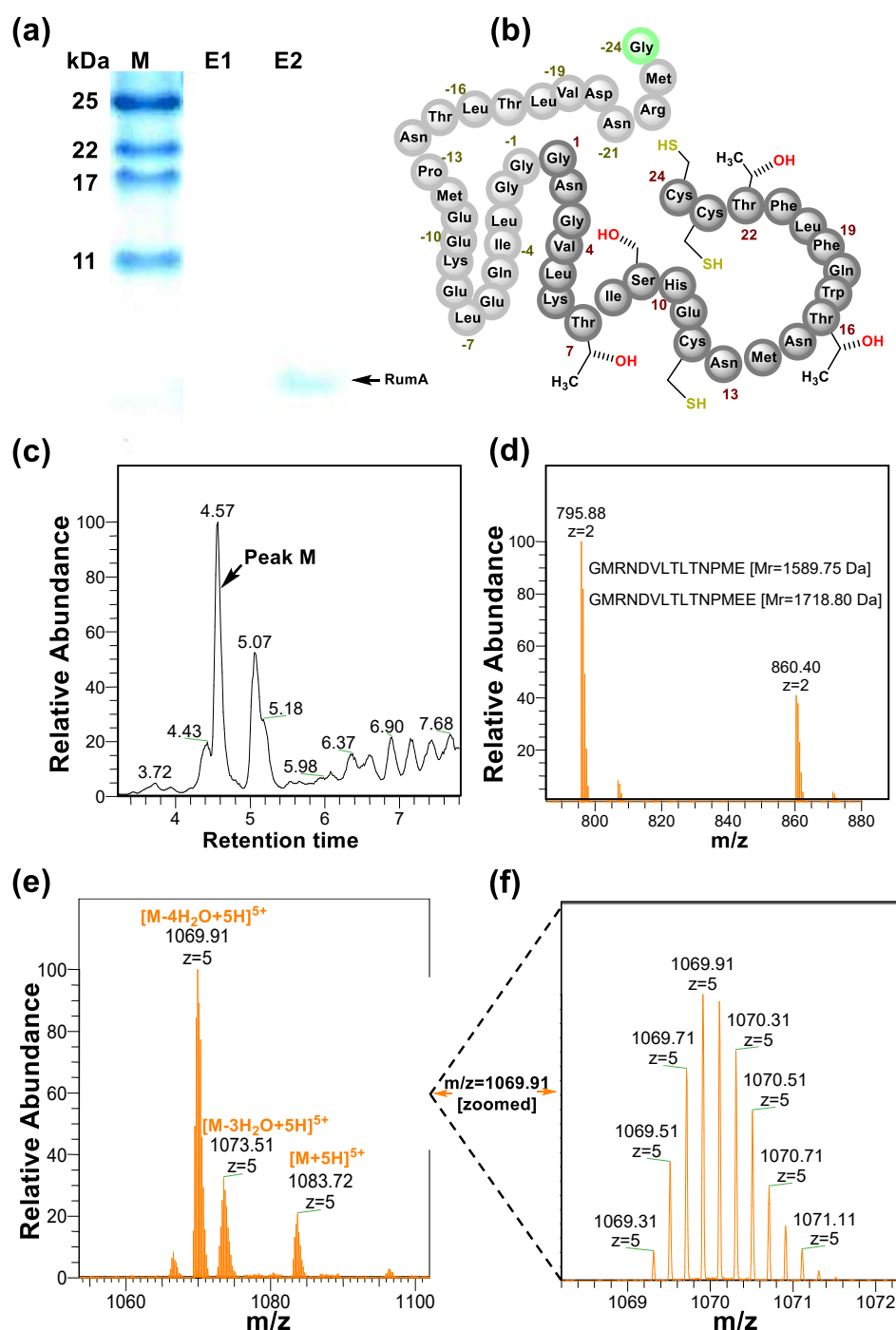
To evaluate the real nature of the issue, all four bands in Figure 4.10e (lanes Q and R) were excised and processed for identification via MS. They all showed completely different amino acid sequence structures with both bands of lane Q and the lower band of lane R showing evidences of truncated His6-GFP-TEV-preRumA. Only the upper band in lane R contained full-length His6-GFP-TEV-preRumA (data not shown). We inserted tandem stop codons into the preRumA encoding gene and a downstream  $\lambda$ t0 terminator (section 3.5.10.4) to ensure that

these two bands were not the result of read-through translation (Ryoji et al., 1983). Since neither of these procedures resolved the issue, we concluded that it was not a read-through translation problem, narrowing down the possibilities to degradation only; where the C-terminal protease-prone precursor peptide were readily digested by host proteases.

#### 4.2.6 LC-ESI-MS analyses of preRumA from WLEOgrA and WLEOgrA/M

TEV-digested His6-GFP-TEV-preRumA purified from WLEOgrA and WLEOgrA/M were extracted with butanol (section 3.7.4). The extracts were resolved by SDS-PAGE and results showed no band for extracts from WLEOgrA and again, a single band for extracts from WLEOgrA/M (Figure 4.11a, lanes E1 and E2 respectively). The E1 and E2 extracts were dried and re-dissolved in ACN/H<sub>2</sub>O/formic acid (50:50:0.2%). LC-ESI-MS analysis of E1 did not produce any charged ion corresponding to unmodified preRumA (schematically represented in Figure 4.11b), which further corroborated results obtained from SDS-PAGE. Major peaks resolved in the reversed-phase chromatogram (Figure 4.11c) was analyzed by ESI-MS and found to contain charged ion corresponding to truncated preRumA fragments (Figure 4.11d), indicating that the peptide may have undergone host protease degradation as suggested in the previous section.

Interestingly, analysis of extract E2 yielded a mixture of fourfold dehydrated, triple-dehydrated and non-dehydrated preRumA (Figure 4.11e). The latter observation suggests that RumM may have conferred some form of stabilizing role on the precursor peptide since we could not identify non-dehydrated preRumA in extracts purified from the strain expressing His6-GFP-TEV-preRumA alone. This results further supported our suggestion that biophysical interactions between His6-preRumA and His6-RumM may have facilitated the solubility of the former without necessarily introducing the desired modifications in the peptide (see section 4.2.4). Nevertheless, mixtures of partially dehydrated products are possible since investigations revealed that the coupling of cysteinyl thiol to Dha/Dhb via Michael addition cyclization can prevent further modification of serine and/or threonine residues (Kuipers et al., 2008, Lubelski et al., 2009), although the formation of Dha and Dhb are independent of one another. The peak corresponding to fourfold dehydrated preRumA was further zoomed to reveal its isotopic distribution (Figure 4.11f). The measured exact masses showed very tight consistencies with the calculated exact masses with an error margin of  $\pm 0.05$  Da (see Table 3.3 for the accuracy of measurements expressed in parts per million [ppm]). These results supplied enough evidence to demonstrate that biosynthesis and modification of RumA is achievable in *E. coli*.



**Figure 4. 11** Butanol extraction and LC-ESI-MS. (a) Butanol extracts of TEV-digested His6-GFP-TEV-preRumA purified from WLEOgrA (lane E1) and WLEOgrA/M (lane E2). (b) The primary structure of preRumA showing threonine, serine and cysteine side chains that are targeted for modifications by RumM, as well as Gly-24 which is added to the N-terminus of the peptide following TEV cleavage. (c) Reverse-phase chromatogram indicating a major elution peak (peak M). (d) Mass spectrum showing charged ions mass peaks obtained from MS scan of peak M. (e) ESI mass spectrum indicating a mixture of mass peaks associated with the loss of four H<sub>2</sub>O molecules (5344.55 Da), three H<sub>2</sub>O molecules (5362.55 Da) and non-dehydrated preRumA (5413.60 Da). The calculated exact masses of the fourfold, threefold and non-dehydrated preRumA were 5344.54, 5362.50 and 5413.55 respectively. (f) Isotopic distribution of fourfold dehydrated preRumA peak.

In this section, we showed that fusing preRumA to GFP linked via a TEV cleavage site and co-expressing the chimeric construct simultaneously with the RumA lanthionine synthetase RumM on separate plasmids resulted in *in vivo* modifications of the preRumA core peptide. Non-PTMs bacteriocins like enterocin P and enterocin A have been produced in *E. coli* using intein chitin-binding domain as fusion partner (Ingham et al., 2005, Klocke et al., 2005). Other fusion tags like GST, SUMO, MBP and TRX have been successfully used to express AMPs in *E. coli* (Li et al., 2009, Li, 2011, Bell et al., 2013, Pane et al., 2016). Although GST falls within the same size range as GFP, application of the latter as a fusion partner in heterologous AMP production is rather unpopular. Even MBP which has a larger molecular weight compared to both GST and GFP, has been applied for peptide production in *E. coli* (Li and Leong, 2011, Tay et al., 2011). However, reports where GFP has been utilized as a fusion partner for AMP production are rare in the literature. Additionally, no study has reported the fusion of larger protein carriers to the N-terminus of a lanthipeptide precursor to achieve active *E. coli* production. This study therefore constitutes the first report of its kind for lanthipeptide. Since the modifying enzyme requires the leader peptide as a docking site to direct its activity on the core structure, results here also demonstrate that a larger attachment to the N-terminus of the leader peptide does not obstruct *in vivo* processivity of LanM enzymes in interacting and modifying the core peptide.

To the best of our knowledge, this is the first time that a lanthipeptide from a strictly anaerobic Gram-positive microbe is heterologously expressed in a Gram-negative host. Our approach here differed from earlier studies where either the whole biosynthesis cluster was isolated and inserted into a plasmid and subsequently expressed in *E. coli* (Caetano et al., 2011), or where the genes for the structural peptide and the modifying enzymes were constructed in single polycistronic plasmids (Shi et al., 2011, Shi et al., 2012, Tang and van der Donk, 2012, Kuthning et al., 2015).

Modification was not achieved in the combinatorial system expressing both His6-preRumA and His6-RumM. In contrast, a similar system expressing His6-GFP-TEV-preRumA and His6-RumM, modification was successful. Since the fusion partner used here exhibits high expression and solubility in *E. coli*, we deduced that GFP offered supportive role to the precursor peptide allowing enough time for RumM to act. It was also interesting to observe that non-dehydrated preRumA was identified in extracts obtained from strain *WLEOgrA/M*, but not in extracts obtained from *WLEOgrA*. This observation prompted us to further suggest that the presence of RumM *in vivo* fostered a cooperative molecular interactions with both the leader sequence and the core peptide or preRumA, preventing it from degradation by host protease.

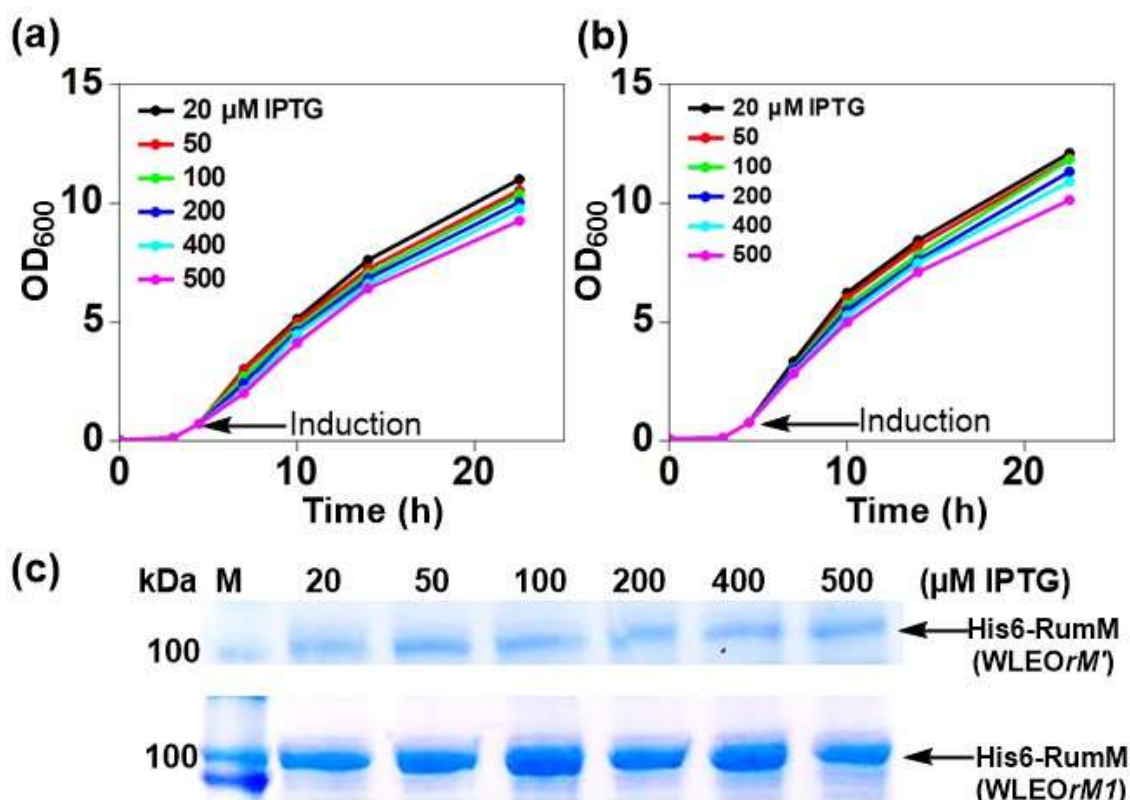
Another reason for also noticing higher degradation of His6-GFP-TEV-preRumA in the combinatorial system may be linked to the amount of His6-RumM enzyme produced by the weakly expressing vector. The ratio of His6-RumM to His6-GFP-TEV-preRumA expression was too low which may have not been optimal for complete modification of the precursor, thus, the reason why we also observed a mixture of fully dehydrated and intermediate products. However, such drawback may be averted by reengineering the host vector to increase RumM expression.

## 4.3 Quality enhancement, characterization & activation of preRumA

### 4.3.1 Optimized vector for His6-RumM expression

The chimeric His6-GFP-TEV-preRumA encountered some expression challenges as reported in the previous section. We decided to enhance *rumM* expression by optimizing the host vector. To do this, we designed an expression vector named pCUT7 that hosted the *CU* promoter and the ribosomal binding site (RBS) of gene 10 of bacteriophage T7 (section 3.5.8). This vector was designed to replace the *CTU* promoter in pCTUT7 vector with *CU* promoter because *rumM* expression under the *CTU* promoter which is considered the strongest in that series (Kraft et al., 2007) was inefficient (section 4.2.2). *rumM* PRC amplicons were digested and inserted into pCUT7 vector to yield pLEOrM1 (see section 3.5.10.6).

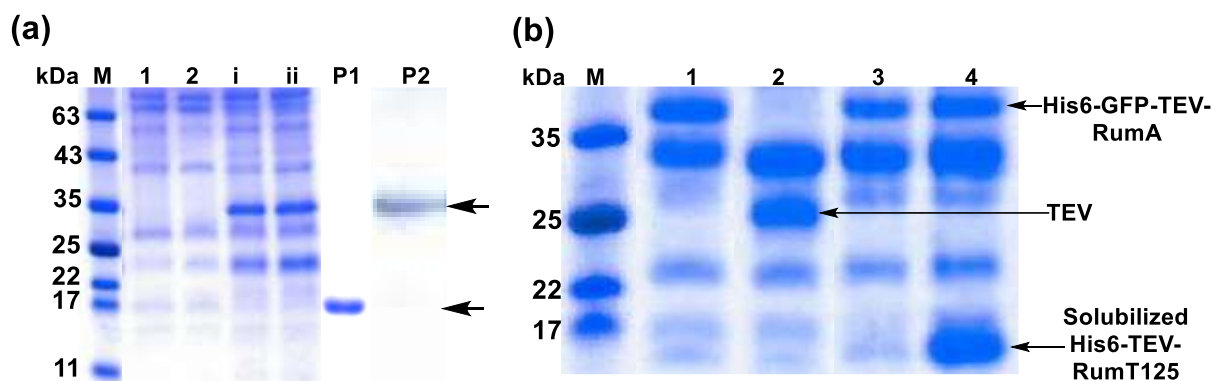
Transformation of *E. coli* W3110 with pLEOrM1 resulted in strain WLEOrM1. WLEOrM1 [expressing His6-RumM under control of the *CU*\_(+T7 RBS) promoter] and WLEOrM' [expressing His6-RumM under control of the *CTU*\_(+T7 RBS) promoter] were grown in TB medium as described in the methods part (section 3.6.2) except for the fact that the 100-ml culture was distributed into 24-well deep-well plate at the point of induction. Varying IPTG concentrations were applied to one row of the wells and then replicated for the remaining three rows. ODs of the cultures were measured at different time points. Results in Figures 4.12a and 4.12b show that both strains shared similar growth characteristics. His6-RumM production was drastically improved in the WLEOrM1 system (Figure 4.12b). Since there was no significant influence on growth for both plasmids, we can only conclude that the *CU* promoter and T7 RBS combination favoured expression of His6-RumM. There was no evidence of aggregation in both systems expressing His6-RumM but minor signs of degradation were apparent after the IMAC purification.



**Figure 4.12** Growth and product optimization. Growth curves of WLEOrM1 (a) and WLEOrM' (b) cultures induced with varying IPTG concentrations in TB. (b) Expression of His6-RumM in pCTUT7 (WLEOrM' strain) versus pCTUT7 (WLEOrM1 strain). The concentrations of IPTG are indicated in the plots and at the top of each lane on the gel. The optimal IPTG concentration range is 100-200 μM.

### 4.3.2 Expression of RumT peptidase domain

As a general characteristic for class II lanthipeptides, the ABC-type transporter/activator protein RumT cleaves off the leader peptide for preRumA after PTMs have been installed and export the activated RumA out of the cell (Gomez et al., 2002, Chatterjee et al., 2005, Furgerson Ihnken et al., 2008, Bierbaum and Sahl, 2009, Nishie et al., 2011). Here, the nucleotide sequence encoding the first 125 amino acid residues of the N-terminus of RumT was codon-optimized for *E. coli* expression and cloned in pCTUT7 and pCTUT7-SUMO vectors to yield pLEOrT125 and pLEOrsT125 (see section 3.5.10.8 for details). The plasmids were subsequently used for transformation of different *E. coli* strains and tested for expression. His6-SUMO-TEV-RumT125 was solubly produced while the His6-TEV-RumT125 went into inclusion bodies (Figure 4.13a). We resolubilized and refolded the His6-TEV-RumT125 inclusion bodies. All attempts to cleave off the SUMO tag from His6-SUMO-TEV-RumT125, or the Histag from resolubilized His6-TEV-RumT125, using TEV protease were unsuccessful.



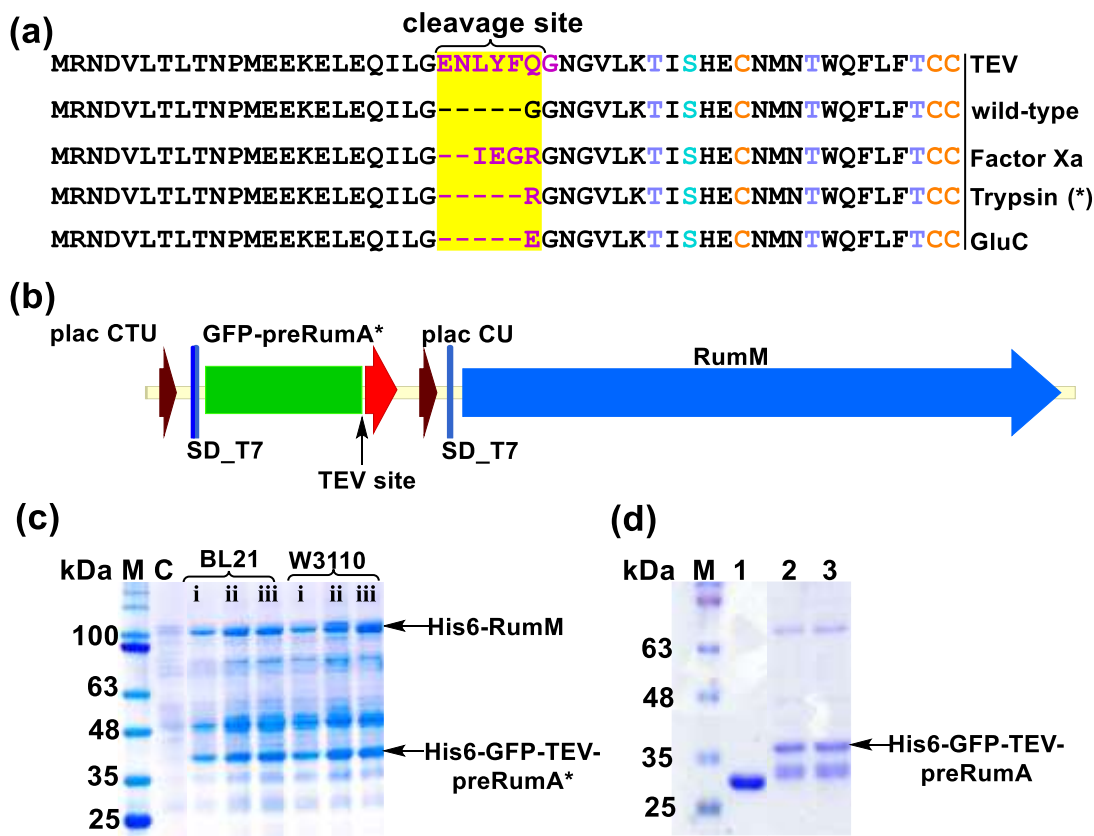
**Figure 4.13** Expression *RumT* and activity evaluation. (a) SDS-PAGE showing whole cell soluble extracts of *E. coli* W3110 (lane 1) and BL21 (lane 2) expressing pLEOrA, compared with *E. coli* W3110 (lane i) and BL21 (lane ii) expressing pLEOrT; as well as purified His6-TEV-RumT125 inclusion bodies (lane P1) and purified soluble His6-SUMO-TEV-RumT125 (lane P2). The bands representing His6-TEV-RumT125 (~17.13 kDa) and His6-SUMO-TEV-RumT125 (~29.12 kDa) are indicated by bold arrows. (b) SDS-PAGE showing purified His6-GFP-TEV-preRumA (lane 1), activity of TEV on His6-GFP-TEV-preRumA (lane 2), no activity for purified His6-SUMO-TEV-RumT125 (lane 3) and resolubilized/refolded His6-TEV-RumT125 (lane 4) on His6-GFP-TEV-preRumA.

We tested both constructs for their ability to cleave His6-GFP-TEV-preRumA at the Gly-Gly motif at the C-terminus of the leader peptide of preRumA and there was no evidence of any activity (Figure 4.13b). Such inactivity as well as non-specificity were also reported when a GST tag was fused to the N-terminus of LctT peptidase domain (Furgerson Ihnken et al., 2008). Here, the activity of the protease was evaluated only based on an expected shift in the mobility of the upper band of His6-GFP-TEV-preRumA on SDS-PAGE as demonstrated for TEV in lane 2 (Figure 4.13b). Further investigations on this aspect and/or optimization of the constructs were put on hold since the rationale for such experiments did not reflect much on the focus of the current study.

The basis for requesting further experiments arise from the fact that leader peptide cleavage by LanT proteins does not have a general mechanism. For instance, lactacin 481 transporter LctT cleaves both modified and unmodified preLctA indiscriminately (Furgerson Ihnken et al., 2008) while nukacin ISK-1 transporter NukT is only active in the presence of membrane vesicles and requires modified preNukA to serve as its substrate (Nishie et al., 2009). Whereas a cooperative action between peptidase and ATP binding domains of NukT (section 4.1.2) are intricately involved in preNukA leader peptide processing (Nishie et al., 2011), the N-terminal peptidase domains of LagD, ComA, CvaB and LctT (all AMS proteins) do not require ATP binding domain for activity (Havarstein et al., 1995, Wu and Tai, 2004, Ishii et al., 2006, Furgerson Ihnken et al., 2008).

### 4.3.3 Plasmid-encoded bicistronic operon

Introducing alternative leader peptide cleavage sites in preRumA was necessary to activate the peptide *in vitro* since expression of the peptidase domain of RumT was unsuccessful (section 4.3.2). TEV, factor Xa, Trypsin and GluC cleavage sites were incorporated into preRumA via site-directed mutagenesis (section 3.5.9). The sequence descriptions are illustrated in Figure 4.14a.



**Figure 4.14** Site-directed mutagenesis and bicistronic vector construction. (a) preRumA peptide sequences, showing the various cleavage sites engineered in the peptide via site-directed mutagenesis and residues targeted for PTMs. Position -1 is indicated by the yellow shading. The symbols \*, represents the Gly-1/Arg preRumA mutant construct that will be widely used for subsequent investigations in this work. (b) The arrangement of genes encoding the His6-GFP-TEV-preRumA\* and His6-RumM on pCUT7 vector. (c) Analysis of His6-GFP-TEV-preRumA\* and His6-RumM extracted from *E. coli* W3110 and BL21 expressing the bicistronic vector pLEOgrA\*M1. The letter C, control extract from WLEOrA strain (section 4.2.2). The symbols i, 20  $\mu$ M; ii, 50  $\mu$ M; iii, 100  $\mu$ M IPTG. (d) Analysis of purified GFP (lane 1); His6-GFP-TEV-preRumA encoded by the bicistronic vector pLEOgrAM in W3110 (lane 2) and BL21 (lane 3).

To further reduce the metabolic burden imposed by the presence of two distinct plasmids in the same expression host, we decided to clone the two genes encoding the chimeric preRumA

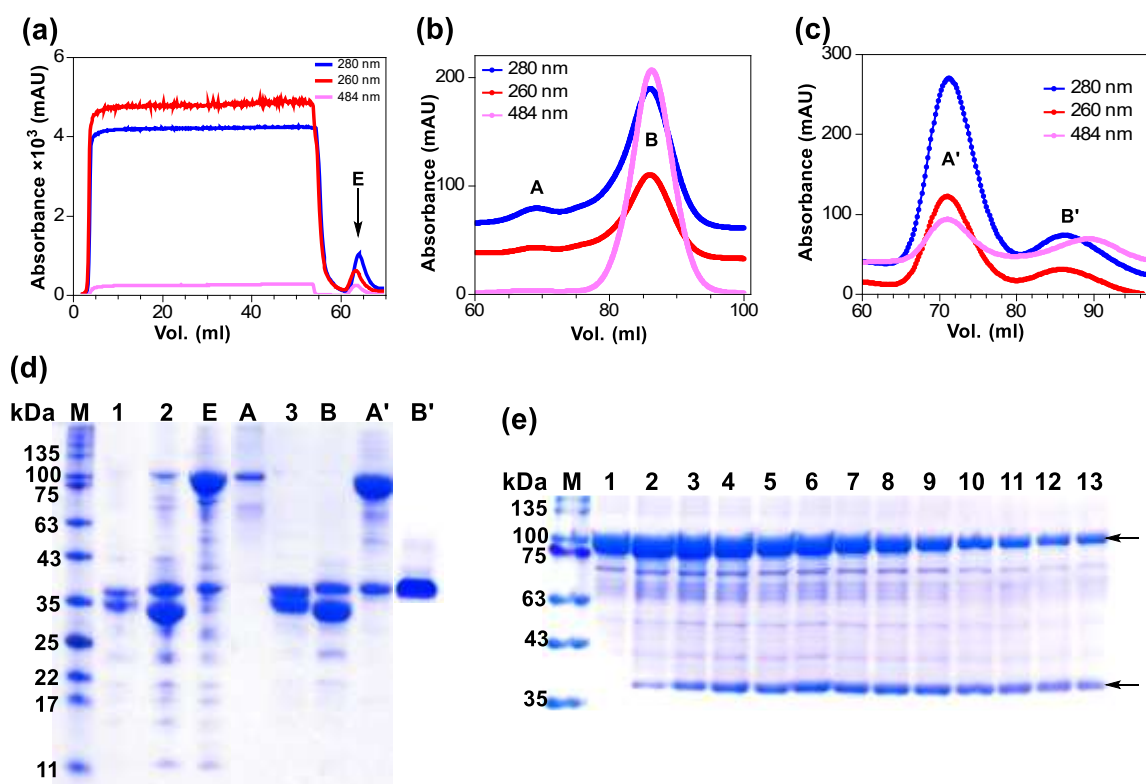


construct and the modifying enzyme RumM on a single-plasmid bicistronic expression vector. The plasmids were constructed as described in the experimental part (section 3.5.10.7). All plasmids were used to transform *E. coli* W3110 or BL21 and tested for expression. All the six constructs showed positive production of the desired product via SDS-PAGE analysis (data not shown). One of the *E. coli* W3110 strains (named WLEOgrA\*M1) hosting the bicistronic pLEOgrA\*M1 vector, and expressing both His6-GFP-TEV-preRumA\* and His6-RumM, exhibited impressive protein yields. Note here that the symbol (\*) denotes the variant containing trypsin cleavage site at position -1 of preRumA which was generated by replacing Gly-1 with Arg. For clarity, this mutant is referred to as preRumA\* while the wild-type remains preRumA throughout this report.

The bicistronic operon controlled by the *CTU* and *CU* promoters are arranged as schematically represented in Figure 4.14b. Expression tests demonstrated IPTG dependent expression of both targets, with very good yields estimated by the thickness of bands on SDS-PAGE (Figure 4.14c). It is important to note that the single plasmid-encoded bicistronic system obtained by cloning the RumM encoding gene under control of the *CTU* promoter (i.e. promoter/rumM fragment amplified from pLEOrM') did not resolve the degradation issues associated with His6-GFP-TEV-preRumA (Figure 4.14d). However, expression results from the pLEOgrA\*M1 vector system (where *rumM* is under control of the *CU* promoter) indicated almost complete eradication of this problem. Thus, reports in the following sections will focus more on the product of this system.

#### 4.3.4 Purification of His6-GFP-TEV-preRumA\* and TEV digestion

Bacterial lysates obtained from 500-ml WLEOgrA\*M1 or WLEOgrAM culture were purified via IMAC. Figure 4.15a represents the chromatogram of a single IMAC run, indicating a sharp elution peak. Elution fractions representing this peak were subsequently collected and pooled together and further purified via gel filtration. The elution chromatograms showed UV signals at 260 and 280 nm (representing protein absorption) and at 484 nm (representing fluorescence absorption).



**Figure 4.15** Expression and purification of GFP-preRumA fusion constructs. Bacterial lysates were purified by IMAC followed by size exclusion chromatography. (a) Chromatogram of IMAC purification of lysate extracted from WLEOgrA\*M1. (b) Size exclusion chromatogram showing elution of His6-RumM (peak A) and His6-GFP-TEV-preRumA (peak B) both purified from WLEOgrAM strain. (c) Size exclusion chromatogram showing elution of His6-RumM/His6-GFP-TEV-preRumA\* complex (peak A') and His6-GFP-TEV-preRumA\* (peak B'). (d) SDS-PAGE analyses of the elution peaks, IMAC purifications of lysate from WLEOgrA\* (1) and WLEOgrAM (2), as well as gel filtration purification of IMAC extract from WLEOgrA\* (lane 3). Letters atop the gel; A, B, A', B' and E are pooled fractions from peak labels in (a), (b) and (c). (e) SDS-PAGE analysis of individual fractions that constitute the chromatographic peak A'.

The peaks marked “A” represented elution of His6-RumM and peak B represented elution of His6-GFP-TEV-RumA (Figure 4.15 b). Furthermore, peak A' represented elution of a complex mixture of His6-RumM and His6-GFP-TEV-preRumA\* while peak B' belonged to His6-GFP-TEV-preRumA\* (Figure 4.15 c). The pooled fraction collected for each of these elution peaks were analyzed via SDS-PAGE (section 3.6.5) and results clearly indicated that the system with enhanced RumM production resolved the degradation problem since product purified from strain WLEOgrA\*M1 did not show multiple bands corresponding to His6-GFP-TEV-preRumA\* as did constructs purified from WLEOgrA\* or WLEOgrAM (Figure 4.15d). Overexpression of His6-RumM was also visibly apparent in samples purified from WLEOgrA\*M1 strain. The concentrations of the chimeric construct purified from various systems are displayed in Table 3.2.

His6-GFP-TEV-preRumA\* was degraded when produced without or in the presence of low amount of RumM, but not when high amount of RumM was present. These results are very important since they clearly show that the presence of sufficient amount of the modifying enzyme *in vivo* stabilizes preRumA fused to the C-terminus of GFP. It should be noted that we have earlier shown that the degradation occurs within the preRumA sequence (section 4.2.6). Results from SDS-PAGE analyses of eluted fractions collected from chromatographic peak A' also supported the fact that His6-GFP-TEV-preRumA\* and His6-RumM were coeluted as a single complex (Figure 4.15e). It is worthy to note that the amount of His6-GFP-TEV-preRumA\* stays fairly constant from lane 4 to around lane 9 while the amount of His6-RumM decreases. This observation was further investigated by native PAGE (see section 4.3.5).

**Table 3.2** Purification yields of total protein per litre

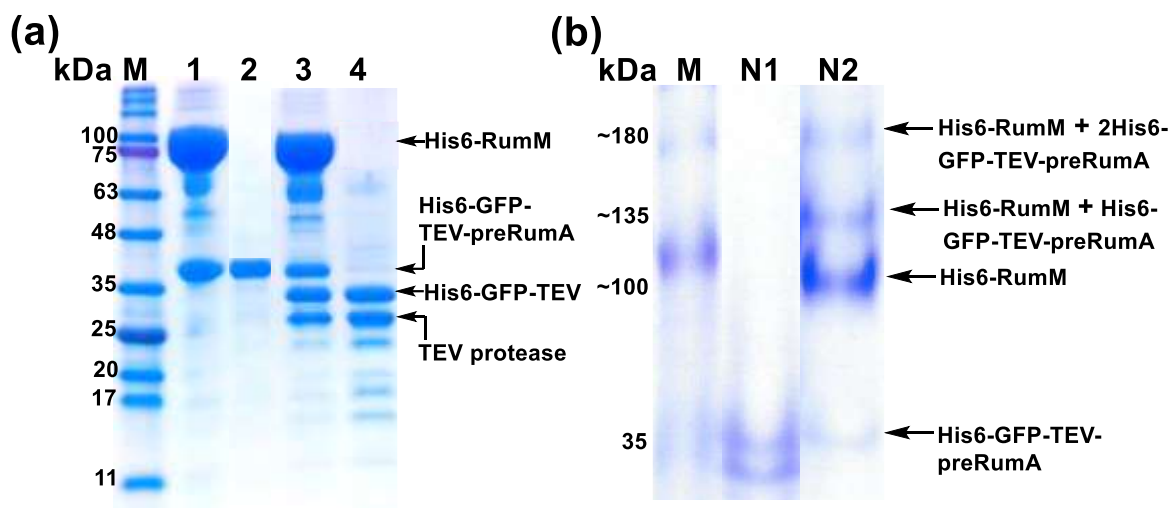
Purification step	His6-GFP-TEV-preRumA (mg)	His6-GFP-TEV-preRumA* (mg)	<sup>‡</sup> His6-GFP-TEV-preRumA* (mg)	preRumA (mg)	preRumA* (mg)	<sup>‡</sup> preRumA* (mg)
IMAC	84.53	97.61	97.61	-	-	
Size exclusion	65.32	39.01	79.40	-	-	
TEV-digested IMAC FT	-	-	-	3.23	4.45	6.14
TEV-digested butanol extract	-	-	-	1.34	2.50	4.75

The symbol <sup>‡</sup>, pooled fractions from peaks A' and B'. This specifically applies to His6-GFP-TEV-preRumA\* which appeared in two separate chromatographic peaks during purification. FT; flow through

#### 4.3.5 Interactions between His6-RumM and His6-GFP-TEV-preRumA\*

His6-RumM and His6-GFP-TEV-preRumA\* were coeluted together and appeared as a single peak during size exclusion chromatographic purification as shown in the previous section (section 4.3.4) despite their enormous size difference (~73 kDa). Samples were obtained from the combined eluted fractions that represented the chromatographic peaks and digested with TEV. Results in Figure 4.16a show that His6-GFP-TEV-preRumA\* complexed with His6-RumM was only partially susceptible to TEV cleavage while the same construct not complexed with His6-RumM was readily digested during the overnight reaction. These results suggested possible molecular interactions between His6-RumM and the preRumA fused to the C-terminus of GFP. Most highly, such interactions could possibly shield the cleavage site making it inaccessible to TEV.

The presence of both digested and undigested products in the TEV cleavage reaction involving the His6-RumM/His6-GFP-TEV-preRumA\* complex further suggests that the interactions between His6-GFP-TEV-preRumA\* and His6-RumM may be either cooperative or noncooperative. Cooperativity here would refer to interactions with both the core peptide and the leader peptide at the same time. Assuming that RumM interacts noncooperatively with the core peptide only, the leader peptide which is closest to the TEV site would be free, creating enough space for TEV to engage. However, if RumM interacts cooperatively with both leader and core segments, or noncooperatively with the leader peptide alone, there is a chance that the TEV site would be buried and rendered inaccessible to the protease. The formation of complexes between precursor peptide and modifying enzymes is well-established in the literature (Khusainov et al., 2011, Mavaro et al., 2011, Repka et al., 2017).



**Figure 4.16** TEV digestion and Native PAGE analysis of His6-RumM/His6-GFP-TEV-preRumA\* complexes. (a) SDS-PAGE analysis of His6-RumM/His6-GFP-TEV-preRumA\* complex (lane 1), noncomplexed (free) His6-GFP-TEV-preRumA\* (lane 2), overnight TEV digests of complexed His6-GFP-TEV-preRumA\* (lane 3) and free His6-GFP-TEV-preRumA\* (lane 4). (b) Native PAGE of purified His6-GFP-TEV-preRumA from WLEOgrAM strain (lane N1) and complexed His6-GFP-TEV-preRumA\* (lane N2). The latter shows multiple bands above the expected molecular weight of His6-RumM corresponding to 1 or 2 molecules of His6-GFP-TEV-preRumA\* complexed to one molecule His6-RumM

A sample from the His6-RumM/His6-GFP-TEV-preRumA\* complex was obtained and run on native PAGE employing purified His6-GFP-TEV-preRumA from WLEOgrAM strain (with low RumM expression) as a control. Result in Figure 4.16b shows the usual two bands reported for the latter, while four distinct bands were obtained for the former belonging to His6-GFP-TEV-preRumA\*, His6-RumM and possible His6-RumM/His6-GFP-TEV-preRumA complexes

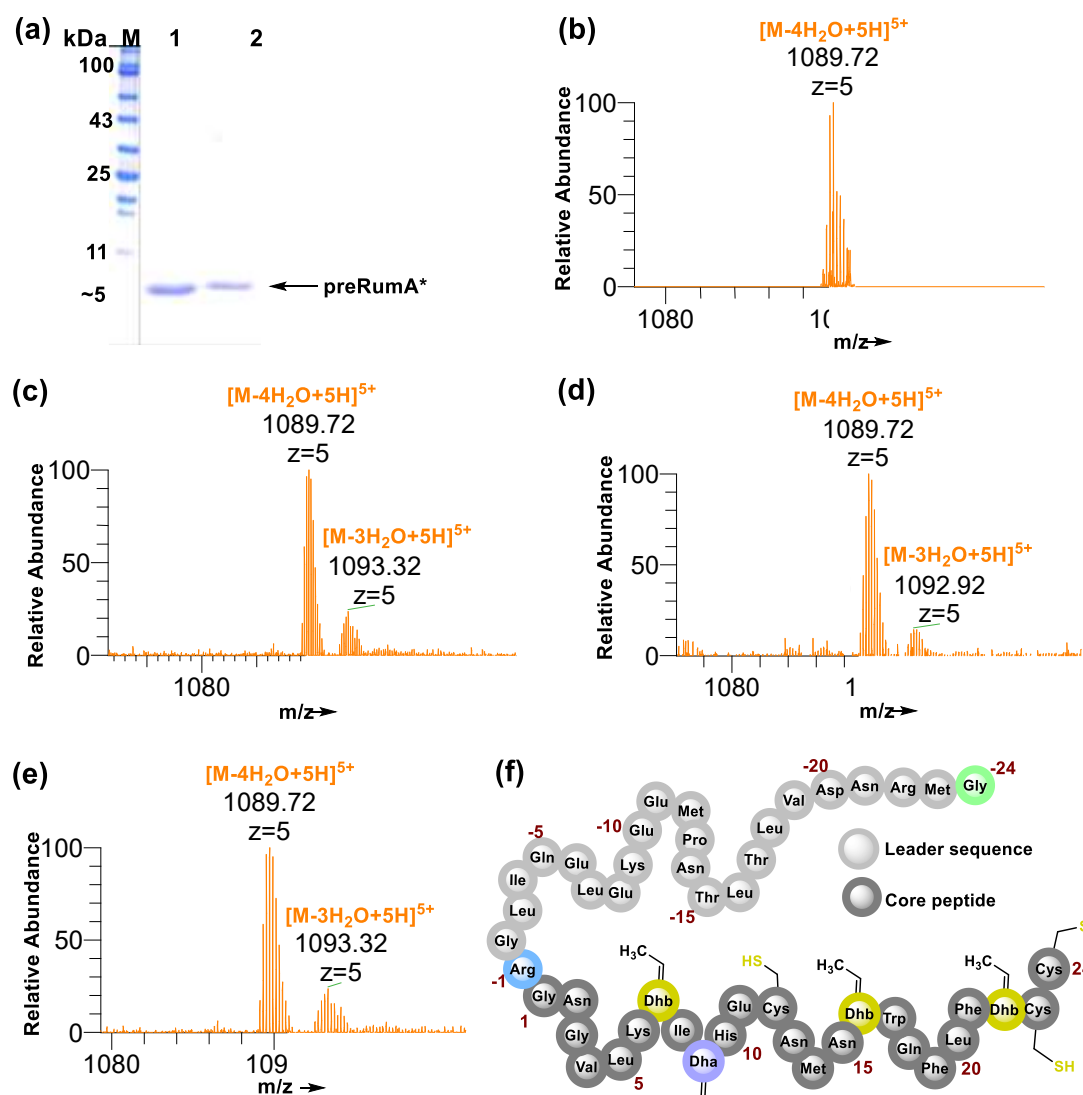
as indicated on the gel. The result further suggests that the complex formation may have occurred between His6-RumM and dehydrated preRumA\* since Mavaro et al. (2011) demonstrated that the interactions between modifying enzyme and dehydrated precursor peptide is much stronger. It would be interesting to determine if this is actually the case with the His6-RumM/His6-GFP-TEV-preRumA\* complex.

The aspect of interactions between modifying enzymes and precursor peptide has been investigated in detail for the prototype lanthipeptide nisin. Biophysical characterization of purified NisA and NisB indicated that the leader peptide creates a docking motif that allows NisB to bind and interact very strongly with the unmodified, dehydrated or mature NisA (Mavaro et al., 2011). Cooperative interactions involving both the core and the leader segments of the precursor peptide, and the modifying enzyme were previously observed (Khusainov et al., 2011).

#### 4.3.6 Extraction and nLC-ESI-MS analyses of preRumA\*

Having a simple method to isolate the peptide may facilitate characterization of several mutant constructs. In this study, TEV-digested samples were processed via a Ni<sup>2+</sup>-NTA column to remove the His6-GFP-TEV component and/or extracted with butanol as described in the methods part (section 3.7.4). The concentrations of the respective extracts and purified products are presented in Table 3.2. Dried samples from -20 °C were redissolved in water and processed for SDS-PAGE analysis. Purified preRumA extracts are shown in lanes 1 and 2 of Figure 4.17a.

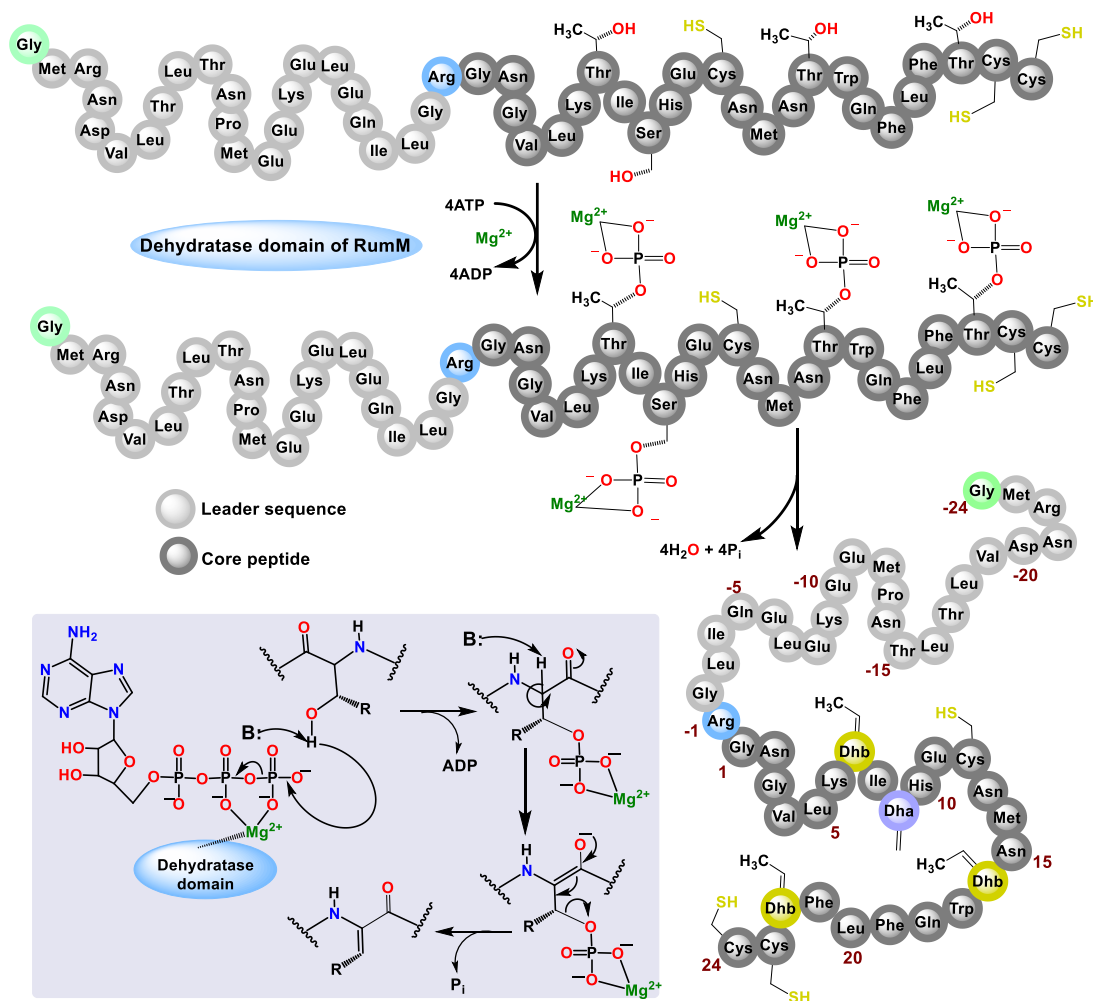
Dried samples from -20 °C were equally resuspended in 0.1 % formic acid and processed using the C18 pipette tips (section 3.10.1) and analyzed by mass spectrometry. The Orbitrap Fusion nLC-ESI-MS analysis (section 3.10.2) of both butanol and IMAC flow through extracts, as well as direct ZipTip purification of the crude digest and the aqueous phase of the butanol extract produced the mass spectra in Figures 4.17b-e, indicating fourfold dehydrated preRumA\* (structure is schematically represented in Figure 4.17f). The mass-to-charge (m/z) ratios of the resulting five times charged ion  $[M-4H_2O+5H]^{5+}$  measured for preRumA\* in all cases were consistent with the calculated m/z of 1089.72 corresponding to preRumA\* including the loss of four water molecules.



**Figure 4.17** TEV-digested His6-GFP-TEV-preRumA\*, extraction of preRumA\* and nLC-ESI-MS analyses. (a) SDS-PAGE analysis of IMAC flow through extract (lane 1) and butanol extract (lane 2) from purified His6-GFP-TEV-RumA\* TEV digest. EIS-MS spectrum of preRumA\* extracted from TEV-digested His6-GFP-TEV-RumA\* using butanol (b), ZipTip purification of the aqueous phase of butanol extract (c), direct ZipTip purification of crude digest (d) and IMAC flow through (e). The calculated molecular mass of the unmodified peptide  $[M+5H]^{5+} = 1104.12$ . (f) Schematic structure of preRumA\*, indicating the dehydroamino acids, an extra residue (Gly-24) that is included in the peptide following cleavage by TEV and Arg that replaced Gly-1 in wild-type preRumA to enable trypsin cleavage.

We deduced from the mass spectra of samples purified via the different methods that butanol extraction isolated only the four times dehydrated product since traces of three times dehydrated preRumA\* were present in all the other extracts but absent in the butanol extract. Furthermore, applying the ZipTip procedure to purify the aqueous phase of the butanol extract indicated that the organic solvent extraction isolated only a fraction of the fully modified peptide as traces of fourfold and threefold dehydrated products were also visibly apparent in the mass

spectrum when the aqueous phase extract was analyzed. The accuracies of measurements are expressed in [ppm] and recorded in Table 3.3. These results suggest activity of RumM as dehydration of Ser/Thr in the core peptide of preRumA\* are detectable. Comparing this data with the literature pointed, that these losses occurred at Thr7, Thr16 and Thr22 to yield three Dhb, and Ser9 to produce Dha as illustrated in Figure 4. 17f.



**Figure 4.18** Proposed reaction mechanism through which ruminococcin-A synthetase or other LanM enzymes convert Ser/Thr residues in the core peptide into dehydroamino acids via phosphorylation and subsequent phosphate elimination from pSer/pThr

The scheme displayed in Figure 4.18 illustrates a suggested mechanism for the reaction catalyzed by the dehydratase domain of RumM. According to this scheme, RumM, like other LanM enzymes including LctM, BovM and NukM, requires ATP hydrolysis to supply phosphate for phosphorylation using  $Mg^{2+}$  as cofactor, followed by an elimination reaction which also utilizes ADP as a cofactor to convert the phosphate esters of Ser/Thr (pSer/pThr) to Dha/Dhb (Xie et al., 2004, Knerr and van der Donk, 2012, Ma et al., 2014, Shimafuji et al., 2015, Repka

et al., 2017). Earlier investigations revealed that the dehydratase introduces pSer/pThr into the core peptide via a unidirectional N- to C-terminal catalytic mode (Lee et al., 2009), while the leader peptide plays a determinant role in ensuring catalytic efficiency (Thibodeaux et al., 2016).

Our data shows that the dehydratase domain of RumM successfully catalyzed the installation of dehydroamino acids into the core peptide of preRumA in *E. coli*. These PTMs were identified by multiple charged ions in the mass spectra corresponding to the molecular mass of the precursor peptide, with an equivalent loss of four water molecules. These data supplied incontestable evidence that *in vivo* biosynthesis and modification of the (Gly-1/Arg)preRumA mutant is also achievable in *E. coli*. The information provided herein may then be used to design *in vitro* experiments to characterize the activity of RumM.

**Table 3.3** Predicted and empirical molecular masses measured in the LC-ESI-MS and Orbitrap Fusion nLC-ESI-MS & MS<sup>2</sup> spectra of *in vivo* *E. coli* synthesized preRumA

Molecular mass <sup>[c]</sup>	Chemical Formula	Exact mass calculated	Mass found	Error [ppm]	Charge
<b>Wild-type preRumA</b>					
<i>M</i>	C232H369N63O74S6	1083.7143	1083.7201	5.35	5
<i>M</i> -3H <sub>2</sub> O	C232H363N63O71S6	1073.5085	1073.512	3.26	5
<i>M</i> -4H <sub>2</sub> O	C232H361N63O70S6	1069.9064	1069.9128	5.98	5
<b>preRumA*</b>					
<i>M</i>	C236H378N66O74S6	1104.1208	-	-	-
<i>M</i> -3H <sub>2</sub> O	C236H372N66O71S6	1093.3244	1093.3249	0.45	5
<i>M</i> -4H <sub>2</sub> O	C236H370N66O70S6	1089.7223	1089.7211	1.10	5
<b>preRumA* [b-ions]</b>					
b3 <sup>+</sup>	C13H25N6O3S	345.178	345.1775	1.44	1
b4 <sup>+</sup>	C17H30N8O5S	459.2131	459.2125	1.30	1
b5 <sup>+</sup>	C21H35N9O8S1	574.2401	574.2392	1.56	1
b6 <sup>+</sup>	C26H44N10O9S1	673.3085	673.3075	1.48	1
b7 <sup>+</sup>	C32H55N11O10S2	786.3925	786.3916	1.14	1
b8 <sup>+</sup>	C36H62N12O12S1	887.4402	887.4412	1.12	1
b9 <sup>+</sup>	C42H73N13O13S1	1000.5243	1000.5228	1.49	1
b10 <sup>+</sup>	C46H80N14O15S1	1101.5717	1101.5712	0.45	1
b11 <sup>+</sup>	C50H86N16O17S1	1215.6147	1215.6128	1.56	1
b12 <sup>2+</sup>	C55H93N17O18S1	656.8374	656.8366	1.21	2



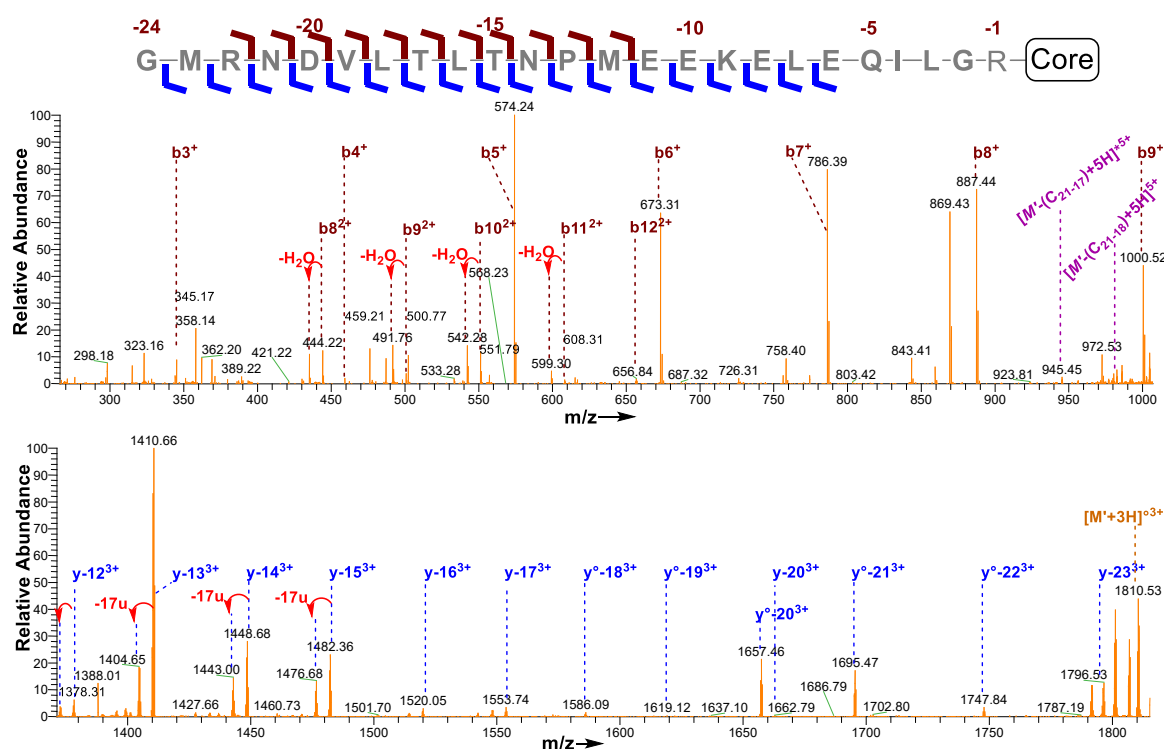
**preRumA\* [y-ions]**

y-23 <sup>3+</sup>	C234H367N65O69S6	1796.5264	1796.5266	0.11	3
y <sup>o</sup> -22 <sup>3+</sup>	C229H357N64O67S5	1747.8464	1747.8414	2.86	3
y <sup>o</sup> -21 <sup>3+</sup>	C223H343N59O67S5	1695.4767	1695.4703	3.77	3
y-20 <sup>3+</sup>	C219H340N58O65S5	1662.7982	1662.7937	2.70	3
y <sup>o</sup> -19 <sup>3+</sup>	C215H333N57O61S5	1619.12	1619.1201	0.06	3
y <sup>o</sup> -18 <sup>3+</sup>	C210H324N56O60S5	1586.0968	1586.0925	2.71	3
y-17 <sup>3+</sup>	C204H315N55O60S5	1553.7383	1553.7401	1.15	3
y-16 <sup>3+</sup>	C200H308N54O58S5	1520.0557	1520.0538	1.24	3
y-15 <sup>3+</sup>	C194H297N53O57S5	1482.361	1482.3607	0.20	3
y-14 <sup>3+</sup>	C190H290N52O55S5	1448.6784	1448.6804	1.38	3
y-13 <sup>3+</sup>	C186H284N50O53S5	1410.664	1410.6622	1.27	3
y-12 <sup>3+</sup>	C181H277N49O52S5	1378.313	1378.3112	1.30	3
y-11 <sup>5+</sup>	C176H268N48O51S4	1334.6329	1334.6313	1.19	3
y-10 <sup>5+</sup>	C171H261N47O48S4	1292.6181	1292.6145	2.78	3
y-9 <sup>5+</sup>	C166H254N46O45S4	1248.2709	1248.2687	1.76	3
y-8 <sup>5+</sup>	C160H242N44O44S4	1205.5726	1205.5714	0.99	3
y-7 <sup>5+</sup>	C155H235N43O41S4	1162.5584	1162.557	1.20	3
y-6 <sup>5+</sup>	C149H224N42O40S4	1124.5298	1124.5311	1.15	3

**4.3.7 Mass Spectrometric Fragmentation and MS<sup>2</sup> Analysis of PreRumA\***

The processing of the precursor peptide by some class III lanthipeptide biosynthesis enzymes have been shown to proceed in a C→N-terminal mode (Krawczyk et al., 2012, Jungmann et al., 2014), establishing an inclusive biosynthesis model involving phosphorylation, elimination and cyclization (Jungmann et al., 2014). We have demonstrated in the previous sections (sections 3.2.6 and 3.3.5) that the ruminococcin-A lanthionine synthetase RumM is able to introduce dehydroamino acids into preRumA. However, this data did not contain information about cyclization of the peptide because both the linearly dehydrated peptide (with no thioether rings) and the cyclized version (with thioether rings) are identical with respect to their molecular mass (they both have an average molecular mass of 5444.27 Da). Furthermore, if the C-terminus of RumM is dysfunctional, accumulation of dehydrated preRumA would still be possible as reported for nisin where nisin cyclase was inactivated (Lubelski et al., 2009). Thus, it would be premature to assume that data presenting evidences of dehydration directly indicate the presence of thioether rings in the precursor peptide. Here, we report the characterization of modified full-length preRumA\* using tandem MS analysis.

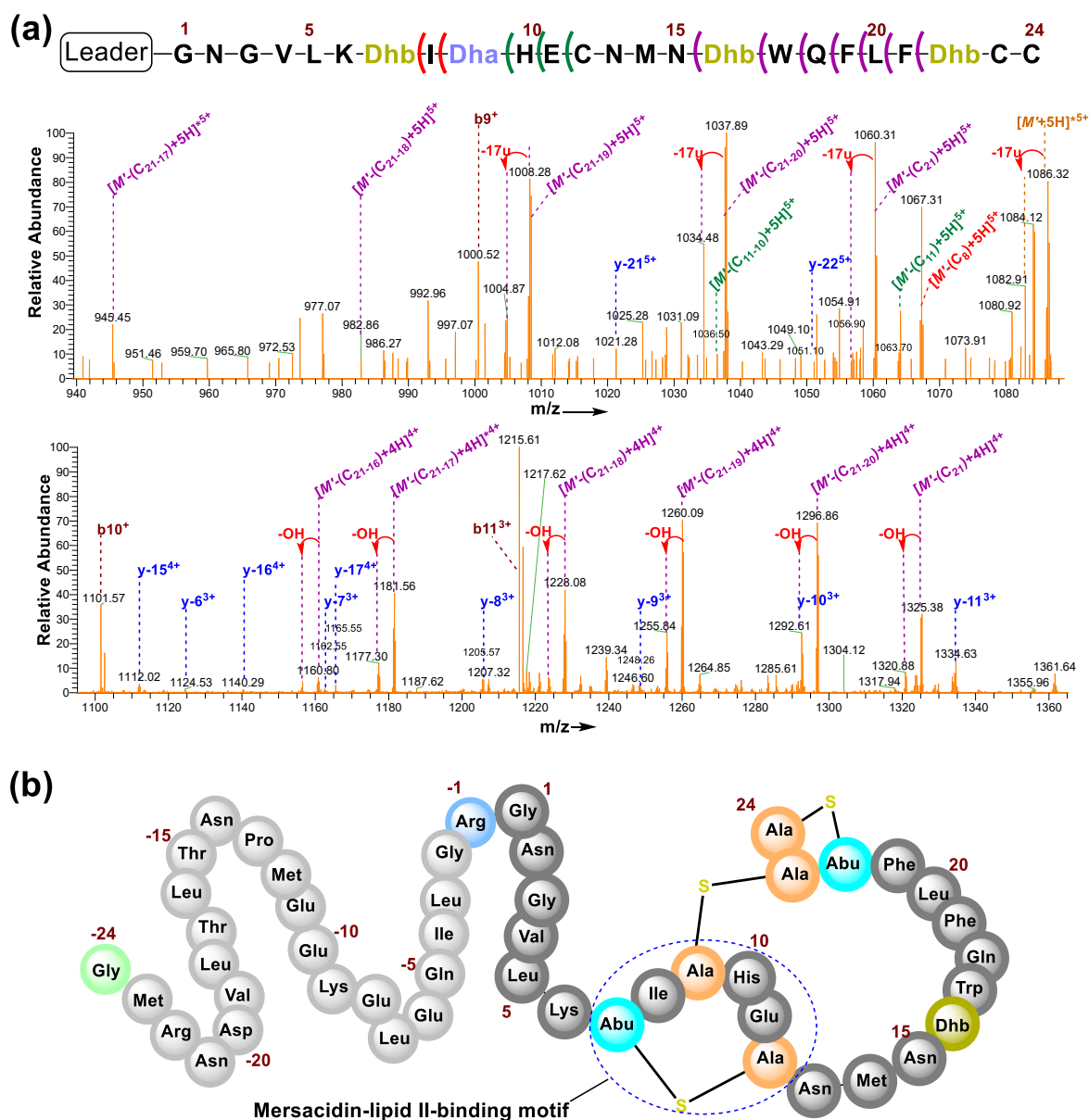
A combination of CID, HCD and ETD at different fragmentation energies (in Volts) was used to generate different fragmentation patterns (Appendix 6.6). In the CID and HCD spectra, high-intensity b- and y-type product ions were identified, that fitted perfectly to the first 19 amino acid residues of the leader peptide (Figure 4.19). With respect to the primary structure of the peptide, no high-intensity b- or y-type ions were easily identified at the C-terminus, associated with fragmentation of the core peptide. This is reasonable since data from the literature reveal that the presence of cyclic thioether cross-bridges in the core peptide confer resistance to fragmentation (Goto et al., 2010, Meindl et al., 2010, Müller et al., 2010, Sambeth and Süßmuth, 2011, Krawczyk et al., 2012, Völler et al., 2012, Férir et al., 2013, Krawczyk et al., 2013, Völler et al., 2013). However, a series of intensive product ions that fitted to successive loss of residues from the core peptide were observed (Figure 4.20a).



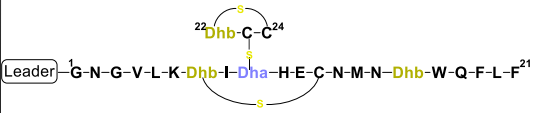
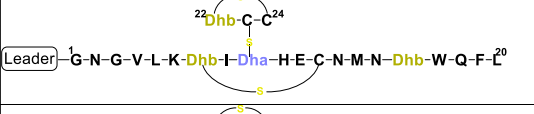
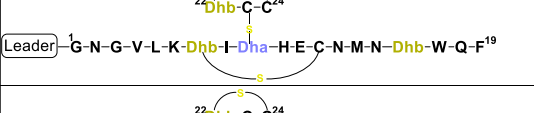
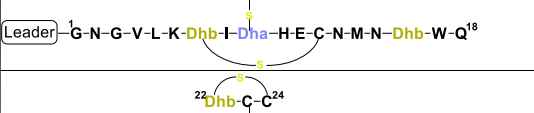
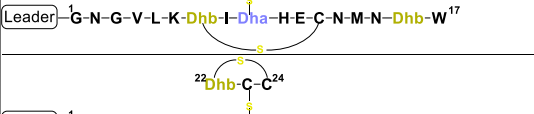
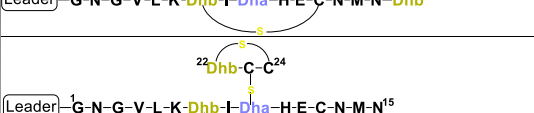
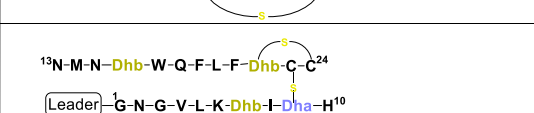
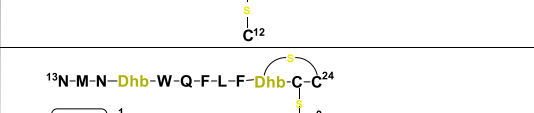
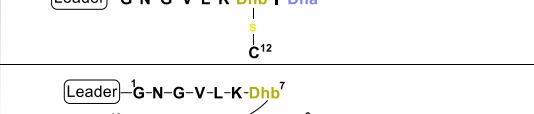
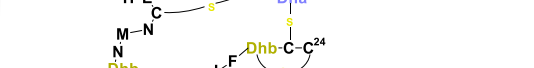
**Figure 4.19.** Tandem MS<sup>2</sup>-experiments and product ions assignment for the leader peptide of preRumA\*. Orbitrap Fusion MS<sup>2</sup> spectra representing ion series with intensive peaks produced from fragmentation of the N-terminal leader peptide of preRumA\*.  $[M'+5H]^{5+} = 1089.72$ .

Note that in labelling the product ion peaks in Figures 4.19 and 4.20 the following notations were considered: C<sub>n</sub> denotes the core peptide residue (C) at specific position (n); while C<sub>n-n'</sub> denotes core peptide residues from position n to position n' following a C- to N-terminal direction. No further high-intensity product ions were identified within the core peptide. The

current data describes the structure of preRumA\* schematically represented as shown in Figure 4.20b. Figure 4.21 describes the structure associated with product ions generated as a result of the loss of each successive core peptide residue. The measured molecular masses nicely fitted to the structures of the fragment presented in the first column.



**Figure 4.20** MS<sup>2</sup>-experiments and product ions assignment for preRumA\* core peptide. (a) nLC-MS<sup>2</sup> spectra showing intensive ion series resulting from fragmentation of C-terminal core peptide of preRumA\*.  $[M+5H]^{5+} = 1089.72$  RumA\*.  $C_n$  denotes the position of the core peptide amino acid residue (e.g.  $C_{21-17}$  denote residues 21 to 17) (b) Proposed schematic structure of modified preRumA\*.

Peptide Fragment	Residues removed	Exact calculated mass	Measured mass	Charge	Error [ppm]
	None	1089.7223	1089.7211	+5	1.10
	Phe21	1325.1344	1325.1327	+4	1.28
	Phe21, Leu20	1296.863	1296.8616	+4	1.38
	Phe21, Leu20, Phe19	1260.0963	1260.0946	+4	1.34
	Phe21, Leu20, Phe19, Tyr18	1228.0816	1228.0799	+4	1.38
	Phe21, Leu20, Phe19, Tyr18, Trp17	1181.5617	1181.5607	+4	0.84
	Phe21, Leu20, Phe19, Tyr18, Trp17, Dhb16	1160.8024	1160.8003	+4	1.80
	Glu11	1063.7137	1063.7017	+5	11.28
	Glu11, His10	1036.5020	1036.5087	+5	6.46
	Ile8	1067.1055	1067.1045	+5	0.93

**Figure 4.21** Analysis of fragmented product ions obtained from C-terminal core peptide of preRumA\*. This figure describes the possible cross-linked structures that gave rise to the resulting ions identified in the mass spectrum. The charges and measurement precisions in parts per million [ppm] are also indicated.

To further expatiate on the structural description of the various fragments that gave rise to the intensive product ion peaks identified in the mass spectrum, different structures were proposed of which those presented in Figure 4.21 fitted precisely to the measured  $m/z$ . Notice that each structure represent an internal double cleavage of the polypeptide backbone from the C-terminus that resulted in product ions associated with the successive loss of amino acids in the

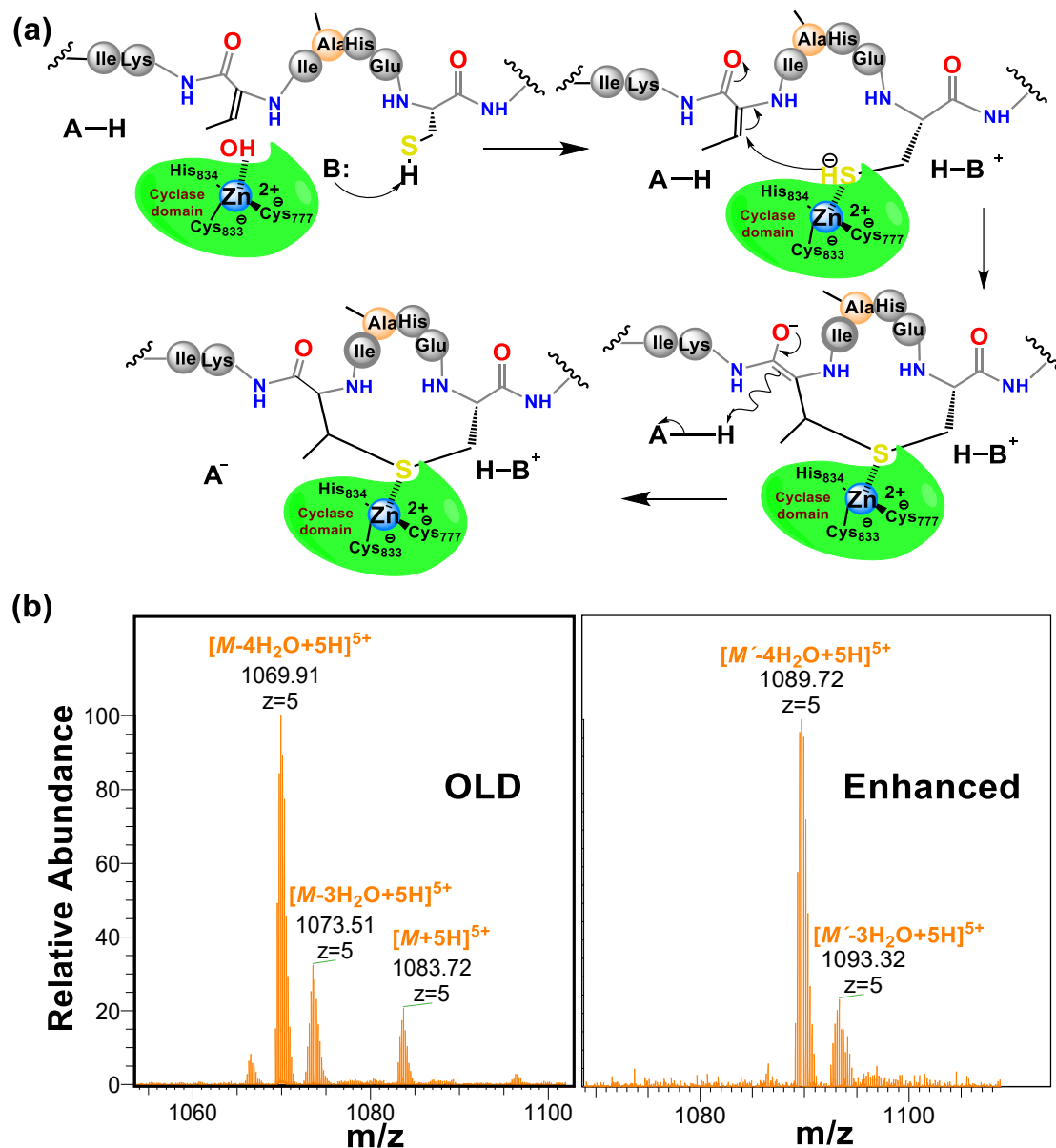
direction indicated by the arrow. Any loss of an internal residue should automatically result in b- and/or y-type ion(s) or other associated ion types like a, c, x and z. This was not the case observed with the fragmentation of C-terminal preRumA\*. Not even the x- and z-type which are observed during de novo peptide sequencing were identified. Instead, the double cleavage fragments hosted residues present upstream and downstream of the cleavage position. This is only possible if there is a bond linking the C-terminal fragment to the N-terminal segment. The presence of cysteine or Dhb/Dha residue adjacent to these cleavage positions allowed us to allocate thioether cross-linkages. Moreover, Figure 4.21 shows details of how the various structures were associated to their respective molecular masses, with very low mass errors (in ppm).

Whereas the present data are consistent with those obtained by Dabard and coworkers (Dabard et al., 2001), our findings further indicated the presence of a third MeLan ring formed between Thr7 and Cys12. It is worth noting that Dabard et al. (2001) concluded in their study that the formation of thioether cross-bridge between these residues was not possible. Their conclusion was however, based on Edman degradation assay which has some limitations in studying the structure of lanthipeptides since it is blocked by Lan/MeLan rings as well as Dhb/Dha residues, and no sequence information can be obtained thereafter (Lohans and Vederas, 2014). We confidently state that the final structure determined herein represents an improved structure of modified ruminococcin-A precursor peptide because the third MeLan ring identified in this study also constitutes the mersacidin-lipid II-binding motif which appears to be common in all class II lanthipeptides (Knerr and van der Donk, 2012).

The RumM-catalyzed Michael addition cyclisation reactions may follow the scheme proposed in Figure 4.22. The scheme is similar to the mechanism proposed elsewhere for NisC (Li et al., 2006). Generally, there is an overwhelmingly low sequence identity between constituent family members of lanthionine-generating enzymes. However, the C-terminus of RumM possesses conserved active site residues which have been identified in other lanthionine cyclases to be critically implicated in the catalytic process (section 4.1.1). Hence, by inference, there exist possible structure-function similarities and closely similar mode of action as discussed elsewhere in the literature (Knerr and van der Donk, 2012, Repka et al., 2017).

The scheme illustrates that dedicated active site residues of the cyclase domain of RumM coordinates a hydrated  $Zn^{2+}$  cofactor whose water molecule is displaced by the cysteinyl thiol (Cys12) which is targeted for conjugation with Dhb7 in the core peptide of preRumA. As a result, the thiol group by itself becomes activated, creating a nucleophilic active center. The displacement of the water molecule or the activation of the cysteinyl thiol may be facilitated by His778 (Figure 4.1, section 4.1.1) which probably plays the role of an acid/base catalyst like

the Tyr304 in SpaC (Helfrich et al., 2007). The electrophilic carbon atom of the Dhb7 now launches an attack on the activated Cys12, forming an enolate intermediate which is subsequently protonated to generate the thioether cross bridge.



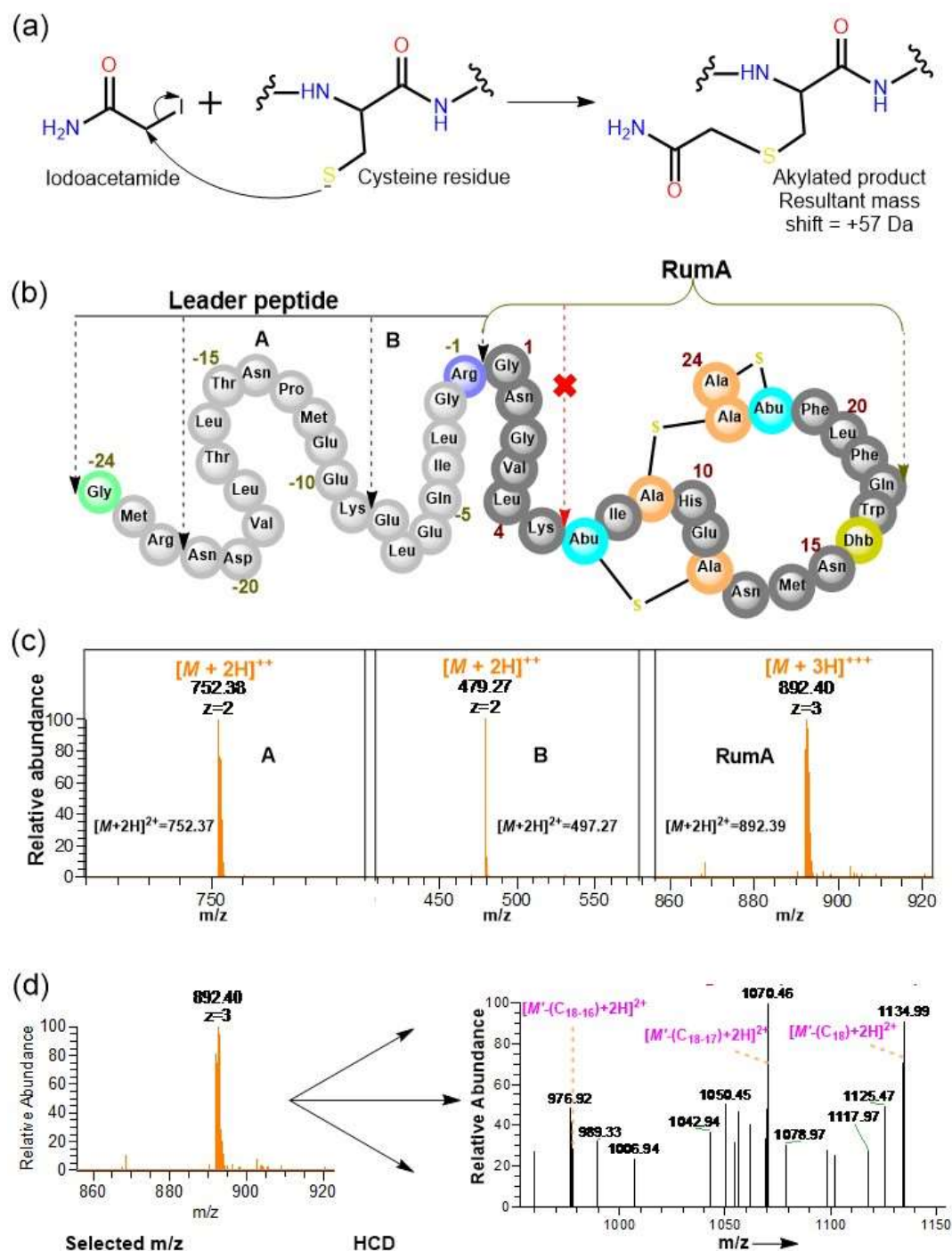
**Figure 4.22** Proposed mechanism of Michael-type addition reaction catalyzed by RumM and catalytic efficiency. (a) The C-terminus of this bifunctional enzyme possesses the conserved residues that have been shown to coordinate  $Zn^{2+}$  ion in the active site of some lanthionine cyclase. The scheme illustrates the formation of the first N-terminal lanthionine ring of RumA. (b) Comparison between the old and the new enhanced system with respect to products intermediates formed. Note that the differences in masses arise from the fact the new old system expressed the wild-type while the enhanced system expressed the Gly-1/Arg mutant.

The WLEOgrA\*M1 strain did not only resolve the degradation issues observed with the chimeric GFP fusion construct, but it also improved yields of the desired products drastically. In addition, only traces of triple-dehydrated preRumA\* intermediate were observed in the mass spectrum, and no unmodified product as was the case in the WLEOgrA/M system (Figure 4.22b). Although reports indicate that the formation of thioether cross-linkage may circumvent the dehydratase from further converting Thr/Ser to Dhb/Dha (Kuipers et al., 2008, Lubelski et al., 2009), we cannot at this time report whether the intermediate products observed contained at least one of the rings or not. Nevertheless, the data presented here show that low levels of RumM may have been responsible for the inefficiency of the WLEOgrA/M system.

Results from the tandem MS characterization show that RumM catalyzes the dehydration of Thr/Ser residues in the core peptide of preRumA into dehydroamino acids, and subsequently conjugate the Dhb/Dha residues with the sulfhydryl groups of specific cysteine side chains to produce a modified version of preRumA containing three MeLan/Lan cross-linkages and an additional Dhb. The dehydratase reactions involve Thr7, Ser9, Thr16 and Thr22 while the cyclase reactions involve the conjugation of the following pair of residues; Dhb7/Cys12, Dhb22/Cys24 and Dha9/Cys23 via Michael-type addition cyclization reactions.

#### 4.3.8 Alkyl derivatization and trypsin digestion of preRumA\*

Naturally, trypsin does not cleave at position 6 of the core peptide due to modification on the adjacent Thr7 (Dabard et al., 2001, Gomez et al., 2002, Gomez et al., 2002). The PTMs on Thr7 render the lysine residue inaccessible to trypsin. It was on this basis that we selected trypsin to serve as an alternative leader peptide processing strategy. Prior to trypsin cleavage, preRumA\* was first derivatized with iodoacetamide (see section 3.9). The alkylating agent reacts irreversibly with the sulfhydryl group of the cysteine side chain according to the scheme illustrated in Figure 4.23a. Assuming one cysteine side chain were to be involved in this reaction, the mass spectrum of the derivatized sample would show an alkylated species with an overall mass shift of +57 Da corresponding to the alkyl derivative. Therefore, it was expected that if preRumA\* did not contain any Lan/MeLan rings, all three cysteines in the core peptide would have been free to engage in the alkylation reaction. Interestingly, we were unable to identify charged ions associated with either single, double or triple alkylation of full-length preRumA\* and the various tryptic peptide fragments, including the mature RumA labelled in Figure 4.23b. Note that all the three cysteine residues in preRumA\* peptide sequence are present in the core peptide. Thus, it was expected that the alkyl modifications would occur in this segment of the peptide but results plainly displayed no evidence to suggest that there were any alkyl derivatives in the sample.



**Figure 4.23** Alkyl derivatization and trypsin digestion of preRumA\*. (a) Scheme showing how iodoacetamide react with thiol group of cysteine side chain. (b) Schematic representation of preRumA\* showing five different fragments expected from the tryptic digestion. (c) Charged ions observed in nLC-ESI-MS spectrum, representing the various tryptic fragments of preRumA\*. The predicted exact masses for each of the fragments are labelled in their respective spectra. (d) Tandem MS experiments and assignment of trypsin-activated RumA fragment ion peaks. Major product ions in the spectrum were identified to fit successive loss of Gln18, Trp17 and Dhb16.  $[M' + 3H]^{3+} = 892.40^*$ .



These outcomes supplemented results from the previous section by supplying more data to corroborate the fact that all cysteine residues in the core peptide were involved in lanthionine cross-bridge formation. This also further verifies the activity of RumM in catalyzing PTMs installation in preRumA\* since one could accurately state that the cysteinyl thiols were inaccessible to the iodoacetamide derivatization because they were already blocked by thioether cross-linkages generated during the cyclization step of the biosynthesis pathway. Furthermore, when the same experiments were performed with unmodified His6-preRumA (section 4.2.4), all the cysteine residues were S-carboxymethylated (Figure 4.9).

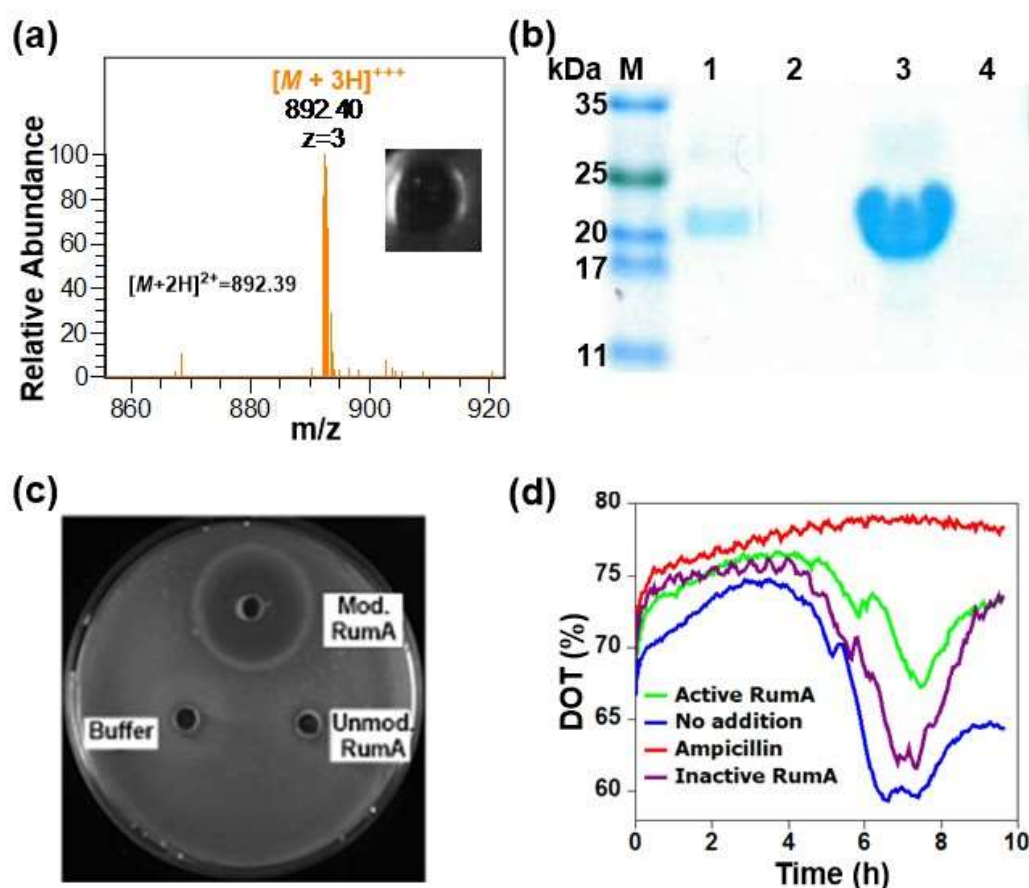
Results from trypsin digestion also supported what Dabard and colleagues already reported with regards to cleavage at Lys6. The mass spectrum of the trypsin-digested sample indicated charged ion species corresponding to the various tryptic peptides (Figure 4.23c). We were unable to identify fragments resulting from cleavage at Lys6, indicating that this position was not accessible to cleavage by trypsin as modifications at Thr7 forbid this. Fragmentation of the mature core peptide (RumA) precursor ion yielded a fragmentation spectrum with very little information due to low concentration. Nevertheless, we also deduced that the MS<sup>2</sup> spectrum contained characteristic peaks that were identical to those of preRumA\* (Figure 4.23d).

So far, we have shown that ruminococcin-A lanthionine synthetase RumM can catalyze *in vivo* installation of PTMs including 2 MeLan rings, one Lan ring and an  $\alpha,\beta$ -didehydrobutyrine into preRumA when the two proteins are coproduced simultaneously in *E. coli*. The data here demonstrates that a larger fusion partner to the N-terminus of the leader peptide does not interfere with *in vivo* processivity of RumM in catalyzing both the dehydratase and cyclization reactions. We also determined that the fully modified peptide host three thioether cross-bridges and possessed a specific lipid II-binding motif that is common to most class II peptides. This study, to the best of our knowledge reports the first successful heterologous production of fully modified lanthipeptide originally isolated from a strictly anaerobic Gram-positive microbe, in a Gram-negative host. With regards to product titre, we report here a total yield of approximately 6 mg of cyclic-modified preRumA per litre of *E. coli* culture. This may reduce to about 1-2 mg of pure active RumA per litre of culture when the leader peptide is removed since both the leader peptide and the core peptide have almost the same size with respect to the number of amino acid residues in their sequences. This yield is very important as it overwhelmingly exceeds that obtained earlier (Dabard et al., 2001) in the order of approximately 10<sup>4</sup>.

#### 4.4 Bioassay analysis of trypsin activated RumA

Biological assays were assembled as described in the experimental part (section 3.11). 50  $\mu$ l of the tryptic digest of preRumA\* containing the active RumA (section 4.3.8) was obtained and

dispensed into a punched well initially created in an agar plate spread with *Bacillus subtilis* ATCC 6633. The set-up was subsequently incubated overnight at 30 °C. Results in Figure 4.24a shows a distinct zone of inhibition, suggesting that the trypsin-digested sample contained a growth inhibitory component with activity against *B. subtilis* ATCC 6633. In this case RumA was the only possible suspect.



**Figure 4.24** In vitro activation of preRumA\* and bioassay. (a) Mass spectrum showing prominent ion peak corresponding to active RumA purified from the trypsin digest of cyclized preRumA\* (section 4.3.8), as well as the activity of the resulting product (b) SDS-PAGE analysis of IMAC-purified modified (lane 1) and non-modified (lane3); as well as trypsin-digested modified (lane 2) and non-modified (lane 4) His6-SUMO-preRumA\*. (c) The bioactivity of modified and non-modified trypsin digests against *B. subtilis* ATCC 6633 grown on a culture plate. (d) The effect of active RumA on *B. subtilis* ATCC 6633 liquid culture.

PreRumA\* was also fused to a SUMO tag as described in the experimental part (section 3.5.10.5). The resulting plasmids (pLEOsrA\* and pLEOsrA\*M) were used to transform *E. coli* W3110 and the resulting strains were cultivated in 100 ml TB medium. Purified products from both systems were digested with trypsin to yield a mixture of peptide fragments including

the putative mature RumA. One half of the digested products was used for SDS-PAGE (Figure 4.24b) while 50  $\mu$ l of the other half was pipetted into punched wells in an agar plate spread with *B. subtilis* ATCC 6633. Following an overnight incubation at 30 °C, a distinct zone of inhibition was observed for digested His6-SUMO-preRumA\* purified from the system expressing pLEOsrA\*M but no activity was apparent for extracts purified from the pLEOsrA\* expressing host (Figure 4.24c), suggesting that only the modified product has growth inhibitory activity. These results show that preRumA fused to the C-terminus of SUMO was also modified by RumM as does the GFP fusion constructs, indicating that SUMO tag may also be used as fusion partner instead of GFP. Although we did not perform MS characterization of products derived from the SUMO-fused constructs, the growth inhibitory activity demonstrated by the modified construct and none for the non-modified one allowed us to draw a positive conclusion with regards to the functionality of the system.

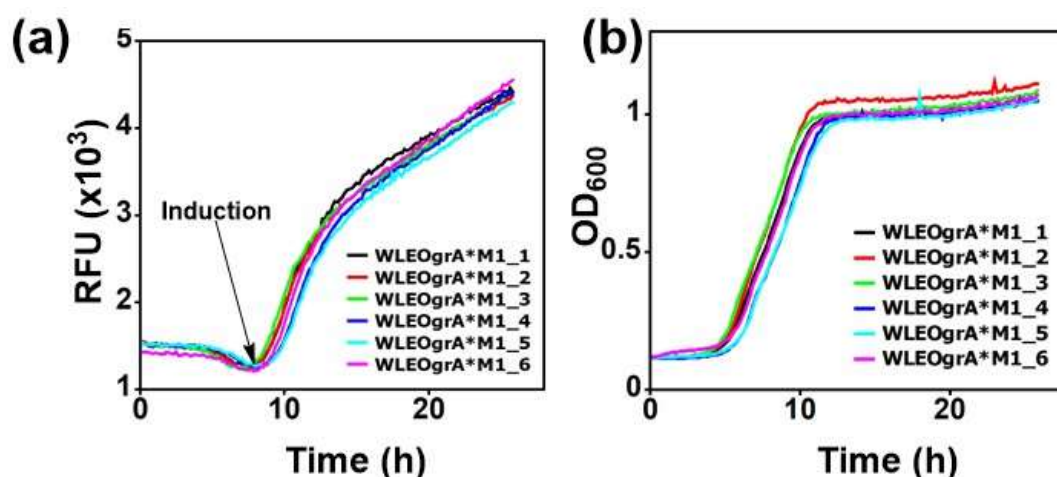
The ability of the tryptic digest of His6-SUMO-preRumA\* to inhibit growth of *B. subtilis* ATCC 6633 in liquid cultures was also evaluated by monitoring the DOTs of different cultures online using the PreSens OxoDish OD24. *B. subtilis* ATCC 6633 was grown to OD<sub>600</sub> of ~0.5 and then distributed into the 24 wells of the OxoDish. Ampicillin, modified and non-modified tryptic digests of His6-SUMO-preRumA\* were added to the wells in four replicates and the DOT of the respective wells were measured over time. Results show that the modified extract exerted a significant overall effect on growth while the non-modified construct fairly followed the DOT profile of the cultures without any supplement (Figure 4.24d). In the positive control, where the culture was supplemented with ampicillin, the DOT fairly remained constant throughout the cultivation, indicating no cell growth. This data suggests a simple method to screen for biological activity of several active compounds or RumA mutants online. This strategy may reduce the experimentation time and enable easy identification of interesting candidates.

We cannot say much about the mechanism of action of RumA at this time, but binding to lipid II and inhibiting cell wall biosynthesis (Götz et al., 2014), inhibition of spore proliferation (Gut et al., 2011), formation of pores in the phospholipid membranes (Breukink et al., 1999, Hasper et al., 2004, Hasper et al., 2006) and membrane disruption (Bakhtiary et al., 2017) are some of the general mechanisms of the antimicrobial activity of lantipeptides. Furthermore, RumA also possess a distinct lipid II-binding motif which has been shown in other class II lantibiotics to interact with the cell wall precursor without forming pores in the membrane (Islam et al., 2012) but rather induce intense cell wall stress responses in their targets (Sass et al., 2008).

## 4.5 Microtiter plate cultivations

### 4.5.1 Colony screening in microtiter plate using GFP as a reporter

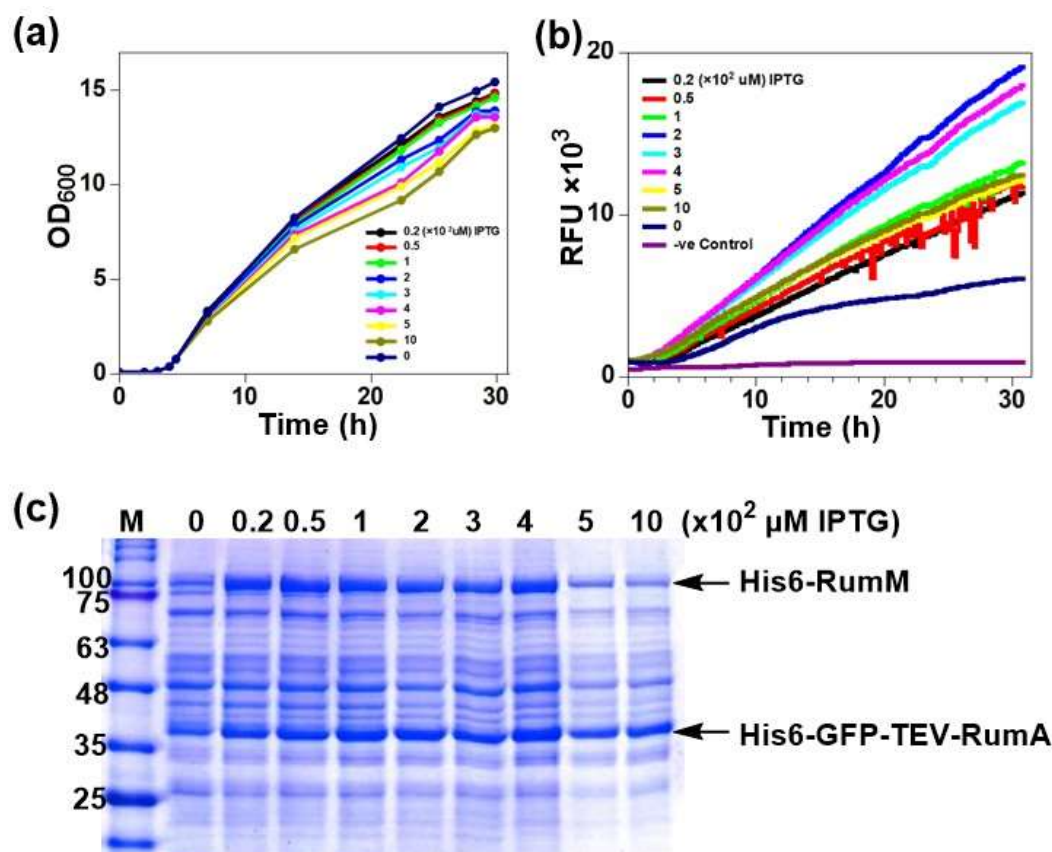
Besides colony PCR technique (section 3.5.4.3), online GFP fluorescence signals were also used to identify positive clones. In this procedure, single colonies were selected from LB agar plate containing *E. coli* Top10 transformed with pLEOgrA\*M1 and used for inoculating 96-well microtiter plate containing LB medium supplemented with the appropriate antibiotics. These experiments were performed using the Hamilton Microlab STAR Liquid-handling station as described in the methods part (section 3.6.3). The same amount of IPTG was automatically supplied to all wells at the point of induction, approximately 9 h after cultivation started. Results indicated that not all colonies produced GFP fluorescence (see Appendix 6.7), meaning that some of the strains growing on the agar plate did not harbour the correct vector or perhaps the gene of interest was wrongly inserted in the vector.



**Figure 4.26** Online GFP fluorescence and ODs of WLEOrA\*M1 cultivated in 96-well microtiter plate. (a) GFP fluorescence curves representing six of the cultures in the microtiter plate (b) Corresponding online OD curves of the cultures.

Since data from reaction kinetics studies of lanthipeptide synthetases revealed that these enzymes require enough time to catalyze PTMs formation in their substrate (Thibodeaux et al., 2014), having information about the stability of the substrate over the cultivation time may help to estimate the optimal duration of cultivation. We further selected the strain that demonstrated the most prominent characteristics in terms of growth and GFP fluorescence (see Appendix 6.7, Figure A8) and used it for further experiments. The pLEOgrA\*M1 plasmid was purified from this strain, sequenced and further transferred into *E. coli* W3110. In a similar procedure like the colony identification experiment, we isolated single colonies from LB agar plate containing the WLEOrA\*M1 culture and inoculated a 96-well microtiter plate. Results showed

that all colonies exhibited similar growth and fluorescence emission characteristics (see Appendix 6.7, Figure A9). Six of the 96 cultures are presented in Figure 4.26. One of the WLEOgrA\*M1 cultures was then randomly selected and used for subsequent experiments which have been described in the previous sections



**Figure 4.27** Performance of WLEOgrA\*M1 in TB medium. (a) Growth of TB medium cultures of WLEOgrA\*M1 induced with varying concentrations of IPTG. (e) GFP fluorescence signals of corresponding cultures. Concentrations of IPTG are indicated in the plots. (c) SDS-PAGE analysis of samples extracted from the cultures. IPTG concentrations are labelled at the top of each gel lane. For the ODs and fluorescence measurements, the cultures were set up in three replicates.

#### 4.5.2 Cultivation of WLEOgrA\*M1 in 24-well plate using TB medium

Microcultivations were performed in 24-well microtiter plate with varying IPTG concentrations (20 to 1000 μM). The growth of WLEOgrA\*M1 showed slight but consistent responses to IPTG concentrations (Figure 4.27a), recording a final OD<sub>600</sub> between 13 and 15 for all the cultures. The influence of IPTG on product formation (estimated by GFP fluorescence signal) was not uniform (Figure 4.27b). For instance, higher production rates were observed in cultures induced with intermediate IPTG concentrations in the range 200–400 μM, while cultures induced with higher IPTG concentrations like 500–1000 μM generated production rates

comparable to those induced in the lower concentration range (20-100  $\mu$ M IPTG). Thus, there was no direct correlation between growth and fluorescence intensities measured for the different cultures.

His6-RumM and His6-GFP-TEV-preRumA\* were highly produced (Figure 4.27c). No degradation of His6-GFP-TEV-preRumA\* was observed. We deduced from the multi-scale parallel screening that the concentration of IPTG necessary for optimal growth and expression lies between 50 and 100  $\mu$ M estimated by the thickness of bands on SDS-PAGE.

## 4.6 Strain optimization using multicultivation & screening strategies

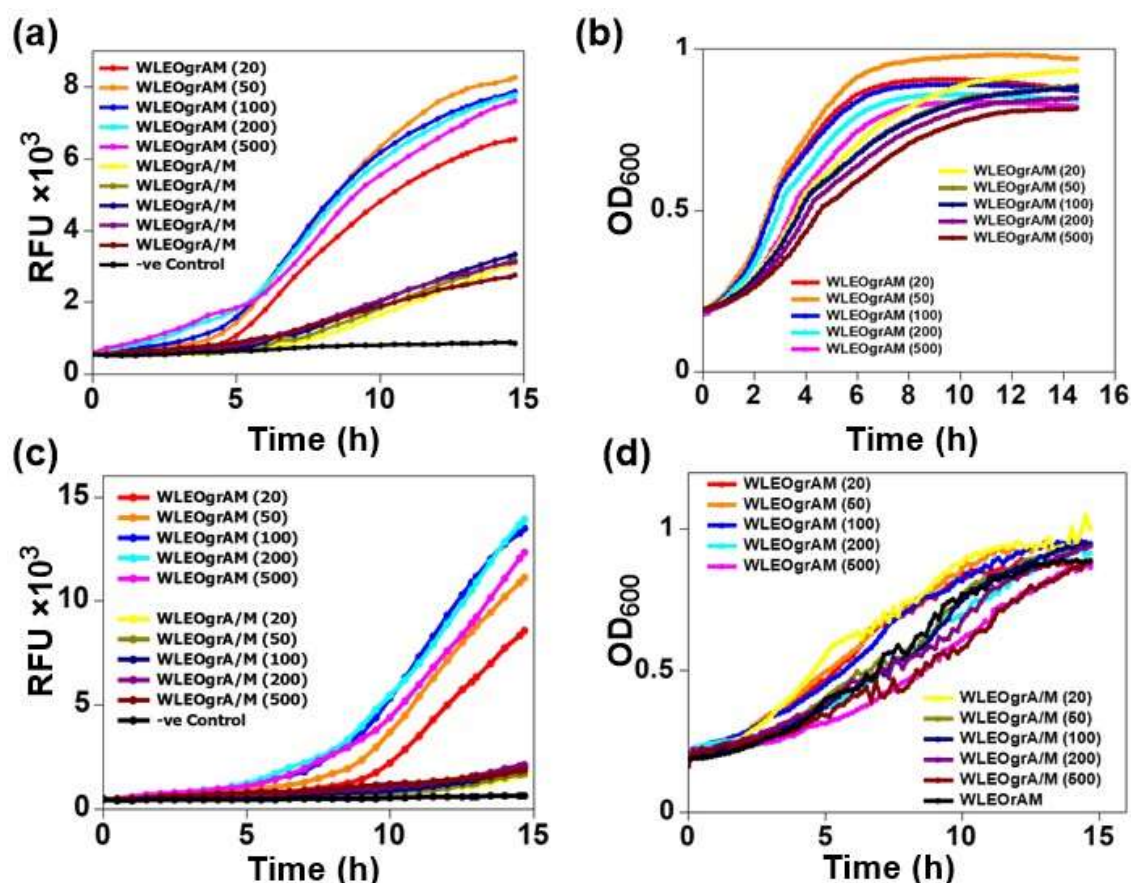
### 4.6.1 Characteristics of the bicistronic and two-vector systems

The strains WLEOgrAM (expressing the bicistronic vector pLEOgrAM) and WLEOgrA/M (harbouring pLEOgrA and pLEOrM) were cultivated in LB and the enzyme-based fed-batch EnPresso B medium. ODs and GFP fluorescence signals were measured online. Results indicated quite similar growth and production characteristics for all the cultures grown in the complex medium, with obvious and noticeable effects associated with variations in IPTG concentrations (Figures 4.28a and 4.28b). Although the WLEOgrA/M grew significantly, GFP fluorescence measurement remained comparatively low despite the fact that it was producing both His6-GFP-TEV-preRumA and His6-RumM like the WLEOgrAM system. The EnPresso cultivation system produced a quasi-linear growth characteristic for all cultures. Whereas intensive fluorescence signals were recorded for the WLEOgrAM strain, almost no fluorescence signal was noticeable with the WLEOgrA/M until about 8 h after induction (Figures 4.28c and 4.28d).

The results reported here indicate that the lower GFP fluorescence observed in the two-plasmid system may be attributed to metabolic burden incurred by hosting the two plasmids, and not dependent on the cultivation medium. Furthermore, production seems to reach maximum earlier in the complex medium cultures than in the defined EnPresso medium. This is understandable because complex media are deficient in divalent cations which are necessary to maintain membrane stability and cells grown in them are therefore susceptible to lysis (Wee and Wilkinson, 1988). These cations especially  $Mg^{2+}$  have very low content in LB broth: 30 to 40  $\mu$ M (Papp-Wallace and Maguire, 2008) and we cannot rule out the possibility that this may contribute to the early stationary phase observed with *E. coli* growth in LB cultures. Furthermore, *E. coli* grown in complex media is associated with alterations in carbon nutrition in which the easier-to-utilize ones get depleted easily and the cultured organism then



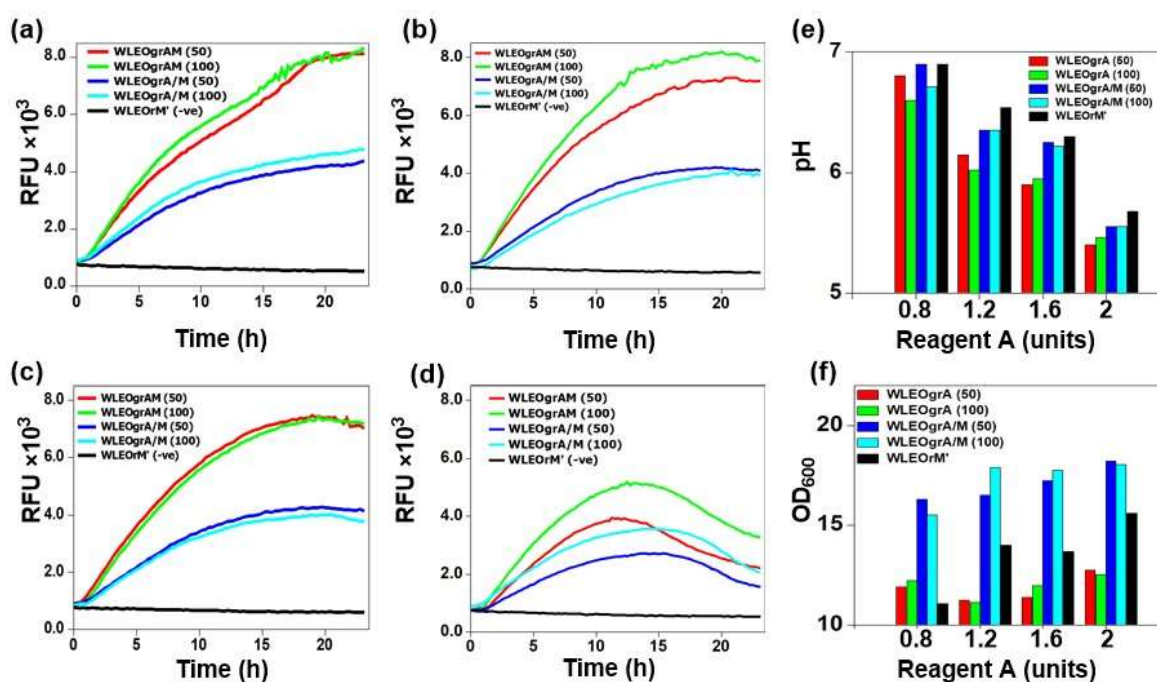
switches to the difficult-to-utilize carbon sources (Sezonov et al., 2007). These observations and the fact that the medium composition vary from batch-to-batch make LB broth very unreliable to obtain reproducible parameters in a miniature cultivation set-up that are amenable to large-scale production.



**Figure 4.28** Cultivation of WLEOgrAM and WLEOgrA/M in different media conditions. (a) Online GFP fluorescence signals measured in LB cultures. (b) Online ODs measured during LB cultivation. (c) Fluorescence signals measured in EnPresso B cultures. (d) Online ODs measured during EnPresso cultivation.

To investigate the reproducibility of the EnPresso cultures, we applied different glucose releasing rates by supplying varying amounts of reagent A to the cultures at the point of induction, without boosting. Panels (a) to (d) of Figure 4.29 clearly demonstrate that supplying an excess of glucose to the culture is detrimental to product formation. Most especially, when the amount of the reagent A was more than double the usual recommended amount of 0.8 units, production began to decline just about 12 h after induction and the final RFUs of the cultures decreased to about half the maximum attained during the entire period of cultivation. Moreover, the final pH of the cultures displayed a decreasing trend with increasing amount of

reagent A added (Figure 4.29e); meanwhile the final OD<sub>600</sub> increased as the amount of reagent A added to the culture also increased (Figure 4.29f).



**Figure 4.29** Effects of different glucose-releasing rates on the production of His6-GFP-TEV-preRumA by WLEOgrAM and WLEOgrA/M. Two separate IPTG-induction concentrations were used for each culture supplied with 0.8 unit (a), 1.2 units (b), 1.6 units (c) and 2 units (d) of reagent A, at the point of induction (No boosting). The final pH and OD<sub>600</sub> of the cultures were measured and plotted in (e) and (f) respectively. Values in brackets represent  $\mu$ M IPTG concentrations.

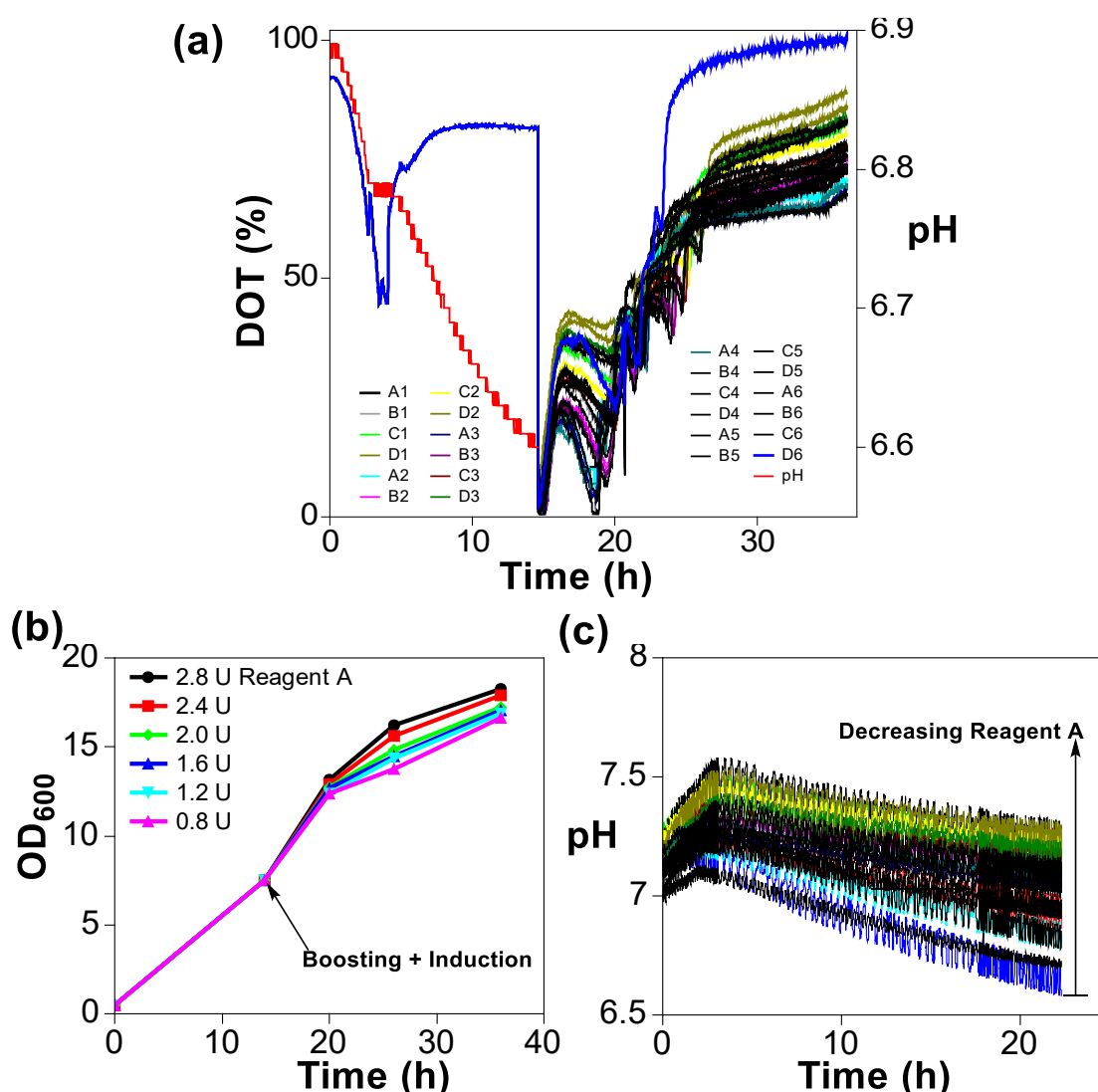
We also noticed that although the fluorescence signals produced by WLEOgrAM remained exceedingly higher in comparison to those produced by WLEOgrA/M, the final ODs of cultures growing the latter strain were more than 1.5 times higher than those growing the former. These outcomes indicated that the fluorescence signals were not dependent on the optical density of the culture as already suggested by results from TB medium cultures (section 3.6.2). Whereas reagent A appeared to cause only slight changes in the overall growth, its influence on the final pH of the cultures were rather significant. These aspects are explained further in the next section.



#### 4.6.2 Optimization of WLEOgrA\*M1 using process-relevant parameters

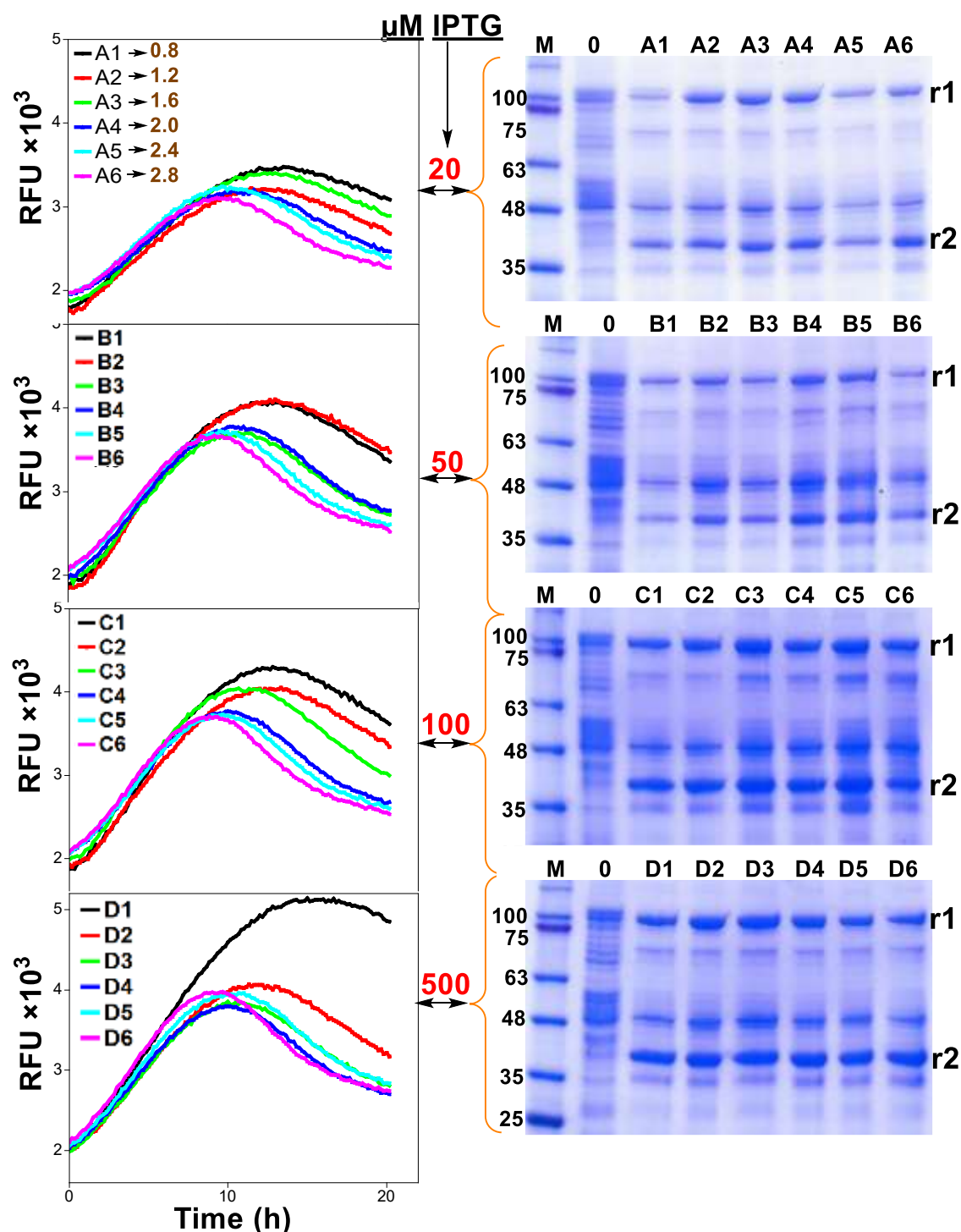
To evaluate the behavior of WLEOgrA\*M1 under process-relevant conditions, parallelized microtiter plates were used to screen process-relevant parameters like dissolved oxygen tension (DOT) and pH. In this case, the EnPresso B cultivation system was applied to study how DOT and pH may vary during a standard fermentation process. The WLEOgrA\*M1 strain was first cultivated in a PreSens shake flask containing an oxygen and a pH sensor integrated into the polycarbonate material at the bottom of the flask. The first 14 h of cultivation indicated a typical growth profile of *E. coli* (Figure 4.30a). During the first 5 h of cultivation, the cells took up free glucose in the EnPresso medium (free glucose is derived from nonspecific degradation of the complex polymer component of EnPresso) and exhibited characteristics that were typical for a batch cultivation, demonstrated by the constant drop in DOT until about 40 %. Within 2 h after this drop, the DOT rapidly increased to 75 % where it remained at a stable level. This rapid transition indicated a switch to glucose-limited growth. The pH decreased slowly throughout the cultivation period.

After 14 h of cultivation, the culture was split into three 24-well microtiter plate containing varying concentrations of IPTG and reagent A. Measurements were performed as described in the method part (see section 3.12). Plate 1 was used to monitor the DOT, plate 2 (24-well deep-well plate) to monitor pH changes and plate 3 to measure GFP fluorescence signals over time. The optical densities of cultures in the first row of plate 2 were also measured. At the start of plate measurements, all 24 cultures were nearly in an anaerobic state of growth, with initial DOT values around 0 %. This rapidly changed as DOT increased to about 25-40 % depending on the amount of reagent A that each of the cultures received (Figure 4.30a). This transient change occurred because there was no aeration when the culture was distributed into the wells of the plates, but aeration resumed once the plates were placed into the shaker. It is worthy to note that between 14 and 20 h, growth was rapid due to excess glucose present, supplied by the additional reagent A (Figures 4.30a and 4.30b). It seems that between 19.5 and 27 h most of the cultures experienced transient alterations between batch and glucose-limited growth before attaining a steady DOT level. The pH was controlled by nitrogenous sources in the booster tablets added to the culture prior to splitting in the plates. However, the pH profiles shown in Figure 4.30c indicated slight decreases with increasing amount of reagent A but maintaining the physiological range more tightly than the non-boosted cultures reported earlier (section 4.7.1).



**Figure 4.30** Screening fermentation-relevant parameters. (a) DOT of cultures before and after inductions, as well as pH before induction (red line). (b) Growth of cultures in the first row of the 24-well plate induced with 20  $\mu$ M IPTG (see Figure A10, Appendix 6.7) (c) pH of cultures after boosting/induction.

Whereas there appeared to be no significant difference in the behaviour of the different cultures with respect to DOT, culture in well D6 which received maximum amount of reagent A and the highest concentration of IPTG appeared to display a pronounced deviation (blue curve in Figure 4.30a). We also noticed that the pH of this particular culture was most of the time relatively low compared to the others during the cultivation (blue curve in Figure 4.30c). It is worth mentioning that increasing the amount of reagent A consequently increases the rate of glucose release into the medium and excess of the reagent also supplies excess glucose to the culture, causing acetate accumulation.



**Figure 4.31** Production optimization. The left panel indicates online GFP fluorescence measurements of WLEOgra\*M1 strain grown in the 24-well microtiter plate. Each graph represents a row on the plate and the well numbers are indicated in the graphs. The concentration of reagent A (in Units) added to each well of a row from left to right, are labelled in the top-left graph. The IPTG concentrations added to the wells in each row are indicated by the double arrows in between the graphs and the SDS gels. The right panel shows the SDS-PAGE analyses of corresponding His6-RumM (r1) and His6-GFP-TEV-preRumA (r2) extracts obtained at the end of the cultivation. See Figure A10 in the Appendix for reagent A concentrations added to each well.

Furthermore, applying high concentrations of IPTG to induce a recombinant gene strongly inhibits *E. coli*'s glucose uptake capacities and their abilities to respire optimally leading to formation and accumulation of acetate (Neubauer et al., 2003). Accumulation of excess acetate may inhibit growth of *E. coli* (Roe et al., 2002, Neubauer et al., 2003), even though the organism may later oxidize the organic acid as an alternative carbon source. Acetate accumulation may explain why the pH of the cultures decreased with increasing concentration of reagent A (Figure 4.30a). These results further corroborated those reported in section 4.7.1.

Examining the GFP fluorescence signals recorded for all the 24 cultivation conditions, we could easily identify that product formation was greatly influenced by the IPTG concentration. The trend is visible as one navigates vertically through the wells in each column from A to D as depicted in the left panel of Figure 4.31. Conversely, increasing the concentration of reagent A resulted in a significant decrease in fluorescence signals starting 10 h after induction. Extracts from all the wells were purified with HisTrap spin columns and analyzed via SDS-PAGE. Results show that the production levels of targets increased with increasing IPTG concentrations (right panel, Figure 4.31). The inducer concentration increases vertically from the gel at the top to the bottom. Interestingly, we also observed the same trend on SDS-PAGE as for the measured GFP fluorescence signals.

The final product extracted from the wells demonstrated higher production levels if more reagent A was added but excess amount of the reagent (like in column 6 of the plate) instead negatively affected expression especially in cultures where the concentration of the inducer was low (rows A and B: 1<sup>st</sup> and 2<sup>nd</sup> gels from top-right panel). However, we took notice of the fact that some of the end GFP fluorescence signals do not directly correlate to the final amount of products (estimated by the thickness of bands on SDS-PAGE). Therefore, *in vivo* GFP fluorescence may be used to qualitatively determine expression but may not be applied as a quantitative method to assess product formation. Nevertheless, it would be interesting to determine product formation at the point where the fluorescence signals begin to decline (around 18 h) to ascertain if the same conclusion can still be drawn in that situation.

Results here demonstrate that supplying excess glucose to the EnPresso B culture increases growth rate and consequently produces higher biomass yields, but the same procedure has contrary effects with regards to target protein expression levels. If the fluorescence results were to be taken as a standard, then cultivation should be stopped after 10 h induction since product formation begins to decline. However, this would imply producing a mixture of partially modified product since the modifying enzyme may require longer time to catalyze formation of PTMs (Thibodeaux et al., 2014, Repka et al., 2017). Our results show that an IPTG

concentration of 100  $\mu$ M in combination with up to 1.6 units of reagent A (i.e. double amount of the recommended standard) added to the cultures at the point of induction/boosting is safe to ensure optimal production of His6-GFP-TEV-RumA\* and His6-RumM using strain WLEOgrA\*M1. Our data also present growth and production characteristics which may be easily adapted to larger scale cultivations.

## 4.7 Overall discussion of results

In this work, we isolated the biosynthetic machinery of the lanthipeptide ruminococcin-A from the natural producer and heterologously reconstituted the pathway in *E. coli*. We successfully expressed and modified the peptide as a chimeric fusion product together with GFP. Fusing the structural gene for the precursor peptide to the gene encoding GFP and co-expressing the chimeric construct simultaneously with the dedicated lanthionine synthetase RumM did not alter the formation of thioether cross-bridges. This was surprising since installation of PTMs in the core peptide by RumM requires the N-terminal leader sequence as a docking site to direct its activity. One may expect that employing a larger fusion partner as such would influence the activity of the enzyme but this appears not to be the case. We obtained fully modified peptide possessing thioether bridges and an additional dehydroamino acid residue. The results of this study have been submitted and is currently under consideration for publication (Ongey et al., 2018). This study, to the best of our knowledge, is the first reported case illustrating that the precursor peptide can still be correctly modified with such a large fusion partner to the leader sequence. The data reported herein supplies an alternative experimental design to gain more insights into mechanistic events that drive the generation of MeLan/Lan rings as well as dehydroamino acids in lanthipeptides.

One of the main objectives of this study was to improve the production strategy thereby enabling further biotechnological development and engineering possibilities. The current system recorded immense performance in the regard. The amount of modified precursor peptide obtained was 6 mg per litre of *E. coli* culture. Upon removing the leader peptide and subsequent purification of the active RumA, we expect to record a yield ranging between 1 and 2 mg of pure product, taking into consideration the size ratio of the leader-to-core segment (~1:1). This product concentration surpasses the amount of RumA initially purified from the natural source by several thousand-fold. Dabard et al. (2001) only succeeded in achieving a yield of 0.665 microgram ruminococcin-A per litre of *R. gnavus* E1 culture (Dabard et al., 2001). Perhaps the nature of *R. gnavus* E1 with respect to growth and optimization contributed largely to this outcome. It is improbable that regulating cultivation parameters and engineering

cultivation media would produce a meaningful improvement. Assuming that would be the case, the overall production cost would be skyrocketed by additional expenses incurred by supplementing the cultivation medium with trypsin and hemin.

Being able to produce the peptide now in *E. coli* which is simple, easy to manage and have a fast generation time, we expect that transferring the current system to process scale production would require less time and efforts. It may be worthwhile considering other approaches like performing further optimization including strain engineering or employing the use of smaller fusion partners to achieve higher production yields like other systems reported in the literature (Shi et al., 2011, Shi et al., 2012, Tang and van der Donk, 2012, Kuthning et al., 2015). Employing a single-plasmid tricistronic operon to simultaneously express preRumA, RumM and RumT, may enable biosynthesis and secretion of the peptide. Of course this would reduce some steps in the purification procedure that may consequently produce a positive impact on the yield. Our study did not include such approach because we were unable to express RumT separately. Nevertheless, it is possible that if there exist a biosynthesis complex (involving all three components) like the one proposed herein, it may mask the toxicity of RumT, allowing the peptide to be produced and exported. However, for the purpose of studying peptide engineering and LanM catalysis, the current system represents a sufficient tool. Additionally, expanding and/or increasing the antimicrobial activity of RumA or other related compounds may be easier to apply with the present system. An example of such approach includes the use of amber stop codon suppression to incorporate non-canonical amino acids into the peptide (Steiner et al., 2008, Chatterjee et al., 2012, Oldach et al., 2012).

With respect to structural characterization of modified preRumA\*, data presented herein are consistent with those obtained by Dabard and coworkers, with the exception of one additional methyllanthionine ring between Thr7 and Cys12 which was never reported. In fact, Dabard et al. (2001) concluded in their study that the MeLan ring formation between these two residues was not possible. Nevertheless, they arrived at this conclusion using information derived from Edman degradation assay. However, the Edman sequencing method is not very accurate in studying the structure of lanthipeptides because of limitations associated with the fact that any modification at the N-terminus of the peptide sequence will block the experiment. Additionally, Edman degradation is blocked by Lan/MeLan rings as well as Dhb/Dha residues, and no sequence information can be obtained thereafter (Lohans and Vederas, 2014). We may conclude with certainty that the structure described herein represents the actual structure of modified ruminococcin-A precursor peptide because the third ring identified herein also constitutes the core active feature referred to as the mersacidin-lipid II-binding motif which is common in all class II lanthipeptides (Knerr and van der Donk, 2012).

This implies that in *E. coli*, the dedicated ruminococcin-A lanthionine synthetase RumM is capable of catalytically installing three thioether cross-linkages and one additional  $\alpha$ ,  $\beta$ -unsaturated amino acid into the core structure of preRumA. The dehydratase domain of RumM catalyzes the dehydration of Thr7, Thr16 and Thr22 to three Dhb, and Ser9 to Dha. Subsequently, the cyclase domain engages in a Michael-type addition-cyclisation reactions involving Dhb7 and Dhb22, and activated thiol groups of Cys12 and Cys24 to produce two MeLan rings, while the Dha9 and Cys23 produce a Lan ring. We have proposed mechanisms for RumM-catalyzed reactions which may be studied further. For instance, utilizing mutagenesis studies in conjunction with mass spectrometry and/or bioassay analyses may be employed to characterize the two domains of RumM. Additionally, the crystal structure of the full-length enzyme may be more resourceful.

In this study, we initially considered using two plasmids to separately control biosynthesis of the precursor peptide preRumA and the modifying enzyme RumM. We designed the plasmids so that *rumM* expression was controlled by a weaker *CU* promoter while *rumA* was under control of the stronger *CTU* promoter. This strategy was conceived to limit metabolic burden and to increase cellular availability of the unmodified preRumA substrate for RumM to catalyze PTMs formation. However, this approach failed to produce the desired modifications in preRumA. Although just recently, Basi-Chipalu and coworkers employed a similar approach to modify pseudomycoicidin in *E. coli* (Basi-Chipalu et al., 2015), we cannot say exactly why preRumA was not modified when His6-preRumA was coexpressed simultaneously with His6-RumM. We anticipate that the enhanced solubility (and no modification) of preRumA in the WLEOrA/M strain was facilitated by simple binding interactions between His6-preRumA and His6-RumM since interactions between nonmodified peptide precursor and its modifying enzyme have been reported (Lubelski et al., 2009, Khusainov et al., 2011, Mavaro et al., 2011, Repka et al., 2017).

Furthermore, a mixture of modified and non-modified preRumA were identified in extracts obtained from strain WLEOgrA/M (expressing His6-GFP-TEV-preRumA and His6-RuM on separate plasmids), and non-modified preRumA was not apparent in extract from WLEOgrA (expressing His6-GFP-TEV-preRumA alone), but rather truncated products of preRumA. These outcomes also obliged the reasoning that RumM may have fostered some form of molecular interactions that in turn conferred a stabilizing role on the precursor peptide since it is believed that the enzyme forms a complex with the precursor peptide (Lubelski et al., 2008, Lubelski et al., 2009). Existing data also suggest a constant interaction with both the non-modified and the modified substrate (Yang and van der Donk, 2015). This Result further support our suggestion that biophysical interactions between His6-preRumA and His6-RumM may have facilitated the solubility of the former without necessarily introducing the desired

modifications in the peptide. Nevertheless, mixtures of partially dehydrated products are possible since investigations revealed that the coupling of cysteinyl thiol to Dha/Dhb via Michael addition cyclization can prevent further modification of serine and/or threonine residues (Kuipers et al., 2008, Lubelski et al., 2009). However, we cannot say whether the intermediate products observed contained at least one of the rings or not. Nevertheless, the data presented here show that low levels of RumM may have been responsible for the inefficiency of the WLEOgrA/M system.

The amount of enzyme produced by the weakly expressing vector may not be sufficient for complete modification of the precursor peptide, which may explain why unmodified preRumA was identified in extracts from the strain WLEOgrA/M. This appeared to be the case as this drawback was mitigated by reengineering the host vector to favour adequate production of RumM. It is important to note that applying the single-plasmid bicistronic expression vector system did not improve the quality of His6-GFP-TEV-preRumA when RumM was expressed under control of the *CU* promoter and RBS from the lactose operon. However, by replacing the RBS from the lactose operon with that of gene 10 of bacteriophage T7, RumM expression was enhanced and the degradation problem was solved. In fact the WLEOgrA\*M1 strain produced only minute traces of three-fold dehydrated peptide and no unmodified peptide compared to WLEOgrA/M. Therefore, it is important to ensure adequate expression of the modifying enzyme when trying to develop a combinatorial system for the production of lanthipeptides in *E. coli*.

The strong interactions observed between His6-GFP-TEV-preRumA\* and His6-RumM may be further investigated to supply more insights into the nature of complexing. It may be worthwhile to crystallize the complex to provide data required to characterize these complexes, which may reveal some structural features in LanM enzymes yet to be known. Our data suggests that dehydrated preRumA\* may be predominant in the His6-GFP-TEV-preRumA\*/His6-RumM complex since another report showed that the interactions between modifying enzyme and dehydrated precursor peptide is much stronger (Mavaro et al., 2011). This may be true since a LanM enzyme was shown to catalyze reversible opening of the thioether ring (Yang and van der Donk, 2015) and thus, RumM may constantly supply the dehydrated precursor, and subsequently engages the free substrate to form complexes.

Lastly, our optimization experiments produced results to demonstrate that supplying excess of glucose to *E. coli* cultures expressing the chimeric GFP fusion product simultaneously with the modifying enzyme, may produce a positive response on growth, but constitute counter effects with regards to target protein expression levels. Furthermore, our data indicated that strength



of GFP fluorescence signal may not be employed as a tool to determine when to end the cultivation as the signals start decreasing at a point when it is not expected that full modifications must have occurred in the preRumA. Additionally, our results indicated an optimal IPTG concentration in the range of 100-200  $\mu$ M for expression of His6-GFP-TEV-RumA\* and Hs6-RumM using the strain WLEO*grA*\*M1. These data may be easily adapted to larger scale cultivations.

## 5. Conclusions and outlook

### 5.1 Conclusions

We successfully constructed the biosynthesis pathway of ruminococcin-A in *E. coli* and showed for the first time that fusing the precursor peptide to a larger carrier like GFP and subsequent coexpression with the lanthionine synthetase resulted in high-level *in vivo* production of the active compound. Since the modifying enzyme requires the leader peptide as a docking site to direct its activity on the core structure, our results demonstrate that a larger attachment to the N-terminus of the leader sequence does not hamper the activity of the lanthionine-generating enzymes. Interactions between the modifying enzyme and the GFP fusion construct yielded some insights into the mechanism of RumM catalysis. No report exist that describes the fusion of large protein carriers to the N-terminus of the leader peptide and so the current system may be helpful in providing further mechanistic insights on the modification machinery. Our data further indicated that the fully modified compound possess a total of three thioether cross-bridges (two MeLan and one Lan) instead of two as previously reported. The third ring identified in this study actually constitutes the mersacidin-lipid II-binding motif which is common in all class II lanthipeptides.

Sequence–structure–function relationship studies showed that lanthionine-generating enzymes of class II lanthipeptides share distinct domain homologies within their respective family members, with the C-terminal domain of RumM sharing >60 % similarity with the class I nisin cyclase. In the absence of RumM, the precursor peptide preRumA was largely expressed as insoluble product but regained its solubility when it was coexpressed simultaneously with RumM. However, modifications were observed only in preRumA expressed as a fusion partner to a larger protein tag and not to His-tag alone.

Microscale cultivations enabled us to determine the optimal growth and expression conditions that were easily adapted to larger scale cultivations. We attained a product yield of approximately 6 mg of cyclized preRumA per litre of *E. coli* culture. This yield corresponds to the modified precursor peptide and we estimated that upon removing the leader peptide and subsequent purification of the active product, this amount may reduce to about 1-2 mg of pure and active RumA per litre of culture. This is because the leader peptide and the core peptide share a ratio of 1:1 with respect to their average molecular weights. The yield reported here overwhelmingly exceeds values reported earlier by  $10^4$  order. Alternative strategies were used to remove the leader peptide and consequently activating preRumA. Growth inhibitory activity

of trypsin-activated RumA against *B. subtilis* ATCC 6633 was achieved, suggesting that the modified product possess biological activity.

The heterologous production system developed herein, offers irrefutable advantages over *R. gnavus* E1 (native host) since scale-up studies and reaching economically feasible production are practicable. Complicated and expensive experimental set-ups are not required. Furthermore, this work opens up prospective studies in the area of peptide engineering aimed at expanding and/or increasing the antimicrobial activity of RumA or generating novel compounds with enhanced biological functions.

## 5.2 Outlook

The system herein may be exploited as a new generational approach to identify and characterize new peptide variants with improved therapeutic efficiency via peptide engineering. This is possible by combining specific mutations in the peptide together with the promiscuous catalytic properties of the class II lanthionine-generating enzyme RumM. Other tailoring modifications like the N-glycosylation and non-canonical amino acids may also help to increase stability and enhance biological functions by providing a broad distribution of chemical functionalities that would strengthen interactions with biological target components. For example, specific residues in the core peptide preRumA may be replaced by unnatural amino acids like those containing  $\alpha$ -haloacetamide via genetic code engineering to allow expansion of the chemical reactivity space and broaden the activity spectrum of the peptide.

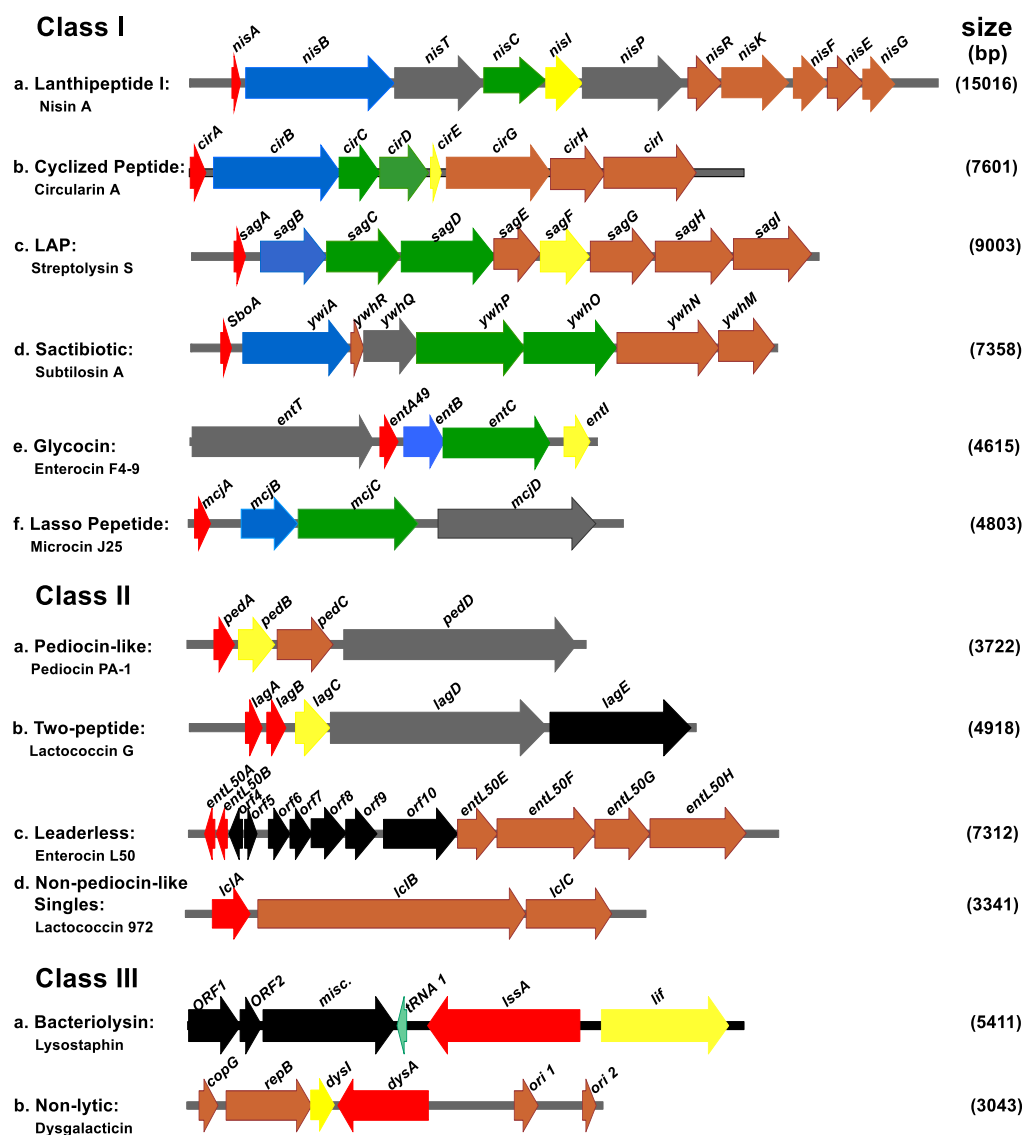
This study also opens avenues for large-scale studies and to investigate the mode of action of RumA since we now have the peptide in an easy-to-cultivate host and preliminary assessment of the production level seems very encouraging. Some vectors that were developed during this work were not fully characterized e.g. pLEOrA<sup>TEV</sup>M1 and pLEOsrA<sup>Xa</sup>M1. It would be interesting to further evaluate these constructs containing TEV or Factor Xa sites at position -1 of preRumA in order to evolve with a system that would facilitate purification of the active peptide.

We have suggested a couple of mechanisms like those involving the dehydratase and cyclase domains of RumM that require empirical evaluation by conducting the appropriate experiments. Additionally, we postulate that the biosynthesis and export of RumA in *R. gnavus* E1 require the synergistic involvement of RumM and RumT functioning as a complex machinery during the maturation process. Adequate information is supplied that may be used to design *in vivo* and *in vitro* experiments to evaluate these concepts. Further optimization, expression and characterization of the peptidase domain of RumT would be important to add to the growing list of characterized class II LanTs. Additionally, engineering a polycistronic expression vector

that simultaneously expresses preRumA, RumM and RumT may foster *in vivo* expression and secretion of active RumA that may easily be purified from the culture broth.

## 6. Appendix

### 6.1 Structure-function analyses, description of the biosynthesis operons, classification and physicochemical characteristics of randomly selected bacteriocins



**Figure S1.1 Schematic organization of biosynthesis gene clusters involved in production bacteriocins (drawn to scale).** Bacteriocins structural genes, red; modification genes, blue & green; maturation and transporter genes, dark grey; light green, immunity genes, yellow; accessory/secretory or other immunity genes, brown; genes with unknown functions, black.

**Table S1** Sources and physicochemical characteristics of selected examples of different classes of bacteriocins

Class of bacteriocin	Example	Native producer	AA residues	MW (kDa)	Reference
<b>Ia</b> <b>(Lanthipeptides)</b>	Elgicins	<i>Paenibacillus elgii</i> B69	45	4.5 – 4.8	(Teng et al., 2012)
	Nisin	<i>Lactococcus lactis</i>	34	3.4	(Field et al., 2015)
	BacCH91	<i>Staphylococcus aureus</i> CH-91 DSM 26258	22	2.07	(Wladyka et al., 2013)
	NAI-107	<i>Microbispora</i> sp.	24	2.2	(Castiglione et al., 2008)
	Haloduracin	<i>Bacillus halodurans</i> C-125	28; 30	2.3; 3.0	(McClerren et al., 2006)
	Pseudomycoicidin	<i>Bacillus pseudomycooides</i> DSM 12442	26	2.7	(Basi-Chipalu et al., 2015)
	Ruminococcin-A	<i>Ruminococcus gnavus</i> E1	24	2.6	(Dabard et al., 2001)
<b>Ib</b> <b>Circular peptides</b>	Amylocyclin	<i>Bacillus amyloliquefaciens</i> FZB42	64	6.3	(Scholz et al., 2014)
	Circularin A	<i>C. beijerinckii</i> ATCC 25752	69	7.6	(Kemperman et al., 2003)
	Enterocin 4	<i>E. faecalis</i>	70	7.1	(Joosten et al., 1996)
	Butyrivibriocin AR10	<i>Butyrivibrio fibrisolvens</i>	51	6.0	(Kalmokoff et al., 2003)
	Subtilisin	<i>B. subtilis</i>	32	3.4	(Babasaki et al., 1985)
	Uberolysin	<i>S. uberis</i>	70	7.0	(Wirawan et al., 2007)
	Garvicin ML	<i>L. garvieae</i>	60	6.0	(Borrero et al., 2011)
	Leucocyclin Q	<i>Leuconostoc mesenteroides</i> TK41401	61	6.1	(Masuda et al., 2011)
	Carnocyclin A	<i>Carnobacterium maltaromaticum</i> UAL307	60	5.8	(Martin-Visscher et al., 2008)
	Gasserin A	<i>Lb. gasseri</i>	58	5.6	(Kawai et al., 1998, Kawai et al., 2001)
	AS-48	<i>E. faecalis</i>	70	7.1	(Gálvez et al., 1986)
	Ent53B	<i>E. faecium</i>	64	6.3	(Himeno et al., 2015)
	Plantazolicin A	<i>B. amyloliquefaciens</i> FZB42	14	1.3	(Scholz et al., 2011)

<b>Ic LAPs</b>	Microcin B17		43	3.1	(Baquero and Moreno, 1984)
	Streptolysin S	Group A streptococcus (GAS)	30	2.7	(Nizet et al., 2000)
	azolemycins A–D	<i>Streptomyces sp. FXJ1.264</i>	7	0.68	(Liu et al., 2015)
	Goadsporin	<i>Streptomyces sp. TP-A0584</i>	19	1.6	(Igarashi et al., 2001)
<b>Id Sactibiotics</b>	Hyicin 4244	<i>Staphylococcus hyicus 4244</i>	35	<10	(de Souza Duarte et al., 2017)
	Subtilisin A	<i>B. subtilis 168</i>	32	3.3	(Babasaki et al., 1985)
	Thuricin CD	<i>B. thuringiensis DPC 6431</i>	30 & 31	2.7, 2.8	(Rea et al., 2010)
	Thurincin H	<i>Bacillus thuringiensis SF361</i>	31	3.1	(Lee et al., 2009)
	Propionicin F	<i>P. freudenreichii</i>	43	4.3	(Brede et al., 2004)
	Glycocin F	<i>Lactobacillus plantarum KW30</i>	43	5.2	(Stepper et al., 2011)
<b>Ie Glycocins</b>	Thurandacins A & B	<i>Bacillus thuringiensis serovar andalousiensis BGSC 4AW1</i>	42	~10.5 – 11.0	(Zwick et al., 2012, Wang et al., 2013)
	enterocin F4-9	<i>Enterococcus faecalis F4-9</i>	47	5.5	(Maky et al., 2015)
	Sublancin 168	<i>B. subtilis</i>	37	3.8	(Paik et al., 1998, Garcia De Gonzalo et al., 2014)
	capistruin	<i>Burkholderia thailandensis E264</i>	19	2.04	(Knappe et al., 2008)
<b>If Lasso peptides</b>	Sviceucin	<i>Streptomyces svicens ATCC 20983</i>	20	2.08	(Kersten et al., 2011, Li et al., 2015)
	Lassomycin	<i>Lentzea kentuckyensis sp</i>	16	1.8	(Gavrish et al., 2014)
	Microcin j25	<i>Escherichia coli</i>	21	2.14	(Bayro et al., 2003)
	cattlecin	<i>Streptomyces cattleya</i>	20	2.68	(Sugai et al., 2017)
	Pediocin PA-1	<i>Pediococcus acidilactici</i>	44	4.6	(Chikindas et al., 1993)
<b>Ila Pediocin-like peptides</b>	Enterocin A	<i>Enterococcus faecium</i> ,	47	4.8	(Aymerich et al., 1996)
	Bacteriocin T8	<i>Enterococcus faecium T8</i> ,	36	4.2	(De Kwaadsteniet et al., 2006)
	Bacteriocin 31	<i>Enterococcus faecalis</i>	43	4.9	(Tomita et al., 1996)
	Enterocin P	<i>Enterococcus faecium</i>	44	4.6	(Cintas et al., 1997)
	Leucocin A-UAL 187	<i>Leuconostoc gelidum</i> .	37	3.9	(Hastings et al., 1991)

<b>IIb</b> <b>Two-peptide</b> <b>bacteriocins</b>	Maltaricin CPN	<i>Carnobacterium maltaromaticum</i> CPN	44	4.4	(Hammi et al., 2016)
	Mesentericin Y105	<i>Leuconostoc mesenteroides</i>	36	2.5 – 3.0	(Hécharde et al., 1992)
	Enterocin HF	<i>Enterococcus faecium</i> M3K31	43	4.3	(Arbulu et al., 2015)
	Sakacin P	<i>Lb sake</i> Lb706	41	4.4	(Tichaczek et al., 1992)
	Curvacin A	<i>Lb curvatus</i> LTH1174	38-41	4.3	(Tichaczek et al., 1992)
	Lactococcin G	<i>Lc lactis</i>	α;39	4.3	(Nissen-Meyer et al., 1992)
	Plantaricin EF	<i>Lb plantarum</i> C11	β;35	4.1	(Diep et al., 1996, Ekblad et al., 2016)
			PlnE;32	3.5	
	Thermophilin 13	<i>Streptococcus thermophilus</i>	PlnF;34	3.7	(Marciset et al., 1997)
			ThmA;62	5.7	
	Plantaricin JK		ThmB;43	3.9	(Diep et al., 1996)
			PlnJ;25	2.9	
	Enterocin C	<i>E. faecalis</i> C901	PlnK;32	3.5	(Maldonado-Barragán et al., 2009)
			α;39	4.2	
	lactococcin Q	<i>Lc. lactis</i>	β;35	3.8	(Zendo et al., 2006)
			α;39	4.2	
	Enterocin X	<i>E. faecium</i> KU-B5	β;35	4.0	(Hu et al., 2010)
			α;40	4.4	
	Carnobacteriocin XY	<i>Carnobacteria</i>	β;37	4.0	(Acedo et al., 2017)
			X; 33	3.5	
<b>IIc</b> <b>Leaderless</b> <b>bacteriocins</b>	Aureocin A53	<i>Staphylococcus aureus</i> A53	51	6.0	(Netz et al., 2002)
	Aureocin A70	<i>Staph. aureus</i>	AurA: 31	2.92	(Netz et al., 2001)
			Aur B: 30	2.7	
			AurC: 31	2.95	
			AurD:31	3.0	
	Enterocin L50	<i>E. faecium</i>	L50A: 44	5.19	(Cintas et al., 1998)
			L50B: 43	5.17	
	Enterocin Q		34	3.9	(Cintas et al., 2000)
	Lacticin Q	<i>Lc. lactis</i>	53	5.9	(Fujita et al., 2007)
	LsbB	<i>Lc. lactis</i>	30	3.4	(Gajic et al., 2003)
	Garvieacin Q	<i>L. garvieae</i> BCC 43578	70	5.3	(Tosukhowong et al., 2012)
	Enterocin EJ97	<i>E. faecalis</i>	44	5.3	(Gálvez et al., 1998)
	Garvicin A	<i>L. garvieae</i>	44	4.6	(Maldonado-Barragán et al., 2013)
	Enterocin K1	<i>E faecium</i> EnGen0026	37	-	(Ovchinnikov et al., 2017)



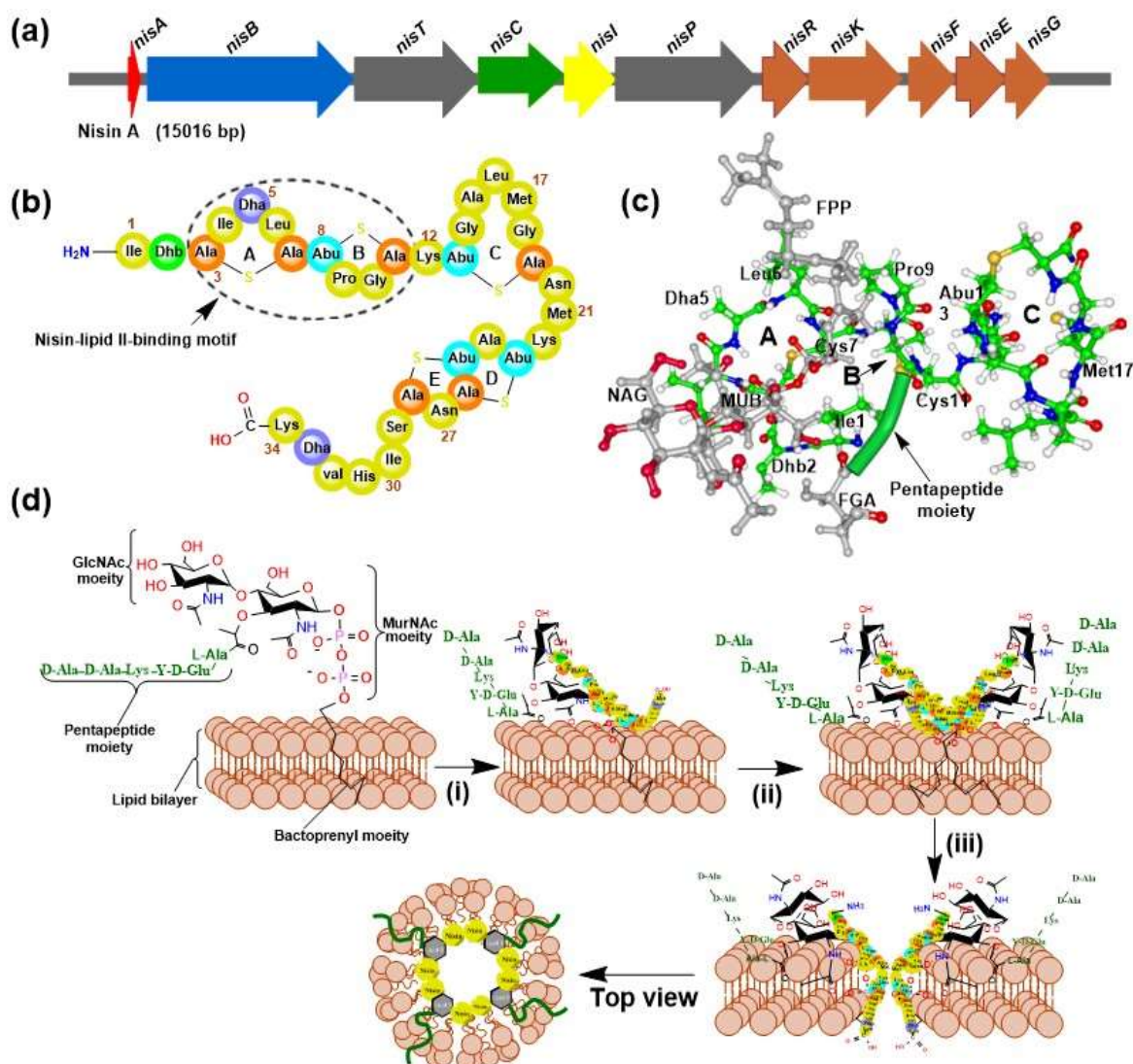
<b>II d</b> Non-pediocin-like single peptides	Laterosporulin	<i>Brevibacillus sp.</i> <i>strain GI-9.</i>	50	5.6	(Singh et al., 2012, Singh et al., 2015)
	Lactococcin A	<i>Lactococcus lactis</i> <i>subsp. cremoris LMG</i> <i>2130</i>	54	5.7	(Holo et al., 1991)
	Lactococcin 972	<i>Lactococcus lactis</i> <i>IPLA 972</i>	66	7.5	(Martínez et al., 1999)
	Lactococcin B	<i>Lactococcus lactis</i> <i>subsp. cremoris 9B4,</i>	47	5.3	(van Belkum et al., 1992)
	Enterocin B	<i>Enterococcus</i> <i>faecium T I 36</i>	53	5.4	(Casas et al., 1997)
	Microcin V	<i>E. coli</i>	88	8.7	(Håvarstein et al., 1994)
	Epidermicin NI01	<i>Staphylococcus</i> <i>epidermidis strain</i> <i>224</i>	51	6.0	(Sandiford and Upton, 2012)
	Lacticin Q	<i>Lactococcus lactis</i> <i>QU 5</i>	53	5.9	(Fujita et al., 2007)
<b>III</b> <b>Large heat-labile</b> <b>bacteriocins</b>	<b>Lytic</b>				
	Enterolysin A	<i>E. faecalis LMG</i> <i>2333</i>	316	34.5	(Nilsen et al., 2003)
	Zoocin A	<i>Streptococcus equi</i> <i>subsp.</i> <i>zooepidemicus</i>	262	27.9	(Simmonds et al., 1997)
	Millericin B	<i>Streptococcus milleri</i> <i>NMSCC 061</i>	259	28.	(Beukes et al., 2000)
	Lysostaphin	<i>Staphylococcus sp.</i>	246	~27	(Schindler and Schuhardt, 1964, Bastos et al., 2010)
	<b>Non-lytic</b>				
	Dysgalactacin	<i>S. dysgalactiae</i> <i>subsp. equisimilis</i>	192	21.5	(Heng et al., 2006)
	Caseicin	<i>Lactobacillus casei</i> <i>strain 138</i>	-	~42	(Müller and Radler, 1993)
	Helveticin-J	<i>Lactobacillus</i> <i>helveticus</i>	255	37.5	(Joerger and Klaenhammer, 1990)
<b>Unsorted</b> <b>bacteriocins</b>	Clostin 574	<i>C. tyrobutyricum</i>	82	7.0	(Kemperman et al., 2003)
	PAMP	<i>P. jensenii</i>	64	6.3	(Faye et al., 2002)
	Plantaricin 163	<i>Lb. plantarum 163</i>	32	3.5	(Hu et al., 2013)
	Bactofencin A	<i>Lb. salivarius</i> <i>DPC6502</i>	22	2.7	(O'Shea et al., 2013)
	plantaricin LpU4	<i>Lb. plantarum LpU4</i>	-	4.8	(Miloni et al., 2015)
	Enterocin 7B	<i>Enterococcus</i> <i>faecalis</i>	43	5.2	(Martín-Platero et al., 2006)

### 6.1.1 Class I: Posttranslationally modified peptides (<10 kDa)

#### 6.1.1.1 Ia—The lanthipeptides

The *nisA* biosynthesis cluster encodes an operon consisting of 11 genes (Figure S1.2a). *nisA* operon is regulated by a two-component quorum sensing (QS) system involving *nisRK* that codes for a regulator protein and a histidine kinase which is sensitive to nisin concentration (Lubelski et al., 2008). Nisin cluster encodes the gene for the structural peptide, *nisA*; a dehydratase, *nisB*; a cyclase, *nisC*; a transporter for secretion of the peptide, *nisT*; an immunity factor for self-immunization of the producer, *nisI*; and a protease that cleaves off the leader sequence to activate peptide, *nisP*. An ATP-binding cassette (ABC) transporter system which is also responsible for immunity is encoded by *nisFEG* (Lubelski et al., 2008). In some cases, the activation protein may not be located in the cluster (e.g., subtilin). Pep5 for example is activated intracellularly while nisin is activated after secretion. The situation is different for type II which is activated by the transporter prior to secretion (Knerr and van der Donk, 2012, Repka et al., 2017).

Most class I lantibiotics exert their biological functions by binding to lipid II. For example, gallidermin may interact with the bactoprenylpyrophosphate moiety of lipid II to inhibit transglycosylation (Götz et al., 2014). Studies on the mode of action of nisin began way back in the 80's (Ruhr and Sahl, 1985). However, actual data that specifically demonstrate this has only been generated during the last decade. The A/B ring structure (Figure S1.2b) appears to be a peculiar feature observed in several other members of class I lantipeptides such as epidermin, gallidermin, microbisporicin and mutacin 1140. NMR studies of Nisin complexed to lipid II shows that the N-terminal A and B rings provide the amide backbone (Figure S1.2c) that coordinates intramolecular hydrogen bonding with the pyrophosphate moiety of the lipid intermediate (Hsu et al., 2004). Nisin binding to lipid II triggers membrane insertion, creating pores via a stoichiometric assembly of eight nisin and four lipid II molecules as illustrated in Figure S1.2d (Breukink et al., 1999, Hasper et al., 2004, Hasper et al., 2006). Mechanistic studies on self-immunity of *L. lactis* to nisin show that the C-terminus of NisI protects Lipid II from nisin binding and hence prevent pore formation (AlKhatib et al., 2014).



**Figure S1.2 Production and structure-function relationship of prototype lanthipeptides nisin.**

(a) Schematic drawings of nisin biosynthesis gene clusters (drawn to scale); structural genes, red; modifications, blue & green; leader peptide cleavage and/or transport, dark grey; immunity gene, yellow; accessory or other immunity genes, brown; (b) Schematic structure of nisin; (c) solution structure of the nisin, showing the A, B & C rings, and the lipid II-[C-terminal A/B rings] complex (lipid II is coloured grey). Abbreviations: Dha, dehydroalanine; Abu, 2-aminobutyric acid; Dhb, dehydrobutyrine; NAG, N-acetylglucosamine; FPP, farnesyl diphosphate; FGA, fibrinogen alpha chain; MUB, membrane-anchored ubiquitin-fold. (d) Mechanism of action of nisin, illustrating binding (I), membrane insertion (II) and pore formation (III). Nisin-lipid II complex was edited using solution NMR data for nisin (PDB: 1WCO) and Protein Workshop toolkit (Moreland et al., 2005).

Nisin also inhibits outgrowth of spores and disrupts the propagation of vegetative cells (Gut et al., 2011). These events may be coordinated by covalent interactions between the lantibiotic and the spores as suggested elsewhere (Morris et al., 1984). Other experimental data identify the dehydroalanine at position 5 of nisin to be responsible for its spore inhibitory role (Chan et al., 1996). Same is true for subtilin (Liu and Hansen, 1993). However, recent results from fluorescently labeled nisin analogues produced no evidence for a covalent mechanism of inhibition (Gut et al., 2011). Given that these effects are based on distinct structure-activity

relationships, the biological effects of such compounds should therefore proceed via two distinct molecular mechanisms. Furthermore, molecular basis that supports the ability of the two-peptide lanthipeptide lactacin 3147 to rapidly disrupt target cell membrane in the absence lipid II was recently characterized and results indicate that precise ring geometries in the two peptides are essential for synergistic activity (Bakhtiary et al., 2017). Although their pore-forming potentials may be weakened by the thickness and composition of the lipid bilayer membrane of the target strain (Wiedemann et al., 2006), some type II lantibiotics interact via a distinct lipid II-binding motif but do not form pores (Islam et al., 2012). Rather, they induce intense cell wall stress responses in their targets (Sass et al., 2008).

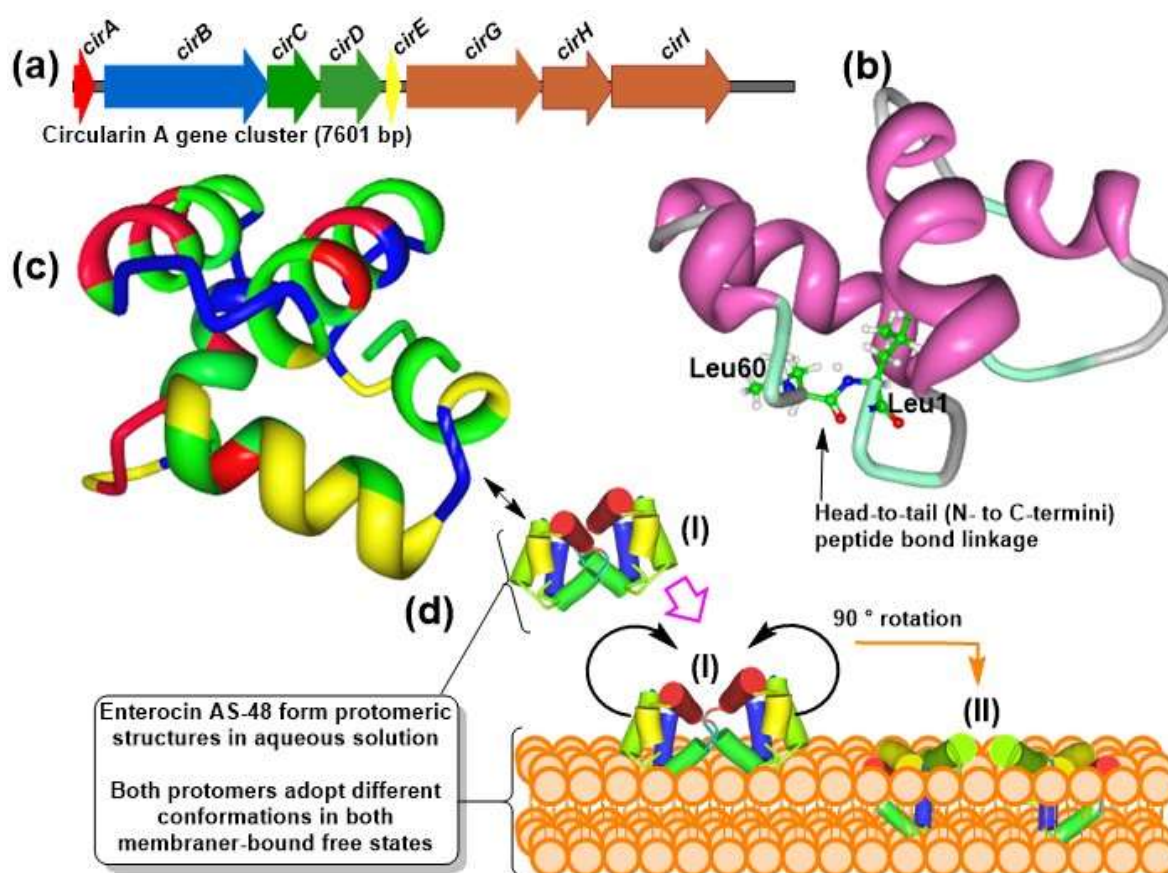
There is no record of types III and IIV exhibiting antimicrobial activity, however, they have shown activities in other aspects such as the activity of labyrinthopeptin A1 against HIV/AIDS and herpes simplex virus (HSV) (Férir et al., 2013), and the efficacy of labyrinthopeptin A2 against a modelled mouse inflicted with neuropathic pain (Meindl et al., 2010).

#### **6.1.1.2 Ib—Circular bacteriocins**

The main feature of this group of RiPP is the head-to-tail (N- and C-termini) peptide bond linkage which generate the cyclic nature of the molecules (Maqueda et al., 2008). All members possess  $\alpha$ -helices of similar sizes that fold into a tertiary structure having a central pore surrounded by a compact globular bundle comparable to the saposin folds (Montalbán-López et al., 2012, Acedo et al., 2015, Himeno et al., 2015). Note that other circular peptide antibiotics also exist like vancomycin and gramicidin S, but they are non-ribosomally produced.

Circularin A and enterocin AS-48 are two prominent examples of cyclized bacteriocins. The latter is produced by a cluster of 10 genes arranged on a 68 kb pheromone-responsive conjugative plasmid in a two-component operon system (Martínez-Bueno et al., 1998, Diaz et al., 2003). These genes include *as-48A*, *as-48B*, *as-48C*, *as-48C1D*, *as-48D1* and *as-48EFGH* which respectively codes for the structural peptide, a putative cyclase, a DUF95 protein implicated in immunity and biosynthesis, a putative ABC transporter, immunity protein, and additional ABC transporter responsible for self-immunization (Maqueda et al., 2008, Mu et al., 2014). The equivalence of *as-48ABCDD1* in circularin A was determined to be the *cirABCDE* (Figure S1.3a), and constituted the minimal set of genes required for active biosynthesis (Maqueda et al., 2008). The additional immunity genes are non-essential and play just a minor role since they appear to be absent in some known gene clusters (Maqueda et al., 2008, Gabrielsen et al., 2014). *In silico* analysis of a putative circular bacteriocin cluster in *Streptococcus pneumoniae* also identified a putative regulator (Bogaardt et al., 2015). Studies

on how the various components of the biosynthesis operon cooperate to produce active AS-48 have been reported extensively (Sanchez-Hidalgo, 2011, Cebrián et al., 2014), and other data point to the fact that the biosynthesis mechanism is not necessarily the same for all circular bacteriocins (Gabrielsen et al., 2014).



**Figure S1.3 Production machinery, structure and function of circular bacteriocins.** (a) Schematic representation of circularin A biosynthesis gene cluster (drawn to scale): structural gene, red; modifications, blue & green; immunity gene, yellow; accessory genes, brown. Ribbon representation of the structures of (b) carnocyclin A and (c) enterocin AS-48, showing saposin folds. The first (Leu1) and the last (Leu60) residues of carnocyclin A are shown to illustrate the C- to N-termini linkage. (d) Structural model for enterocin AS-48 molecular mechanism, illustrating how two protomeric units interact to produce two dimeric forms. The free-state dimeric form (I) transforms to dimeric form (II) at the membrane's surface, burying the hydrophobic helices into the hydrophobic lipid core while the polar helices interact with the polar heads. Structures (b) and (c) were edited using the solution NMR data for carnocyclin A (PDB: 2KJF) and the X-ray crystallographic data for enterocin AS-48 (PDB: 1O82).

ATPase activity may be necessary to supply the energy required to form the head-to-tail peptide bond during transport since there is no C-terminal extension and an additional processing step seems to be necessary to remove the leader peptide (Montalbán-López et al., 2012, Gabrielsen et al., 2014, Scholz et al., 2014, Alvarez-Sieiro et al., 2016).

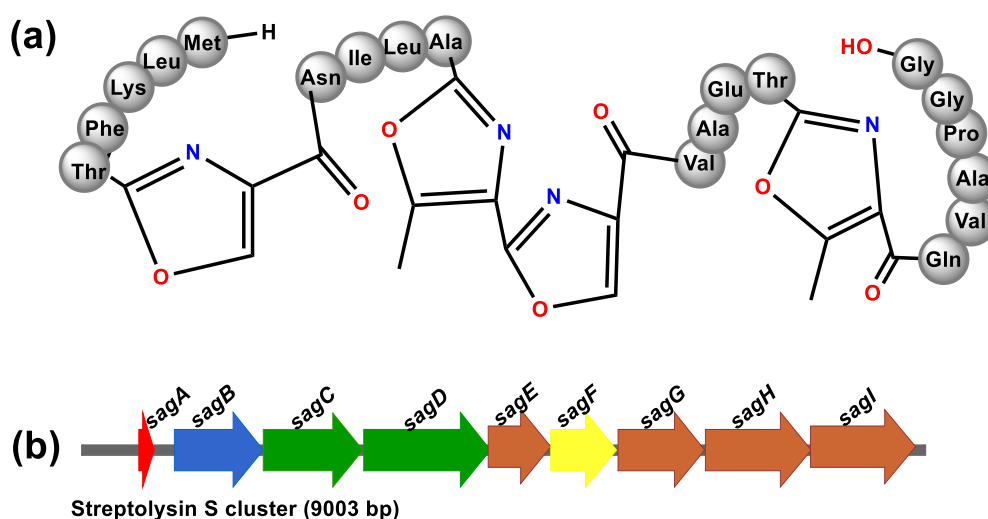
Reports suggest that circular bacteriocins may have multiple mechanisms of action, however, a vast majority of their members possess a common structural motif that is similar to the fold of mammalian saposins (Figure S1.3b & 1.3c) and hence may have a direct disruptive effect on the membranes or activate lipid-degrading enzymes (Martin-Visscher et al., 2009). Nevertheless, some deviation from this common principle exist. For example, enterocin AS-48 forms pores in lipid membranes and cause leakage of ions and low-molecular-weight compounds (Gálvez et al., 1991) whereas carnocyclin A forms voltage-dependent pores that are selective to anions (Gong et al., 2009). The formation of multimeric structures in aqueous solution which subsequently undergoes conformational readjustment to bury the non-polar core in the lipid membrane is characteristic of AS-48 as illustrated in Figure S1.3d (Cruz et al., 2013, Cebrián et al., 2015).

Whether or not cyclized AMPs in general require a docking molecule to exert their inhibitory effects on target organisms have been a subject under considerable debates in recent times. This was laid to rest when Gabrielsen and colleagues demonstrated increase sensitivity of *L. lactis* to garvicin ML resulting from the expression of a maltose ABC transporter complex (Gabrielsen et al., 2012). This study recorded the first putative target receptor for a circular bacteriocin. Furthermore, receptor-independent bactericidal effects was observed with high concentrations of enterocin AS-48 and carnocyclin A, indicating a concentration-dependent mode of action and that nonspecificity occur at higher bacteriocin concentrations while receptor on target cells maintain specific activity at the lower extremes (Gabrielsen et al., 2014).

#### **6.1.1.3 *lc*—Linear azol(in)e-containing peptides (LAPs)**

The biosynthesis pathway of some bacteriocins involves ATP-dependent cyclodehydration of threonine, cysteine and serine to produce flavin-dependent substrates which subsequently undergo dehydrogenation to form a mixture heterocyclic thiazole and (methyl)oxazole within the peptide chain (Melby et al., 2011). These types of peptides are referred to as Linear azol(in)e-containing peptides (Figure S1.4a).





**Figure S1.4 Structural features and production machinery of streptolysin S.** (a) Schematic structure of streptolysin S. (b) Organization of genes in streptolysin S biosynthesis operon (drawn to scale); structural gene, red; modifications, blue & green; immunity gene, yellow; accessory genes, brown.

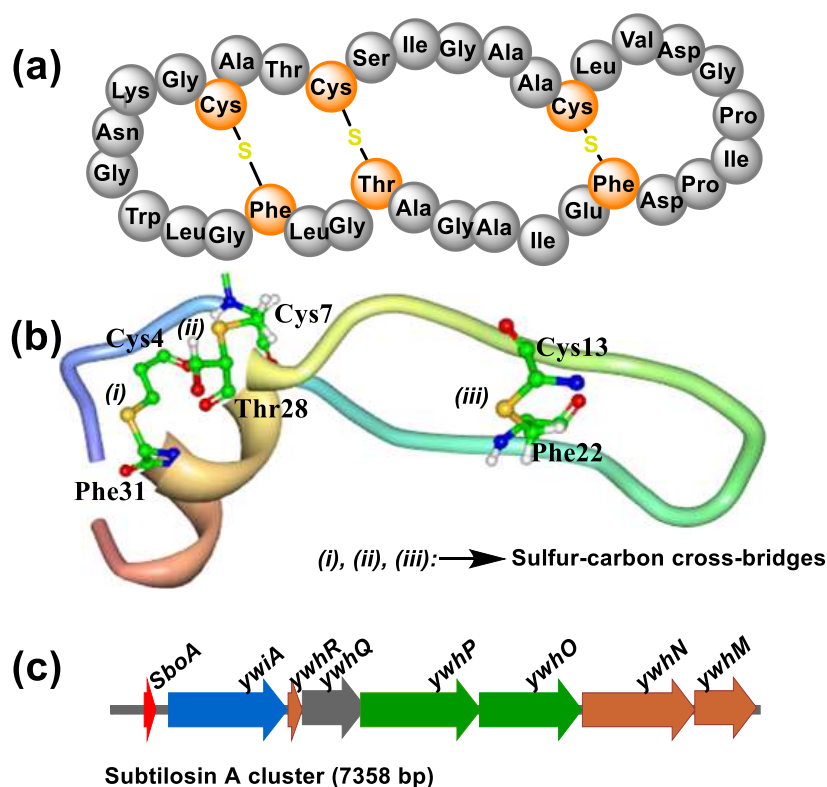
The biosynthesis machinery of Streptolysin S which happens to be the most prominent member of this class composed of *sagA*, *sagB*, *sagCD* and *sagE* (Figure S1.4b) which respectively codes for the structural peptide, a dehydrogenase, a cyclodehydratase and a protease (Cox et al., 2015). The ABC-type transporter *SagGHI* and *SagF*, may be responsible for immunity (Lee et al., 2008). With respect to their mode of action, LAPs remain extensively uncharacterized (Alvarez-Sieiro et al., 2016)

#### 6.1.1.4 Id—Sactipeptides

Sactipeptides are bacteriocins that contain sulphur-to- $\alpha$ -carbon linkages (Arnison et al., 2013, Mathur et al., 2015). Those that possess antimicrobial activities are also called sactibiotics. Extensive studies performed with subtilisin A show that the carbon-sulfur linkages and hairpins are common structural elements (Kawulka et al., 2003, Maqueda et al., 2008, Murphy et al., 2011). Subtilisin A has three sulfur-to- $\alpha$ -carbon cross-linkages (Figure S1.5a and S1.5b) and demonstrate wide activity spectrum against a variety of bacterial strains (Montalban-Lopez et al., 2011, Mathur et al., 2015). There are also two-component sactibiotic like thuricin CD with enhanced activity against *Clostridium difficile* (Rea et al., 2010). Thuricin CD and thuricin H are both produced by *Bacillus thuringiensis*, but the latter is a single peptide containing four sulfur-to- $\alpha$ -carbon linkages (Mathur et al., 2015).

The biosynthesis cluster of subtilisinA constitutes 8 open reading frames (ORF) including *sboA*, *ywiA* and *yRQPONM* most of whose molecular functions are unknown (Figure S1.5c).

*sboA* encodes the structural peptide, *ywiA* codes for modification protein, while *yRQPONM* are believed to function in processing, immunity and export of the peptide (Zheng et al., 1999, Stein et al., 2004). YwiA is an S-adenosylmethionine (SAM) enzymes containing the CXXXCXXC conserved motif that forms a [4Fe-4S] clusters necessary for the reductive cleavage of SAM into methionine and a 5'-deoxyadenosyl radical (Flühe et al., 2013).



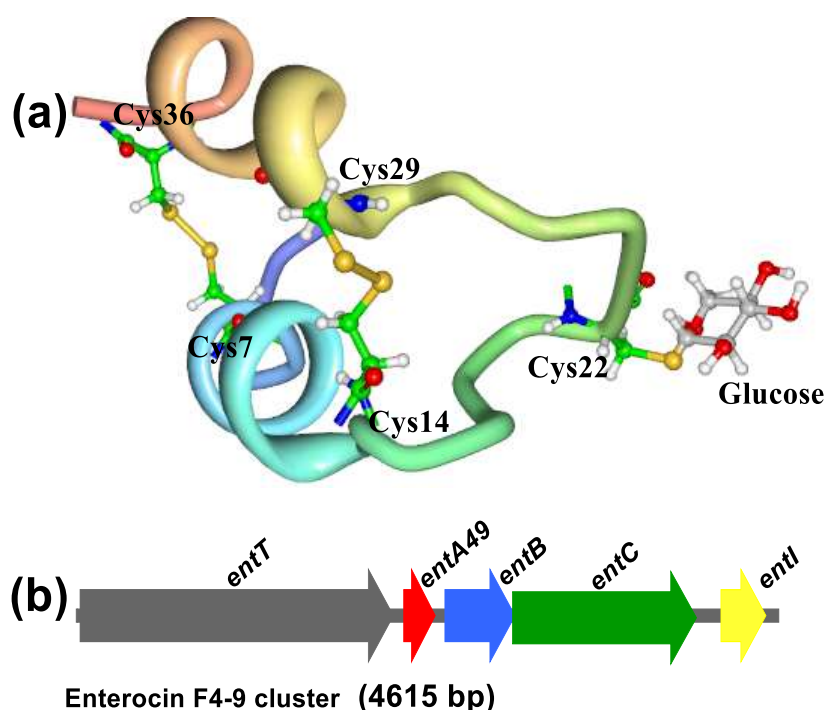
**Figure S1.5 Biosynthesis operon and structure of subtilisin A.** (a) schematic structure of subtilisin A and (b) its solution structure, showing the coordination of the sulfur- $\alpha$ -carbon bridges. (c) Organization of genes in the biosynthesis cluster of subtilisin A (drawn to scale); structural gene, red; modifications, blue & green; transport, dark grey; accessory genes, brown. Structure (b) was edited using the solution NMR data for subtilisin A (PDB: 1PXQ).

Unlike thurincin H which seems not to affect the permeability of the phospholipid bilayer membrane (Wang et al., 2014), the bactericidal mechanism of subtilisin A involves partial insertion of the peptide into the hydrophobic core of target cell membrane. The buried peptide ultimately causes a disarray within the area leading to the creation of transient pores (Noll et al., 2011).



### 6.1.1.5 *le*—Glycocins

Glycocins are a group of RiPPs with one or more residues in the peptide chain linked to a carbohydrate moiety (Arnison et al., 2013). Glycocin F produced by *Lactobacillus plantarum* is a prototype example of this class to be structurally characterized (Stepper et al., 2011). Other examples like thurandacin A & B (Wang et al., 2013), Sublancin 168 (Garcia De Gonzalo et al., 2014) and enterocin F4-9 (Maky et al., 2015) have also been described. Sublancin 168 and thurandacin A are  $\beta$ -S-linked glycosylated bacteriocins; enterocin F4-9 contains  $\beta$ -O-linked glucose moiety (Figure S1.6), while glycocin F and thurandacin B contain both  $\beta$ -S-linked and  $\beta$ -O-linked glycosylated moieties. Glucose, N-acetylglucosamine and N-acetylhexosamine are the carbohydrate compounds that have been identified so far.



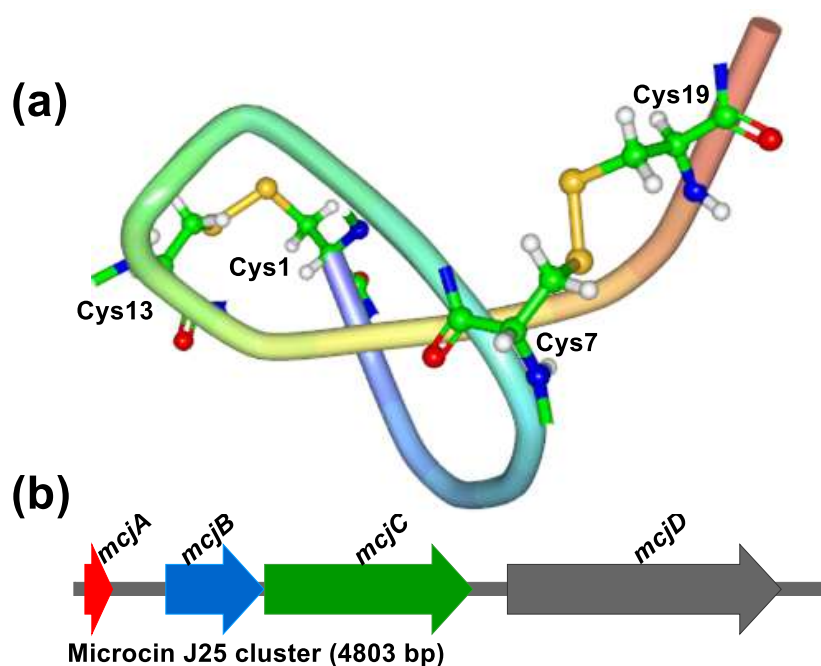
**Figure S1.6 Structural features and biosynthesis of glycopeptide bacteriocins.** (a) NMR structure of sublancin 168 (PDB: 2MIJ), showing a glucose moiety linked to Cys22 and two intramolecular disulfide bridges. (b) Organization of genes in the biosynthesis cluster of enterocin F4-9 (drawn to scale); structural gene, red; modifications, blue & green; transport, dark grey; immunity gene, yellow. Structure in (a) was edited using the solution NMR data for sublancin 168 (PDB: 2MIJ).

Enterocin F4-9 is a bacteriostatic peptide (unlike glycocin F) and its biosynthesis pathway comprises five genes including *entT*, *entA49*, *entB*, *entC* and *entI*, which respectively encodes a putative ABC-transporter; the structural peptide, the glycosyltransferase, a thioldisulfide isomerase and a self-protecting protein (Maky et al., 2015).

Although data on the mechanism of action are rare, results with O-deglycosylated glycin F shows that the O-linked N-acetylglucosamine may be involved in a reversible interaction with target cells (Stepper et al., 2011).

#### 6.1.1.6 If—Lasso peptides

Lasso peptides are bacteriocins that possess a ring formed via an amide bond between the first residue of the core peptide and a negatively charged core residue at positions +7, +8 or +9; after which the ring then embodies the linear C-terminus of the sequence (Arnison et al., 2013, Hegemann et al., 2015, Alvarez-Sieiro et al., 2016). Additional modifications such as intramolecular disulfide bridges like in the case of svieceucin (Figure S1.7a) are also possible. Such structural arrangement confers high stability to the peptides and hence they may be used as peptide scaffolds (Alvarez-Sieiro et al., 2016).



**Figure S1.7 Structural features and biosynthesis of lasso peptides.** (a) NMR structure of svieceucin, showing two intramolecular disulfide bridges between Cys1 and Cys13, and Cys7 and Cys19. (b) Organization of genes in the biosynthesis cluster of microcin J25 (drawn to scale); structural gene, red; modifications, blue & green; immunity gene, yellow; transport, dark grey. Structure in (a) was edited using the solution NMR data for svieceucin (PDB: 2LS1).

Microcin J25, from *Escherichia coli* was the first lasso peptide with antimicrobial activities to be characterized. Genome mining approach on 87 different proteobacterial strains recently

identified 108 putative biosynthetic gene clusters for lasso peptides and heterologous expression of 12 members from this list was successfully performed in *E. coli* (Hegemann et al., 2013). The biosynthesis cluster of microcin J25 constitutes four genes namely; *mcjA*, *mcjB*, *mcjC* and *mcjD* (Figure S1.7b) which respectively encodes the structural peptide, the peptidase, the cyclase and the ABC transporter (Yan et al., 2012).

With respect to the mode of action, it has earlier been established that Microcin J25 enters target bacterial cells via FhuA receptor in the outer membrane (Salomón and Fariás, 1993) as well as the SbmA protein in the inner membrane (Salomón and Farias, 1995) to inhibit RNA polymerase (Mathavan et al., 2014), triggers the production of reactive oxygen species that facilitates the inhibition process (Chalon et al., 2009). Capistrin also inhibits bacterial RNA polymerase (Kuznedelov et al., 2011), unlike lassomycin which targets *Mycobacterium tuberculosis* by inhibiting proteases (Gavriš et al., 2014). Like other antimicrobial peptides, biological activities of lasso peptides expand beyond antimicrobials to include antiviral or anticancer (Maksimov et al., 2012).

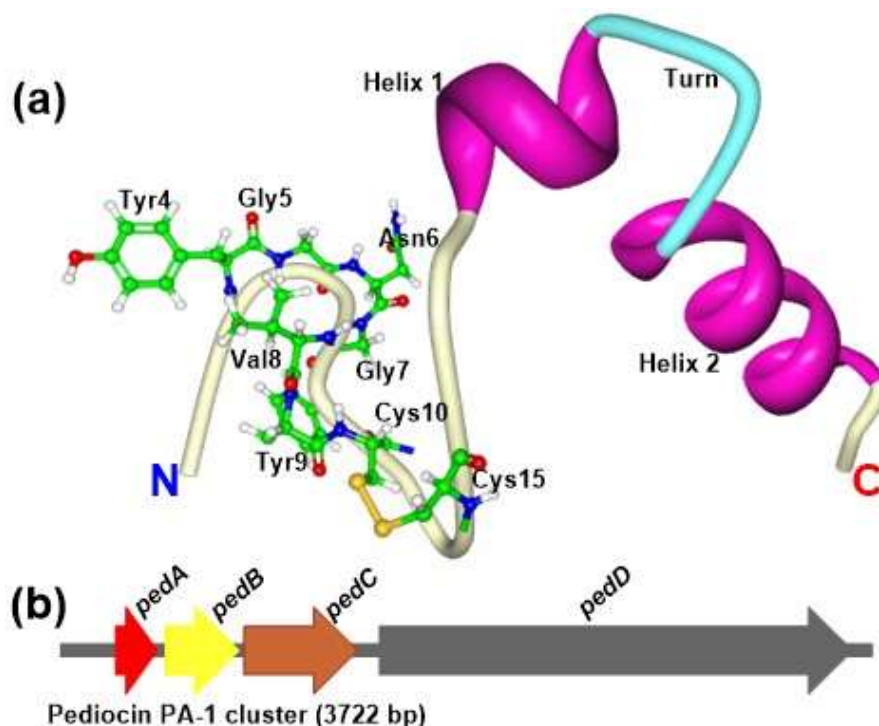
## 6.1.2 Class II: Unmodified bacteriocins (<10 kDa)

### 6.1.2.1 Ila—pediocin-like

Pediocin-like bacteriocins constitute the largest bacteriocin group, heat-stable and are produced by a variety of LAB (Cui et al., 2012). They are broad spectrum antimicrobials with strong activity against listeria (Kjos, 2011). Class Ila peptides has two distinct regions in their structures separated by a flexible hinge (Haugen et al., 2008). The N-terminus has an overall positive charge, a disulfide bridge formed between two cysteine residues and a conserved YGNGVXC consensus motif (Figure S1.8a), which may be actively involved in target recognition and the killing process (Cui et al., 2012, Perez et al., 2014). The disulfide bridge may play a stability role and not necessary directly involved in the killing process since its replacement by hydrophobic interaction did not abolish the activity of leucocin A (Sit et al., 2012).

Pediocin PA-1 is the prototype member of this class and has been extensively studied. Its biosynthesis cluster is harboured on a plasmid (Ennahar et al., 1999). As shown in Figure S1.8b, the *ped* operon is encoded by four genes, namely; *pedA* and *pedB*, encoding the structural peptide and the immunity protein respectively, and *pedC* and *pedD*, encoding the accessory factor component that constitutes part of the ABC transport systems and the ABC transporter respectively (Miller et al., 2005). The leader sequence in the structural peptide serves the purpose of recognition and directs the unprocessed molecule to a dedicated ABC transport systems where it is processed and secreted (Alvarez-Sieiro et al., 2016). Additionally,

sec-dependent translocation pathway is employed by a few class IIa like the bacteriocin T8, enterocin P and bacteriocin 31 (De Kwaadsteniet et al., 2006). Efforts to improve class IIa bacteriocins production which is usually controlled by the QS system enabled the cloning and expression in new hosts that allowed constitutive overexpression (Ennahar et al., 1999).



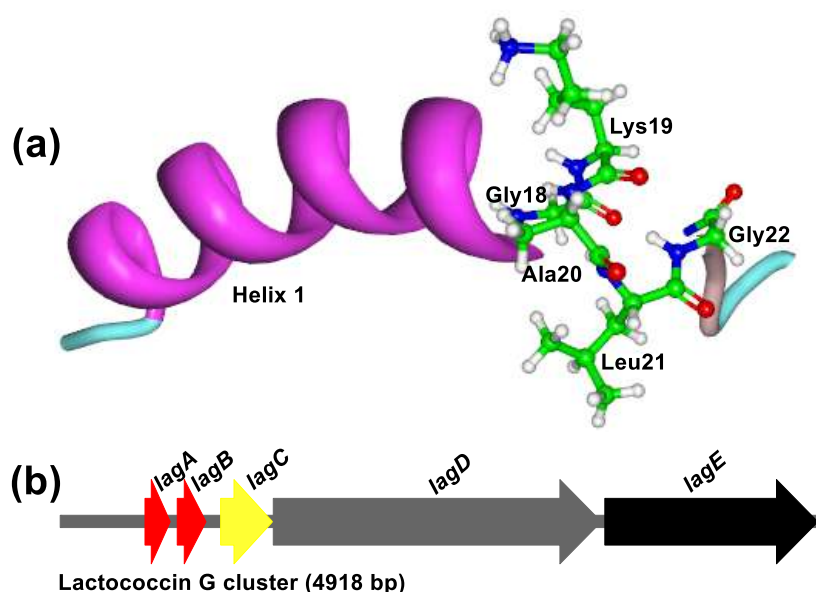
**Figure S1.8 Structural features and biosynthesis of pediocin-like bacteriocins.** (a) NMR structure of Curvacin A, showing the N-terminal YGNGVXC conserved motif, the N-terminal intramolecular disulfide bridge between to Cys10 and Cys15, and the distinct C-terminal helical region. (b) Schematic representation of genes in the biosynthesis cluster of pediocin PA-1 (drawn to scale); structural gene, red; immunity gene, yellow; transport, dark grey; accessory gene, brown. Structure in (a) was edited using the solution NMR data for Curvacin A (PDB: 2A2B).

For Class IIa bacteriocins to target a specified bacterium, a mannose phosphotransferase system (ManPTS) receptor is required as recently shown in *Enterococcus faecium* that disrupting a specific ManPTS is responsible for its resistance against class IIa bacteriocins (Geldart and Kaznessis, 2017). Additionally, altering the cell wall and lipid membrane composition in combination with malfunctioned ManPTS confer resistance in *Listeria monocytogenes* (Masias et al., 2017). The killing mechanism, of class IIa bacteriocins encompasses binding of the peptide to ManPTS receptors, followed by insertion into the phospholipid bilayer membrane and subsequently creating a pore complex (Diep et al., 2007).

### 6.1.2.2 *IIb—Two-peptide bacteriocins*

Class IIb bacteriocins are also referred to as two-peptide bacteriocins and they are identified based on both  $\alpha$ - and  $\beta$ - components synergistically complementing each other to exhibit full bioactivity, but show little or no activity in separate fractions (Nissen-Meyer et al., 2010). They usually have a conserved GxxxG-motifs (Figure S1.9a), however, examples like as plantaricin S $\beta$  and plantaricin NC8 $\beta$  possess instead an AxxxA- or SxxxS-motif respectively (Nissen-Meyer et al., 2010). Self-immunization of the native host to both peptides is sufficiently conferred by one immunity protein and the biosynthesis operon encodes the two structural genes in succession (Nes et al., 2007). Individual peptides may exhibit high antimicrobial activity by themselves like in the case of thermophilin 13 (Marciset et al., 1997), but the effectiveness would be synergistically enhanced upon addition of its complementary peptide. This does not hold true for lactococcin G which requires the presence of both peptides for full activity to be achieved (Nissen-Meyer et al., 1992). Additionally, the situation is different with enterocin X where single and/or synergistic activities are dependent on the indicator strain (Hu et al., 2010). Furthermore, the complementary peptides may also come from homologous members (Oppegård et al., 2007)

The biosynthesis cluster of two-peptide bacteriocins require the products of at least five genes operating in separate or on the same operons (Alvarez-Sieiro et al., 2016). For example, a single operon encodes the production of lactococcin G (Figure S1.9b). This includes the genes that encode the structural peptides *lagA* and *lagB*, which are succeeded by the immunity determinant *lagC*; a dedicated ABC transporter *lagD* and an accessory factor component of the ABC transporter system *lagE*, whose actual function remains to be determined (Oppegård et al., 2010). The structural genes are arranged in succession along the operon together with one immunity gene that offer self-protection from both peptides to the host (Nissen-Meyer et al., 2009, Rogne et al., 2009), suggesting a functional synergism of the complementary components. A three-components QS system regulates the biosynthesis of class IIb bacteriocins involving an inducer peptide (IP) a membrane-associate histidine kinase, and response regulators. A good example is Plantaricin JK and plantaricin EF which are both induced by plantaricin A in *L. plantarum* C11 (Diep et al., 2003)

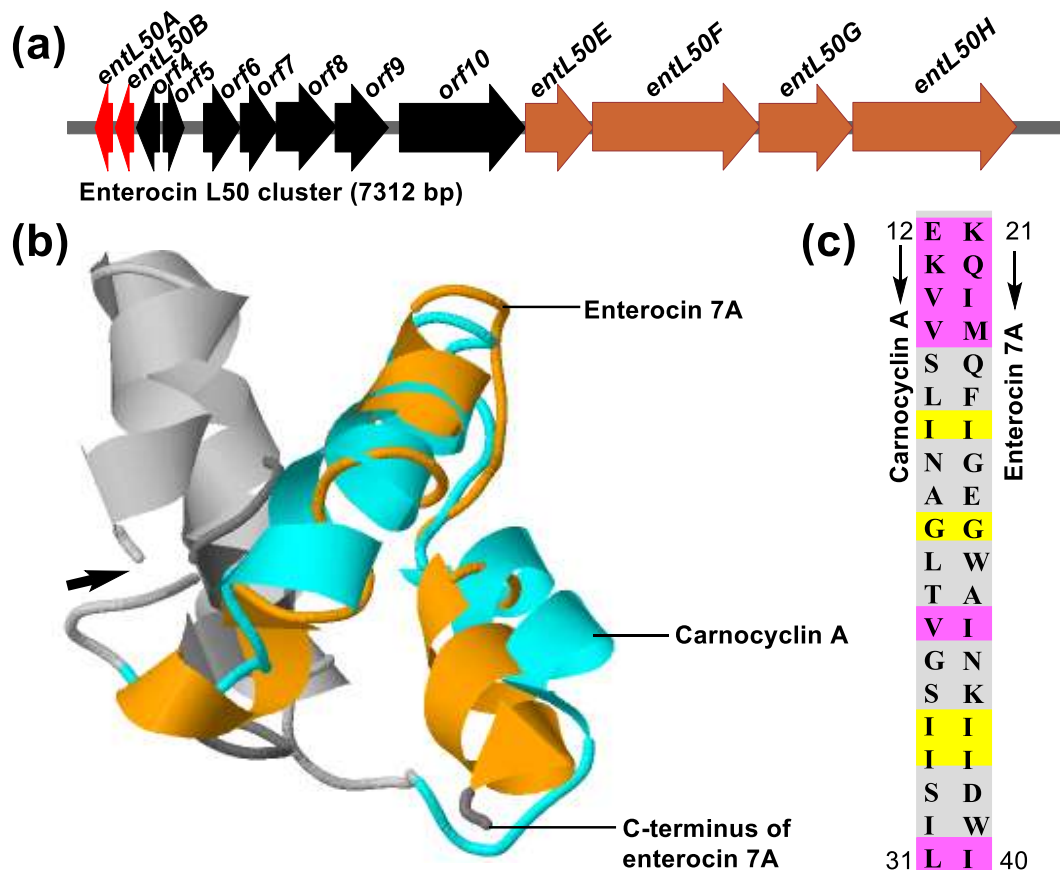


**Figure S1.9 Structural features and biosynthesis of two-peptide bacteriocins.** (a) NMR structure of Lactococcin G- $\alpha$ , showing the N-terminal region with one  $\alpha$ -helix and the conserved GxxxG motif. (b) Schematic representation of genes in the biosynthesis cluster of Lactococcin G (drawn to scale); structural genes, red; immunity gene, yellow; transport, dark grey; unknown gene, black. Structure in (a) was edited using the solution NMR data for Lactococcin G- $\alpha$  (PDB: 2JPJ).

A killing model for class IIb bacteriocins initially suggested that the  $\alpha$ - and  $\beta$ -peptides complex form a pore across the membrane of the target bacteria, triggering the efflux of monovalent cations including  $K^+$  ions and eventual death of the microbe (Moll et al., 1996). This was further investigated using lactococcin G and findings also supported the initial assertion while maintaining that the GxxxG-motifs of the peptides may form a membrane-penetrating helix–helix-interacting structure that engages a membrane receptor especially UppP protein (Kjos et al., 2014) on the susceptible bacteria membrane (Rogne et al., 2008, Nissen-Meyer et al., 2010), causing monovalent cations efflux from the cell.

### 6.1.2.3 IIc—Leaderless bacteriocins

Class IIc bacteriocins are referred to as leaderless bacteriocins because they have no N-terminal leader sequence or signal peptide, which in other classes plays a signaling role in the secretion and modification pathways, as well as maintaining the peptide inert inside the producer cell (Cintas et al., 2000, Liu et al., 2011, Masuda et al., 2012). To date, more than 20 different leaderless bacteriocins have been isolated and characterized.



**Figure S1.10 Biosynthesis operon and structural properties of leaderless bacteriocins.** (a) Schematic representation of genes involved in enterocin L50 biosynthesis (drawn to scale); structural genes, red; accessory genes, brown; unknown gene, black. (b) Structural overlay of enterocin 7A (orange) and carnocyclin A (cyan). Bold arrow on the left side of the overlay denotes position where the head-to-tail (N- to C-termini) peptide bond of carnocyclin A is formed. (c) Peptide sequence alignment of a segment of carnocyclin A and enterocin 7A. Yellow colour indicates conserved residues, pink indicates residues with very similar properties, and the other residues in light grey are weakly similar. The overlay in (b) was edited using RCSB PDB Protein Comparison Tool and the solution NMR data for enterocin 7A (PDB: 2M5Z) and carnocyclin A (PDB: 2KJF).

The biosynthesis operon includes a dedicated immunity gene, but sometimes immunity for the producers may be conferred by the ABC transporter that is also involved in their secretion (Gajic et al., 2003, Criado et al., 2006). However, this is still not clear since most their clusters do not seem to contain the immunity encoding genes (Iwatani et al., 2012). Sometimes the operon consists of consecutive genes encoding non-identical peptide copies (Nes et al., 2007), which strongly suggests evolutionary minor sequence alterations that potentiate production efficiency and strengthens antimicrobial activity via synergistic action. The most studied example is the two-peptide leaderless peptide enterocin L50 produced by *E. faecium* L50 (Cintas et al., 1998) whose cluster is made of 13 open ORFs (Figure S1.10a) including the structural genes *entL50A* and *entL50B*, helper proteins and the *entL50EFGH* which is



homologous to the second ABC transporter system of AS-48 bacteriocin *as-48EFGH* (Franz et al., 2007).

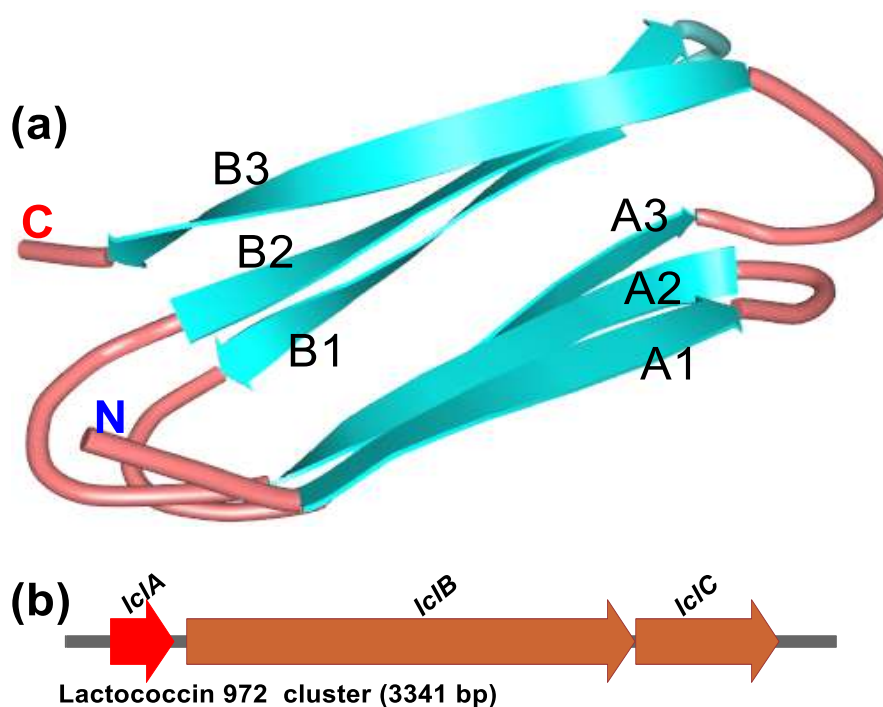
Enterocins 7A and 7B possess distinguishable high level of sequence homology to enterocins L50A and L50B respectively (Liu et al., 2011). Interestingly, enterocins L50A and L50B adopt a defined folding conformation in aqueous solution, with the final NMR solution structure of enterocins 7A showing an unexpected structural homology to the circular bacteriocin carnocyclin A (Figures S1.10b and S1.10c) (Lohans et al., 2013). Furthermore, although enterocins 7A and 7B are shorter than Aureocin A53 & lacticin Q, structural homology modeling showed that there exist a conserved structural motif and some similarity to circular bacteriocins, which may be common for all leaderless bacteriocins (Acedo et al., 2016).

The “huge toroidal pore” mechanism has been reported for lacticin Q where the cationic peptide rapidly interacts electrostatically and binds to outer membrane interface, forms pores and diffuses to the inner membrane causing membrane leakage (Yoneyama et al., 2009). Nevertheless, this mechanism is limited by the accumulation of hydroxyl radicals serving as activity inducing factor (Li et al., 2013). Furthermore, the leaderless bacteriocin LsbB uses a zinc-dependent membrane metallopeptidase receptor for its biological activity (Uzelac et al., 2013).

#### **6.1.2.4 *lld*—Non-pediocin-like single-peptide bacteriocins**

Non-pediocin-like single-peptide bacteriocins is composed of varieties of linear peptide that have very little in common including dissimilarities in their structures, different transport mechanisms and mode of action like lactococcin A, lactococcin 972, and enterocin B (Franz et al., 2007, Alvarez-Sieiro et al., 2016). The plasmid-encoded biosynthesis cluster of the sec-dependent, heat-labile and pH-stable peptide Lactococcin 972 has been extensively studied (Martínez et al., 1999) and its solution structure has been determined (Turner et al., 2013). NMR solution structure (Figure S1.11a) shows that lactococcin 972 has two three-stranded antiparallel  $\beta$ -sheets that folds into a  $\beta$ -sandwich (Turner et al., 2013). Its biosynthesis gene cluster (Figure S1.11b) consist of *lclA*, encoding the structural peptide; *lclB* and *lclC*, both encoding a putative dedicated ABC transporter system which is also involved in self-protection (Martínez et al., 1999, Campelo et al., 2014).





**Figure S1.11 Structural features and biosynthesis of non-pediocin-like single-peptide bacteriocins.** (a) NMR solution structure of lactococcin 972. Structure was edited using the solution NMR data for lactococcin 972 (PDB: 2LGN). (b) Schematic representation of the biosynthesis cluster of lactococcin 972 (drawn to scale); structural gene, red; accessory genes, brown.

Binding of lipid II precursor and inhibition of cell wall biosynthesis in diverse species of lactococci is the major route through which lactococcin 972 exerts its biological action (Martínez et al., 2000). Alternatively, lactococcin A uses ManPTS as receptor on sensitive cells to permeabilize the cytoplasmic membrane, causing cell death (Diep et al., 2007). It is surprising that lactococcin A utilize ManPTS receptors on target cells and kills exclusively other lactococci, while pediocin-like bacteriocins which uses the same target do not kill lactococci at all (Diep et al., 2007).

### 6.1.3 Class III: Unmodified high molecular weight bacteriocins

Class III bacteriocins are unmodified multidomain bacteriocins with molecular weights >10 kDa, they are sensitive to heat and exhibiting either bacteriolytic or non-lytic mode of action (Alvarez-Sieiro et al., 2016). An example is enterolysin A whose sequence analysis shows that it comprises of an N-terminal endopeptidase domain and a C-terminal domain. The latter serves as a substrate recognition motif with significant homology to zoocin A (Lai et al., 2002, Nilsen et al., 2003).

The gene cluster of the prototype member lysostaphin (Schindler and Schuhardt, 1964) is harboured on plasmid pACK1 (Heath et al., 1987) and consists of *lssA* and *lif* genes which respectively encode the structural peptides and the immunity protein (Thumm and Götz, 1997). The lysostaphin immunity determinant has an opposite orientation with respect to the *LssA* (Appendix 6.1). Similarly, the non-lytic dysgalacticin from *Streptococcus pyogenes* is also plasmid-encoded and its cluster comprises two genes, the genes for structural precursor (*dysA*) and the immunity determinant (*dysI*) (Heng et al., 2006). In the case of zoocin A, the immunity factor introduces L-alanine into the peptidoglycan cross-linkages of *Streptococcus equi subsp. Zooepidemicus*, thereby eliminating the ability of the bacteriocin to degrade the cell wall (Gargis et al., 2009). Another example is the murein hydrolase millericin B whose cluster encodes *milB*, *milF*, *milT* for the expression of precursor, immunity factor and a transporter respectively (Beukes et al., 2000).

The mode of action of dysgalacticin involves binding to the glucose- and/or mannose-PTS, obstructing sugar uptake as well as causing the efflux of small molecules out of the cell (Swe et al., 2009). Alternatively, the lytic bacteriocins enterolysin A targets peptidoglycan of susceptible cells by performing intrapeptide cleavage between the L-alanine and D-glutamic acid of the stem peptide as well as interpeptide cleavage between L-lysine and D-aspartic acid of the bridge (Khan et al., 2013). Caseicin on the other hand inhibits DNA replication, transcription and/or translation (Müller and Radler, 1993).

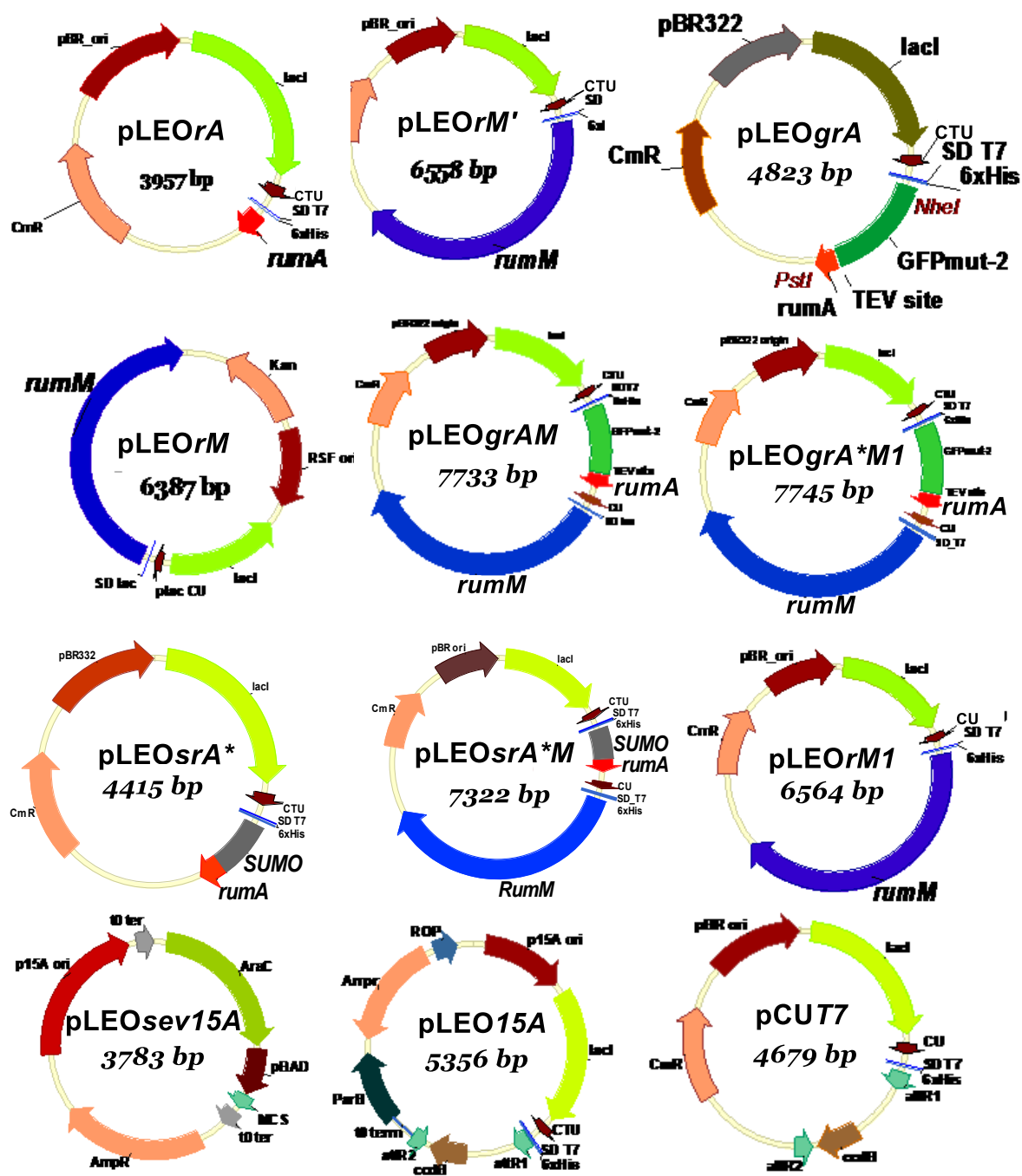
## 6.2 Vector descriptions and vector maps

**Table A1** Name and description of plasmids

Vectors	Size (kb)	Properties	Marker	Ref.
pCTUT7_His	4.7	plac_CTU promoter, SD_T7 RBS, N-terminal His-tag, <i>attR sites</i> , <i>ccdB</i> , <i>lacI</i> , <i>pBR ori</i>	CmR	Horn group, HKI Jena
pCTUT7-SUMO	5	plac_CTU promoter, SD_lac RBS, N-terminal His-Sumo-tag <i>attR sites</i> , <i>ccdB</i> , <i>lacI</i> , <i>pBR ori</i>	CmR	Horn group, HKI Jena
pCTUT7_Trx	5	plac_CTU promoter, SD_T7 RBS, N-terminal His-Trx-tag, <i>attR sites</i> , <i>ccdB</i> , <i>lacI</i> , <i>pBR ori</i>	CmR	Horn group, HKI Jena
pJL01	4.7	plac_CTU promoter, SD_T7 RBS, N-terminal His-tag, <i>attR sites</i> , <i>ccdB</i> , <i>lacI</i> , <i>pBR ori</i> , <i>SpeI</i> site	CmR	Li, (2013)

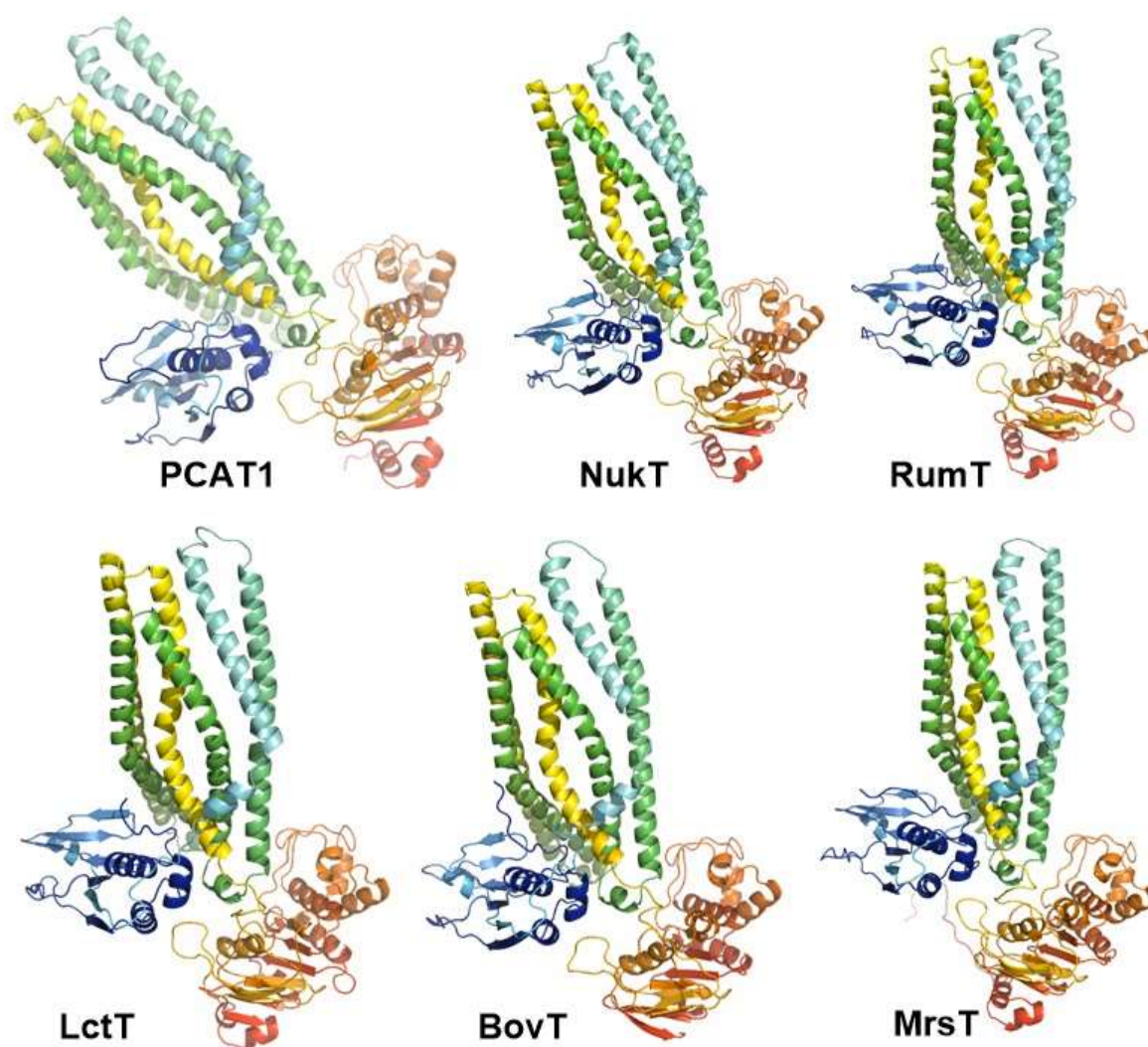
pJL02	4.8	plac_CTU promoter, SD_T7 RBS, N-terminal His-tag, <i>attR sites</i> , <i>ccdB</i> , <i>lacI</i> , <i>p15A ori</i> , <i>SpeI</i> site	CmR	Li, (2013)
pJL07	5.3	plac_CTU promoter, SD_T7 RBS, N-terminal His-tag, <i>attR sites</i> , <i>ccdB</i> , <i>lacI</i> , <i>parB</i> , <i>rop</i> , <i>λt0 terminator</i> , <i>pBR ori</i>	CmR	Li, (2013)
pJL10	3.8	plac_CU promoter, SD_T7 RBS, N-terminal His-tag, <i>attR sites</i> , <i>ccdB</i> , <i>lacI</i> , <i>pRSF ori</i>	KanR	Li, (2013)
pRSF-1b	3.6	T7 promoter His-tag, <i>S-tag</i> , <i>T7 terminator</i> , <i>RSF ori</i> , <i>lacI</i>	KanR	Novagen
pCUT7	4.6	plac_CU promoter, SD_T7 RBS, N-terminal His-tag, <i>attR sites</i> , <i>ccdB</i> , <i>lacI</i> , <i>pBR ori</i>	CmR	This work
pLEOp15A	4.7	plac_CTU promoter, SD_T7 RBS, N-terminal His-tag, <i>attR sites</i> , <i>ccdB</i> , <i>lacI</i> , <i>parB</i> , <i>rop</i> , <i>λt0 terminator</i> , <i>p15A</i>	CmR	This work
pLEOrC2	8.1	plac_CTU promoter, SD_T7 RBS, N-terminal His-tag, <i>attR2</i> , <i>rumA1A2A3</i> , <i>rumM</i> , <i>lacI</i> , <i>pBR ori</i>	CmR	This work
pLEOrA	3.9	plac_CTU promoter, SD_T7 RBS, <i>His-rumA1</i> , <i>attR2</i> , <i>lacI</i> , <i>pBR ori</i>	CmR	This work
pLEOrA*	3.9	plac_CTU promoter, SD_T7 RBS, <i>His-rumA1</i> ( <i>Trypsin site</i> ), <i>attR2</i> , <i>lacI</i> , <i>pBR ori</i>	CmR	This work
pLEOrA <sup>GluC</sup>	3.9	plac_CTU promoter, SD_T7 RBS, <i>His-rumA1</i> ( <i>GluC site</i> ), <i>attR2</i> , <i>lacI</i> , <i>pBR ori</i>	CmR	This work
pLEOrA <sup>Xa</sup>	3.9	plac_CTU promoter, SD_T7 RBS, <i>His-rumA1</i> ( <i>Factor Xa site</i> ), <i>attR2</i> , <i>lacI</i> , <i>pBR ori</i>	CmR	This work
pLEOrA <sup>TEV</sup>	3.9	plac_CTU promoter, SD_T7 RBS, <i>His-rumA1</i> ( <i>TEV site</i> ), <i>attR2</i> , <i>lacI</i> , <i>pBR ori</i>	CmR	This work
pLEOrA <sup>TEVM</sup> <sub>1</sub>	6.6	plac_CTU promoter, SD_T7 RBS, <i>His-rumA1</i> ( <i>TEV site</i> ), <i>His-rumM</i> , <i>attR2</i> , <i>lacI</i> , <i>pBR ori</i>	CmR	This work
pLEOrM	6.3	plac_CU promoter, SD_T7 RBS, <i>His-rumM</i> , <i>attR2</i> , <i>lacI</i> , <i>pRSF ori</i>	KanR	This work
pLEOrM'	6.5	plac_CTU promoter, SD_T7 RBS, <i>His-rumM</i> , <i>attR2</i> , <i>lacI</i> , <i>pBR ori</i>	CmR	This work
pLEOrM1	6.5	plac_CU promoter, SD_T7 RBS, <i>His-rumM</i> , <i>attR2</i> , <i>lacI</i> , <i>pBR ori</i>	CmR	This work
pRSF-rM	3.6	T7 promoter His-tag, <i>S-tag</i> , <i>T7 terminator</i> , <i>RSF ori</i> , <i>lacI</i> , <i>His-rumM</i>	KanR	This work
pLEOgrA	4.8	plac_CTU promoter, SD_T7 RBS, <i>His-gfpmut2-tev-rumA1</i> , <i>attR2</i> , <i>lacI</i> , <i>pBR ori</i>	CmR	This work
pLEOgrAM	7.7	plac_CTU promoter1, plac_CU promoter2, SD1_T7 RBS, SD2_lac RBS, <i>His-gfpmut2-tev-rumA1</i> , <i>His-rumM</i> , <i>lacI</i> , <i>pBR ori</i>	CmR	This work
pLEOgrA*	4.8	plac_CTU promoter, SD_T7 RBS, N-terminal His-tag, <i>attR2</i> , <i>gfpmut2</i> , <i>rumA1</i> ( <i>Trypsin site</i> ), <i>lacI</i> , <i>pBR ori</i>	CmR	This work

pLEOgrA*M	7.7	plac_CTU promoter1, plac_CTU promoter2, SD1_T7 RBS, SD2_lac RBS, His-gfpmut2-tev-rumA1( <i>Trypsin site</i> ), His-rumM, <i>lacI</i> , <i>pBR ori</i>	CmR	This work
pLEOgrA*M1	7.7	plac_CTU promoter1, plac_CU promoter2, SD1_T7 RBS, SD2_T7 RBS, N-terminal His-tag, <i>gfpmut2-tev-rumA1(Trypsin site)</i> , His-rumM, <i>lacI</i> , <i>pBR ori</i>	CmR	This work
pLEOt0grA	6.1	plac_CTU promoter, SD_T7 RBS, N-terminal His-tag, <i>attR2</i> , <i>gfpmut2-tev-rumA1</i> , <i>lacI</i> , <i>parB</i> , <i>rop</i> , <i>λt0 terminator</i> , <i>pBR ori</i>	AmpR	This work
pLEOsrA*	4.4	plac_CTU promoter, SD_lac RBS, His-Sumo-rumA1( <i>Trypsin site</i> ), <i>attR2</i> , <i>lacI</i> , <i>pBR ori</i>	CmR	This work
pLEOsrA <sup>Xa</sup>	4.4	plac_CTU promoter, SD_lac RBS, His-Sumo-rumA1( <i>Factor Xa site</i> ), <i>attR2</i> , <i>lacI</i> , <i>pBR ori</i>	CmR	This work
pLEOsrA <sup>TEV</sup>	4.4	plac_CTU promoter, SD_lac RBS, His-Sumo-rumA1( <i>TEV site</i> ), <i>attR2</i> , <i>lacI</i> , <i>pBR ori</i>	CmR	This work
pLEOsrA*M	7.3	plac_CTU promoter1, plac_CTU promoter2, SD1_T7 RBS, SD2_lac RBS, His-Sumo-rumA1( <i>Trypsin site</i> ), <i>attR2</i> , His-rumM, <i>lacI</i> , <i>pBR ori</i>	CmR	This work
pLEOsrA*M1	7.3	plac_CTU promoter1, plac_CU promoter2, SD1_T7 RBS, SD2_T7 RBS, His-Sumo-rumA1( <i>Trypsin site</i> ), <i>attR2</i> , His-rumM, <i>lacI</i> , <i>pBR ori</i>	CmR	This work
pLEOsrA <sup>TEVM</sup>	7.3	plac_CTU promoter1, plac_CTU promoter2, SD1_T7 RBS, SD2_T7 RBS, His-Sumo-rumA1( <i>TEV site</i> ), <i>attR2</i> , His-rumM, <i>lacI</i> , <i>pBR ori</i>	CmR	This work
pLEOsrA <sup>XaM</sup>	7.3	plac_CTU promoter1, plac_CTU promoter2, SD1_T7 RBS, SD2_T7 RBS, His-Sumo-rumA1( <i>factorXa site</i> ), <i>attR2</i> , His-rumM, <i>lacI</i> , <i>pBR ori</i>	CmR	This work
pLEOrT125	4.3	plac_CTU promoter, SD_T7 RBS, N-terminal His-tag, <i>attR2</i> , <i>rumT125</i> , <i>lacI</i> , <i>pBR ori</i>	CmR	This work
pLEOsrT125	4.6	plac_CTU promoter, SD_lac RBS, N-terminal His-Sumo-tag <i>attR2</i> , <i>rumT125</i> , <i>lacI</i> , <i>pBR ori</i>	CmR	This work
pRK793	7.4	tac promoter, <i>rrnB t1 terminator</i> , <i>lacI</i> , AmpR, <i>pBR</i> , MBP, TEV,	AmpR	Budisa's Group
pSEVA1810	3.8	pBAD promoter, <i>λt0 terminator</i> , AraC, AmpR, pUC ori	AmpR	Budisa's Group
pSEVA15A	3.7	pBAD promoter, <i>λt0 terminator</i> , AraC, AmpR, p15A ori,	AmpR	This work
pUA66	4.5	<i>lacZ</i> promoter, pSC101 ori, <i>gfpmut2</i> ,	KanR	Frank Delvigne's Group



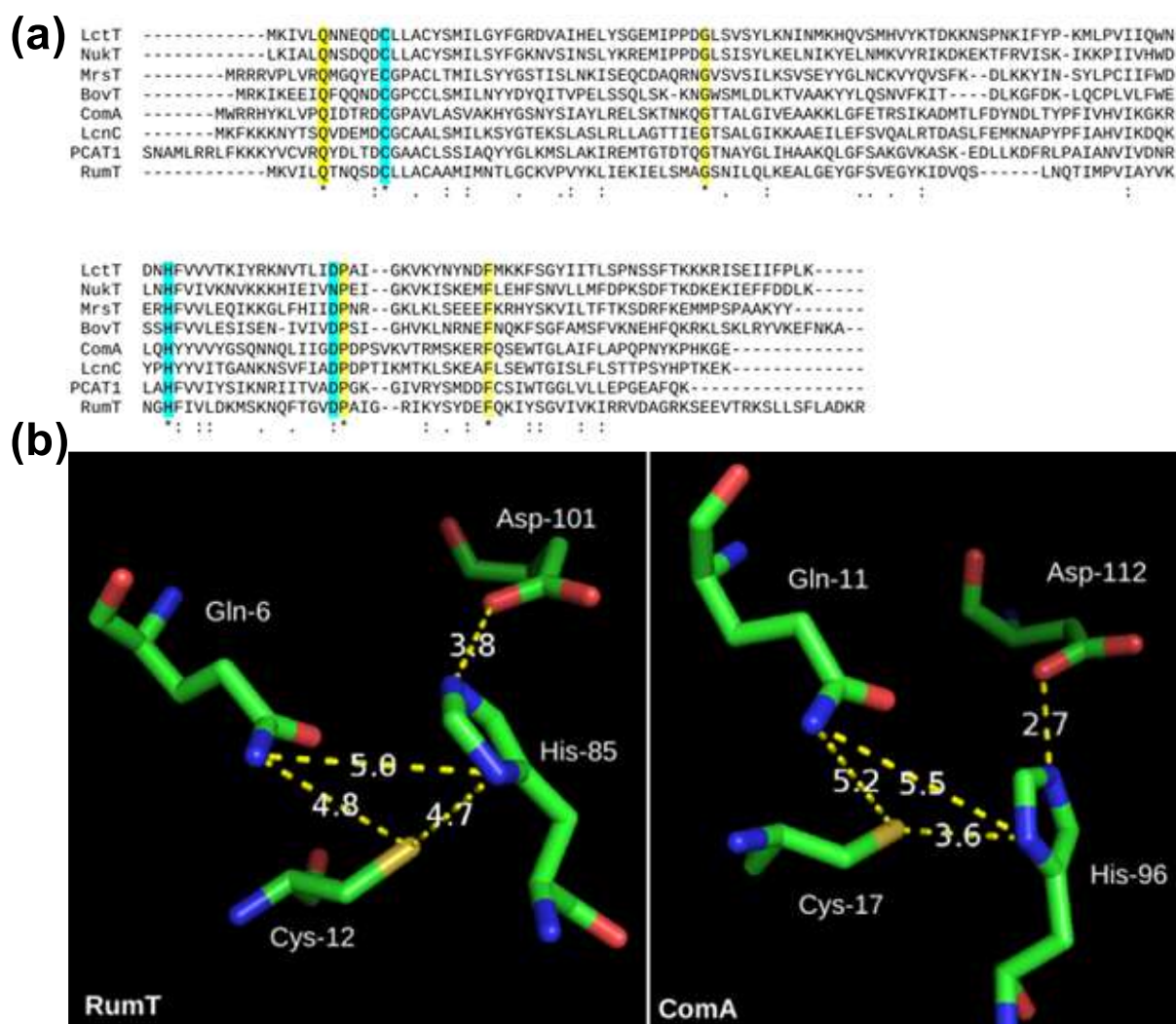
**Figure A1 Expression vectors description.** Maps of expression plasmids indicating all necessary features

### 6.3 Structure prediction and sequence similarity



**Figure A2 Predicted structures of AMS proteins.** All five class II LanT proteins show strong structural homology to PCAT1. Structures were predicted using Raptor X tool and subsequently edited with PyMOL 2.1 (Schrödinger)





**Figure A3** Conserve catalytic site residues of the peptidase domains of AMS protein and hydrogen bond predictions. (a) Multiple alignments of the peptidase domains of seven AMS proteins using CLUSTALW. Cyan, putative catalytic triad; yellow, other conserved residues. (b) Predictions of hydrogen bonding between catalytic site residues in putative RumT protease domain in comparison with actual structural data obtained from ComA (Ishii et al., 2010). The bond distances between atoms of the catalytic residues in the predicted structure of RumT (left) are comparably similar to those of the known crystal structure of ComA (right).

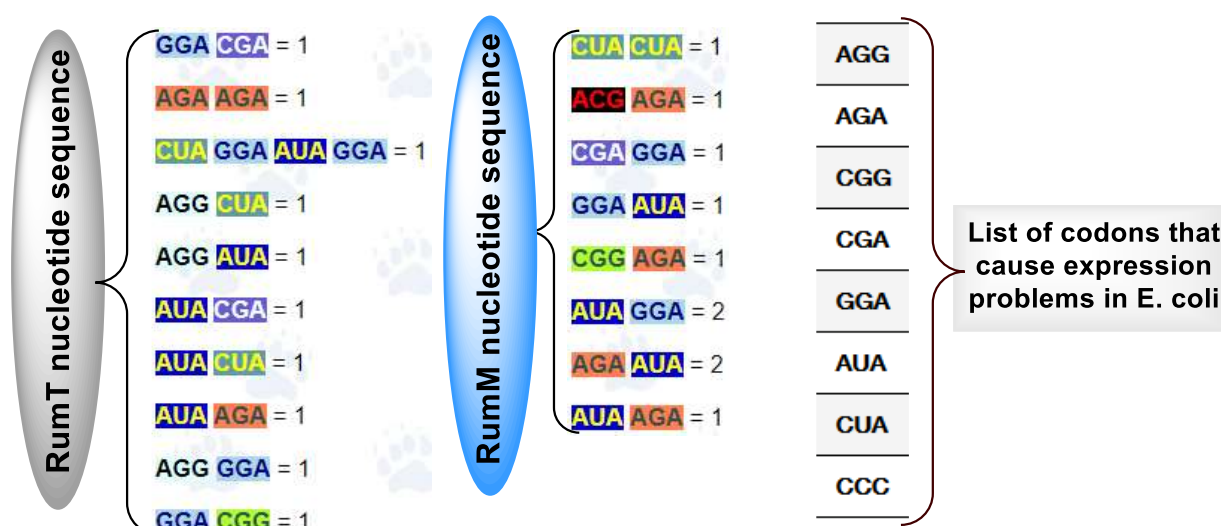


Figure A4 Repeated and/or Consecutive Rare Codons

## 6.4 Codon-optimized gene sequence of the peptidase-encoding domain of RumT (UniProt: Q93JP5)

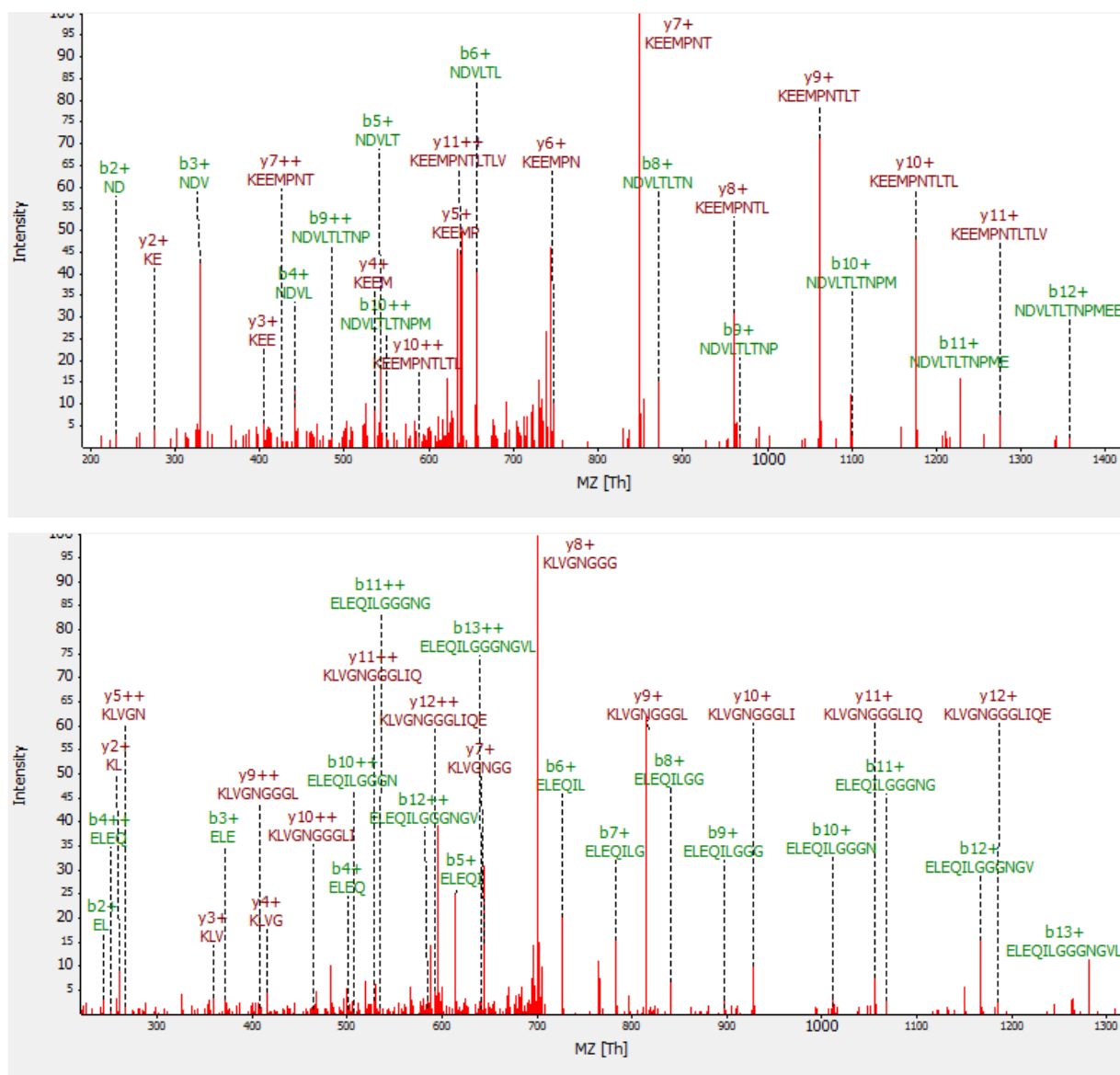
5'-

ATGCATCACCATCACCATCACGAAAACCTGTATTTTCAGGGCATGAAGGTGATTCTGCAA  
ACCAATCAGAGCGATTGTCTGCTGGCATGTGCAGCAATGATTATGAATACCCTGGGTTG  
TAAAGTGCCGGTGTATAAACTGATCGAAAAAATCGAACTGAGCATGGCAGGTAGCAATA  
TTCTGCAACTGAAAGAAGCACTGGGCGAATACGGTTTTAGCGTTGAAGGCTATAAAATC  
GATGTTTCAGAGCCTGAATCAGACCATTATGCCGGTTATTGCCTATGTGAAAAACGGCCA  
TTTTATCGTGCTGGACAAAATGAGCAAAAACAGTTTACCGGTGTTGATCCGGCAATTG  
GTCGTATCAAATATAGCTATGATGAATTCAGAAAAATCTATAGCGGTGTGATCGTGAAAA  
TCTAATAA-3'

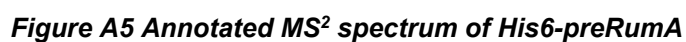
The yellow-shaded region indicates His6+TEV tag included to the sequence.

## 6.5 MS<sup>2</sup> assignments of His6-preRumA fragments

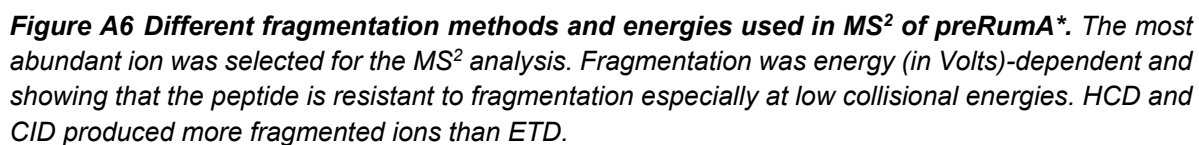


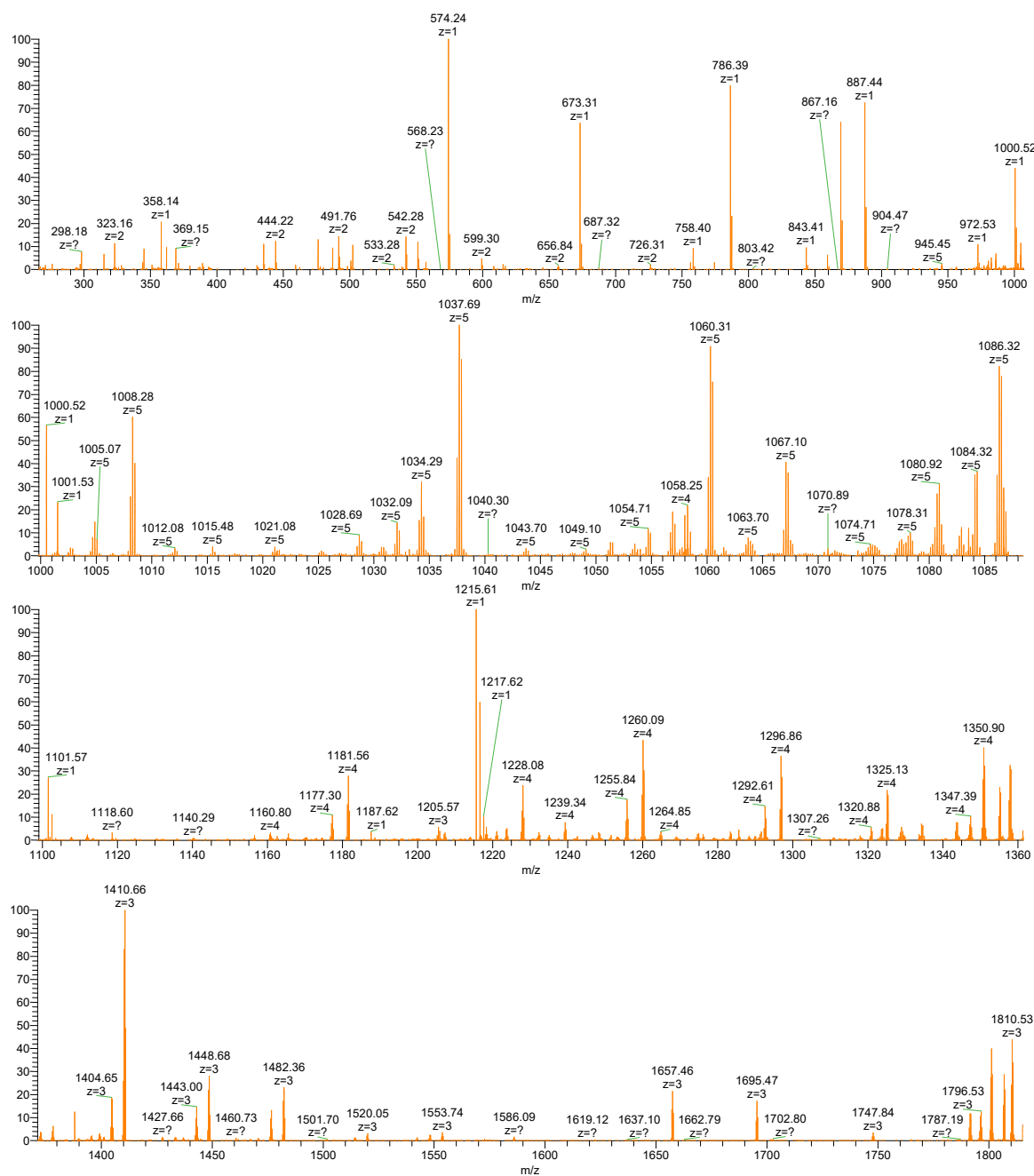


**Figure A5 Annotated MS<sup>2</sup> spectrum of His6-preRumA**



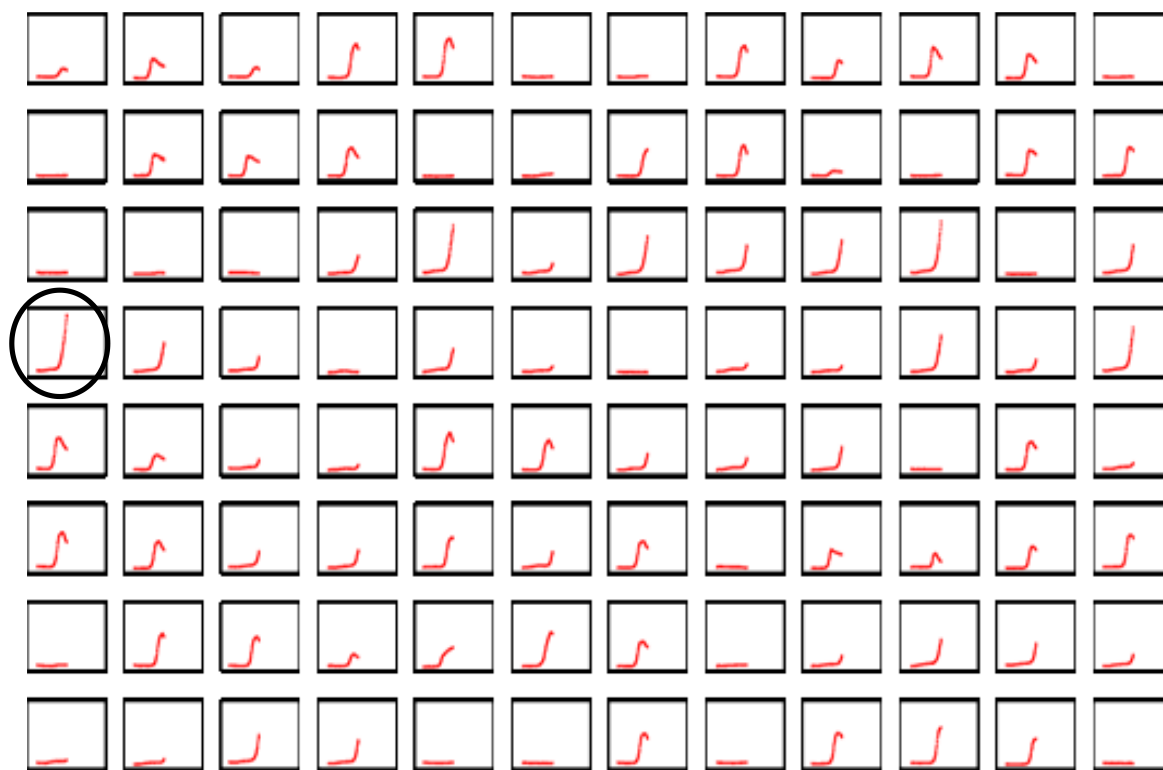
## 6.6 MS<sup>2</sup> characterization of preRumA\*





**Figure A7 MS<sup>2</sup> analysis of preRumA\*. Cleavage occurs at peptide bonds and the peaks in the mass spectrum represent charged ions produced as a result. The charge of each ionic product is also shown. The charges were calculated in cases where they were not detectable.**

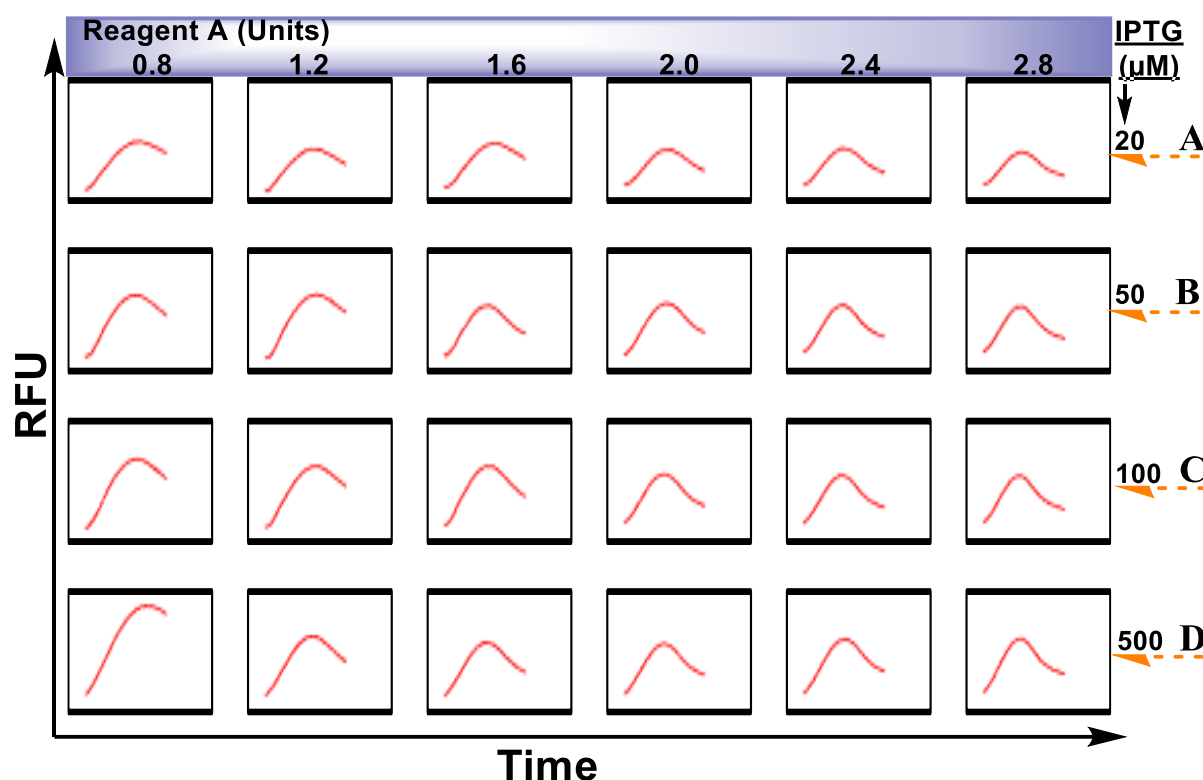
## 6.7 Multi-parallel cultivations



**Figure A8 Online GFP fluorescence curves showing identification of the most prominent strain expressing the pLEOgrA\*M1 plasmid from 96 selected colonies.** Cultures that did not produce any fluorescence signals were considered not to possess the correct plasmid. The strain that was selected for subsequent processing is circled in the figure.



**Figure A9 Online curves showing GFP fluorescence and OD<sub>600</sub> measurements.** (a) GFP fluorescence signals of colonies selected from LB agar plate growing WLEOgrA\*M1, and WLEOrA as control (column one only). (b) ODs of the respective cultures whose fluorescence signals are measured in (a)



**Figure A10** Experimental design and online curves displaying GFP fluorescence measurement of WLEOgrA\*M1 strain cultivated at different glucose feeding rates and IPTG induction concentrations.

## 6.8 The concept of mass spectrometry (MS)

### 6.8.1 Mass spectrometry and peptide sequencing

Proteomic studies engage wide application of mass spectrometry (MS) as a unique technique to identify and quantify diverse synthetic and natural compounds. The molecular structures of these compounds can also be determined using data obtained from MS analyses. As discussed earlier, the structures of many lanthipeptides and the catalytic mechanisms of their PTM enzymes have been characterized using MS data. Four major principles govern the design and utilization of an MS instrument for such analytical purposes namely; ionization of the analyte, separation of the ions based on their mass-to-charge ( $m/z$ ) ratios, fragmentation of selected ions and detection. The Fourier transforms of the signals with given  $m/z$  ratios are detected and presented as a mass spectrum (Marshall and Verdun, 2016). For this to be actualized, the mass spectrometer must have an ion source that introduces charges on the

vaporized analyte, a mass analyzer that separates the ions with respect to differences in their  $m/z$  ratios as they converge via a magnetic field and a detector that measures, amplifies and displays the output signals (Dass, 2007).

Most mass spectrometers that are used in studying the structure of larger molecular weight compounds like peptides are usually coupled to a high-performance liquid chromatographic (HPLC) systems to enhance separation and make both instruments more efficient (Arpino, 1992). HPLC like all other chromatographic techniques has a mobile phase (solvent) and a stationary phase (column). The analyte is injected into the mobile phase and a constant pressure is applied across the column to enable the mobile phase to move through. The analyte then interacts with the uniformly dispersed nanoporous silica or other polymers that embedded in the column to enable separation. Properties like hydrophobicity/hydrophilicity, size/conformation, primary structure and detection wavelengths are the principal determinants of separation principle of peptide and protein molecules by HPLC (Mant and Hodges, 2017). This study engaged the use of the reversed-phase chromatographic technique via an HPLC coupled to a mass spectrometer. This instrumental set-up was used for the electrospray ionization (ESI)-LC-MS analysis of extracted preRumA. The stationary phase of the instrument consisted of Poroshell 120 EC-C8 ( $2.1 \times 50$  mm,  $2.1 \mu\text{m}$ ) which are alkylated silica particles used for both the column and pre-column. The polar mobile phase was made up of water or buffer (Aguilar 2004). For this study, a gradient of water/acetonitrile/0.1 % formic acid mixture was used.

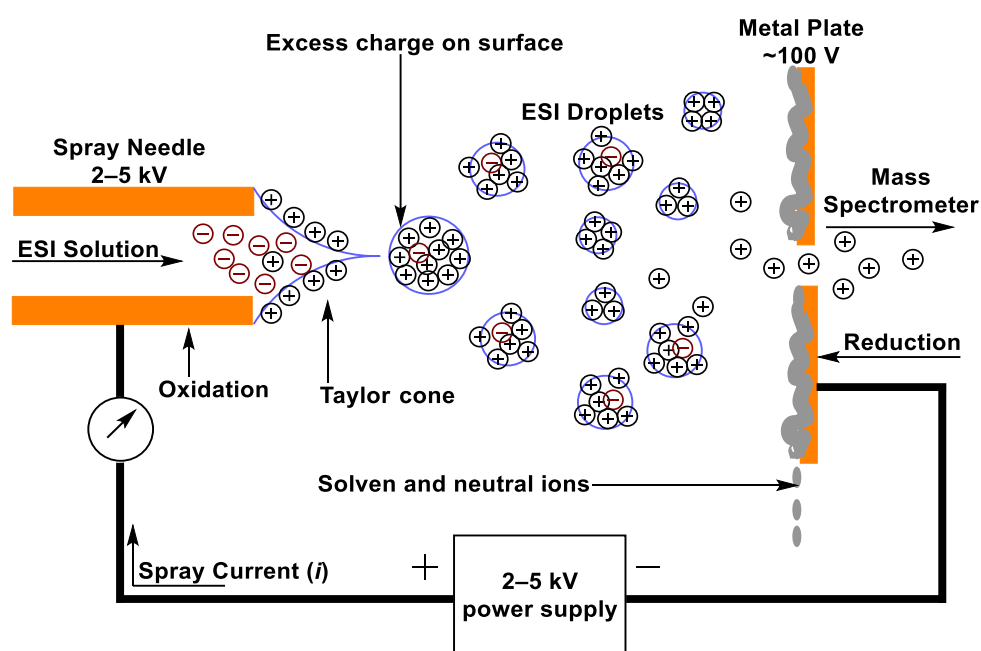
### 6.8.2 Electrospray ionization

There are two techniques that are commonly used to analyze liquid and solid biological samples namely; ESI (Fenn et al., 1989) and matrix-assisted laser desorption/ionization (MALDI) (Karas et al., 1987). ESI is the most common ionization method used in combined HPLC-MS (Pitt, 2009), and it involves applying a high voltage in the range of 2–5 kV to the liquid analyte passing through a capillary tube. ESI produces stable intact molecular ion species which are either protonated ( $[M + H]^+$ ) or deprotonated ( $[M + H]^-$ ) without any fragmentation occurring in the molecules during the ionization process (Watson and Sparkman, 2007). When set to the positive ion mode, the strong electric field allows positive ions to accumulate at the liquid surface present at the tip of the capillary, whereas in the negative ion mode, negative ions are propelled towards the inside region of the capillary.

The repulsive force within the enriched positive ions at the surface together with the pulling force exerted by the electric field on the positive ions overwhelm the surface tension of the



liquid which eventually expands into a cone called the “Taylor cone” (Kearle, 2000). The cone is very unstable such that it elongates into a liquid filament until the filament begins to pinch off into individual charged droplets (Figure A5.1). A gas sheath like hot nitrogen, flowing co-axially to the ESI solution facilitates the formation of droplets and solvent evaporation. Since ESI is performed at atmospheric pressure, the ions produced usually go through stepwise pressure reduction stages to create the level of vacuum required for mass analysis (Cech and Enke, 2001). The fast scanning speeds of ESI-MS enable measurements of multiple analytes in one analytical run (Pitt, 2009). In this study, the positive ion mode was used.

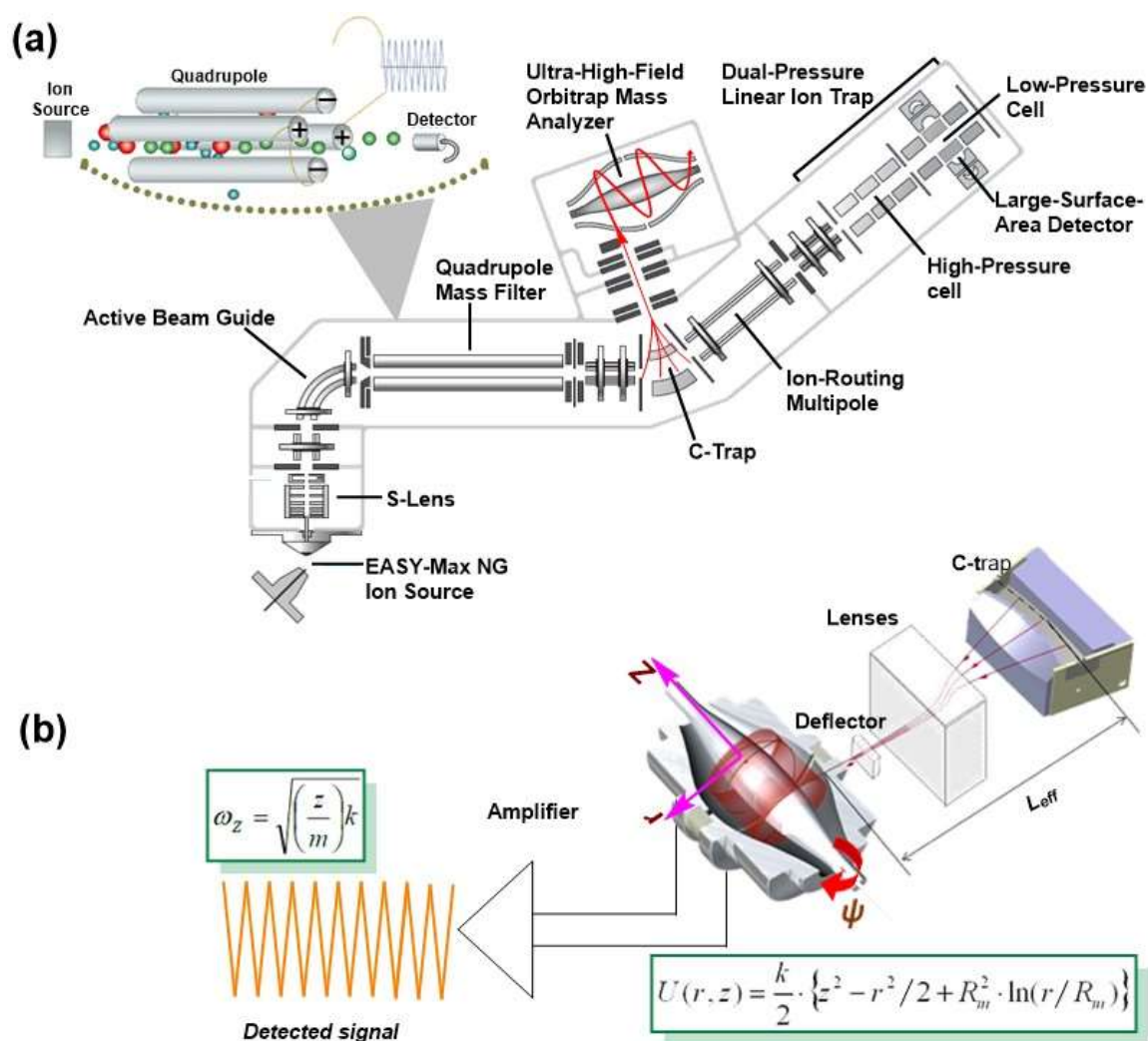


**Figure A5.1** The concept of Electrospray ionization as conceived by Cech and Enke (2001).

### 6.8.3 The Orbitrap Fusion Tribrid mass spectrometer

The schematic diagram of the Orbitrap Fusion Tribrid MS is shown in Figure A5.2. The mass spectrometer incorporates quadrupole, Orbitrap, and linear trap analyzers in a single system (Senko et al., 2013). Ions from an ESI source are directed through a stacked ring ion guide (S-lens) (Shaffer et al., 1997) which captures, focuses and transmits the ions to an electromagnetic active beam guide (ABG) made up of a curved quadrupole with an axial direct-current (DC) field (Senko et al., 2013). The ABG reduces noise by circumventing neutral charges from entering the quadrupole. While in the quadrupole mass filter, precursor ions are selected, allowing multiple parallel data acquisitions. The electromagnetic field of the

quadrupole mass filter also allows efficient ion transmission to the ion-routing multipole (Senko et al., 2013). In the full scan analysis, the precursor ions efficiently accumulate in ion-routing multipole where they are routed to the C-trap (Makarov et al., 2006). The C-trap then focuses the ions and inject them into the high scan speed orbitrap mass analyzer (Makarov et al., 2009, Michalski et al., 2012) for analysis and detection.



**Figure A5.2** Schematic description of the Orbitrap Fusion Tribrid mass spectrometer, showing (a) the Quadrupole mass filter as described by Glish and colleague (Glish and Vachet, 2003), (b) Measurement principle of the Orbitrap mass analyzer.

For MS<sup>2</sup> scans, one species of the precursor ion is selected in the quadrupole mass filter and directed to the ion-routing multipole where it is fragmented using higher-energy collisional dissociation (HCD) and routed back to the orbitrap mass analyzer for detection. The detection principle involves a central electrode with enhanced Fourier transform (Lange et al., 2014).

Additionally, the selected precursor may be forwarded to the dual-pressure linear ion trap (Pekar Second et al., 2009) where it is fragmented using collision-induced dissociation (CID) or electron transfer dissociation (ETD). The fragmented ions from CID and ETD are detected using a large-surface-area detector (Scientific, 2016). Furthermore, the instrument can also be used to perform MS<sup>n</sup> scans (Stratton et al., 2013).

### 6.8.4 The quadrupole mass filter

In the quadrupole mass analyzer, a DC and a radio-frequency (RF) potential are applied to a square symmetric array of four parallel metal rods, each rod having opposite polarities to its adjacent constituent. Figure A5.2a describes this arrangement from a top-view (Glish and Vachet, 2003). Whereas residual ions with varied  $m/z$  ratios cannot maintain uniform trajectories when DC and RF voltages are applied, accurate magnitudes of these energies supplied to the rods would allow ions of the same  $m/z$  to sustain stable trajectories from their source right up to the detector. Hence the quadrupole serve as a bandpass mass filter (Miller and Denton, 1986). The bandpass width is determined by the DC/RF ratio which can also be kept constant while altering their respective potentials to obtain a mass spectrum. Other operation modes are also available like the total-ion mode where only the RF is altered by making the potential of the DC component constant (zero). This allows the quadrupole to behave as a broadband mass filter enabling wide varieties of ion species with different  $m/z$  ratios to go through.

### 6.8.5 Ion trap mass analyzer

There are only few functional difference between ion traps and quadrupoles. The first derives from the ability of the ion trap to maintain ions in stable trajectories within the traps for a certain period of time before they are released for measurement. This therefore allows separation of masses as a function of time and not in space as for the quadrupoles. Secondly, the ion trap has oscillating field which can be adjusted to trap ions of designated  $m/z$  ratios (also referred to as mass-selective stability) or to trap all ion species and reject those with certain  $m/z$  ratios (also called mass-selective instability). Several types exist but here, we are going to discuss only the Orbitrap and the linear trap quadrupole (LTQ) which were used in this study.

The Orbitrap consist of a barrel-like electrode that shields a coaxial spindle-like electrode in the inside (Figure A5.2b). This combination traps and maintain ions in an orbital motion spinning around the spindle (Makarov, 2000, Hu et al., 2005). The motion of the trapped ions induces a current on the outer electrodes at a certain angular frequency ( $\omega_z$ ). Two symmetrical pick-up sensors are connected from the outer electrode to a differential amplifier which

converts the Fourier transform of this angular frequency signals to a mass spectrum (Figure A5.2b).

Ions are injected into the Orbitrap by ramping down RF voltages while applying DC gradients across the C-trap (a curved linear trap) to produce a laser-like ion source that are routed to the Orbitrap. Important to note is that in the Orbitrap, the inertia of the ions in the inner electrode balances their electrostatic attraction to the electrode causing them to be trapped, thereby forcing the ions to maintain elliptical trajectories with seemingly helical spatial arrangement (Makarov, 2000). The axial motion of these ions is harmonic and depends only on their  $m/z$  ratios. The angular frequency of the axial motion ( $\omega_z$ ) is expressed as shown in Figure A5.2b, where  $k$  represents the force constant of the potential. Since the ions move with the same axial frequency, but different rotational frequencies, the ions produce rings of specific  $m/z$  ratios which induce current on the outer electrode at specific frequencies.

An LTQ possesses similar features to the conventional quadrupole, but its own set of quadrupole rods are designed to possess three axial sections as illustrated in Figure A5.3a. This construction forces the ions to assume a radial topology, while a static electrical potential on each end of the electrodes confine the ions in an axial field (Douglas et al., 2005). The DC and RF of the three sections of the linear trap can be adjusted individually to allow it to function as a selective mass filter or trap ions in the center section (March, 2000).

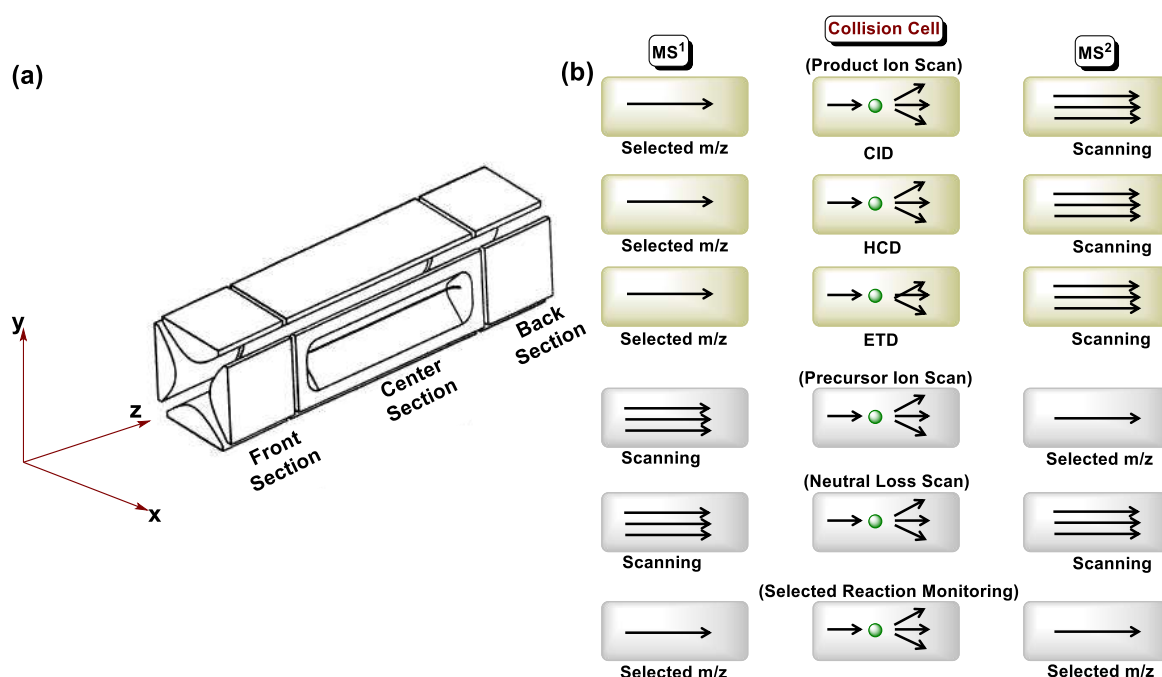
### 6.8.6 Tandem mass spectrometry (MS<sup>2</sup>)

Tandem MS is very important in studies involving complexly modified RiPPs like lanthipeptides as reported in many cases (Goto et al., 2010, Meindl et al., 2010, Müller et al., 2010, Sambeth and Süssmuth, 2011, Krawczyk et al., 2012, Völler et al., 2012, Férrir et al., 2013, Krawczyk et al., 2013, Völler et al., 2013). This is because data obtained from MS<sup>2</sup> experiments contain information that may help to determine the positions of the thioether cross-bridges and additional PTMs in the compound. Additional information like stereochemical configurations may not be easily derived from MS<sup>2</sup> data but chiral gas chromatography-mass spectrometry (GC/MS) (Garg et al., 2012, Garg et al., 2016) or crystallization (Meindl et al., 2010) or NMR in cases of overlapping rings (Hsu et al., 2004, Ekkelenkamp et al., 2005, Garg et al., 2012) may be used.

Tandem MS technique involves the combination of two mass analyzers in multiple mass analysis or selection steps, with fragmentation occurring at any given stage in between. It is also called MS/MS and it offers ideal analytical resources to characterize, identify, elucidate

and quantify compounds. In the first stage of analysis in a tandem mass spectrometer ( $MS^1$ ), ions are differentiated based on their  $m/z$  ratios. Specific ions with certain  $m/z$  ratio called the precursor ions are selected and fragmented using a collision cell made up of a mixture of reactive and collisional gasses. In modern instrumentation like the Orbitrap Fusion Tribrid Mass Spectrometer, product ions from fragmentation experiments are generated by CID, HCD or ETD. The resulting product ions from  $MS^1$  are further separated and detected in  $MS^2$  (the second stage). The pattern in which the product ions of a particular precursor ion are generated is peculiar to that particular precursor and this information can be used to identify a target compound in a complex mixture of different compounds in a sample. Varieties of acquisition modes can be used in  $MS^2$  (Figure A5.3b).

Product Ion Scan is a very common scanning technique where mass analyzer 1 basically functions as a mass filter, selecting only ions with certain  $m/z$ . These precursor ions are then fragmented by either methods and analyzed in mass analyzer 2 to produce the  $MS^2$  spectrum. Data obtained from this scan mode are useful for structural elucidation. Product Ion Scan was used in this study to characterize the structure of modified preRumA.



**Figure A5.3** (a) Schematic diagram of a 2D linear ion trap showing the three axial sections which can be independently adjusted using DC and RF voltages. (b) Modes of data acquisition in tandem MS

Another scanning mode whose data can supply information about closely related compounds in a sample mixture is the Precursor Ion Scan where all resulting fragmented product ions in  $MS^2$  are scanned from  $MS^1$ . Neutral loss scans is a scanning mode that may acquire

information useful in different situations like the Precursor Ion Scan. Finally, Selected Reaction Monitoring (SRM) is a data acquisition used to quantify specific analytes in a complex mixture. In this case, precursor ion with a certain  $m/z$  are selected in  $MS^1$  and only these ions are fragmented and detected in  $MS^2$ , meanwhile only product ions with a characteristic pattern are monitored. Quantification is achieved by comparing the peak area of the specific product ion to an internal standard.

## **6.9 Tables and Figures for materials and methods section**

**Table A5.1** List of biological reagents/kits

<b>Enzymes and markers</b>	<b>Source</b>
Benzonase	Merck, Darmstadt, Germany
DreamTaq Green DNA Polymerase	Thermo Scientific, Waltham, USA
FastDigest Restriction Enzymes	Thermo Scientific, Waltham, USA
FastAP Thermosensitive Alkaline Phosphatase	Thermo Scientific, Waltham, USA
GeneRuler 100 bp Plus DNA Ladder	Thermo Scientific, Waltham, USA
GeneRuler 1 kb Plus DNA Ladder	Thermo Scientific, Waltham, USA
Lambda Mix Marker, 19	Thermo Scientific, Waltham, USA
Lysozyme	Sigma-Aldrich, St. Louis, USA
Phusion High-Fidelity DNA Polymerase	New England Biolabs (NEB), Frankfurt am Main, Germany
Q5 High-Fidelity DNA Polymerase	NEB, Massachusetts, USA
Roti-Mark TRICOLOR	Carl Roth (Karlsruhe, Germany)
Spectra Multicolor Broad Range Protein Ladder	Thermo Scientific, Waltham, USA
T4 DNA Ligase	Thermo Scientific, Waltham, USA
T4 DNA Ligase	NEB, Frankfurt am Main, Germany
<b>Antibodies Name</b>	<b>Manufacturer</b>
Penta His Antibody, mouse monoclonal	Qiagen, Hilden
Anti-Mouse IgG, Alkaline Phosphatase labeled	Sigma-Aldrich, Hamburg
<b>Nucleic acid stains</b>	<b>Source</b>
GelRed Nucleic Acid Gel Stain	Biotium, Hayward, USA
Ethidium bromide	Carl Roth, Karlsruhe, Germany
<b>Kits</b>	<b>Source</b>
HiYield PCR Clean-up/Gel Extraction Kit	Süd-Laborbedarf, Gauting, Germany
GenElute™ Bacterial Genomic DNA Kit	Sigma-Aldrich (St. Louis, USA)
Invisorb Spin Plasmid Mini Two Kit	Stratec, Berlin, Germany
Q5 Site-Directed Mutagenesis Kit	Frankfurt am Main, Germany

### 6.9.1 Instrumentation used

All instruments used are listed in Table A5.2

**Table A5.2** List of instruments

<b>Instrument</b>	<b>Name</b>	<b>Manufacturer</b>
Analytical balance	Sartorius	Sartorius AG, Göttingen
Autoclave	Systec V95	Systec GmbH, Wettenberg
Centrifuge	Himac/CT15RE	VWR, Leuven
Centrifuge	Eppendorf centrifuge 5810R	Eppendorf, Hamburg

Culture monitoring online	SDR SensorDish Reader	PreSens, Regensburg
DNA electrophoresis units	Biozym	Biozym, Hessisch Oldendorf
Electroblotting device	semiDRY	Cleaver Scientific, UK
Electroporator	electroporator 2510	Eppendorf, Hamburg
Gel documentation	GenoPlex	VWR, Leuven
HPLC	Agilent 1290 Infinity UHPLC	Agilent Technologies, Waldbronn
HPLC column	Zorbax Eclipse Plus C18 RRHD 2.1 x 50 mm, 1.8 µm	Agilent Technologies, Waldbronn
HPLC column	Poroshell 120 EC-C8 (2.1 × 50 mm, 2.1 µm)	Agilent Technologies, Waldbronn
Incubation shaker	Infors HT	Infors, Switzerland
Incubation shaker	Kühner LT-X	Adolf Kühner, Switzerland
Incubator	Heraeus	Heraeus, Berlin
Liquid-handling Robot	Microlab STAR	Hamilton, Martinsried
Mass spectrometer	Agilent 6460 TripleQuad LC/MS	Agilent Technologies, Waldbronn
Mass spectrometer	Orbitrap Fusion Tribrid	Thermo Scientific, Waltham, USA
Mass spectrometer	LTQ Orbitrap XL	Thermo Scientific, Waltham, USA
Mass spectrometer	QTrap2000	AB SCIEX, Darmstadt
PCR Thermal cycler	Mastercycler gradient	Eppendorf, Hamburg
pH Meter	Mettler Toledo	Mettler-Toledo, Giessen
Plate reader	Synergy Mx	BioTek Instruments, Vermont, USA
Protein electrophoresis unit	omniPAGE mini	Cleaver Scientific, UK
Protein IMAC column	Ni-NTA HisTrap FF Crude 1-ml column	GE Healthcare, Freiburg
Protein purification system	Äkta avant 25	GE Healthcare, Freiburg
Scanner	Konica Minolta	KONICA MINOLTA, Tokyo, J
Shaker water bath	GFL	Labortechnik, Burgwedel
Size exclusion column	HiLoad 16/60 Superdex 200	GE Healthcare, Freiburg
Sonication bath	Bandelin Sonorex	Schalltec, Mörfelden-Walldorf
Spectrophotometer	Ultrospec 3300	Amersham, Germany
Spectrophotometer	NanoDrop ND-1000	Peqlab Biotechnologie, Erlangen
Speed vacuum dryer	Concentrator Plus vacuum concentrator	Eppendorf, Hamburg
Thermomixer	Thermomixer comfort	Eppendorf, Hamburg
Ultrasonicator	UP200S	Hielscher Ultrasonics, Teltow
UV Transilluminator	GenoView	VWR, Leuven
Vortex	Vortex Mixer	Gemmy Industrial, Taiwan



## 6.9.2 Expendables

Single-use materials are listed in Table A5.3

**Table A5.3** List of expendables

<b>Name</b>	<b>Manufacturer</b>
24-Well Deep Well Plates	Ritter GmbH, Schwabmünchen
96-Well Microplate,	Costar, New York, USA
AirOTop Enhanced Seals	BioSilta Oy, Oulu, Finland
Amicon Ultra centrifugation tubes	Merk Millipore, Darmstadt
Breathable Film	Starlab, Hamburg
Centrifuge tubes, conical (15 ml, 50 ml)	VWR, USA
Cryo tubes (2 ml)	VWR, Leuven
Electroporation cuvettes (1 mm)	VWR, Leuven
His SpinTrap columns	GE Healthcare, Freiburg
Micro test tubes (2.0 ml, 1.5 ml)	Carl Roth, Karlsruhe
Nitrocellulose membrane Protran BA 85	Whatman, UK
PCR tubes / strips 0.2 ml	Biozym, Hessisch Oldendorf
Petri dishes (ø 90 mm)	Greiner Bio-One, Frickenhausen
PreSens HydroDish HD24-well	PreSens, Regensburg
PreSens OxoDish OD24-well	PreSens, Regensburg
Roti-Store cryo vials	Carl Roth, Karlsruhe
Syringe filter CME 0.45 µm	Carl Roth, Karlsruhe
Syringes (5 ml, 10 ml, 20 ml)	Terumo Corporation, Tokyo, Japan
UltraYield Flasks (125 ml, 250 ml, 500 ml, 2.5 l)	BioSilta Oy, Oulu, Finland
Whatman 3 MM Filter paper	Whatman, UK
ZipTip pipette tips	Merck, Darmstadt

## 6.9.3 Software

Software used in this work are listed in Table A5.4

**Table A5.4** List of software

<b>Name</b>	<b>Manufacturer</b>
Äkta UNICORN 6.1	GE Healthcare, Freiburg
Analyst 1.4.2	AB SCIEX, Darmstadt
ChemBioDraw Ultra 14.0	PerkinElmer, Waltham, US
Gen5	Biotek, Vermont, US
MassHunter Qualitative Analysis Software	Agilent Technologies, Waldbronn
MASCOT	Matrix Science, UK
MaxQuant	Cox's Group, Max Planck Institute of Biochemistry
MicrolabSTAR VENUS one	Hamilton, Martinsried
PreSens Measurement Studio 2 version 2.1.0	PreSens, Regensburg
PyMOL 2.1	Schrödinger, New York, US
SDR version 4.0	PreSens, Regensburg
SigmaPlot 12	Systat Software, Erkrath

Vector NTI Advance 11.5  
Xcalibur

Invitrogen, Karlsruhe  
Thermo Scientific, Waltham, USA

**Table A5.5** Description of major microbial strains used in this study

Name of strain	Description	Purpose	Source
<i>E. coli</i> W3110	wild type like <i>E. coli</i> ; Sex: F- Chromosomal markers: $\lambda^-$ <i>rph-1</i> INV( <i>rrnD</i> , <i>rmE</i> )	expression host	Stock center, Yale
<i>E. coli</i> ccdB+	F <sup>-</sup> <i>mcrA</i> $\Delta$ ( <i>mrr-hsdRMS-mcrBC</i> ) $\phi$ 80 <i>lacZ</i> $\Delta$ M15 $\Delta$ <i>lacX74</i> <i>recA1</i> <i>ara</i> $\Delta$ 139 $\Delta$ ( <i>ara-leu</i> )7697 <i>galU</i> <i>galK</i> <i>rpsL</i> (Str <sup>R</sup> ) <i>endA1</i> <i>nupG</i> <i>fhuA::IS2</i>		Invitrogen
<i>Ruminococcus gnavus</i> E1	wild-type	isolation of genomic DNA	(Dabard et al., 2001)
<i>B. subtilis</i> ATCC 6633	—	biological activity test	DSM 347
<i>E. coli</i> Top10	F <sup>-</sup> <i>mcrA</i> $\Delta$ ( <i>mrr-hsdRMS-mcrBC</i> ) $\phi$ 80 <i>lacZ</i> $\Delta$ M15 $\Delta$ <i>lacX74</i> <i>recA1</i> <i>araD</i> 139 $\Delta$ ( <i>ara leu</i> )7697 <i>galU</i> <i>galK</i> <i>rpsL</i> (Str <sup>R</sup> ) <i>endA1</i> <i>nupG</i> , $\lambda^-$	cloning & plasmid storage	Invitrogen
<i>E. coli</i> DH5 $\alpha$	F <sup>-</sup> <i>endA1</i> <i>glnV44</i> <i>thi-1</i> <i>recA1</i> <i>relA1</i> <i>gyrA96</i> <i>deoR</i> <i>nupG</i> <i>purB20</i> $\phi$ 80 <i>dlacZ</i> $\Delta$ M15 $\Delta$ ( <i>lacZYA-argF</i> )U169, <i>hsdR17</i> ( <i>r<sub>K</sub><sup>-</sup> m<sub>K</sub><sup>+</sup></i> ), $\lambda^-$	cloning	Invitrogen
<i>E. coli</i> BL21	F <sup>-</sup> <i>ompT</i> <i>gal</i> <i>dcm</i> <i>lon</i> <i>hsdS<sub>B</sub></i> ( <i>r<sub>B</sub><sup>-</sup> m<sub>B</sub><sup>-</sup></i> ) [ <i>malB</i> <sup>+</sup> ] <sub>K-12</sub> ( $\lambda^S$ )	expression host	Novagen
<i>E. coli</i> Rosetta 2	F <sup>-</sup> <i>ompT</i> <i>hsdS<sub>B</sub></i> ( <i>r<sub>B</sub><sup>-</sup> m<sub>B</sub><sup>-</sup></i> ) <i>gal</i> <i>dcm</i> <i>pRARE2</i> (Cam <sup>R</sup> )	rare codon expression host	Novagen
<i>E. coli</i> Rosetta 2 (DE3)	F <sup>-</sup> <i>ompT</i> <i>hsdS<sub>B</sub></i> ( <i>r<sub>B</sub><sup>-</sup> m<sub>B</sub><sup>-</sup></i> ) <i>gal</i> <i>dcm</i> (DE3) <i>pRARE2</i> (Cam <sup>R</sup> )		
WLEOrA	Cm <sup>R</sup> , W3110 carrying pLEOrA	expression of RumA	This work

WLEOrM'	Cm <sup>R</sup> , W3110 carrying pLEOrM'	expression of RumM	This work
WLEOrM1	Cm <sup>R</sup> , W3110 carrying pLEOrM1	expression of RumM	This work
WLEOrA/M	Kan <sup>R</sup> , Cm <sup>R</sup> , W3110 carrying pLEOrA and pLEOrM	expression of mRumA & RumM	This work
WLEOgrA	Cm <sup>R</sup> , W3110 carrying pLEOgfrA	expression of GFP-RumA	This work
WLEOgrA/M	Kan <sup>R</sup> , Cm <sup>R</sup> , W3110 carrying pLEOgfrA and pLEOrM	expression of mGFP-RumA & RumM	This work
WLEOgrAM	Cm <sup>R</sup> , W3110 carrying pLEOgfrAM	expression of mGFP-RumA & RumM	This work
WLEOgrA*M1	Cm <sup>R</sup> , W3110 carrying pLEOgrA*M1	expression of mRumA* and RumM	This work
WLEOsrA*	Cm <sup>R</sup> , W3110 carrying pLEOsrA*	expression of SUMO-RumA*	This work
WLEOsrA*M	Cm <sup>R</sup> , W3110 carrying pLEOsrA*M	Expression of mSUMO-RumA* & RumM	This work

**Table A5.6** Growth media recipes**LB Medium (Bertani 1951)**

Yeast extract	5 g l <sup>-1</sup>
Tryptone	10 g l <sup>-1</sup>
NaCl	10 g l <sup>-1</sup>
pH 7.5	
(Agar)	1.50%

**M+ medium**

Peptone	5.0 g l <sup>-1</sup>
Meat extract	3.0 g l <sup>-1</sup>
MnSO <sub>4</sub>	0.02 g l <sup>-1</sup>
(Agar)	1.50%
pH 7.0	

**Terrific Broth (Tartoff and Hobbs 1987)**

Tryptone	12 g l <sup>-1</sup>
Yeast extract	24 g l <sup>-1</sup>
Glycerol	0.40%
Autoclaved separately:	
K <sub>2</sub> HPO <sub>4</sub>	0.072 M
KH <sub>2</sub> PO <sub>4</sub>	0.017 M

**SOC Medium (Hanahan 1983)**

Yeast extract	0.5 % (w/v)
Tryptone	2 % (w/v)
NaCl	10 mM
KCl	2.5 mM
MgSO <sub>4</sub>	20 mM
Glucose	20 mM

**Brain Heart Infusion (BHI)-YH Broth**

Solid BHI	6 g l <sup>-1</sup>
-----------	---------------------

**Superbroth (Atlas 2004)**

Tryptone	35 g l <sup>-1</sup>
----------	----------------------

Yeast extract	5 g l <sup>-1</sup>	Yeast extract	20 g l <sup>-1</sup>
Hemin	5 g l <sup>-1</sup>	NaCl	5 g l <sup>-1</sup>
		NaOH (1N)	0.50%
		pH 7.0	

**Table A5.7** Antibiotics used in cultivation media

Antibiotic	Dissolved in	Stock solution	Concentration in media
Chloramphenicol (Cm)	Ethanol	34 mg ml <sup>-1</sup>	34 µg ml <sup>-1</sup>
Kanamycin (Kan)	ddH <sub>2</sub> O	50 mg ml <sup>-1</sup>	50 µg ml <sup>-1</sup>
Ampicillin (Amp)	ddH <sub>2</sub> O	100 mg ml <sup>-1</sup>	100 µg ml <sup>-1</sup>

**Table A5.8** Oligonucleotides used in this work

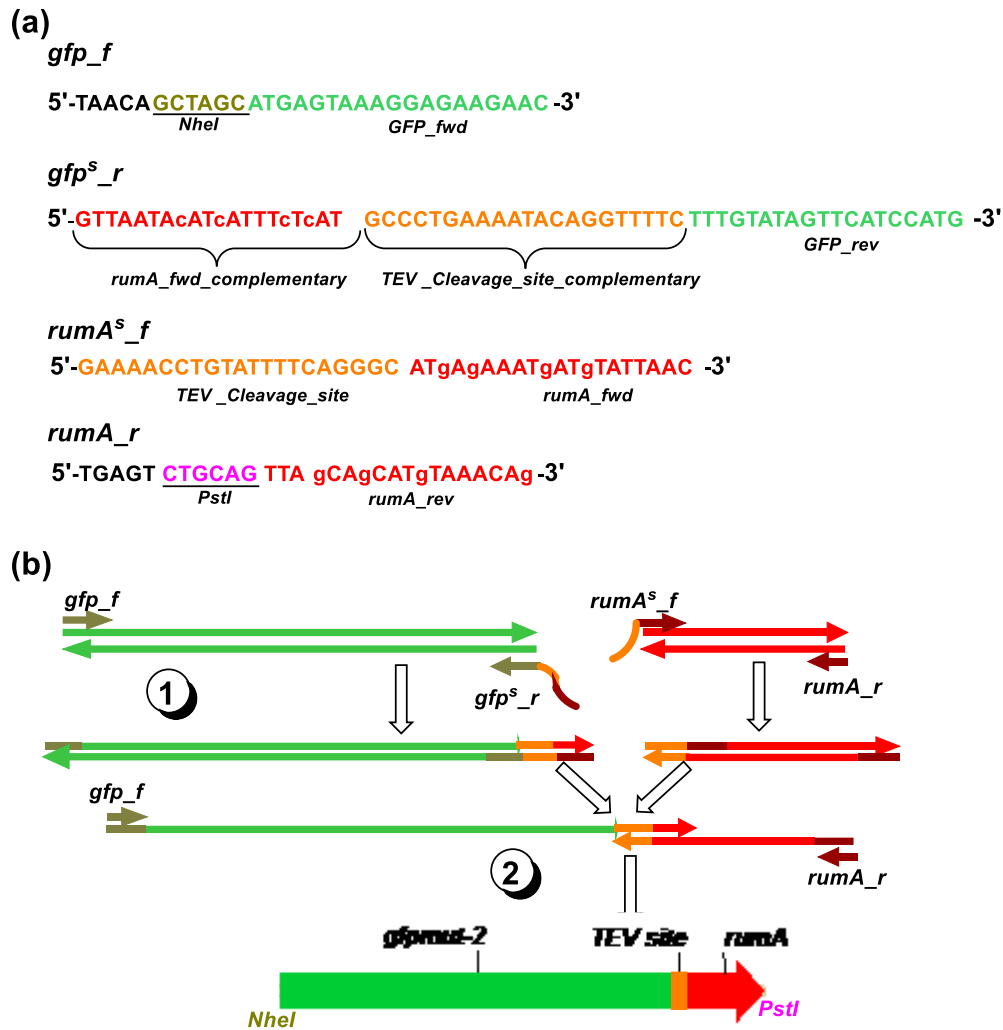
Name_	Oligonucleotide sequence	Restriction Enzyme
orientation		
<b>Isolation of <i>ra1A2A3M</i> from <i>R. gnavus</i> E1 genome</b>		
<i>rumClus_f</i>	ATGAGAAATGATGTATTAACATTAACAAAC	—
<i>rumClus_r</i>	TTACACAGTTTCAAGCATTAAACAGAG	—
<b>Cloning of <i>rumA</i></b>		
<i>rumA_f</i>	AACAG <b>CTAGC</b> ATGAGAAATGATGTATTAAC	<i>NheI</i>
<i>rumA_r</i>	TGAGT <b>CTGCAG</b> TTAGCAGCATGTAAACAG	<i>PstI</i>
<i>rumA_r2</i>	TGAGT <b>AAGCTT</b> TTAT <b>TTA</b> GCAGCATGTAAACAG	<i>HindII</i>
<b>Cloning of <i>rumM</i></b>		
<i>rumM_f</i>	TCGA <b>GGTACC</b> ATGCATAAGAAGTTTTGTGG	<i>KpnI</i>
<i>rumM_f1</i>	TCGA <b>GCTAGC</b> ATGCATAAGAAGTTTTGTGG	<i>NheI</i>
<i>rumM_f3</i>	TCGAG <b>GGATCC</b> ATGCATAAGAAGTTTTGTGG	<i>BamHI</i>
<i>rumM_r</i>	AAGT <b>CTGCAG</b> TTACACAGTTTCAAGCATTAA	<i>PstI</i>
<i>rumM_r2</i>	GAGT <b>AAGCTT</b> TTACACAGTTTCAAGCATTAAAC	<i>HindII</i>
<b>Cloning <i>rumM</i> and <i>rumA</i> or <i>gfpmut2-rumA</i> on the same vector</b>		
<i>prumM_f</i>	AACA <b>CTGCAG</b> AACGTTACTGGTTTCACATTC	<i>PstI</i>
<i>prumM_r</i>	GAGT <b>AAGCTT</b> TTACACAGTTTCAAGCATTAA	<i>HindII</i>
<i>rumM_f2</i>	AACA <b>CTGCAG</b> GAGCGGTACCCGATAAAAG	<i>PstI</i>
<i>rumM_r2</i>	GAGT <b>AAGCTT</b> TTACACAGTTTCAAGCATTAAAC	<i>HindII</i>
<b>Fusing <i>rumA</i> to <i>gfpmut2</i></b>		
<i>gfp_f</i>	TAACAG <b>CTAGC</b> ATGAGTAAAGGAGAAGAAC	<i>NheI</i>

<i>gfp<sup>+</sup>_r</i>	GTTAATACATCATTTCTCATGCCCTGAAAATACAGGTTTT CTTTGTATAGTTCATCCATG	—
<i>rumA<sup>+</sup>_f</i>	GAAAACCTGTATTTTCAGGGCATGAGAAATGATGTATTA AC	—
<b>Site directed mutagenesis primers</b>		
<i>pTrypsin_f</i>	GATCTTGGGT <b>AGA</b> GGTAATGGTG	—
<i>pTrypsin_r</i>	TGCTCCAGTTCTTTTTCTTC	—
<i>pGluC_f</i>	GATCTTGGGT <b>GAA</b> GGTAATGGTG	—
<i>pFactor Xa_f</i>	<b>GGTAGA</b> GGTAATGGTGTGTAAAAAC	—
<i>pFactor Xa_r</i>	<b>CTCTAT</b> CAAGATCTGCTCCAGTTC	—
<i>pTEV<sub>f</sub><sup>opt</sup></i>	<b>TATTTTCAG</b> GGTAATGGTGTGCTGAAAAC	—
<i>pTEV<sub>r</sub><sup>opt</sup></i>	<b>CAGGTTTTCT</b> AAAAATTTGTTCCAGTTCATTTTC	—
<b>Vectors modification</b>		
<i>SDT7_f</i>	TTGT <b>ACTAGT</b> TCACCTGCCCCGCTTTCCAG	<i>SpeI</i>
<i>SDT7_r</i>	GAGT <b>CATATG</b> TATATCTCCTTCTCTAGAAATTCGTAATC	<i>NdeI</i>
<i>P15A_f</i>	AT <b>GGCCGGCC</b> GCTCTAGCGGAGTGTATACT	<i>FseI</i>
<i>P15A_r</i>	<b>AGGCGCGCC</b> ACAACTTATATCGTATGGGGC	<i>AscI</i>
<i>AmpR_f</i>	GTGGT <b>AAGCTT</b> CTTGACCTGTG	<i>HindIII</i>
<i>AmpR_r</i>	TGAGCAAAC <b>TGGCCTCAG</b>	<i>Bsu36I</i>
<b>Cloning of N-terminal domain of RumT</b>		
<i>rumT125<sub>f</sub><sup>opt</sup></i>	AACAG <b>CTAGC</b> ATGCATCACCATCACCATC	<i>NheI</i>
<i>rumT125<sub>r</sub><sup>opt</sup></i>	TGAGT <b>CTGCAG</b> TTATTAGATTTTCACGATCACAC	<i>PstI</i>

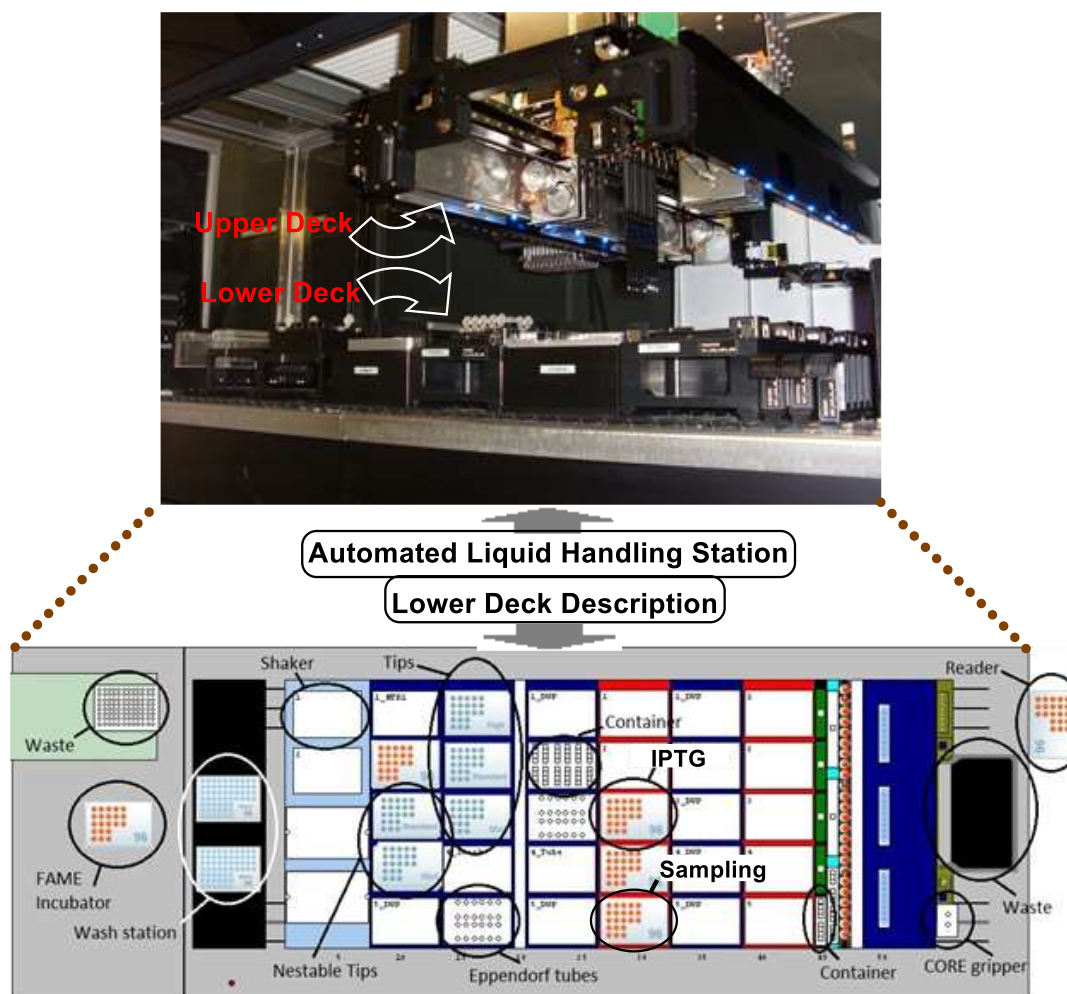
---

Forward primer, *f*; reverse primer, *r*; splicing primers, <sup>+</sup>; optimized gene, <sup>opt</sup>; restriction sites, bold; insertion sequences, bold green-coloured; tandem stop codon, bold underlined.

---



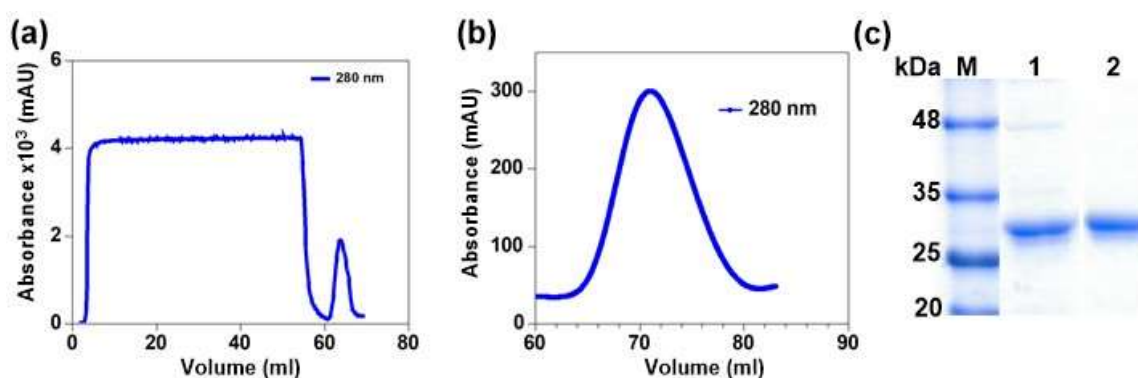
**Figure A5.4** Fusion primers design and SOE PCR. (a) Primer design for SOE experiments. (b) Schematic description of SOE experimental procedure (1) Both *gfpmut2* and *rumA* sequences were amplified separately using the primers labelled on the diagram. (2) The products of the two reactions were mixed in a megaprimer PCR reaction to generate an expressible chimeric construct



**Figure A5.5** The Hamilton liquid handling station. The lower deck of robot has compartments that can handle different forms of containers and expendables including microtiter plates, Eppendorf tubes, pipette tips and waste handlers. Also incorporated to the station are a shaker, PCR thermocycler, a plate reader and an incubator.

## 6.10 Purification of TEV

TEV was used in different aspects of this work as already described in previous sections. *E. coli* W3110 was transformed with the expression vector pRK793 and subsequently cultivated in TB medium. Extracted lysates were purified by IMAC followed by size exclusion chromatography using the Äkta purification system. Figure A5.6a depicts IMAC Ni<sup>2+</sup>-NTA purification while Figure A5.6b shows size exclusion chromatographic purification. We obtained pure TEV after the gel filtration run as indicated by a single band on SDS-PAGE (Figure A5.6c). 100 µl aliquots of the protein (eluted with gel filtration buffer containing 20 % glycerol) were obtained and stored at -20 °C for further use.



**Figure A5.6** TEV purification. (a) Chromatogram representing IMAC Ni<sup>2+</sup>-NTA purification, (b) Size exclusion chromatographic elution of TEV and (c) SDS-PAGE analyses of IMAC-purified (lane1) and gel filtration-purified (lane2) samples



## References

- Acedo JZ, van Belkum MJ, Lohans CT, McKay RT, Miskolzie M, Vederas JC (2015) Solution structure of acidocin B, a circular bacteriocin produced by *Lactobacillus acidophilus* M46. *Applied and Environmental Microbiology* 81:2910-2918
- Acedo JZ, van Belkum MJ, Lohans CT, Towle KM, Miskolzie M, Vederas JC (2016) Nuclear magnetic resonance solution structures of lacticin Q and aureocin A53 reveal a structural motif conserved among leaderless bacteriocins with broad-spectrum activity. *Biochemistry* 55:733-742
- Acedo JZ, Towle KM, Lohans CT, Miskolzie M, McKay RT, Doerksen TA, Vederas JC, Martin-Visscher LA (2017) Identification and three-dimensional structure of carnobacteriocin XY, a class IIb bacteriocin produced by Carnobacteria. *FEBS Letters* 591:1349-1359
- Agustin J, Rosenstein R, Wieland B, Schneider U, Schnell N, Engelke G, Entian K-D, Götz F (1992) Genetic analysis of epidermin biosynthetic genes and epidermin-negative mutants of *Staphylococcus epidermidis*. *European Journal of Biochemistry* 204:1149-1154
- AlKhatib Z, Lagedroste M, Fey I, Kleinschrodt D, Abts A, Smits SH (2014) Lantibiotic immunity: Inhibition of nisin mediated pore formation by Nisl.
- Altschul SF, Gish W, Miller W, Myers EW, Lipman DJ (1990) Basic local alignment search tool. *Journal of Molecular Biology* 215:403-410
- Alvarez-Sieiro P, Montalbán-López M, Mu D, Kuipers OP (2016) Bacteriocins of lactic acid bacteria: extending the family. *Applied Microbiology and Biotechnology* 100:2939-2951
- Arbulu S, Lohans CT, van Belkum MJ, Cintas LM, Herranz C, Vederas JC, Hernández PE (2015) Solution structure of enterocin HF, an antilisterial bacteriocin produced by *Enterococcus faecium* M3K31. *Journal of Agricultural and Food Chemistry* 63:10689-10695
- Arnison PG, Bibb MJ, Bierbaum G, Bowers AA, Bugni TS, Bulaj G, Camarero JA, Campopiano DJ, Challis GL, Clardy J (2013) Ribosomally synthesized and post-translationally modified peptide natural products: overview and recommendations for a universal nomenclature. *Natural Product Reports* 30:108-160
- Arpino P (1992) Combined liquid chromatography mass spectrometry. Part III. Applications of thermospray. *Mass Spectrometry Reviews* 11:3-40
- Aymerich T, Holo H, Håvarstein LS, Hugas M, Garriga M, Nes IF (1996) Biochemical and genetic characterization of enterocin A from *Enterococcus faecium*, a new antilisterial bacteriocin in the pediocin family of bacteriocins. *Applied and Environmental Microbiology* 62:1676-1682
- Babasaki K, Takao T, Shimonishi Y, Kurahashi K (1985) Subtilosin A, a new antibiotic peptide produced by *Bacillus subtilis* 168: isolation, structural analysis, and biogenesis. *Journal of Biochemistry* 98:585-603
- Bakhtiary A, Cochrane SA, Mercier P, McKay RT, Miskolzie M, Sit CS, Vederas JC (2017) Insights into the Mechanism of Action of the Two-Peptide Lantibiotic Lacticin 3147. *Journal of the American Chemical Society*
- Bakkes PJ, Moll GN, Driessen AJ (2008) Distinct contributions of the nisin biosynthesis enzymes NisB and NisC and transporter NisT to prenisin production by *Lactococcus lactis*. *Applied and Environmental Microbiology* 74:5541-5548
- Bali V, Panesar PS, Bera MB, Kennedy JF (2016) Bacteriocins: recent trends and potential applications. *Critical Reviews in Food Science and Nutrition* 56:817-834
- Baquero F, Moreno F (1984) The microcins. *FEMS Microbiology Letters* 23:117-124
- Basi-Chipalu S, Dischinger J, Josten M, Szekat C, Zweynert A, Sahl H-G, Bierbaum G (2015) Pseudomycoicidin, a class II lantibiotic from *Bacillus pseudomycolides*. *Applied and Environmental Microbiology* 81:3419-3429
- Bastos MdCdF, Coutinho BG, Coelho MLV (2010) Lysostaphin: a staphylococcal bacteriolysin with potential clinical applications. *Pharmaceuticals* 3:1139-1161

- Bayro MJ, Mukhopadhyay J, Swapna G, Huang JY, Ma L-C, Sineva E, Dawson PE, Montelione GT, Ebright RH (2003) Structure of antibacterial peptide microcin J25: a 21-residue lariat protoknot. *Journal of the American Chemical Society* 125:12382-12383
- Bell MR, Engleka MJ, Malik A, Strickler JE (2013) To fuse or not to fuse: What is your purpose? *Protein Science* 22:1466-1477
- Benson DA, Clark K, Karsch-Mizrachi I, Lipman DJ, Ostell J, Sayers EW (2014) GenBank. *Nucleic Acids Research* 42:D32-D37
- Beukes M, Bierbaum G, Sahl H-G, Hastings J (2000) Purification and partial characterization of a murein hydrolase, millericin B, produced by *Streptococcus milleri* NMSCC 061. *Applied and Environmental Microbiology* 66:23-28
- Bierbaum G, Sahl HG (2009) Lantibiotics: mode of action, biosynthesis and bioengineering. *Current Pharmaceutical Biotechnology* 10:2-18
- Blin K, Medema MH, Kazempour D, Fischbach MA, Breitling R, Takano E, Weber T (2013) antiSMASH 2.0—a versatile platform for genome mining of secondary metabolite producers. *Nucleic Acids Research* 41:W204-W212
- Blin K, Kazempour D, Wohlleben W, Weber T (2014) Improved lanthipeptide detection and prediction for antiSMASH. *PLoS One* 9:e89420
- Boakes S, Cortés J, Appleyard AN, Rudd BA, Dawson MJ (2009) Organization of the genes encoding the biosynthesis of actagardine and engineering of a variant generation system. *Molecular Microbiology* 72:1126-1136
- Bogaardt C, Tonder AJ, Brueggemann AB (2015) Genomic analyses of pneumococci reveal a wide diversity of bacteriocins—including pneumocyclicin, a novel circular bacteriocin. *BMC Genomics* 16:554
- Bommarius B, Jenssen H, Elliott M, Kindrachuk J, Pasupuleti M, Gieren H, Jaeger K-E, Hancock R, Kalman D (2010) Cost-effective expression and purification of antimicrobial and host defense peptides in *Escherichia coli*. *Peptides* 31:1957-1965
- Borrero J, Brede DA, Skaugen M, Diep DB, Herranz C, Nes IF, Cintas LM, Hernández PE (2011) Characterization of garvicin ML, a novel circular bacteriocin produced by *Lactococcus garvieae* DCC43, isolated from mallard ducks (*Anas platyrhynchos*). *Applied and Environmental Microbiology* 77:369-373
- Bradford MM (1976) A rapid and sensitive method for the quantitation of microgram quantities of protein utilizing the principle of protein-dye binding. *Analytical Biochemistry* 72:248-254
- Brede DA, Faye T, Johnsborg O, Ødegård I, Nes IF, Holo H (2004) Molecular and genetic characterization of propionicin F, a bacteriocin from *Propionibacterium freudenreichii*. *Applied and Environmental Microbiology* 70:7303-7310
- Breukink E, Wiedemann I, Van Kraaij C, Kuipers O, Sahl H-G, De Kruijff B (1999) Use of the cell wall precursor lipid II by a pore-forming peptide antibiotic. *Science* 286:2361-2364
- Caetano T, Krawczyk JM, Mösker E, Süßmuth RD, Mendo S (2011) Heterologous expression, biosynthesis, and mutagenesis of type II lantibiotics from *Bacillus licheniformis* in *Escherichia coli*. *Chemistry & Biology* 18:90-100
- Caetano T, Krawczyk JM, Mösker E, Süßmuth RD, Mendo S (2011) Lichenicidin biosynthesis in *Escherichia coli*: *licFGEHI* immunity genes are not essential for lantibiotic production or self-protection. *Applied and Environmental Microbiology* 77:5023-5026
- Campelo AB, Roces C, Mohedano ML, López P, Rodríguez A, Martínez B (2014) A bacteriocin gene cluster able to enhance plasmid maintenance in *Lactococcus lactis*. *Microbial cell factories* 13:77
- Candiano G, Bruschi M, Musante L, Santucci L, Ghiggeri GM, Carnemolla B, Orecchia P, Zardi L, Righetti PG (2004) Blue silver: a very sensitive colloidal Coomassie G-250 staining for proteome analysis. *Electrophoresis* 25:1327-1333
- Cao P, Yu J, Lu W, Cai X, Wang Z, Gu Z, Zhang J, Ye T, Wang M (2010) Expression and purification of an antitumor-analgesic peptide from the venom of *Mesobuthus martensii*

- Karsch by small ubiquitin-related modifier fusion in *Escherichia coli*. *Biotechnology Progress* 26:1240-1244
- Casaus P, Nilsen T, Cintas LM, Nes IF, Hernández PE, Holo H (1997) Enterocin B, a new bacteriocin from *Enterococcus faecium* T136 which can act synergistically with enterocin A. *Microbiology* 143:2287-2294
- Cascales E, Buchanan SK, Duché D, Kleanthous C, Lloubes R, Postle K, Riley M, Slatin S, Cavard D (2007) Colicin biology. *Microbiology and Molecular Biology Reviews* 71:158-229
- Castiglione F, Lazzarini A, Carrano L, Corti E, Ciciliato I, Gastaldo L, Candiani P, Losi D, Marinelli F, Selva E (2008) Determining the structure and mode of action of microbisporicin, a potent lantibiotic active against multiresistant pathogens. *Chemistry & Biology* 15:22-31
- Cebrián R, Rodríguez-Ruano S, Martínez-Bueno M, Valdivia E, Maqueda M, Montalbán-López M (2014) Analysis of the promoters involved in enterocin AS-48 expression. *PloS one* 9:e90603
- Cebrián R, Martínez-Bueno M, Valdivia E, Albert A, Maqueda M, Sánchez-Barrena MJ (2015) The bacteriocin AS-48 requires dimer dissociation followed by hydrophobic interactions with the membrane for antibacterial activity. *Journal of structural biology* 190:162-172
- Cech NB, Enke CG (2001) Practical implications of some recent studies in electrospray ionization fundamentals. *Mass Spectrometry Reviews* 20:362-387
- Chalon MC, Bellomio A, Solbiati JO, Morero RD, Farias RN, Vincent PA (2009) Tyrosine 9 is the key amino acid in microcin J25 superoxide overproduction. *FEMS Microbiology Letters* 300:90-96
- Chan W, Dodd H, Horn N, Maclean K, Lian L, Bycroft B, Gasson M, Roberts G (1996) Structure-activity relationships in the peptide antibiotic nisin: role of dehydroalanine 5. *Applied and Environmental Microbiology* 62:2966-2969
- Chatterjee A, Xiao H, Schultz PG (2012) Evolution of multiple, mutually orthogonal prolyl-tRNA synthetase/tRNA pairs for unnatural amino acid mutagenesis in *Escherichia coli*. *Proceedings of the National Academy of Sciences* 109:14841-14846
- Chatterjee C, Paul M, Xie L, van der Donk WA (2005) Biosynthesis and mode of action of lantibiotics. *Chemical Reviews* 105:633-684
- Chen E, Chen Q, Chen S, Xu B, Ju J, Wang H (2017) Mathermycin, a lantibiotic from the marine Actinomycete *Marinactinospora thermotolerans* SCSIO 00652. *Applied and Environmental Microbiology* 83:e00926-00917
- Chen YQ, Zhang SQ, Li BC, Qiu W, Jiao B, Zhang J, Diao ZY (2008) Expression of a cytotoxic cationic antibacterial peptide in *Escherichia coli* using two fusion partners. *Protein Expression and Purification* 57:303-311
- Chikindas ML, García-Garcera MJ, Driessen A, Ledebøer AM, Nissen-Meyer J, Nes IF, Abee T, Konings WN, Venema G (1993) Pediocin PA-1, a bacteriocin from *Pediococcus acidilactici* PAC1. 0, forms hydrophilic pores in the cytoplasmic membrane of target cells. *Applied and Environmental Microbiology* 59:3577-3584
- Chikindas ML, Weeks R, Drider D, Chistyakov VA, Dicks LM (2018) Functions and emerging applications of bacteriocins. *Current Opinion in Biotechnology* 49:23-28
- Cintas LM, Casaus P, Håvarstein LS, Hernandez PE, Nes IF (1997) Biochemical and genetic characterization of enterocin P, a novel sec-dependent bacteriocin from *Enterococcus faecium* P13 with a broad antimicrobial spectrum. *Applied and Environmental Microbiology* 63:4321-4330
- Cintas LM, Casaus P, Holo H, Hernandez PE, Nes IF, Håvarstein LS (1998) Enterocins L50A and L50B, two novel bacteriocins from *Enterococcus faecium* L50, are related to staphylococcal hemolysins. *Journal of Bacteriology* 180:1988-1994
- Cintas LM, Casaus P, Herranz C, Håvarstein LS, Holo H, Hernández PE, Nes IF (2000) Biochemical and genetic evidence that *Enterococcus faecium* L50 produces enterocins L50A and L50B, thesec-dependent enterocin P, and a novel bacteriocin secreted without an N-terminal extension termed enterocin Q. *Journal of Bacteriology* 182:6806-6814
- Clarke IV TF, Clark PL (2008) Rare codons cluster. *PloS one* 3:e3412

- Consortium U (2016) UniProt: the universal protein knowledgebase. *Nucleic Acids Research* 45:D158-D169
- Cotter PD, Hill C, Ross RP (2005) Bacteriocins: developing innate immunity for food. *Nature Reviews Microbiology* 3:777-788
- Cotter PD, Ross RP, Hill C (2013) Bacteriocins—a viable alternative to antibiotics? *Nature Reviews: Microbiology* 11:95-105
- Cox CL, Doroghazi JR, Mitchell DA (2015) The genomic landscape of ribosomal peptides containing thiazole and oxazole heterocycles. *BMC Genomics* 16:778
- Cox J, Mann M (2008) MaxQuant enables high peptide identification rates, individualized ppb-range mass accuracies and proteome-wide protein quantification. *Nature Biotechnology* 26:1367-1372
- Criado R, Diep DB, Aakra Å, Gutiérrez J, Nes IF, Hernández PE, Cintas LM (2006) Complete sequence of the enterocin Q-encoding plasmid pCIZ2 from the multiple bacteriocin producer *Enterococcus faecium* L50 and genetic characterization of enterocin Q production and immunity. *Applied and Environmental Microbiology* 72:6653-6666
- Crost E, Ajandouz E, Villard C, Geraert P, Puigserver A, Fons M (2011) Ruminococcin C, a new anti-*Clostridium perfringens* bacteriocin produced in the gut by the commensal bacterium *Ruminococcus gnavus* E1. *Biochimie* 93:1487-1494
- Cruz VL, Ramos J, Melo MN, Martinez-Salazar J (2013) Bacteriocin AS-48 binding to model membranes and pore formation as revealed by coarse-grained simulations. *Biochimica et Biophysica Acta (BBA)-Biomembranes* 1828:2524-2531
- Cui Y, Zhang C, Wang Y, Shi J, Zhang L, Ding Z, Qu X, Cui H (2012) Class IIa bacteriocins: diversity and new developments. *International journal of molecular sciences* 13:16668-16707
- Dabard J, Bridonneau C, Phillipe C, Anglade P, Molle D, Nardi M, Ladire M, Girardin H, Marcille F, Gomez A (2001) Ruminococcin-A, a new lantibiotic produced by a *Ruminococcus gnavus* strain isolated from human feces. *Applied and Environmental Microbiology* 67:4111-4118
- Dass C (2007) Basics of Mass Spectrometry. *Fundamentals of Contemporary Mass Spectrometry*:1-14
- Davidson AL, Chen J (2004) ATP-binding cassette transporters in bacteria. *Annual Review of Biochemistry* 73:241-268
- De Kwaadsteniet M, Fraser T, Van Reenen C, Dicks L (2006) Bacteriocin T8, a novel class IIa sec-dependent bacteriocin produced by *Enterococcus faecium* T8, isolated from vaginal secretions of children infected with human immunodeficiency virus. *Applied and Environmental Microbiology* 72:4761-4766
- de Oliveira Junior AA, de Araújo Couto HGS, Barbosa AAT, Carnelossi MAG, de Moura TR (2015) Stability, antimicrobial activity, and effect of nisin on the physico-chemical properties of fruit juices. *International Journal of Food Microbiology* 211:38-43
- de Souza Duarte AF, Ceotto-Vigoder H, Barrias ES, Souto-Padrón TCBS, Nes IF, de Freire Bastos MdC (2017) Hyicin 4244, the first sactibiotic described in staphylococci, exhibits an anti-staphylococcal biofilm activity. *International Journal of Antimicrobial Agents*
- Diaz M, Valdivia E, Martínez-Bueno M, Fernández M, Soler-González AS, Ramírez-Rodrigo H, Maqueda M (2003) Characterization of a new operon, as-48EFGH, from the as-48 gene cluster involved in immunity to enterocin AS-48. *Applied and Environmental Microbiology* 69:1229-1236
- Diep DB, Håvarstein LS, Nes IF (1996) Characterization of the locus responsible for the bacteriocin production in *Lactobacillus plantarum* C11. *Journal of Bacteriology* 178:4472-4483
- Diep DB, Johnsborg O, Risøen PA, Nes IF (2001) Evidence for dual functionality of the operon *plnABCD* in the regulation of bacteriocin production in *Lactobacillus plantarum*. *Molecular Microbiology* 41:633-644

- Diep DB, Myhre R, Johnsborg O, Aakra Å, Nes IF (2003) Inducible bacteriocin production in *Lactobacillus* is regulated by differential expression of the *pln* operons and by two antagonizing response regulators, the activity of which is enhanced upon phosphorylation. *Molecular Microbiology* 47:483-494
- Diep DB, Skaugen M, Salehian Z, Holo H, Nes IF (2007) Common mechanisms of target cell recognition and immunity for class II bacteriocins. *Proceedings of the National Academy of Sciences* 104:2384-2389
- Dischinger J, Basi Chipalu S, Bierbaum G (2014) Lantibiotics: promising candidates for future applications in health care. *International Journal of Medical Microbiology* 304:51-62
- Douglas DJ, Frank AJ, Mao D (2005) Linear ion traps in mass spectrometry. *Mass Spectrometry Reviews* 24:1-29
- Drider D, Bendali F, Naghmouchi K, Chikindas ML (2016) Bacteriocins: not only antibacterial agents. *Probiotics and antimicrobial proteins* 8:177-182
- Duetz WA, Rüedi L, Hermann R, O'Connor K, Büchs J, Witholt B (2000) Methods for intense aeration, growth, storage, and replication of bacterial strains in microtiter plates. *Applied and Environmental Microbiology* 66:2641-2646
- Dunny GM, Leonard BA (1997) Cell-cell communication in gram-positive bacteria. *Annual Reviews in Microbiology* 51:527-564
- Ekblad B, Kyriakou PK, Opegård C, Nissen-Meyer J, Kaznessis YN, Kristiansen PE (2016) Structure-function analysis of the two-peptide bacteriocin plantaricin EF. *Biochemistry* 55:5106-5116
- Ekkelenkamp MB, Hanssen M, Hsu S-TD, de Jong A, Milatovic D, Verhoef J, van Nuland NA (2005) Isolation and structural characterization of epilancin 15X, a novel lantibiotic from a clinical strain of *Staphylococcus epidermidis*. *FEBS Letters* 579:1917-1922
- Engelke G, Gutowski-Eckel Z, Hammelmann M, Entian K (1992) Biosynthesis of the lantibiotic nisin: genomic organization and membrane localization of the NisB protein. *Applied and Environmental Microbiology* 58:3730-3743
- Ennahar S, Sonomoto K, Ishizaki A (1999) Class IIa bacteriocins from lactic acid bacteria: antibacterial activity and food preservation. *Journal of Bioscience and Bioengineering* 87:705-716
- Etayash H, Azmi S, Dangeti R, Kaur K (2016) Peptide bacteriocins-structure activity relationships. *Current Topics in Medicinal Chemistry* 16:220-241
- Faye T, Brede DA, Langsrud T, Nes IF, Holo H (2002) An antimicrobial peptide is produced by extracellular processing of a protein from *Propionibacterium jensenii*. *Journal of Bacteriology* 184:3649-3656
- Fenn JB, Mann M, Meng CK, Wong SF, Whitehouse CM (1989) Electrospray ionization for mass spectrometry of large biomolecules. *Science* 246:64-71
- Férir G, Petrova MI, Andrei G, Huskens D, Hoorelbeke B, Snoeck R, Vanderleyden J, Balzarini J, Bartoschek S, Brönstrup M (2013) The lantibiotic peptide labyrinthopeptin A1 demonstrates broad anti-HIV and anti-HSV activity with potential for microbicidal applications.
- Field D, Cotter PD, Hill C, Ross R (2015) Bioengineering lantibiotics for therapeutic success. *Frontiers in microbiology* 6
- Field D, Cotter PD, Ross RP, Hill C (2015) Bioengineering of the model lantibiotic nisin. *Bioengineered* 6 00-00
- Flühe L, Burghaus O, Wieckowski BM, Giessen TW, Linne U, Marahiel MA (2013) Two [4Fe-4S] clusters containing radical SAM enzyme SkfB catalyze thioether bond formation during the maturation of the sporulation killing factor. *Journal of the American Chemical Society* 135:959-962
- Franz CM, Van Belkum MJ, Holzapfel WH, Abriouel H, Gálvez A (2007) Diversity of enterococcal bacteriocins and their grouping in a new classification scheme. *FEMS Microbiology Reviews* 31:293-310
- Freund S, Jung G, Gutbrod O, Foikers G, Gibbons WA, Allgaier H, Werner R (1991) The solution structure of the lantibiotic gallidermin. *Biopolymers* 31:803-811

- Fuhrmann M, Hausherr A, Ferbitz L, Schödl T, Heitzer M, Hegemann P (2004) Monitoring dynamic expression of nuclear genes in *Chlamydomonas reinhardtii* by using a synthetic luciferase reporter gene. *Plant Molecular Biology* 55:869-881
- Fujita K, Ichimasa S, Zendo T, Koga S, Yoneyama F, Nakayama J, Sonomoto K (2007) Structural analysis and characterization of lacticin Q, a novel bacteriocin belonging to a new family of unmodified bacteriocins of gram-positive bacteria. *Applied and Environmental Microbiology* 73:2871-2877
- Funk MA, Van Der Donk WA (2017) Ribosomal natural products, tailored to fit. *Accounts of Chemical Research* 50:1577-1586
- Furgerson Ihnken L, Chatterjee C, Van Der Donk WA (2008) *In vitro* reconstitution and substrate specificity of a lantibiotic protease. *Biochemistry* 47:7352-7363
- Gabere MN, Noble WS (2017) Empirical comparison of web-based antimicrobial peptide prediction tools. *Bioinformatics*:btx081
- Gabrielsen C, Brede DA, Hernández PE, Nes IF, Diep DB (2012) The maltose ABC transporter in *Lactococcus lactis* facilitates high-level sensitivity to the circular bacteriocin garvicin ML. *Antimicrobial Agents and Chemotherapy* 56:2908-2915
- Gabrielsen C, Brede DA, Nes IF, Diep DB (2014) Circular bacteriocins: Biosynthesis and mode of action. *Applied and Environmental Microbiology* 80:6854-6862
- Gabrielsen C, Brede DA, Salehian Z, Nes IF, Diep DB (2014) Functional genetic analysis of the GarML gene cluster in *Lactococcus garvieae* DCC43 gives new insights into circular bacteriocin biosynthesis. *Journal of Bacteriology* 196:911-919
- Gajic O, Buist G, Kojic M, Topisirovic L, Kuipers OP, Kok J (2003) Novel mechanism of bacteriocin secretion and immunity carried out by lactococcal multidrug resistance proteins. *Journal of Biological Chemistry* 278:34291-34298
- Gálvez A, Maqueda M, Valdivia E, Quesada A, Montoya E (1986) Characterization and partial purification of a broad spectrum antibiotic AS-48 produced by *Streptococcus faecalis*. *Canadian Journal of Microbiology* 32:765-771
- Gálvez A, Maqueda M, Martínez-Bueno M, Valdivia E (1991) Permeation of bacterial cells, permeation of cytoplasmic and artificial membrane vesicles, and channel formation on lipid bilayers by peptide antibiotic AS-48. *Journal of Bacteriology* 173:886-892
- Gálvez A, Valdivia E, Abriouel H, Camafeita E, Mendez E, Martínez-Bueno M, Maqueda M (1998) Isolation and characterization of enterocin EJ97, a bacteriocin produced by *Enterococcus faecalis* EJ97. *Archives of Microbiology* 171:59-65
- Garcia De Gonzalo CV, Zhu L, Oman TJ, Van Der Donk WA (2014) NMR structure of the S-linked glycopeptide sublancin 168. *ACS Chemical Biology* 9:796-801
- Garg N, Tang W, Goto Y, Nair SK, van der Donk WA (2012) Lantibiotics from *Geobacillus thermodenitrificans*. *Proc Natl Acad Sci USA* 109:5241-5246
- Garg N, Salazar-Ocampo LMA, van der Donk WA (2013) *In vitro* activity of the nisin dehydratase NisB. *Proceedings of the National Academy of Sciences of the United States of America* 110:7258-7263
- Garg N, Goto Y, Chen T, van der Donk WA (2016) Characterization of the stereochemical configuration of lanthionines formed by the lanthipeptide synthetase GeoM. *Peptide Science* 106:834-842
- Gargis SR, Heath HE, Heath LS, LeBlanc PA, Simmonds RS, Abbott BD, Timkovich R, Sloan GL (2009) Use of 4-sulfophenyl isothiocyanate labeling and mass spectrometry to determine the site of action of the streptococcolytic peptidoglycan hydrolase zoocin A. *Applied and Environmental Microbiology* 75:72-77
- Gavrish E, Sit CS, Cao S, Kandror O, Spoering A, Peoples A, Ling L, Fetterman A, Hughes D, Bissell A (2014) Lassomycin, a ribosomally synthesized cyclic peptide, kills *Mycobacterium tuberculosis* by targeting the ATP-dependent protease ClpC1P1P2. *Chemistry & Biology* 21:509-518
- Geldart K, Kaznessis YN (2017) Characterization of class IIa bacteriocin resistance in *Enterococcus faecium*. *Antimicrobial Agents and Chemotherapy* 61:e02033-02016

- Gibbs GM, Davidson BE, Hillier AJ (2004) Novel expression system for large-scale production and purification of recombinant class IIa bacteriocins and its application to piscicollin 126. *Applied and Environmental Microbiology* 70:3292-3297
- Glish GL, Vachet RW (2003) The basics of mass spectrometry in the twenty-first century. *Nature Reviews Drug Discovery* 2:140
- Gomez A, Ladiré M, Marcille F, Fons M (2002) Trypsin mediates growth phase-dependent transcriptional regulation of genes involved in biosynthesis of ruminococcin A, a lantibiotic produced by a *Ruminococcus gnavus* strain from a human intestinal microbiota. *Journal of Bacteriology* 184:18-28
- Gomez A, Ladiré M, Marcille F, Nardi M, Fons M (2002) Characterization of ISRgn1, a novel insertion sequence of the IS3 family isolated from a bacteriocin-negative mutant of *Ruminococcus gnavus* E1. *Applied and Environmental Microbiology* 68:4136-4139
- Gong X, Martin-Visscher LA, Nahirney D, Vederas JC, Duszyk M (2009) The circular bacteriocin, carnocyclin A, forms anion-selective channels in lipid bilayers. *Biochimica et Biophysica Acta (BBA)-Biomembranes* 1788:1797-1803
- Goto Y, Li B, Claesen J, Shi Y, Bibb MJ, van der Donk WA (2010) Discovery of unique lanthionine synthetases reveals new mechanistic and evolutionary insights. *PLoS Biology* 8:e1000339
- Goto Y, Ökesli Ae, van der Donk WA (2011) Mechanistic studies of Ser/Thr dehydration catalyzed by a member of the LanL lanthionine synthetase family. *Biochemistry* 50:891-898
- Götz F, Perconti S, Popella P, Werner R, Schlag M (2014) Epidermin and gallidermin: staphylococcal lantibiotics. *International Journal of Medical Microbiology* 304:63-71
- Gut IM, Blanke SR, van der Donk WA (2011) Mechanism of inhibition of *Bacillus anthracis* spore outgrowth by the lantibiotic nisin. *ACS Chemical Biology* 6:744-752
- Hammami R, Zouhir A, Le Lay C, Hamida JB, Fliss I (2010) BACTIBASE second release: a database and tool platform for bacteriocin characterization. *BMC Microbiology* 10:22
- Hammi I, Delalande F, Belkhou R, Marchioni E, Cianferani S, Ennahar S (2016) Maltaricin CPN, a new class IIa bacteriocin produced by *Carnobacterium maltaromaticum* CPN isolated from mould-ripened cheese. *Journal of Applied Microbiology* 121:1268-1274
- Hasper HE, de Kruijff B, Breukink E (2004) Assembly and stability of nisin-lipid II pores. *Biochemistry* 43:11567-11575
- Hasper HE, Kramer NE, Smith JL, Hillman J, Zachariah C, Kuipers OP, De Kruijff B, Breukink E (2006) An alternative bactericidal mechanism of action for lantibiotic peptides that target lipid II. *Science* 313:1636-1637
- Hastings J, Sailer M, Johnson K, Roy K, Vederas J, Stiles M (1991) Characterization of leucocin A-UAL 187 and cloning of the bacteriocin gene from *Leuconostoc gelidum*. *Journal of Bacteriology* 173:7491-7500
- Haugen HS, Kristiansen PE, Fimland G, Nissen-Meyer J (2008) Mutational analysis of the class IIa bacteriocin curvacin A and its orientation in target cell membranes. *Appl Environ Microbiol* 74:6766-6773
- Havarstein LS, Diep DB, Nes IF (1995) A family of bacteriocin ABC transporters carry out proteolytic processing of their substrates concomitant with export. *Molecular Microbiology* 16:229-240
- Håvarstein LS, Holo H, Nes IF (1994) The leader peptide of colicin V shares consensus sequences with leader peptides that are common among peptide bacteriocins produced by gram-positive bacteria. *Microbiology* 140:2383-2389
- Heath LS, Heath H, Sloan GL (1987) Plasmid-encoded lysostaphin endopeptidase gene of *Staphylococcus simulans biovar staphylolyticus*. *FEMS Microbiology Letters* 44:129-133
- Hécharde Y, Dériard B, Letellier F, Cenatiempo Y (1992) Characterization and purification of mesentericin Y105, an anti-Listeria bacteriocin from *Leuconostoc mesenteroides*. *Journal of General Microbiology* 138:2725-2731
- Hegarty JW, Guinane CM, Ross RP, Hill C, Cotter PD (2016) Bacteriocin production: a relatively unharnessed probiotic trait? *F1000Research* 5

- Hegemann JD, Zimmermann M, Zhu S, Klug D, Marahiel MA (2013) Lasso peptides from proteobacteria: genome mining employing heterologous expression and mass spectrometry. *Peptide Science* 100:527-542
- Hegemann JD, Zimmermann M, Xie X, Marahiel MA (2015) Lasso peptides: an intriguing class of bacterial natural products. *Accounts of Chemical Research* 48:1909-1919
- Helfrich M, Entian K-D, Stein T (2007) Structure-function relationships of the lanthionine cyclase SpaC involved in biosynthesis of the *Bacillus subtilis* peptide antibiotic subtilin. *Biochemistry* 46:3224-3233
- Heng NC, Ragland NL, Swe PM, Baird HJ, Inglis MA, Tagg JR, Jack RW (2006) Dysgalacticin: a novel, plasmid-encoded antimicrobial protein (bacteriocin) produced by *Streptococcus dysgalactiae* subsp. *equisimilis*. *Microbiology* 152:1991-2001
- Hetrick KJ, van der Donk WA (2017) Ribosomally synthesized and post-translationally modified peptide natural product discovery in the genomic era. *Current Opinion in Chemical Biology* 38:36-44
- Himeno K, Rosengren KJ, Inoue T, Perez RH, Colgrave ML, Lee HS, Chan LY, Henriques SnT, Fujita K, Ishibashi N (2015) Identification, characterization, and three-dimensional structure of the novel circular bacteriocin, enterocin NKR-5-3B, from *Enterococcus faecium*. *Biochemistry* 54:4863-4876
- Holo H, Nilssen Ø, Nes I (1991) Lactococcin A, a new bacteriocin from *Lactococcus lactis* subsp. *cremoris*: isolation and characterization of the protein and its gene. *Journal of Bacteriology* 173:3879-3887
- Hsu S-TD, Breukink E, Tischenko E, Lutters MA, de Kruijff B, Kaptein R, Bonvin AM, van Nuland NA (2004) The nisin-lipid II complex reveals a pyrophosphate cage that provides a blueprint for novel antibiotics. *Nature Structural & Molecular Biology* 11:963-967
- Hu C-B, Malaphan W, Zendo T, Nakayama J, Sonomoto K (2010) Enterocin X, a novel two-peptide bacteriocin from *Enterococcus faecium* KU-B5, has an antibacterial spectrum entirely different from those of its component peptides. *Applied and Environmental Microbiology* 76:4542-4545
- Hu M, Zhao H, Zhang C, Yu J, Lu Z (2013) Purification and characterization of plantaricin 163, a novel bacteriocin produced by *Lactobacillus plantarum* 163 isolated from traditional Chinese fermented vegetables. *Journal of Agricultural and Food Chemistry* 61:11676-11682
- Hu Q, Noll RJ, Li H, Makarov A, Hardman M, Graham Cooks R (2005) The Orbitrap: a new mass spectrometer. *Journal of Mass Spectrometry* 40:430-443
- Huo L, van der Donk WA (2016) Discovery and characterization of bicereucin, an unusual D-amino acid-containing mixed two-component lantibiotic. *J Am Chem Soc* 138:5254-5257
- Iftime D, Jasyk M, Kulik A, Imhoff JF, Stegmann E, Wohlleben W, Süssmuth RD, Weber T (2015) Streptocollin, a type IV lanthipeptide produced by *Streptomyces collinus* Tü 365. *ChemBioChem* 16:2615-2623
- Igarashi Y, Kan Y, Fujii K, Fujita T, Harada K-I, Naoki H, Tabata H, Onaka H, Furumai T (2001) Goadsporin, a chemical substance which promotes secondary metabolism and morphogenesis in *Streptomyces*. *The Journal of Antibiotics* 54:1045-1053
- Ingham A, Sproat K, Tizard M, Moore R (2005) A versatile system for the expression of nonmodified bacteriocins in *Escherichia coli*. *Journal of Applied Microbiology* 98:676-683
- Iorio M, Sasso O, Maffioli SI, Bertorelli R, Monciardini P, Sosio M, Bonezzi F, Summa M, Brunati C, Bordoni R (2014) A glycosylated, labionin-containing lanthipeptide with marked antinociceptive activity. *ACS Chemical Biology* 9:398-404
- Ishii S, Yano T, Hayashi H (2006) Expression and characterization of the peptidase domain of *Streptococcus pneumoniae* ComA, a bifunctional ATP-binding cassette transporter involved in quorum sensing pathway. *Journal of Biological Chemistry* 281:4726-4731



- Ishii S, Yano T, Ebihara A, Okamoto A, Manzoku M, Hayashi H (2010) Crystal structure of the peptidase domain of *Streptococcus* ComA, a bifunctional ATP-binding cassette transporter involved in the quorum-sensing pathway. *Journal of Biological Chemistry* 285:10777-10785
- Islam MR, Nishie M, Nagao J-i, Zendo T, Keller S, Nakayama J, Kohda D, Sahl H-G, Sonomoto K (2012) Ring A of nukacin ISK-1: a lipid II-binding motif for type-A (II) lantibiotic. *Journal of the American Chemical Society* 134:3687-3690
- Iwatani S, Yoneyama F, Miyashita S, Zendo T, Nakayama J, Sonomoto K (2012) Identification of the genes involved in the secretion and self-immunity of lacticin Q, an unmodified leaderless bacteriocin from *Lactococcus lactis* QU 5. *Microbiology* 158:2927-2935
- Jamaluddin N, Stuckey DC, Ariff AB, Faizal Wong FW (2017) Novel approaches to purifying bacteriocin: a review. *Critical Reviews in Food Science and Nutrition*
- Joerger M, Klaenhammer T (1990) Cloning, expression, and nucleotide sequence of the *Lactobacillus helveticus* 481 gene encoding the bacteriocin helveticin J. *Journal of Bacteriology* 172:6339-6347
- Joosten H, Nunez M, Devreese B, Van Beeumen J, Marugg JD (1996) Purification and characterization of enterocin 4, a bacteriocin produced by *Enterococcus faecalis* INIA 4. *Applied and Environmental Microbiology* 62:4220-4223
- Jungmann NA, Krawczyk B, Tietzmann M, Ensle P, Süßmuth RD (2014) Dissecting reactions of nonlinear precursor peptide processing of the class III lanthipeptide curvopeptin. *Journal of the American Chemical Society* 136:15222-15228
- Källberg M, Wang H, Wang S, Peng J, Wang Z, Lu H, Xu J (2012) Template-based protein structure modeling using the RaptorX web server. *Nature protocols* 7:1511
- Kalmokoff M, Cyr T, Hefford M, Whitford M, Teather R (2003) Butyrivibriocin AR10, a new cyclic bacteriocin produced by the ruminal anaerobe *Butyrivibrio fibrisolvens* AR10: characterization of the gene and peptide. *Canadian Journal of Microbiology* 49:763-773
- Karas M, Bachmann D, Bahr U, Hillenkamp F (1987) Matrix-assisted ultraviolet laser desorption of non-volatile compounds. *International Journal of Mass Spectrometry and Ion Processes* 78:53-68
- Kaur S, Kaur S (2015) Bacteriocins as potential anticancer agents. *Frontiers in pharmacology* 6
- Kawai Y, Saito T, Suzuki M, Itoh T (1998) Sequence analysis by cloning of the structural gene of gasserin A, a hydrophobic bacteriocin produced by *Lactobacillus gasseri* LA39. *Bioscience, Biotechnology, and Biochemistry* 62:887-892
- Kawai Y, Ishii Y, Uemura K, Kitazawa H, Saito T, Itoh T (2001) *Lactobacillus reuteri* LA6 and *Lactobacillus gasseri* LA39 isolated from faeces of the same human infant produce identical cyclic bacteriocin. *Food Microbiology* 18:407-415
- Kawulka K, Sprules T, McKay RT, Mercier P, Diaper CM, Zuber P, Vederas JC (2003) Structure of Subtilisin A, an Antimicrobial Peptide from *Bacillus s ubtilis* with Unusual Posttranslational Modifications Linking Cysteine Sulfurs to  $\alpha$ -Carbons of Phenylalanine and Threonine. *J Am Chem Soc* 125:4726-4727
- Kearle P (2000) A brief overview of the present status of the mechanisms involved in electrospray mass spectrometry. *Journal of Mass Spectrometry* 35:804-817
- Kemperman R, Kuipers A, Karsens H, Nauta A, Kuipers O, Kok J (2003) Identification and characterization of two novel clostridial bacteriocins, circularin A and closticin 574. *Applied and Environmental Microbiology* 69:1589-1597
- Kersten RD, Yang Y-L, Xu Y, Cimerancic P, Nam S-J, Fenical W, Fischbach MA, Moore BS, Dorrestein PC (2011) A mass spectrometry-guided genome mining approach for natural product peptidogenomics. *Nature Chemical Biology* 7:794-802
- Khan H, Flint S, Yu PL (2013) Determination of the mode of action of enterolysin A, produced by *Enterococcus faecalis* B9510. *Journal of Applied Microbiology* 115:484-494
- Khusainov R, Heils R, Lubelski J, Moll GN, Kuipers OP (2011) Determining sites of interaction between prenisin and its modification enzymes NisB and NisC. *Molecular Microbiology* 82:706-718

- Kim S, Lee SB (2006) Rare codon clusters at 5'-end influence heterologous expression of archaeal gene in *Escherichia coli*. *Protein Expression and Purification* 50:49-57
- Kirst HA (2013) Developing new antibacterials through natural product research. *Expert Opinion on Drug Discovery* 8:479-493
- Kjos M (2011) Target recognition, resistance, immunity and genome mining of class II bacteriocins from Gram-positive bacteria. *Microbiology* 157:3256-3267
- Kjos M, Oppegård C, Diep DB, Nes IF, Veening J-W, Nissen-Meyer J, Kristensen T (2014) Sensitivity to the two-peptide bacteriocin lactococcin G is dependent on UppP, an enzyme involved in cell-wall synthesis. *Molecular Microbiology* 92:1177-1187
- Klaenhammer TR (1993) Genetics of bacteriocins produced by lactic acid bacteria. *FEMS Microbiology Reviews* 12:39-85
- Klein C, Kaletta C, Schnell N, Entian K (1992) Analysis of genes involved in biosynthesis of the lantibiotic subtilin. *Applied and Environmental Microbiology* 58:132-142
- Klocke M, Mundt K, Idler F, Jung S, Backhausen JE (2005) Heterologous expression of enterocin A, a bacteriocin from *Enterococcus faecium*, fused to a cellulose-binding domain in *Escherichia coli* results in a functional protein with inhibitory activity against *Listeria*. *Applied Microbiology and Biotechnology* 67:532-538
- Knappe TA, Linne U, Zirah S, Rebuffat S, Xie X, Marahiel MA (2008) Isolation and structural characterization of capistruin, a lasso peptide predicted from the genome sequence of *Burkholderia thailandensis* E264. *Journal of the American Chemical Society* 130:11446-11454
- Knerr PJ, van der Donk WA (2012) Discovery, biosynthesis, and engineering of lantipeptides. *Annual Review of Biochemistry* 81:479-505
- Kodani S, Hudson ME, Durrant MC, Buttner MJ, Nodwell JR, Willey JM (2004) The SapB morphogen is a lantibiotic-like peptide derived from the product of the developmental gene ramS in *Streptomyces coelicolor*. *Proceedings of the National Academy of Sciences of the United States of America* 101:11448-11453
- Kotake Y, Ishii S, Yano T, Katsuoka Y, Hayashi H (2008) Substrate Recognition Mechanism of the Peptidase Domain of the Quorum-Sensing-Signal-Producing ABC Transporter ComA from *Streptococcus*. *Biochemistry* 47:2531-2538
- Kraft M, Knüpfer U, Wenderoth R, Kacholdt A, Pietschmann P, Hock B, Horn U (2007) A dual expression platform to optimize the soluble production of heterologous proteins in the periplasm of *Escherichia coli*. *Applied Microbiology and Biotechnology* 76:1413-1422
- Krawczyk B, Ensle P, Müller WM, Süßmuth RD (2012) Deuterium labeled peptides give insights into the directionality of class III lantibiotic synthetase LabKC. *Journal of the American Chemical Society* 134:9922-9925
- Krawczyk B, Völler GH, Völler J, Ensle P, Süßmuth RD (2012) Curvopeptin: A new lanthionine-containing class III lantibiotic and its co-substrate promiscuous synthetase. *ChemBioChem* 13:2065-2071
- Krawczyk JM, Völler GH, Krawczyk B, Kretz J, Brönstrup M, Süßmuth RD (2013) Heterologous expression and engineering studies of labyrinthopeptins, class III lantibiotics from *Actinomadura namibiensis*. *Chemistry & Biology* 20:111-122
- Krebber A, Burmester J, Plückthun A (1996) Inclusion of an upstream transcriptional terminator in phage display vectors abolishes background expression of toxic fusions with coat protein g3p. *Gene* 178:71-74
- Kuipers A, de Boef E, Rink R, Fekken S, Kluskens LD, Driessen AJ, Leenhouts K, Kuipers OP, Moll GN (2004) NisT, the transporter of the lantibiotic nisin, can transport fully modified, dehydrated, and unmodified prenisin and fusions of the leader peptide with non-lantibiotic peptides. *Journal of Biological Chemistry* 279:22176-22182
- Kuipers A, Meijer-Wierenga J, Rink R, Kluskens LD, Moll GN (2008) Mechanistic dissection of the enzyme complexes involved in biosynthesis of lactacin 3147 and nisin. *Applied and Environmental Microbiology* 74:6591-6597

- Kuipers OP, Beerthuyzen MM, de Ruyter PG, Luesink EJ, de Vos WM (1995) Autoregulation of nisin biosynthesis in *Lactococcus lactis* by signal transduction. *Journal of Biological Chemistry* 270:27299-27304
- Kulikova T, Akhtar R, Aldebert P, Althorpe N, Andersson M, Baldwin A, Bates K, Bhattacharyya S, Bower L, Browne P (2006) EMBL nucleotide sequence database in 2006. *Nucleic Acids Research* 35:D16-D20
- Kurland C, Gallant J (1996) Errors of heterologous protein expression. *Current Opinion in Biotechnology* 7:489-493
- Kuthning A, Mösker E, Süssmuth RD (2015) Engineering the heterologous expression of lanthipeptides in *Escherichia coli* by multigene assembly. *Applied Microbiology and Biotechnology*:1-11
- Kuthning A, Durkin P, Oehm S, Hoesl MG, Budisa N, Süssmuth RD (2016) Towards biocontained cell factories: an evolutionarily adapted *Escherichia coli* strain produces a new-to-nature bioactive lantibiotic containing thienopyrrole-alanine. *Scientific reports* 6
- Kuznedelov K, Semenova E, Knappe TA, Mukhamedyarov D, Srivastava A, Chatterjee S, Ebright RH, Marahiel MA, Severinov K (2011) The antibacterial threaded-lasso peptide capistrin inhibits bacterial RNA polymerase. *Journal of Molecular Biology* 412:842-848
- Lagedroste M, Smits SH, Schmitt L (2017) Substrate Specificity of the Secreted Nisin Leader Peptidase NisP. *Biochemistry* 56:4005-4014
- Lai A-Y, Tran S, Simmonds R (2002) Functional characterization of domains found within a lytic enzyme produced by *Streptococcus equi subsp. zooepidemicus*. *FEMS Microbiology Letters* 215:133-138
- Lange O, Damoc E, Wiegand A, Makarov A (2014) Enhanced Fourier transform for Orbitrap mass spectrometry. *International Journal of Mass Spectrometry* 369:16-22
- Lee H, Churey JJ, Worobo RW (2009) Biosynthesis and transcriptional analysis of thurincin H, a tandem repeated bacteriocin genetic locus, produced by *Bacillus thuringiensis* SF361. *FEMS Microbiology Letters* 299:205-213
- Lee MV, Ihnken LAF, You YO, McClerren AL, Donk WAvd, Kelleher NL (2009) Distributive and directional behavior of lantibiotic synthetases revealed by high-resolution tandem mass spectrometry. *Journal of the American Chemical Society* 131:12258-12264
- Lee SW, Mitchell DA, Markley AL, Hensler ME, Gonzalez D, Wohlrab A, Dorrestein PC, Nizet V, Dixon JE (2008) Discovery of a widely distributed toxin biosynthetic gene cluster. *Proceedings of the National Academy of Sciences* 105:5879-5884
- Letzel A-C, Pidot SJ, Hertweck C (2014) Genome mining for ribosomally synthesized and post-translationally modified peptides (RiPPs) in anaerobic bacteria. *BMC Genomics* 15:983
- Li B, Yu JPJ, Brunzelle JS, Moll GN, Van der Donk WA, Nair SK (2006) Structure and mechanism of the lantibiotic cyclase involved in nisin biosynthesis. *Science* 311:1464-1467
- Li B, Cooper LE, van der Donk WA (2009) *In vitro* studies of lantibiotic biosynthesis. *Methods in Enzymology* 458:533-558
- Li B, Sher D, Kelly L, Shi Y, Huang K, Knerr PJ, Joewono I, Rusch D, Chisholm SW, van der Donk WA (2010) Catalytic promiscuity in the biosynthesis of cyclic peptide secondary metabolites in planktonic marine cyanobacteria. *Proc Natl Acad Sci* 107:10430-10435
- Li H, Xu H, Zhou Y, Zhang J, Long C, Li S, Chen S, Zhou J-M, Shao F (2007) The phosphothreonine lyase activity of a bacterial type III effector family. *Science* 315:1000-1003
- Li J (2013) Multi-scale optimization for heterologous biosynthesis of the nonribosomal peptide antibiotic valinomycin in *Escherichia coli*.
- Li J, Jaitzig J, Theuer L, Legala OE, Süssmuth RD, Neubauer P (2015) Type II thioesterase improves heterologous biosynthesis of valinomycin in *Escherichia coli*. *Journal of Biotechnology* 193:16-22
- Li JF, Zhang J, Song R, Zhang JX, Shen Y, Zhang SQ (2009) Production of a cytotoxic cationic antibacterial peptide in *Escherichia coli* using SUMO fusion partner. *Applied Microbiology and Biotechnology* 84:383-388

- Li JF, Zhang J, Zhang Z, Ma HW, Zhang JX, Zhang SQ (2010) Production of bioactive human beta-defensin-4 in *Escherichia coli* using SUMO fusion partner. The protein journal 29:314-319
- Li M, Yoneyama F, Toshimitsu N, Zendo T, Nakayama J, Sonomoto K (2013) Lethal hydroxyl radical accumulation by a lactococcal bacteriocin, lacticin Q. Antimicrobial Agents and Chemotherapy 57:3897-3902
- Li X, Leong SSJ (2011) A chromatography-focused bioprocess that eliminates soluble aggregation for bioactive production of a new antimicrobial peptide candidate. Journal of Chromatography A 1218:3654-3659
- Li Y (2011) Recombinant production of antimicrobial peptides in *Escherichia coli*: a review. Protein Expression and Purification 80:260-267
- Li Y, Ducasse R, Zirah S, Blond A, Goulard C, Lescop E, Giraud C, Hartke A, Guittet E, Pernodet J-L (2015) Characterization of svicucin from *Streptomyces* provides insights into enzyme exchangeability and disulfide bond formation in lasso peptides. ACS Chemical Biology
- Lin DY-w, Huang S, Chen J (2015) Crystal structures of a polypeptide processing and secretion transporter. Nature 523:425
- Liu N, Song L, Liu M, Shang F, Anderson Z, Fox DJ, Challis GL, Huang Y (2015) Unique post-translational oxime formation in the biosynthesis of the azolemycin complex of novel ribosomal peptides from *Streptomyces* sp. FXJ1. 264. Chemical Science 7:482-488
- Liu W, Hansen JN (1993) The antimicrobial effect of a structural variant of subtilin against outgrowing *Bacillus cereus* T spores and vegetative cells occurs by different mechanisms. Applied and Environmental Microbiology 59:648-651
- Liu X, Vederas JC, Whittall RM, Zheng J, Stiles ME, Carlson D, Franz CM, McMullen LM, van Belkum MJ (2011) Identification of an N-terminal formylated, two-peptide bacteriocin from *Enterococcus faecalis* 710C. Journal of Agricultural and Food Chemistry 59:5602-5608
- Lohans CT, Towle KM, Miskolzie M, McKay RT, van Belkum MJ, McMullen LM, Vederas JC (2013) Solution structures of the linear leaderless bacteriocins enterocin 7A and 7B resemble carnocyclin A, a circular antimicrobial peptide. Biochemistry 52:3987-3994
- Lohans CT, Li JL, Vederas JC (2014) Structure and biosynthesis of carnolysin, a homologue of enterococcal cytolyisin with D-amino acids. Journal of the American Chemical Society 136:13150-13153
- Lohans CT, Vederas JC (2014) Structural characterization of thioether-bridged bacteriocins. The Journal of Antibiotics 67:23-30
- Lubelski J, Rink R, Khusainov R, Moll G, Kuipers O (2008) Biosynthesis, immunity, regulation, mode of action and engineering of the model lantibiotic nisin. Cellular and Molecular Life Sciences 65:455-476
- Lubelski J, Khusainov R, Kuipers OP (2009) Directionality and coordination of dehydration and ring formation during biosynthesis of the lantibiotic nisin. Journal of Biological Chemistry 284:25962-25972
- Ma H, Gao Y, Zhao F, Wang J, Teng K, Zhang J, Zhong J (2014) Dissecting the catalytic and substrate binding activity of a class II lanthipeptide synthetase BovM. Biochemical and Biophysical Research Communications 450:1126-1132
- Makarov A (2000) Electrostatic axially harmonic orbital trapping: a high-performance technique of mass analysis. Analytical Chemistry 72:1156-1162
- Makarov A, Denisov E, Kholomeev A, Balschun W, Lange O, Strupat K, Horning S (2006) Performance evaluation of a hybrid linear ion trap/orbitrap mass spectrometer. Analytical Chemistry 78:2113-2120
- Makarov A, Denisov E, Lange O (2009) Performance evaluation of a high-field Orbitrap mass analyzer. Journal of the American Society for Mass Spectrometry 20:1391-1396
- Maksimov MO, Pan SJ, Link AJ (2012) Lasso peptides: structure, function, biosynthesis, and engineering. Natural Product Reports 29:996-1006

- Maky MA, Ishibashi N, Zendo T, Perez RH, Doud JR, Karmi M, Sonomoto K (2015) Enterocin F4-9, a novel O-linked glycosylated bacteriocin. *Applied and Environmental Microbiology* 81:4819-4826
- Maldonado-Barragán A, Caballero-Guerrero B, Jiménez E, Jiménez-Díaz R, Ruiz-Barba JL, Rodríguez JM (2009) Enterocin C, a class IIb bacteriocin produced by *E. faecalis* C901, a strain isolated from human colostrum. *International Journal of Food Microbiology* 133:105-112
- Maldonado-Barragán A, Cárdenas N, Martínez B, Ruiz-Barba JL, Fernández-Garayzábal JF, Rodríguez JM, Gibello A (2013) Garvicin A, a novel class IIc bacteriocin from *Lactococcus garvieae* that inhibits septum formation in *L. garvieae* strains. *Applied and Environmental Microbiology* 79:4336-4346
- Mant CT, Hodges RS (2017) High-performance liquid chromatography of peptides and proteins: separation, analysis, and conformation. CRC press
- Maqueda M, Sánchez-Hidalgo M, Fernández M, Montalbán-López M, Valdivia E, Martínez-Bueno M (2008) Genetic features of circular bacteriocins produced by Gram-positive bacteria. *FEMS Microbiology Reviews* 32:2-22
- March RE (2000) Quadrupole ion trap mass spectrometry: a view at the turn of the century. *International Journal of Mass Spectrometry* 200:285-312
- Marciset O, Jeronimus-Stratingh MC, Mollet B, Poolman B (1997) Thermophilin 13, a nontypical antilisterial poration complex bacteriocin, that functions without a receptor. *Journal of Biological Chemistry* 272:14277-14284
- Marshall AG, Verdun FR (2016) Fourier transforms in NMR, optical, and mass spectrometry: a user's handbook. Elsevier
- Martín-Platero AM, Valdivia E, Ruiz-Rodríguez M, Soler JJ, Martín-Vivaldi M, Maqueda M, Martínez-Bueno M (2006) Characterization of antimicrobial substances produced by *Enterococcus faecalis* MRR 10-3, isolated from the uropygial gland of the hoopoe (*Upupa epops*). *Applied and Environmental Microbiology* 72:4245-4249
- Martin-Visscher LA, van Belkum MJ, Garneau-Tsodikova S, Whittall RM, Zheng J, McMullen LM, Vederas JC (2008) Isolation and characterization of carnocyclin A, a novel circular bacteriocin produced by *Carnobacterium maltaromaticum* UAL307. *Applied and Environmental Microbiology* 74:4756-4763
- Martin-Visscher LA, Gong X, Duszyk M, Vederas JC (2009) The three-dimensional structure of carnocyclin A reveals that many circular bacteriocins share a common structural motif. *Journal of Biological Chemistry* 284:28674-28681
- Martínez-García E, Aparicio T, Goñi-Moreno A, Fraile S, de Lorenzo V (2014) SEVA 2.0: an update of the Standard European Vector Architecture for de-/re-construction of bacterial functionalities. *Nucleic Acids Research* 43:D1183-D1189
- Martínez-Bueno M, Valdivia E, Gálvez A, Coyette J, Maqueda M (1998) Analysis of the gene cluster involved in production and immunity of the peptide antibiotic AS-48 in *Enterococcus faecalis*. *Molecular Microbiology* 27:347-358
- Martínez B, Fernández Ma, Suárez JE, Rodríguez A (1999) Synthesis of lactococcin 972, a bacteriocin produced by *Lactococcus lactis* IPLA 972, depends on the expression of a plasmid-encoded bicistronic operon. *Microbiology* 145:3155-3161
- Martínez B, Rodríguez A, Suárez JE (2000) Lactococcin 972, a bacteriocin that inhibits septum formation in lactococci. *Microbiology* 146:949-955
- Masias E, Dupuy FG, da Silva Sanches PR, Farizano JV, Cilli E, Bellomio A, Saavedra L, Minahk C (2017) Impairment of the class IIa bacteriocin receptor function and membrane structural changes are associated to enterocin CRL35 high resistance in *Listeria monocytogenes*. *Biochimica et Biophysica Acta (BBA)-General Subjects*
- Masuda Y, Ono H, Kitagawa H, Ito H, Mu F, Sawa N, Zendo T, Sonomoto K (2011) Identification and characterization of leucocyclin Q, a novel cyclic bacteriocin produced by *Leuconostoc mesenteroides* TK41401. *Applied and Environmental Microbiology* 77:8164-8170
- Masuda Y, Zendo T, Sonomoto K (2012) New type non-lantibiotic bacteriocins: circular and leaderless bacteriocins. *Beneficial microbes* 3:3-12

- Mathavan I, Zirah S, Mehmood S, Choudhury HG, Goulard C, Li Y, Robinson CV, Rebuffat S, Beis K (2014) Structural basis for hijacking siderophore receptors by antimicrobial lasso peptides. *Nature Chemical Biology* 10:340-342
- Mathur H, C Rea M, D Cotter P, Hill C, Paul Ross R (2015) The sactibiotic subclass of bacteriocins: an update. *Current Protein and Peptide Science* 16:549-558
- Mavaro A, Abts A, Bakkes PJ, Moll GN, Driessen AJ, Smits SH, Schmitt L (2011) Substrate recognition and specificity of the NisB protein, the lantibiotic dehydratase involved in nisin biosynthesis. *Journal of Biological Chemistry* 286:30552-30560
- McClerren AL, Cooper LE, Quan C, Thomas PM, Kelleher NL, Van Der Donk WA (2006) Discovery and *in vitro* biosynthesis of haloduracin, a two-component lantibiotic. *Proceedings of the National Academy of Sciences* 103:17243-17248
- Meindl K, Schmiederer T, Schneider K, Reicke A, Butz D, Keller S, Gühring H, Vértesy L, Wink J, Hoffmann H (2010) Labyrinthopeptins: a new class of carbacyclic lantibiotics. *Angewandte Chemie International Edition* 49:1151-1154
- Melby JO, Nard NJ, Mitchell DA (2011) Thiazole/oxazole-modified microcins: complex natural products from ribosomal templates. *Current Opinion in Chemical Biology* 15:369-378
- Méndez C, Salas JA (2001) The role of ABC transporters in antibiotic-producing organisms: drug secretion and resistance mechanisms. *Research in Microbiology* 152:341-350
- Michalski A, Damoc E, Lange O, Denisov E, Nolting D, Müller M, Viner R, Schwartz J, Remes P, Belford M (2012) Ultra high resolution linear ion trap Orbitrap mass spectrometer (Orbitrap Elite) facilitates top down LC MS/MS and versatile peptide fragmentation modes. *Molecular & Cellular Proteomics* 11:O111. 013698
- Milioni C, Martínez B, Degl'Innocenti S, Turchi B, Fratini F, Cerri D, Fischetti R (2015) A novel bacteriocin produced by *Lactobacillus plantarum* LpU4 as a valuable candidate for biopreservation in artisanal raw milk cheese. *Dairy science & technology* 95:479-494
- Miller K, Ray P, Steinmetz T, Hanekamp T, Ray B (2005) Gene organization and sequences of pediocin AcH/PA-1 production operons in *Pediococcus* and *Lactobacillus* plasmids. *Letters in Applied Microbiology* 40:56-62
- Miller PE, Denton MB (1986) The quadrupole mass filter: basic operating concepts. *Journal of Chemical Education* 63:617
- Moll G, Ubbink-Kok T, Hildeng-Hauge H, Nissen-Meyer J, Nes IF, Konings WN, Driessen A (1996) Lactococcin G is a potassium ion-conducting, two-component bacteriocin. *Journal of Bacteriology* 178:600-605
- Montalbán-López M, Sanchez-Hidalgo M, Valdivia E, Martinez-Bueno M, Maqueda M (2011) Are bacteriocins underexploited? Novel applications for old antimicrobials. *Current Pharmaceutical Biotechnology* 12:1205-1220
- Montalbán-López M, Sánchez-Hidalgo M, Cebrián R, Maqueda M (2012) Discovering the bacterial circular proteins: bacteriocins, cyanobactins, and pilins. *Journal of Biological Chemistry* 287:27007-27013
- Montalbán-López M, van Heel AJ, Kuipers OP (2016) Employing the promiscuity of lantibiotic biosynthetic machineries to produce novel antimicrobials. *FEMS Microbiology Reviews*:fuw034
- Montalbán-López M, Deng J, van Heel AJ, Kuipers OP (2018) Specificity and Application of the Lantibiotic Protease NisP. *Frontiers in microbiology* 9:160
- Moreland JL, Gramada A, Buzko OV, Zhang Q, Bourne PE (2005) The Molecular Biology Toolkit (MBT): a modular platform for developing molecular visualization applications. *BMC Bioinformatics* 6:21
- Morris SL, Walsh RC, Hansen J (1984) Identification and characterization of some bacterial membrane sulfhydryl groups which are targets of bacteriostatic and antibiotic action. *Journal of Biological Chemistry* 259:13590-13594
- Mu F, Masuda Y, Zendo T, Ono H, Kitagawa H, Ito H, Nakayama J, Sonomoto K (2014) Biological function of a DUF95 superfamily protein involved in the biosynthesis of a

- circular bacteriocin, leucocyclicin Q. *Journal of Bioscience and Bioengineering* 117:158-164
- Müller E, Radler F (1993) Caseicin, a bacteriocin from *Lactobacillus casei*. *Folia Microbiologica* 38:441-446
- Müller WM, Schmiederer T, Ensle P, Süßmuth RD (2010) *In Vitro* Biosynthesis of the Prepeptide of Type III Lantibiotic Labyrinthopeptin A2 Including Formation of a C-C Bond as a Post-Translational Modification. *Angewandte Chemie International Edition* 49:2436-2440
- Murphy K, O'Sullivan O, Rea MC, Cotter PD, Ross RP, Hill C (2011) Genome mining for radical SAM protein determinants reveals multiple sactibiotic-like gene clusters. *PloS one* 6:e20852
- Nagao J-i, Harada Y, Shioya K, Aso Y, Zendo T, Nakayama J, Sonomoto K (2005) Lanthionine introduction into nukacin ISK-1 prepeptide by co-expression with modification enzyme NukM in *Escherichia coli*. *Biochemical and Biophysical Research Communications* 336:507-513
- Nagao J-i, Asaduzzaman SM, Aso Y, Okuda K-i, Nakayama J, Sonomoto K (2006) Lantibiotics: insight and foresight for new paradigm. *Journal of Bioscience and Bioengineering* 102:139-149
- Nagao J-i, Morinaga Y, Islam MR, Asaduzzaman SM, Aso Y, Nakayama J, Sonomoto K (2009) Mapping and identification of the region and secondary structure required for the maturation of the nukacin ISK-1 prepeptide. *Peptides* 30:1412-1420
- Nakamura Y, Gojobori T, Ikemura T (2000) Codon usage tabulated from international DNA sequence databases: status for the year 2000. *Nucleic Acids Research* 28:292-292
- Nes IF, Diep DB, Håvarstein LS, Brurberg MB, Eijsink V, Holo H (1996) Biosynthesis of bacteriocins in lactic acid bacteria. *Antonie Van Leeuwenhoek* 70:113-128
- Nes IF, Yoon S, Diep DB (2007) Ribosomally synthesized antimicrobial peptides (bacteriocins) in lactic acid bacteria: a review. *Food Science and Biotechnology* 16:675
- Netz DJA, Sahl H-G, Marcolino R, dos Santos Nascimento Jn, de Oliveira SS, Soares MB, de Freire Bastos MdC (2001) Molecular characterisation of aureocin A70, a multi-peptide bacteriocin isolated from *Staphylococcus aureus*. *Journal of Molecular Biology* 311:939-949
- Netz DJA, Pohl R, Beck-Sickinger AG, Selmer T, Pierik AJ, de Freire Bastos MdC, Sahl H-G (2002) Biochemical characterisation and genetic analysis of aureocin A53, a new, atypical bacteriocin from *Staphylococcus aureus*. *Journal of Molecular Biology* 319:745-756
- Neubauer P, Lin H, Mathiszik B (2003) Metabolic load of recombinant protein production: inhibition of cellular capacities for glucose uptake and respiration after induction of a heterologous gene in *Escherichia coli*. *Biotechnology and Bioengineering* 83:53-64
- Nilsen T, Nes IF, Holo H (2003) Enterolysin A, a cell wall-degrading bacteriocin from *Enterococcus faecalis* LMG 2333. *Applied and Environmental Microbiology* 69:2975-2984
- Nishie M, Shioya K, Nagao J-i, Jikuya H, Sonomoto K (2009) ATP-dependent leader peptide cleavage by NukT, a bifunctional ABC transporter, during lantibiotic biosynthesis. *Journal of Bioscience and Bioengineering* 108:460-464
- Nishie M, Sasaki M, Nagao J-i, Zendo T, Nakayama J, Sonomoto K (2011) Lantibiotic transporter requires cooperative functioning of the peptidase domain and the ATP binding domain. *Journal of Biological Chemistry* 286:11163-11169
- Nissen-Meyer J, Holo H, Håvarstein L, Sletten K, Nes I (1992) A novel lactococcal bacteriocin whose activity depends on the complementary action of two peptides. *Journal of Bacteriology* 174:5686-5692
- Nissen-Meyer J, Rogne P, Oppegård C, Haugen H, Kristiansen P (2009) Structure-function relationships of the non-lanthionine-containing peptide (class II) bacteriocins produced by gram-positive bacteria. *Current Pharmaceutical Biotechnology* 10:19-37
- Nissen-Meyer J, Oppegård C, Rogne P, Haugen HS, Kristiansen PE (2010) Structure and mode-of-action of the two-peptide (Class-IIb) bacteriocins. *Probiotics and Antimicrobial Proteins* 2:52-60

- Nizet V, Beall B, Bast DJ, Datta V, Kilburn L, Low DE, De Azavedo JC (2000) Genetic locus for streptolysin S production by group A streptococcus. *Infection and Immunity* 68:4245-4254
- Noll KS, Sinko PJ, Chikindas ML (2011) Elucidation of the molecular mechanisms of action of the natural antimicrobial peptide subtilisin against the bacterial vaginosis-associated pathogen *Gardnerella vaginalis*. *Probiotics Antimicrob Proteins* 3:41-47
- O'Shea EF, O'Connor PM, O'Sullivan O, Cotter PD, Ross RP, Hill C (2013) Bactofencin A, a new type of cationic bacteriocin with unusual immunity. *MBio* 4:e00498-00413
- Oldach F, Al Toma R, Kuthning A, Caetano T, Mendo S, Budisa N, Süssmuth RD (2012) Congeneric lantibiotics from ribosomal *in vivo* peptide synthesis with noncanonical amino acids. *Angewandte Chemie International Edition* 51:415-418
- Ongey EL, Neubauer P (2016) Lanthipeptides: chemical synthesis versus *in vivo* biosynthesis as tools for pharmaceutical production. *Microb Cell Fact* 15:97
- Ongey EL, Yassi H, Pflugmacher S, Neubauer P (2017) Pharmacological and pharmacokinetic properties of lanthipeptides undergoing clinical studies. *Biotechnology Letters*:1-10
- Ongey EL, Giessmann R, Fons M, Rappsilber J, Adrian L, Neubauer P (2018) Heterologous biosynthesis and structural characterization of ruminococcin-A, a lanthipeptide from the gut bacterium *Ruminococcus gnavus* E1, in *E. coli*. *Front Microbiol* 9:1688. doi: 10.3389/fmicb.2018.01688
- Ongey EL, Pflugmacher S, Neubauer P (2018) Bioinspired Designs, Molecular Premise and Tools for Evaluating Ecological Importance of Antimicrobial Peptides. *Pharmaceuticals*, 11(3), 68; <https://doi.org/10.3390/ph11030068>
- Oppegård C, Fimland G, Thorbæk L, Nissen-Meyer J (2007) Analysis of the two-peptide bacteriocins lactococcin G and enterocin 1071 by site-directed mutagenesis. *Applied and Environmental Microbiology* 73:2931-2938
- Oppegård C, Emanuelsen L, Thorbek L, Fimland G, Nissen-Meyer J (2010) The lactococcin G immunity protein recognizes specific regions in both peptides constituting the two-peptide bacteriocin lactococcin G. *Appl Environ Microbiol* 76:1267-1273
- Ortega MA, Velásquez JE, Garg N, Zhang Q, Joyce RE, Nair SK, van der Donk WA (2014) Substrate specificity of the lanthipeptide peptidase ElxP and the oxidoreductase ElxO. *ACS Chemical Biology* 9:1718-1725
- Ortega MA, Hao Y, Zhang Q, Walker MC, Van Der Donk WA, Nair SK (2015) Structure and mechanism of the tRNA-dependent lantibiotic dehydratase NisB. *Nature* 517:509-512
- Ovchinnikov KV, Kristiansen PE, Straume D, Jensen MS, Aleksandrak-Piekarczyk T, Nes IF, Diep DB (2017) The leaderless bacteriocin enterocin K1 is highly potent against *Enterococcus faecium*: a study on structure, target spectrum and receptor. *Frontiers in microbiology* 8
- Paik SH, Chakicherla A, Hansen JN (1998) Identification and characterization of the structural and transporter genes for, and the chemical and biological properties of, sublancin 168, a novel lantibiotic produced by *Bacillus subtilis* 168. *Journal of Biological Chemistry* 273:23134-23142
- Pane K, Durante L, Pizzo E, Varcamonti M, Zanfardino A, Sgambati V, Di Maro A, Carpentieri A, Izzo V, Di Donato A (2016) Rational design of a carrier protein for the production of recombinant toxic peptides in *Escherichia coli*. *PloS one* 11:e0146552
- Papp-Wallace K, Maguire M (2008) Magnesium Transport and Magnesium Homeostasis. *EcoSal Plus* 3
- Paul M, Patton GC, Van Der Donk WA (2007) Mutants of the zinc ligands of lactacin 481 synthetase retain dehydration activity but have impaired cyclization activity. *Biochemistry* 46:6268-6276
- Pekar Second T, Blethrow JD, Schwartz JC, Merrihew GE, MacCoss MJ, Swaney DL, Russell JD, Coon JJ, Zabrouskov V (2009) Dual-pressure linear ion trap mass spectrometer improving the analysis of complex protein mixtures. *Analytical Chemistry* 81:7757-7765



- Perez RH, Zendo T, Sonomoto K (2014) Novel bacteriocins from lactic acid bacteria (LAB): various structures and applications. *Microb Cell Fact* 13:S3
- Pitt JJ (2009) Principles and applications of liquid chromatography-mass spectrometry in clinical biochemistry. *The Clinical Biochemist Reviews* 30:19
- Quintana VM, Torres NI, Wachsman MB, Sinko PJ, Castilla V, Chikindas M (2014) Antiherpes simplex virus type 2 activity of the antimicrobial peptide subtilisin. *Journal of Applied Microbiology* 117:1253-1259
- Ramare F, Nicoli J, Dabard J, Corring T, Ladire M, Gueugneau A-M, Raibaud P (1993) Trypsin-dependent production of an antibacterial substance by a human *Peptostreptococcus strain* in gnotobiotic rats and *in vitro*. *Applied and Environmental Microbiology* 59:2876-2883
- Rea MC, Sit CS, Clayton E, O'Connor PM, Whittall RM, Zheng J, Vederas JC, Ross RP, Hill C (2010) Thuricin CD, a posttranslationally modified bacteriocin with a narrow spectrum of activity against *Clostridium difficile*. *Proceedings of the National Academy of Sciences* 107:9352-9357
- Repka LM, Chekan JR, Nair SK, van der Donk WA (2017) Mechanistic understanding of lanthipeptide biosynthetic enzymes. *Chem Rev* 117:5457-5520
- Rince A, Dufour A, Le Pogam S, Thuault D, Bourgeois C, Le Pennec J (1994) Cloning, expression, and nucleotide sequence of genes involved in production of lactococcin DR, a bacteriocin from *Lactococcus lactis subsp. lactis*. *Applied and Environmental Microbiology* 60:1652-1657
- Rink R, Arkema-Meter A, Baudoin I, Post E, Kuipers A, Nelemans S, Akanbi MHJ, Moll G (2010) To protect peptide pharmaceuticals against peptidases. *Journal of Pharmacological and Toxicological Methods* 61:210-218
- Rios AC, Moutinho CG, Pinto FC, Del Fiol FS, Jozala A, Chaud MV, Vila MMDC, Teixeira JA, Balcão VM (2016) Alternatives to overcoming bacterial resistances: state-of-the-art. *Microbiological Research* 191:51-80.
- Roe AJ, O'Byrne C, McLaggan D, Booth IR (2002) Inhibition of *Escherichia coli* growth by acetic acid: a problem with methionine biosynthesis and homocysteine toxicity. *Microbiology* 148:2215-2222
- Rogne P, Fimland G, Nissen-Meyer J, Kristiansen PE (2008) Three-dimensional structure of the two peptides that constitute the two-peptide bacteriocin lactococcin G. *Biochimica et Biophysica Acta (BBA)-Proteins and Proteomics* 1784:543-554
- Rogne P, Haugen C, Fimland G, Nissen-Meyer J, Kristiansen PE (2009) Three-dimensional structure of the two-peptide bacteriocin plantaricin JK. *Peptides* 30:1613-1621
- Ruhr E, Sahl H-G (1985) Mode of action of the peptide antibiotic nisin and influence on the membrane potential of whole cells and on cytoplasmic and artificial membrane vesicles. *Antimicrobial Agents and Chemotherapy* 27:841-845
- Ryoji M, Hsia K, Kaji A (1983) Read-through translation. *Trends in Biochemical Sciences* 8:88-90
- Salomón RA, Farías RN (1993) The FhuA protein is involved in microcin 25 uptake. *Journal of Bacteriology* 175:7741-7742
- Salomón RA, Farias RN (1995) The peptide antibiotic microcin 25 is imported through the TonB pathway and the SbmA protein. *Journal of Bacteriology* 177:3323-3325
- Sambeth GM, Süssmuth RD (2011) Synthetic studies toward labionin, a new  $\alpha$ ,  $\alpha$ -disubstituted amino acid from type III lantibiotic labyrinthopeptin A2. *Journal of Peptide Science* 17:581-584
- Sambrook J, Russell DW (2001) Molecular cloning: a laboratory manual 3rd edition. Coldspring-Harbour Laboratory Press, UK
- Sanchez-Hidalgo M (2011) AS-48 bacteriocin: close to perfection. *Cellular and Molecular Life Sciences* 68:2845-2857
- Sandiford S, Upton M (2012) Identification, characterization, and recombinant expression of epidermicin NI01, a novel unmodified bacteriocin produced by *Staphylococcus epidermidis* that displays potent activity against staphylococci. *Antimicrobial Agents and Chemotherapy* 56:1539-1547

- Sass P, Jansen A, Szekat C, Sass V, Sahl H-G, Bierbaum G (2008) The lantibiotic mersacidin is a strong inducer of the cell wall stress response of *Staphylococcus aureus*. *BMC Microbiology* 8:186
- Schägger H (2006) Tricine–sds-page. *Nature protocols* 1:16
- Schindler CA, Schuhardt V (1964) Lysostaphin: a new bacteriolytic agent for the *Staphylococcus*. *Proceedings of the National Academy of Sciences* 51:414-421
- Schnell N, Entian K-D, Schneider U, Götz F, Zähner H, Kellner R, Jung G (1988) Prepeptide sequence of epidermin, a ribosomally synthesized antibiotic with four sulphide-rings. *Nature Publishing Group* 333:276-278
- Schnell N, Engelke G, Augustin J, Rosenstein R, Ungermann V, Götz F, ENTIAN KD (1992) Analysis of genes involved in the biosynthesis of lantibiotic epidermin. *European Journal of Biochemistry* 204:57-68
- Scholz R, Molohon KJ, Nachtigall J, Vater J, Markley AL, Süßmuth RD, Mitchell DA, Borriss R (2011) Plantazolicin, a novel microcin B17/streptolysin S-like natural product from *Bacillus amyloliquefaciens* FZB42. *Journal of Bacteriology* 193:215-224
- Scholz R, Vater J, Budiharjo A, Wang Z, He Y, Dietel K, Schwecke T, Herfort S, Lasch P, Borriss R (2014) Amylocyclicin, a novel circular bacteriocin produced by *Bacillus amyloliquefaciens* FZB42. *Journal of Bacteriology* 196:1842-1852
- Scientific T (2016) Thermo Scientific Orbitrap Fusion Tribrid Mass Spectrometer. Thermo Fisher Scientific Inc
- Senko MW, Remes PM, Canterbury JD, Mathur R, Song Q, Eliuk SM, Mullen C, Earley L, Hardman M, Blethrow JD (2013) Novel parallelized quadrupole/linear ion trap/Orbitrap tribrid mass spectrometer improving proteome coverage and peptide identification rates. *Analytical Chemistry* 85:11710-11714
- Sezonov G, Joseleau-Petit D, D'Ari R (2007) *Escherichia coli* physiology in Luria-Bertani broth. *Journal of bacteriology* 189:8746-8749
- Shaffer SA, Tang K, Anderson GA, Prior DC, Udseth HR, Smith RD (1997) A novel ion funnel for focusing ions at elevated pressure using electrospray ionization mass spectrometry. *Rapid Communications in Mass Spectrometry* 11:1813-1817
- Sharp PM, Li W-H (1987) The codon adaptation index—a measure of directional synonymous codon usage bias, and its potential applications. *Nucleic Acids Research* 15:1281-1295
- Shevchenko A, Tomas H, Havli J, Olsen JV, Mann M (2006) In-gel digestion for mass spectrometric characterization of proteins and proteomes. *Nature protocols* 1:2856-2860
- Shi Y, Yang X, Garg N, van der Donk WA (2011) Production of lantipeptides in *Escherichia coli*. *Journal of the American Chemical Society* 133:2338-2341
- Shi Y, Bueno A, van der Donk WA (2012) Heterologous production of the lantibiotic Ala (0) actagardine in *Escherichia coli*. *Chemical Communications* 48:10966-10968
- Shimafuji C, Noguchi M, Nishie M, Nagao J-i, Shioya K, Zendo T, Nakayama J, Sonomoto K (2015) *In vitro* catalytic activity of N-terminal and C-terminal domains in NukM, the post-translational modification enzyme of nukacin ISK-1. *Journal of Bioscience and Bioengineering* 120:624-629
- Shin YC, Bischof GF, Lauer WA, Desrosiers RC (2015) Importance of codon usage for the temporal regulation of viral gene expression. *Proceedings of the National Academy of Sciences* 112:14030-14035
- Si L-G, Liu X-c, Lu Y-Y, Wang G-y, Li W-M (2007) Soluble expression of active human beta-defensin-3 in *Escherichia coli* and its effects on the growth of host cells. *Chinese medical journal* 120:708-713
- Simmonds RS, Simpson WJ, Tagg JR (1997) Cloning and sequence analysis of zooA, a *Streptococcus zooepidemicus* gene encoding a bacteriocin-like inhibitory substance having a domain structure similar to that of lysostaphin. *Gene* 189:255-261

- Singh PK, Sharma V, Patil PB, Korpole S (2012) Identification, purification and characterization of laterosporulin, a novel bacteriocin produced by *Brevibacillus* sp. strain GI-9. *PLoS One* 7:e31498
- Singh PK, Solanki V, Sharma S, Thakur KG, Krishnan B, Korpole S (2015) The intramolecular disulfide-stapled structure of laterosporulin, a class IId bacteriocin, conceals a human defensin-like structural module. *The FEBS journal* 282:203-214
- Sit CS, Lohans CT, van Belkum MJ, Campbell CD, Miskolzie M, Vederas JC (2012) Substitution of a Conserved Disulfide in the Type Ila Bacteriocin, Leucocin A, with L-Leucine and L-Serine Residues: Effects on Activity and Three-Dimensional Structure. *ChemBioChem* 13:35-38
- Šiurkus J, Panula-Perälä J, Horn U, Kraft M, Rimšeliene R, Neubauer P (2010) Novel approach of high cell density recombinant bioprocess development: Optimisation and scale-up from microlitre to pilot scales while maintaining the fed-batch cultivation mode of *E. coli* cultures. *Microbial Cell Factories* 9
- Skosyrev VS, Kuleskiy EA, Yakhnin AV, Temirov YV, Vinokurov LM (2003) Expression of the recombinant antibacterial peptide sarcotoxin IA in *Escherichia coli* cells. *Protein Expression and Purification* 28:350-356
- Stein T, Düsterhus S, Stroth A, Entian K-D (2004) Subtilisin production by two *Bacillus subtilis* subspecies and variance of the sbo-alb cluster. *Applied and Environmental Microbiology* 70:2349-2353
- Steiner T, Hess P, Bae JH, Wiltschi B, Moroder L, Budisa N (2008) Synthetic biology of proteins: tuning GFPs folding and stability with fluoroproline. *PLoS One* 3:e1680
- Stepper J, Shastri S, Loo TS, Preston JC, Novak P, Man P, Moore CH, Havlíček V, Patchett ML, Norris GE (2011) Cysteine S-glycosylation, a new post-translational modification found in glycopeptide bacteriocins. *FEBS Letters* 585:645-650
- Stratton T, Ding C, Henrich C, Grensemann H, Pfaff H, Strohm M, Kusch S (2013) Using HRAM survey analysis combined with rapid MS<sup>2</sup> data to develop a fragmentation based detection workflow for structure ID acquisition. Thermo Fisher Scientific Inc
- Strohm M, Kavan D, Novak P, Volny M, Havlicek V (2010) mMass 3: a cross-platform software environment for precise analysis of mass spectrometric data. *Analytical Chemistry* 82:4648-4651
- Sugai S, Ohnishi-Kameyama M, Kodani S (2017) Isolation and identification of a new lasso peptide cattlecin from *Streptomyces cattleya* based on genome mining. *Applied Biological Chemistry* 60:163-167
- Sutyak KE, Anderson RA, Dover SE, Feathergill KA, Aroutcheva AA, Faro S, Chikindas ML (2008) Spermicidal activity of the safe natural antimicrobial peptide subtilisin. *Infectious Diseases in Obstetrics and Gynecology* 2008
- Swe PM, Cook GM, Tagg JR, Jack RW (2009) Mode of action of dysgalactin: a large heat-labile bacteriocin. *Journal of Antimicrobial Chemotherapy* 63:679-686
- Tang W, van der Donk WA (2012) Structural characterization of four prochlorosins: a novel class of lantipeptides produced by planktonic marine cyanobacteria. *Biochemistry* 51:4271-4279
- Tay DKS, Rajagopalan G, Li X, Chen Y, Lua LH, Leong SSJ (2011) A new bioproduction route for a novel antimicrobial peptide. *Biotechnology and Bioengineering* 108:572-581
- Teng Y, Zhao W, Qian C, Li O, Zhu L, Wu X (2012) Gene cluster analysis for the biosynthesis of elgicins, novel lantibiotics produced by *Paenibacillus elgii* B69. *BMC Microbiology* 12:45
- Thibodeaux CJ, Ha T, Van Der Donk WA (2014) A price to pay for relaxed substrate specificity: a comparative kinetic analysis of the class II lanthipeptide synthetases ProcM and HalM2. *Journal of the American Chemical Society* 136:17513-17529
- Thibodeaux CJ, Wagoner J, Yu Y, van der Donk WA (2016) Leader Peptide Establishes Dehydration Order, Promotes Efficiency, and Ensures Fidelity During Lactacin 481 Biosynthesis. *Journal of the American Chemical Society* 138:6436-6444
- Thumm G, Götz F (1997) Studies on prollystaphin processing and characterization of the lysostaphin immunity factor (Lif) of *Staphylococcus simulans* biovar staphylolyticus. *Molecular Microbiology* 23:1251-1265

- Tichaczek PS, Nissen-Meyer J, Nes IF, Vogel RF, Hammes WP (1992) Characterization of the bacteriocins curvacin A from *Lactobacillus curvatus* LTH1174 and sakacin P from *L. sake* LTH673. *Systematic and Applied Microbiology* 15:460-468
- Tomita H, Fujimoto S, Tanimoto K, Ike Y (1996) Cloning and genetic organization of the bacteriocin 31 determinant encoded on the *Enterococcus faecalis* pheromone-responsive conjugative plasmid pYI17. *Journal of Bacteriology* 178:3585-3593
- Tosukhowong A, Zendo T, Visessanguan W, Roytrakul S, Pumpuang L, Jaresitthikunchai J, Sonomoto K (2012) Garvieacin Q, a novel class II bacteriocin from *Lactococcus garvieae* BCC 43578. *Applied and Environmental Microbiology* 78:1619-1623
- Tulini FL, Lohans CT, Bordon KC, Zheng J, Arantes EC, Vederas JC, De Martinis EC (2014) Purification and characterization of antimicrobial peptides from fish isolate *Carnobacterium maltaromaticum* C2: Carnobacteriocin X and carnolysins A1 and A2. *International Journal of Food Microbiology* 173:81-88
- Turner DL, Lamosa P, Rodríguez A, Martínez B (2013) Structure and properties of the metastable bacteriocin Lcn972 from *Lactococcus lactis*. *J Mol Struct* 1031:207-210
- Urban JH, Moosmeier MA, Aumüller T, Thein M, Bosma T, Rink R, Groth K, Zulley M, Siegers K, Tissot K (2017) Phage display and selection of lanthipeptides on the carboxy-terminus of the gene-3 minor coat protein. *Nature Communications* 8:1500
- Uzelac G, Kojic M, Lozo J, Aleksandrak-Piekarczyk T, Gabrielsen C, Kristensen T, Nes IF, Diep DB, Topisirovic L (2013) A Zn-dependent metallopeptidase is responsible for sensitivity to LsbB, a class II leaderless bacteriocin of *Lactococcus lactis* subsp. *lactis* BGMN1-5. *Journal of Bacteriology* 195:5614-5621
- van Belkum MJ, Kok J, Venema G (1992) Cloning, sequencing, and expression in *Escherichia coli* of lcnB, a third bacteriocin determinant from the lactococcal bacteriocin plasmid p9B4-6. *Applied and Environmental Microbiology* 58:572-577
- van Heel AJ, de Jong A, Montalban-Lopez M, Kok J, Kuipers OP (2013) BAGEL3: automated identification of genes encoding bacteriocins and (non-) bactericidal posttranslationally modified peptides. *Nucleic Acids Research* 41:W448-W453
- van Heel AJ, Kloosterman TG, Montalban-Lopez M, Deng J, Plat A, Baudu B, Hendriks D, Moll GN, Kuipers OP (2016) Discovery, Production and Modification of Five Novel Lantibiotics Using the Promiscuous Nisin Modification Machinery. *ACS synthetic biology* 5:1146-1154
- Völler GH, Krawczyk JM, Pesic A, Krawczyk B, Nachtigall J, Süßmuth RD (2012) Characterization of new class III lantibiotics—erythreapeptin, avermipeptin and griseopeptin from *Saccharopolyspora erythraea*, *Streptomyces avermitilis* and *Streptomyces griseus* demonstrates stepwise N-terminal leader processing. *ChemBioChem* 13:1174-1183
- Völler GH, Krawczyk B, Ensle P, Süßmuth RD (2013) Involvement and unusual substrate specificity of a prolyl oligopeptidase in class III lanthipeptide maturation. *Journal of the American Chemical Society* 135:7426-7429
- Vukomanović M, Žunič V, Kunej Š, Jančar B, Jeverica S, Suvorov D (2017) Nano-engineering the antimicrobial spectrum of lantibiotics: activity of nisin against gram negative bacteria. *Scientific reports* 7:4324
- Wang G, Feng G, Snyder AB, Manns DC, Churey JJ, Worobo RW (2014) Bactericidal thurincin H causes unique morphological changes in *Bacillus cereus* F4552 without affecting membrane permeability. *FEMS Microbiology Letters* 357:69-76
- Wang G, Li X, Wang Z (2016) APD3: the antimicrobial peptide database as a tool for research and education. *Nucleic Acids Research* 44:D1087-D1093
- Wang H, Van Der Donk WA (2012) Biosynthesis of the class III lanthipeptide catenulipeptin. *ACS Chemical Biology* 7:1529-1535
- Wang H, Oman TJ, Zhang R, Garcia De Gonzalo CV, Zhang Q, Van Der Donk WA (2013) The glycosyltransferase involved in thurandacin biosynthesis catalyzes both O- and S-glycosylation. *Journal of the American Chemical Society* 136:84-87

- Watson JT, Sparkman OD (2007) Introduction to mass spectrometry: instrumentation, applications, and strategies for data interpretation. John Wiley & Sons
- Wee S, Wilkinson BJ (1988) Increased outer membrane ornithine-containing lipid and lysozyme penetrability of *Paracoccus denitrificans* grown in a complex medium deficient in divalent cations. *Journal of bacteriology* 170:3283-3286
- Welch M, Govindarajan S, Ness JE, Villalobos A, Gurney A, Minshull J, Gustafsson C (2009) Design parameters to control synthetic gene expression in *Escherichia coli*. *PloS one* 4:e7002
- WHO (2014) Antimicrobial resistance: global report on surveillance 2014. In Geneva, Switzerland
- Wiedemann I, Böttiger T, Bonelli RR, Schneider T, Sahl H-G, Martínez B (2006) Lipid II-based antimicrobial activity of the lantibiotic plantaricin C. *Applied and Environmental Microbiology* 72:2809-2814
- Willey JM, van der Donk WA (2007) Lantibiotics: peptides of diverse structure and function. *Annual Review of Microbiology* 61:477-501
- Willey JM, Gaskell AA (2010) Morphogenetic signaling molecules of the streptomycetes. *Chemical Reviews* 111:174-187
- Wirawan RE, Swanson KM, Kleffmann T, Jack RW, Tagg JR (2007) Uberolysin: a novel cyclic bacteriocin produced by *Streptococcus uberis*. *Microbiology* 153:1619-1630
- Wladyka B, Wielebska K, Wloka M, Bochenska O, Dubin G, Dubin A, Mak P (2013) Isolation, biochemical characterization, and cloning of a bacteriocin from the poultry-associated *Staphylococcus aureus* strain CH-91. *Applied Microbiology and Biotechnology* 97:7229-7239
- Wu G, Culley DE, Zhang W (2005) Predicted highly expressed genes in the genomes of *Streptomyces coelicolor* and *Streptomyces avermitilis* and the implications for their metabolism. *Microbiology* 151:2175-2187
- Wu K-H, Tai PC (2004) Cys32 and His105 are the critical residues for the calcium-dependent cysteine proteolytic activity of CvaB, an ATP-binding cassette transporter. *Journal of Biological Chemistry* 279:901-909
- Xie L, Chatterjee C, Balsara R, Okeley NM, van der Donk WA (2002) Heterologous expression and purification of SpaB involved in subtilin biosynthesis. *Biochemical and Biophysical Research Communications* 295:952-957
- Xie L, Miller LM, Chatterjee C, Averin O, Kelleher NL, Van Der Donk WA (2004) Lactacin 481: *in vitro* reconstitution of lantibiotic synthetase activity. *Science* 303:679-681
- Xu Y, Li X, Li R, Li S, Ni H, Wang H, Xu H, Zhou W, Saris PE, Yang W (2014) Structure of the nisin leader peptidase NisP revealing a C-terminal autocleavage activity. *Acta Crystallographica Section D: Biological Crystallography* 70:1499-1505
- Yan KP, Li Y, Zirah S, Goulard C, Knappe TA, Marahiel MA, Rebuffat S (2012) Dissecting the maturation steps of the lasso peptide microcin J25 *in vitro*. *ChemBioChem* 13:1046-1052
- Yang X, van der Donk WA (2013) Ribosomally synthesized and post-translationally modified peptide natural products: New insights into the role of leader and core peptides during biosynthesis. *Chemistry-A European Journal* 19:7662-7677
- Yang X, van der Donk WA (2015) Michael-type cyclizations in lantibiotic biosynthesis are reversible. *ACS Chemical Biology* 10:1234-1238
- Yang X, van der Donk WA (2015) Post-translational introduction of D-Alanine into ribosomally synthesized peptides by the dehydroalanine reductase NpnJ. *Journal of the American Chemical Society* 137:12426-12429
- Yoneyama F, Imura Y, Ohno K, Zendo T, Nakayama J, Matsuzaki K, Sonomoto K (2009) Peptide-lipid huge toroidal pore, a new antimicrobial mechanism mediated by a lactococcal bacteriocin, lactacin Q. *Antimicrobial Agents and Chemotherapy* 53:3211-3217
- You YO, Van Der Donk WA (2007) Mechanistic investigations of the dehydration reaction of lactacin 481 synthetase using site-directed mutagenesis. *Biochemistry* 46:5991-6000
- Yu G, Baeder D, Regoes R, Rolff J (2017) Predicting drug resistance evolution: antimicrobial peptides vs. antibiotics. *bioRxiv*:138107

- Zambaldo C, Luo X, Mehta AP, Schultz PG (2017) Recombinant macrocyclic lanthipeptides incorporating non-canonical amino acids. *Journal of the American Chemical Society* 139:11646-11649
- Zendo T, Koga S, Shigeri Y, Nakayama J, Sonomoto K (2006) Lactococcin Q, a novel two-peptide bacteriocin produced by *Lactococcus lactis* QU 4. *Applied and Environmental Microbiology* 72:3383-3389
- Zhang L-j, Gallo RL (2016) Antimicrobial peptides. *Current Biology* 26:R14-R19
- Zhang Q, Yu Y, Vélasquez JE, Van Der Donk WA (2012) Evolution of lanthipeptide synthetases. *Proceedings of the National Academy of Sciences* 109:18361-18366
- Zhang Q, Yang X, Wang H, van der Donk WA (2014) High divergence of the precursor peptides in combinatorial lanthipeptide biosynthesis. *ACS Chemical Biology* 9:2686-2694
- Zhao X, van der Donk WA (2016) Structural characterization and bioactivity analysis of the two-component lantibiotic Flv system from a ruminant bacterium. *Cell chemical biology* 23:246-256
- Zheng G, Yan LZ, Vederas JC, Zuber P (1999) Genes of the *sbo-alb* locus of *Bacillus subtilis* are required for production of the antilisterial bacteriocin subtilisin. *Journal of Bacteriology* 181:7346-7355
- Zhou L, van Heel AJ, Montalban-Lopez M, Kuipers OP (2016) Potentiating the activity of nisin against *Escherichia coli*. *Frontiers in cell and developmental biology* 4
- Zouhir A, Hammami R, Fliss I, Hamida JB (2010) A new structure-based classification of Gram-positive bacteriocins. *The protein journal* 29:432-439
- Zwick ME, Joseph SJ, Didelot X, Chen PE, Bishop-Lilly KA, Stewart AC, Willner K, Nolan N, Lentz S, Thomason MK (2012) Genomic characterization of the *Bacillus cereus sensu lato* species: backdrop to the evolution of *Bacillus anthracis*. *Genome Research* 22:1512-1524

## Acknowledgements

A lot of wonderful people contributed to the achievements that I have made during the last four years leading to this dissertation, and I would like to seize this opportunity to extend my heartfelt appreciations.

Firstly, I am deeply thankful to my supervisor Prof. Dr. Peter Neubauer and honored to have made the decision to stay back with him in Berlin to complete my PhD thesis. He has acted from day one like a father to me, giving me encouragements in all aspects of life. His enthusiastic support and the zeal to push towards success ignited a new spirit in me which served as a driving force throughout this work. He introduced me into this interesting research field and together, we crafted the research project. Let me also extend deepest appreciations to Prof. Dr. Roderich Süßmuth, Prof. Dr. Juri Rappsilber and Prof. Dr. Lorenz Adrian for their collaboration and for providing their mass spectrometry facilities to measure my samples. Most especially, to Prof. Dr. Lorenz Adrian who provided immense support in data analysis.

I would also like to appreciate the inputs of Robert Giessmann worked together with me for a short while during the project, providing insightful thoughts on how we could analyze the activity and/or masses of resulting products. Additionally, I want to extend my deepest gratitude to my Master's student Rosine Mezatio for her resourceful contributions towards this project. I am also grateful to Dr. Nicolas Cruz-Bournazou, Dr. Stefan Junne, Dr. Matthias Gimpel, Dr. Anke Wagner, Dr. Florian Glauche, Dr. Christian Reiltz for their assistance in experimental design and planning during this project.

I am most grateful to the Berlin International Graduate School of Natural Sciences and Engineering (BIG-NSE) and to the UniCat "Cluster of Excellence" for awarding me a position, as well as the German Academic Exchange Service (DAAD) for supplying the funding that carried me through this project. I want extend great thanks to Dr. Jean-Philippe Lonjaret for all his support at the BIG-NSE and for organizing the much needed social and recreational activities. I also want to thank my BIG-NSE batch mates (batch WS2014) for their collaborative support, not also forgetting BIG-NSE members from other batches who were very helpful.

Let me also seize this opportunity to extend my gratitude to the current and former members of Prof. Neubauer's group (BVT). Special thanks to Irmgard Maue-Mohn, Brigitte Burckhardt and Thomas Högl for their technical support. Special thanks to Mrs. Sabine Lühr-Müller for helping me with all the administrative paper works. Heba Yehia and Funda Cansu Ertem for their interesting jokes and discussions, Prof. Dr. Mario Birkholz for interesting global reviews. I also want to thank Maryke Fehlau, Erich Kielhorn, Anika Bockisch, Klaus Pellicer, Dr. Anna

Maria Marba, Sarah Kamel, Emmanuel Anane, Andri Hutari, Dr. Juan Antonio Arzate, Dr. Howard Ramirez-Malule and all other colleagues who constituted a great working environment.

Finally, I want to thank my family, especially my wife Faith N. Ongey, who have braced all the odds to stand by me all step of the way. Their understanding, unconditional love and support encouraged me to forge ahead even in difficult times. I will forever remain indebted to you!





## Short Resume

### Publications

#### Peer-reviewed articles

**Ongey, E. L.;** Giessmann, R.; Fons, M.; Rappsilber, J.; Adrian, L. and Neubauer, P. (2018) Heterologous biosynthesis, modifications and structural characterization of ruminococcin-A, a lanthipeptide from a gut bacterium, in *E. coli*. *Front Microbiol* 9:1688.

<http://doi.org/10.3389/fmicb.2018.01688>

**Ongey E. L.,** Pflugmacher S., Neubauer P. (2018) Bioinspired designs, molecular premise and tools for evaluating ecological importance of antimicrobial peptides. *Pharmaceuticals*, 11(3), 68; <https://doi.org/10.3390/ph11030068>

**Ongey, E. L.;** Yassi, H.; Pflugmacher, S.; Neubauer, P. (2017) Pharmacological and pharmacokinetic properties of lanthipeptides undergoing clinical studies. *Biotechnol Lett* 39: 473. <https://doi.org/10.1007/s10529-016-2279-9>

**Ongey, E. L.;** Neubauer, P. (2016) Lanthipeptides: chemical synthesis versus *in vivo* biosynthesis as tools for pharmaceutical production. *Microb Cell Fact* 15: 97.

<https://doi.org/10.1186/s12934-016-0502-y>

Li, J.; Jaitzig, J.; Theuer, L.; **Ongey, E. L.;** Süssmuth, R. D. and Neubauer, P. (2015) Type II thioesterase improves heterologous biosynthesis of valinomycin in *Escherichia coli*. *J. Biotechnol* 193: 16-22. <http://doi.org/10.1016/j.jbiotec.2014.10.037>

Narwal, Mohit, Jarkko Koivunen, Teemu Haikarainen, Ezeogo Obaji, **Ongey E. Legala**, et al. (2013) Discovery of tankyrase inhibiting flavones with increased potency and isoenzyme selectivity. *J Med Chem* 56 (20): 7880-7889. <http://doi.org/10.1021/jm401463y>

#### Conference contributions

##### Oral presentation

Ongey E. L., Giessmann R., Fons M., Rappsilber J., Süssmuth R. D. and Neubauer. Production of Ruminococcin-A, a lanthipeptide isolated from a strict anaerobic Gram-positive bacterium, in *E. coli*. International Conference on Recombinant Protein Production (RPP9) 2017, April 23-25, Dubrovnik, Croatia. Web: <https://fems-microbiology.org/opportunities/recombinant-protein-production-9-comparative-view-host-physiology/>

## Poster presentation

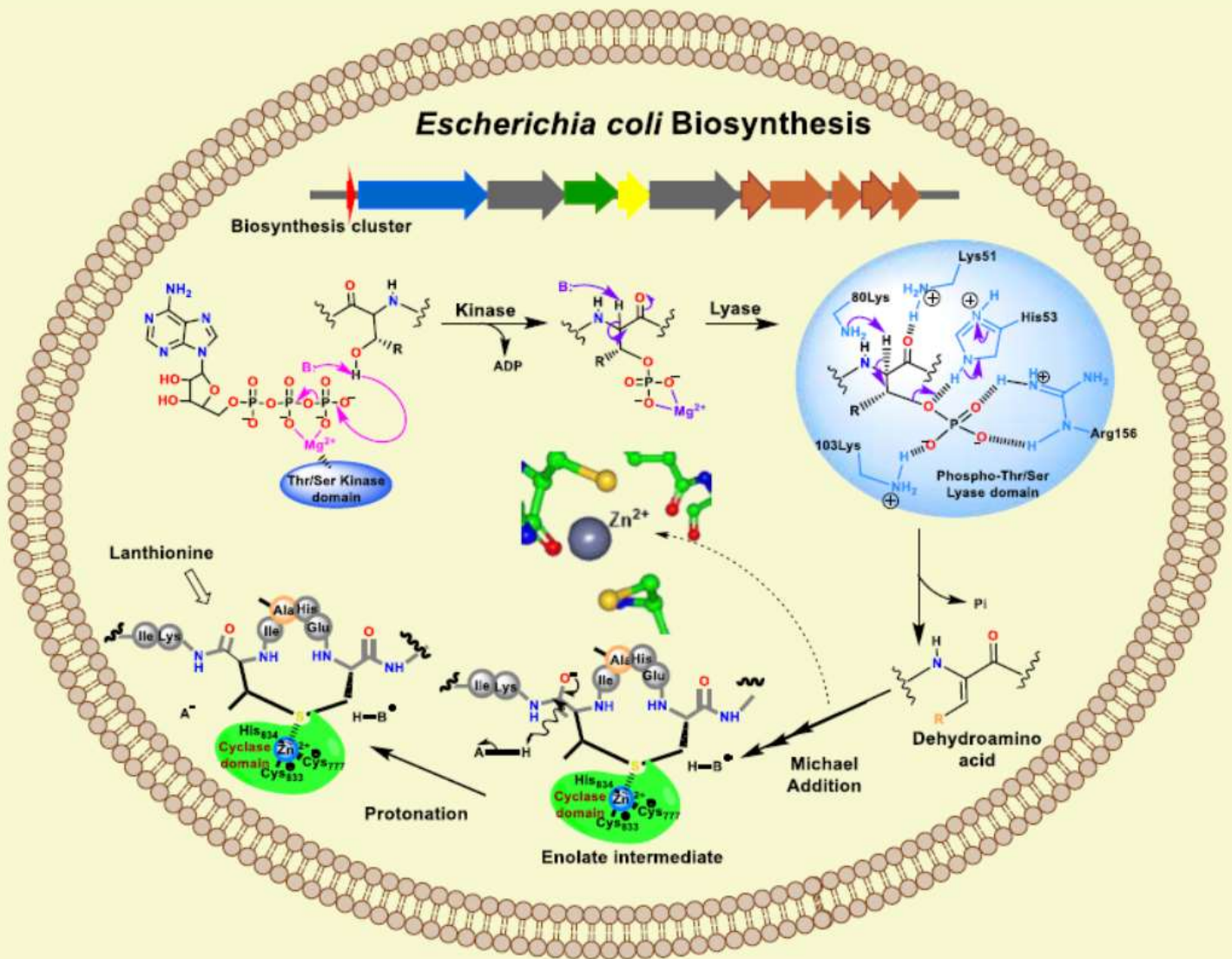
Ongey E. L., Giessmann R., Fons M., Rappsilber J., Süssmuth R. D. and Neubauer. Biosynthesis and modification of Ruminococcin-A, in *E. coli*. Systems Biology meets Synthetic Biology workshop 2017, May 2-3, Germany

web: [https://dechema.converia.de/frontend/index.php?page\\_id=4201](https://dechema.converia.de/frontend/index.php?page_id=4201)

Ongey E. L., Giessmann R., Fons M., Rappsilber J., Süssmuth R. D. and Neubauer. Production of Ruminococcin-A, a lanthipeptide isolated from a strict anaerobic Gram-positive bacterium, in *E. coli*. International Conference on Recombinant Protein Production (RPP9) 2017, April 23-25, Dubrovnik, Croatia web: <https://fems-microbiology.org/opportunities/recombinant-protein-production-9-comparative-view-host-physiology/>

Ongey E. L., Fons M. and Neubauer P. Reconstructing the biosynthetic machinery of Ruminococcin A, a lanthipeptide identified from a strict anaerobic Gram-positive bacterium, in *E. coli*. 8th International Congress on Biocatalysis 2016, August 28-September 1, Hamburg, Germany web: <http://www.biocascades.eu/event/8th-international-congress-on-biocatalysis/>

Ongey E. L., Fons M. and Neubauer P. Systematic approach towards biosynthesis of polycyclic microbial peptides in *E. coli*. Biotechnology Conference: 4th BioProScale Symposium 2016, April 6-8, Berlin, Germany web: <http://biotechnologie.ifgb.de/seminars-conventions/bioproscale-conference-2016.html>



Printed with the support of the German Academic Exchange Service