Investigating Structural Properties of Antibacterial Peptides

by

Sorina Chiorean

A thesis submitted in partial fulfillment of the requirements for the degree of

Doctor of Philosophy

Department of Chemistry
University of Alberta

Abstract

The race towards uncovering and developing novel antibiotics requires a multipronged, interdisciplinary approach. Highlighted in this thesis are biochemical and structural studies on peptides, small naturally occurring molecules that have innate bioactivities, such as antimicrobial characteristics. These include: faerocin MK, a potent anti-listerial molecule; teixobactin, a peptide with the ability to inhibit Gram-positive bacteria; tridecaptin A_1, an inhibitor of Gram-negative bacteria; and microcin J25, a lasso peptide with activity against Gram-negative foodborne pathogens.

Faerocin MK was uncovered after the genome sequencing of a newly discovered probiotic bacterium, *Enterococcus faecium* M3K31. The sequence encoding for faerocin MK contained two components: a structural gene and an immunity protein. Heterologous expression studies led to the production, isolation, and characterization of this new peptide. These studies showed that this peptide is active against Gram-positive organisms, with potent activity against *Listeria* species, a genus of foodborne pathogens responsible for listeriosis.

Teixobactin is a peptide isolated in 2015 and has gained much attention due to its potent activity against Gram-positive bacteria and unique structural features. With the use of analogues, the binding of teixobactin and its bacterial target, lipid II, a peptidoglycan precursor present within bacterial membranes, was investigated. Furthermore, the activity of teixobactin was expanded to reach Gram-negative organisms through synergistic assays with membrane-disrupting peptides.

Tridecaptin A_1 is part of a class of peptides which are potent against Gram-negative bacteria. They are known to bind to lipid II, the same peptidoglycan precursor targeted by teixobactin and a number of other antimicrobial peptides. In order to better understand the tridecaptin A_1 mechanism of action, in particular its interaction with lipid II at the bacterial membrane surface, nuclear magnetic resonance experiments were employed. The structures of tridecaptin A_1 and lipid II were studied in a new series of membrane-mimicking micelles. These micelle-forming dodecylphosphocholine compounds were synthesized, with heteroatom replacements at strategic positions, and structural data were compared amongst analogues.

Microcin J25 is a lasso peptide from *Escherichia coli* active against Gram-negative, foodborne pathogens such as *Salmonella* species and *E. coli.* The unique structure of this class of peptides, wherein the C-terminus is threaded through an N-terminal macrocycle, confers on them desirable attributes such as thermal stability and protease resistance, along with potent innate activities against bacteria and viruses. The lasso shape has long been regarded as a valuable scaffold for pharmaceutical development, yet these peptides have thus far been accessed through biological means as a facile synthetic approach has not yet been developed. A method to synthesize this interlocked, stable structure was initiated, providing a potential means of accessing microcin J25 and other lasso peptides in the future.

Preface

Parts of this thesis have previously been published while some work is still ongoing and therefore unpublished research. Sections of Chapter 2 have been published as: Chiorean S, Vederas JC, and van Belkum MJ. Identification and heterologous expression of the *sec*-dependent bacteriocin faerocin MK from *Enterococcus faecium M3K31*. *Probiotics Antimicrob Proteins*, **2018**, *10*, 142-147. The work was mostly completed by myself (75%) and Dr. van Belkum (25%), who assisted in the heterologous gene cloning and operon annotation.

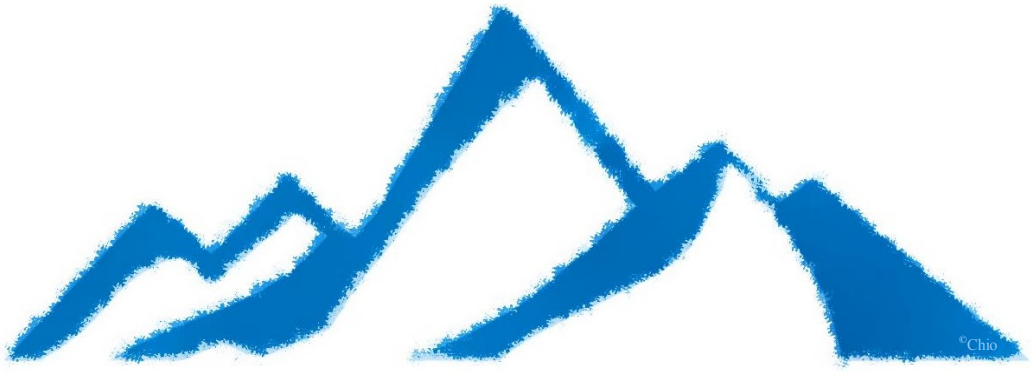
The majority of Chapter 3 has been published as Chiorean S, Antwi I, Carney DW, Kotsogianni I, Giltrap AM, Alexander FM, Cochrane SA, Payne RJ, Martin NI, Henninot A, and Vederas JC. Dissecting the binding interactions of teixobactin with the bacterial cell wall precursor lipid II. *ChemBioChem*, **2020**, *21*, 789-792. The research was performed by myself (75%) with synthetic contributions from Drs. Carney and Henninot (10%) for teixobactin analogues along with Mr. Antwi (10%) and Dr. Cochrane (5%) who both aided in the synthesis of lipid II and analogues.

At the time of this thesis preparation, Chapter 4 and Chapter 5 remain unpublished. Chapter 4 was a collaborative effort with Mr. Antwi; I performed the structural elucidation experiments and data analysis using materials synthesized by Mr. Antwi. The modelling in Chapter 5 was performed with advice from Dr. Kaitlyn Towle and the outlined residue synthesis was completed by myself.

Dedication

I would like to dedicate this thesis to my parents, Lili & Sorin Chiorean,

who gave me the autonomy to pursue my own adventure.



"What she tackles, she conquers."

– Richard Gilmore

Acknowledgements

I would need to compile a second thesis to truly acknowledge all the people who have supported me in this endeavour. Foremost, I would like to thank John Vederas, a brilliant chemist who has supported not only my scientific research but my growth as a well-rounded academic. His contagious enthusiasm for science, thirst for knowledge, and earnest support are unparalleled – our countless short chats will not be forgotten. Thank you for helping me climb the first of many mountains!

The group of researchers I had the pleasure to work alongside made the Vederas Group a wonderful learning environment from my first day to my last. I would like to thank Kaitlyn Towle for her patient mentorship and warm friendship and Marco van Belkum for taking the time to share his expertise and push me to excel as a scientist and an academic writer. Thank you to Fabricio Mosquera Guagua, my first Edmonton friend and constant squash buddy, and Isaac Antwi, my hardest working collaborator – you both made the late nights and frustrating times in the lab bearable. Thank you also to Jonathan Beadle for the priceless life advice and to Tess Lamer for reviewing this document.

My research was only possible on account of the fantastic faculty and staff I have been privileged to learn from. Thank you to Jing Zheng, Angelina Morales Izquierdo, Béla Reiz, Randy Whittal, Wayne Moffat, Jennifer Jones, Gareth Lambkin, Ryan McKay, Mike Barteski, Ryan Lewis, Todd Lowary, Connor Prat, and Hayley Wan for your support and willingness to answer my multitude of questions. Also, a mountain-sized thank you to Mark Miskolzie for the patience, guidance, and delight you brought to all my NMR work.

Outside of the lab, I had the pleasure of meeting a group of inspirational people that came together to form the University of Alberta Working for Inclusivity in Chemistry initiative. Thank you to Meagan Oakley for diving into this adventure with me and for being my friend through it all. Thank you to Jocelyn Sinclair for the patience and wise advice, to Sarah Milliken for ensuring I maintained a social life, and a special thank you to Rochelin Dalangin for trying to keep me sane with overpriced coffee and witty repartee. I am immensely proud of what we have achieved – I am excited to see where our paths will take us next!

To my dear friends Michelle MacMillan, Volition La, and Laura Brenner: your friendship during the last decade has meant more than I can express. Thank you for celebrating my wins and bringing me up from my lows.

Last and certainly not least, I would not be where I am today if it was not for the support of my family. A heartfelt thank you to my devoted parents, Lili and Sorin, and my sister, Andreea, who have put up with me for the last quarter century and never once told me there was something I couldn't achieve. This journey would have been impossible without the love and encouragement you have unconditionally given me my entire life. My degree would not have been the same without the companionship of Randy Sanichar. Thank you for your unwavering support, brilliant synthetic insights, and continuous friendship over these past few years. I look forward to sharing the next pages of my life with you.

This thesis is a testament to the incredible individuals I have met along the way and who, in big ways or small, have contributed to its completion. From the bottom of my heart, I thank you.

Contents

List of Tables

List of Figures

List of Schemes

List of Symbols, Units & Abbreviations

$[\alpha]^{25}_{D}$	specific rotation at 25 ^{o}C
Ac	acetyl
Ace	acetone
ACN	acetonitrile
AMP	antimicrobial peptide
AMR	antimicrobial resistance
ATP	adenosine triphosphate
BLAST	Basic Local Alignment Search Tool
Boc	*tert*-butyloxycarbonyl
Bu	butyl
c	concentration in g/mL
CD	circular dichroism
COSY	correlated spectroscopy, ^{1}H-^{1}H
d	doublet
Dab	2,4-diaminobutyric acid
DAP	diaminopimelic acid
DIC	*N*,*N*'-diisoproprylcarbodiimide
DCM	dichloromethane
DIBAlH	diisobutylaluminum hydride
DIPEA	*N, N*-diisopropylethylamine
DMAP	4-(dimethylamino)pyridine
DMF	dimethylformamide

DMSO	dimethyl sulfoxide
DNA	deoxyribonucleic acid
DPC	dodecylphosphocholine
DSS	3-(trimethylsilyl)propane-1-sulfonate, sodium salt
DTT	1,4-dithiothreitol
EDCI	1-ethyl-3-(3-dimethylaminopropyl)carbodiimide
End	enduracididine
equiv	equivalent(s)
ESI	electrospray ionization
Et	ethyl
FaeMK	faerocin MK
Fmoc	fluorenylmethoxycarbonyl
FTIR	Fourier transform infrared spectroscopy
HATU	1-[bis(dimethylamino)methylene]-1H-1,2,3-triazolo[4,5-b]pyridinium 3-oxide hexafluorophosphate,
HBTU	2-(1H-benzotriazol-1-yl)-1,1,3,3-tetramethyluronium hexafluorophosphate
Hex	hexanes
HFIP	hexafluoroisopropanol
HOAt	1-hydroxy-7-azabenzotriazole
HOBt	hydroxybenzotriazole
HPLC	high performance liquid chromatography
HRMS	high-resolution mass spectrometry
Ile	isoleucine

IPTG	isopropyl β-D-1-thiogalactopyranoside
J	coupling constant, in Hertz
K_d	dissociation constant
kDa	kilodaltons
LCMS	liquid chromatography mass spectrometry
LCMS/MS	liquid chromatography with tandem mass spectrometry
LII	lipid II
m	multiplet
m/z	mass to charge ratio
MALDI	matrix assisted laser desorption/ionization
MALDI-TOF	MALDI-time-of-flight
MccJ25	microcin J25
Me	methyl
MS	mass spectrometry
Ms	methylsulfonyl, mesyl
MQ	Milli-Q, Millipore Quality filtered
Ni-NTA	nickel-nitrilotriacetic acid
NMP	*N*-methyl-2-pyrrolidone
NMR	nuclear magnetic resonance
NOESY	nuclear Overhouser effect spectroscopy, ^{1}H-^{1}H
NRP	nonribosomal peptide
NRPS	nonribosomal peptide synthetase
Ns	nitrobenzenesulfonyl, nosyl

NTP	nucleoside triphosphate
Ph	phenyl
ppm	parts per million
PyBOP	(benzotriazol-1-yloxytripyrrolidinophosphonium hexafluorophosphate)
q	quartet
quant.	quantitative yield
RBF	round bottom flask
RNA	ribonucleic acid
RP-HPLC	reverse phase-high performance liquid chromatography
s	singlet
SDS-PAGE	sodium dodecyl sulfate polyacrylamide gel electrophoresis
SPPS	Solid-phase peptide synthesis
SUMO	small ubiquitin-like modifier
t	triplet
Tf	triflyl
TFA	trifluoroacetic acid
TFE	2,2,2-trifluoroethanol
THF	tetrahydrofuran
TIPS	triisopropylsilane
TLC	thin layer chromatography
TMS	tetramethylsilane
TOCSY	total correlation spectroscopy, ^{1}H-^{1}H
$TriA_1$	tridecaptin A_1

Txb	teixobactin
UDP	uridine diphosphate
UTP	uridine triphosphate
UV	ultraviolet spectroscopy
δ	chemical shift in parts per million

Chapter 1: Introduction

1.1 Bacteria: The good, the bad, and the useful

Since the discovery of "animalcules" in 1676 by Antonie van Leeuwenhoek,[1] this domain of organisms have left their mark on humankind. From the devastating effects of infectious diseases to the indulgence of wine and cheese, the role bacteria play in our society is an intricate balance of good and bad, dangerous yet beneficial. More recently, studies looking at large bacterial aggregations have yielded interesting results. Investigating the community of cells found in the human gastrointestinal tract (gut microbiome) and the network bacteria can form when adhered to a surface (biofilms) has reinforced the idea that bacteria play a useful role in our society.[2]

Studies have linked the gut microbiome, and the natural products produced by these organisms, to a person's physiological and psychological wellbeing. Gut bacteria believed to contribute positively to human health, commonly termed probiotics, arrive in the gastrointestinal tract *via* the food we consume, including breast milk as infants. For example, a probiotic formulation containing *Lactobacillus helveticus* has been shown to reduce anxiety in mouse models as well as in human volunteers.[2a] Often viewed as harmful and difficult to remove, biofilms and the mechanisms by which bacteria colonies form them can be hijacked and repurposed. An interesting facet of research has looked at manipulating the contents of the extracellular matrix excreted by bacteria to create purposefully designed biofilms. These biofilms can be used to form biomaterials for the textile industry, clean polluted waters, perform desired biocatalytic reactions, and even colonize the gastrointestinal tract in patients with chronic inflammatory diseases to reduce swelling.[2c]

Bacteria have evolved mechanisms to assemble the basic building blocks of life and create natural products with a kaleidoscope of targeted functions. Amino acids, lipids, sugars, nucleic acids, and small molecules have been anabolized by bacteria into secondary metabolites and natural products with specific purposes. Through the isolation, characterization, and study of these natural products, scientists have been able to apply and adapt the compounds' original function to benefit society. Bacterial natural products have proven useful in several industries including agriculture, food, beauty, medicine, and research.[3]

One such native purpose of these natural products is to target and kill competing bacteria vying for the same resources in their ecological niche. This innate activity can be exploited and applied in areas such as bacterial infection remedies and food preservation.[4] This avenue of research is of particular interest as the rise in antimicrobial resistance (AMR), the ability of bacteria to survive antibiotic treatments, is a growing concern in modern medicine.[5]

Since the discovery of penicillin by Alexander Fleming in 1928,[6] a golden age of antibiotic natural products emerged and subsequently vanished. Each decade since the 1940s, when treatment of World War II soldiers with penicillin showed success in preventing bacteria-infected wounds,[7] has seen the discovery of 20 – 30 new drug candidates.[5] Regrettably, bacteria found ways to fight back and developed resistance, prompting scientists to expand upon the natural products and produce synthetic derivatives of penicillin and other antibiotics.[5] In response, bacteria once again evolved pathways to withstand these newly developed antibiotics, now at seemingly faster rates. Presently, it is estimated by the Infectious Diseases Society of America

that more than 70% of pathogenic bacteria have become resistant to antibiotics currently on the market (Figure 1.1)[5a] *Enterococcus faecium*, *Staphylococcus aureus*, *Klebsiella pneumoniae*, *Acinetobacter baumannii*, *Pseudomonas aeruginosa*, and Enterobacter species, known as the ESKAPE pathogens, have been classified by the World Health Organization (WHO) as the most alarming due to their multidrug resistance.[5c]

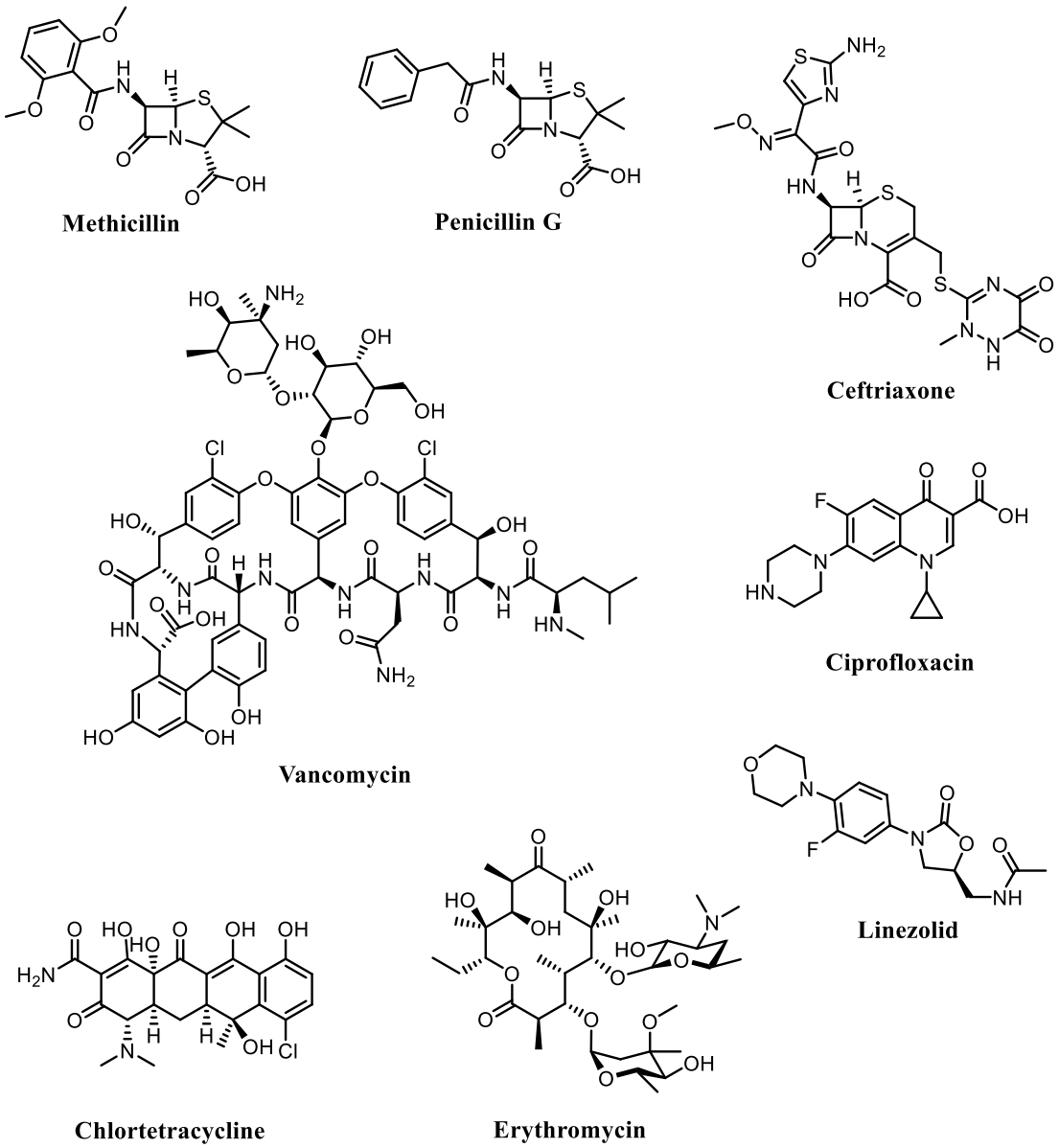


Figure 1.1 Antibiotics on the market with documented resistance

The consequences of leaving this issue unaddressed are dire – it is projected that over 10 million annual deaths will be caused by untreatable bacterial infections; a fourteen-fold increase from the 700,000 annual deaths observed worldwide in 2019.[5c] Concurrently, the world economy is also impacted by the strain caused to medical systems. Current trends predict an annual deficit of 10 billion USD if the issue of antimicrobial resistance is not addressed.[5] Peptides, a subclass of bacterial natural products, are a promising alternative to currently available antibiotics.[8] They are compatible with biological systems and have specific bacterial targets, making them generally recognized as safe for human consumption.[8]

1.2 Antibacterial peptides

Peptides with the ability to kill bacteria are often called antimicrobial peptides (AMPs) or, more specifically, antibacterial peptides. AMPs have been isolated from various organisms, including bacteria, fungi, and mammals. This designation encompasses a broad range of peptide-derived compounds which have various sizes, shapes, properties, and targeted activities. While different, these compounds do share a few common features. AMPs tend to be amphiphilic or hydrophilic and cationic, characteristics which help guide them towards bacteria's anionic membranes, and of relatively small size (< 60 amino acids).[9] Though a detailed mode of action has yet to be elucidated for each peptide, it is proposed most of these peptides interact with the membrane or receptors on the membrane surface, disrupt the membrane, and form pores.[10] This membrane perturbation results in cell components leaking out, which in turn leads to cell death. In certain cases, these peptides can also target a particular

cellular function that inhibits the target bacteria's ability to survive. For example, a peptide can bind within the bacterial ribosome complex, prevent the biosynthesis of essential proteins, and eventually lead to the cell's demise.[11]

Due to the similar topology of bacterial membranes and the parallels between their cellular function, AMPs synthesized by bacteria have the potential to be self-harming and attack the producer. Therefore, the producing organism has a clever counter measure: an immunity protein that is capable of defending the producer against its own AMP.[12] Although not much is known about each protein's exact mechanism, the immunity protein is able to counteract the effects exerted by the AMP.

In order to utilize these peptides in relevant applications, one must first study and understand their properties, their structures, and their mode of action. Interdisciplinary approaches which combine native production of peptides with synthetic methods and spectroscopic techniques fill a valuable crevice in the quest to uncover structural insights of AMPs. Of focus in this thesis are peptides which are biosynthesized by bacteria, in one of two ways described in the following pages.

1.2.1 Ribosomally synthesized AMPs

Ribosomally synthesized peptides (RSP) are widespread throughout nature; they are present in bacteria as well as fungi, plants, and animals.[13] In line with more traditional biosynthetic machinery of larger proteins, ribosomally synthesized peptides are made from the standard 20 amino acids with the help of the ribosome and can then undergo

a series of post-translational modifications to further diversify and embellish their structures.

1.2.1.1 Biosynthesis of RSPs

In short, upon the detection of a stimulus, the encoding DNA, which may be located on the chromosome or a plasmid,[14] is transcribed to messenger RNA (mRNA).[15] The ribosome then reads and translates the mRNA into the encoded protein material.[15] The number of genes required for the biosynthesis of AMPs varies. In Gram-positive bacteria they range from as few as two to as many as twenty-one genes,[14,16] depending on the complexity of the downstream processing required. At minimum, two genes encode for the precursor peptide and the immunity protein. The precursor peptide (Figure 1.2) contains the core peptide, which will be modified into the final, active peptide, and sometimes a leader sequence, typically found at the N-terminus.[13]

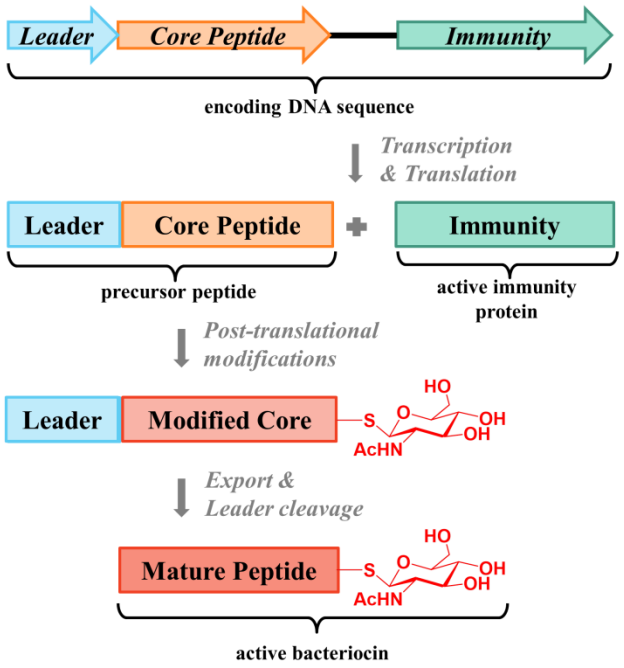
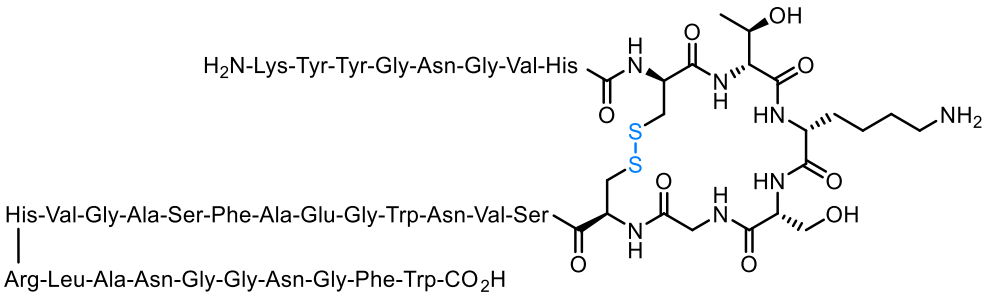
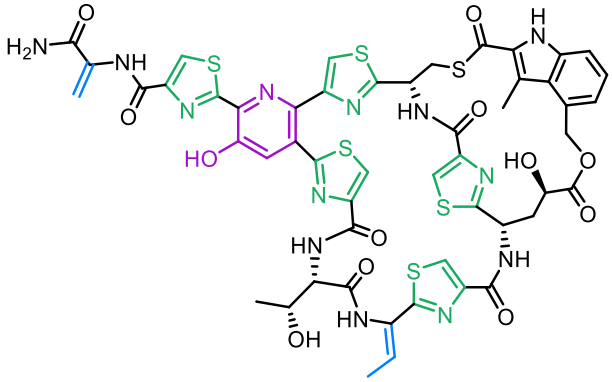


Figure 1.2 Example of a biosynthetic pathway of ribosomally synthesized AMPs

The leader peptide helps the sequence to be recognized by the enzymes which perform the post-translational modifications, often by the export proteins which will secrete the active peptide into the external environment.[17] The additional genes account for all the modifying enzymes required to process the core peptide to the desired product.[13] As previously alluded to, some RSP gene clusters do not contain any modifying enzymes and therefore undergo little to no post-translational enzymatic processing. For example, the formation of a disulfide bridge between two Cys can be an enzyme-less peptide modifications. On the other hand, some peptides require several enzymes and become almost unrecognizable from their original primary amino acid sequence (Figure 1.3).[13,14]



Leucocin A



Nosiheptide

Figure 1.3 Examples of unmodified (top) and highly modified (bottom) RSPs

Additionally, in some cases, a series of auxiliary proteins and enzymes required for the secretion of the mature peptide are also encoded within the peptide's gene cassette. If RSPs cannot be excreted by the *sec*-pathway, the general secretory pathway ubiquitously used by bacteria to transport the majority of their protein to or outside of the membrane,[18] an alternative mechanism must be used. A common alternative used by RSPs is translocation facilitated by the adenosine triphosphate (ATP)-binding cassette (ABC) transporters.[19] The ABC transporters are a family of membrane transport protein complexes which use ATP hydrolysis to actively transport molecules across membranes.[20] In addition to exporting RSPs to the extracellular matrix, the ABC transporters may also have catalytic functions. The transporters recognize and cleave a double Gly sequence at the C-terminus of the leader peptide, freeing up the core peptide.[21] This keeps the peptide in its precursor, inactive form while within the cell, only becoming the mature, active peptide once the ABC transporter has recognized it and proceeded to expel it outside of the cell.

1.2.1.2 Structural properties and classification of RSPs

The group of AMPs that are ribosomally synthesized by bacteria are commonly referred to as bacteriocins, which are further subdivided based on whether they are produced by Gram-positive or Gram-negative bacteria as well as their structural features. Literature has bestowed several designations to these molecules over the decades, some of which are now obsolete as new RPSs are uncovered and their structures elucidated. For this thesis, the designations in Table 1.1 will be introduced and used throughout.[22]

Table 1.1 Classifications of bacteriocins[22]

Gram-positive origin		
Designation	**Structural features and properties**	**Example**
Class I (modified)		
Lantibiotics	(Methyl)lanthionine residues	Nisin A
Lipolanthines	N-terminal fatty acid, avionin moiety (aminovinylcysteine-labionin hybrid)	Avionin
Linear azol(in)e-containing bacteriocins	Thiazole and (methyl)oxazole rings, linear backbone	Plantazolicin
Thiopeptides	6-membered nitrogen heterocycle, azol(in)e rings, dehydro residues	Nosiheptide
Bottromycins	Macrocyclic amidine, decarboxylated C-terminal thiazole, β-methylated residues	Bottromycin A2
Sactibiotics	Cysteine sulfur to α-carbon bridges	Subtilosin A
Lasso peptides	N-terminal amine to γ-acid residue cyclization, C-terminal tail threaded through ring	Sviceucin
Glycocins	Contain glycosylated residue(s)	Glycocin F
Head-to-tail cyclized bacteriocins	N- to C-terminal cyclization	Carnocyclin A
Class II (unmodified)		
YGNG-motif containing bacteriocins	YGNG consensus motif, minimum one disulfide bridge, antilisterial	Faerocin MK, Leucocin A
Two-peptide bacteriocins	Two peptides acting together for synergistic activity	Carnobacteriocin XY
Leaderless bacteriocins	Produced with no leader sequence	Aureocin A53
Other linear bacteriocins	Not YGNG-like, linear peptides	Lactococcin 972
Class III (large, heat-labile)		
Bacteriolysins	Large polypeptides, bacteriolytic	Lysostaphin
Non-lytic large bacteriocins	Large polypeptides, not bacteriolytic	Caseicin 80
Tailocins	Multi-protein complex, bacteriophage tail-like	Diffocin
Gram-negative origin		
Designation	**Structural features and properties**	**Example**
Modified	Contain post-translational modifications	Microcin J25
Unmodified	Unmodified, generally 10 kDa or smaller	Microcin V
Large	Generally larger than 20 kDa	Colicin A
Tailocins	Multi-protein complex, bacteriophage tail-like	Pyocin

Bacteriocins from Gram-positive organisms are filed under three main classes based on whether they are post-transitionally modified (Class I), unmodified (Class II), or large and heat-labile (Class III).[22] Class I contains the bacteriocins which require extensive enzymatic help for the core peptide to reach its final, mature form. At present, nine subclasses exist, including lasso peptides such as microcin J25. Class II contains bacteriocins which have no or minimal (ie. removal of leader peptide) post-translational modifications. Four subclasses exist, including YGNG-motif containing bacteriocins such as faerocin MK and leucocin A. Class III contains the bacteriocins which are heat-labile and large, approaching protein-level molecular weights. Three subclasses exist, including tailocins which are phage-like protein complexes that form pores in bacterial membranes in a similar manner as bacteriophages.[23]

Bacteriocins from Gram-negative organisms contain a less complex classification system, dividing the RSPs into four categories: modified, unmodified, large, and multi-protein tailocins. This is largely due to their lower abundancy compared to Gram-positive bacteriocins; of the many reported bacteriocins, only about 10% are of Gram-negative origin.[24] These are microcins (low molecular weight, $\leq$ 10 kDa) which are either modified or unmodified,[25] colicins (medium molecular weight, 20 to 80 kDa),[26] and tailocins (high molecular weight, multi-peptide complexes).[27] Structurally, some of these peptides resemble those in the Gram-positive classification and were in fact isolated from Gram-negative organisms first before their identification in Gram-positive bacteria; microcin J25 is one such peptide. Bacteriophage-like tailocins is another example of a class that has been found in both bacterial types.

This thesis will discuss two bacteriocins: faerocin MK in Chapter 2 (YGNG-containing bacteriocin of Gram-positive origin), and microcin J25 in Chapter 5 (a lasso peptide of Gram-negative origin) in more detail. An advantage of RSPs is that genetic manipulation can be employed to design and express native peptides as well as related analogues, or to isotopically label peptides *via* relatively inexpensive means through bacteria grown on labelled media. Ribosomal synthesis, however, limits the functional groups present in peptides to those accessible by this biosynthetic pathway.

1.2.2 Nonribosomally synthesized AMPs

Another route bacteria may take to biosynthesize antibacterial peptides is *via* nonribosomal peptide synthetases (NRPSs), multi-domain enzyme complexes, which produce correspondingly named nonribosomal peptides (NRPs). Arguably the most famous compound which falls under this label is penicillin, the first uncovered antibiotic.[1] The existence of NRPSs was first alluded to in 1963 with a study into the biosynthesis of tyrocidine, a cyclic antimicrobial peptide isolated from *Bacillus brevis*.[28] This cyclic decapeptide contains ornithine and D-amino acids, features which piqued the curiosity of researchers to investigate the matter further as these residues were not previously identified elements in known proteins. This initial study fed the producing bacterium aureomycin and chloramphenicol (ribosome inhibitors) and concluded the production of tyrocidine must be autonomous from the known protein synthesis relying on the ribosome. Further work, mostly involving gene knockouts, enzyme isolations, and feeding assays with isotopically labelled

substrates, allowed for a clearer picture to emerge regarding this ribosome-independent biosynthetic pathway.[29]

1.2.2.1 Biosynthesis of NRPs

The popularity of genome sequencing and, more recently, genome mining, has allowed for the development of powerful database tools to search and identify gene clusters within various organisms that encode for NRPSs.[30] Numerous bacteria, as well as fungi (Table 1.2), contain genes that encode for these large enzyme complexes responsible for producing antimicrobial NRPs.[29]

Table 1.2 Antimicrobial nonribosomal products from bacteria and fungi

Bacteria		
Class	**Species**	**Example**
Actinobacteria	*Amycolatopsis orientalis*	Vancomycin
	Streptomyces antibioticus	Actinomycin
	Streptomyces roseosporus	Daptomycin
	Streptomyces venezuelae	Chloramphenicol
Bacilli	*Bacillus subtilius*	Bacitracin & Surfactin
	Brevibacillus brevis	Tyrocidine & Gramicidin
	Paenibacillus polymyxa	Polymyxin & Tridecaptin
Fungi		
Phylum	**Genus**	**Example**
Ascomycota	*Penicillium*	Penicillin G
	Acremonium	Cephalosporin

Often all the elements required to synthesize and secrete a single molecule are within one gene cluster. This includes the enzymes responsible for subunit synthesis and attachment of other functionalities, secretion proteins, a self-protecting immunity factor, and of course at least one NRPS.[31] NRPS are mega protein complexes

composed of a series of modules with dedicated catalytic roles performed in a collinear biosynthetic process. Each elongation module links a designated amino acid to the growing peptide chain in the assembly line and performs any necessary modifications to that residue. In order for the module to perform its predictated task, it has a set of domains (minimum 3), each responsible for part of the job. The required domains are: an activation domain (A) which initiates the process by adenylation of the desired amino acid with ATP, a thiolation domain (T, also known as PCP or peptidyl carrier protein) which keeps the growing chain of the peptide tethered to the NRPS, and a condensation domain (C) which ligates the next residue *via* an amide bond to the growing chain. The first module in the sequence lacks a condensation domain and the last module in the sequence contains a thioesterase (TE) domain, which releases the peptide from the T domain though hydrolysis or macrocyclization. Additionally, modification domains such as epimerization (E), cyclization (Cy), oxidation (Ox), and methylation (M) can also be found within the corresponding module.[31] Outside of the NRPSs responsible for assembling the core peptide, a series of other tailoring enzymes may also be present. This allows for modifications such as halogenation, oxidative crosslinking, lipidation, and glycosylation.[32]

Vancomycin, a glycopeptide isolated from *Amycolatopsis orientalis* (previously known as *Streptomyces orientalis*), is a nonribosomally synthesized antimicrobial peptide that requires three NRPSs, encoded for by *VpsABC*, containing a total of seven distinct modules (Figure 1.4).[29,31,33] Interestingly, a single halogenase installs both chlorine atoms on residues 2 and 6; a process which has been proposed to occur

in parallel with the NRPS assembly, prior to the completion of module 7.[33] Additionally, three distinct P450 enzymes are responsible for the oxidative cross-linking of the linear peptide. These enzymes, encoded for by *OxyB*, *OxyA*, and *OxyC* link residues 6 & 4, 4 & 2, and lastly 7 & 5, respectively, to supply the polycyclic product. Finally, a set of glycosyltransferases install the two saccharide units at the hydroxyl group on residue 4 to yield the active vancomycin.

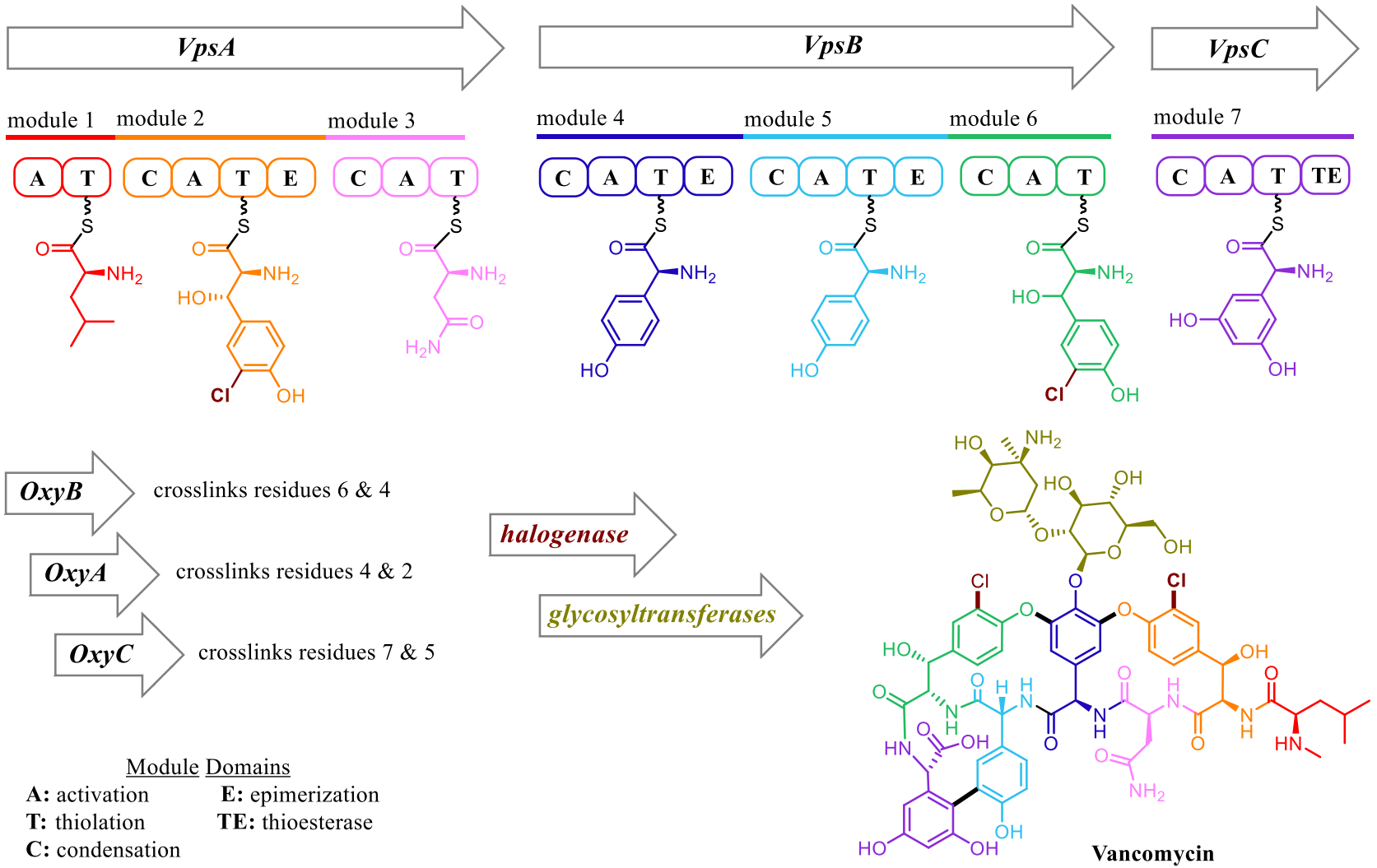


Figure 1.4 Biosynthetic pathway of vancomycin, a NRP

1.2.2.2 Structural diversity and classification of NRPs

A significant number of peptide drugs currently on the market stem from NRPs and are used as antimicrobials as well as immunosuppressants, antifungals, and antitumor treatments.[29] As alluded to by the plethora of NRPSs and tailoring enzymes, the variety of structural features present in NRPs is significantly more diverse than that found in RSPs. The incorporation of elements such as heterocycles and D-amino acids confer a higher degree of variance compared to those achievable by the 20 canonical amino acids. These features help facilitate NRP activity against bacteria (Figure 1.5).

The largest group is often considered to be cyclic NRPs, which include head-to-tail cyclized peptides (gramicidin S) and macrocycle containing peptides (bacitracin). Peptide ring constraints have been shown to be both stringent, in that they form a predefined structure primed for target selectivity, and flexible, which allows for malleability that can optimize binding surface interactions with their target.[35] Another major type of NRPs are lipopeptides, peptides that possess an acylated lipid to their N-terminus, often including both D- and L-residues, and can exist as linear peptides (tridecaptin A_1) or with a cyclic component (polymyxin B). Lipidation, particularly in the context of membrane-interacting AMPs, is believed to endow attributes that promote peptide insertion into bacterial membranes as well as confer stability.[36] More complex NRP classes include β-lactams (penicillin G and its many derivatives) and glycopeptides (vancomycin), which are both engineered by nature to suit their bioactivity with precision. β-lactams mimic their target's natural substrate and irreversibly bind in its active site.[37] In the case of vancomycin, the saccharides align the peptide hydrogen bonding network to maximize target interaction.[38]

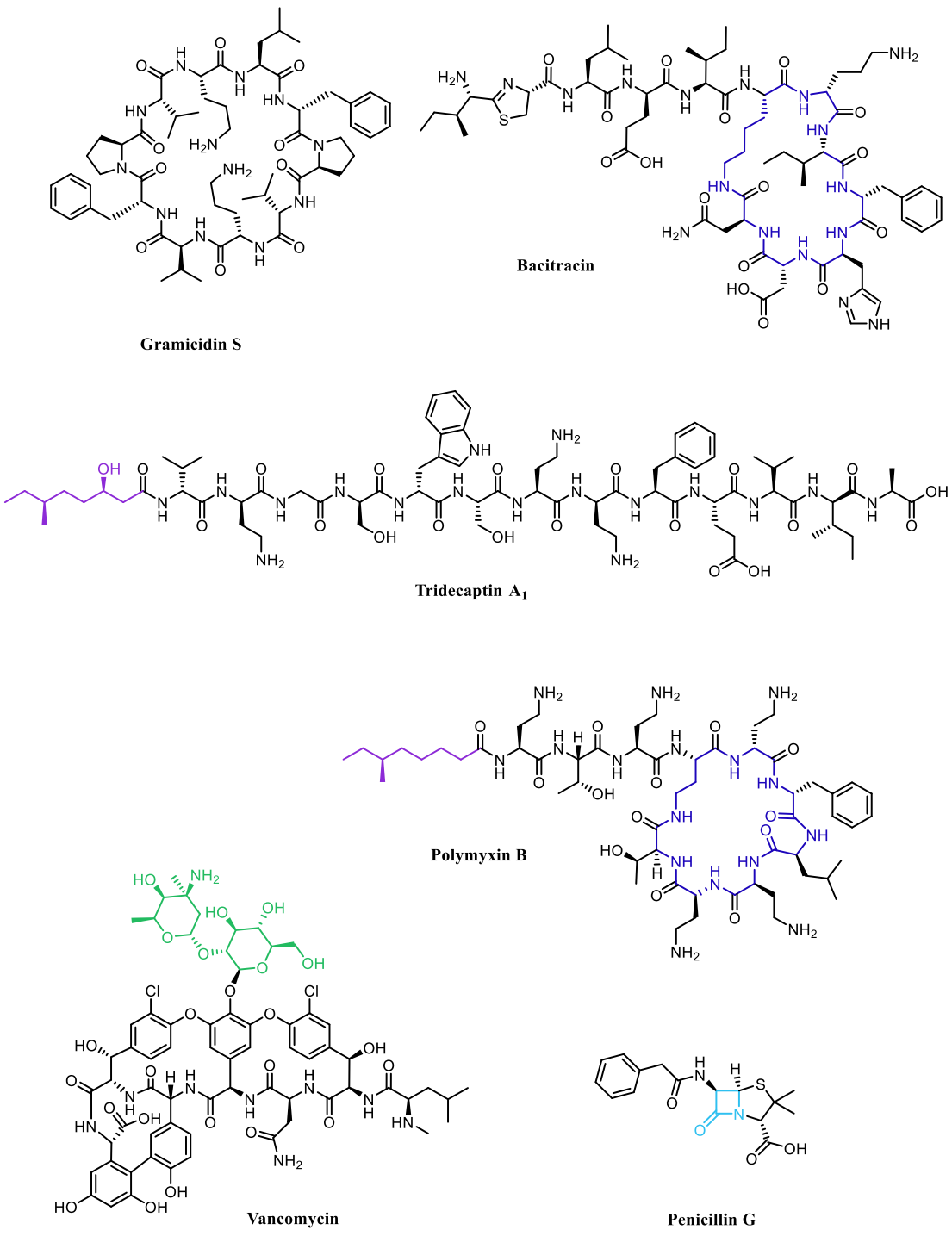


Figure 1.5 Representative structures highlighting diverse NRP features

A literature consensus on a concrete classification system has not been reached as new classes are continually being uncovered. Unlike RSP-encoding genes, which can be relatively straightforward to transcribe and predict, NRPSs and tailoring enzymes are not as intuitively perceived. Genome sequencing, genome mining, and bioinformatics tools have allowed for previously unidentified NRPs' biosynthetic gene clusters to be analyzed.[39] The 'great plate count anomaly' is another factor that contributes to the ever-growing classes of NRPs. The ability to culture microorganisms in laboratory settings where they can be studied is limited to known techniques, established practices, and available media and materials.[40] It has been predicted that in a given sample obtained from the natural environment, current methods have successfully cultured only about 1-2% of organisms, with as little as 0.01% for marine-origin samples.[41] This limits the biodiversity of bacteria isolated and sequenced, and therefore the classes of known and characterized NPRs.

Advances in computational tools have cleared the way for more facile predictions of NRPS-encoding gene products and the design and the use of metagenome library screening has enabled the discovery of genes from bacteria without the need to culture the producer.[42] Through the identification, heterologous expression, and purification of predicted NRPs, a series of new antimicrobial peptides have been uncovered in the last five years alone (Figure 1.6).[43] A metagenome screening technique was developed which uses a tag sequence to uncover novel epoxyketones, potent protease inhibitors.[43a] After the success of this initial study, the technique was used to expand the search, leading to the discovery of calcium-dependent, acidic, antimicrobial lipopeptide families malacidins and cadasides.

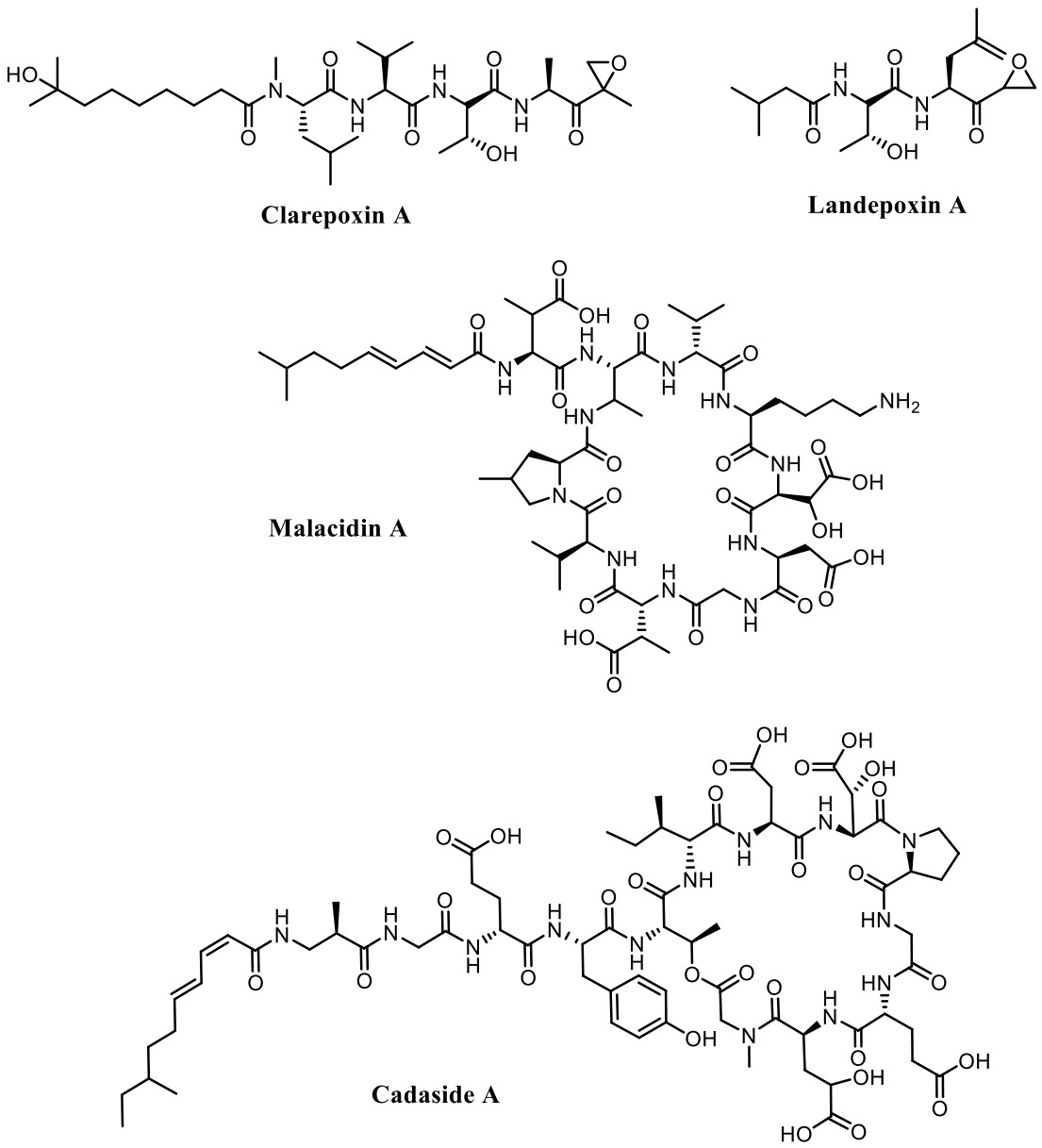
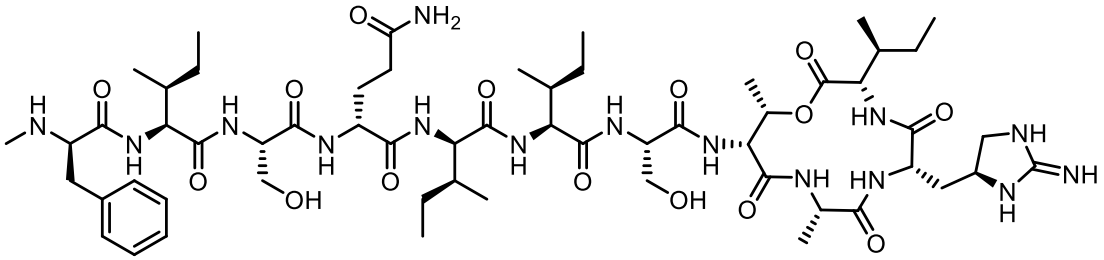


Figure 1.6 Antimicrobial NRPs discovered from metagenome library screening

Although proven a fantastic tool in recent years, the drawback of this screening requires prior knowledge of the NRPS system. For example, to uncover the malacidins, the A-domain responsible for installing the first amino acid in the core motif of other acidic lipopeptides was determined to be the most conserved and used as the tag to search for other, related NRPSs.[43b]

New methods of collecting, purifying, and culturing previously uncultured bacteria have also been developed and proved useful in uncovering novel NRPs and other secondary metabolites. One such technique is the isolation chip (iChip) technology which uses several diffusion chambers to collect bacteria, along with nutrients from their native environment, and segregate each cell into its own individual channel.[44] This allows the bacterium to be purified, but not isolated from its much needed native nutrients, and to grow into a colony with adequate DNA available for sequencing. Although not perfect, this method is able to culture close to 50% of bacteria from their native environment, which is a significant increase over the usual 1-2% achieved previously.[44] In 2015, the iChip was used in the discovery of teixobactin (Txb), a NRP from *Eleftheria terrae*, a newly identified species of β-proteobacteria.[45] Teixobactin is an AMP with a strong activity against Gram-positive organisms and interesting structural features (Figure 1.7). This depsipeptide contains four D-amino acids out of eleven residues, a methylated N-terminus, a 13-membered lactone formed by the C-terminal residue Ile_{11} and Thr_8, and an enduracididine at position 10.



Teixobactin

Figure 1.7 Structure of teixobactin

Teixobactin (Chapter 3) and tridecaptin (Chapter 4) are NRPs which will be described in this thesis. The structural complexity afforded by this biosynthetic pathway

increases the chemical space that these peptides can fill compared to RSPs. Adding unique features, such as noncanonical amino acids, heterocyclic rings, *N*-methylated amino acids, saccharides, and fatty acids, often increases the peptide stability and improves desired targeted activity.[31] This has brought NRPs to the forefront as starting point scaffolds for pharmaceutical development to combat microbial infections as well as tumors and autoimmune complications.[29] However, unlike ribosomally synthesized peptides that can be made by readily-available, recyclable cellular machinery, each NRP requires the expression of a unique mega enzyme complex; a process that adds a costly metabolic load to the producing organism.

1.3 Limitations of natural antimicrobial peptides

Although nature has evolved for thousands of years to produce these unique peptides with diverse structures and activities, there are numerous limitations which need to be addressed in order for these compounds to have practical applications in our world. Of emphasis here are AMPs which may be promising alternatives to currently administered antibiotics. As highlighted earlier, the growing threat of resistance to existing antibiotics on the market is a continuous battle fought in the last several decades.[8] As a result of overuse and misuse of antibiotics in our society, resistance is now developing at faster rates than previously documented.[5b] Additionally, the transition of these compounds from a research setting to a feasible drug candidate requires other parameters, such as pharmacokinetics, to be optimized.[46]

1.3.1 Bacterial antibiotic resistance mechanisms

Even prior to human discovery, bacteria have been engaging in vicious biological warfare, creating substances to advance their own survival by diminishing the livelihood of competitors vying for the same ecological niche. In response to this attack, bacteria have developed mechanisms to counteract the effects of antimicrobial compounds. Antimicrobials target different cellular processes, leading to different approaches adapted by bacteria to become resistant.[47] These approaches can be categorized in one of three main mechanisms of resistance (Figure 1.8).

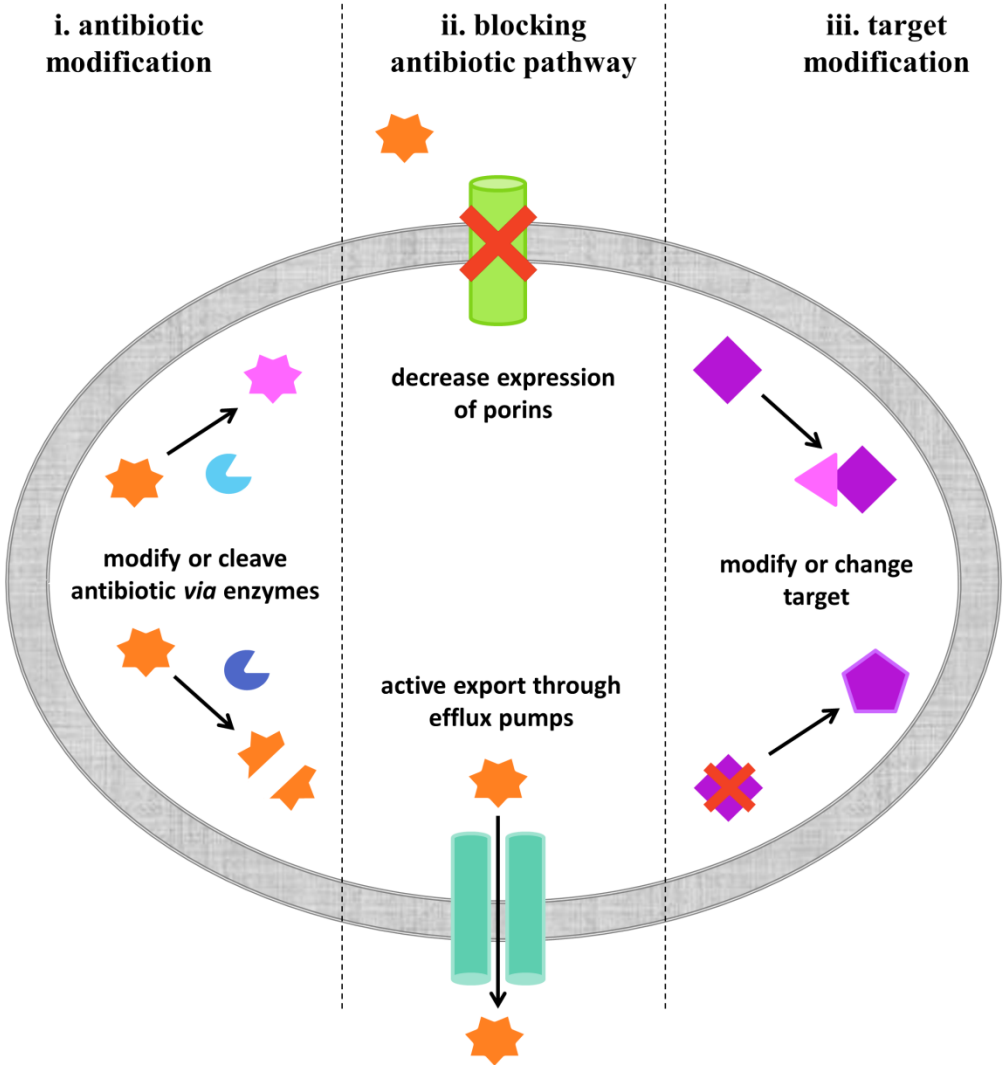


Figure 1.8 Mechanisms of antibiotic resistance developed by bacteria

i. Antimicrobial agent modification

Due to their specificity and relatively small size, any slight change or alternation of an antimicrobial may render it completely or partially inactive. Several examples of modifications facilitated by enzymes have been uncovered; bacteria may decorate intrusive antimicrobials with additional functional groups.[47] This leads to disruptions in the compound-target binding interface, likely due to steric hindrance and interfering interactions from the newly introduced moieties. For example, a slew of modifying enzymes have been identified for aminoglycoside antibiotics: acyltransferases, phosphotransferases, and acetyltransferases attach their corresponding moiety to the hydroxyl or amino group of aminoglycoside to render the antimicrobial inactive.[48] Moreover, bacteria may have genes that encode for β-lactamases, a family of enzymes that recognize and hydrolyze β-lactam rings, destroying antimicrobials such as penicillins, cephalosporins, and carbapenems.[49] These genes are often encoded on plasmids, allowing for bacterial strains to share this mechanism of resistance with other bacteria through horizontal gene transfer.[47]

ii. Inhibiting the antimicrobial's ability to reach the target

The first requirement of the antimicrobial to exert its activity is the ability of the compound to reach its intended target. Natural protectors, such as the double membrane of Gram-negative organisms, prevent certain antimicrobials from penetrating into the cells. The expression of porins, transmembrane proteins used by β-lactams and tetracyclines to enter cells, can be downregulated to limit the influx of unwanted antimicrobial intruders.[50] In addition to preventing passive permeability,

certain cells have also evolved efflux pumps which actively transport toxins out of the cell.[51] There are nominally five families of transmembrane pumps, differentiated by the type of bacterial membrane they are found in (Gram-positive or Gram-negative), the conformation they take, the energy source they require, and the varieties of compounds they are capable of exporting (broad or narrow spectrum).[52] A plethora of genes that encode efflux pumps which can remove some β-lactams, nisin, tetracyclines, erythromycin, and linezolid have been reported.[47,53]

iii. Target modification

The last stage of the antimicrobial journey is binding to its target and causing the adverse effects it was designed to instigate. Bacteria have developed ways of modifying the target beyond the antimicrobial recognition or simply bypassing the use of that target altogether.[47] Modification of the target can be achieved through genetic changes, which result in fluctuations to protein expression levels or mutated proteins, or through chemical alterations catalyzed by enzymes. Both of these approaches result in targets that are no longer susceptible to antimicrobials or targets with decreased affinity for their antimicrobial ligand.[53]

Resistance to ciprofloxacin, a fluoroquinolone that disrupts DNA replication by binding to important enzymes in the pathway, has developed at the chromosomal level. Mutations in the genes encoding for the target enzymes, DNA gyrase and DNA topoisomerase IV, result in enzymes with mutated subunits. These mutant enzymes are no longer recognized by ciprofloxacin, lowering or even abolishing its activity.[54] Similar systems have been uncovered for linezolid, a oxazolidinone which attaches at

the peptidyl transferase center of the 50S bacterial ribosomal subunit and causes cell death by preventing protein synthesis.[55] The chromosomal modifications result in mutant ribosome proteins that are no longer high affinity targets for the linezolid family of antimicrobials. Erythromycin, an antimicrobial which binds to the peptide exit channel of the 50S ribosomal subunit, has also become redundant against some species due to target modifications. Bacteria have evolved several enzymes that can interfere with erythromycin's capability of binding in this pocket of the 50S subunit. Methyltransferases catalyse the addition of mono- or di-methyl units to an adenine residue (A2058) in the 23S rRNA that is located at this binding site, thus impairing the antimicrobial activity of erythromycin.[56]

In some cases, bacteria can entirely bypass the need of a certain target. The prevalence of methicillin-resistant *Staphylococcus aureus* (MRSA) has risen in part due to these bacteria utilizing an exogenous protein in place of penicillin binding proteins.[57] Vancomycin, along with a series of other peptides, targets a peptidoglycan precursor known as lipid II. The antimicrobial specifically recognizes and binds to a dipeptide region (D-Ala, D-Ala) present on this lipid.[58] To circumvent this, vancomycin-resistant enterococci (VRE) have biosynthesized lipid II analogues to bypass the need for the original target. Bacteria created these new molecules by replacing the targeted D-Ala with D-lactate or D-Ser, which completely or partially prevent vancomycin from binding, respectively.

Several previously mentioned antimicrobials and the corresponding mechanisms of resistance developed by various bacterial strains are listed in Table 1.3. [49-61]

Table 1.3 Antimicrobials and their mechanisms of resistance

Antimicrobial Peptides		
Antimicrobial	**Target**	**Mechanism of Resistance**
Penicillins	penicillin binding proteins (DD trans- carboxy- endo-peptidases)	Agent modification, decreased cell permeability, target modification
Methicillin		Agent modification, decreased cell permeability, bypass target
Ceftriaxone		Agent modification, decreased cell permeability, target modification
Vancomycin	Peptidoglycan (D-Ala, D-Ala)	Target modification
Nisin A	Peptidoglycan (pyrophosphate)	Agent modification, efflux pumps, decreased cell permeability, target modification
Other Antimicrobial Agents		
Antimicrobial	**Target**	**Mechanism of Resistance**
Ciprofloxacin	DNA topoisomerase (gyrase)	Decreased cell permeability, target modification
Chlortetracycline	30S ribosomal subunit (A-site)	Decreased cell permeability, efflux pumps, target modification
Erythromycin	50S ribosomal subunit (exit tunnel)	Efflux pumps, target modification
Linezolid	50S ribosomal subunit (peptidyl transferase center)	Efflux pumps, target modification

As showcased in the table above, bacteria often use more than one approach outlined to neutralize the effect of antimicrobials. In addition to navigating through the various resistance tactics, antimicrobials must also display favourable pharmacokinetics with human bodies to be considered as drug candidates.

1.3.2 Implementing AMPs as effective therapeutics

In addition to bacteria finding mechanisms of resistance, there are a series of considerations that must also be addressed to successfully bring a peptide therapeutic to market. Nominally the behaviour of these compounds in biological systems is not

ideally suited to that of a drug candidate. Peptides may be inherently unstable; they are susceptible to proteolytic degradation by mammalian proteases, they may be oxidized or hydrolyzed, and they have the tendency to self-aggregate. [62] In assessing administration pathways, the oral availability of peptides is also a hurdle as the gastrointestinal tract is filled with proteolytic enzymes. Peptides also have high renal clearance, short half-lives in blood plasma, and low membrane permeability.[62] Challenges also exist in the production of these peptides, with the cost of production and purification being relatively high. [63]

Despite these few drawbacks, peptides are one of the fastest-growing contenders in the therapeutic drug market.[64] In 2011, approximately 60 peptide drugs had been approved with 140 others in clinical trials and 500 in preclinical development. This constituted a global market worth 14.1 billion USD at the time.[46] In under a decade, this number has reached 197 approved drugs with another 807 peptides undergoing clinical trials, in a market worth 28.15 billion USD as of 2019. Current trends have predicted this market will continue to increase and reach a net worth of over 60 billion USD in the next five years.[64]

The keen interest growing in these compounds as therapeutic candidates is not unwarranted. Peptides, especially when compared with larger proteins, tend to be less immunogenic, making them safe and tolerated in mammalian systems.[65] The amino acid residues are readily metabolized and cleared, ensuring low tissue attrition and lower toxicity.[62] AMPs have evolved to optimally interact and bind to their designated target, making them highly potent and selective compounds with limited adverse off-target effects.[63] Advances in chemical strategies to synthesize and modify

these peptides has lowered the cost of their production and increased the scope of achievable functionalities.[65]

1.3.3 Synthetic AMPs

The previous section outlined the two ways in which bacteria can biosynthesize native AMPs. Synthetic peptides, compounds made through the chemical manipulations of starting materials into desired polymers, are necessary to address concerns about resistance and pharmacokinetic behaviour as well as to optimize or fine-tune properties.[66] Strategic chemical modifications of key structural elements of peptides has been shown to improve the original activity and to add new, desired properties (Figure 1.9).

As mentioned beforehand, the penicillin class of molecules has been a victim of numerous resistance pathways limiting its use. Efforts to design analogues which can bypass the resistance protocols have produced a series of compounds such as flucloxacillin.[67] This compound, an isoxazole penicillin, is active against β-lactamase-producing Staphylococci. Arylomycins, a class of AMPs latently active against Gram-positive bacteria, has also been extensively studied in this regard. A series of analogues which can expand the activity scope to include Gram-negative organisms through variation of the lipid tail have been synthesized.[68] Another commonly administered antibiotic, vancomycin, has also seen upgraded analogues on the market which can withstand developed resistance.[69]

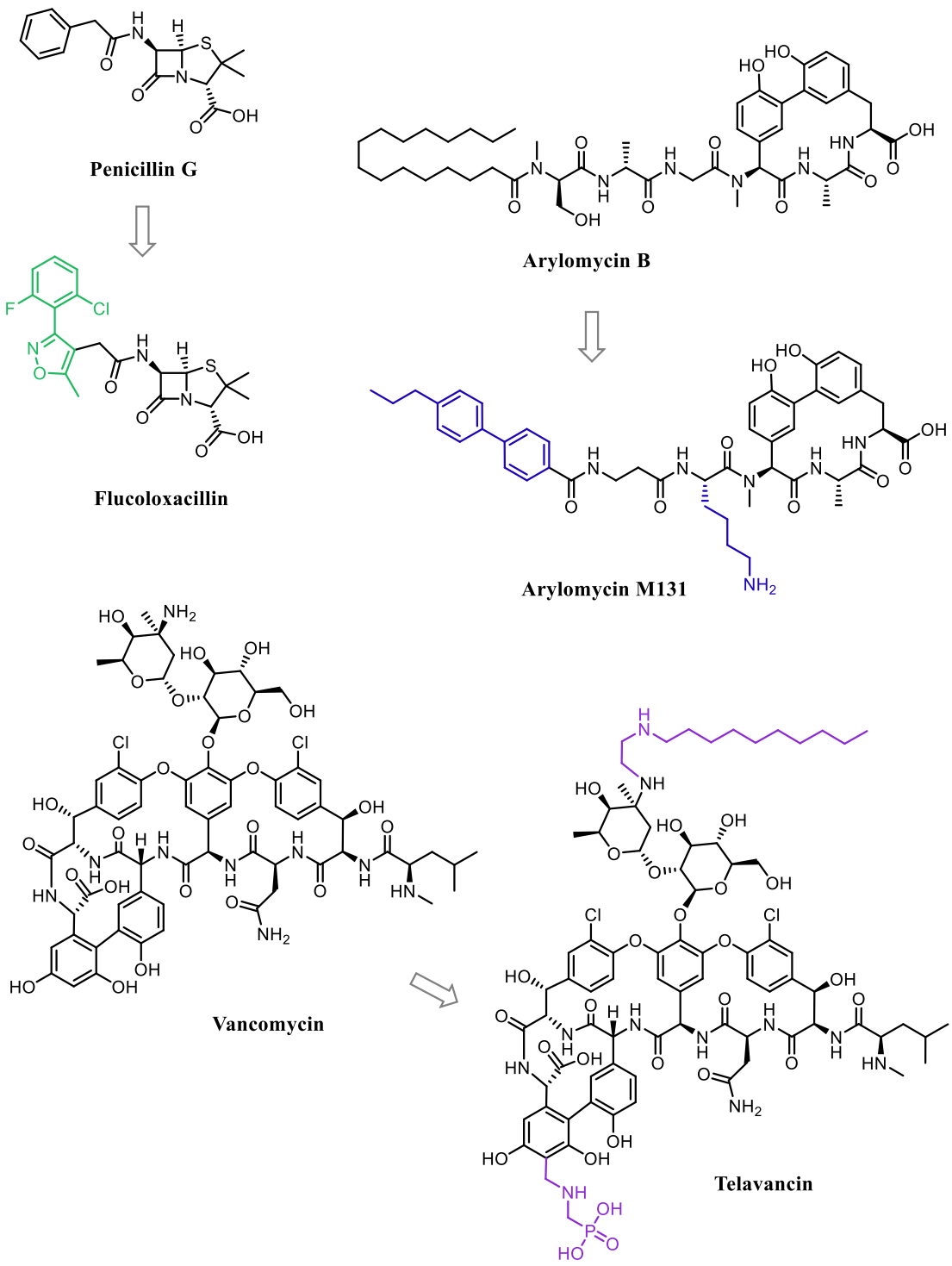


Figure 1.9 Native AMPs and their modified synthetic counterparts

Optimizing the features of native peptides has become a well-defined workflow through the use of Structure-Activity Relationship (SAR) studies.[46] Once a natural product has been identified, whether through isolation from nature or a library screening, a series of chemical approaches can be taken to better understand the peptide composition and modify it as desired.[46] The first step is usually to determine the minimum sequence needed to maintain activity by creating a series of truncated analogues. This would ensure extra residues and moieties which may be expensive or time consuming to synthesize can be omitted if they do not contribute to the desired activity. From the sequence which remains, the critical residues can be pinpointed by synthesizing a series of analogues, with each residue in turn being replaced by a non-offensive residue, such as Ala. This is now commonly known as an alanine scan and a routine experiment performed to probe the biological significance of each amino acid residue in peptides and proteins.[70]

If the AMP will be applied in a mammalian system, the roadblocks mentioned above need to be circumvented by increasing the stability of the peptide in the human body. To prevent degradation by exopeptidases, the termini of the peptide can be protected. This can be achieved through the acylation of the N-terminus or installing a primary amide at the C-terminus.[71] Endopeptidases, particularly ones which target amide bonds, are also an issue. Backbone modifications such as methylation of the amide bonds, using D-residues over L-residues, or entirely replacing the amide bond in favour of another non-native linkage can decrease proteolytic degradation by preventing recognition from endogenous proteases.[72] There have also been many

examples where cyclization, either of the entire molecule head-to-tail or macrocycles, tends to lend itself to stability and increased endurance against proteases. [73,35]

Lastly, the peptide can be optimized for activity by installing unconventional residues and functional groups at positions which are not detrimental to activity. If information about the target or binding pocked is available, rational design can help increase the number of attracting forces, such as hydrogen bonding or hydrophobic interactions, to strengthen binding.[74] Another factor that can be addressed to optimize activity is solubility, ensuring the compound can dissolve in biological media at high enough concentrations to exert its desired effect. The addition of hydrophilic charged residues or the limitation of hydrophobic regions not necessary for activity can help in this regard.[46]

1.4 Thesis overview

Nature has brought us this far; now it is up to researchers to build on the scaffolds we have been provided and design peptides with improved efficacy, stability, and which can be readily accessible. To take advantage of this innate antimicrobial activity, one must first understand the modus operandi of AMPs. Multipronged, interdisciplinary studies can help paint the bigger picture by giving insights into peptide expression systems, target binding affinities, and three-dimensional structures. The following chapters will recount in chronological order the advancements my projects have made in our understanding by analyzing the properties, activities, and structures of AMPs.

Chapter 2 will introduce faerocin MK (FaeMK), a newly uncovered peptide with potent antilisterial activity. A putative YGNG-containing peptide gene was observed in the genome of a bacterium isolated by our group and the sequence was used to express, purify, and study this peptide. Structure elucidation efforts are also outlined.

Chapter 3 will summarize the studies on teixobactin (Txb), a depsipeptide with an intriguing sequence and strong activity against Gram-positive organisms. The study of Txb and its receptor, lipid II (LII) were achieved using a series of analogues and isothermal titration calorimetry. Additionally, efforts to expand Txb's activity to include Gram-negative organisms are discussed.

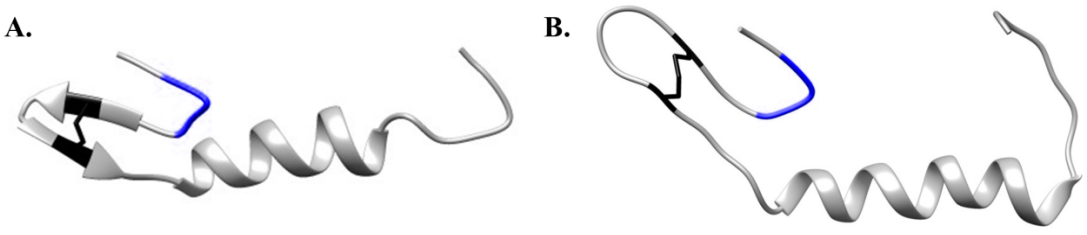
Chapter 4 will highlight the endeavours of NMR structure elucidation of tridecaptin A_1 ($TriA_1$) with its target, LII, in lipid membrane mimics. Often these membrane-interacting peptides undergo a structural change when they encounter a lipid-based solvent system or their intended target. Changes in the conformation of the peptide as it moves from an aqueous system to a lipid environment can give insights into the structural components that play key roles in membrane recognition, receptor binding, and peptide activity.

Chapter 5 will outline a unique approach to the synthesis of lasso peptide microcin J25 (MccJ25). Chemical access to the lasso-shaped peptides has been so far unreported as previous attempts led to unthreaded, and therefore inactive, peptides. The synthesis of two new building block residues allows for the formation of a temporary disulfide bridge, forming the hairpin-shaped peptide on resin. Future plans to cyclize and desulfurize the final peptide will also be proposed.

Chapter 2: Isolation and characterization of faerocin MK

2.1 Background

As highlighted earlier, ribosomally synthesized and post-translationally modified peptides are a large class of natural products that may have potent antimicrobial activity. Those biosynthesized by bacteria that are antimicrobial, are appropriately termed bacteriocins and fall under defined classes. YGNG-motif containing bacteriocins, also called type IIa or pediocin-like bacteriocins in previous literature, are a subtype of class II unmodified bacteriocins. The characteristics of this group are well defined: they are overall cationic, possess an N-terminal amphipathic α-helix, contain a conserved N-terminal region of residues (YGNG), and have at least one disulfide bond five amino acids apart.[22] These bacteriocins also possess a non-conserved C-terminal tail that is believed to fold back onto the helix and this area is often a flexible random coil. A couple select examples of this class are shown below in Figure 2.10, where the blue marks the conserved YGNG-motifs and the black marks the disulfide bridges.



Leucocin A
(PDB: 1CW6)

Enterocin HF
(PDB: 2N4K)

Figure 2.10 Structures of (A) leucocin A and (B) enterocin HF, two YGNG-motif containing bacteriocins

The YGNG-motif containing bacteriocins are generally excreted by an ABC transporter which cleaves the leader peptide (recognizing a Gly-Gly sequence). Leaders that do not contain the double Gly sequence are proposed to be a signal peptide for the *sec*-pathway. Their antimicrobial activity is exclusively observed against Gram-positive bacteria, with strong, nanomolar inhibition towards Listeria.[75]

Often the discovery of new AMPs in recent years comes from genome mining and dataset searching using software such as Basic Local Alignment Search Tool (BLAST). Isolating a novel genome and sequencing it allows one to search previously identified and annotated genomes in deposited databanks against a data set of interest. This comparison can be used to begin deciphering the metabolites of the bacterium under investigation prior to performing further studies.

The efforts of our group's collaborative work led to the isolation and genome sequencing of a novel bacterium, *Enterococcus faecium* M3K31, isolated from *Gyps fulvus* (Griffon vultures) in 2015.[76] Analysis of these data indicated the genome contained sequences encoding for three bacteriocins: enterocin HF, enterocin P, and an unknown peptide. The sequence of this unknown peptide, which I named faerocin MK (FaeMK), showed an 85% similarity to a previously reported YGNG-motif containing bacteriocin, SRCAM 602.[77]

SRCAM 602 was isolated from *Paenibacillus polymyxa* NRRL B-30509. However, the genome of this *Paenibacillus* strain was sequenced by our group and the corresponding gene was not discerned.[78,79] Rather, it appears as if the activity originally attributed to SRCAM 602 was from a lipopeptide.[80] Further studies discussed below support this hypothesis.

Sequence similarities between FaeMK and known YGNG-motif containing bacteriocins, such as the presence of concerted residues YGNG followed by two Cys residues at n and n+5 in the N-terminal region, [22] led us to believe FaeMK is also a bacteriocin of this class. This group of peptides has been extensively studied and is proposed to interact with the mannose phosphotransferase system (Man-PTS), a transmembrane sugar transport protein complex present within the cytoplasmic membrane of bacteria.[81] The results below summarize the efforts taken to confirm the classification of FaeMK by determining its activity and structure.

2.2 Results and discussion

2.2.1 Peptide sequence analyses

A comparison of FaeMK and SRCAM 602 mature peptide sequences with a series of other bacteriocins of this class was performed using Clustal Omega software.[82] The sequence alignments of select bacteriocins can be found in Figure 2.11 below, with the conserved YGNG region at the N-terminus highlighted in blue and the two Cys residues at n and n+5 positions in green (the leader sequences are omitted).

Carnobacteriocin B2: VN**YGNG**VS**C**SKTK**C**SVNWGQAFQERYTAGINSFVSGVASGAGSIGRR
Durancin GL: ATY**YGNG**VY**C**NKQE**C**WVDWNKASKEIGKIIVNGWVQHGPWAPR
Hiracin JM79: ATY**YGNG**LY**C**NKEK**C**WVDWNQAKGEIGKIIVNGWVNHGPWAPRR
Faerocin MK: ATY**YGNG**VY**C**NKQK**C**WVDWNKASKEIGKIIVNGWVQHGPWAPR
SRCAM 602: ATY**YGNG**LY**C**NKQKHYTWVDWNKASREIGKITVNGWVQH

Figure 2.11 Peptide sequence alignment of YGNG-motif containing bacteriocins

The bacteriocins chosen include a peptide studied previously by our group, carnobactericin B2,[83] as well as durancin GL[84] and hiracin JM79.[85] Both of these

showed very high sequence homology, 98% and 86% respectively, with the proposed FaeMK sequence. While the conserved sequence at the N-terminus is present in all five peptides shown above, SRCAM 602 is lacking a second Cys residue at position n+5. The lack of the second Cys residue at this position, or anywhere else along the reported sequence, means this peptide would be unable to form a disulfide bridge. The presence of a disulfide bridge, or a non-covalent interaction such as pi stacking, is key in this class of bacteriocins.[22] Without this interaction, and presumably the 3D structure conformation that results from it, the peptides are largely inactive.[86]

2.2.2 Antimicrobial activity bioassays

To determine if the activity profile matches that of known bacteriocins of this class, both FaeMK and SRCAM 602 were chemically synthesized and purified for these trials. Using a solid-phase approach on an automated peptide synthesizer, the two peptides were obtained in moderately good yields and purified by HPLC to approximately 95% purity. Next, assays were performed on both peptides to determine the minimum inhibitory concentrations (MICs) for each; MICs are reported in μM in Table 2.4.[14] Previously mentioned YGNG-motif containing bacteriocins are known to have potent activities against Gram-positive bacteria, specifically Listeria, with activity seen at low μM and nM concentrations.[75]A panel of bacterial indicator strains were selected for spot-on-lawn assays and encompassed both Gram-negative and Gram-positive bacteria, including two Listeria species.

Table 2.4 Minimum inhibitory concentrations of FaeMK and SRCAM 602

Indicator Strain	MIC (μM)	
	FaeMK	**SRCAM 602**
Gram-positive organisms		
Brochothrix thermosphacta ATCC 11509	–	–
Carnobacterium divergens LV13	0.125	–
Carnobacterium maltaromaticum UAL26	0.25	–
Enterococcus faecalis ATCC 7080	0.5	–
Enterococcus faecium BFE 900	1	–
Latilactobacillus sakei 706	2	
Lactococcus lactis ssp. *cremoris* HP	16	–
Listeria monocytogenes ATCC 15313	0.031	–
Listeria monocytogenes UAFM1	0.125	–
Staphylococcus aureus ATCC 25923	–	–
Gram-negative organisms		
Escherichia coli DH5α	–	–
Salmonella enterica serovar Typhimurium ATCC 13311	–	–
Salmonella enterica serovar Typhimurium ATCC 23564	–	–

– , no noticeable inhibition at 64 μM (highest tested concentration)

The initial screening contained a maximum concentration of 32 μM and only six organisms: three Gram-negative and three Gram-positive bacteria. Neither peptide was active against the Gram-negative bacteria, but FaeMK was active against all three Gram-positive strains tested in this initial trial. These results supported our hypothesis that the lack of disulfide bond would render SRCAM 602 inactive. To ensure this was not an isolated result, the modest scope of indicator strains was extended to include all those listed in Table 2.4 above and the maximum concentration tested was increased to 64 μM. The activity of FaeMK matched those of other bacteriocins in this class, with low MIC values against most Gram-positive bacteria and potent (MIC of 31 nM) activity against Listeria. However, no inhibitory activity was observed for SRCAM 602 towards any of the bacterial strains tested. Encouraged by these results, the structure of these peptides became my next point of investigation.

2.2.3 Circular dichroism

Optical activity can be used to characterize the relative portions of proteins and peptides which are folded into the two most commonly recognized secondary structures, α-helix and β-sheets, as well as random coil.[87] Proteins, in particular this class that is believed to interact with a membrane-bound receptor, can remain relatively unstructured in aqueous systems and undergo conformational changes when they encounter their target.[88] To this end, 2,2,2-trifluoroethanol (TFE) has been used as a structure inducing solvent as it promotes the formation of secondary structures of peptides and proteins.[89] It has been proposed that this occurs due to the ability of TFE to compete for water molecules and form hydrogen bond networks, freeing up the peptide amide backbone and side chain residues to self-interact and fold into their predisposed secondary structures.[90]

Circular dichroism curves were collected at room temperature over wavelength scans from 185 to 250 nm for both peptides (Figure 2.12) using three solvent systems. These peptides were predicted to be α-helical like others in the class and the relative α-helicity of each peptide was calculated for all solvents systems using the equation $\frac{3000-\theta_{222}}{39000} \times 100\%$.[91] For example, in the 50% mixture of TFE in water (purple trace spectrum), FaeMK demonstrated 28% α-helicity while SRCAM 602 only showed a 16% α-helical character. This further supported our proposal that SRCAM 602 does not take the expected structure, resulting in a lack of bioactivity. On the other hand, FaeMK proved to have characteristics similar to other YGNG-motif containing peptides.

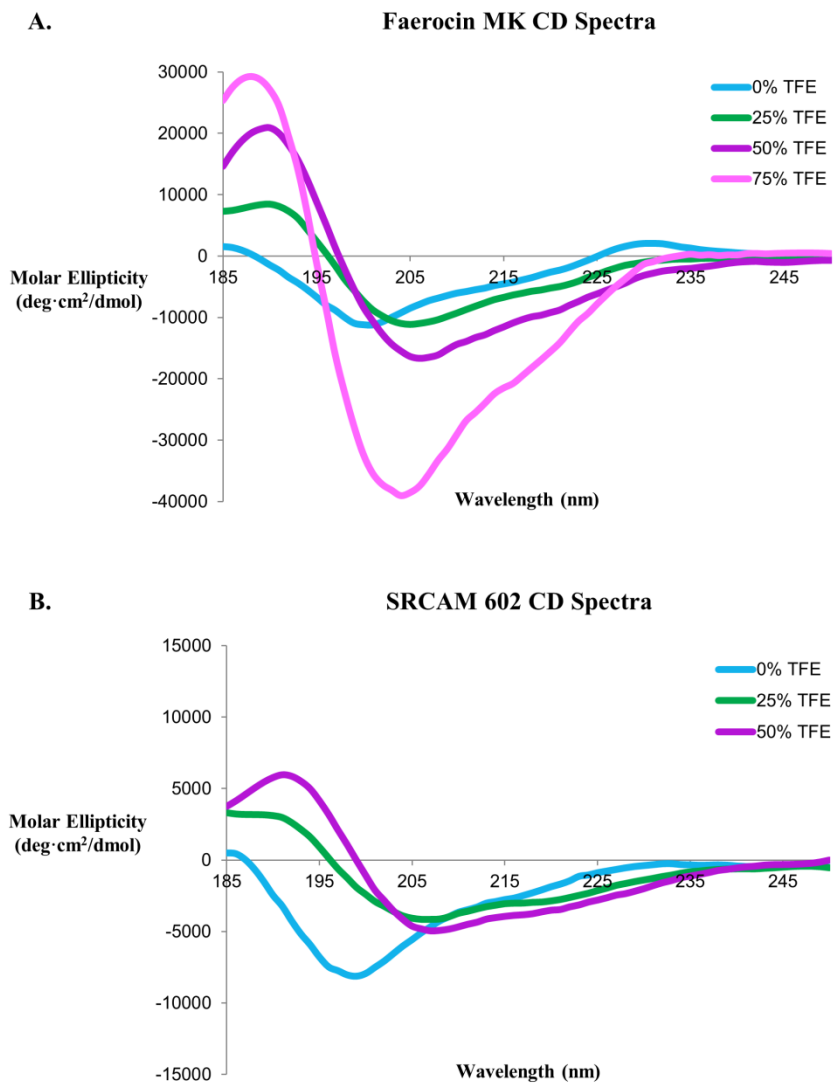


Figure 2.12 Circular dichroism spectra for (A) faerocin MK and (B) SRCAM 602

To ensure the observed trend is true, a fourth solvent system (75% TFE in water) was assayed for FaeMK. Combining this with earlier data showed an increase from 12% to 19% to 28% to 39% α-helicity at 0, 25, 50, 75% TFE in water, respectively, for FaeMK. Although some α-helical character is calculated for SRCAM 602, this does not compare with the strong and clear trend observed for FaeMK.

2.2.4 NMR structure elucidation trials

Encouraged with activity and CD data, I embarked on trying to determine the 3D structure of FaeMK using multi-dimensional NMR techniques. A solution of 1:3 TFE_{d3}:water was chosen to be the solvent system as the TFE could help induce the formation of the desired α-helix while the water would ensure any exchangeable protons, particularly those in the amide backbone, would remain non-deuterated.

The initial data collected were a one-dimensional proton NMR (^{1}H NMR), which showed some overlap of peaks, but relatively sharp signals and a very pure sample. Next, a temperature scan was performed using ^{1}H NMR from 20 ^{o}C to 50 ^{o}C. Although protein unfolding can occur at higher temperatures, bacteriocins, particularly class II bacteriocins, have been known to survive undamaged at high temperatures and highly acidic conditions.[92] It was decided that 40 ^{o}C provided the best peak distribution and further experiments were performed at this temperature. Optimizations were attempted with different solvent systems as well, including dodecylphosphocholine$_{d38}$ (DPC_{d38}) micelles; however, this resulted in broad peaks that increased the observed overlap so TFE:water was the preferred system.

A selection of two-dimensional experiments (^{1}H-^{1}H COSY, TOCSY, NOESY) were performed to acquire data required to assign the protons of FaeMK and allow for the calculation of its structure. Regrettably, only approximately 70% of the resonances could be assigned using 2D techniques due to the persistent peak overlap; the assignment table is available in Chapter 7. This led us to return to the genome and attempt to obtain isotopically labelled peptide from heterologous expression as a

fusion product. This would allow for advanced 3D NMR techniques to be employed in the quest for peptide structural elucidation.

2.2.5 Operon annotations

Based on the previously sequenced genome and with comparison to previously mentioned YGNG-motif containing bacteriocins, the FaeMK operon was annotated with aid from Dr. Marco van Belkum and is shown in Figure 2.13. The putative -35 and -10 promoter regions as well as the ribosome binding site (RBS) are underlined along the nucleotide sequence. An inverted repeat, 13 base pairs in length and shown in italics with a double underline, was observed in between the -10 promoter and the RBS. Although its function was undetermined in this study, these inverted repeats often result in a hairpin due to self-association of the DNA strand. This is often a regulation mechanism used to control the expression of genes within that operon as desired.[93]

The figure also includes the deduced peptide sequence of the two gene products present in this operon: *faeMK* and *faeI*, the structural precursor peptide and the immunity protein, respectively. The green arrow indicates the cleavage site of the faeMK parent peptide to produce the core peptide product. These annotations, particularly for the predicted immunity protein and structural peptide, were confirmed using heterologous expression as will be described in the following sections.

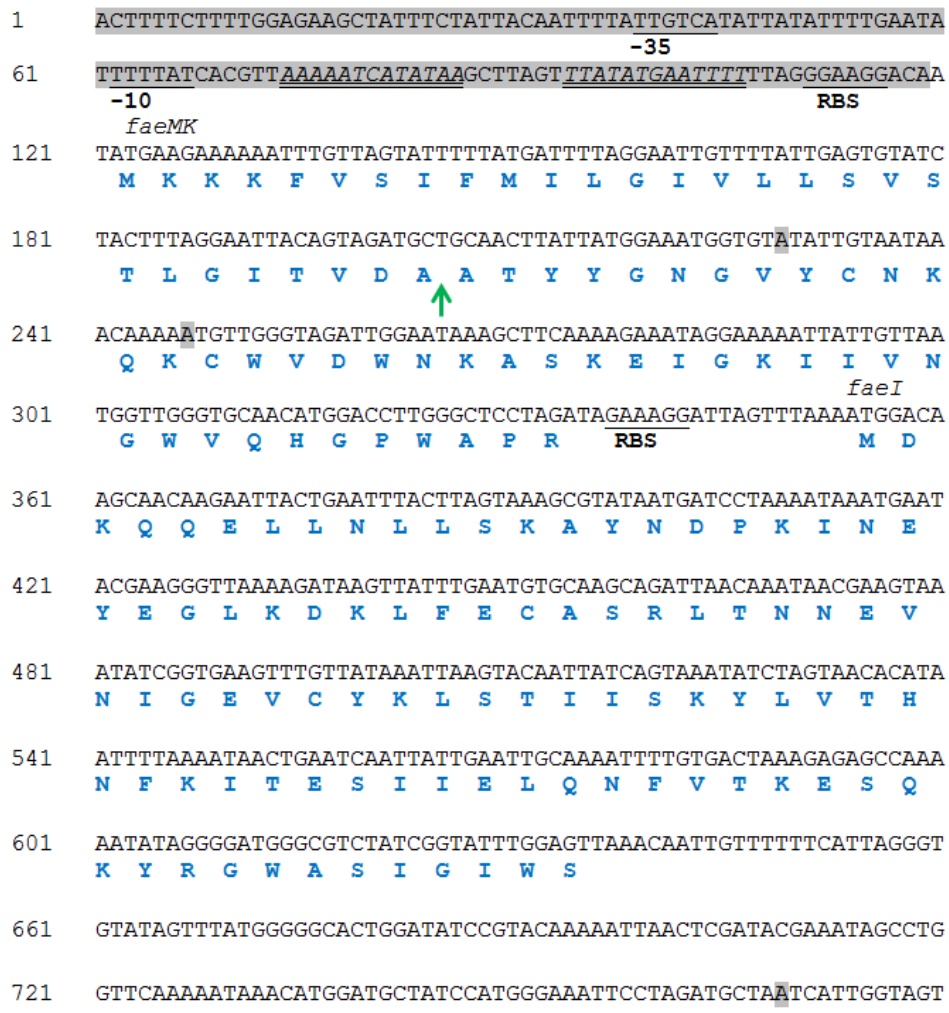


Figure 2.13 Annotated nucleotide sequence of faerocin MK

The areas highlighted in grey show the nucleic acid bases which are different from the durancin GL encoding operon, an extremely closely related YGNG-motif containing bacteriocin.[83] A single point mutation E14K differentiates FaeMK from durancin GL. Additionally, the gene encoding for durancin GL is present on a plasmid isolated from *Enterococcus durans* while *faeMKI* is present within the chromosome of its producing strain.

2.2.6 Recombinant FaeMK expression using SUMO

The next task was to express the peptide in a host organism grown in labelled medium in hopes of accessing isotopically ^{15}N and ^{13}C labelled FaeMK in a relatively economical way. With a peptide of 43 residues in length, regular synthetic approaches using solid-phase would require large amounts of pricey, isotopically labelled amino acid building blocks, making biological expression a promising alternative.

The expression system chosen for this was an *Escherichia coli* expression vector with the peptide of interest fused at the C-terminus of a SUMO (small ubiquitin-like modifier) protein. Proteins, particularly small peptides, are often expressed as fusion proteins to improve the expression and increase the stability and solubility of the peptide, or with tags to aid in their purification.[94] This method makes use of a highly specific SUMO protease which cleaves the SUMO sequence from the peptide of interest without the addition or deletion of residues onto a relatively small peptide. Also, the addition of a His_6 tag to both the SUMO sequence and the SUMO protease permits for facile affinity chromatography purifications using Ni-NTA resin.

With the aid of Dr. Marco van Belkum, the *faeMK* core peptide gene was cloned into pET SUMO vector and the plasmid was transformed into *E. coli* BL21(DE3) *via* electroporation. Bacteria were grown to desired optical density and expression of the recombinant SUMO-FaeMK fusion protein was induced using IPTG. Purification of the fusion product using Ni-NTA affinity chromatography (Figure 2.14A) yielded approximately 14 mg SUMO-FaeMK per 1 L of cultured cells. This fusion protein was treated with SUMO protease under various conditions (Figure 2.14B) to maximize the yield of FaeMK, free from its SUMO tag.

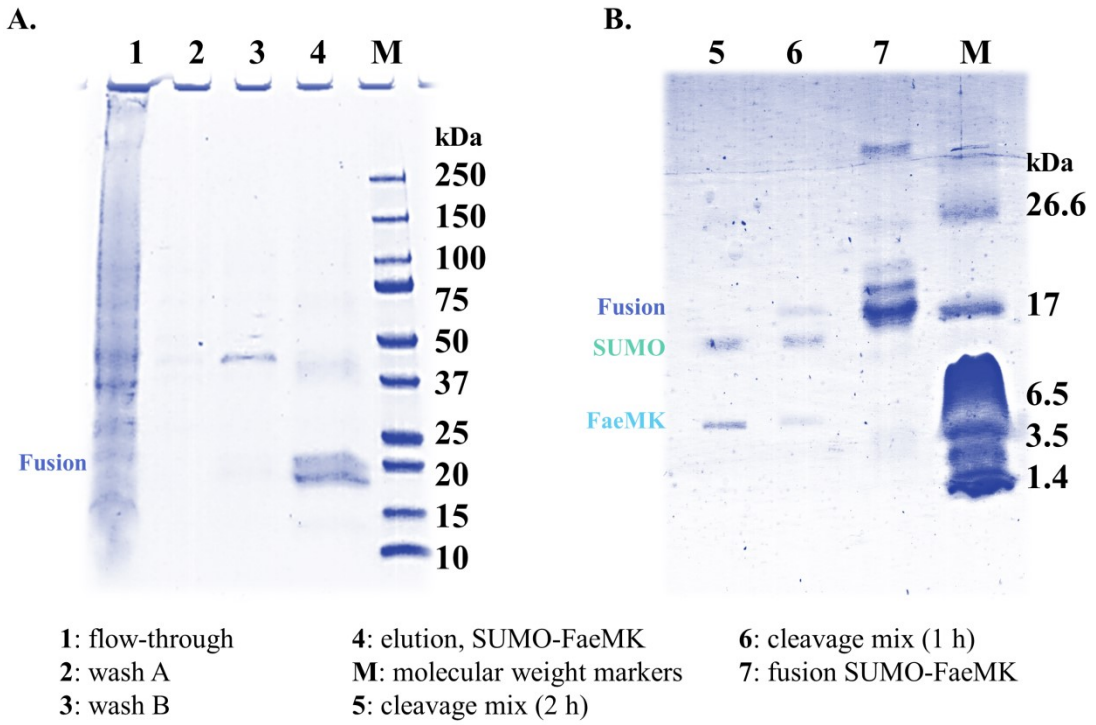


Figure 2.14 SDS-PAGE of (A) fusion protein and (B) cleavage trials

Along with SDS-PAGE monitoring, mass spectrometry analysis (MALDI-TOF) was performed to ensure the molecular weights of the fragments observed in the gels corresponded to those expected. The SUMO-FaeMK fusion protein was calculated to weigh slightly over 18 kDa, with the FaeMK just under 5 kDa and SUMO tag a bit over 13 kDa. The results of a 1 h cleavage can be found in Figure 2.15A, which shows the fusion as well as the two desired fragments. The cleavage mixture was filtered through a second Ni-NTA affinity column, this time with the unwanted peptides, the SUMO fragment and the uncleaved SUMO-FaeMK, sticking to the column and the FaeMK eluting out in the first wash (Figure 2.15B). Unfortunately, it was at this stage that issues were noticed; in addition to the expected 5 kDa peptide, two additional entities were observed in the mass spectra.

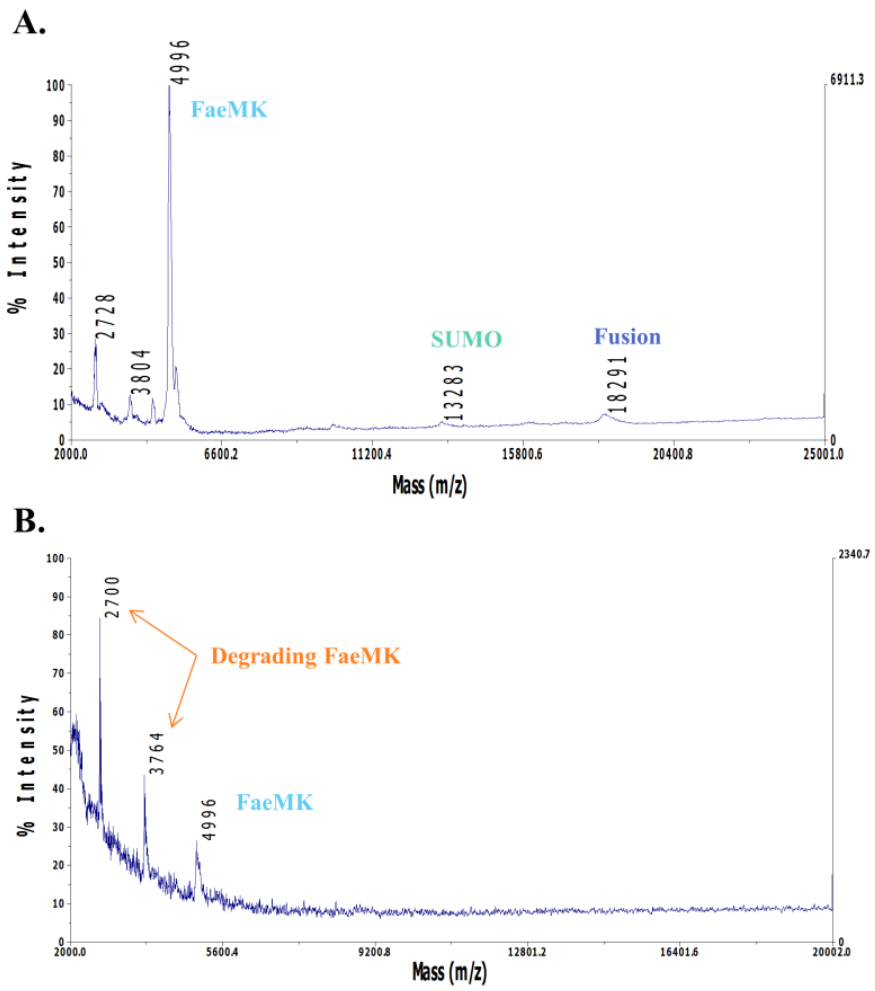


Figure 2.15 MALDI-TOF MS spectra of (A) cleavage mixture and (B) post Ni-NTA column purification of FaeMK

These entities, corresponding to 3.8 kDa and 2.7 kDa, were purified using RP-HPLC on a C18 column. Each fraction was analyzed using LCMS/MS and it was realized that the impurities are fragments of FaeMK with C-terminal deletions (Figure 2.16).

```
4966 m/z: ATYYGNGVYCNKQKCWVDWNKASKEIGKIIVNGWVQHGPWAPR
3764 m/z: ATYYGNGVYCNKQKCWVDWNKASKEIGKIIVNG
2700 m/z: ATYYGNGVYCNKQKCWVDWNKAS
```

Figure 2.16 Peptide sequences of fragmented FaeMK impurities

Several avenues for optimization of the expression and purification process were explored to try and limit these degradations. Numerous trials with varying incubation times, temperatures, acidity, buffer contents, protease inhibitors, and purification techniques were performed. Through extensive HPLC purification, a small amount of purified FaeMK could be obtained, approximately 0.2 mg per 1 L of culture. Regrettably, these yields would require an astronomical price tag for the media needed to produce enough labelled material for NMR structural elucidation.

The origin strain, *E. faecium* M3K31 was also cultured in hopes of isolating FaeMK from the producing organism. However, the production levels were either too low for MALDI-TOF detection limit or the expression was not induced, so this avenue was not pursued further.

2.2.7 Heterologous expression

Concurrently, the FaeMK operon was investigated using the cloning vector pMG36c for heterologous expression. The secretion pathway and the putative immunity protein for FaeMK were probed using two bacterial strains: *Carnobacterium maltaromaticum* UAL26 and *Enterococcous faecium* BFE 900. These two strains were chosen as they are readily made into competent cells and showed high sensitivity to FaeMK in my aforementioned bioassays. With the aid of Dr. Marco van Belkum, the *faeMKI* operon was inserted behind the constitutive promoter of pMG36c vector and transformed into the two strains. As a control, the same vector bearing no relevant gene insert was also transformed into the same two bacterial strains.

The YGNG-motif containing bacteriocins are most often secreted with the help of a dedicated ABC transporter, which also cleaves the leader sequence producing the active, mature peptide.[75] However, the operon for this peptide only contained two genes: *faeMK* and *faeI*. This led us to wonder if other bacteria, when provided with these two genes, could produce and secrete active FaeMK. An assay was conducted to test this theory using the corresponding bacterial strain, without any vector transformation.

First, cells of *C. maltaromaticum* UAL26 transformed with pMG36c vector (top) or with pMG36c-*faeMKI* (bottom) were spotted onto a hard agar plate and allowed to grow overnight (Figure 2.17). FaeMK-sensitive *C. maltarimaticum* UAL26 was used to inoculate and overlay soft media onto the plates, which were once again grown overnight. In the case of the colony containing only the empty vector, void of any FaeMK gene inserts, the presence of the spotted colony had no effect on the bacterial lawn underneath. However, in the case of the colony transformed with pMG36c-*faeMKI*, a very distinct zone of inhibition was observed. This zone was approximately 9 mm in diameter and formed a fairly even, circular shape around the colony. This indicated that these bacteria were able to produce and secrete enough FaeMK to inhibit the growth of any bacterial neighbours. The plates were kept for an additional few days and the observed inhibition zone persisted.

This experiment was repeated with *E. faecium* BFE 900 to confirm our observations with *C. maltaromaticum* UAL26. The results were similar, with an inhibition zone of 7 mm in diameter observed for the *faeMKI* containing colony; these results are found in the right panel of Figure 2.17. The ability of both these bacteria to produce active

FaeMK implies that leader cleavage and disulfide bond formation is something that can occur through generic cellular machinery or perhaps through spontaneous oxidation. The lack of ABC transporter gene, and the fact that the leader peptide of the precursor does not resemble the double Gly type leader peptide, led us to deduce that FaeMK is secreted through the general *sec*-pathway commonly used by bacteria.[18]

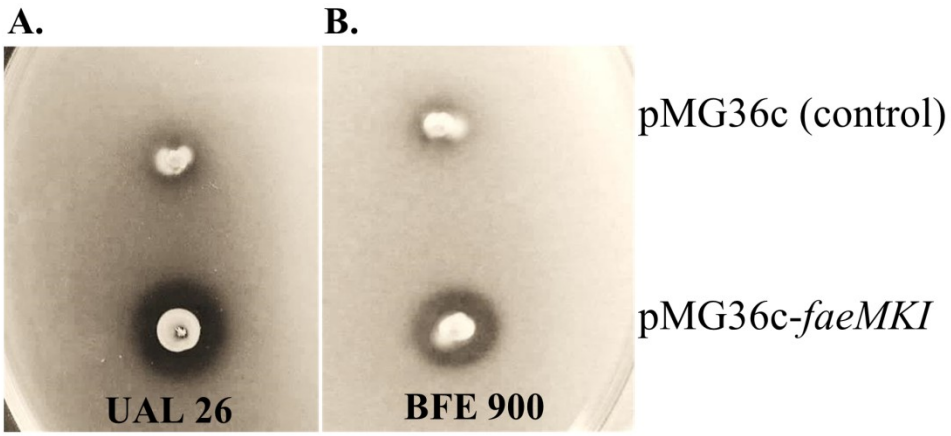


Figure 2.17 Inhibition zones observed on plates growing (A) C. maltaromaticum *UAL26 & (B)* E. faecium *BFE 900 with empty vector or faeMKI-containing vector*

The immunity protein properties were analyzed using the same set of four transformed bacterial strains. Each of the bacteria were grown separately and plated on solid media for a spot-on-lawn inhibition assay. A serial dilution of synthetically obtained FaeMK was then spotted overtop. Zones of inhibition were analyzed after 24 h and the MICs (in μM) are reported in the Table 2.5. The last column highlights the differences in sensitivity to FaeMK between cells which had the immunity protein gene *faeI* and those that did not. In the case of *E. faecium* BFE 900, the presence of *faeI* contributed to a 32-fold decrease in susceptibility to FaeMK. This trial supported

our hypothesis that the gene product of *faeI* is an immunity protein which can deter the activity of FaeMK towards its producing host.

Table 2.5 Minimum inhibitory concentrations (μM) of vector transformed cells

Vector / Indicator Strain	pMG36c	pMG36c-*faeMKI*	Difference
C. maltaromaticum UAL26	2	16	8×
E. faecium BFE 900	4	128	32×

It is important to note that the act of transforming a vector into a cell, done here through electroporation, can have a detrimental effect on the cell's membrane integrity. For that reason, although still low, the pMG36c transformed cells which survived electropermeabilization exhibit slightly elevated MICs (approximately 4× higher) than those reported previously in Table 2.4 for untransformed cells. Thus, these controls were used as the threshold for comparison between sensitivity levels.

2.3 Conclusions and future directions

FaeMK was a suspected YGNG-motif containing bacteriocin encoded within the genome of *E. faecium* M3K31, a novel bacterium isolated and sequenced in 2015. Through the synthesis, activity assays, and circular dichroism studies of this gene product, it was confirmed that this peptide belongs to the YGNG-motif containing class of bacteriocins. The activity profile, strong activity towards Listeria and other Gram-positive organisms, matches that of this bacteriocin class. The primary sequence homology and the α-helicity character observed from CD spectra also support the categorization of FaeMK into this group. Though unsuccessful in a full

structure determination using multidimensional NMR, a partial assignment was possible. The core FaeMK peptide was also heterologously expressed in *E. coli* as a recombinant SUMO fusion protein and purified using an affinity tag. Additionally, the gene products of *faeMK* and *faeI* were investigated to pinpoint the *sec*-pathway as the secretion method of this peptide and the ability of the immunity protein to confer resistance to the host, respectively. In tandem, the reported SRCAM 602 sequence was demonstrated to be inactive and did not possess significant α-helicity compared to other peptides of this class. It is proposed the cause of this to be the lack of disulfide bond, which contributes to non-random coil structure and, in turn, the function of YGNG-motif containing bacteriocins.

Though the suspected target of this bacteriocin class is the Man-PTS, there has not been extensive research done into the exact mechanism by which these peptides bind their target and exert their antimicrobial properties. Future work could be dedicated at analyzing the interactions between YGNG-motif containing bacteriocins and the outer matrix protruding portions of Man-PTS. Future studies could also perform a deeper analysis of the *E. faecium* M3K31 genome. In addition to FaeMK, the bacterium encodes two other bacteriocins of the same class. A potent cocktail from controlled gene expression of this bacterial strain could provide a bacteriocin-producing probiotic. Although not studied at length here, probiotic bacteria are attractive species to perform desired functions, such as excreting antimicrobials, as they are innately safe for human use. This makes them great additives for consumables such as food and beverage preservatives, as well as other topical applications such as creams and serums in the beauty and health industry.

Chapter 3: Activity and binding interactions of teixobactin

3.1 Background

Teixobactin (Txb) is a NRP produced by *Eleftheria terrae* that was identified as a result of a new method of culturing bacteria, the isolation chip (iChip) technology.[44] The iChip is a device by which bacteria from a certain environment can be gathered and cultivated without necessarily removing the organisms from their native environment. In the case of teixobactin discovery, a soil sample was collected and diluted across the iChip, which has 192 tiny chambers. The dilution and the size of the chambers force approximately one bacteria cell per chamber, which facilitates easier purification for DNA analyses later on. The chambers are then covered with a semipermeable membrane. The small pores of this membrane permit the passage of nutrients and minerals required for the bacteria's growth, while keeping the bacterial cell trapped within its chamber. This device is then returned to the original environment, in this case the soil of the state of Maine, and the bacteria were incubated until a colony had grown.[45]

The device was then removed and collected colonies were tested for antimicrobial resistance against *Staphylococcus aureus*. From the approximately 10,000 organisms cultivated with this device, one showed potent activity. The bacterial strain within this chamber was isolated, the genome was sequenced, and it was uncovered that the organism in question is a Gram-negative strain from a new genus closely related to *Aquabacteria.*[95] This group of organisms had previously not been reported to contain any antimicrobial peptides. However, activity-guided purification led to the isolation of a compound belonging to a new class of antimicrobials: teixobactin.

Genome sequencing confirmed the presence of two NRPS encoding genes, *txo1* and *txo2*, whose putative product matched the mass of the compound isolated from active fractions. The biosynthetic pathway was thoroughly studied and an outline can be found in Figure 3.18.[96] It was confirmed that *txo1* encodes for a NRPS with six modules, responsible for the first six amino acids as well as the methylation of the initial residue, D-Phe$_1$. The *txo2* gene encodes for a NRPS with five modules,

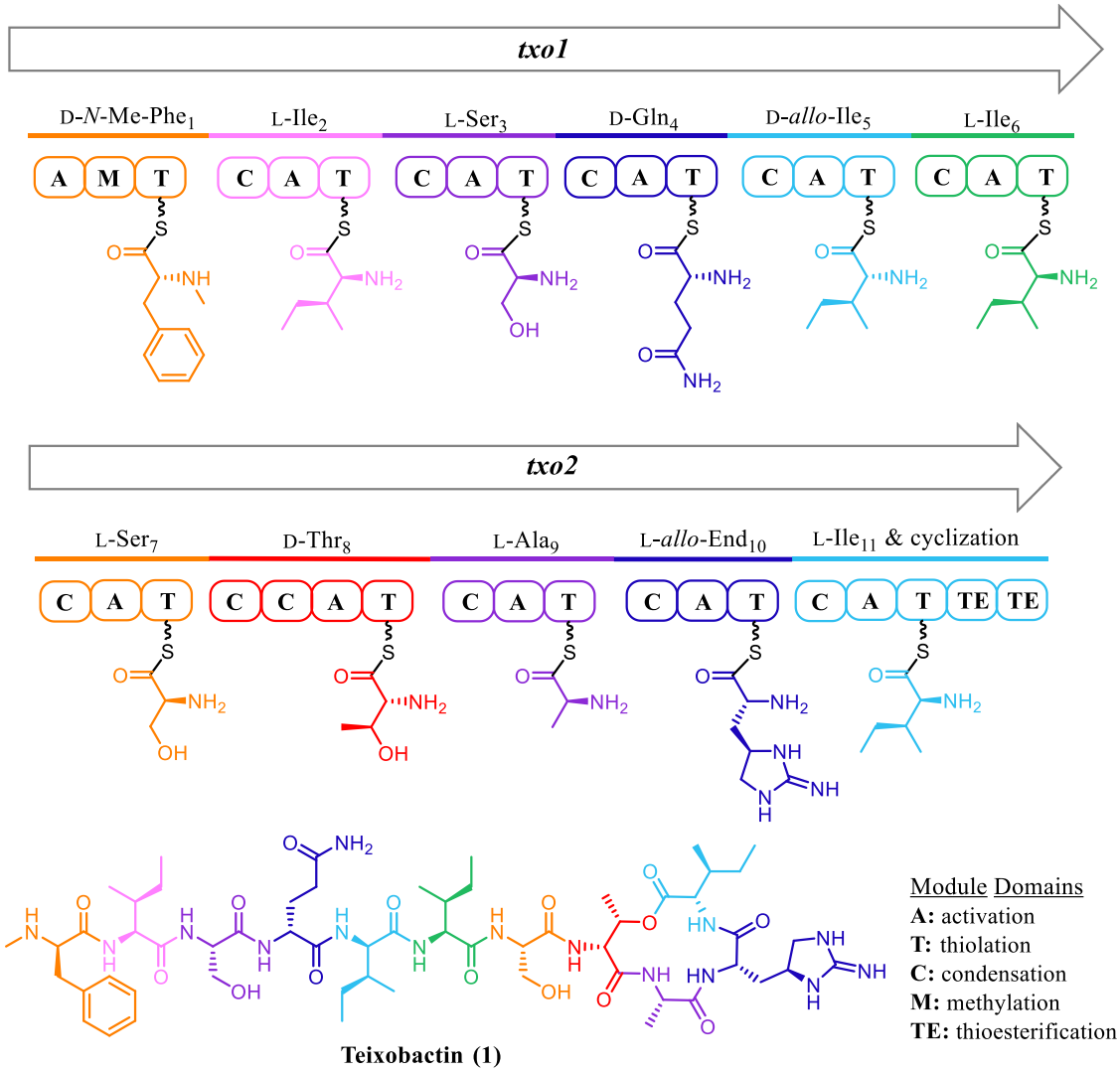


Figure 3.18 Biosynthetic pathway of teixobactin

responsible for the last five amino acids, as well as the cyclization using two tandem thioesterase domains on the terminal module. Interestingly, a second condensation domain with an unclear function exists on module 8. This module installs Thr_8, the amino acid responsible for the lactone portion of the macrocyclization.

More in-depth antimicrobial assays revealed that Txb is active against numerous Gram-positive organisms, including methicillin-resistant *S. aureus* and vancomycin-resistant enterococci such as *E. faecalis*.[45] Developing antibiotics against these two bacteria is of high priority as they are part of the ESKAPE pathogen group of concern.[97] Additionally, Txb was found to be active against *Mycobacterium tuberculosis* and *Clostridioides difficile*, bacteria that cause tuberculosis and gastrointestinal infections, respectively. Due to a potent, broad spectrum of inhibition against Gram-positive bacteria, a more general mechanism of action was proposed rather than a specific protein target. Indirect incubation studies with Txb and cellular components were performed and, from these bioassays, it was proposed that Txb binds a series of closely related lipid molecules.[45]

These lipid targets, namely lipid I, II, and III, are precursor molecules in the biosynthetic pathway of peptidoglycan (Figure 3.19). Peptidoglycan is a semipermeable polymer comprised of amino acids (peptido-) and sugars (-glycan) that surrounds bacterial membranes and is vital for their survival.[98] Through a NTP-dependant enzymatic process, bacteria are able to anabolize and catabolize their peptidoglycan layer. Briefly, an undecaprenyl phosphate molecule (C_{55}-P) is enzymatically linked to activated *N*-acetylmuramic acid pentapeptide phosphate, a compound synthesized in the cytosol by a cascade of six enzymes, to form lipid I.[99]

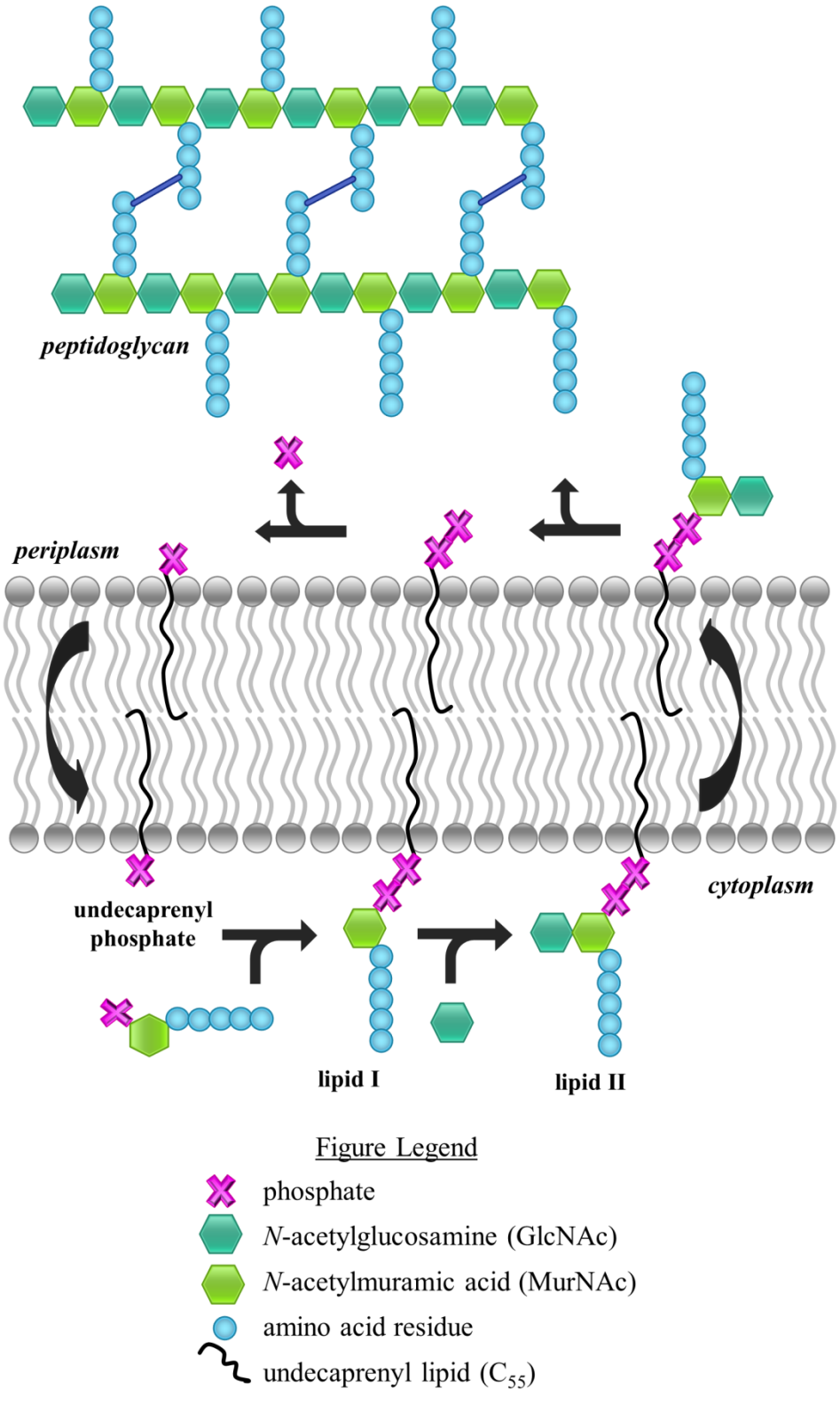


Figure 3.19 Peptidoglycan biosynthetic pathway

A translocase then attaches a UDP-activated *N*-acetylglucosamine to lipid I to yield lipid II (Figure 3.20), a molecule frequently targeted by AMPs. Lipid II (LII) is translocated across the lipid bilayer membrane to the periplasm, a relocation believed to occur with the aid of a flippase.[99] Here a series of oligomerization enzymes, glycosyltransferases for the disaccharides and transpeptidases (also known as penicillin-binding proteins) for the amino acids, work together to form the peptidoglycan.[100] Carboxypeptidases also further functionalize the growing polymer according to the uniqueness of each bacterial species, often removing the terminal Ala.[98] The peptide stem has been found to vary,[98] especially as new resistant organisms are uncovered. Nevertheless, in general, Gram-positive (**2a**) and Gram-negative (**2b**) organisms have similar LII molecules, save a difference at position 3: a Lys or *meso*-diaminopimelic acid (DAP) residue, respectively.

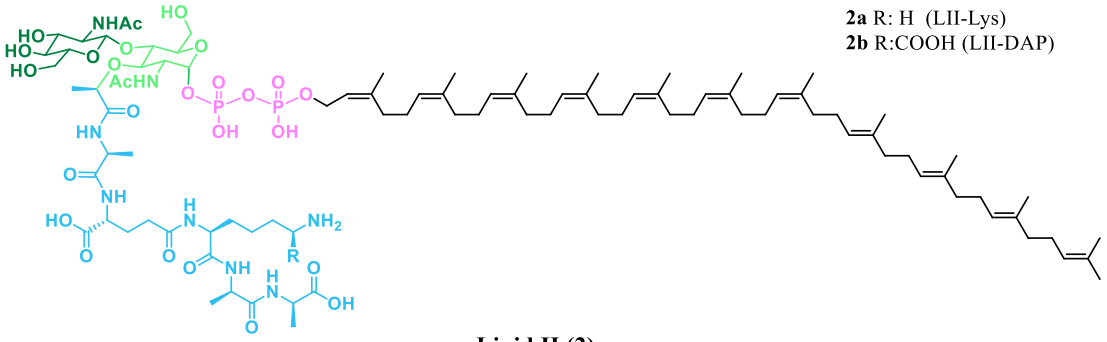


Figure 3.20 Full structure of lipid II, a peptidoglycan precursor

Once the undecaprenyl pyrophosphate has unloaded its peptidoglycan cargo, the terminal phosphate group is removed and C_{55}-P is translocated back into the cytoplasm side of the membrane where it can be recycled.[100] Interruptions to this assembly line can be detrimental to cells, making this system a common target for AMPs. Molecules that bind and inhibit the function of these lipid precursors, such as

Txb, are attractive compounds to investigate as resistance is believed less likely to develop,[45] due to cells' reliance on these lipids to construct its peptidoglycan layer.

Interest in Txb also arose from its unique structural features, indicating the discovery of a new class of antibiotics not previously reported (Figure 3.21). Txb is an eleven-residue peptide with a 13-membered lactone macrocycle formed between residues Thr_8 and Ile_{11} at the C-terminal end (green). The peptide also has a methylated N-terminus (green), effectively protecting both termini of the peptide. Additionally, four of the amino acids possess D configuration (blue) and the unnatural amino acid *allo*-enduracididine (*allo*-End) is present at position 10 (purple).

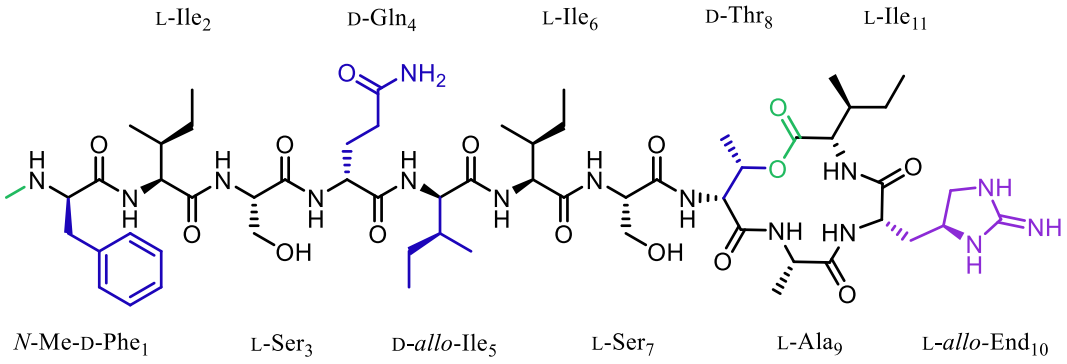
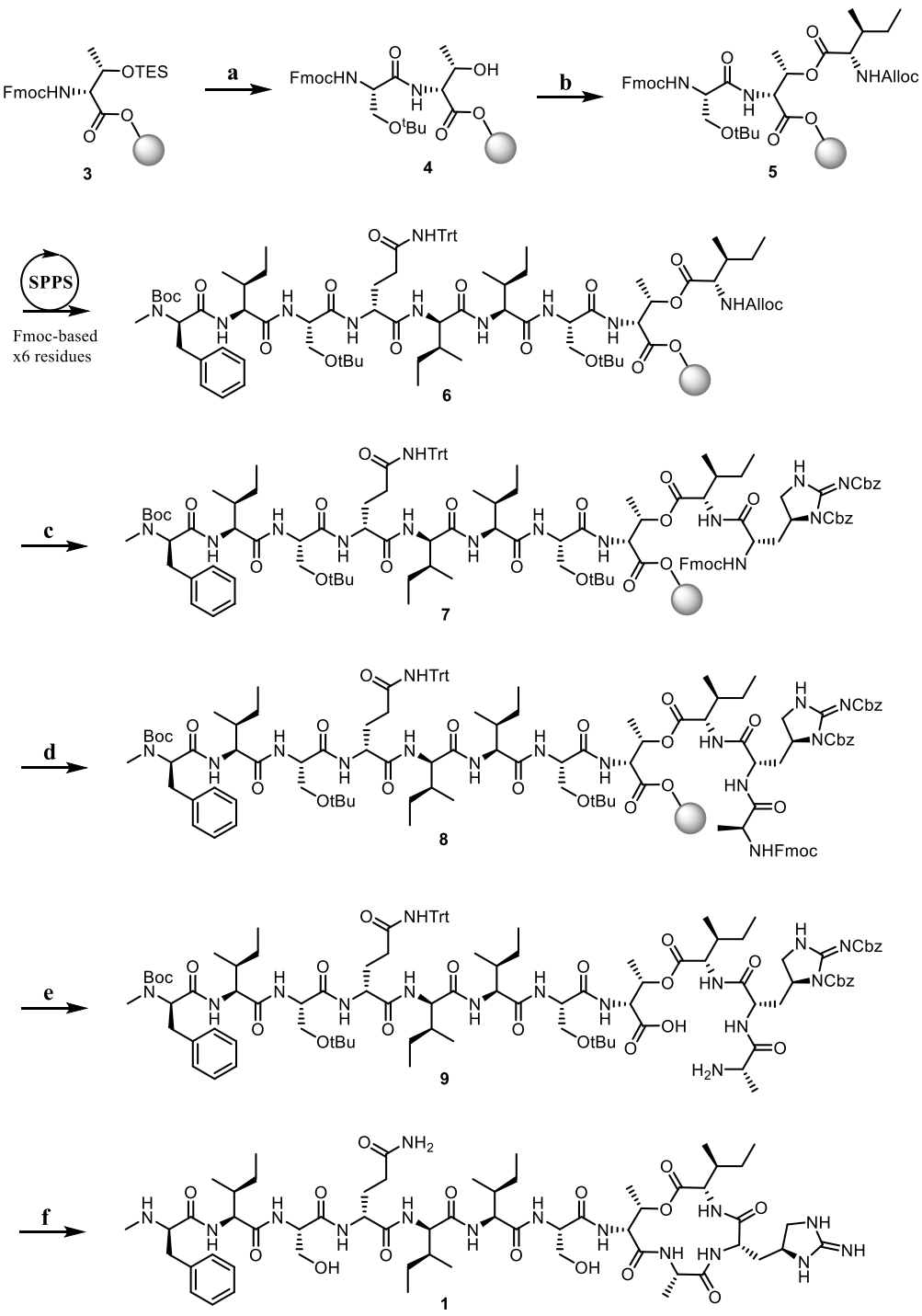


Figure 3.21 Unique structural features of teixobactin

In just five years since its discovery, 14 different synthetic routes have been developed and hundreds of Tbx analogues have been synthesized.[101] The first report of native Txb total synthesis used a linear approach, with the macrocycle bridge residue, Thr_8, attached on resin and the remaining residues linked using stepwise fluorenylmethoxycarbonyl-based solid-phase peptide synthesis (Fmoc-SPPS) as outlined in Scheme 3.1.[102] A seven-step synthesis was used to access a di-carboxybenxyl (Cbz) protected L-*allo*-End and incorporated it at the appropriate



a: 1. 10% piperidine, DMF 2. TBAF, AcOH, DCM:THF 3. Fmoc-Ser(tBu)-OH, PyBOP, NMM, DMF
b: Alloc-Ile-OH, DIC, DMAP, DCM:THF
c: 1. $Pd(PPh_3)_4$, $PhSiH_3$ 2. Fmoc-End$(Cbz)_2$-OH, HATU, HOAt, DIPEA, DMF
d: 1. 10% piperidine, DMF 2. Fmoc-Ala-OH, PyBOP, NMM, DMF
e: 1. 10% piperidine, DMF 2. 1% TFA, DCM
f: 1. DMTMM-BF_4, DIPEA, DMF 2. TFA:thioanisole:TfOH:*m*-cresol (70:10:12:8), 3.5 mM HCl

Scheme 3.1 First total synthesis of teixobactin

position in the SPPS process. The peptide was removed from resin using mildly acidic conditions to preserve protecting groups. Lactamization between the carboxylic acid of Thr_8 and amine of Ala_9 was achieved using 4-(4,6-dimethoxy-1,3,5-triazin-2-yl)-4-methylmorpholinium tetrafluoroborate (DMTMM-BF_4) followed by global deprotection to afford native teixobactin (**1**).

A year previous, a simpler Txb analogue was synthesized which bypassed the need for *allo*-End by replacing this residue with Arg, a linear analogue of End.[103] It was discovered that although stereochemistry is important, the identity of the residue at position 10 is less so – activity is maintained with amino acids such as Arg or Lys, which are still cationic and hydrophobic. Multifaceted studies of various analogues have methodically uncovered the roles of different residues in Txb (Figure 3.22). Results of Ala scans and other mutation studies are summarized as low, mild, and high sensitivity to modification in green, blue, pink, respectively.[101] Additionally, studies also looked at the replacement of the oxygen in the lactone to nitrogen and formed active azateixobactin, a lactam Txb derivative.[104] Other backbone modifications, such as amide methylation, were not very well tolerated and some entirely abolished activity.[105] It was recently elucidated that the backbone amines highlighted in orange facilitate hydrogen bonding with the pyrophosphate of LII.[106]

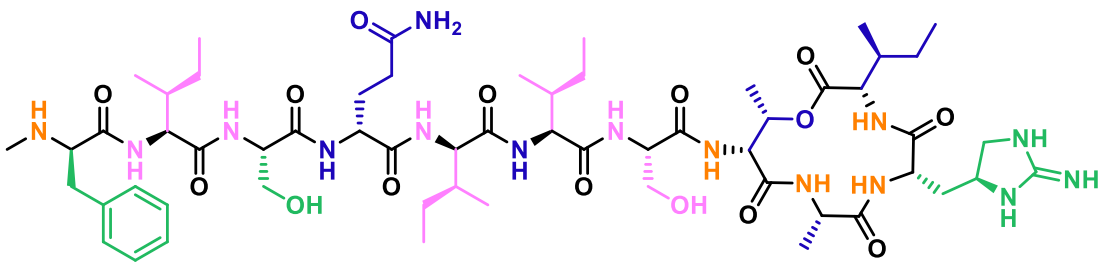


Figure 3.22 Modification sensitivity of teixobactin structure

At the time this work was initiated, little was known about the interactions of this new and exciting peptide, Txb, with its lipid ligands. It was believed to interact with lipid I, II and III but the exact binding conformations and the mechanism of action were not yet revealed. This project was undertaken in hopes of elucidating the 3D solution-phase NMR structure of teixobactin bound to its lipid targets and understanding more about their interaction by developing a series of Txb and LII analogues. Along the way, other properties such as binding affinities and antimicrobial synergistic activities were investigated.

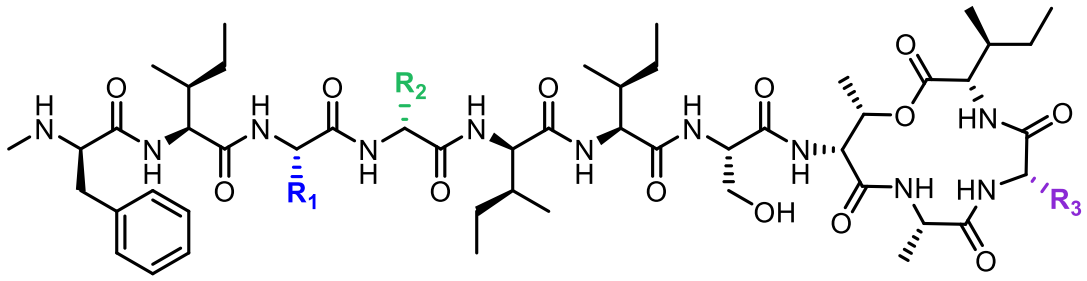
3.2 Results and discussion

3.2.1 Syntheses of necessary analogues

3.2.1.1 Teixobactin analogues

A total of four Tbx analogues (**10-13**) were synthesized during the course of this study (Figure 3.32). Due to ease of synthesis and lower financial burden and labour, analogues without enduracididine, $TxbArg_{10}$ (**10**) and $TxbLys_{10}$ (**11**), were synthesized. These showed comparable bioactivities to native Txb (**1**) and the first round of assays was performed with these three compounds.

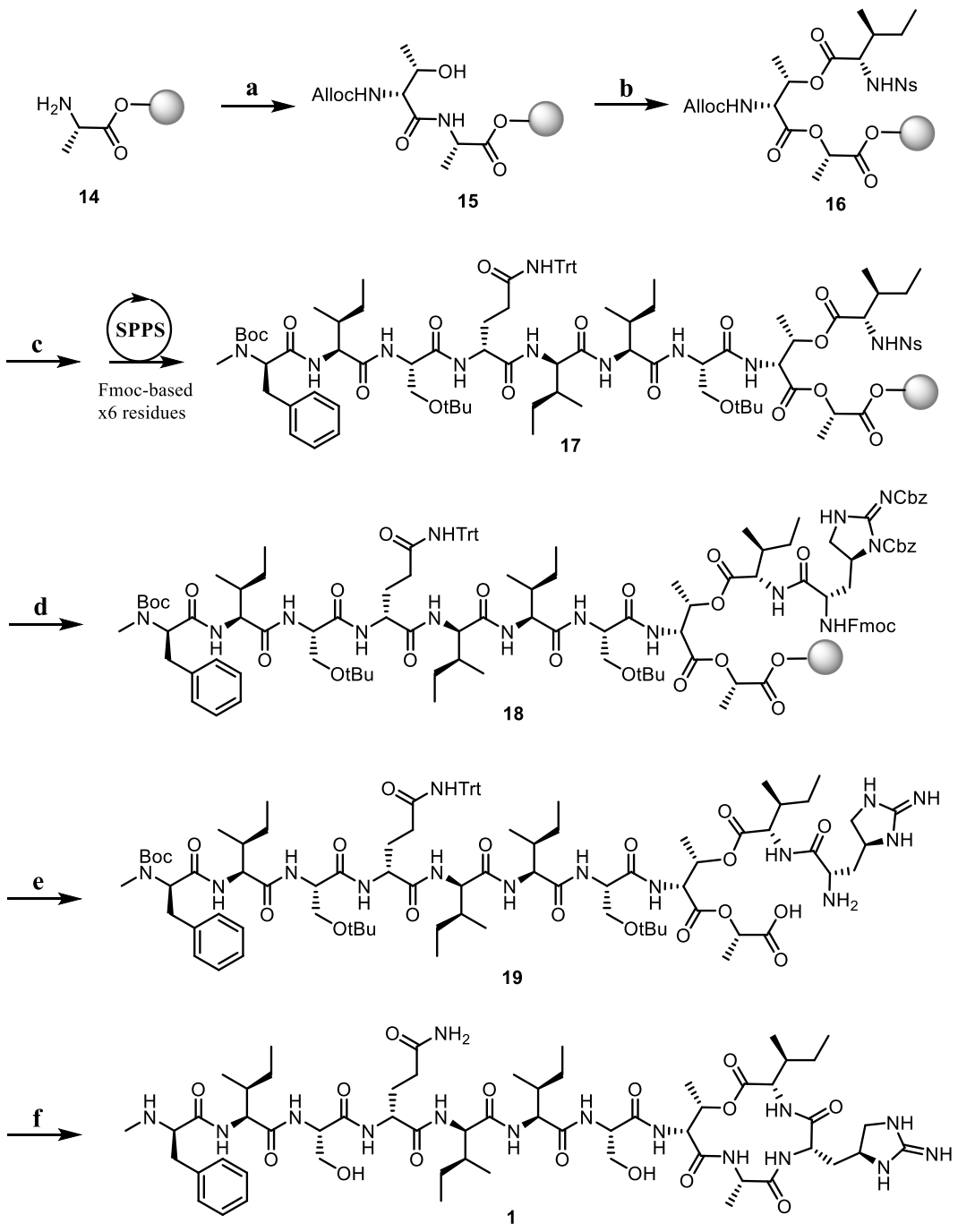
As illustrated above, $_{L}$-Ser_3 and, to a lesser extent, $_{D}$-Glu_4 are positions quite amenable to modifications and substitutions. To deal with solubility issues encountered, two additional Lys analogues, $TxbLys_3Lys_{10}$ (**12**) and $TxbLys_4Lys_{10}$ (**13**) were synthesized as well, with both modifications maintaining the original configuration.



1 Txb **R**$_1$:Ser **R**$_2$:D-Gln **R**$_3$:*allo*-End
10 TxbArg$_{10}$ **R**$_1$:Ser **R**$_2$:D-Gln **R**$_3$:Arg
11 TxbLys$_{10}$ **R**$_1$:Ser **R**$_2$:D-Gln **R**$_3$:Lys
12 TxbLys$_3$Lys$_{10}$ **R**$_1$:Lys **R**$_2$:D-Gln **R**$_3$:Lys
13 TxbLys$_4$Lys$_{10}$ **R**$_1$:Ser **R**$_2$:D-Lys **R**$_3$:Lys

Figure 3.23 Teixobactin and its analogues studied

The synthesis of teixobactin analogues was undertaken by Dr. Antoine Henninot following a slightly different method than the one illustrated previously (Scheme 3.2). This approach also relied on Fmoc-SPPS strategy; however, in this case the residue loaded on resin was Ala$_9$. The chain was extended to the orthogonally protected Thr$_8$ and Ile$_{11}$ (**16**) on resin, where the Fmoc-protected amine of Thr was substituted for allyloxycarbonyl (Alloc) and amine of Ile was protected by nitrobenzenesulfonyl (nosyl, Ns). The Alloc protection on Thr was removed and the chain was extended by six residues using automated Fmoc-based SPPS to complete the linear N-terminal portion. Next, the Ile amine was deprotected and di-Cbz protected, Fmoc-L-*allo*-End (or the desired cationic residue) was attached. Mildly acidic conditions were employed to cleave the peptide from resin while still maintaining orthogonal sidechain protecting groups. The macrocyclization was achieved by coupling the amine of End$_{10}$ to the carboxylic acid of Ala$_9$ using 1-ethyl-3-(3-dimethylaminopropyl)carbodiimide (EDCI). Global deprotection with trifluoroacetic acid (TFA) yielded the fully deprotected Txb and its analogues.

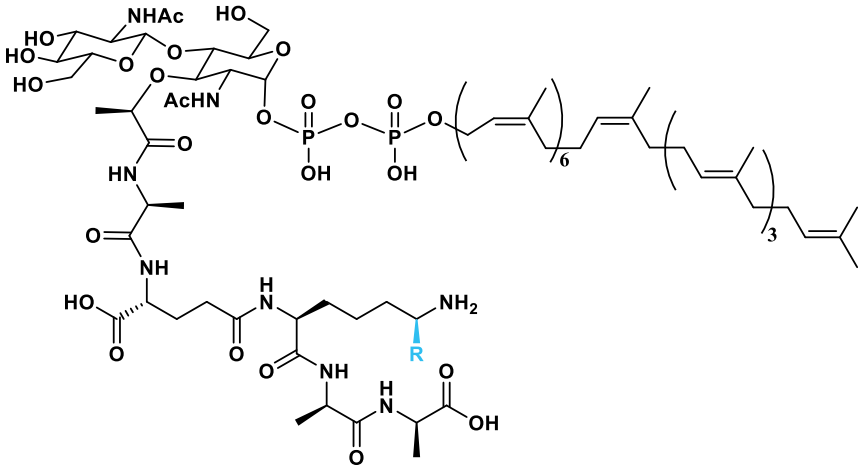


a: 1. Fmoc-D-Thr-OH, HBTU, Et_3N, DCM:DMF 2. 25% piperidine, DMF 3. AllocCl, Et_3N, DCM:DMF
b: 1. Fmoc-Ile-OH, DIC, DMAP, DCM:NMP 2. 25% piperidine, DMF 3. NsCl, collidine, NMP
c: 1. $Pd(PPh_3)_4$, $PhSiH_3$, DCM
d: 1. 5% thiophenol, DMF 2. Fmoc-L-*allo*-End(Cbz_2)-OH, HATU, HOAt, DMF
e: 1. 10% piperidine, DMF 2. 30% HFIP, DCM 3. H_2, 10% Pd/C, MeOH, AcOH
f: 1. EDCI, HOAt, DCM:DMF 2. TFA:TIPS (98:2)

Scheme 3.2 Synthesis of teixobactin and its analogues

3.2.1.2 Lipid II analogues

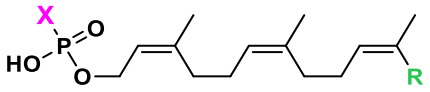
Two major forms of LII exist in nature (Figure 3.24), those found in most Gram-positive organisms (**2a**) and a majority of Gram-negative bacteria (**2b**). These two analogues were synthesized by Isaac Antwi following literature protocols.[107]



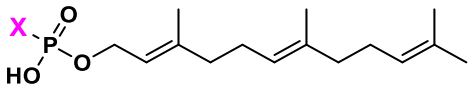
2a LII-Lys (Gram-positive) **R:** H
2b LII-DAP (Gram-negative) **R:**CO_2H

Figure 3.24 Lipid II found in Gram-positive and Gram-negative organisms

The documented promiscuity of Txb for lipid I and III prompted us to investigate the minimum moiety required for Txb to recognize and bind its target. A series of truncated analogues (**20 – 26**) were synthesized for this purpose (Figure 3.25).



20 **R**:CH_3 **X**:OH
22 **R**:$C_{41}H_{89}$ **X**:OH
23 **R**:CH_3 **X**:OPO_3H_2
25 **R**:$C_{41}H_{89}$ **X**:OPO_3H_2
26 **R**:CH_3 **X**:OPO_3H-GlcNAc

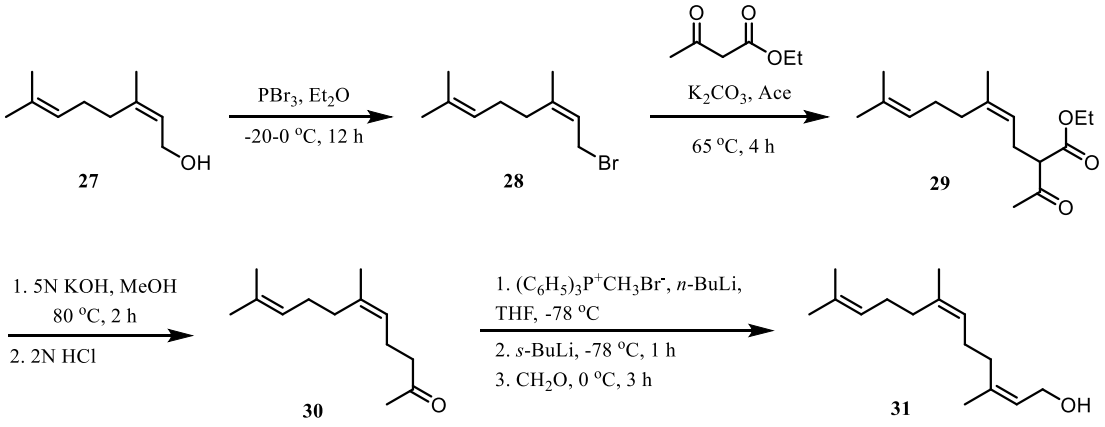


21 **X**:OH

24 **X**:OPO_3H_2

Figure 3.25 Lipid II analogues synthesized for this study

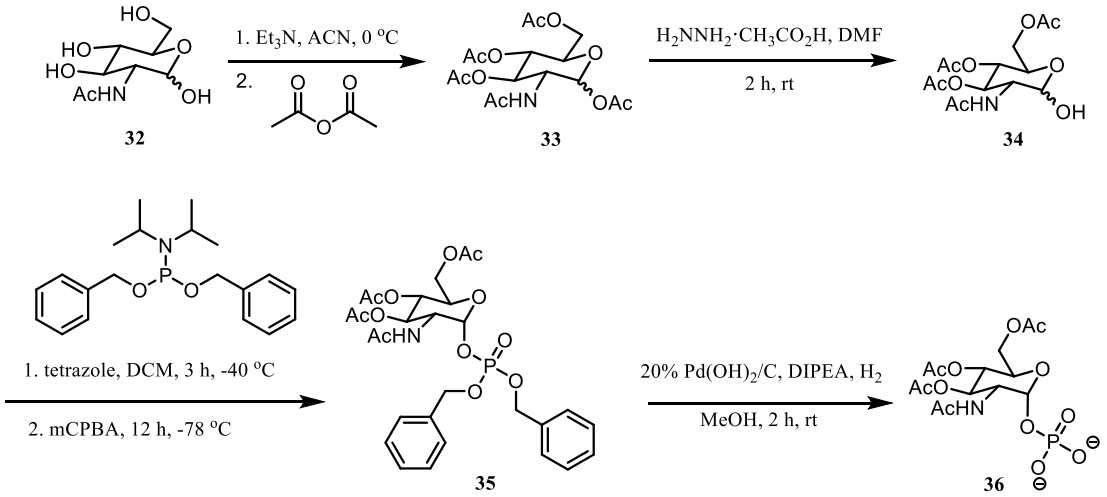
The undecaprenyl tail of LII consists of 55 carbons or 11 isoprene units, the first 7 of these units possess a *cis* configuration across the double bond and the terminal 4 are *trans*. In addition to the natural undecaprenyl analogues (synthesized with undecaprenol extracted by Dr. Stephen Cochrane and coworkers), shorter C_{15} lipid tails were chosen for these analyses. These shorter, farnesyl lipids can circumvent the issues that arise from LII solubility in aqueous solvent systems[107] as well as the relatively slower tumbling observed for large molecules in NMR experiments.[108] Although the *Z*,*Z*-farnesol is needed to mimic the natural LII, the relative price of this isomer is substantial. Commercial asking price for *Z*,*Z*-farnesol was approximately 1,631 CAD per 1 mg (now discontinued) while the *trans* isomer, *E*,*E*-farnesol, is only 42 CAD per 1 g. Trials were first done with the *E*,*E* isomer while the *Z*,*Z* isomer (**31**) was synthesized from *Z*-nerol (**27**) over four steps (Scheme 3.3).[109]



Scheme 3.3 Synthesis of Z,Z-farnesol

The first set of truncated LII analogues, the lipid monophosphates (**20 – 22**), were obtained by activating the corresponding alcohol with trichloroacetonitrile and adding tetrabutylammonium phosphate.[110] The lipid pyrophosphates (**23 – 25**) were obtained

by converting the desired lipid alcohol to a chloro-intermediate and reacting this with tris(tetrabutylammonium) hydrogen pyrophosphate.[111] The last truncated analogue (**26**), containing a monosaccharide, was obtained by reacting the corresponding monophosphate lipid (**20**) with a deprotected sugar-phosphate synthesized according to an established literature protocol (Scheme 3.4).[112]



Scheme 3.4 Synthesis of protected N*-acetylglucosamine phosphate*

This collection of LII analogues was subjected to binding affinity assays against Txb and the four Txb analogues discussed in the previous section.

3.2.2 Binding assays using isothermal titration calorimetry

Isothermal titration calorimetry (ITC) measures the heat released or absorbed by a system upon the titration of one compound into a solution of a second. ITC is an advantageous technique to study binding parameters between two interacting molecules as it can determine multiple thermodynamic properties, such as affinity, enthalpy, and stoichiometry, from just one experiment.[113] ITC was previously used by

our group to study another AMP which targets LII, tridecaptin A_1, and a majority of instrument parameters and buffer conditions were based on this experimental design.[114]

Experiments were conducted at room temperature, in aqueous Tris buffer at pH 6.5, with LII titrated into a solution containing Txb. Each titration was performed in triplicate and a series of controls was included. Raw data were collected over a period of approximately 3 h and parameters were calculated by Origin 7 software. An example of data collected can be seen in Figure 3.26; the titration curve of LII-Lys with Txb (A) and the buffer solution with Txb (B) as a control, which does not show any significant temperature changes throughout.

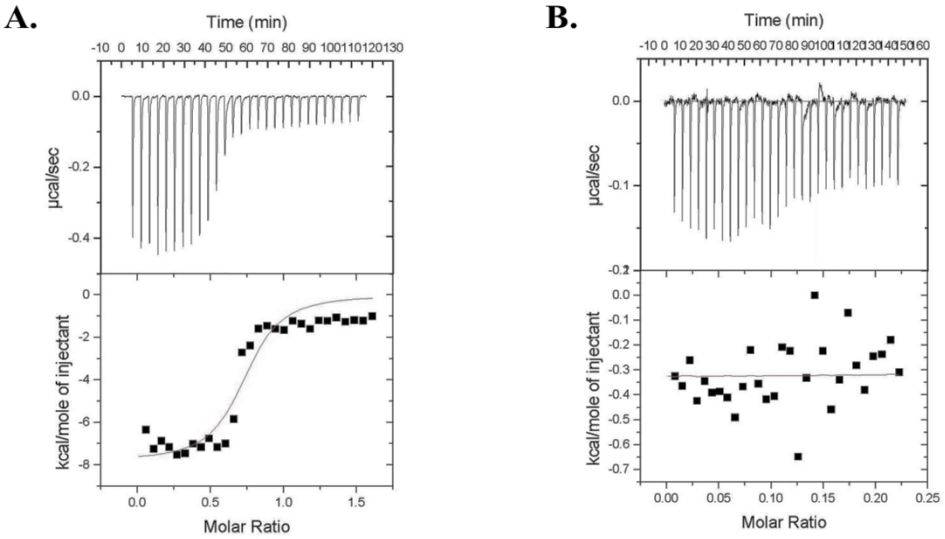


Figure 3.26 ITC data for (A) LII-Txb binding curve and (B) a negative control

The binding stoichiometry, or molar ratio, calculated varied from 1 to 2 with an average of approximately 1.6:1 Txb:LII. This indicates that the binding is not an even 1:1 ratio, rather each LII molecule can bind more than one Txb peptide. These results are consistent with previous data indicating two Txb molecules bind one lipid.[45]

Dissociation constants (K_d) for the ITC experiments of Txb with various truncated lipid analogues, as well as native Gram-positive and Gram-negative LII were calculated and are summarized in Table 3.6.

Table 3.6 Dissociation constants (μM) for Txb with LII and truncated analogues

Lipid	Txb (1)
Z,Z-C_{15}-P (**20**)	11.26
E,E-C_{15}-P (**21**)	–
C_{55}-P (**22**)	7.69
Z,Z-C_{15}-PP (**23**)	0.92
E,E-C_{15}-PP (**24**)	1.04
C_{55}-PP (**25**)	0.82
Z,Z-C_{15}-PP-GlcNAc (**26**)	0.89

–, no thermal changes detected

The *Z,Z*-lipid monophosphates **20** and **22** had comparatively high K_d values while the *E,E*-lipid monophosphate **21** had no significant thermal changes, with measurements comparable to that of the negative buffer controls. While Txb could bind to the *Z,Z* lipids (the configuration observed in the native LII), the monophosphate did not facilitate as strong of an interaction as the pyrophosphates. All three pyrophosphate lipids, **23 – 25**, showed K_d values within the same range, regardless of the unsaturated lipid tail configuration. This coordination with the pyrophosphate of LII has been documented in other LII-binding AMPs, with the peptide macrocycles forming a 'pyrophosphate cage' through hydrogen bonding of the amide backbone.[115] The addition of a monosaccharide (**26**) resulted in a K_d of 0.89 μM. This is very similar to the same lipid without the sugar (**23**), which has a K_d of 0.92 μM. The presence or absence of sugar did not play a significant role in improving binding, indicating that Txb is likely drawn to this pyrophosphate region of LII. This assumption was further

supported by the fact that neither the configuration nor length of the lipid tails had a significant effect on binding, as long as two phosphate groups were present. Inversely, with the monophosphate lipid ITC trials, the configuration and length played a more significant role in binding affinity. These findings are in line with reports of Txb binding other peptidoglycan precursors such as lipid I and lipid III, which both possess an undecaprenyl pyrophosphate.

The next set of lipid analogues to be analyzed was the native LII variants, LII-Lys and LII-DAP. Dissociation constants (K_d) for LII-Lys and LII-DAP with Txb and the four Txb analogues were calculated in μM and are summarized in Table 3.7.

Table 3.7 Dissociation constants (μM) of Txb analogues and native LII

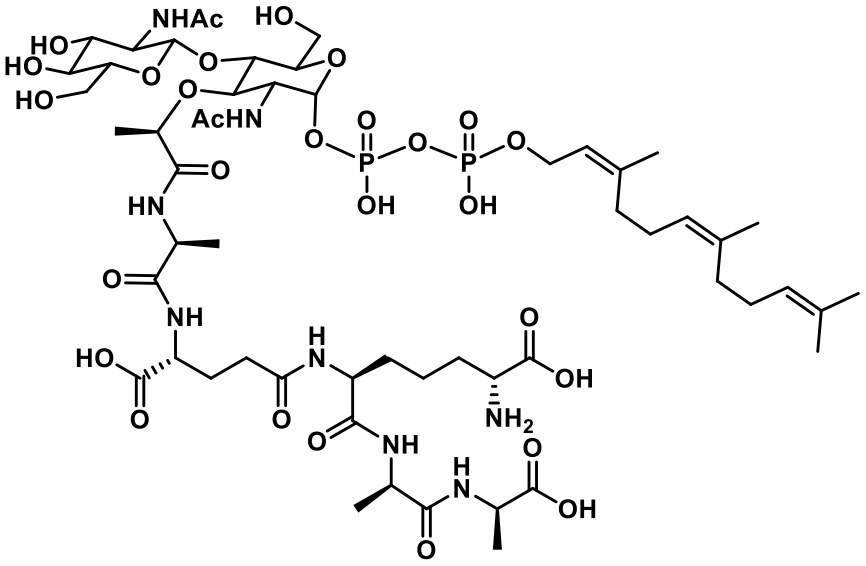
Compound	Txb (**1**)	$TxbArg_{10}$ (**10**)	$TxbLys_{10}$ (**11**)	$TxbLys_3Lys_{10}$ (**12**)	$TxbLys_4Lys_{10}$ (**13**)
LII-Lys (**2a**)	0.43	4.13	0.60	63.01	37.86
LII-DAP (**2b**)	1.36	0.06	0.90	1.68	2.30

As standard practice, stock solutions of peptides are made, degassed, and then diluted out for each ITC trial. Regrettably, the aqueous buffer used for previous LII work caused some solubility issues, particularly with $TxbArg_{10}$. As $TxbLys_{10}$ showed comparable bindings to native Txb and was more soluble, two additional analogues **12** and **13** were made that use Lys at position 10 rather than *allo*-End and also add an additional cationic residue to aid with solubility. These two analogues were indeed more soluble; however, they showed less potent binding towards both LII-DAP and, more significantly, LII-Lys. These results were a bit surprising as substitutions at these positions with Lys have shown comparable activities against Gram-positive organisms and were used here for that reason.[116]

Overall, the binding affinities for LII-DAP, the Gram-negative variant, proved to be comparable and in some cases stronger than those for LII-Lys, the Gram-positive variant. Txb had been shown to be active against Gram-positive organisms but not in Gram-negative bacteria. This was later explored through synergistic bioactivity assays of the five Txb compounds. As LII-Lys and Txb had strong binding in these ITC experiments, the next step was to uncover more details about their binding conformations using NMR assisted structure elucidation.

3.2.3 NMR structure elucidation attempts

The original structure elucidation of Txb was determined in 100% $DMSO_{d6}$[45] while most NMR studies with LII were performed in aqueous phosphate buffers.[114] A few considerations had to be made for these experiments. Rather than the undecaprenyl native LII, the truncated *Z,Z*-C_{15} LII-DAP (**37** in Figure 3.27) was chosen for NMR studies to prevent issues arising from lipid insolubility or larger macromolecule tumbling speeds. Although peptide NMR is ideally conducted in acidic conditions to limit the hydrogen solvent exchange of the amide backbone and sidechains,[117] the prolonged exposure of the LII in acidic conditions resulted in pyrophosphate cleavage. This was observed in trials with LII outlined in Chapter 4 and characterized; the pyrophosphate was fragmented into the sugar-phosphate and lipid phosphate segments at pH values lower than 6.5. For this reason, the pH of the solvent systems tested was kept between 5.8 and 6.8.



***Z,Z*-C_{15} LII-DAP (37)**

Figure 3.27 LII variant conducive to NMR experiments

Txb had already shown some solubility issues in the ITC aqueous buffer solutions. Of note, the concentrations required for ITC are 10- or 100-fold less than required for NMR; solutions of 100 μM lipid to 10 μM Txb were used for the titrations. To obtain good structural data in NMR experiments, a minimum of 1 mM of peptide sample is preferable. A larger complex, such as LII bound Txb, would ideally be a more concentrated sample. Additionally, as it is believed binding occurs in a 2:1 ratio, the intention was to reach this concentration with at least 2 mM:1 mM Txb:LII.

The first solvent system attempted was 1:1 water:DMSO pH 6.5, in the hopes that both LII and Txb would be soluble and stable in their respective solvents. Unfortunately, this system resulted in Txb precipitating out of the solution. Increasing the DMSO concentration, up to 1:3 water:DMSO, was attempted with no success.

Trials varying the temperature and pH were also not fruitful. A selection of other unsuccessful solvent systems and conditions tested are listed in Table 3.8.

Table 3.8 Solvent system tested for Txb:LII NMR sample preparation

D_2O:H_2O (1:9)	$DMSO_{d6}$	CD_3CN	CD_3OH	C_3D_7OH	$CDCl_3$	pH	Temp (^oC)
1	1					6.5	25
1	3					6.5	25
1	1					6.5	40
1	1					5.8	25
1		1				6.5	25
	1		1			6.5	25
	1			1		6.5	25
	1				1	6.5	25
1	1		2			6.8	40

During these ongoing tribulations, a collaborative effort of another research group reported the structure of LII and Txb using solid-state NMR (ssNMR).[118] This approach used ^{13}C and ^{15}N labelled Txb and *Z,Z*-C_{15} LII-DAP (truncated lipid analogue, although not the Gram-positive LII variant). Solution-phase NMR experiments of Txb were performed in DPC_{d38} micelles rather than simpler solvent systems. The addition of LII to Txb in solution caused large aggregates to form, resulting in too many missing peaks to allow for full structure elucidation from the NMR data. Trials with various DMSO concentrations (as high as 1:4 water:DMSO) were also attempted with similar results. Thus, the use of ssNMR was required for the complete structure elucidation of the complex which could not be achieved in solution NMR. They also proposed that the binding ratio is indeed greater than one, likely at 2:1 Txb:LII.

Subsequently, a series of studies investigating Txb and Txb analogues from Nowick and coworkers indicated Txb aggregates as amyloids composed of antiparallel β-sheets.[119] Figure 3.28, recreated here from the deposited Protein Data Bank accession number 6E00, shows two $TxbLys_{10}$ molecules aligned as antiparallel β-sheets, with Phe_1 (purple) and the macrocycle (blue, Lys_{10} shown) coming together. The amide backbones of these residues have now been shown to hydrogen bond with the pyrophosphate of LII through the use of ssNMR.[106]

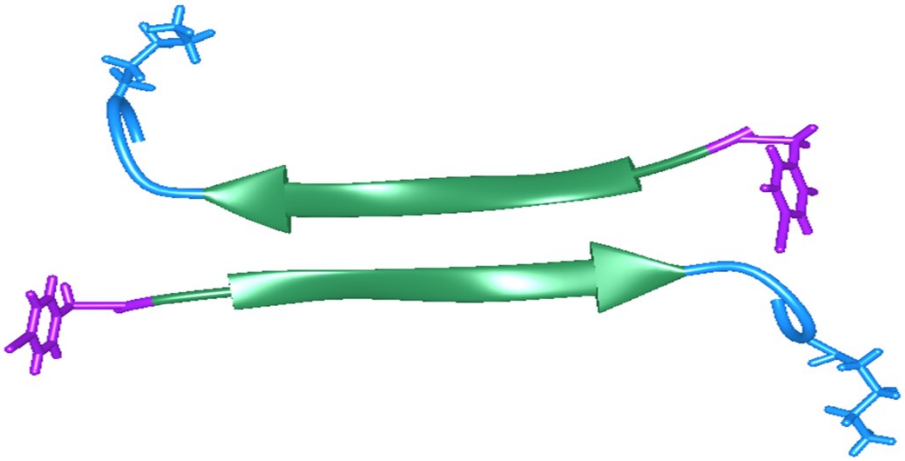
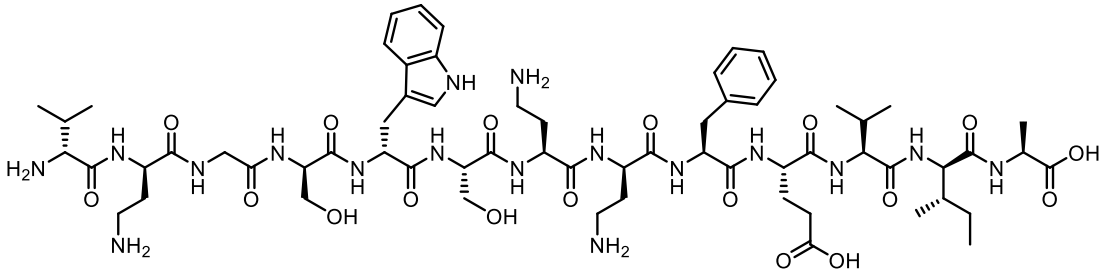


Figure 3.28 Antiparallel β-sheets of $TxbLys_{10}$ amyloids

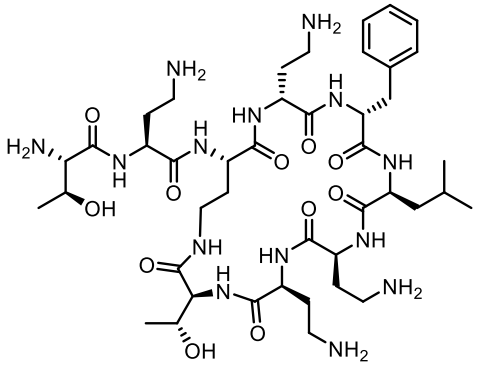
3.2.4 Synergistic assays to target Gram-negative organisms

The results of the ITC experiments indicated that Txb can bind both LII-Lys and LII-DAP with comparable affinities, K_d values of 0.43 and 1.36 μM, respectively, even though activity has only been documented for Txb against Gram-positive (LII-Lys) organisms. This led to the hypothesis that if Txb is able to penetrate the outer membrane of Gram-negative organisms and reach LII, it could also be an effective antimicrobial against Gram-negative pathogens.

Rather than using mechanical methods or harsher reagents to disturb the outer membrane, a more chemical approach was chosen. Previous work with truncated antimicrobial lipopeptides has shown that if the lipid tail is removed, the peptides are no longer antimicrobial. Rather, they have only perturbative effects on the membrane but cannot kill the bacteria cells. Two such membrane-disruptive peptides, unacylated tridecaptin A_1 (H-TriA_1)[120] and commercially available polymyxin B nonapeptide (PMBN)[121] were selected for synergistic bioassays (Figure 3.29). Both of these peptides are known to be highly potent against Gram-negative organisms; however, if the lipid tail is removed or not added during synthesis, the peptides lose their potency.



unacylated tridecaptin, H-TriA_1 (38)



polymyxin B nonapeptide, PMBN (39)

Figure 3.29 Structures of outer membrane-disruptive peptides H-TriA_1 and PMBN

H-TriA_1 was synthesized using Fmoc-based SPPS and purified following methodology used previously for tridecaptin and tridecaptin analogues.[120] PMBN is a commercially available compound. The peptide's identity and purity were confirmed using LCMS and HRMS. It was then used for these studies without further purification.

Synergistic effects of Txb with either H-TriA_1 or PMBN were determined by collecting and comparing MICs using a microplate dilution assay. Synergy was assessed using a 96-well plate, with a serial dilution of Txb performed in the X-axis and a serial dilution of the membrane-disruptive peptide in the Y-axis. Bacterial growth was monitored by optical density (OD_{600}) and assessed after the appropriate incubation time for each strain. MICs were determined as the lowest concentration at which no growth was observed, as compared to a negative control.

Table 3.9 shows a summary of synergistic assay results obtained for teixobactin and its four analogues, along with vancomycin as a positive control, with membrane-disruptive peptides against a series of Gram-negative bacteria. The MIC values reported are an average of triplicates.

Table 3.9 Comprehensive MICs from synergistic assays with H-TriA_1 and PBMN

Organism	Antibiotic	MIC at [H-TriA_1] in µg/mL					MMD for H-TriA_1	MIC at [PMBN] in µg/mL					MMD for PMBN
		0	1.56	3.13	6.25	12.5		0	3.25	7.50	15.0	30.0	
E coli ATCC 25922	Txb	22.5	11.3	1.41	1.41	0.70	32×	22.5	11.3	5.63	5.63	5.63	4×
	$TxbArg_{10}$	90	45	22.5	22.5	22.5	4×	90	90	90	90	45	2×
	$TxbLys_{10}$	45	45	22.5	22.5	22.5	2×	45	45	45	45	22.5	2×
	$TxbLys_3Lys_{10}$	22.5	11.3	5.63	2.81	1.41	16×	22.5	11.3	11.3	11.3	2.81	8×
	$TxbLys_4Lys_{10}$	22.5	11.3	11.3	11.3	11.3	2×	22.5	22.5	22.5	22.5	22.5	0
	Vancomycin	200	50	25	25	25	8×	200	100	50	25	25	8×
E. coli DH5α	Txb	22.5	11.3	5.63	5.63	2.81	8×	22.5	11.3	5.63	5.63	5.63	4×
	$TxbArg_{10}$	45	22.5	11.3	11.3	11.3	8×	45	22.5	22.5	22.5	22.5	2×
	$TxbLys_{10}$	45	11.3	11.3	11.3	5.63	8×	45	22.5	22.5	22.5	11.3	4×
	$TxbLys_3Lys_{10}$	22.5	11.3	5.63	2.81	1.41	16×	22.5	11.3	11.3	11.3	11.3	2×
	$TxbLys_4Lys_{10}$	22.5	5.63	5.63	5.63	2.81	8×	22.5	5.63	5.63	5.63	5.63	4×
	Vancomycin	50	25	12.5	6.25	3.13	16×	50	12.5	6.25	6.25	3.13	16×
S. enterica serovar Typhimurium ATCC 13311	Txb	22.5	11.3	5.63	1.41	0.18	125×	45	11.3	11.3	11.3	11.3	4×
	$TxbArg_{10}$	90	22.5	11.2	1.41	0.09	1024×	90	45	22.5	22.5	22.5	4×
	$TxbLys_{10}$	45	22.5	5.63	1.41	0.09	512×	45	22.5	22.5	22.5	22.5	2×
	$TxbLys_3Lys_{10}$	22.5	11.3	5.63	1.41	0.70	32×	22.5	11.3	11.3	11.3	11.3	2×
	$TxbLys_4Lys_{10}$	22.5	22.5	22.5	11.3	1.41	16×	22.5	5.63	5.63	5.63	5.63	4×
	Vancomycin	100	50	12.5	25	6.25	16×	100	12.5	12.5	12.5	12.5	8×

Organism	Antibiotic	MIC at [H-TriA$_1$] in µg/mL					MMD for H-TriA$_1$	MIC at [PMBN] in µg/mL					MMD for PMBN
		0	1.56	3.13	6.25	12.5		0	3.25	7.50	15.0	30.0	
S. enterica serovar Typhimurium ATCC 23564	Txb	45	22.5	22.5	11.3	11.3	4×	45	5.63	5.63	5.63	5.63	8×
	TxbArg$_{10}$	–	90	45	22.5	11.3	> 8×	–	22.5	22.5	22.5	22.5	> 4×
	TxbLys$_{10}$	90	90	45	22.5	22.5	4×	90	22.5	22.5	22.5	22.5	4×
	TxbLys$_3$Lys$_{10}$	22.5	11.3	11.3	5.63	5.63	4×	22.5	11.3	11.3	11.3	11.3	2×
	TxbLys$_4$Lys$_{10}$	45	22.5	11.3	11.3	5.63	8×	45	5.63	5.63	5.63	5.63	8×
	Vancomycin	100	50	50	12.5	6.25	16×	100	25	12.5	12.5	12.5	>8×
K. pneumoniae ATCC 13883	Txb	45	22.5	11.3	5.63	5.63	8×	45	11.3	11.3	11.3	11.3	4×
	TxbArg$_{10}$	45	5.63	1.41	1.41	1.41	32×	45	22.5	22.5	22.5	22.5	2×
	TxbLys$_{10}$	45	2.81	2.81	2.81	2.81	16×	45	45	45	45	45	-
	TxbLys$_3$Lys$_{10}$	45	22.5	11.3	5.63	1.41	32×	45	11.3	11.3	2.81	2.81	16×
	TxbLys$_4$Lys$_{10}$	45	2.81	2.81	1.41	1.41	4×	45	45	45	45	45	-
	Vancomycin	100	50	12.5	1.56	0.78	128×	100	25	25	25	25	4×
A. baumannii ATCC 19606	Txb	22.5	22.5	22.5	22.5	22.5	0	22.5	22.5	22.5	22.5	22.5	0
	Vancomycin	–	–	–	–	100	> 0	–	100	100	100	50	> 2×
A. baumannii ATCC BAA-1605	Txb	22.5	22.5	22.5	22.5	22.5	0	22.5	22.5	22.5	22.5	22.5	0
	Vancomycin	200	200	200	200	100	2×	200	100	100	100	100	2×

–, no inhibition observed at the highest soluble concentration of peptide tested; MMD: maximum MIC difference caused by the addition of a membrane-disruptive peptide (H-TriA$_1$ in blue, PMBN in green)

Initial trials used *Escherichia coli* DH5α and *Salmonella enterica* serovar Typhimurium ATCC 23564, both known to be linked with foodborne infections. With promising results, other strains of these two bacteria were also assayed, *E. coli* ATCC 25922 and *S. enterica* serovar Typhimurium ATCC 13311. The scope was then extended to include Gram-negative organisms that are part of the ESKAPE pathogens of concern, such as *Klebsiella pneumoniae* and *Acinetobacter baumannii*.[5c]

Every strain except for *A. baumannii* proved to be more sensitive towards Txb in synergy with either membrane-disruptive peptide. *S. enterica* ATCC 13311 was the most sensitive, particularly in combination with H-TriA_1. Txb and analogue TxbArg_{10} showed a 125- and 1024-fold decrease in MIC, respectively, when 12 μg/mL of H-TriA_1 was added. Intriguingly, the presence of H-TriA_1 was overall a more significant contributor to lowering the MIC over PMBN. The most pronounced improvement in MIC noted with PMBN resulted in a modest 8-fold enhancement. Additionally, H-TriA_1 appeared to have a proportional decrease in MIC with increasing H-TriA_1 concentrations, while PMBN appeared to have a more consistent effect regardless if concentration of the membrane-disruptive peptide was increased. The more potent synergy of H-TriA_1 could be attributed to the relatively different mechanisms of action. Previous work indicates that H-TriA_1 binds with its target, precursor of the peptidoglycan layer LII, in a concentration-dependant manner.[120] On the other hand, PBMN on a cell surface appears to reach a saturation point, a maximum concentration after which more PBMN molecules do not bind to the same bacteria cell, thus its synergistic effect would be limited to this concentration.[122]

Overall, synergistic studies showed that in most cases Txb can be made active against Gram-negative organisms at MICs that are pharmacologically relevant. In particular, H-TriA_1 can dramatically enhance Txb activity against foodborne pathogen *S. enterica* as well as be quite potent against *K. pneumoniae*, an ESKAPE pathogen of priority for the World Health Organization.

3.3 Conclusions and future directions

This work was aimed at better understanding the binding interaction between Txb and LII through thermodynamic measurements, spectroscopic structure elucidation, and bioactivity assays. A series of analogues of both peptide and lipid were synthesized, with an emphasis on modifications to improve compound solubility, for these investigations. Using these analogues, a number of ITC experiments were conducted to determine the binding parameters of Txb against its lipid target. Native Txb was titrated against a selection of truncated LII analogues to determine the minimum moiety required for recognition. The binding affinities obtained from ITC indicated that Txb is able to recognize the pyrophosphate head group of the LII. Although the peptide prefers the *Z,Z*-lipid isomer, the pyrophosphate was the more significant contributor to binding and Txb bound to *E,E*-lipid pyrophosphate isomer with comparable magnitude. Based on averages from all ITC trials, it appears that Txb binds LII in a 2:1 molar ratio.

Additionally, the LII variants found in most Gram-positive organisms (LII-Lys) and in Gram-negative organisms (LII-DAP) both bound Txb and its analogues with comparable affinities. This prompted a study using membrane-disruptive peptides,

unacylated tridecaptin A_1 (H-TriA_1) and polymyxin B nonapeptide (PBMN), as synergistic additives. Results showed that the MICs could be dramatically reduced with H-TriA_1 and moderately reduced with PBMN for a panel of tested Gram-negative organisms, including *S. enterica* and *K. pneumoniae*, a foodborne disease-causing bacteria and a pathogen deemed by the WHO as one of concern, respectively.

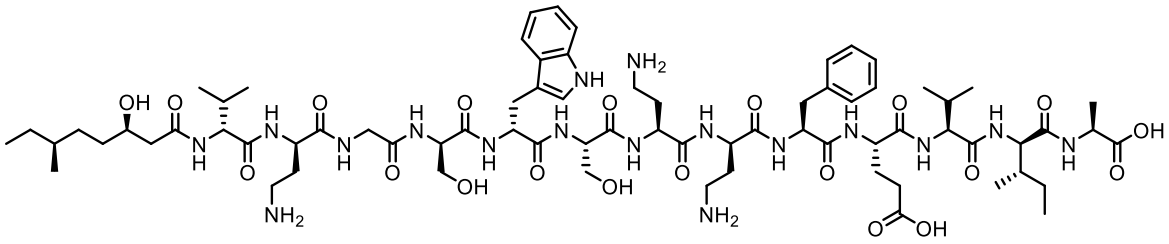
Regrettably, due to solubility issues caused by Txb precipitating out of solution, a NMR solvent system that would keep both Txb and LII stable was not found. In the meantime, other groups had bypassed this solubility problem by moving their spectroscopy studies to solid state NMR. Through this means, the structure of LII bound Txb as well as interesting scaffolds formed by Txb antiparallel β-sheets stacking into amyloid fibrils were elucidated.[119]

The issues encountered with Txb aggregation or precipitation out of solution prompted our collaborator, Dr. Antoine Henninot, and coworkers to develop new analysis methods with practical industry applications.[123] Since the conclusion of this work, they have designed a high-throughput assay to determine aqueous solubilities for proteins and peptides which have a high propensity to form gels or aggregates.[123] This assay can be used to efficiently determine the conditions which instigate gelation through rapid viscosity measurement screening. It has been used as an additional tool for assessing peptides in the pharmaceutical workflow for issues related to solubility.

Chapter 4: NMR structure investigations of tridecaptin A_1

4.1 Background

Tridecaptins are a class of linear non-ribosomally synthesized lipopeptides found to be active against Gram-negative bacteria. Interest in their studies arises from their potent activity towards Gram-negative ESKAPE pathogens such as *S. aureus*, *K. pneumoniae*, *P. aeruginosa* as well as foodborne pathogens *Salmonella*, *E. coli*, and *Campylobacter jejuni*. This group of AMPs encompasses several 13-residue lipopeptides, which differ in their amino acid composition or the lipid tail. Tridecaptin A_1 ($TriA_1$), the most studied member of this class, was first identified in 1978 from a *Bacillus polymyxa* strain[124] and the full structural assignment, including the configuration of the lipid tail, was more recently determined by our group (Figure 4.30).[125] Similar to the NRP teixobactin discussed previously, $TriA_1$ contains a few noncanonical features such as three diaminobutyric acid (Dab) residues, six D-amino acids, and a hydroxylated lipid tail at the N-terminus.



Tridecaptin A_1 (40)

Figure 4.30 Structure of native tridecaptin A_1

The putative biosynthetic gene cluster of tridecaptin A_1 was proposed[125] following the genome analysis of *Paenibacillus terrae* NRRL B-30644 by our efforts.[79] The gene cluster contains five proteins: a putative thioesterase (TriA), two ABC transporter proteins involved in secretion (TriB and TriC), and two NRPSs (TriD and TriE).[125] The role of the thioesterase (TriA) is not yet clear, although similar enzymes have been found in NRPS biosynthetic gene clusters of surfactin and bacitracin AMPs.[126] In these pathways, thioesterases are believed to ensure the correct lipid has been installed within the NRPS's thiolation domain. If an incorrect lipid is detected, the enzymes cleave the lipid and regenerate the misacylated NRPS.

The two NRPSs, TriD and TriE, are responsible for installing the first ten and the last three amino acids, respectively (Figure 4.31). Along with the usual activation, condensation, and thiolation domains, an epimerization domain is found within several modules. This domain assists in the isomerization of desired residues into the D-amino acids. Interestingly, perhaps as a result of vestigial evolution, domain 3 which installs a Gly residue also contains an epimerization domain, even though this amino acid is achiral. The acylation of the N-terminal residue with (3*R*,6*S*)-3-hydroxy-6-methyloctanoic acid is believed to be an enzyme-assisted process though this has not been currently investigated to date.

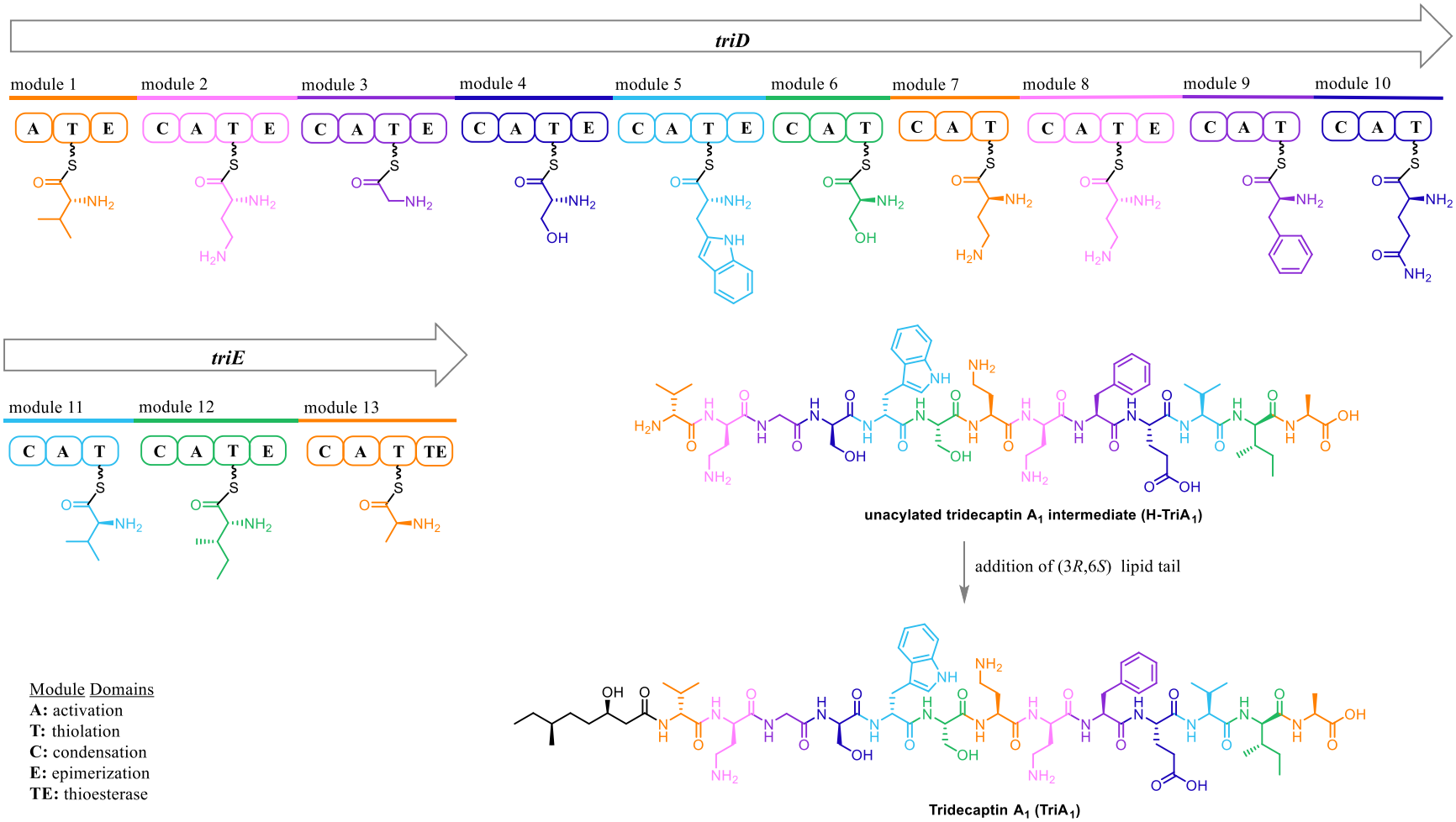
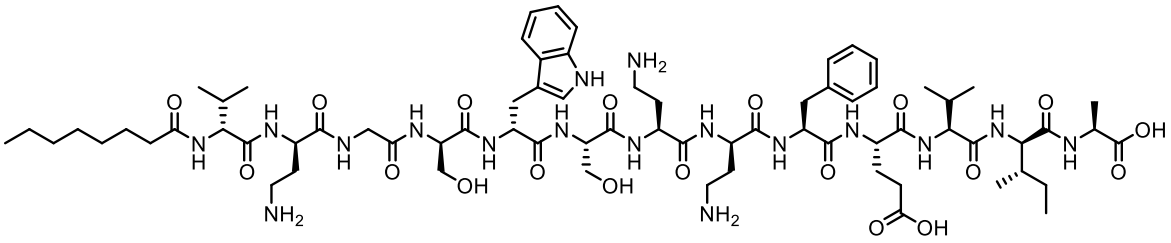


Figure 4.31 NRPSs involved in the biosynthesis of TriA$_1$

The initial study from our group involved the synthesis of all four possible stereoisomers of $TriA_1$'s lipid tail, derived from the two stereogenic centres at the 3-hydroxyl and 6-methyl positions, to confirm the configuration of the peptide isolated from *P. tarrae*. Subsequent studies further explored tolerable substitutions of this N-terminal acylation through the synthesis and activity testing of various analogues.[127] This yielded H-$TriA_1$, the unacylated peptide used in the previous chapter as an outer membrane-disrupting peptide, which has a greater than 30-fold decrease in activity compared to native $TriA_1$. Although a lipid tail appears vital for activity, its substituents and exact configuration seem to play a lesser role. An unsubstituted eight-carbon acyl chain was used to synthesize octyl-tridecaptin A_1 (Oct-$TriA_1$, **41**), a simpler and cheaper analogue with comparable activity to that of native $TriA_1$ (Figure 4.32).



Octyl-Tridecaptin A_1 (41)

Figure 4.32 Structure of a simpler $TriA_1$ derivative, Oct-$TriA_1$

During the same study, our group also showed $TriA_1$ is capable of disrupting the proton motive force across a cell's membrane which eventually leads to cell death.[114] The essential nature of the lipid tail noted in activity assays supported the proposal that $TriA_1$ targets the cell membrane, an interaction greatly enhanced by the presence of the acyl chain. Bioassays showed a difference in activity for the four lipid tail isomers, implying that chirality may play a role in target recognition and the mode of

action involves a specific target rather than a general mechanism. It was shown that, like several AMPs, the bacterial target of $TriA_1$ is the peptidoglycan layer precursor lipid II (LII). Contrasting teixobactin which is promiscuous amongst LII analogues, $TriA_1$ selectively binds LII-DAP, the variant of LII found in most Gram-negative organisms.[114]

Structural NMR studies of $TriA_1$, conducted with and without LII-DAP, in DPC_{d38} micelles showed a conformational change (Figure 4.33). Upon exposure to LII, $TriA_1$ moves from a closely packed U-shaped strand with all hydrophobic residues on one face to a more open V-shape which brings residues Trp_5 and Phe_9 (green) in close enough proximity to engage in π-stacking.[114] This covalent brace likely supports the peptide's more open conformation for optimal binding with LII. Dab_8 (blue) has been pinpointed as a critical residue[128] and docking studies with LII-DAP suggest $TriA_1$'s Dab_8 amine interacts with the ε-carboxylate of DAP_3.[114] This would explain the lack of binding observed for LII-Lys, which has no carboxylate at this position.

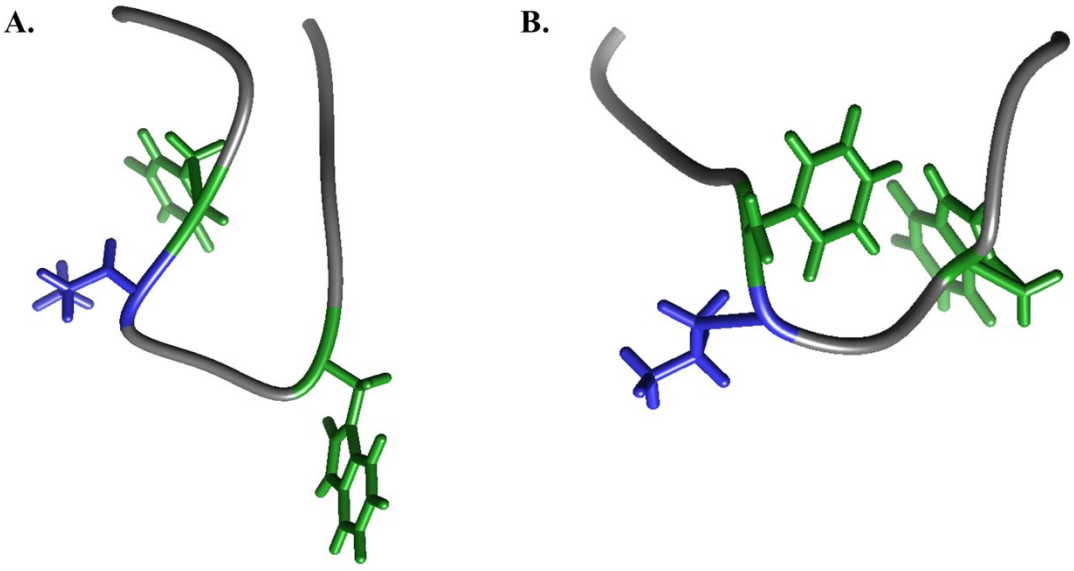


Figure 4.33 Structure of $TriA_1$ in DPC (A) without and (B) with LII

Although it is postulated the lipid tails of $TriA_1$ and LII are inserted within the DPC membrane mimic, the exact position of this peptide-target complex within the membrane is not known. A common mechanism of action found for AMPs is insertion into the membrane, whether through binding a target or a general interaction, and causing the formation of a pore. As a continuation of work previously done within our group, we next sought to investigate the orientation of $TriA_1$-LII within DPC micelles. Our approach aimed to install heteroatoms along the DPC lipid tail, creating a series of depths within the micelle. These modifications would result in a change to the local environment experienced by molecules interacting at the surface or inserting within the micelle (Figure 4.34). We opted to exchange methylene carbons with oxygen in hopes of promoting changes experienced by the protons of the $TriA_1$-LII complex interacting with the DPC micelles, changes which would be detectable *via* NMR chemical shift monitoring.

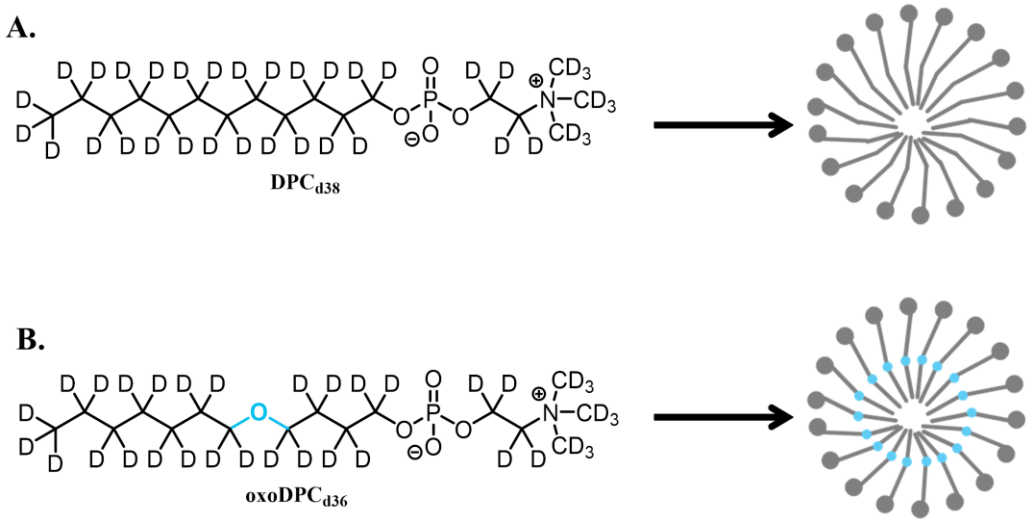


Figure 4.34 DPC forming micelles (A) without and (B) with oxygen depth gauge

4.2 Results and discussion

4.2.1 Syntheses of required compounds

Syntheses of truncated Gram-negative LII variant (*Z,Z*-C_{15} LII-DAP, **37**) and simplified C_8 acylated $TriA_1$ (Oct-$TriA_1$, **41**) were completed by Isaac Antwi following previously developed methods (Figure 4.34).[114,127] These analogues were preferentially used as they are simpler and cheaper to make or are more conducive to NMR studies, while conferring similar activities to wild-type compounds.

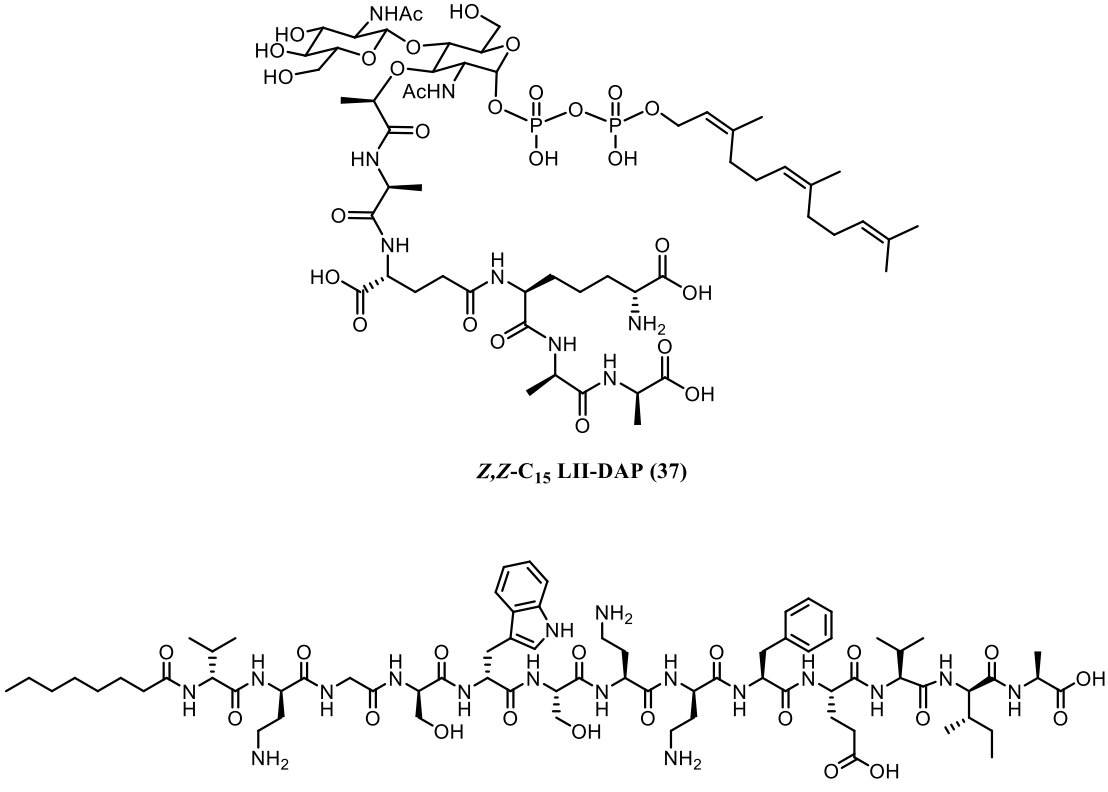


Figure 4.35 Structures of LII and $TriA_1$ analogues used for NMR studies

Next, a series of deuterated DPC analogues were synthesized by Isaac Antwi with a designated methylene group along the lipid chain replaced with an oxygen atom (Figure 4.36). The deuterated oxygen DPC analogues ($oxoDPC_{d36}$) chosen were $3oxoDPC_{d36}$ (**42**), $5oxoDPC_{d36}$ (**43**), $6oxoDPC_{d36}$ (**44**), and $10oxoDPC_{d36}$ (**45**). The oxygen label of these compounds was strategically placed near the head group (the surface of the micelle), in the middle, and near the tail (close to the centre of the micelle). Additionally, analogues **43** and **44**, differing by a single position, would theoretically be present on the two different faces of each DPC molecule. These analogues would ensure a comprehensive survey of the spatial arrangement within a DPC micelle formed by these analogues.

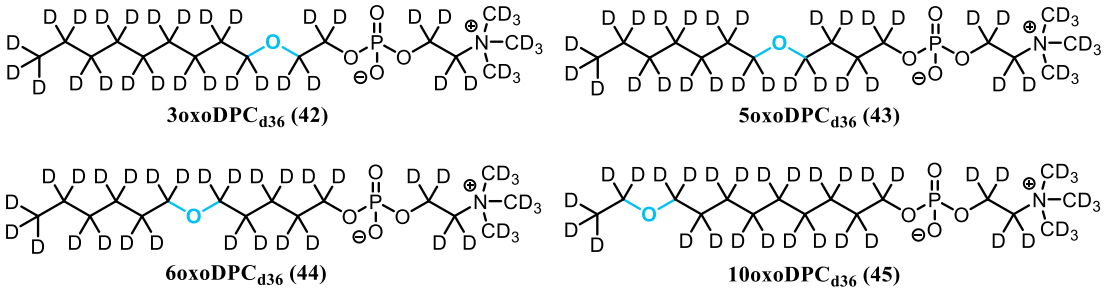
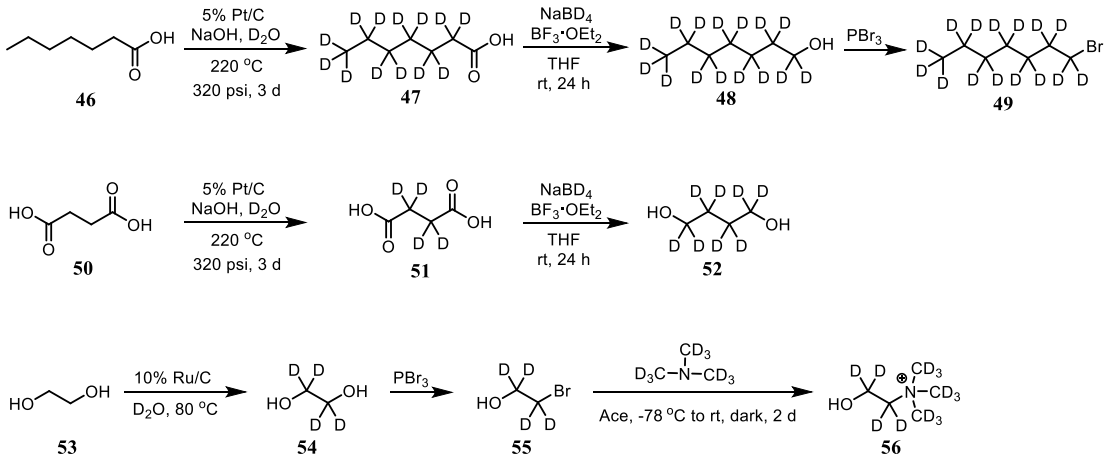


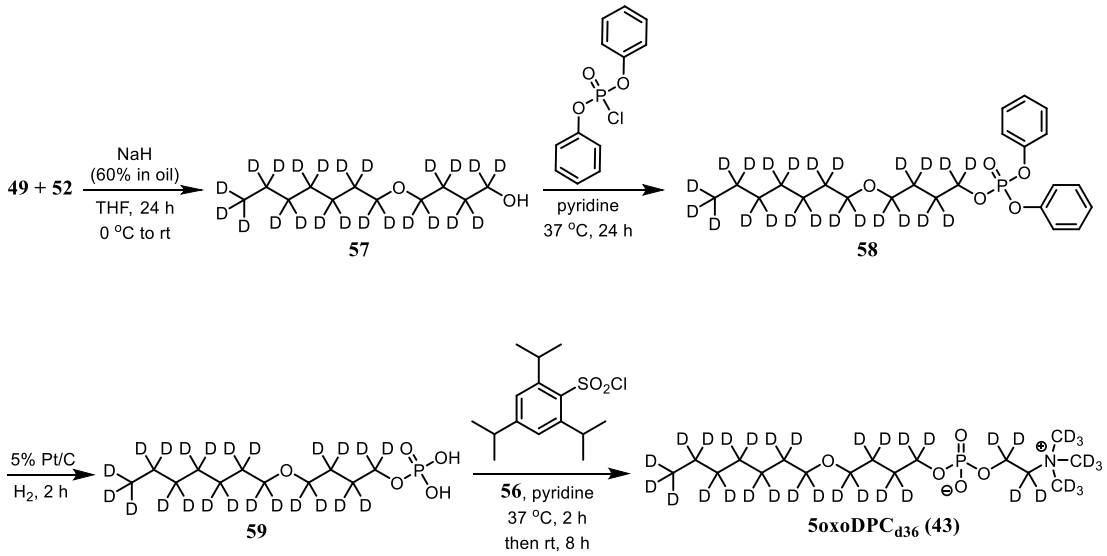
Figure 4.36 Structures of the four oxygen-embellished DPC_{d36} analogues

The synthesis of these four analogues was completed by Isaac Antwi.[129] The initial steps required a metal-catalyzed deuteration of each lipid chain segment. Taking $5oxoDPC_{d36}$ as an example, this would require the synthesis of deuterated 1-bromoheptane **49**, butane-1,4-diol **52**, and 2-hydroxy-*N,N,N*-trimetylethane-1-aminium **56**; this was done as outlined in Scheme 4.5. Starting from the appropriate chain lengths, the segments required for the other three analogues were also synthesized by this approach, with a 97% average deuteration for the whole molecule.



Scheme 4.5 Synthesis of necessary deuterated lipid segments of DPC

With the deuterated segments in hand, the assembly of the full-length DPC was done by Isaac Antwi as shown in Scheme 4.6. The bromine of deuterated 1-bromoheptane **49** was displaced by butane-1,4-diol **52** to yield the 11-carbon DPC tail **57**. The terminal alcohol of **57** was reacted with diphenylphosphoryl chloride to yield **58**, which was deprotected to yield the free phosphate DPC tail **59**. Finally, deuterated 2-hydroxy-*N,N,N*-trimetylethane-1-aminium **56** was coupled to **59** using triisopropylbenzenesulfonyl chloride to yield 5oxoDPC$_{d36}$ **43**.



*Scheme 4.6 Synthesis of 5oxaDPC$_{d36}$ analogue **43***

4.2.2 Oct-TriA$_1$ in deuterated oxoDPC analogues

Samples were prepared in an aqueous phosphate buffer system (pH 6.1), with 4 mM Oct-TriA$_1$ and 180 mM DPC. Circular dichroism scans of Oct-TriA$_1$ in the presence of all four oxo analogues, conducted by Isaac Antwi, showed relatively similar spectra compared to the native DPC. This indicated that Oct-TriA$_1$ maintains a similar structural conformation in the presence of all four analogues. Encouraged by this consistency, samples were prepared with the same concentration ratio (1:45 peptide:lipid) in 300 µL of 9:1 H_2O:D_2O for NMR experiments.

Unfortunately, two issues arose during NMR data acquisition. The oxoDPC analogues, although highly deuterated, still had a small percentage of protons along the lipid chain. Due to the large amount of DPC required to form micelles in NMR studies, these aliphatic proton signals overlapped with the chemical shifts of the octyl tail as well as the hydrophobic residues of TriA$_1$. In particular, the middle of the C_8 tail and Val$_{11}$, *allo*-Ile$_{12}$, Ala$_{13}$ had significant overlaps and could not be discerned at times.

Secondly, the prepared NMR samples had a tendency to form gels over time. Initial experiments, particularly with 6oxoDPC$_{d36}$ and 10oxoDPC$_{d36}$, had no significant issues. However, in the case of 3oxoDPC$_{d36}$ (the analogue with the oxygen closest to the micelle surface) the gels formed almost immediately. Freeze-drying and subsequent resuspension allowed for TOCSY and NOESY experiments to be conducted within a 10 h timeframe, but 2D carbon-correlation experiments were not possible. High temperature scans up to 60 ^oC were also conducted; however, this approach was not useful in maintaining a gel-free solution-phase sample in the

3oxoDPC$_{d36}$ matrix. Another tactic, to lower the amount of DPC, was also employed. Using CD, a scan of Oct-TriA$_1$ in various DPC concentrations (145, 120, 90, and 70 mM DPC) was performed. At 70 mM, the spectrum shape of Oct-TriA$_1$ changed, thus a DPC concentration above this threshold would be more ideal. Between the 90 and 120 mM trials, which appeared similar based on CD scans, the observed gel-formation timeline was similar. It was decided to adjust the initial ratio and use a 120 mM solution of DPC (making a 1:30 peptide:DPC ratio) in further NMR experiments. Interestingly, the addition of Oct-TriA$_1$ to commercial DPC$_{d38}$ showed the formation of a gel over time as well, suggesting the gel formation may be linked to the exposure of TriA$_1$ to zwitterionic lipids in this aqueous buffer system rather than to the specific modifications of the oxo analogues.

Nevertheless, the structure of Oct-TriA$_1$ in DPC micelles was determined for the four oxoDPC$_{d36}$ analogues using a series of NMR experiments. The full table of assignments for Oct-TriA$_1$ in the four oxoDPC$_{d36}$ analogues as well as commercial DPC$_{d38}$ can be found in Chapter 7. The chemical shifts of Oct-TriA$_1$ in each analogue were compared to the chemical shifts of Oct-TriA$_1$ in commercial DPC$_{d38}$ (Figure 4.37). Changes are shown throughout this chapter in shades of blue: dark navy blue indicates a difference of 0.01 ppm, medium ocean blue indicates a change of 0.01 – 0.04 ppm, and differences of 0.05 ppm or higher are shown in light aquamarine blue.

As highlighted in Figure 4.37, the octyl tail showed modest differences in all of the oxoDPC$_{d36}$ systems, making it hard to ascertain how deep the lipid tail penetrates into the micelle. Also, it appears that, on average, Oct-TriA$_1$ chemical shifts are most different in 5oxoDPC$_{d36}$ micelles. Discounting 3oxoDPC$_{d36}$ due to solubility issues,

this DPC analogue is found closest to the surface of the micelle. The overall widespread distribution of chemical shift changes observed for Oct-TriA$_1$ could be indicative of a more fluid system, where the peptide is dynamic in its affiliation with the micelle, rather than a more static, well-defined binding surface interface.

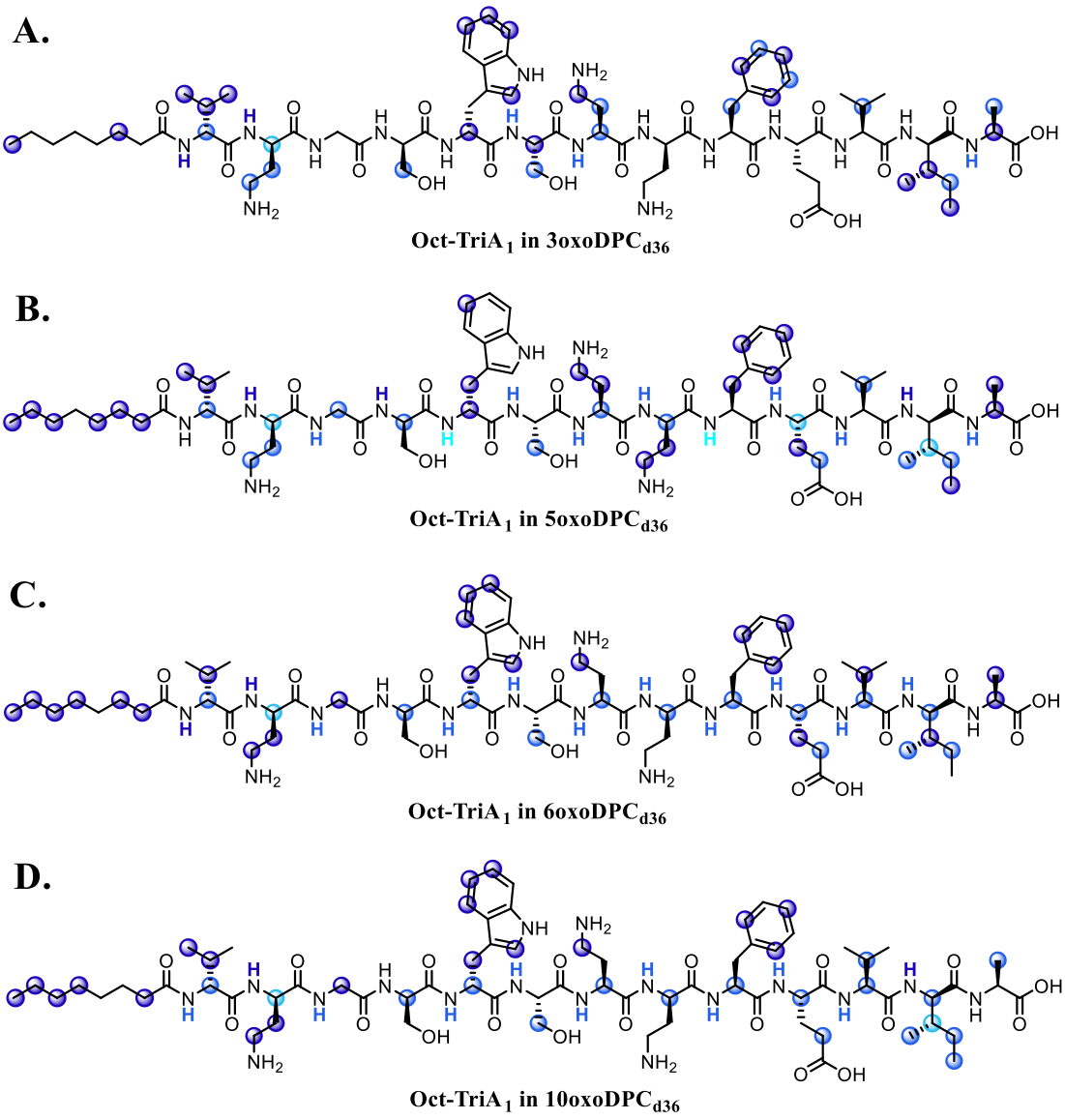


Figure 4.37 2D representations of changes in Oct-TriA$_1$ chemical shifts observed in (A) 3oxoDPC$_{d36}$ (B) 5oxoDPC$_{d36}$ (C) 6oxoDPC$_{d36}$ and (D) 10oxoDPC$_{d36}$ micelles

4.2.2.1 Oct-TriA$_1$ in 3oxoDPC$_{d36}$

In the case of 3oxoDPC$_{d36}$, where the propensity of gel formation was greatest, a partial assignment was possible. Residues Gly$_3$ and Glu$_{10}$, the lipid tail, as well as the alpha protons of Ser$_4$, Dab$_8$, Phe$_9$, and Val$_{11}$ were not discerned. It could be possible that these residues are involved in interactions which led to the formation of the observed gel, making the protons inaccessible by solution-phase NMR.

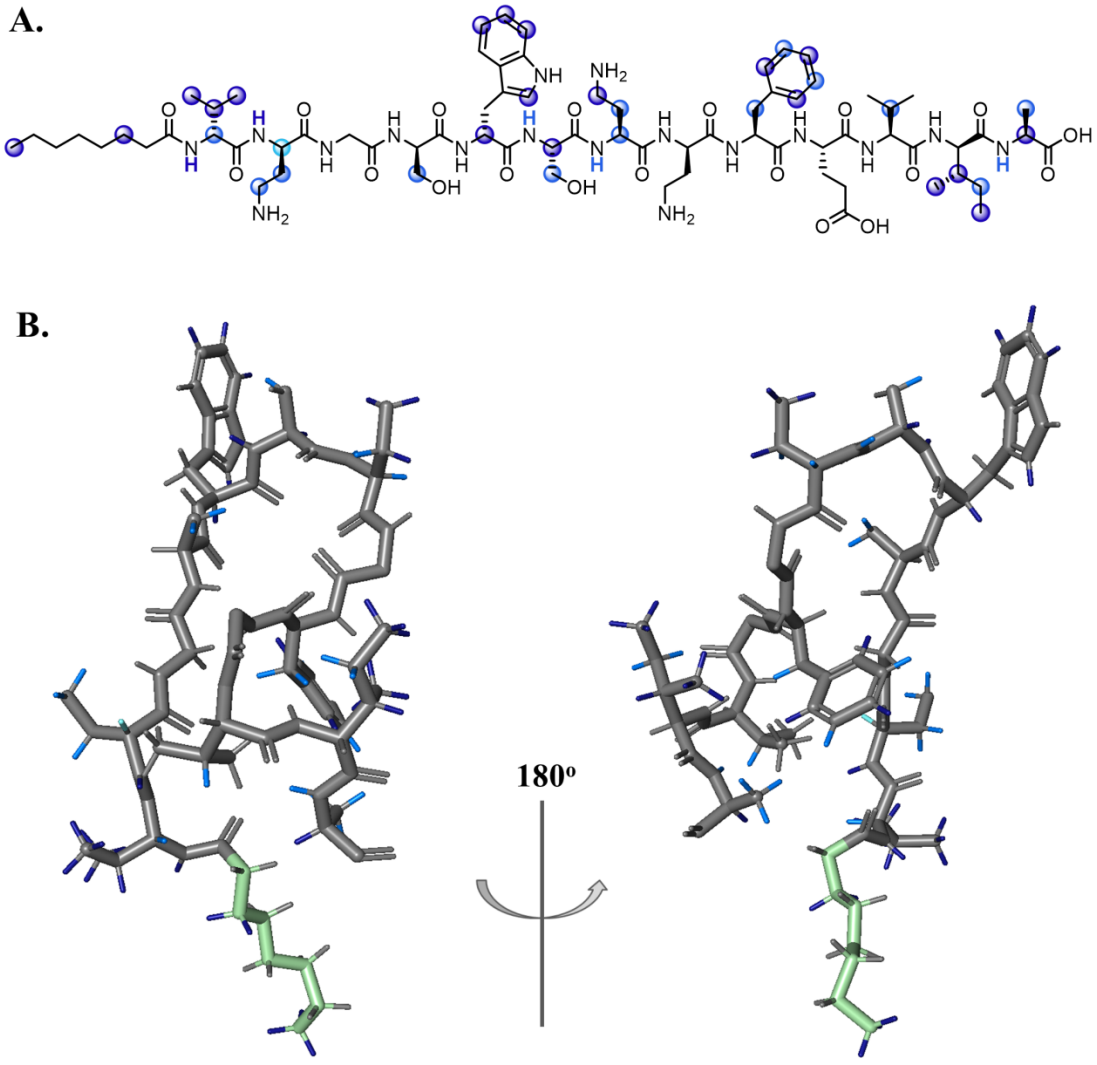


Figure 4.38 Oct-TriA$_1$ in 3oxoDPC$_{d36}$ as (A) 2D and (B) 3D representations

Once again using shades of blue to indicate differences (lighter blue denoting greater chemical shift change), Figure 4.38 illustrates a two- and three-dimensional

representation of Oct-TriA$_1$ in 3oxoDPC$_{d36}$; the lipid tail was coloured in pale green in all structures for ease of orientation. One of the biggest observed changes occurs at the alpha protons of the Val$_1$ and Dab$_2$ residues, with modest changes observed for the rest of the attached sidechains. These two amino acids are the ones closest to the lipid tail, and could be the closest membrane-interacting segments of Oct-TriA$_1$.

4.2.2.2 Oct-TriA$_1$ in 5oxoDPC$_{d36}$

The chemical shift changes observed for Oct-TriA$_1$ in 5oxoDPC$_{d36}$ are illustrated in Figure 4.39. This analogue of oxoDPC exerted the largest overall changes observed in Oct-TriA$_1$ compared to the other analogues, as is illustrated by a higher occurrence of light (Δ 0.05 or greater ppm) and medium blue (Δ 0.01 – 0.04 ppm). In addition to the change observed for Val$_1$ and Dab$_2$ as in the 3oxoDPC$_{d36}$ discussed previously, there are also larger changes observed in Gly$_3$, the backbone amides of Trp$_5$ and Phe$_9$, and the alpha proton of Glu$_{10}$ and sidechain of *a*Ile$_{12}$. Additionally, modest changes are observed in the areas surrounding the amide protons of the aromatic residues, particularly in Ser$_4$ and Dab$_8$. Looking at the three-dimensional fold (Figure 4.39B), the noted chemical shift changes correspond to protons on one face of the peptide. Rotating the structure another 90^o allows for a side-view (last panel) of Oct-TriA$_1$ where it can be noted that the left side of the structure contains more grey hydrogens while the right hand side is noticeably more colourful, with several hydrogens coloured in medium or light blue. It could be possible that, after the lipid tail of the Oct-TriA$_1$ is inserted into the micelle, the rest of the tridecaptin folds into a 'boat' structure; wherein one surface of the molecule floats on top of the membrane, while

the other face (which experienced significant chemical shift differences) becomes submerged within the micelle lipids (and therefore closer to the oxygen label). This scenario would have the lipid tail on one side and Trp_5 on the other as two membrane-anchors, keeping the tridecaptin within this orientation.

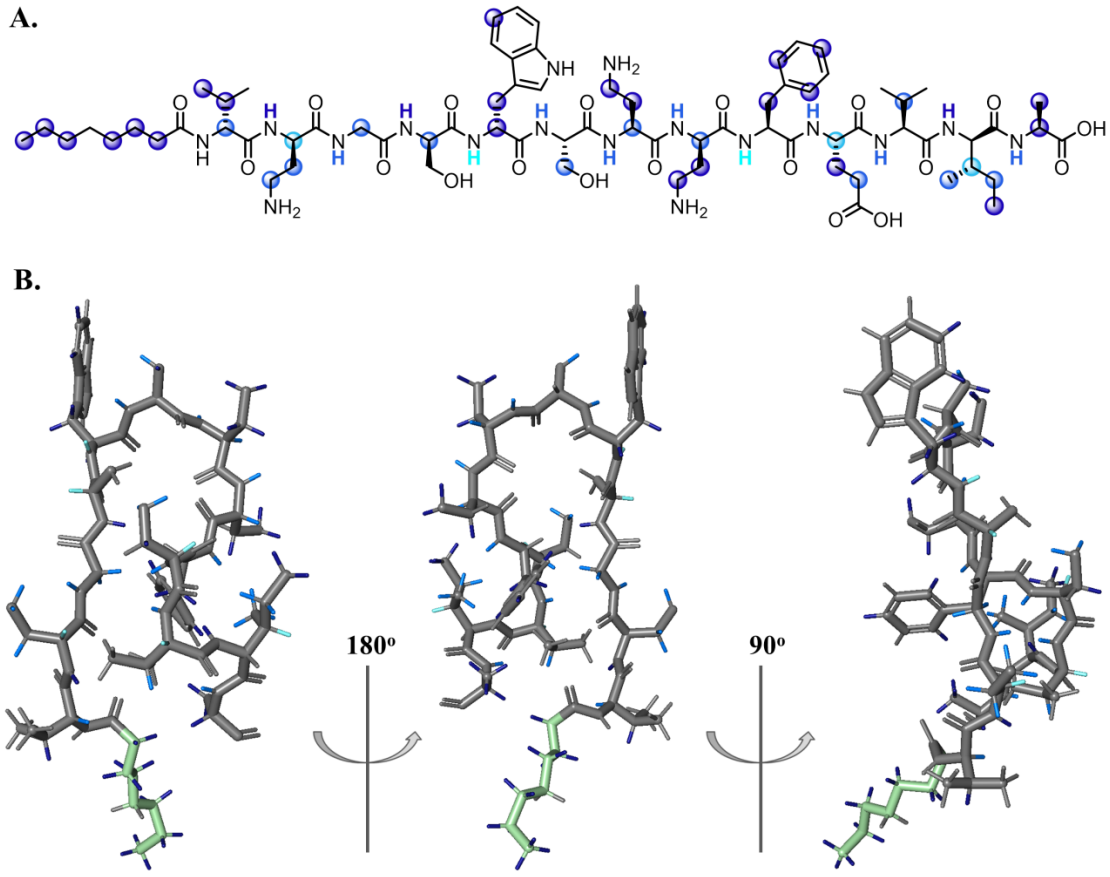


Figure 4.39 Oct-TriA$_1$ in 5oxoDPC$_{d36}$ as (A) 2D and (B) 3D representations

4.2.2.3 Oct-TriA$_1$ in 6oxoDPC$_{d36}$

The next analogue in the sequence, in 6oxoDPC$_{d36}$, contains the oxygen atom modification just one position further into the micelle than 5oxoDPC$_{d36}$, yet shows surprisingly different results. The proton chemical shifts of Oct-TriA$_1$ in 6oxoDPC$_{d36}$ do not show changes as significant as those observed in the previous analogue (Figure

4.40). The aromatic residues noted in the previous analogue, Trp_5 and Phe_9, still maintain some modest changes, with the amide and alpha protons of these two residues as well as the backbone protons of neighbouring residues showing modest (0.01 – 0.04 ppm) changes. Compared with the $5oxoDPC_{d36}$, there are also fewer substantial differences noticed in the chemical shifts of residues Glu_{10} and $aIle_{12}$. The changes in shifts observed at Val_1 and Dab_2 are consistent throughout these studies, supporting the notion that the lipid tail of Oct-$TriA_1$ inserts into the membrane for all $oxoDPC_{d36}$ analogues.

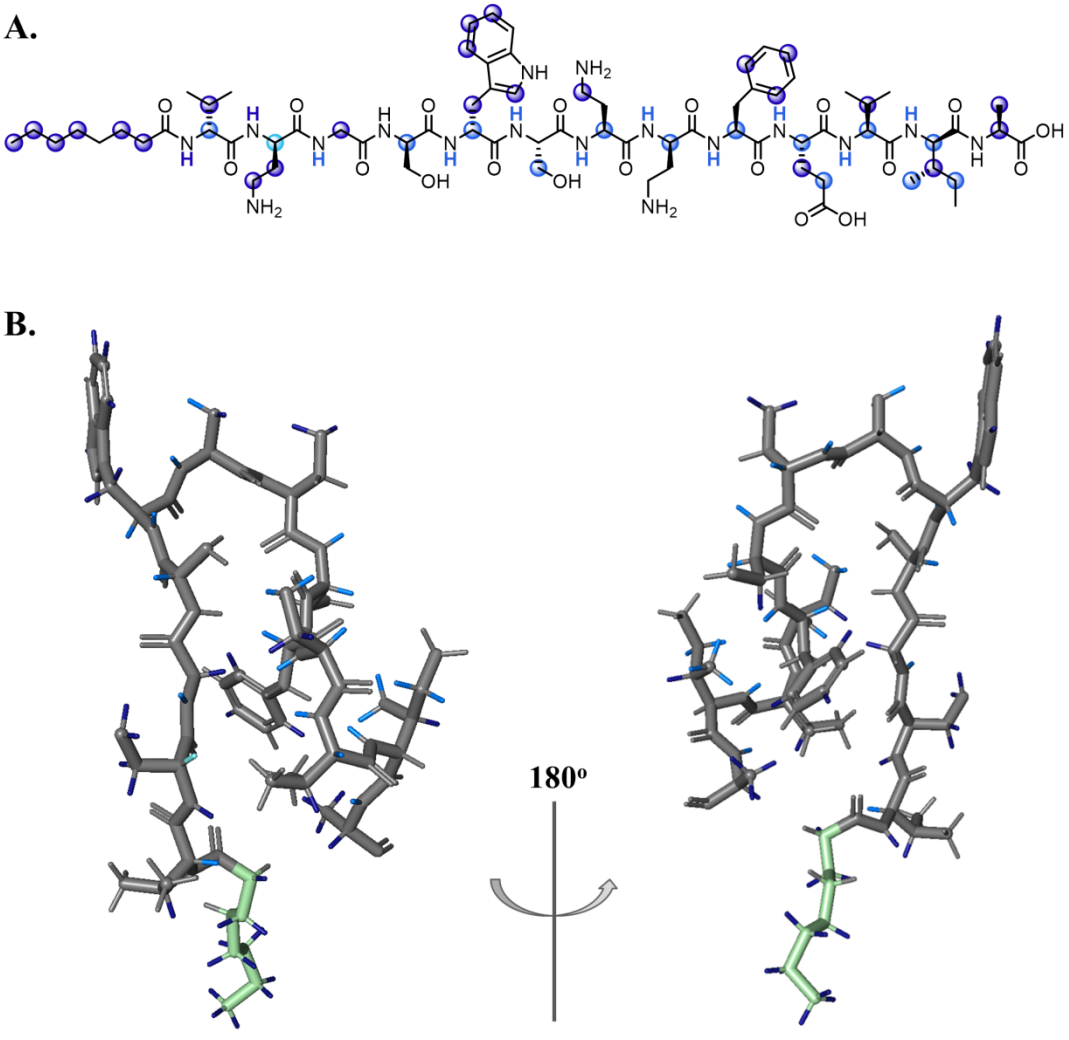


Figure 4.40 Oct-$TriA_1$ in $6oxoDPC_{d36}$ as (A) 2D and (B) 3D representations

4.2.2.4 Oct-TriA_1 in 10oxoDPC_{d36}

Differences in proton chemical shifts observed for Oct-TriA_1 in the last analogue, 10oxoDPC_{d36}, are displayed in Figure 4.41. Overall, the changes noticed in this case showed a relatively similar pattern to that observed for 6oxoDPC_{d36}. Although the oxygen label is believed to be significantly further from the surface and close to the centre of the micelle, the chemical shifts differences noted in Oct-TriA_1 were moderate when compared to the peptide in commercial DPC_{d38}.

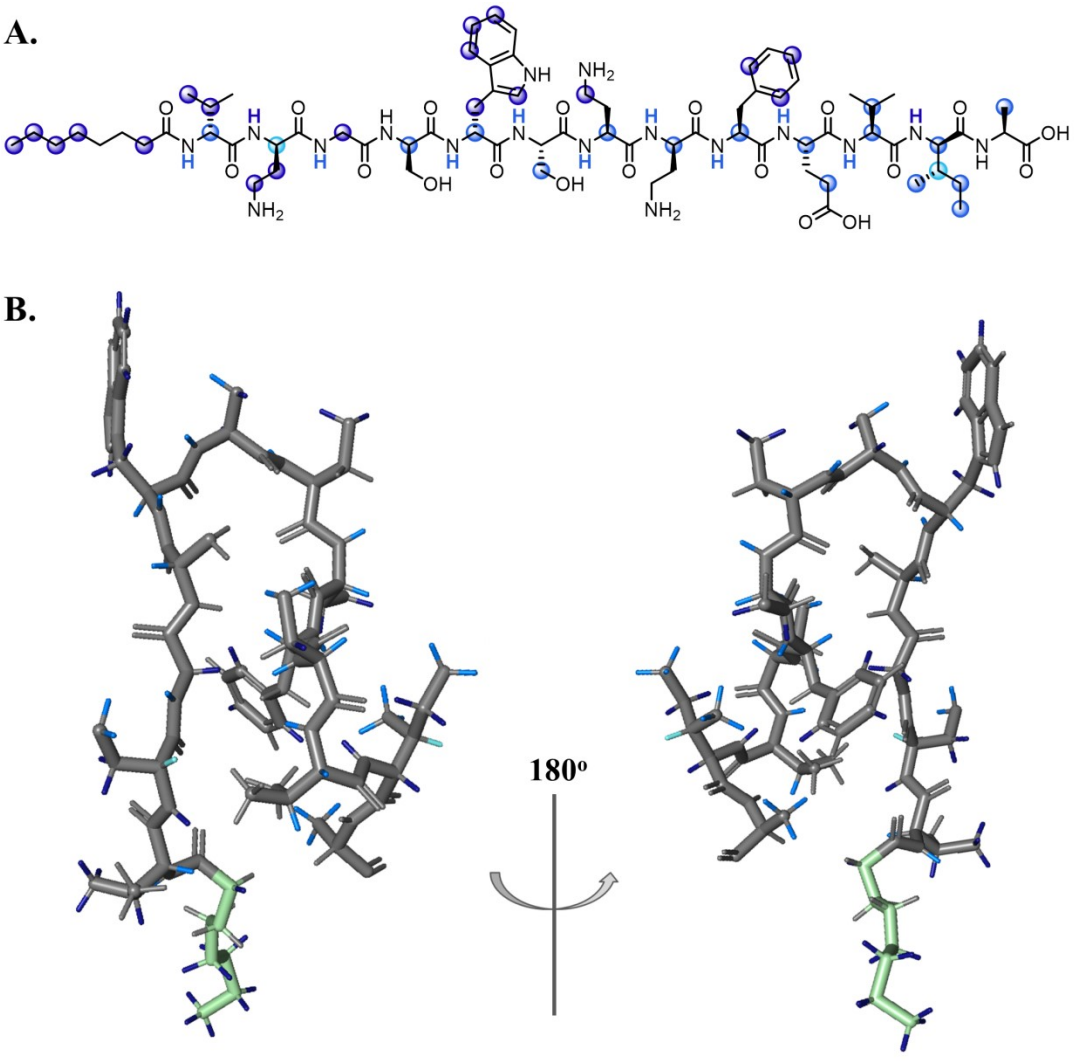


Figure 4.41 Oct-TriA_1 in 10oxoDPC_{d36} as (A) 2D and (B) 3D representations

The most significant changes observed in Oct-TriA_1 chemical shifts throughout this study were when the peptide was immersed in 5oxoDPC_{d36} micelles. This analogue contains the oxygen label in the top third of the micelle, making it the second closest to the micelle surface. It is possible greater changes would have been observed for 3oxoDPC_{d36}, however in this analogue the peptide was too prone to gel formation to allow for a full structural assignment.

4.2.3 *Z,Z*-C_{15} LII-DAP in deuterated oxoDPC analogues

Unlike the Oct-TriA_1 experiments, the addition of LII to the oxoDPC micelles did not propagate the formation of a gel, even after extended periods of time. However, a different issue was observed with the LII after prolonged exposure to the slightly acidic solvent system. The LII molecule appears to be cleaving at the pyrophosphate connection, splitting the molecule into the pentapeptide-disaccharide phosphate and the C_{15} lipid phosphate fragments (confirmed by HPLC and HR-MS). A pH scan of the LII in phosphate buffer was conducted and it was determined that an optimal pH for LII is above 6.7 in aqueous buffer systems. However, as the hope was to study the Oct-TriA_1 and LII complex, moving the experiments closer to the neutral was not ideal as this would broaden the chemical shift peaks of the peptide amide backbone.

NMR experiments of *Z,Z*-C_{15} LII-DAP in the four oxoDPC_{d36} analogues as well as commercial DPC_{d38} were conducted and tables of assignments can be found in Chapter 7. Figure 4.42 indicates the changes in chemical shift observed for each analogue as compared to *Z,Z*-C_{15} LII-DAP in commercial DPC_{d38}. Differences are shown in blue of increasing brightness: dark blue means a change of 0.01 ppm,

medium blue indicates a difference of 0.01 – 0.04 ppm, and changes of 0.05 ppm or greater are in light blue.

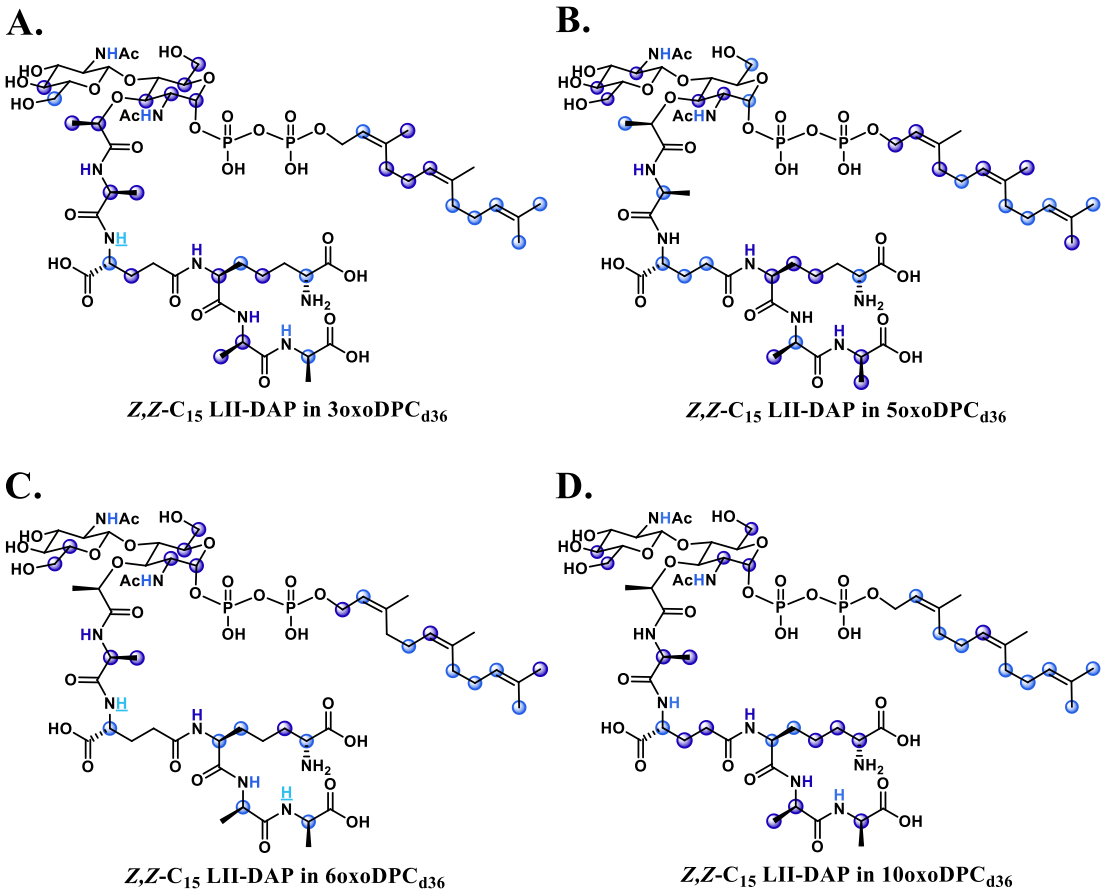


Figure 4.42 2D representations of changes in Z,Z-C_{15} LII-DAP chemical shifts observed in (A) 3oxoDPC$_{d36}$ (B) 5oxoDPC$_{d36}$ (C) 6oxoDPC$_{d36}$ and (D) 10oxoDPC$_{d36}$ micelles

The disaccharide moiety of the *Z,Z*-C_{15} LII-DAP appears to be uniformly behaved in terms of chemical shift differences across all experiments with the various oxo analogues. On the other hand, there are some differences observed in the pentapeptide and terpene tail when exposed to the four oxoDPC$_{d36}$ analogues.

4.2.3.1 *Z,Z*-C_{15} LII-DAP in 3oxoDPC$_{d36}$

Figure 4.43 shows the changes observed in *Z,Z*-C_{15} LII-DAP chemical shifts when in 3oxoDPC$_{d36}$ vs DPC$_{d38}$; the sugars are coloured in pale yellow and the phosphate groups are in pale pink to help in visualizing the transition from two- to three-dimensional representation. As this analogue is found closest to the micelle surface, it is not surprising that there are some changes observed within the terpene tail as, presumably, this portion is embedded within the micelle. The pentapeptide, in particular the backbone amide of Glu_2 and Ala_5 also show some differences in shift.

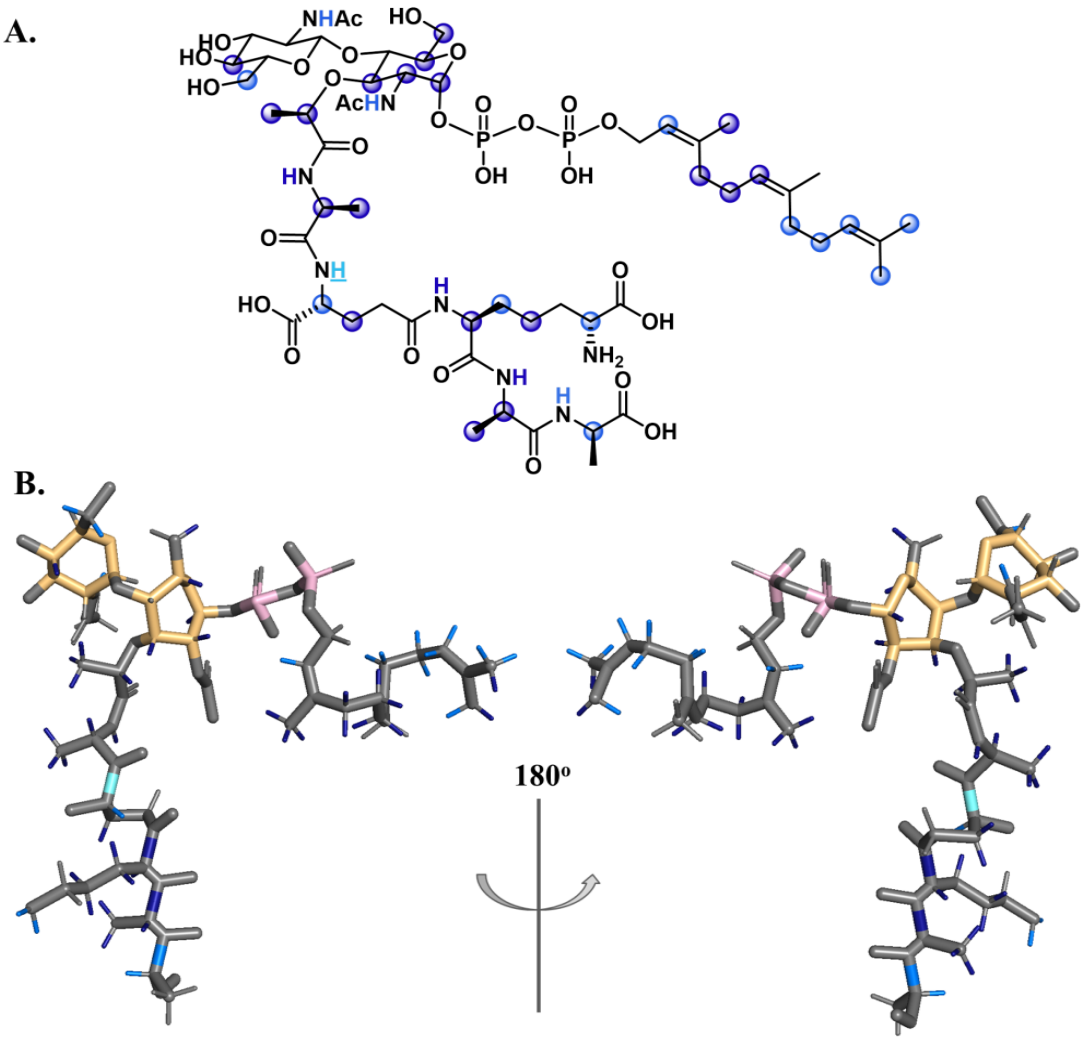


Figure 4.43 Z,Z-C_{15} LII-DAP in 3oxoDPC$_{d36}$ as (A) 2D and (B) 3D representations

4.2.3.2 *Z,Z*-C_{15} LII-DAP in 5oxoDPC$_{d36}$

Chemical shift differences noted for *Z,Z*-C_{15} LII-DAP when in 5oxoDPC$_{d36}$ are relatively widespread (Figure 4.44). The terpene tail shifts, when compared to the previous analogue, show that the methylene at C6 and methyl on C9 are affected more by this analogue. In contrast, the amide backbone is relatively less affected in 5oxoDPC$_{d36}$ than in 3oxoDPC$_{d36}$, with the changes at Glu_2 and Ala_5 not as prominent. Instead, the alpha proton of Ala_4 and the sidechain of Glu_2 show moderate changes.

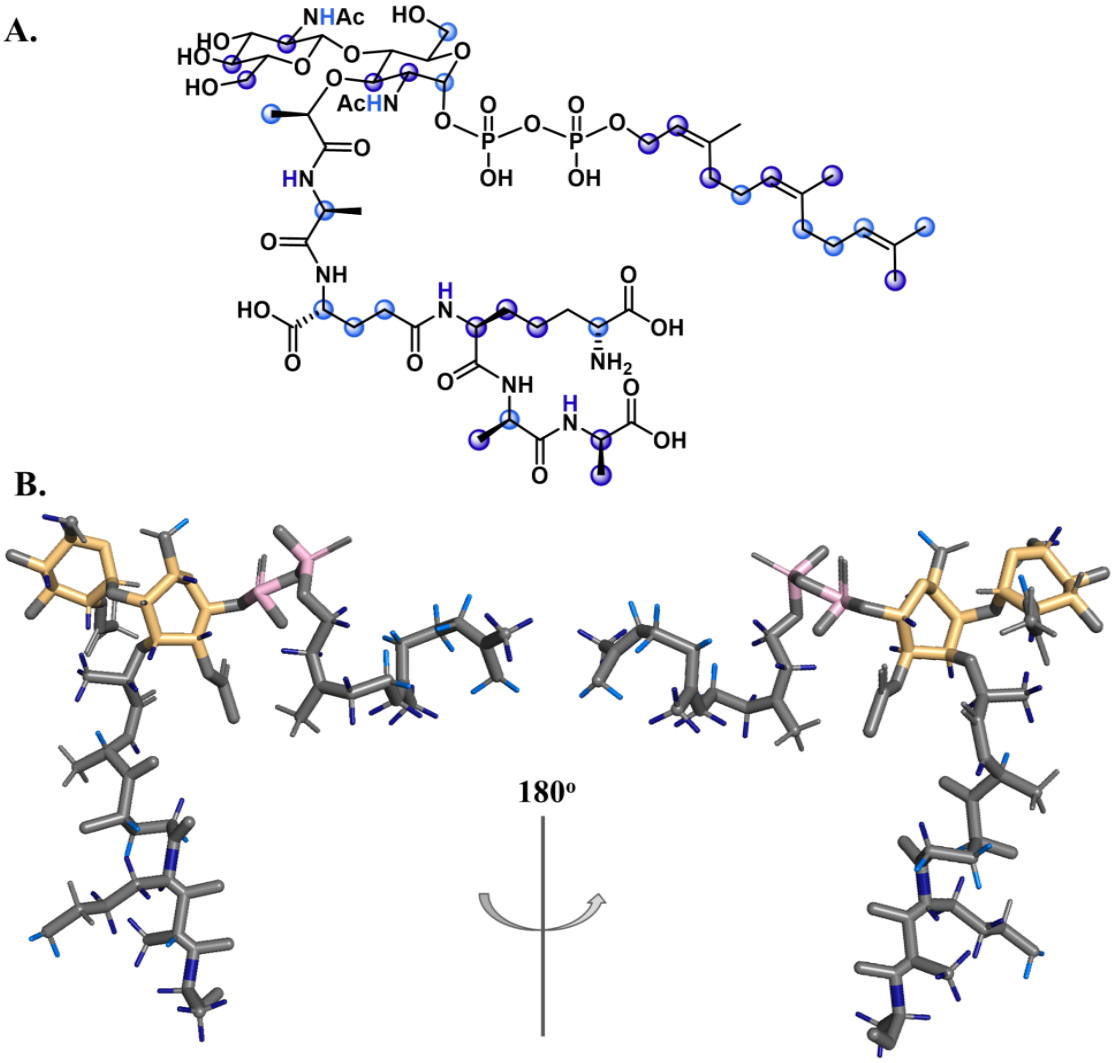


Figure 4.44 Z,Z-C_{15} LII-DAP in 5oxoDPC$_{d36}$ as (A) 2D and (B) 3D representations

4.2.3.3 *Z,Z*-C_{15} LII-DAP in 6oxoDPC$_{d36}$

The most significant changes observed for *Z,Z*-C_{15} LII-DAP chemical shifts occurred in experiments with 6oxoDPC$_{d36}$ (Figure 4.45). As observed for 3oxoDPC$_{d36}$, chemical shift difference exists along the terpene tail with the exception of C4 and C9 methyl groups. The backbone protons of all five residues in the peptide chain showed some changes, with the most significant being observed for the amides of Glu_2 and Ala_5, once again similar to the data gathered of *Z,Z*-C_{15} LII-DAP in 3oxoDPC$_{d36}$.

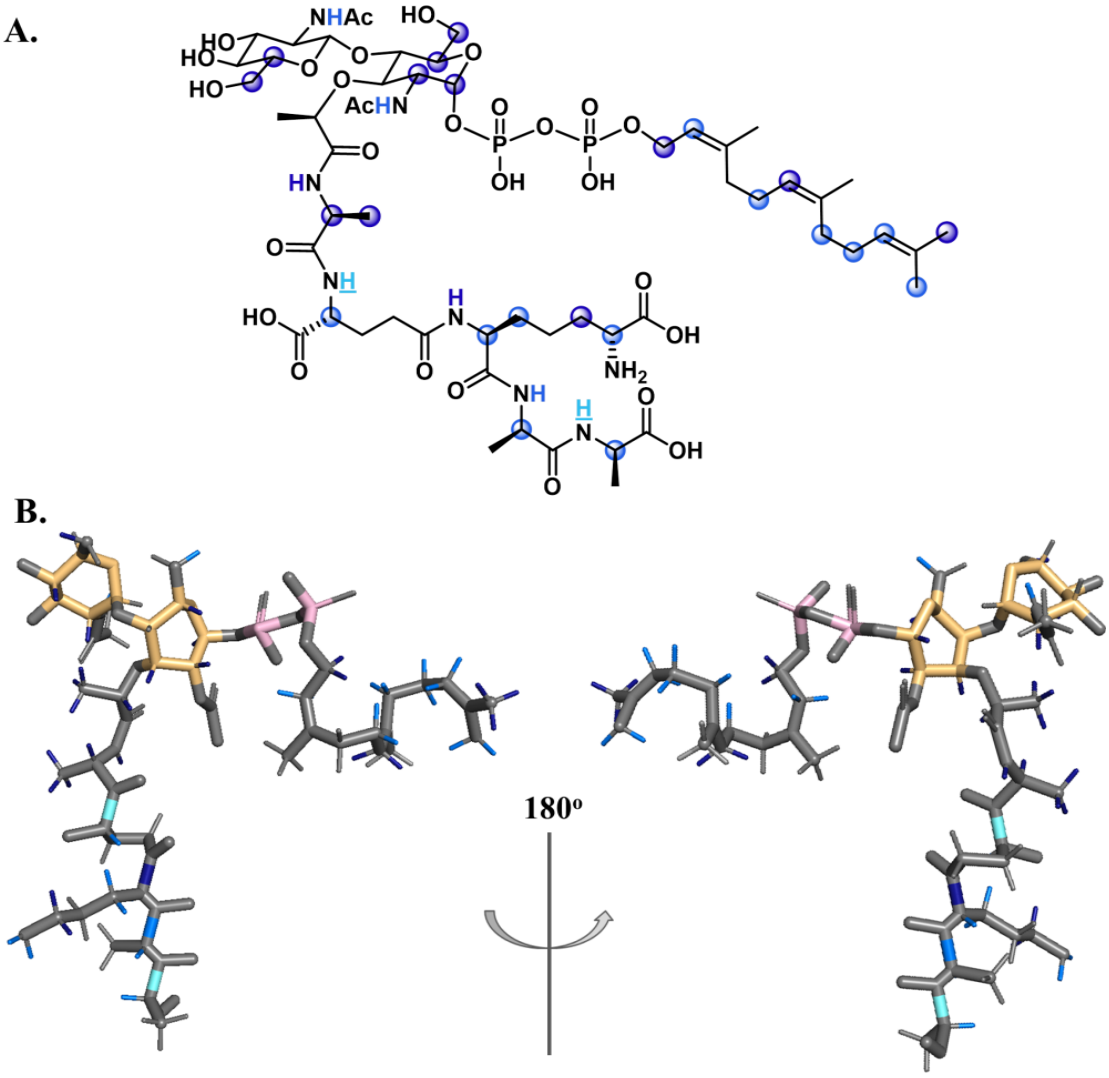


Figure 4.45 Z,Z-C_{15} LII-DAP in 6oxoDPC$_{d36}$ as (A) 2D and (B) 3D representations

4.2.3.4 *Z,Z*-C_{15} LII-DAP in 10oxoDPC$_{d36}$

Although it would have been expected that *Z,Z*-C_{15} LII-DAP in 10oxoDPC$_{d36}$, the analogue with oxygen furthest away from the micelle surface, would contain the most modest changes, that was not the case. Chemical shift differences observed for *Z,Z*-C_{15} LII-DAP in 10oxoDPC$_{d36}$ as compared to the commercially available unoxygenated DPC$_{d38}$ are shown in Figure 4.46. Changes along the terpene tail and within the pentapeptide are consistent with shifts observed in 5oxoDPC$_{d36}$, with the exception of Glu_2 and Ala_5 amide protons which here have a medium change.

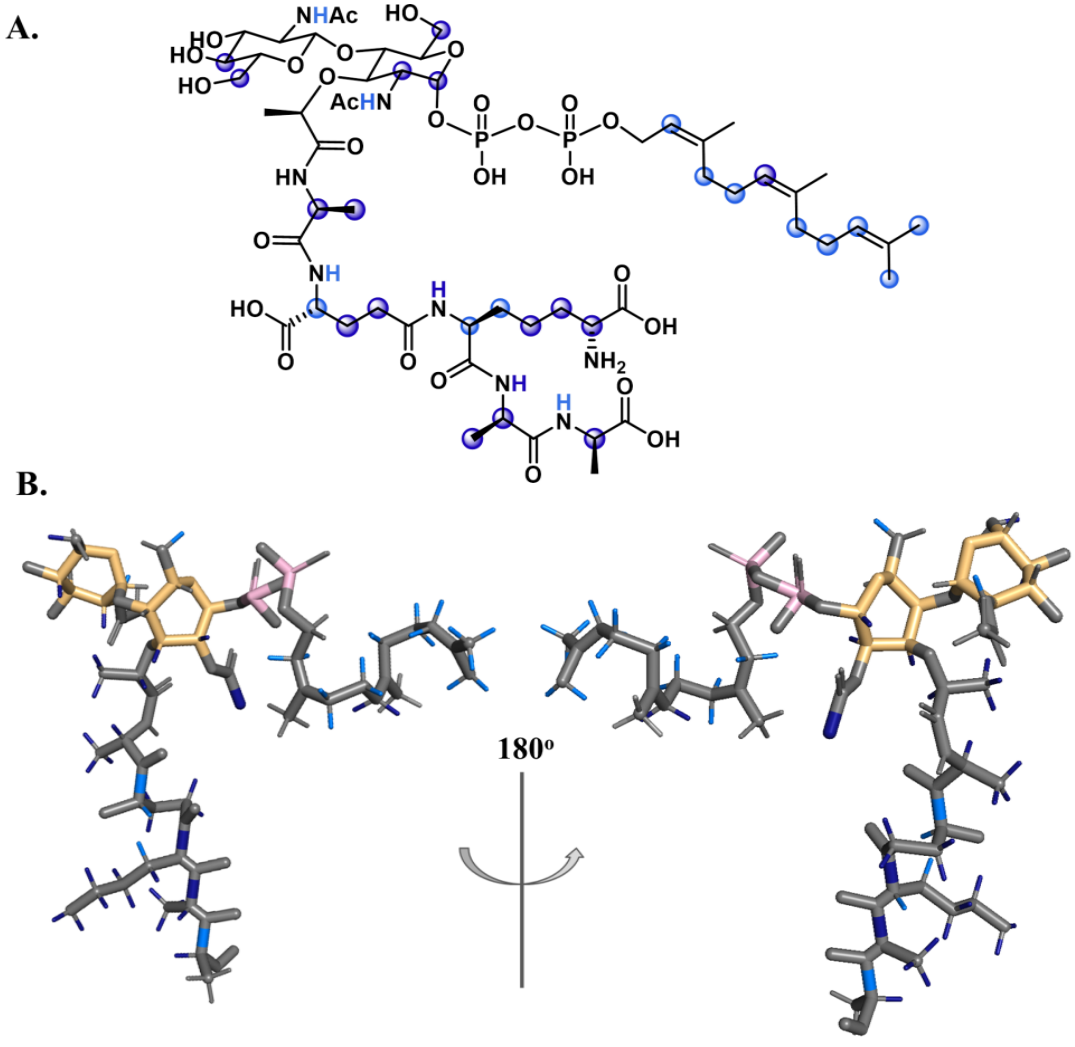


Figure 4.46 Z,Z-C_{15} LII-DAP in 10oxoDPC$_{d36}$ as (A) 2D and (B) 3D representations

4.2.4 Oct-TriA$_1$ and *Z,Z*-C$_{15}$ LII-DAP in deuterated oxoDPC analogues

Following the data acquisition of the peptide and lipid in the four oxo analogues individually, the subsequent step was to perform titration experiments of Oct-TriA$_1$ into *Z,Z*-C$_{15}$ LII-DAP containing oxoDPC$_{d36}$ micelles. The goal was to titrate tridecaptin at three various concentrations, and track any changes using NMR at these three different molar ratios. A tridecaptin:lipid ratio of 1:2 was first tried in 10oxoDPC$_{d36}$, keeping the oxoDPC$_{d36}$ concentration consistent at 120 mM, with 2 mM of Oct-TriA$_1$ being added to 4 mM of *Z,Z*-C$_{15}$ LII-DAP. These data were relatively tough to distinguish and assign as the concentration of lipid overpowered the tridecaptin signals. The second concentration set, a 1:1 ratio of tridecaptin:lipid of 4 mM each, provided much better data and could be used to assign the structure of each component in the following section. Lastly titration of a 2:1 ratio of tridecaptin:lipid was analyzed; however, by the time all the data were acquired for this sample, significant degradation of the *Z,Z*-C$_{15}$ LII-DAP was observed.

Thus, subsequent experiments were performed with a 1:1 ratio of tridecaptin:lipid in the remaining oxoDPC$_{d36}$ analogues. Data were gathered for 10oxoDPC$_{d36}$ and 6oxoDPC$_{d36}$ and assignment tables of Oct-TriA$_1$ and *Z,Z*-C$_{15}$ LII-DAP can be found in Chapter 7. Regrettably, trials performed in 5oxoDPC$_{d36}$ and 3oxoDPC$_{d36}$ resulted in poor data due to gel-formation and lipid degradation occurring faster than the timescale required for data acquisition of two-dimensional NMR experiments.

4.2.4.1 Oct-TriA$_1$ & *Z,Z*-C$_{15}$ LII-DAP in 6oxoDPC$_{d36}$

The chemical shift changes observed for Oct-TriA$_1$ and *Z,Z*-C$_{15}$ LII-DAP complex in 6oxoDPC$_{d36}$ as compared to the solo compounds can be found in Figures 4.47 and 4.58, respectively. As done previously, protons are coloured in increasingly lighter shades of blue to depict the scale of change observed. Dark navy blue means a change of 0.01 ppm, medium ocean blue indicates a difference within 0.01 – 0.04 ppm, and light aquamarine blue highlights changes of 0.05 ppm or greater. In the case of structure analyses performed in 6oxoDPC$_{d36}$, which showed significant changes, pink circles highlight areas of note. Tridecaptin's lipid tail is shown in pale green; LII's disaccharides are in yellow and its phosphates are in pale pink.

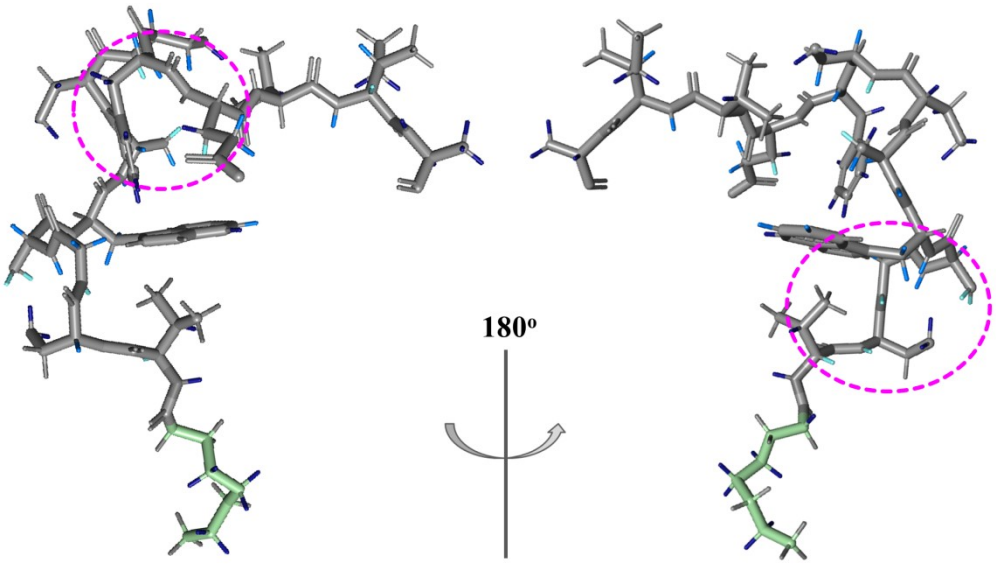


Figure 4.47 Oct-TriA$_1$ in 6oxoDPC$_{d36}$ with Z,Z-C$_{15}$ LII-DAP

Two major pockets within the three-dimensional structure of Oct-TriA$_1$ showed a noticeable change, with great (light blue, 0.05 or more ppm) differences observed for several protons in these regions. In addition to the expected shifts in Dab$_8$, the residue

modelled to interact with the carboxylic acid of LII,[114] changes are also observed in residue Glu_{10} and the area closest to the lipid tail. Specifically the alpha proton of Val_1, the amide protons of Dab_2 and Gly_3, and the sidechain methylene of Ser_4, all of which face towards each other in this pocket, experience a substantial chemical shift change. Other protons around this pocket also show moderate (0.02 – 0.04 ppm) chemical shift differences in the presence of *Z,Z*-C_{15} LII-DAP in 6oxoDPC$_{d36}$.

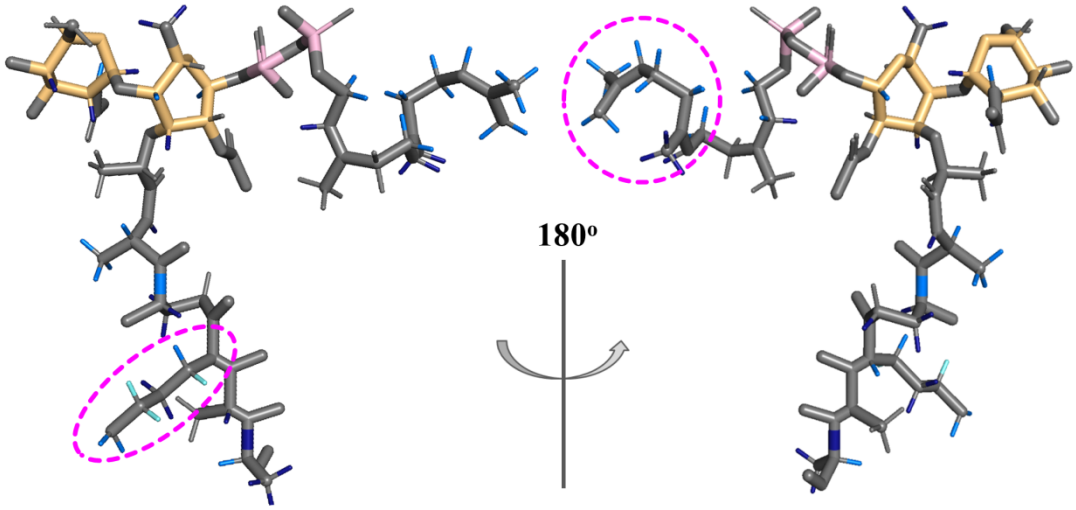


Figure 4.48 Z,Z-C_{15} LII-DAP in 6oxoDPC$_{d36}$ with Oct-TriA_1

Changes observed for *Z,Z*-C_{15} LII-DAP are also relatively significant for this analogue. In particular residues Glu_2 and, more significantly, DAP_3 experience significant changes upon interacting with tridecaptin. This is consistent with the models previously outlined, which predicted that DAP_3 interacts with Dab_8, accounting for the selectivity observed for Gram-negative vs Gram-positive LII (LII DAP vs LII-Lys).[114] The terpene tail of the lipid shows some significant changes upon binding of Oct-TriA_1. Although this region is believed to be embedded into the

membrane and not interact with the peptide, it is possible that the lipid tail of tridecaptin, also believed to be in the membrane, may find itself in close proximity.

4.2.4.2 Oct-TriA$_1$ & *Z,Z*-C$_{15}$ LII-DAP in 10oxoDPC$_{d36}$

The chemical shift changes observed for Oct-TriA$_1$ and *Z,Z*-C$_{15}$ LII-DAP in 10oxoDPC$_{d36}$ can be found in Figures 4.49 and 4.50, respectively. On average, these changes were less significant as those observed for the previous analogue.

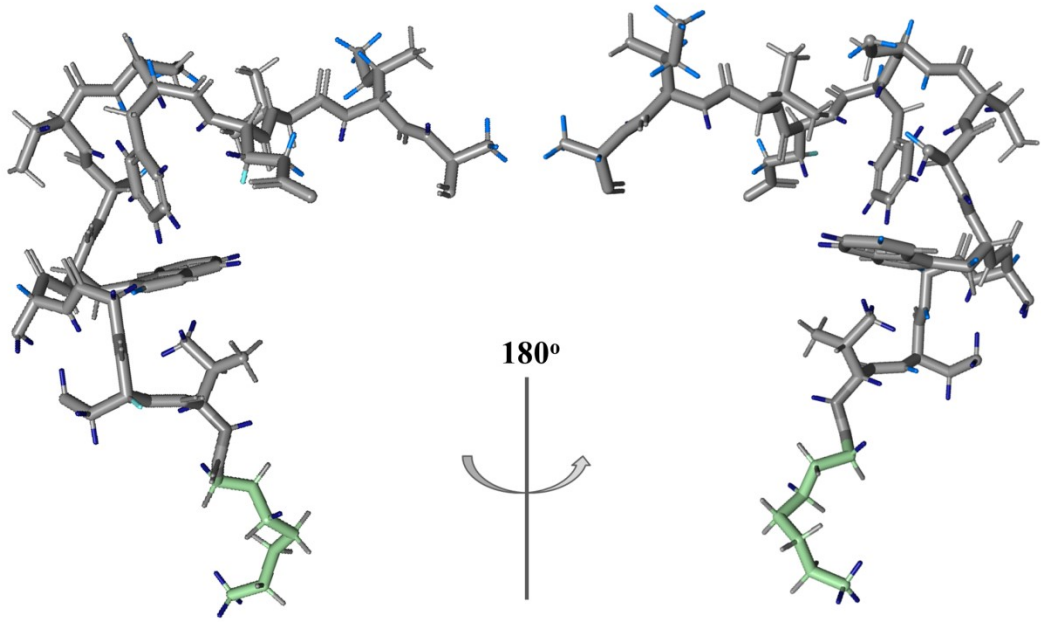


Figure 4.49 Oct-TriA$_1$ in 10oxoDPC$_{d36}$ with Z,Z-C$_{15}$ LII-DAP

The chemical shift changes of Oct-TriA$_1$ with and without *Z,Z*-C$_{15}$ LII-DAP constitute a similar pattern as those observed for the compounds in DPC$_{d38}$. Changes in shift are observed along Val$_1$ and Dab$_2$, residues close to the lipid tail, and in the C-terminal end at residues *a*Ile$_{12}$.

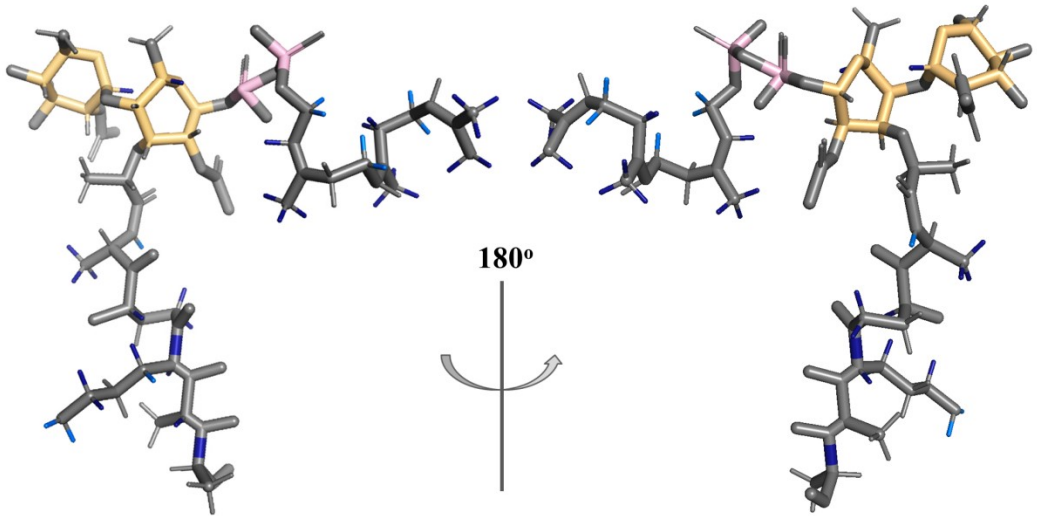


Figure 4.50 Z,Z-C_{15} LII-DAP in 10oxoDPC$_{d36}$ with Oct-TriA$_1$

The chemical shift changes observed for *Z,Z*-C_{15} LII-DAP in the presence of Oct-TriA$_1$ were not very significant for 10oxoDPC$_{d36}$ micelles. The amide and sidechain protons of DAP_3 showed the biggest difference, with medium (0.01 – 0.04 ppm) or small (0.01 ppm) changes.

4.3 Conclusions and future directions

TriA$_1$ is a lipopeptide with potent activity against Gram-negative ESKAPE and foodborne pathogens. It is known to selectively interact with the Gram-negative variant of LII, LII-DAP, located within bacterial cell membranes. The structure of Oct-TriA$_1$ undergoes a conformational change when exposed to LII in DPC membrane-mimicking micelles. Four deuterated DPC analogues with oxygen replacing a methylene at various locations were synthesized for this purpose. NMR structure elucidations of Oct-TriA$_1$ and *Z,Z*-C_{15} LII-DAP were completed in the four

oxoDPC$_{d36}$ analogues. The structures of both components in 6oxoDPC$_{d36}$ and 10oxoDPC$_{d36}$ were also elucidated.

Regrettably, the TriA$_1$-LII system came with a few challenges which prevented the complete study of the complex in all four oxoDPC analogues. Stability and solubility issues, perhaps coupled with a very dynamic relationship with the micelles, limited a comprehensive structure elucidation in DPC analogues that possess an oxygen label closer to the surface. An assumption made for this study is that micelles contain uniformly distributed DPC molecules. While the lipids would likely maintain the general orientation of having the lipid chain pointing towards the inside of the micelle and the phosphate head group located on the micelle surface, the location of the oxygen atom of the synthesized analogues may not be at a uniform depth throughout. Unsaturated hydrocarbon chains may kink or bend and the oxygen atoms may repel one another, preventing uniform stacking within the micelle.

The use of oxoDPC analogues to analyze the spatial arrangement of membrane-interacting peptides with receptors may yet prove viable. An alternate approach, using a larger, more structurally-rigid, lipid-free, isotopically labelled peptide, could be tried in the future. For example, leucocin A, a YGNG-motif containing bacteriocin which is proposed to interact with the maltose phosphotransferase protein complex, has been previously expressed as a $^{13}C,^{14}N$ labelled peptide by our group.[88] Its structure is well-defined in DPC micelles and the peptide is robust and quite stable for long-interval NMR experiments. Furthermore, the labelled material would be more easily distinguished from any potential interference encountered from chemical shifts overlap with oxoDPC.

Chapter 5: Steps towards microcin J25 total synthesis

5.1 Background

Lasso peptides are a class of ribosomally synthesized natural products found in bacteria, named after their threaded lasso, or lariat knot, 3D structural conformation.[22] This unique topology stems from the formation of a N-terminal macrocycle through which the C-terminal tail is threaded, creating a lasso shape. Currently known lasso peptides vary in function, finding potential applications as antibacterial and antiviral agents or as receptor antagonists linked to diabetes and blood pressure.[130]

Peptides of this class undergo post-translational modifications to obtain their active, lasso structure. The N-terminus, which is often a Gly or in some cases Cys or Ala, is suspected to be exposed following enzymatic cleavage of the leader sequence from the precursor peptide.[131] The core peptide is then enzymatically modified to form a macrolactam ring between the N-terminal amine and the carboxylate of a Glu or Asp residue at position 7, 8, or 9.[130] This yields the final, active lasso peptide, which possesses a C-terminal tail hairpin encompassed by this N-terminal macrocycle.

Microcin J25 (MccJ25), arguably the most studied lasso peptide bacteriocin, was first discovered in 1992 as a product of an *Escherichia coli* strain cultured from infant feces.[132] The peptide showed activity towards several Gram-negative bacteria, including foodborne pathogens *E. coli*, *Salmonella*, and *Shigella*. MccJ25 was found to be interacting with and inhibiting the activity of *E. coli* RNA polymerase.[133] Without this enzyme, the bacteria are unable to translate their DNA and eventually die. Unexpectedly, this mode of action is not the cause of death for *Salmonella*

newport. MccJ25 has quite a different approach for *S. newport*, interfering with the membrane permeability, disrupting the membrane potential, and producing superoxide.[134] Heightened superoxide production was later also found in *E. coli*, illustrating that MccJ25 appears to have two separate modes of action: MccJ25 interacts with the membrane or is transported into the cytosol and binds RNA polymerase.[135]

MccJ25 was initially thought to contain just 20 amino acids, have a blocked N-terminus, and found to be extremely stable to a wide range of pH (2 to 12) and high temperatures (including autoclaving at 120 ^{o}C for 15 min). [132] Subsequent work less than a decade later revealed the N-terminus was blocked due to a cyclization, at the time suspected as an N-to-C terminal or head-to-tail cyclization, and the amino acid sequence was revealed to contain 21 residues.[136] The unusual characteristics, particularly the stability of this cyclic peptide yet relatively simple amino acid primary sequence (nearly 30% of the peptide being Gly), prompted researchers to synthesized the head-to-tail peptide.[137] Surprisingly, this compound did not possess a well-defined structure in numerous solvent systems and was inactive against Gram-negative strains. Further investigations into the naturally isolated compound led to the elucidation of MccJ25's true structure (Figure 5.51).[137]

Insights obtained through proton NMR spectra analyses indicated that the N-terminal Gly_1 (black) is found in close proximity to the sidechain of Glu_8 (blue), suggesting that a backbone-to-sidechain linkage may be the true form of MccJ25. Interactions between C-terminal tail and macrocycle residues (purple) were also observed in NMR studies, supporting the idea of a threaded loop.[137] Bulky residues along the C-terminal

tail, Phe_{19} and Tyr_{20} (green), on either side of the macrocycle are suspected to act as blockers or plugs which prevent the unthreading of the lasso peptide.[137] This proposed structure of MccJ25 was supported by digestion and mass spectrometry analyses, and this lasso motif has since been found in a plethora of other lasso peptides analyzed since the initial unveiling of MccJ25's true strucure.[131]

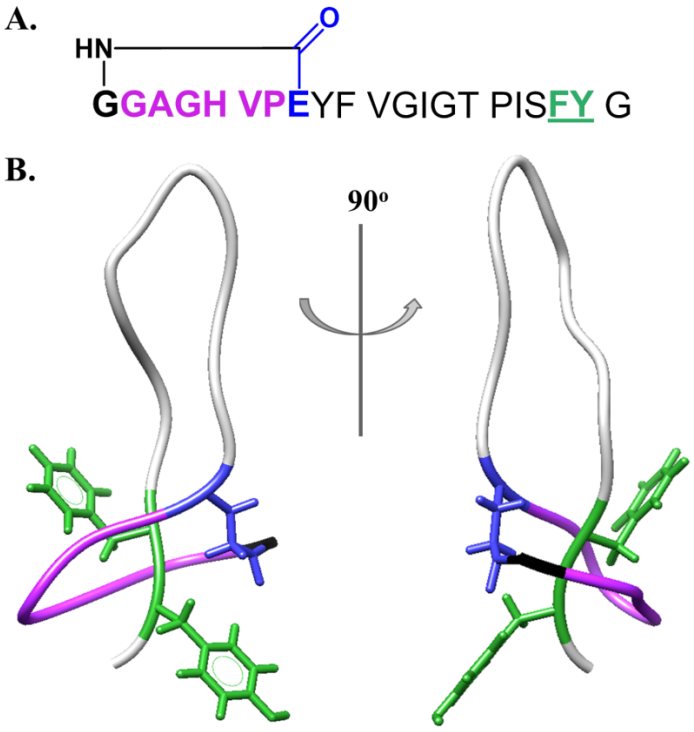


Figure 5.51 MccJ25 sequence (A) and structure (B)

The biosynthetic gene cluster of MccJ25 was found to contain four gene products, encoded for by *mcjABCD* (Figure 5.52).[138] The precursor peptide is produced from the structural gene *mcjA* as a 58 residue polypeptide, with a 37 residue N-terminal leader sequence and a 21 residue core peptide. Two enzymes, the gene products of *mcjB* and *mcjC*, are required for the production of mature, active MccJ25.[139] The linear peptide is first modified by mcjB, a cysteine protease with high homology to

transglutaminase, which cleaves the leader sequence. The core peptide is cyclized by mcjC, an ATP-binding enzyme that activates the carboxylate of Glu_8 for nucleophilic attack by the now free N-terminal amine of Gly_1.[139] The last component of this gene cluster, *mcjD*, encodes for an efflux pump which is part of the ABC transporter protein family.[138] Unlike other ABC transporters used by bacteriocins that can recognize the inactive precursor peptide, mcjD recognizes the mature, active peptide for transport. As such, this protein also acts as a self-immunity protein by expelling MccJ25 from the producer cell through active transport.

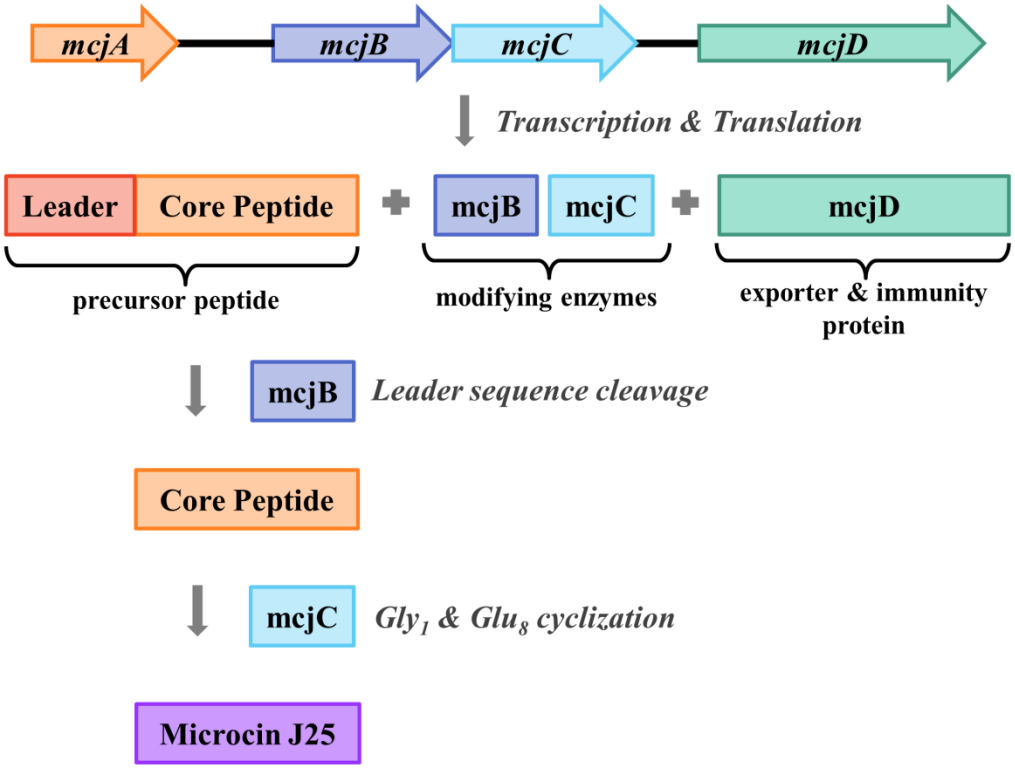


Figure 5.52 Biosynthetic pathway of MccJ25

Identification of the two modifying enzymes mcjB and mcjC, along with the rest of the gene-cluster, has invoked a herculean genome mining effort to uncover lasso peptides and other post-translationally modified peptides.[140] Close to 1500 putative

lasso peptide gene clusters have been uncovered, several which have been heterologously expressed to confirm the production and identity of the suspected lasso peptide.[141] Notably, genome mining only provides the genetic information of possible lasso peptides, it does not give any hints as to whether these peptides will have useful biological applications. Their activities and potencies can only be ascertained through expression and isolation of the predicted structural proteins.

Most efforts towards the assembly of lasso peptides with desirable activities have involved extensive genome manipulations as well as trial and error through targeted mutagenesis. Certain enzymes required in lasso peptide biosynthesis, such as mcjC from MccJ25 biosynthesis, are quite amenable to sequence changes, tolerating mutations as long as the two positions required for cyclization, Gly_1 and Glu_8, are undisturbed.[142] These mutations may lead to partially active peptides but are limited in their modification scope to the 20 canonical amino acids and existing functionalities.

The biosynthesis of a lasso peptide with modified activity has put this structural motif in the spotlight as a scaffold of interest for pharmaceutical development. Heterologous expression of a chimera MccJ25 peptide, which contained a three-amino acid mutation, produced a lasso peptide with antagonistic activity towards integrins.[143] A later example showed the fusion of lasso peptides to other proteins, to produce stable, active recombinant proteins using the native assembly enzymes.[144] Despite the tremendous advancement in genome mining efforts and recombinant expression systems,[141] the design of the peptide is ultimately at the mercy of the biosynthetic machinery of host organisms. The ability to access the lasso scaffold

through synthetic means would provide a platform to create improved analogues with diverse, noncanonical functionalities and thus more feasible therapeutic applications.

5.2 Results and discussion

5.2.1 Alternative approach to creating hairpin fold

Several studies performed on this class of intriguingly shaped peptides have given insights into the features most vital for activity and stability – the overall structural conformation of the macrocycle encompassed C-tail hairpin. Synthesis of untreated or linearized peptides, though containing the appropriate amino acid sequence, do not spontaneously fold into the desirable shape and are inactive.[145,137] Alternatively, the digestion of MccJ25 by thermolysin, a thermostable metalloproteinase which cleaves the amide bond between Phe_{10} and Val_{11}, yields a cleaved but tightly interlocked two-peptide complex.[146] This complex did not become unthreaded even in extreme thermal conditions. It also maintained partial activity, even though the binding of RNA polymerase is believed to occur around positon Val_{11} of the MccJ25 hairpin, where the peptide had been cleaved by thermolysin.[145]

Using MccJ25 as a model system, we envisioned an approach which folds the linear peptide into the desired hairpin shape (Figure 5.53). By installing a thiol group on two amino acids which are in close proximity, a disulfide bridge can be formed. This bond would bring the middle and the C-terminus of the peptide together, in the inherent hairpin conformation. Macrocyclization can then be facilitated between the N-terminal Gly_1 and the carboxylate of Glu_8, ideally encompassing the pre-folded

hairpin, to yield the desired lasso peptide shape. The final step will require desulfurization of the non-native disulfide bond, a procedure which has been extensively used in peptide synthesis,[147] to yield the native MccJ25.

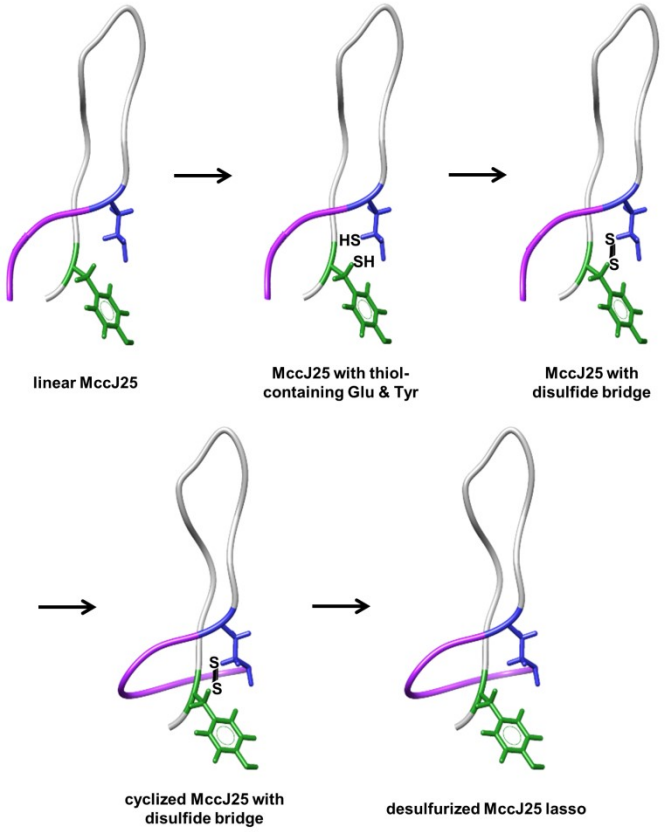


Figure 5.53 Proposed strategy of MccJ25 total synthesis

Looking at the coordinates of the NMR-derived structure of MccJ25 (PBD 1Q71),[137] Dr. Kaitlyn Towle initially analyzed possible sites of modification and identified Glu_8 and Tyr_{20} as the two most promising locations to install the disulfide bridge. Residue Phe_{19} is found too far from Glu_8 and thus adding a linkage here may distort the natural hairpin shape. Using PyMOL, several measurements and models of an inserted disulfide bridge between these two residues were taken and analyzed. Based on the distance between various atoms as well as the predicted space required for a disulfide

bond, the pro-*R*-H of the β-position of Glu_8 and the same location on Tyr_{20} were the two locations chosen for the installation of the thiol groups (Figure 5.54). Within a distance of approximately 4.8 Å, a sulfur to sulfur bond (2.05 Å) as well as the two longer carbon to sulfur bonds vs the native carbon to hydrogen bonds (1.82 Å vs 1.10 Å) could be accommodated without substantial disruption to the native hairpin shape.[149] The γ-position of Glu_8 was also considered; however, the distance was decidedly too short and this would put the disulfide bond in closer proximity to the ε-acid of Glu_8 required for lactamization. Thus, I embarked on synthesizing the required *R*-β-thiol derivatives of Tyr and Glu with Fmoc-SPPS compatible protecting groups as neither residue is commercially available.

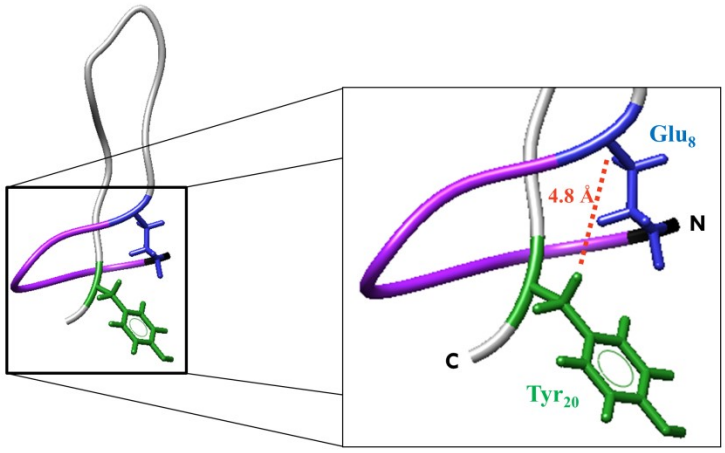
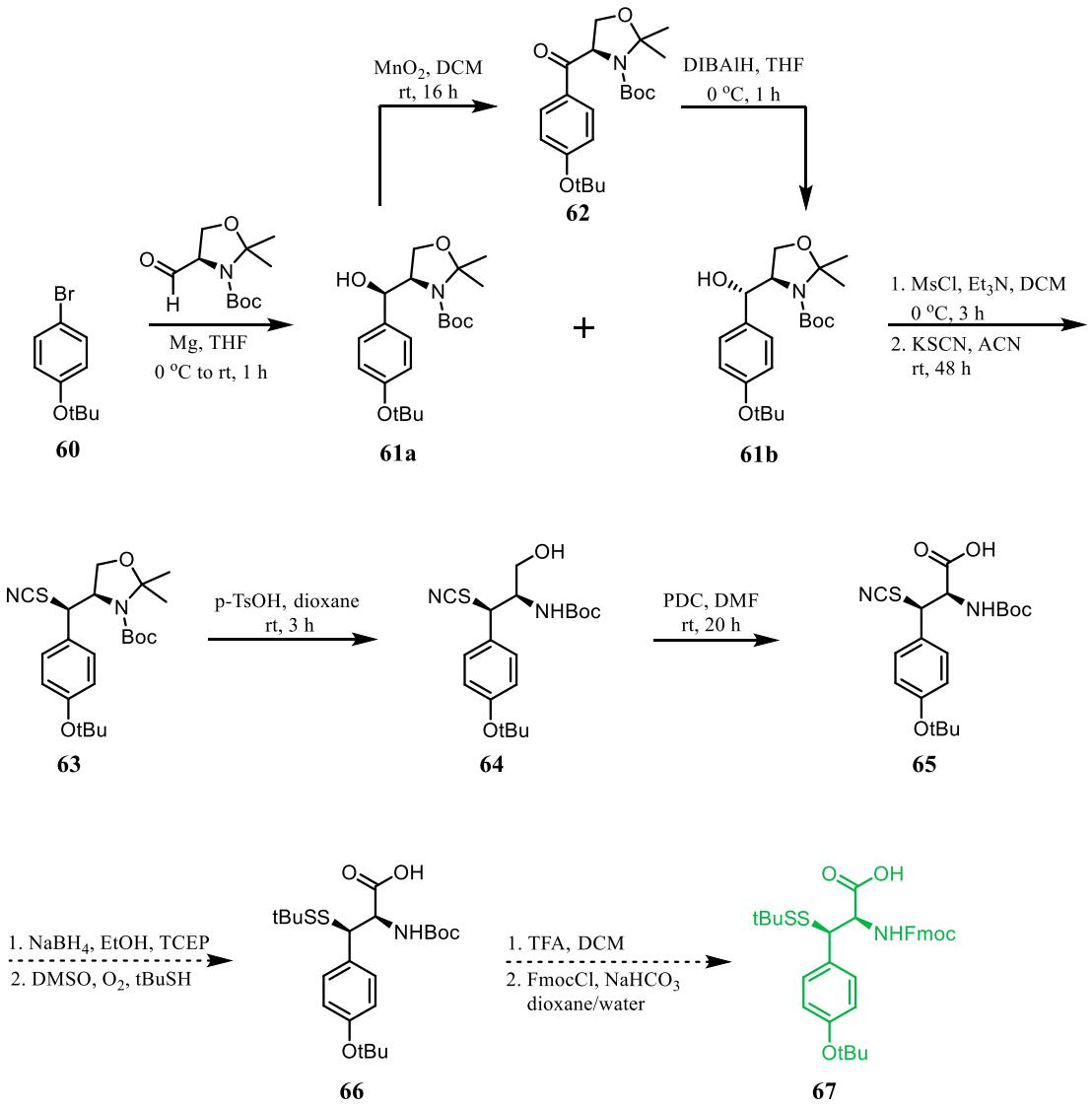


Figure 5.54 Measurements between proposed thiol insertions for MccJ25 residues

5.2.2 Synthesis of protected β-thiol Tyr

The modified amino acid β-thiol Tyr has not been previously synthesized. The closest relative to this was found in the synthesis of β-thiol Phe and this literature protocol was used to develop the synthetic scheme proposed below.[150]



DIBAlH: diisobutylammonium hydride, PDC:pyridinium chlorochromate, TCEP: tris(2-carboxyethyl)phosphine

Scheme 5.7 Proposed synthetic pathway to Fmoc-SPPS compatible β-thiol Tyr

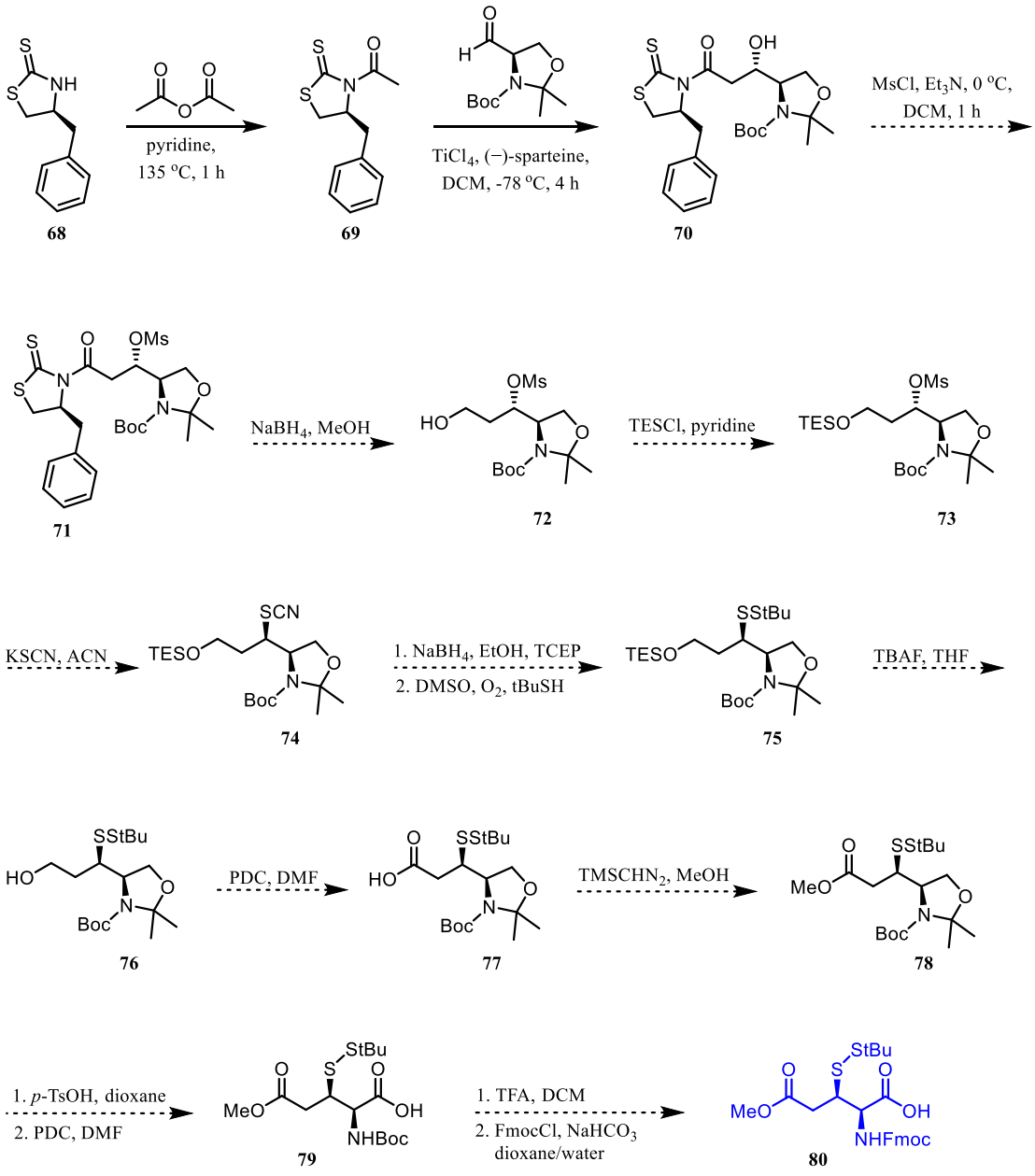
A mixture of β-hydroxyl isomers **61a** and **61b** were accessed from the reaction of Garner's aldehyde, with the *in situ* prepared Grignard reagent from 1-bromo-4-*tert*-butoxybenzene. These two isomers were separated and the undesired isomer, the *syn* β-hydroxyl **61a**, was oxidized using freshly prepared manganese(IV) dioxide to the

ketone intermediate **62**. From there, the bulky DIBAlH reducing agent was used to selectively convert the carbonyl to the desired *anti* β-hydroxyl **61b**. Mesylation of the hydroxyl created a good leaving group, allowing for the $^{-}$SCN to displace the hydroxyl and afford the protected β-thiol intermediate **63** with inversion of configuration. Initially, this substitution reaction was particularly low yielding and several conditions were attempted to optimize this step. Using *para*-toluenesulfonic acid (*p*-TsOH) to hydrolyze the five-membered ring freed the N and C termini of the β-thiol intermediate to yield **64**. This step also required optimization as the Boc-group appeared to be sensitive to *p*-TsOH. The primary alcohol was then oxidized to the carboxylic acid using pyridinium dichromate (PDC) for the synthesis of intermediate **65**. The remaining steps will involve the exchange of protecting groups found on the thiol and the amide. Specifically, orthogonal strategy of using *tert*-butylthiol as a thiol protecting group should allow for selective deprotection with phosphine and formation of disulfide while still attached to resin.[151] Lastly, the Boc protecting group could be removed and replaced with Fmoc to yield the final desired β-thiol Tyr **67** as a fully protected, Fmoc-SPPS compatible residue.

5.2.3 Synthesis of protected β-thiol Glu

The β-thiol Glu was also unknown and existing literature was consulted to propose an approach to its synthesis (Scheme 5.8).[150,152] The use of the Crimmins chiral auxiliary (*S*)-4-benzylthiazolidine-2-thione was employed as it offers good stereo control and can be cleaved under mild conditions as desired. Synthesis of the previously known *N*-acylated auxiliary **69** by acylation of (*S*)-4-benzylthiazolidine-2-thione with acetic

anhydride was achieved in excellent yields. This precursor was then linked to the (*R*)-Garner's aldehyde under a $TiCl_4$ catalyzed aldol condensation to stereoselectively access the β-hydroxyl intermediate **70**.



TCEP: tris(2-carboxyethyl)phosphine, $TMSCHN_2$: trimethylsilyldiazomethane
TESCl: triethylsilyl chloride, TBAF: tetrabutylammonium fluoride

Scheme 5.8 Proposed synthetic pathway to Fmoc-SPPS compatible β-thiol Glu

Subsequently, I propose that mesylation of the β-hydroxyl of **70** will serve to activate the hydroxyl functionality as well as differentiate it from the free hydroxyl of **72**. The Crimmins auxiliary could then be cleaved from the mesylated β-hydroxyl intermediate **71** by reduction with $NaBH_4$ to yield **72**. The primary alcohol may then be protected with TES and the secondary mesylated alcohol should be displaced by SCN to give β-thiol **74**. Exchanging the thiol protecting group with tertbutylthiol should yield the β-SStBu intermediate **75**, which could then be converted to the desired ε-carboxylic acid. The terminal alcohol could be deprotected by tris(2-carboxyehtyl)phosphine (TBAF), followed by oxidation to the carboxylic acid with PDC to afford intermediate **77**. The carboxylic acid may be alkylated with $TMSCHN_2$/MeOH to give a protected β-thiol intermediate **78**. Next, the deprotection of the amino acid could be achieved using the same method as the previously discussed for Tyr. *p*-TsOH could be used to cleave the oxazolidine ring followed by and PDC oxidation of the free alcohol to the α-carboxylic acid to access the Boc-protected β-thiol compound **79**. The Boc-protected amine could be deprotected and protected with FmocCl as outlined previously to yield Fmoc-SPPS compatible fully protected β-thiol Glu **80**.

5.3 Conclusions and future directions

MccJ25 along with other compounds in the lasso peptide family have gained substantial interest in the last decade. Their unique three-dimensional fold and interlocked structure provide a scaffold with potent activity and high stability. The resistance to proteases and high temperatures are attractive features from a

pharmaceutical and food industry perspective. Using MccJ25 as a model system, a novel approach to its total synthesis was undertaken. This method takes advantage of a bond commonly found in peptides and proteins, a disulfide bridge, to form the hairpin shape of the lasso peptide. This would ensure the peptide is pre-folded into its mature form before attempting the macrocyclization.

Based on the structural orientation and distances between atoms in the NMR structure of MccJ25, residues Glu_8 and Tyr_{20} were chosen as the amino acid residues onto which to install the necessary thiols for a disulfide bridge formation. Specifically, it was determined that the optimal location is the β-carbon of both residues and that R-configuration at this position would position the two thiols facing each other. Synthetic schemes for the synthesis of the two required β-thiol amino acids, Tyr and Glu, were proposed as these compounds are not commercially available and were not synthesized previously. The first several compounds of the pathway were synthesized and the methods were optimized for the amino acids of interest; syntheses outlined in the remaining steps are currently underway.

With these two residues in hand, the peptide can be synthesised on resin using Fmoc-SPPS chemistry. The β-thiol groups of Glu_8 and Tyr_{20} can be orthogonally deprotected and the disulfide bridge can be formed on resin. Cleavage off resin and lactamization of Gly_1 and Glu_8 should afford a bicyclic compound, which may possess partial activity if the overall fold is homologous to native MccJ25. Lastly, desulfurization of the peptide should afford the thiol-free native MccJ25 as a product of total synthesis.

The difficulty in handling the Tyr compounds prompts the consideration of an alternate model system. For example, the lasso peptide caulonodin IV, a peptide identified in the genome of *Caulobacter* sp K31,[153] contains a Phe residue in place of Tyr at this position. β-thiol Phe has been successfully synthesized previously and thus could provide a more direct route to testing if our disulfide hairpin approach is feasible.

Chapter 6: Summary and future outlook

The world of natural products has provided scientists with a wealth of fascinating structures and functionalities. Investigating these compounds has led to great discoveries that have contributed to the betterment of society. Antimicrobial peptides are a subset of great interest as they may prove to be the compounds employed to tackle the ever-growing antimicrobial resistance seen throughout the globe. Four peptides with potential in this area were investigated in this thesis.

Faerocin MK was studied for its possible application as a potent anti-listerial peptide, specifically of interest as it was discovered in the genome of a probiotic strain of Enterococci. Based on sequence homology to YGNG-motif containing bacteriocins, it was predicted this peptide may belong to the same class. Activity and structural studies were conducted to confirm that this peptide is active against Gram-positive bacteria and possesses an α-helical conformation. Additionally, a previously reported peptide with high homology, SRCAM 602, was synthesized and tested using the same methods to uncover that it was inactive and unstructured. Structure elucidation using NMR was attempted; however, due to peak overlap and low yielding expression in labelled media, this was unsuccessful. The operon of the relevant FaeMK genes was cloned and heterologously expressed to identify the immunity gene (*faeI*) as well as the parent core peptide (*faeMK*). This led to the deduction that FaeMK is excreted by its producer and heterologous hosts through the general *sec*-pathway. Further studies could look at the genome of *Enterococcus faecium* M3K31, which contains two other bacteriocins, and test its use as a probiotic strain in the food production industry for prevention of listeriosis, a Listeria-born gastrointestinal infection.

Teixobactin is a NRP discovered owing to the development of a novel way of culturing soil bacteria. The potent activity against Gram-positive pathogens and unusual structural features drew the interest of researchers to further study this peptide and its mechanism of action. Through thermodynamic measurements obtained from isothermal titration calorimetry, the binding of Txb and four analogues to lipid II and several truncated analogues was assessed. It was uncovered that Txb recognizes and binds the lipid-pyrophosphate moiety of LII, leading to its promiscuity towards other lipid phosphate intermediates in the biosynthesis of the peptidoglycan layer. Regrettably, due to solubility and capability issues, the structure elucidation of this complex was not possible with solution-phase NMR. Additionally, Txb showed comparable affinity for LII-Lys and LII-DAP, the lipid variants found in most Gram-positive bacteria and a majority of Gram-negative bacteria, respectively. This led to the expansion of Txb's activity spectrum, using outer membrane-disruptive peptides (unacylated tridecaptin A_1 and polymyxin B nonapeptide) to include Gram-negative bacteria such *S. enterica* and *K. pneumoniae.* Future work may investigate alternative application of this peptide as a topical antimicrobial. For example, the observed formation of amyloid gels, while not compatible with a mammalian system, could be used to coat high contact surfaces in hospitals or food-preparation facilities to prevent the spread of susceptible pathogens.

Tridecaptin A_1 is a NRP extensively studied by the Vederas group throughout the last few years following its isolation from *Paenibacillus terrae*. TriA_1 is active against Gram-negative bacteria listed as ESKAPE pathogens of concern by the World Health Organization as well as common foodborne pathogens such as *Escherichia coli*. It is a

selective binder of Gram-negative lipid II and undergoes a conformational change when interacting with its target in DPC micelles. NMR structure elucidation studies were conducted with tridecaptin, lipid II, and a peptide-lipid complex in the presence of DPC micelles formed by oxygen-labelled DPC_{d36} analogues. The strategy of using oxygen-enriched DPC analogues to study membrane interacting peptides and their receptors could be extended to other AMPs such as leucocin A.

Microcin J25 is a staple example of lasso peptides, a class of structurally unique peptides found in bacteria. It is a potent inhibitor of *E. coli* and can tolerate several substitutions along its sequence, while still maintaining its unique lasso structure. Completion of the two modified amino acid synthesis is currently underway. The linear peptide can be assembled on resin through Fmoc-based SPPS and the thiols can be orthogonally deprotected to form the disulfide bridge. Lactamization and subsequent desulfurization will be optimized to form the macrocycle and then remove the additional disulfide to yield the fully synthetic MccJ25. If successful, this approach could be extended to other lasso peptides as well as be used to synthesize lasso peptides with noncanonical amino acids in the future.

This thesis covered four examples of potent antimicrobial peptides and the advances made towards understanding their behaviours, elucidating their structural conformation, and developing a method for their synthesis. A wide variety of interdisciplinary approaches were utilized in hopes of providing a full picture of this interesting class of molecules. Insights into the structural features and binding properties of these compounds continue to provide vital information in the quest to develop potent and safe antibiotics for widespread applications in our society.

Chapter 7: Experimental methods

7.1 General synthetic procedures

7.1.1 Reagents and techniques

Compounds used were purchased from Millipore-Sigma Canada Ltd., Fisher Scientific Ltd., Chem-Impex International Inc, VWR International, Bio-Rad Laboratories Inc., Cambridge Isotope Laboratories, and used as delivered unless the method states otherwise. All solvents were of American Chemical Society (ACS) grade and used without further purification. For moisture-sensitive reactions, solvents were distilled under positive argon pressure immediately prior to use. DCM, dioxane, ACN were distilled over calcium hydride, THF and diethyl ether were distilled over sodium with benzophenone indicator, and MeOH was distilled over magnesium.[154] Solvents used for liquid chromatography were filtered and degassed before use. Filtered water for aqueous reactions was obtained from a MQ water purification system (Millipore Co., Milford, MA). Solutions of $NaHCO_3$, NH_4Cl, HCl, NaOH, brine (sat. NaCl), Tris (tris(hydroxymethyl)aminomethane), NH_4HCO_3, and NaH_2PO_4 refer to aqueous solutions. For air and moisture sensitive reactions, a positive pressure system using flame-dried glassware was prepared and all solvents used were distilled as mentioned previously. Organic solvents were removed *in vacuo* utilizing a Büchi rotary evaporator and aqueous mixtures were concentrated by lyophilization of flash-frozen samples.

7.1.2 Compound purification: Flash chromatography

Flash column chromatography was performed utilizing silica gel (SiliaFlash® Irregular Silica Gel, P60, 40 - 63 μm, 60 Å) and fractions were monitored using thin-layer chromatography (TLC) with silica-coated glass plates (500 microns, Merck 60 F254). Visualization was done with UV absorption and, where required, by staining with permanganate ($KMnO_4$:K_2CO_3:NaOH:H_2O, 1.5 g: 10 g: 120 mg: 200 mL), phosphomolybdic acid in EtOH (10 g in 100 mL), or *para*-anisaldehyde in EtOH (H_2SO_4:*p*-anisaldehyde, 2.5 mL: 15 mL in 250 mL) and heating with a heat gun.

7.1.3 Compound purification: High-performance liquid chromatography

Reverse-phase high-performance liquid chromatography (RP-HPLC) was performed on one of two systems optimized for i) analytical scale and ii) preparative or semi-preparative scale: i) was performed on a Gilson chromatograph equipped with model 322 pumps, a 7725i Rheodyne injector with a 1 mL sample loop, a 171 diode array detector, and a FC 203B fraction collector and ii) was performed on Gilson chromatograph with model 322 pumps, UV/VIS-150 detector, GX-271 liquid handler.

Method A: Was conducted with a Phenomenex C_{18} column at a 10 mL/min flow rate; UV detection monitoring at 220 nm. Gradient elution was obtained using eluent A: MQ H_2O (0.1% TFA) and eluent B: ACN (0.1% TFA). Starting from 20% eluent B for 8 min, ramping up to 55% B over 30 min, ramping up to 95% B over 3 min, maintaining 95% B for 3 min, and ramping back down to 20% B over 2 min, maintaining 20% B for 5 min.

Method B: Was conducted with a GraceVydac Protein and Peptide C_{18} column (100 mm, 10 micron) at a 10 mL/min flow rate; UV detection monitoring at 220 nm. Gradient elution was obtained using buffer A: MQ H_2O (50 mM NH_4HCO_3) and buffer B: MeOH. Starting from 100% buffer A for 8 min, ramping up to 100% B over 35 min, maintaining 100% B for 3 min, and ramping back down to 0% B over 2 min, maintaining 0% B for 5 min.

Monitoring of HPLC fractions was achieved with the use of mass spectrometry. Small molecule compounds were analyzed using LCMS and larger molecules and peptides were analyzed using MALDI-TOF. For MALDI, fractions with high organic content were concentrated by solvent evaporation under vacuum until only 20% or less organic solvent remained. Next, a 0.1% TFA in water solution was prepared and a two-layer matrix method was used. Smaller, non-hydrophobic compounds were analyzed using α-cyano-4-hydroxycinnamic acid (HCCA) while larger, hydrophobic compounds were analyzed with the help of 3,5-dimethoxy-4-hydroxycinnamic acid, commonly called sinapinic acid (SA). Briefly, a thin layer of the matrix 1 (10 mg/mL SA in 4:1 Ace:MeOH or 10 mg/mL HCCA in 9:1 MeOH:H_2O) was pipetted onto a well of the stainless steel plate and allowed to dry. Next, a 1:1 solution of compound:matrix 2 (10 mg/mL SA in 1:1 ACN:H_2O (0.1% TFA) or 10 mg/mL HCCA in 1:2 MeOH:H_2O (0.1% TFA)) was mixed and then pipetted overlaying matrix layer 1.

7.1.4 Compound purification: Ion-exchange chromatography

Compounds with phosphate or pyrophosphate functional groups were purified using an ion-exchange chromatography technique in lieu of silica flash chromatography. Resin (DOWEX 50WX8, 100-200 mesh resin) was presoaked for 1 h in DI water; approximately 1 g of resin was used for 10 mg of the crude mixture to be separated. The resin was then transferred to a glass column with a faucet bottom and peristaltic pump. The column was flushed with water (5 mL/min flow rate) until elute was clear and reached a neutral pH of 7.0. The resin was charged with ammonium counter ions by flushing with NH_4OH (4 × column volume) and then brought back to neutral pH using water. The elution buffer (49:1 25 mM NH_4HCO_3:IPA) was used to prime the column (2 × column volume) and the flow rate was adjusted to 1 mL/min. Crude sample was loaded and appropriate sized fractions were collected. Fractions containing compound of interest was frozen and lyophilized to yield the ammonium salt of the desired compound.

7.1.5 Compound characterization: NMR, IR, MS, OD

Nuclear magnetic resonance (NMR) spectra were obtained on a Varian Inova 400, Varian Inova 500, Varian Inova 600 MHz (equipped with a triple-resonance HCN probe), or Agilent VNMRS 700 MHz spectrometer (equipped with a triple-resonance HCN cold probe and a z-axis pulsed-field gradients) at 27 ^{o}C unless otherwise indicated. Chemical shifts are reported in parts per million (ppm) and spectra were referenced to the residual proton and carbon resonances or internal standards. ^{1}H NMR chemical shifts were referenced to: $CDCl_3$ δ 7.26, $toluene_{d8}$ δ 2.09, $DMSO_{d6}$ δ

2.50 and CD_3OD δ 3.30; ^{13}C NMR chemical shifts were referenced to: $CDCl_3$ δ 77.06, $toluene_{d8}$ δ 20.4, $DMSO_{d6}$ δ 39.5, and CD_3OD δ 49.1. For spectra in aqueous buffers, DSS or TMS were added as internal standards (spectra were referenced to δ 0.0 ppm); water signal was suppressed by excitation sculpting (^1H, COSY, TOCSY, NOESY) or through presaturation during the relaxation delay (HSQC, HMBC).[155]

Mass spectra were recorded on an Agilent Technologies 6220 orthogonal acceleration time-of-flight (TOF) instrument (Santa Clara, CA) equipped with +ve and –ve ion electrospray ionization (ESI) source, and full-scan, high-resolution MS with two-point lock mass correction. Samples were separated on an Agilent Technologies 1200 SL HPLC system with a C_8 or C_{18} analytical column prior to injection into the mass spectrometer. Infrared spectra (IR) were recorded on a Nicolet Magna 750 instrument. Cast film refers to the evaporation of a solution on a NaCl plate. Optical rotations were measured on Perkin Elmer 241 polarimeter with a microcell (10 cm, 1 mL) at 23 ^oC unless otherwise stated.

7.1.6 Solid-phase peptide synthesis

Solid-phase peptide synthesis was completed on a PreludeX Peptide Synthesizer (Gyros Protein Technologies, Uppsala, Sweden) with nitrogen flow and UV monitoring capabilities. Resin was preloaded with the first amino acid (0.1 mmol loading) and proceeding amino acids were prepared as Fmoc-protected solutions (0.2 M) in DMF. Peptide coupling was achieved using PyBOP (0.99 equiv to aa), HOBt (0.99 equiv to aa) and DIPEA (2 equiv to aa) in DMF (2 h shake, rt). Unextended chains were end-capped using a solution of acetic anhydride and DIPEA in DMF

(1:1:3, 30 min shake, rt). Deprotection of Fmoc groups was achieved using 20% piperidine in DMF; UV monitoring (301 nm) confirmed deprotection efficiency by the formation of side product dibenzofulvene-piperidine. A wash of DCM (3 × 5 mL) and DMF (3 × 5 mL) was employed in between each step.

Peptides were cleaved from resin by shaking for 1 h at rt in a TFA:TIPS:H_2O (95:2.5:2.5) solution. Crude mixture was filtered through glass wool into an Eppendorf tube (1.5 mL), rinsed with TFA:TIPS:H_2O solution, and concentrated *in vacuo*. Peptides were precipitated with cold diethyl ether (2 × 1 mL) and tubes were briefly centrifuged (13 200 rpm, 1 min) to pellet the peptide material. The solvent was evaporated and crude peptide was resuspended for further HPLC purification.

7.2 General biological procedures

7.2.1 Media preparation

Unless stated otherwise, all media and media supplements were purchased from Becton Dickinson and Company (BD, Franklin Lakes, NJ). Sterile conditions were used when handling biological experiments; materials were either filter-sterilized (0.22 μm) or autoclaved (121 ^{o}C, 15 min). Liquid or broth media was prepared following the manufacturer's recommendations. Soft agar was produced with 0.75% (w/v) agar dissolved in appropriate broth media and aliquoted into a 5 mL per test tube prior to autoclaving the whole set. Solutions were stored in a 4 ^{o}C fridge until use, when they were warmed to 25 ^{o}C prior to being inoculated. Hard agar was produced with 1.5% (w/v) agar dissolved in appropriate broth media and autoclaved

as a stock solution. The solution was cooled to 60 ^{o}C and approximately 15 mL was dispensed into each sterile Petri dish inside a laminar flow cabinet and allowed to dry and harden for 20 min. Plates were stored inverted in a 4 ^{o}C fridge until use, when they were warmed to optimal bacteria growth temperatures for 30 min.

7.2.2 Glycerol stocks

Glycerol stocks were used to maintain easily accessible bacterial strains. All bacteria strains, indicators, and producers, were stored in -80 ^{o}C freezers as 20% glycerol solutions. Plastic tubes (2 mL) with screw caps containing 80% glycerol (0.8 mL) were autoclaved and allowed to cool. To this tube, a fresh liquid bacterial culture (0.2 mL) was added to a total volume of 1 mL and mixed to give a 20% glycerol stock solution. Tubes were flash-frozen in liquid nitrogen and stored in -80 ^{o}C freezers long term. Inoculation of fresh cultures of desired bacteria was achieved by the aseptic transfer of a bit of frozen stock into fresh liquid media. Thawing of the stocks was limited by using a cooling block to store tubes while in use outside the freezer.

7.2.3 Bacterial culture growth

In general, Gram-positive strains were grown in All Purpose Tween medium (Difco, Sparks, MD) at 25 ^{o}C for 24 h and Gram-negative strains were grown using Luria-Bertani Broth Base, Miller medium (Difco) at 37 ^{o}C for 24 h. *Staphylococcus aureus* strains were grown on tryptic soy broth medium (Difco), *Acinetobacter baumanii* was grown at 30 ^{o}C, and *Klebseilla pneumoniae* was grown for 48 h. Growth conditions

for the various bacterial species are shown in Table 7.10 below. If required, chloramphenicol (5 μg/mL for *Carnobacterium maltaromaticum* and 10 μg/mL for *Enterococcus faecium*) or kanamycin (50 μg/mL for *Escherichia coli*) were added to the media for selective growth of desired plasmid-containing bacteria.

Table 7.10 Growth conditions for indicator bacterial strains

Bacterial Strain	Growth Details	
	Media	**Conditions**
Gram-positive organisms		
Brochothrix thermosphacta ATCC 11509	APT	25 °C, 0 rpm, 24 h
Carnobacterium divergens LV13	APT	25 °C, 0 rpm, 24 h
Carnobacterium maltaromaticum UAL26	APT	25 °C, 0 rpm, 24 h
Enterococcus faecalis ATCC 7080	APT	25 °C, 0 rpm, 24 h
Enterococcus faecium BFE 900	APT	25 °C, 0 rpm, 24 h
Latilactobacillus sakei 706	APT	25 °C, 0 rpm, 24 h
Lactococcus lactis ssp. *cremoris* HP	APT	25 °C, 0 rpm, 24 h
Listeria monocytogenes ATCC 15313	APT	25 °C, 0 rpm, 24 h
Listeria monocytogenes UAFM1	APT	25 °C, 0 rpm, 24 h
Staphylococcus aureus ATCC 25923	*TBS*	25 °C, 0 rpm, 24 h
Gram-negative organisms		
Acinetobacter baumanii ATCC BAA-1605	LB	*30 °C*, 225 rpm, 24 h
Acinetobacter baumanii ATCC 19606	LB	*30 °C,* 225 rpm, 24 h
Escherichia coli ATCC 25922	LB	37 °C, 225 rpm, 24 h
Escherichia coli DH5α	LB	37 °C, 225 rpm, 24 h
Klebsiella pneumoniae ATCC 13833	LB	37 °C, 225 rpm, *48 h*
Pseudomonas aeruginosa ATCC 14207	LB	37 °C, 225 rpm, 24 h
Pseudomonas aeruginosa ATCC 15442	LB	37 °C, 225 rpm, 24 h
Salmonella enterica serovar Typhimurium ATCC 13311	LB	37 °C, 225 rpm, 24 h
Salmonella enterica serovar Typhimurium ATCC 23564	LB	37 °C, 225 rpm, 24 h

7.2.4 Deferred inhibition assays

A liquid culture of desired bacteria was grown overnight and a portion of the liquid culture was transferred onto the hard media by dipping a sterile toothpick in the culture and stabbing the hard agar media. The plate was incubated overnight at the

appropriate temperature. A 5 mL liquid media broth containing desired indicator strain was grown overnight as well. A 100 μL aliquot of culture was transferred to a test tube containing 5 mL of molten soft agar media. The mixture was gently vortexed, poured over top of the stabbed agar plate, and allowed to cool and harden. Plates were incubated overnight at the desired temperature and inhibition zones were observed and measured (in mm) the following day.

7.2.5 Spot-on-lawn assay

Peptide activity was tested in one of two assays. Spot-on-lawn tests were performed using a two-fold series dilution of the peptide of interest. In most cases, the solvent of choice was sterile MQ water. Indicator strains were grown as outlined above; then 100 μL of each strain was used to inoculate soft agar (0.75% w/v) and overlaid on top of hard agar (1.5% w/v). Starting from a stock solution of peptide of interest, the two-fold series dilution was performed to create several different concentrations. 10 μL of each concentration was aliquoted onto the agar plate and the plates were incubated as appropriate for the indicator strain. This was repeated in triplicate to give minimum inhibitory concentrations (MICs) as determined by zones of inhibition observed at the corresponding concentration.

7.2.6 Microdilution assay

The second method to determined MICs was the use of a modified broth microdilution protocol from the Clinical and Standards Laboratory Institute.[156]

Antimicrobial peptides were dissolved in Muller-Hinton Broth (MHB) media and serial dilutions were made across a 96-well plate. In the case of synergistic assays, dilutions were made in the x-axis for peptide A and in the y-axis for peptide B. Control wells containing no antibiotics were also used for each trial. One plate was used for a particular indicator organism and each well was inoculated to reach a final inoculum of 5×10^5 colony forming units per mL. Using OD_{600} readings normalized to a negative control, MICs were recorded as the lowest concentration at which no growth was detected after 24 h, or 48 h for *K. pneumoniae*, incubation. Each assay was performed in triplicate and average MICs are reported. Absorbance readings were taken on a SpectraMax i3x Plate Reader (Molecular Devices, Sunnyvale, CA).

7.2.7 Cell harvesting

Bacterial cells were grown in appropriate media for a desired period of time prior to harvesting. Cell broth was centrifuged (JA-10, 6300 rpm, 5 min, 4 ^{o}C); the supernatant was discarded and cell pellets were collected and weighed. Cells were resuspended in lysis buffer (50 mM NaH_2PO_4, 500 mM NaCl, 10 mM imidazole, 1% glycerol, 0.5% triton X-100, pH 8.0 and lysozyme, DNAse, and protease inhibitors) with 10 mL buffer/1 g pellet.

7.2.8 Cell lysis by cell disruptor

Resuspended cells were lysed through one of two methods. A TS series benchtop cell disruptor (Constant Systems Ltd, Low March, UK) was used to disrupt cell

membranes at high pressure. The system was primed with lysis buffer (20 mL) at 5 kpsi. Next, the bacteria cell mixture was passed through the disruptor at 25 kpsi, twice, and collected in a Falcon conical centrifuge tube (50 mL). Additional lysis buffer (10 mL) was used to wash the lines and also collected into the same tube. Lysate was centrifuged (JA-20, 15000 rpm, 30 min, 4 ^{o}C); supernatant was decanted into a fresh tube and pelleted debris was discarded.

7.2.9 Cell lysis by sonication

When the cell disruptor was unavailable, cells were lysed using sonication by a Branson Sonifier 450 (Branson Ultrasonics Corporation, Danbury, CT). The resuspended cell mixture was subjected to sonication (30 s) followed by cooling on ice (30 s). This cycle was performed thrice and contents were centrifuged (JA-20, 15000 rpm, 30 min, 4 ^{o}C). Supernatant was decanted into a fresh tube and pelleted debris was discarded.

7.2.10 C8 and C18 solid-phase extraction cartridges

Bond Elut C8 (10 g, 60 mL Agilent, Mississauga, ON) and Strata C18-E (20 g, 60 mL, Phenomenex, Torrance, CA, USA) cartridges were used as per manufacturer recommendations to perform solid-phase extractions. Prior to sample loading, the cartridges were activated with methanol (50 mL) and primed with water (2 × 50 mL). A peristaltic pump (Bio-Rad) was used to control flow rate, which was kept at 10 mL/min throughout.

7.2.11 Tris-glycine SDS-PAGE

Two types of gel, 4 – 20% or 12% Mini-PROTEAN TGX Stain-Free precast gels (Bio-Rad, Mississauga, ON), were used to analyze the contents of protein mixtures. Protein samples and standers were mixed with 2× Laemmli Sample Buffer (Bio-Rad) and boiled at 100 ^{o}C for 5 min using a water bath. Samples were loaded into individual gel lanes using elongated micropipette tips. Protein molecular weight ladders of appropriate size ranges were loaded into the last lane and used as markers. Gels were run in TGX buffer (Bio-Rad) at 180 V for 30 min, or until the smallest ladder marker reached the bottom of the gel. The gels were removed and rinsed with MQ water (3 × 10 mL). They were visualized with coomassie brilliant blue stain (0.1% w/v coomassie R-250, 40% ethanol, 10% acetic acid) for 1 h followed by destaining (10% ethanol, 7.5% acetic acid) over night.

7.2.12 Tris-tricine SDS-PAGE

Samples with peptides smaller than 10 kDa were analyzed using a gel made in-house. A 15% (w/v) acrylamide Tris-tricine polyacrylamide gel with a 4% stacking gel was cast for these samples. Samples were mixed with Tris-triscine Buffer (Bio-Rad) boiled at 100 ^{o}C for 5 min using a water bath. Samples and molecular weight ladder markers were loaded into individual lanes. The gel was run at 70 V for 15 min to compress the bands through the stacking gel and then at 120 V for 45 min, or until the smallest ladder marker reached the bottom of the gel. The gels were placed in fixing solution (50% methanol, 40% water, 10% acetic acid) for 1 h and stained with coomassie brilliant blue as described above.

7.2.13 Plasmid isolation

Plasmids were isolated using either a QIAprep Spin Miniprep Kit (Qiagen, Toronto, ON) or a GeneJET Plasmid Miniprep Kit (Fermentas Canada Inc, Burlington, ON). Isolations were performed according to manufacturer's methods and kept at 4 ^{o}C throughout. The plasmids were eluted with Tris-ethylenediaminetetraacetic acid (EDTA) pH 8 buffer and stored at -20 ^{o}C.

7.2.14 Restriction enzyme digest

FastDigest restriction enzymes (ThermoFisher, Waltham, MA, USA) were used according to manufacturer's recommendations. NA digestion was performed in 100 µL containing FastDigest buffer (10×, 10 µL), restriction enzymes (*SacI* and *XbaI*, 2.5 µL each), DNA of interest (20 µL for inserts, 10 µL for vectors), and sterile MQ water (up to 100 µL total volume). Digest mixtures were incubated at 37 ^{o}C for 2 h and DNA was purified using QIAquick PCR Purification Kit (Qiagen). DNA was concentrated *via* ethanol precipitation.

7.2.15 DNA ligation

DNA ligation was achieved using T4 DNA ligase (Invitrogen) as per manufacturer's protocol. A molar ratio of 3:1 insert:vector DNA concentration was used in a volume of 10 µL. Ligation mixtures were incubated at rt for 1 h or at 4 ^{o}C overnight. Resulting plasmids were transformed into desired bacterial cell lines through electroporation or heat shock treatment.

7.2.16 Ethanol precipitation

To a sample containing DNA of interest, buffer (0.5 vol eq, 3 M sodium acetate, pH 5.2) and ice-cold ethanol (2.5 vol eq, 95%) were added, vortexed, and incubated at 4 ^{o}C overnight. The solution was centrifuged (16 000 g, 20 min, 4 ^{o}C) and supernatant was carefully decanted. To the remaining pellet, ethanol (2.5 vol equiv, 70%) was added and the tubes were vortexed. The mixture was centrifuged (16 000 g, 15 min, 4 ^{o}C), the ethanol was decanted, and the pellets were air-dried at rt for 45 min.

7.2.17 Transformation of chemical competent cells

Frozen tubes containing 50 µL of competent *E. coli* cells (New England Biolabs, Ipswitch, MA) were thawed on ice. The appropriate DNA plasmid (2 µL) was transferred, gently mixed, and incubated on ice for 30 min. Next, the tubes were transferred to a 42 ^{o}C water bath for 30 s and then returned to ice for 2 min. Pre-warmed (37 ^{o}C) Super Optimal Broth (250 µL) was added to each tube and the mixture was incubated at 37 ^{o}C for 45 min. Cell cultures were streaked onto LB hard agar plates, which contained the corresponding antibiotic for selective growth of desired plasmid-containing transformants.

Electroporation was also used to transform *C. maltaromaticum* and *E. faecium* cells. Fresh cultures of *C. maltaromaticum* and *E. faecalis* cells were prepared the day before and grown overnight in APT (2% w/v glycine) media. A small aliquot (50 µL) of culture was transferred to an autoclaved Eppendorf tube on ice containing desired DNA plasmid (2 µL). Solutions were transferred to 2 mm cuvettes and placed on ice for 5 min. Individual 2 mm cuvettes were removed from ice, wiped dry, and placed

within the electroporation system (Bio-Rad). A pulse program for 12 ms, 1000 V, 25 μF, 800 Ω was used. Sucrose-enriched media (0.5 M sucrose, 10% w/v glycine, 1 mM $MgCl_2$, 5 mM potassium phosphate, pH 6) was added to the cuvette, mixed gently using a micropipette, and the whole mixture was transferred to a sterile Eppendorf tube. Cells were incubated at the strains' corresponding temperatures for 4 h and then plated as outline above.

7.2.18 Polymerase chain reaction

DNA amplification was achieved by polymerase chain reaction (PCR) using a Techgene Thermal Cycler (Techne Inc., Burlington, NJ). Platinum Tag polymerase (Invitrogen, Carlsbad, CA) was used as outlined by manufacturer instructions. In general, the recipe for PCR mixtures: template DNA (1 μL), forward and reverse primers (1 μL each), PCR buffer (10×, 10 μL), dNTP (10 mM, 2 μL), $MgCl_2$ (2 μL), polymerase (1 μL), sterile MQ water (up to 100 μL total volume). The solution was mixed and placed on ice until amplification was started. A typical amplification cycle involved a denaturing step (45 s, 94 ^{o}C), annealing step (30 s, 55 ^{o}C), and extension step (1 min, 72 ^{o}C). This was repeated for 25 cycles followed by a final extension step (5 min, 72 ^{o}C).

7.2.19 Agarose gel electrophoresis

Agarose gels (1.2% w/v) were prepared using Ultra-Pure Agarose (Invitrogen) and in Tris-borate-EDTA (TBE). Stain used was SYBR Safe DNA gel stain (Life

Technologies, 1/10 000 dilution in DMSO). DNA samples were incubated with 6× sample buffer (ThermoFisher) for 3 min and loaded onto individual lanes; a molecular weight DNA ladder marker was added in the last lane. Gels were run at 84 V for 25 min and bands were visualized using an ImageQuant RT ECL Imager (GE Healthcare Life Sciences, Little Chalfont, UK) or a Dark Reader Transilluminator (Clare Chemical Research, Dolores, CO).

7.2.20 DNA & protein quantification

Concentration of DNA mixtures was measured using an Implen NanoPhotometer P360 (Implen Inc, Westlake Village, CA) at 260 nm. Purity from protein contaminants obtained during digests or ligation was quantified using the 260:280 nm absorbance reading ratios.

Concentrations of protein were measured using the same machine at 280 nm, with a path length of 1 cm, and molar extinction coefficients (ε) were calculated based on sequence using the ProtParam program from ExPASy Proteomics.[157] Concentration of each sample was determined using Beer's Law ($c=A/b\varepsilon$, where c: concentration in mol/L, A: UV absorbance recorded at 280 nm, b with path length in cm, and ε:molar extinction coefficient in $M^{-1} \times cm^{-1}$).

7.2.21 DNA Sanger sequencing

DNA sequencing was performed at the University of Alberta's Molecular Biology Service Unit. A BigDye Terminator v3.1 Cycle Sequencing Kit (Applied Biosystems,

Foster City, CA) was employed for sequencing. A typical cycle involved a denaturing step (40 s, 96 ^{o}C), annealing (15 s, 50 ^{o}C), and extension (1 min, 60 ^{o}C) and usually 30 cycles were performed for sequencing reactions. Ethanol precipitation was used on the resulting mixture and the purified DNA was analyzed using an ABI 3720 DNA Analyzer (Applied Biosystems).

7.2.22 Bioinformatics and structural software

DNA and protein data were analyzed using a series of bioinformatics tools. Predicted and related products of operons were compared using National Center for Biotechnology Information's Basic Local Alignment Search Tool (BLAST, https://blast.ncbi.nlm.nih.gov/Blast.cgi). Sequences of DNA and amino acids were aligned using Clustal Omega (http://www.clustal.org/omega/).[82] Protein and peptide structures were imaged using University of California - San Francisco's Chimera (https://www.cgl.ucsf.edu/chimera/) software or PyMOL[158] and previously determined structures were extracted using their Protein Data Bank (PDB) codes. Peptide properties were obtained by imputing sequence information into Swiss Bioinformatics Resource Portal's ProtParam software on their ExPASy Proteomics server (https://web.expasy.org/protparam/).[157] NMRPipe[159] and NMRView[160] were used to process raw NMR fid data and perform peak assignments, respectively.

7.2.23 BCA colorimetry analysis

To determine more precise protein concentrations for CD studies, a Pierce bicinchoninic acid (BCA) Protein Micro Assay Kit (ThermoFisher) was used as outlined in the manufacturer's manual. Samples were aliquoted into 96-well plates and bovine serum albumin (BSA) was the protein standard. Absorbance readings at 562 nm were measured using a SpectraMax i3x Plate Reader (Molecular Devices).

7.2.24 Circular dichroism spectroscopy

Peptide concentrations were determined using the BCA colorimetry analysis mentioned in the preceding section. An OLIS DSM 17 CD spectrophotometer with a thermally controlled quartz cell and a 0.2-mm path length was used to collect circular dichroism (CD) spectra. Data were collected during five scans from 185 to 255 nm in 1 nm increments at desired temperature. Peptide solutions were prepared in water and the desired co-solvent (such as TFE or DPC micelles). Results are reported in molar ellipticity units (deg × cm^2 / dmol) and the percent α-helicity of the two peptides was calculated using the equation $\frac{3000-\theta_{222}}{39000} \times 100\%$.[91] This calculation can be used to approximate the α-helix composition of a sample based on the minimum reading at 222 nm (θ_{222}), which corresponds to the wavelength at which predominant movement of electron density for the oxygen π → carbon π* transition is observed for backbone amide carbonyl in an α-helix.[161]

7.2.25 MALDI-TOF mass spectrometry

Matrix-assisted laser desorption ionization time-of-flight mass spectrometry (MALDI-TOF MS) was performed using an AB Sciex Voyager Elite System (Voyager, Foster City, CA) with delayed extractions. Both positive and negative mode as well as reflectron and liner methods were used. Matrices chosen were either 3,5-dimethoxy-4-hydroxycinnamic acid (sinapinic acid) or 4-hydroxy-α-cyanocinnamic acid (HCCA) and applied in two layers.[162] Prior to sample acquisition, samples were concentrated, solvent was evaporated to < 20% organic, and if required, solutions were acidified to pH 3-5 (with 0.1% TFA) and purified using a ZipTip pipette tip (C_4 or C_{18}, EMD Millipore) as outlined by the manufacturer.

7.2.26 LCMS/MS

Liquid chromatography tandem mass spectrometry (LCMS/MS) was performed on a quadruple-time of flight (Q-ToF) Premier mass spectrometer (Waters, Milford, MA) coupled to an ultra-performance liquid chromatographer (Waters). Samples were first passed through a Waters Symmetry C_{18} nanoAquity column trap and Atlantis dC18 nanoAcquity analytical column (100 Å pore sizes, 75 µm × 15 cm, 3 µm). The trap was rinsed with solvent A (1% acetonitrile (ACN) and 0.1% formic acid (FA) in water) to desalt the sample prior to passing through the column, with a gradient of 1 to 60% from the desalting solvent A to solvent B (ACN and 0.1% FA) at a rate of 350 nL/min. Data were processed using software program PEAKS 5.1 (Bioinformatics Solutions, Waterloo, ON) to determine protein and peptide sequences.[163]

7.3 Experimental procedures for faerocin MK

7.3.1 Synthesis of FaeMK and SRCAM 602 using SPPS

Wang resin was loaded with Fmoc-L-Arg(Pmc)-OH or Fmoc-L-His(Trt)-OH on a 0.1 mmol scale. To swollen Wang resin (1 g), Fmoc-protected amino acid (0.5 mmol), HOAt (0.5 mmol), DIC (0.5 mmol), DMAP (0.5 mmol) and DCM/DMF (9:1, 15 mL) were added. Reaction vessel was shaken at rt for 3 h. The rest of the resin was capped with acetic anhydride (1 mL) and pyridine (1 mL), shaken at rt for 1 h, and filtered through a sintered glass funnel. Resin beads were washed with DMF (3 × 10 mL), DCM (3 × 10 mL), and MeOH (3 × 10 mL), dried, and stored at -20 ^{o}C. Following the previously outlined protocol for Fmoc-based solid-phase peptide synthesis (SPPS), automated SPPS on a PreludeX instrument was used to obtain both peptides. Both peptides were purified using C_{18} RP-HPLC method A.

SRCAM 602 was lyophilized and used without further modifications. The sequence was confirmed by LCMS/MS and total mass HRMS (ESI) calcd for $C_{211}H_{307}N_{59}O_{58}S$ $[M + 3H]^{+3}$ 1544.0957, found 1544.0943.

Mass analysis of the FaeMK peptide showed both reduced and oxidized forms of the peptide. Thus, the fractions containing FaeMK were lyophilized and resuspended in an oxidizing buffer (8 mM Cys, 1 mM cystine, 100 mM Tris-HCl, pH 8) at approximately 1 μM peptide and stirred for 2 d under argon to yield a disulphide-containing FaeMK.[164] Its sequence was confirmed by LCMS/MS (which also confirmed the presence of a disulfide bond) and total mass HRMS (ESI) calcd for $C_{228}H_{333}N_{63}O_{59}S_2$ $[M + 4H]^{+4}$ 1241.3690, found 1242.1184.

7.3.2 MIC determination using spot-on-lawn assays

Spot-on-lawn tests were performed using a two-fold serial dilution of FaeMK and SRCAM 602. A peptide concentration of 64 μM in sterile MQ water (140 μL) was made and diluted in series to create a total of 16 concentrations. 10 μL of each peptide concentration was aliquoted onto an agar plate streaked with the desired indicator strain; plates were incubated as appropriate for the indicator strain. This was repeated in triplicate to give minimum inhibitory concentrations (MICs) as an average.

7.3.3 Circular dichroism measurements for FaeMK and SRCAM 602

First, protein concentrations were determined using a bicinchoninic acid (BCA) Protein Micro Assay Kit (ThermoFisher). Kit reagents A:B:C were prepared as a 25:24:1 solution as instructed (10 mL). BSA protein standards were prepared from a 2.0 mg/mL stock following a two-fold serial dilution to create 8 different solutions of decreasing concentrations. These, along with a set of water blanks, were aliquoted into a 96-well plate. In the same plate, a two-fold serial dilution was conducted with FaeMK and SRCAM 602 to yield 3 solutions of different concentrations for each peptide. All samples were done in triplicate. Kit reagent solution (150 μL) was dispensed into every well; the plate was shaken slightly to mix and incubated at 37 ^{o}C for 1.5 h. Absorbance readings (562 nm) were measured and normalized to the BSA standards. FaeMK stock solution was determined to have a concentration of 3.31 mg/mL while SRCAM 602 had concentration of 6.11 mg/mL.

Next, these stock solutions were diluted with TFE to achieve solutions of 0, 25, 50, and 75% TFE in water for FaeMK (3.31 mg/mL, 2.48 mg/mL, 1.66 mg/mL, 0.83

mg/mL) and 0, 25, and 50% TFE in water for SRCAM 602 (6.11 mg/mL, 4.58 mg/mL, 3.06 mg/mL). During various NMR trials, FaeMK (1.52 mg/mL) was also analyzed in a DPC micelle system (180 mM DPC in MQ water). These solutions were analyzed using an OLIS DSM 17 CD spectrometer and data were collected during five scans from 185 to 255 nm in 1 m increments at 20 ^{o}C. Results are reported in molar ellipticity units (deg × cm^2 / dmol) and the percent α-helicity of the two peptides was calculated using the equation $\frac{3000-\theta_{222}}{39000} \times 100\%$.[91]

7.3.4 NMR studies of FaeMK

FaeMK (5.96 mg) was dissolved in TFE_{d3}:H_2O (1:3, 600 µL) to a concentration of 2 mM peptide. DSS (0.01%, 6 µL) was used as an internal standard. This solution was transferred to a standard 5 mm NMR tube for analysis. A temperature scan was performed on a Varian 600 MHz spectrometer, with ^{1}H NMR spectra collection at 10 ^{o}C intervals from 10 to 50 ^{o}C. COSY, TOCSY, NOESY, ^{13}C-HSQC spectra were acquired at 40 ^{o}C on a Varian 700 MHz spectrometer. Spectra were processed using NMRPipe,[160] analyzed using NMRViewJ,[159] and chemical shift assignments were made using TOCSY and NOESY datasets (Table 7.11).[165]

Table 7.11 ^{1}H chemical shift assignments for FaeMK

Residue	NH	Hα	Hβ	Hγ	Others
Ala 1	7.71	4.41	1.31		
Thr 2	8.91	4.48	3.95	0.90	
Tyr 3	*8.61*	*4.63*	*2.80*		–
Tyr 4	*8.51*	*4.26*	*3.06*		–
Gly 5	*7.76*	*4.14*			
Asn 6	8.63	4.40	2.95	2.74	NH 7.34, 6.75
Gly 7	8.10	3.99, 3.57			

Val 8	7.48	4.34	1.93	0.78	δ 0.70
Tyr 9	*8.61*	*5.15*	*2.09, 1.43*		–
Cys 10	8.55	5.44	3.13, 3.11		
Asn 11	9.20	5.10	3.59, 3.14		NH 7.84, 7.07
Lys 12	8.42	4.78	2.11, 1.91	1.75, 1.59	δ 1.48, ε 2.94, NH 7.55
Gln 13	8.41	4.45	2.40, 2.38	2.24	NH 6.82, 7.54
Lys 14	8.46	4.11	1.98	1.78, 1.61	δ 1.55
Cys 15	8.16	5.77	2.84, 2.79		
Trp 16	9.54	4.99	3.44, 3.28		δ 6.99, ε 9.91, –
Val 17	8.00	4.67	1.55	0.77	
Asp 18	9.29	5.00	3.20, 2.83		
Trp 19	8.40	4.56	3.47, 3.25		δ 7.37, ε 10.00, –
Asn 20	8.47	4.57	2.97, 2.90		NH 7.79, 6.93
Lys 21	7.85	4.21	1.96, 1.82	1.65, 1.50	δ 1.43, ε 2.94, NH 7.34, 6.93
Ala 22	8.11	4.17	1.43		
Ser 23	8.16	*4.25*	3.88, 3.81		
Lys 24	*7.93*	*4.27*	*1.95*	*1.73, 1.63*	*δ 1.53, ε 3.03, NH 8.03*
Glu 25	*8.03*	–	–		
Ile 26	–	*3.92*	*2.14*	*1.00*	*δ 0.93, 0.91*
Gly 27	8.16	4.25			
Lys 28	*7.70*	*4.13*	*1.92*	*1.72, 1.55*	*δ 1.43, ε 3.00*
Ile 29	7.78	–	–	–	*0.85*
Ile 30	*7.80*	*3.89*	*2.07*	*1.20*	*1.72, 0.92, 0.88*
Val 31	*8.08*	*3.90*	*1.93*	*1.14, 0.81*	
Asn 32	8.11	4.57	2.74, 2.69		NH 7.43, 6.76
Gly 33	*8.03*	–			
Trp 34	*8.11*	4.65	3.40		δ *7.16,* ε *9.82,* –
Val 35	7.79	–	–	–	
Gln 36	8.09	4.22	2.40, 2.34	2.04, 1.92	NH 7.37, 6.73
His 37	8.15	4.75	3.35, 3.19		δ 7.29, ε 8.45
Gly 38	7.96	3.93, 3.70			
Pro 39		4.28	2.05	1.83	δ 3.32, 3.18
Trp 40	7.54	4.66	3.30, 3.21		δ *7.17,* ε *10.00*
Ala 41	7.61	4.53	1.30		
Pro 42		4.37	2.02	1.98	δ 3.60, 3.36
Arg 43	8.02	4.33	1.93	1.77, 1.66	δ 3.20, NH 7.16, 7.04, 7.57, 7.46, 6.77

– , indicates assignments which could not be made due to significant overlap; *italics*, indicates values which were made through partial NOE agreements with less than 100% confidence

7.3.5 Expression and purification of recombinant faeMK SUMO

7.3.5.1 Construction of pET SUMO-FaeMK plasmid

The FaeMK operon nucleotide sequence of 552 bp was ordered from BioBasic Inc (Markham, ON). The DNA encoding mature FaeMK was amplified by Dr. van Belkum through PCR using primers MVB274 (5'-GCA ACT TAT TAT GGA AAT GGT GTA T-3') and MVB275 (5'-CTA TCT AGG AGC CCA AGG TCC A-3'). The amplified gene was cloned into pET SUMO expression vector as outlined in the manufacturer protocols (Invitrogen) and transformed into *E. coli* Mach1. Nucleotide sequencing was performed on plasmid DNA isolated from the transformants to ensure that *faeMK* was in frame with the SUMO protein. Next, the pET SUMO-FaeMK plasmid was isolated from Mach1 cells and transformed into *E. coli* BL21(DE3). Three glycerol stocks were created from the resulting *E. coli* transformants and stored at -80 ^{o}C for future use.

7.3.5.2 Expression of SUMO-FaeMK

A tube containing LB broth (10 mL) was inoculated with *E. coli* BL21(DE3) (pET SUMO-FaeMK) from frozen glycerol stock and incubated overnight at 37 ^{o}C. Terrific Broth (TB) (24 g yeast extract, 12 g tryptone, 2.31 g potassium dihydrogen phosphate, 12.64 g potassium hydrogen phosphate, 4 mL glycerol, pH 7.4, 1 L MQ water) containing kanamycin was prepared and warmed to 37 ^{o}C approximately 1 h prior to inoculation. Starter culture was used to inoculate 400 mL of TB and the mixture was shaken at 225 rpm at 37 ^{o}C until an optical density (OD_{600}) of 0.9-1.0 was noted, approximately 15 h later. OD measurements were taken using an 8351A

Diode Array Spectrophotometer (Hewlett Packard, Palo Alto, CA). Culture flasks were cooled on ice for 15 min following which IPTG (1 mM final concentration) was added and flasks were incubated at 30 ^{o}C for 24 h at 225 rpm. Cells were harvested the following day and lysed using a cell disruption or sonication. Variations to this protocol occurred during optimization trials.

7.3.5.3 Purification of SUMO-FaeMK

Nickel-nitrilotriacetic acid (Ni-NTA) Superflow Resin (Qiagen) was used to purify proteins and peptides with a His_6-tag. Resin slurry (50% suspension in 30% ethanol) was added to cell lysate containing peptide of interest (at a slurry:peptide ratio of 1:20 mL) and gently shaken for 1 h at 8 ^{o}C. The mixture was transferred to an Econo-Pac gravity-flow chromatography column (Bio-rad) and allowed to flow through by gravity filtration. Remaining resin was washed with 20 mL of buffer A (50 mM NaH_2PO_4, 500 mM NaCl, 20 mM imidazole, pH 8) and 10 mL of buffer B (50 mM NaH_2PO_4, 500 mM NaCl, 50 mM imidazole, pH 8). Fusion peptide was eluted using 10 mL of elution buffer (50 mM NaH_2PO_4, 500 mM NaCl, 200 mM imidazole, pH 8). Salts were removed by dialysis for 4 h at 8 ^{o}C against 4 L Tris buffer (20 mM Tris-HCl, 150 mM NaCl, pH 8). All fractions were collected and analyzed using SDS-PAGE. Variations to this protocol occurred during optimization trials, with varying concentrations of imidazole and volumes used for buffer washes and elution.

7.3.5.4 SUMO protease digestion of fusion protein

Digestion of the fusion protein SUMO-FaeMK with SUMO protease (McLab, South San Francisco, CA) resulted in separate His_6-SUMO and FaeMK protein fragments. A cleavage cocktail (500 mM Tris-HCl, 150 mM NaCl, pH 8, 10 mM DTT, 10 U protease/20 μg protein) was prepared and used to achieve complete digestion after 1 h incubation at rt. Changes to cleavage time, temperature, and cocktail components, including the addition of IGEPAL CA-630 (Sigma-Aldrich), were assayed during optimization trials.

7.3.5.5 Purification of FaeMK

Ni-NTA resin (2 mL) was added to the peptide cleavage mixture and incubated for 1 h at 8 ^{o}C. Mixture was transferred to a fritted column and flow-through was collected. Resin was washed with MQ water (3 × 5 mL) and combined fractions were dialyzed against 1 L MQ water (0.1% TFA) for 4 h at 8 ^{o}C. Alternatively, a C8 or C18 cartridge column was performed to eliminate the salts using an increasing concentration of IPA (0.1% TFA) from 0 to 75% in 25% stepwise increases of four 20 mL fractions. Finally, RP-HPLC was performed using method A with FaeMK eluting at 33.2 min. Pure fractions were combined and lyophilized to yield approximately 0.1 mg FaeMK/1 L culture. HRMS and NMR were used to verify the identity of this compound. Additionally, a spot-on-lawn assay was performed to confirm activity, using *L. monocytogenes* as an indicator. Variations to the solvent systems and types of purification techniques occurred during optimization trials.

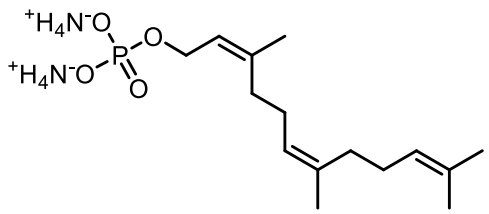
7.3.6 Expression of *faeMKI* gene products in two heterologous hosts

The *faeMKI* gene of interest was ordered from BioBasic Inc (Markham, ON) and amplified by Dr. van Belkum through PCR using primers (5'-ATT TTT GTG AGC TTG GAC TAG A-3') and (5'-AGT ACC ATT TGA CTT TGA ACC T -3'). The 552-bp DNA fragment containing the *FaeMKI* operon (including the gene for the putative immunity protein) and was bookended by *SacI* and *XbaI* restriction enzyme recognition sequences. FastDigest restriction enzymes *SacI* and *XbaI* were used to digest the *faeMKI* DNA as well as the pMG36c cloning vector.[166] After digestion, DNA was purified and the insert was ligated with the digested vector to yield pMG36c-*faeMKI* and transformed into *C. maltaromaticum* UAL26. The resulting plasmid was then isolated and transformed into *E. faecium* BFE 900. As a negative control, these two strains were also separately transformed with pMG36c vector not containing any DNA insert. Transformants were grown in their appropriate incubation conditions and used for deferred inhibition and spot-on-lawn assays.

7.4 Experimental procedures for teixobactin

7.4.1 Syntheses of truncated LII analogues and H-TriA$_1$

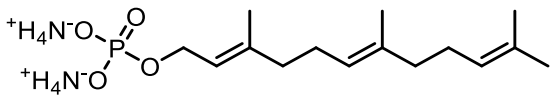
(2*Z*,6*Z*)-3,7,11-trimethyldodeca-2,6,10-trien-1-yl phosphate, ammonium salt (20)



20

This known monophosphate lipid was synthesized according to literature precedent with modification to a smaller scale.[110] To a solution of (*Z,Z*)-3,7,11-trimethyl-2,6,10-dodecatrien-1-ol **31** (2.4 μL, 8.99 μmol) and trichloroacetonitrile (2.2 μL, 21.6 μmol) in ACN (0.15 mL) a mixture of tetrabutylammonium dihydrogen phosphate (9.2 mg, 21.6 μmol) in ACN (0.2 mL) was added dropwise. The mixture was stirred at rt for 8 h after which the solvent was evaporated *in vacuo*. Crude compound mixture was purified using ion-exchange chromatography (DOWEX 50WX8). Compound **20** was isolated as a single peak by C_{18} RP-HPLC using method B (R_t 27.8 min). Sample-containing fractions were combined and lyophilized to yield a fluffy white solid (1.85 mg, 6.11 μmol, 68%). ^{1}H NMR (700 MHz, CD_3OD) δ 5.44 (t, 1H, J = 6.6 Hz), 5.20-5.16 (m, 2H), 4.51 (t, 2H, J = 6.6 Hz), 2.14-2.00 (m, 8H), 1.92 (s, 3H), 1.76 (s, 3H), 1.70 (s, 3H), 1.63 (s, 3H); ^{13}C NMR (126 MHz, CD_3OD) δ 136.7, 132.4, 130.8, 125.8, 125.6, 56.7, 33.3, 32.9, 27.7, 27.6, 26.0, 23.7, 23.0, 18.2; HRMS (ESI) calcd for $C_{15}H_{27}O_4P$ $[M-H]^-$ 301.1574, found 301.1574.

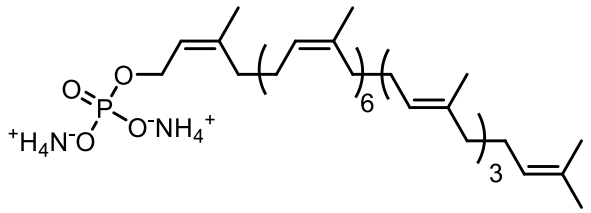
(2*E*,6*E*)-3,7,11-trimethyldodeca-2,6,10-trien-1-yl phosphate, ammonium salt (21)



21

This known compound was synthesized according to literature precedent with the following modifications.[110] To a solution of (*E,E*)-3,7,11-trimethyl-2,6,10-dodecatrien-1-ol (29.6 µL, 0.11 mmol) and trichloroacetonitrile (27.6 µL, 8.99 µmol) in ACN (0.25 mL) a mixture of tetrabutylammonium dihydrogen phosphate (114.5 mg, 0.34 mmol) in ACN (0.5 mL) was added dropwise. The mixture was stirred at rt for 16 h after which solvent was evaporated *in vacuo* and separated as outlined above. Compound **21** was isolated as a single peak by RP-HPLC method B (R_t = 27.2 min) and lyophilized to yield a fluffy white solid (34.1 mg, 0.08 mmol, 71%). ^{1}H NMR (700 MHz, CD_3OD) δ 5.41 (m, 1H), 5.13 (m, 2H), 4.39 (m, 2H), 2.16-2.01 (m, 8H), 1.97 (s, 3H), 1.69 (s, 6H), 1.61 (s, 3H); ^{13}C NMR (126 MHz, CD_3OD) δ 139.6, 135.5, 132.0, 124.3, 123.4, 123.0, 54.8, 33.7, 33.4, 26.7, 26.4, 23.4, 23.0, 22.1, 17.3; ^{31}P NMR (162 MHz, CD_3OD) δ 1.32; HRMS (ESI) calcd for $C_{15}H_{27}O_4P$ $[M-H]^-$ 301.1574, found 301.1568.

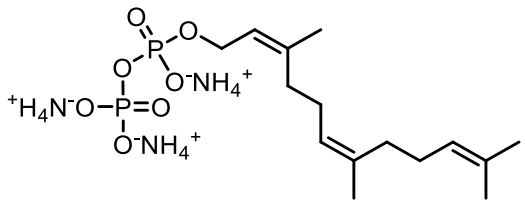
Undecaprenyl phosphate, ammonium salt (22)



22

Known compound **22** was synthesized using the same protocol outlined above.[110] Bay leaf-extracted undecaprenol (2.0 mg, 2.61 μmol) and trichloroacetonitrile (0.6 μL, 6.26 μmol) in ACN (0.1 mL) were added to a reaction vial. A mixture of tetrabutylammonium dihydrogen phosphate (1.8 mg, 5.22 μmol) in ACN (0.5 mL) was added dropwise. The mixture was stirred at rt for 16 h after which solvent was evaporated *in vacuo* an separated as outlined above. Compound **22** was isolated as a single peak by C_{18} RP-HPLC method B (R_t = 16.1 min) and lyophilized to yield a white solid (1.35 mg, 1.59 μmol, 61%). HRMS (ESI) calcd for $C_{55}H_{90}O_4P$ $[M-H]^-$ 8.45.6582, found 845.6591.

(2*Z*,6*Z*)-3,7,11-trimethyldodeca-2,6,10-trien-1-yl diphosphate, ammonium salt (23)

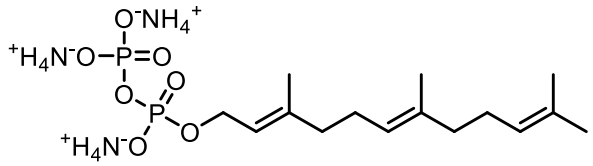


23

This known pyrophosphate lipid was synthesized using a previously reported protocol with the following modifications.[111] To a dry 4 mL vial containing a stir bar and dry

DCM (100 μL) under argon, trimethylamine (10 μL, 72.0 μmol), lithium chloride (2.0 mg, 46.8 μmol), and **20** (2.0 mg, 8.98 μmol) were added and the mixture was cooled to 0 ^{o}C for 15 min. Methanesulfonyl chloride (52.2 μL, 0.67 mmol) was added dropwise and the reaction was stirred at rt for 1 h. DCM (1 mL) and water (1 mL) were added to dilute and quench the reaction. Extractions (3 × 1 mL DCM) and washes with 1 M HCl (1 mL) and brine (1 mL) were followed by drying over $MgSO_4$ for 10 min and solvent was removed *in vacuo*. Crude material was dissolved in ACN (100 μL) and added dropwise to a stirring mixture of Tris(tetrabutylammonium) hydrogen phosphate (24.3 mg, 54.0 μmol) dissolved in ACN (500 μL) at 0 ^{o}C. The mixture was warmed to rt and allowed to stir at rt for 6 h. Solvent was removed *in vacuo* and purification was performed as previously described. Compound **24** was isolated as a single peak by C_{18} RP-HPLC method B (R_t = 34.3 min) as a white solid (1.79 mg, 4.67 μmol, 52%). ^{1}H NMR (700 MHz, CD_3OD) δ 5.45 (m, 1H), 5.14-5.13 (m, 2 H), 4.52 (m, 2H), 2.12-2.02 (m, 8H), 1.73 (s, 3H), 1.70 (s, 3H), 1.68 (s, 3H), 1.61 (s, 3H); ^{13}C NMR (126 MHz, CD_3OD) δ 136.7, 132.4, 130.8, 125.8, 125.6, 56.7, 33.3, 32.9, 27.7, 27.6, 26.0, 23.7, 23.0 18.2; ^{31}P NMR (162 MHz, CD_3OD) δ -7.98 to -10.21 (m, 2P); HRMS (ESI) calcd for $C_{15}H_{27}O_7P_2$ $[M-H]^-$ 381.1237, found 381.1230.

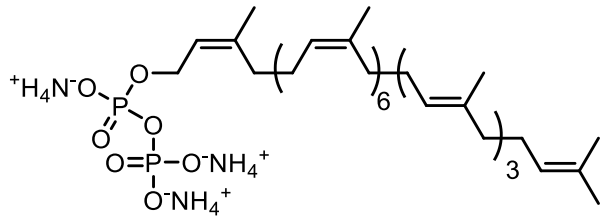
(2*E*,6*E*)-3,7,11-trimethyldodeca-2,6,10-trien-1-yl diphosphate, ammonium salt (24)



24

This known compound was synthesized using a previously reported protocol with the following modifications.[111] To a dry 4 mL vial containing a stir bar and dry DCM (500 μL) under argon, trimethylamine (125 μL, 0.89 mmol), lithium chloride (24.8 mg, 0.58 mmol), and (*E,E*)-3,7,11-trimethyl-2,6,10-dodecatrien-1-ol (29.6 μL, 0.11 mmol) were added and the mixture was cooled to 0 ^{o}C for 15 min. Methanesulfonyl chloride (52.2 μL, 0.67 mmol) was added dropwise and the reaction was stirred at rt for 1 h. DCM (1 mL) and water (1 mL) were added to dilute and quench the reaction. Extractions (3 × 1 mL DCM) and washes with 1 M HCl (1 mL) and brine (1 mL) was followed by drying over $MgSO_4$ for 10 min and solvent was removed *in vacuo*. Crude material was dissolved in ACN (250 μL) and added dropwise to a stirring mixture of Tris(tetrabutylammonium) hydrogen phosphate (304 mg, 0.033 mmol) dissolved in ACN (500 μL) at 0 ^{o}C. The mixture was warmed to rt and allowed to stir at rt for 6 h. Solvent was removed *in vacuo* and purification was performed as previously described. Compound **24** was isolated as a single peak by RP-HPLC method B as a white solid (22.8 mg, 0.06 mmol, 53%). ^{1}H NMR (700 MHz, CD_3OD) δ 5.47 (m, 1 H), 5.17-5.12 (m, 2H), 4.58-4.56 (m 1H), 4.44-4.41(m, 1H), 2.21-2.00 (m, 8H), 1.82 (s, 3H), 1.77 (s, 3H), 1.72 (s, 3H), 1.70 (s, 3H); ^{31}P NMR (162 MHz, CD_3OD) δ -9.08 (d, J = 21.2 Hz), -9.59 (d, J = 18.6 Hz); HRMS (ESI) calcd for $C_{15}H_{27}O_7P_2$ $[M-H]^-$ 381.1237, found 381.1235.

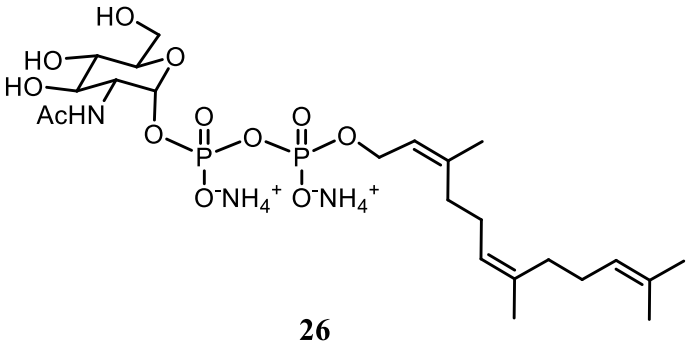
Undecaprenyl pyrophosphate, ammonium salt (25)



25

This known undecaprenyl pyrophosphate was synthesized using a previously reported protocol with the following modifications.[111] To a dry 4 mL vial, DCM (100 µL), trimethylamine (3.1 µL, 21.9 µmol), lithium chloride (0.6 mg, 14.2 µmol), and Bay leaf-extracted undecaprenol (2.1 mg, 2.74 µmol) were added and the mixture was cooled to 0 ^{o}C for 10 min. Methanesulfonyl chloride (13.0 µL, 16.4 µmol) was added dropwise and the reaction was stirred at rt for 1 h. Mixture was quenched and extracted as outlined previously. Compound **25** was purified as above to yield a white solid (1.28 mg, 1.18 µmol, 49%). ^{31}P NMR (162 MHz, CD_3OD) δ -7.98 to -10.21 (m, 2P); HRMS (ESI) calcd for $C_{55}H_{92}O_7P_2$ $[M-H]^-$ 925.6240, found 925.6015.

***Z,Z*-farnesyl pyrophosphate *N*-acetylglucosamine, ammonium salt (26)**

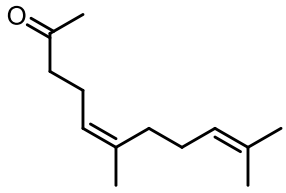


26

Coupling of the sugar phosphate with the lipid phosphate was achieved using a literature protocol to synthesise this known compound with the following modifications.[112] To a dry reaction vessel, 20% w/w $Pd(OH)_2/C$ (15 mg) and MeOH (4 mL) were added. Next, *N*-acetylglucosamine phosphate **35** (90 mg) was dissolved in MeOH (1 mL) and added dropwise, the solution was washed with water (3 × 5 mL), and stirred at rt for 30 min. The mixture was filtered and DIPEA (1 mL) was added to filtrate. Solvent was evaporated and the crude mixture was resuspended in THF (3 mL). DIC (50 mg) was added and mixture was stirred at rt for 2 h. The reaction was quenched with MeOH, stirred for 1 h at rt, and then solvents were evaporated. To a fresh vial, *Z,Z*-farnesyl phosphate **20** (3.2 mg, 8.81 μmol) was added along with the crude mixture (51.8 mg) and allowed to stir for 78 h at rt. A C_{18} solid-phase extraction cartridge (50 mM ammonium bicarbonate to 5% ACN) was used to purify the intermediate; fractions containing desired product were combined and lyophilized to yield a beige solid. Global deprotection was achieved by dissolving the crude in 1,4-dioxane (1.5 mL) and 1 M NaOH (1.5 mL), and allowing the mixture to stir at 37 ^{o}C for 2 h. Crude was purified *via* C_{18} solid-phase extraction cartridge (50 mM ammonium bicarbonate to 50% ACN) to yield **26** as a white fluffy powder (1.60

mg, 2.73 μmol, 31% over 3 steps). IR (cast film, DCM): 3072, 2915, 2546, 1921, 1672, 1621, 1449, 1430, 1329, 1100, 1046, 999, 880, 934, 697, 659 cm^{-1}; ^{1}H NMR (600 MHz, D_2O) δ 5.44 (dd, *J* = 7.0, 3.0 Hz, 1H), 5.41 (t, *J* = 7.0 Hz, 1H), 5.16 (m, 2H), 4.41 (t, *J* = 6.9 Hz, 1H), 3.94 (m, 1H), 3.88 (m, 1H), 3.84 (m, 1H), 3.78-3.74 (m, 2H), 3.50 (t, *J* = 9.4 Hz, 1H), 2.11-2.05 (m, 8H), 2.02 (s, 3H), 1.73 (s, 3H), 1.65 (s, 6H), 1.59 (s, 3H); ^{13}C NMR (126 MHz, D_2O) δ174.8, 143.1, 137.0, 133.7, 124.8, 124.4, 120.5, 94.7, 73.2, 71.2, 69.7, 63.0, 60.5, 53.8, 31.6, 31.1, 26.0, 25.9, 25.0, 22.8, 22.7, 22.2, 17.2; ^{31}P NMR (162 MHz, D_2O) δ -10.70 (d, *J* = 21.0 Hz, 1P), -13.04 (d, *J* = 21.0 Hz, 1P); HRMS (ESI) calcd for $C_{23}H_{40}NO_{12}P_2$ $[M-H]^-$ 584.2031, found 584.2038.

(*Z*)-6,10-dimethylundeca-5,9-dien-2-one (30)

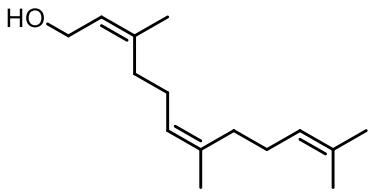


30

This known compound was synthesized using a literature protocol with the following modifications.[167] To a solution of nerol (40.0 g, 0.259 mol) in diethyl ether at -20 °C, phosphorous tribromide (45.7 mL, 0.259 mol) was added dropwise. The reaction mixture was stirred for 2 h, then allowed to warm to 0 °C. Upon completion as detected by TLC, reaction mixture was quenched with ice-cold water (500 mL) and extracted with Hex (3 × 250 mL). Combined organic layers were dried over $MgSO_4$ and concentrated *in vacuo*. Crude neryl bromide (101.8 g, 0.469 mol) was added to

potassium carbonate (84.5 g, 0.610 mol) and ethyl acetoacetate (150 mL, 1.175 mol) in Ace (350 mL) and refluxed for 12 h at 65 ^{o}C. The mixture was cooled to rt, quenched with NH_4Cl (250 mL), poured into water (250 mL) and extracted with diethyl ether (3 × 250 mL). The organic layer was dried over $MgSO_4$ and concentrated *in vacuo.* Crude alkylation product (1.0 g, 3.75 mmol) was dissolved in rt MeOH (50 mL) and 5 M potassium hydroxide (0.694 g in 25 mL water, 10.51 mmol) was added followed by a 2 h reflux at 80 ^{o}C. Reaction was cooled to 0 ^{o}C and quenched by dropwise acidification with 6 M HCl (50 mL), followed by extraction with diethyl ether (3 × 250 mL). The combined organic layers were dried over $MgSO_4$ and concentrated *in vacuo*. The crude product was separated from starting material by flash chromatography (0-10% EtOAc in Hex) and concentrated *in vacuo* to yield **30** (676 mg, 3.48 mmol, 93%). ^{1}H NMR (700 MHz, $CDCl_3$) δ 9.99 (s, 1H), 5.89-5.88 (m, 1H), 5.09-5.07 (m, 2H), 2.22 (s, 4H), 2.04 (s, 4H), 1.68 (s, 6H), 1.60 (3H), 1.56 (s, 3H); ^{13}C NMR (176 MHz, $CDCl_3$) δ 191.3, 163.9, 136.6, 131.5, 127.5, 124.2, 122.5, 40.7, 39.7, 28.5, 26.7, 25.7, 17.7, 17.7, 16.1; HRMS (ESI) calcd for $C_{13}H_{22}O$ $[M + H]^+$ 195.1878, found 195.1876.

(2*Z*,6*Z*)-3,7,11-trimethyldodeca-2,6,10-trien-1-ol (31)

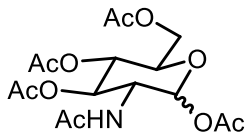


31

This known compound was synthesized using a modified method.[167] In a dry RBF, the Wittig reagent was prepared by dissolving methyltriphenylphosphenyl bromide

(3.4 g, 9.38 mmol) in THF (80 mL). To this solution, *n*-butyllithium (4 mL, 9.38 mmol) was added dropwise and the solution was left to stir for 30 min at 0 ^{o}C. The reaction was cooled to -78 ^{o}C for 10 min, followed by the dropwise addition of (*Z*)-6,10-dimethylundeca-5,9-dien-2-one **30** (2 mL, 8.92 mmol). The reaction as allowed to stir for 1 h after which *sec*-butyllithium (7.7 mL, 11.6 mol) was added dropwise. The mixture was allowed to stir at -78 ^{o}C for 2 h prior to being placed in an ice bath at 0 ^{o}C. Paraformaldehyde solution (70 mL THF, approx. 33.76 mmol formaldehyde) was heated at 120 ^{o}C and transferred *via* cannulation into the reaction vessel and allowed to stir for 30 min at 0 ^{o}C. The reaction was warmed to rt and stirred for an additional 2.5 h before being quenched by water (80 mL). The solution was extracted into Hex (3 × 40 mL), organic layer was dried over $MgSO_4$ and concentrated *in vacuo*. Cold Hex were added to precipitate triphenylphosphine oxide, which was removed by filtration. *In vacuo* concentration was followed by purification *via* flash chromatography (1:99 EtOA:Hex) to yield the desired product **31** as a colourless oil (0.96 g, 4.31 mmol, 46%). IR (neat film): 3326, 2962, 2855, 1666, 1442, 1474, 1001 cm^{-1}; ^{1}H NMR (500 MHz, $CDCl_3$) δ 5.45 (m, 1H), 5.12-5.10 (m, 2H), 4.11 (d, J = 6.9 Hz, 2H), 2.11-2.04 (m, 8H), 1.75 (s, 3H), 1.70 (s, 6H) 1.62 (s, 3H); ^{13}C NMR (126 MHz, $CDCl_3$) δ139.8, 136.1, 131.7, 124.6, 124.5, 124.2, 59.0, 33.2, 31.9, 26.7, 26.3, 25.7, 23.4, 23.3, 17.6; HRMS (ESI) calcd for $C_{15}H_{26}NaO$ $[M+Na]^+$ 245.1876, found 245.1879.

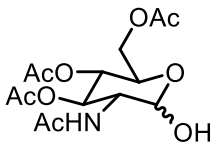
(3*R*,4*R*,5*S*,6*R*)-3-acetamido-6-(acetoxymethyl)tetrahydro-2H-pyran-2,4,5-triyl triacetate (33)



33

Preparation of this known phosphorylated monosaccharide was achieved using a literature method with modifications.[112] To a dry RBF, *N*-acetylglucosamine (2.5 g, 11.3 mmol), ACN (75 mL) and trimethylamine (18.9 mL, 125.6 mmol) were added and the reaction vessel was cooled to 0 ^{o}C for 10 min. Acetic anhydride (22.6 mL, 203.4 mmol) was added dropwise and the reaction was stirred at 0 ^{o}C for 2 h, then warmed to rt and stirred for an additional 2 h. The reaction mixture was once again cooled to 0 ^{o}C and MeOH (25 mL) was added and the resulting mixture was stirred for 30 min. The solvent was evaporated *in vacuo* and resuspended in DCM (75 mL) then washed with 1 M HCl (50 mL), saturated $NaHCO_3$ (2 × 50 mL), brine (50 mL) and dried over Na_2SO_4 for 10 min. Flash chromatography (2:1 to 1:1 Hex:Ace) was used to purify **33** as a white solid (2.23 g, 5.99 mmol, 53%). R_f = 0.74 (1:1 Hex:Ace); $[\alpha]^{25}_D$ = + 78.83 (*c* = 1.16, $CHCl_3$); IR ($CHCl_3$, cast film): 3291, 2955, 1748, 1662, 1239 cm^{-1}; ^{1}H NMR (700 MHz, $CDCl_3$) δ 6.17 (d, *J* = 3.7 Hz, 1H), 5.52 (d, *J* = 9.0 Hz, 1H) 5.25-5.5.20 (m, 2H), 4.49-4.48 (m, 1H), 4.25 (dd, 1H, *J* = 12.5, 4.2 Hz), 4.07 (dd, *J* = 12.5, 2.4 Hz, 1H) 3.99 (m, 1H), 2.20 (s, 3H), 2.09 (s, 3H), 2.06 (s, 3H), 2.05 (s, 3H), 1.94 (s, 3H); ^{13}C NMR (176 MHz, $CDCl_3$) δ 171.8, 170.7, 169.3, 169.1, 168.6, 90.8, 70.8, 69.8, 67.5, 61.6, 51.1, 23.1, 21.0, 20.8, 20.7, 20.6; HRMS (ESI) calcd for $C_{16}H_{23}NO_{10}$ $[M+Na]^+$ 412.1214, found 412.1213.

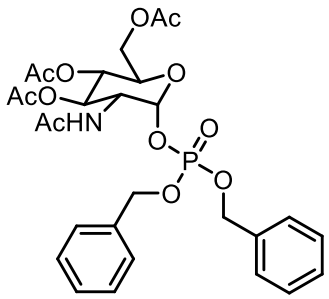
(2*R*,3*S*,4*R*,5*R*)-5-acetamido-2-(acetoxymethyl)-6-hydroxytetrahydro-2*H*-pyran-3,4-diyl diacetate (34)



34

This known monoccharide was synthesized using a literature method with the following modifications.[112] To a solution of peracetylated *N*-acetylglucosamine **33** (4.05 g, 10.4 mmol) dissolved in DMF (50 mL), hydrazine acetate (1.17 g, 12.7 mmol) was added and the solution was stirred at rt for 1.5 h. The solvent was removed *in vacuo* and the mixture was purified using flash chromatography (1:2 Hex:Ace) to yield a yellow oil (3.04 g, 8.74 mmol, 84%). R_f = 0.2 (1:2 Hex:Ace); $[\alpha]^{25}_D$ = + 35.17 (*c* = 0.46, $CHCl_3$); IR ($CHCl_3$, cast film): 3291, 2955, 1748, 1662, 1239 cm^{-1}; ^1H NMR (700 MHz, $CDCl_3$) δ 5.79 (d, *J* = 9.4 Hz, 1H), 5.30-5.28 (m, 2H), 5.14 (appt. d, *J* = 10.1 Hz, 1H), 4.21-4.20 (m, 1H), 4.14-14.13 (m, 2H), 4.11-4.10 (m, 1H), 2.02 (s, 3H), 2.01 (s, 3H), 2.00 (s, 3H), 1.96 (s, 3H), 1.81 (s, 3H); ^{13}C NMR (126 MHz, $CDCl_3$) δ 171.4, 170.8, 170.6, 170.2, 169.3, 91.7, 70.9, 68.2, 67.7, 62.0, 52.2, 25.3, 23.2, 20.7, 20.3; HRMS (ESI) calcd for $C_{14}H_{31}NO_9$ $[M+Na]^+$ 370.1109, found 370.1109.

(2*R*,3*S*,4*R*,5*R*,6*R*)-5-acetamido-2-(acetoxymethyl)-6-((bis(benzyloxy)phosphoryl)oxy)tetrahydro-2H-pyran-3,4-diyl diacetate (35)

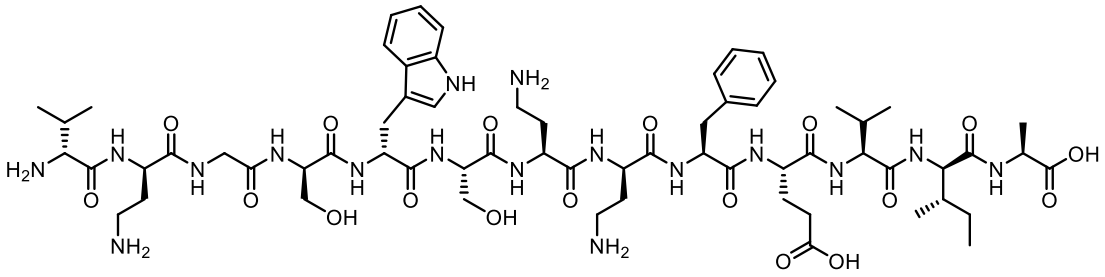


35

Synthesis of this known phosphorylated monosaccharide was achieved using a literature method with modifications.[112] In a flame-dried RBF, anomerically deprotected *N*-acetylglucosamine **34** (0.21 g, 0.59 mmol) and tetrazole solution (13.19 mL, 5.9 mmol) were added to a RBF containing DCM (15 mL). The reaction mixture was placed at -40 ^{o}C for 10 min prior to the dropwise addition of dibenzyl *N,N*-diisopropylphosphoamidite (1.0 mL, 3.0 mmol). The reaction was stirred at rt for 3 h and then cooled to -78 ^{o}C. *Meta*-Chloroperoxybenzoic acid (1.0 g, 5.9 mmol) was first dissolved in DCM (5 mL) then added dropwise to the reaction flask. The reaction was warmed to rt and allowed to stir for 16 h. The reaction mixture was washed with saturated $NaHCO_3$ (15 mL), water (15 mL), and brine (15 mL). The organic layer was dried with $MgSO_4$ and purified using flash chromatography (2% MeOH in DCM) to yield **35** as a white foam (0.26 g, 0.43 mmol, 72% yield), R_f = 0.90 (1:9 MeOH:DCM). $[\alpha]^{25}_D$ = + 53.72 (*c* = 0.34, $CHCl_3$); IR ($CHCl_3$, cast film): 3278, 3067, 2995, 1751, 1682, 1549, 1456, 1368, 1240, 1037 cm^{-1}; ^{1}H NMR (700 MHz, $CDCl_3$) δ 7.40–7.34 (m, 10H), 5.66 (dd, *J* = 3.0, 5.9 Hz, 1H), 5.58 (d, *J* = 9.2, 1H), 5.14-5.03 (m, 6H), 4.39-4.34 (m, 1H), 4.12 (dd, *J* = 4.0, 12.5 Hz, 1H), 4.11-4.10 (m, 1H), 3.95-

3.90 (m, 1H), 3.89 (dd, J = 2.2, 12.5 Hz, 1H), 2.02 (s, 3H), 2.00 (s, 3H), 1.99 (s, 3H), 1.70 (s, 3H); ^{13}C NMR (126 MHz, $CDCl_3$) δ 171.2, 170.6, 170.2, 169.2, 129.0, 128.9, 128.8, 128.2, 128.1, 96.3, 96.2, 70.2, 70.1, 70.0, 69.9, 69.8, 69.7, 67.2, 61.3, 51.9, 51.8, 22.8, 20.7, 20.6, 20.5; ^{31}P NMR (162 MHz, $CDCl_3$) δ -2.35; HRMS (ESI) calcd for $C_{28}H_{34}NNaO_{12}P$ $[M+Na]^+$ 630.1711, found 630.1710.

H-TriA_1, unacylated tridecaptin A_1 (38)



38

Unacylated tridecaptin A_1 was synthesized as previously described by our group through Fmoc-based SPPS.[125] Compound **38** was cleaved from resin and isolated by C_{18} RP-HPLC method B in a single peak (R_t = 23.2 min) as white fluffy solid (8.2 mg, 30%); HRMS (ESI) calcd for $C_{64}H_{100}N_{17}O_{18}$ $[M+H]^+$ 1394.7447, found 1394.7442.

7.4.2 Isothermal titration calorimetry

An MCS isothermal titration calorimeter (Microcal, Northhampton, MA) was used to perform calorimetric experiments of teixobactin (and its four analogues) binding to lipid II (and its analogues). A 10 μM solution of peptide and a 100 μM of lipid were prepared in the same buffer (10 mM Tris-HCl, 150 mM NaCl, pH 7) and degassed by

stirring under vacuum at 15 ^{o}C for 8 min. The lipid solutions were titrated into the peptide solutions at 25 ^{o}C with the following specifications: reference power = 17 μal/s, syringe-stirring speed = 300 rpm, injection volume = 10 μL, number of injections = 29, initial delay = 60 s, time between injections = 300 s. The change in heat during each titration was registered in real time and data were processed using the provided Origin® 7 software. Control experiments were performed with buffer blanks to determine the background observed for heat of dilution during titration. Experiments were performed in triplicate and averages are reported; a standard deviation between 27.2 – 1.7 nM was observed over all completed trials.

7.4.3 NMR trials of Txb with *Z,Z*-C_{15}-LII-Lys

Txb was dissolved in a series of deuterated solvent systems (600 μL) to a concentration of 2 mM peptide. The lipid was added to a 1 mM concentration to preserve the 2:1 expected binding ratio. TMS (0.01%, 6 μL) was used as an internal standard and if exchangeable protons were a concern, H_2O was added to ensure backbone amides would still be detected. pH control was achieved using a nano-pH probe and adjustment with TFA, if required. During the transfer to a standard 5 mm NMR tube for analysis, precipitation was observed at the bottom of the tube. Solvent was decanted and it was confirmed that the precipitate was Txb though mass analysis. Various solvent systems, pH ranges, and temperatures were assayed. A mixture containing the least amount of precipitate (3:1 DMSO:aqueous phosphate buffer) by visual inspection was analyzed using a Varian 600 MHz spectrometer at 26 ^{o}C. This

yielded a good spectrum of lipid but not sufficient Txb stayed dissolved in solution to obtain adequate data for a full assignment.

7.4.4 Synergistic assays

The method to determine MICs for Txb was through a modified broth microdilution protocol from the Clinical and Standards Laboratory Institute.[156] Antimicrobial peptides were dissolved in Muller-Hinton Broth (MHB) medium and serial dilutions were made across a 96-well plate. Txb and analogue stock solutions were 90 μg/mL, H-TriA_1 stock solution was 12.5 μg/mL, PMBN solution was 30 μg/mL, and the vancomycin control was 400 μg/mL. In the case of synergistic assays, dilutions were made in the x-axis for Txb, Txb analogues, and vancomycin and in the y-axis for H-TriA_1 and PMBN. Control wells containing no antibiotics were also used for each trial. One plate was used for each indicator organism and each well was inoculated to reach a final inoculum of 5×10^5 colony forming units per mL. Using OD_{600} readings normalized to a negative control, MICs were recorded as the lowest concentration at which no growth was detected after 24 h, or 48 h for K. *pneumoniae*, incubation. Each assay was performed in triplicate and average MICs are reported. Absorbance readings were taken on a SpectraMax i3x Plate Reader (Molecular Devices, Sunnyvale, CA) at 0, 24, and 48 h.

7.5 Experimental procedures for tridecaptin

7.5.1 NMR sample preparation of $TriA_1$, LII, and co-samples

Both peptide and lipid were assayed individually as well as in combination, at a 1:1 ratio in 120 mM DPC_{d38} or oxo-analogues DPC_{d36}. Oct-$TriA_1$ concentration was 4 mM and *Z,Z*-C_{15} LII-DAP was 4 mM. Samples were dissolved in a 10 mM phosphate buffer D_2O:H_2O (1:9, 300 µL) and pH was adjusted to 6.0. TMS (0.01%, 2 µL) was used as an internal standard. This solution was transferred to a 5 mm Shigemi-D_2O NMR tube for analysis. A temperature scan was performed on a Varian 600 MHz spectrometer, with ^{1}H NMR spectra collection at 10 ^{o}C intervals from 10 to 60 ^{o}C. COSY, TOCSY, NOESY, ^{13}C-HSQC, ^{13}C-HMBC spectra were acquired at 27 ^{o}C on the same spectrometer. To confirm reproducibility, select samples were prepared fresh and duplicate experiments were performed. Spectra were processed using NMRPipe,[160] analyzed using NMRViewJ,[159] and chemical shift assignments were made using TOCSY and NOESY datasets.[165]

The following sections contain the chemical shift assignments of Oct-$TriA_1$ alone in DPC_{d38} micelles as well as in the four oxoDPC_{d36} analogues, the chemical shift assignments of *Z,Z*-C_{15} LII-DAP alone in DPC_{d38} micelles as well as in the four oxoDPC_{36} analogues, and finally the chemical shift assignments of both Oct-$TriA_1$ and *Z,Z*-C_{15} LII-DAP in DPC_{d38} micelles as well as in 6oxoDPC_{d36} and 10oxoDPC_{d36}. Unfortunately, due to the gel formation and the lysis of LII, assignments were not possible in the other two oxo analogues. The blue headings are used for Oct-$TriA_1$ assignment tables and green headings are used for *Z,Z*-C_{15} LII-DAP assignment tables.

7.5.2 Chemical shift assignments for Oct-TriA$_1$

Table 7.12 ^{1}H chemical shift assignments for Oct-TriA$_1$ in DPC$_{d38}$

Residue	NH	Hα	Hβ	Hγ	Others
lipid		2.34, 2.25	1.56, 1.58	1.24	1.23, 1.22, 1.21, 0.81
D-Val1	7.76	4.26	2.09	0.90	
D-Dab2	8.63	4.79	2.17, 2.02	2.96	
Gly3	8.4	4.04, 3.85			
D-Ser4	8.15	4.76	3.67		
D-Trp5	8.43	4.69	3.18		E1 = 10.52, D1 = 7.25, Z2 = 7.41, H2 = 7.01, Z3 = 6.93, E3 = 7.49
Ser6	8.55	4.03	3.73, 3.44		
Dab7	8.5	4.31	2.17, 2.00	3.01	
D-Dab8	8.01	4.44	2.06, 1.83	2.83	
Phe9	8.54	4.80	3.08, 2.93		D1, D2, Z = 7.21, E1, E2 = 7.14
Glu10	8.27	4.54	1.99, 1.86	2.28	
Val11	8.18	4.55	1.91	0.92	
D-*a*Ile12	8.09	4.44	1.91	1.41	γCH_3 = 1.16, δ = 0.90
Ala13	8.19	4.15	1.34		

Table 7.13 ^{1}H chemical shift assignments for Oct-TriA$_1$ in 3oxoDPC$_{d36}$

Residue	NH	Hα	Hβ	Hγ	Others
lipid		2.34, 2.25	1.58, 1.55	1.24	1.22, 0.82
D-Val1	7.77	4.3	2.08	0.89	
D-Dab2	8.62	4.86	2.15, 2.00	2.92	
Gly3	-	-			
D-Ser4	-	-	3.64		
D-Trp5	-	4.75	3.18		E1 = 10.52, D1 = 7.24, Z2 = 7.40, H2 = 7.02, Z3 = 6.92, E3 = 7.49
Ser6	8.58	4.04	3.73, 3.47		
Dab7	8.53	4.33	2.16, 1.97	3.00	
D-Dab8	-	-	-	-	
Phe9	-	-	3.06		D1, D2, Z = 7.20, E1, E2 = 7.11
Glu10	-	-	-	-	
Val11	-	-	1.88	0.92	
D-*a*Ile12		4.44	1.9	1.46	γCH_3 = 1.15, δ = 0.90
Ala13	8.23	4.16	1.32		

Table 7.14 ^{1}H chemical shift assignments for Oct-TriA$_1$ in 5oxoDPC$_{d36}$

Residue	NH	Hα	Hβ	Hγ	Others
lipid		2.35, 2.26	1.58, 1.55	1.25	1.23, 1.21, 1.18, 0.82
D-Val1	7.76	4.29	2.10	0.90, 0.91	
D-Dab2	8.62	4.86	2.15, 2.00	2.93	
Gly3	8.43	4.08, 3.81			
D-Ser4	8.16	4.82	3.67		
D-Trp5	8.51	4.75	3.19		E1 = 10.52, D1 = 7.25, Z2 = 7.41, H2 = 7.01, Z3 = 6.94, E3 = 7.50
Ser6	8.58	4.03	3.74, 3.47		
Dab7	8.53	4.33	2.16, 1.99	3.02	
D-Dab8	7.98	4.48	2.07, 1.83	2.82	
Phe9	8.60	-	3.07		D1, D2, Z = 7.20, E1, E2 = 7.14
Glu10	8.30	4.63	1.99, 1.85	2.25	
Val11	8.21	4.19	1.89	0.92	
D-*a*Ile12	8.10	4.45	1.97	1.45	γCH$_3$ = 1.18, δ = 0.90
Ala13	8.21	4.16	1.33		

Table 7.15 ^{1}H chemical shift assignments for Oct-TriA$_1$ in 6oxoDPC$_{d36}$

Residue	NH	Hα	Hβ	Hγ	Others
lipid		2.35, 2.25	1.55	1.24	1.22, 1.21, 1.20, 0.82
D-Val1	7.75	4.28	2.10	0.90	
D-Dab2	8.62	4.84	2.18, 2.01	2.95	
Gly3	8.43	4.05, 3.84			
D-Ser4	8.15	4.79	3.67		
D-Trp5	8.46	4.73	3.17		E1 = 10.52, D1 = 7.26, Z2 = 7.41, H2 = 7.02, Z3 = 6.94, E3 = 7.50
Ser6	8.58	4.03	3.74, 3.47		
Dab7	8.52	4.33	2.17, 2.00	3.02	
D-Dab8	7.98	4.47	2.06, 1.84	2.83	
Phe9	8.58	4.84	3.08		D1, D2, Z = 7.20, E1, E2 = 7.14
Glu10	8.30	4.57	1.98, 1.85	2.25	
Val11	8.21	4.58	1.9	0.92	
D-*a*Ile12	8.12	4.41	1.9	1.43	γCH$_3$ = 1.18, δ = 0.90
Ala13	8.19	4.16	1.33		

Table 7.16 ^{1}H chemical shift assignments for Oct-TriA$_1$ in 10oxoDPC$_{d36}$

Residue	NH	Hα	Hβ	Hγ	Others
lipid		2.35, 2.25	1.55	1.24	1.22, 1.21, 1.20, 0.82
D-Val1	7.74	4.28	2.10	0.90, 0.89	
D-Dab2	8.61	4.85	2.16, 2.01	2.94	
Gly3	8.40	4.05, 3.84			
D-Ser4	8.15	4.80	3.67		
D-Trp5	8.48	4.76	3.18		E1 = 10.52, D1 = 7.25, Z2 = 7.41, H2 = 7.02, Z3 = 6.94, E3 = 7.50
Ser6	8.58	4.02	3.74, 3.47		
Dab7	8.52	4.33	2.17, 1.99	3.01	
D-Dab8	7.98	4.47	2.07, 1.83	2.82	
Phe9	8.58	4.84	3.08		D1, D2, Z = 7.20, E1, E2 = 7.14
Glu10	8.31	4.57	1.99, 1.85	2.25	
Val11	8.21	4.58	1.89	0.92	
D-*a*Ile12	8.10	4.46	1.96	1.42	γCH$_3$ = 1.18, δ = 0.87
Ala13	8.19	4.15	1.32		

7.5.3 Chemical shift assignments for *Z,Z*-C$_{15}$ LII-DAP

Table 7.17 ^{1}H chemical shift assignments for Z,Z-C$_{15}$ LII-DAP in DPC$_{d38}$

Residue	NH	Hα	Hβ	Hγ	Others
GlcNAc	8.33				H1 = 4.53, H2 = 3.71, H3 = 3.53, H4 = 3.39, H5 = 3.53, CH$_2$ = 3.69 & 3.84, Ac = 2.00
MurNac	8.18				H1 = 5.46, H2 = 3.99, H3 = 3.69, H4 = 3.87, H5 = 3.38, CH$_2$ = 3.89 & 3.62, Ac = 2.00, OCH = 4.42, CH$_3$ = 1.37
Ala1	8.23	4.32	1.39	8.23	
D-Glu2	8.06	4.15	2.35	2.10, 1.91	
DAP3	8.24	4.19	1.89, 1.79	1.42	δ = 1.73, ε = 3.71, NH$_2$ = 8.14, 8.26
D-Ala4	8.34	4.29	1.34		
D-Ala5	7.98	4.16	1.33		
C$_{15}$					H1 = 4.41, H2 = 5.40, H4 = 1.72, H5 = 2.11, H6 = 2.09, H7 = 5.16, H9 = 1.66, H10 = 2.05, H11 = 2.05, H12 = 5.17, H14 = 1.66, H15 = 1.60

Table 7.18 ^{1}H chemical shift assignments for Z,Z-C_{15} LII-DAP in 3oxoDPC$_{d36}$

Residue	NH	Hα	Hβ	Hγ	Others
GlcNAc	8.36				H1 = 4.53, H2 = 3.70, H3 = 3.53, H4 = 3.38, H5 = 3.53, CH_2 = 3.69 & 3.84, Ac = 2.00
MurNac	8.21				H1 = 5.45, H2 = 3.98, H3 = 3.70, H4 = 3.87, H5 =3.38, CH_2 = 3.89 & 3.71, Ac = 2.01, OCH = 4.41, CH_3 = 1.36
Ala1	8.24	4.31	1.38		
D-Glu2	8.12	4.18	2.34	2.10, 1.91	
DAP3	8.23	4.18	1.88, 1.77	1.43	δ = 1.73, ε = 3.69
D-Ala4	8.35	4.28	1.33		
D-Ala5	8.01	4.18	1.33		
C_{15}					H1 = 4.41, H2 = 5.42, H4 = 1.73, H5 = 2.10, H6 = 2.08, H7 = 5.15, H9 = 1.66, H10 = 2.01, H11 = 2.02, H12 = 5.11, H14 = 1.64, H15 = 1.58

Table 7.19 ^{1}H chemical shift assignments for Z,Z-C_{15} LII-DAP in 5oxoDPC$_{d36}$

Residue	NH	Hα	Hβ	Hγ	Others
GlcNAc	8.36				H1 = 4.53, H2 = 3.70, H3 = 3.53, H4 = 3.38, H5 = 3.69, CH_2 = 3.68 & 3.84, Ac = 2.00
MurNac	8.21				H1 = 5.44, H2 = 3.98, H3 = 3.70, H4 = 3.87, H5 =3.38, CH_2 = 3.89 & 3.62, Ac = 2.00, OCH = 4.41, CH_3 = 1.38
Ala1	8.24	4.30	1.39		
D-Glu2	8.06	4.15	2.33	2.09, 1.89	
DAP3	8.23	4.18	1.88, 1.78	1.43	δ = 1.73, ε = 3.69, NH_2 = 8.26
D-Ala4	8.34	4.27	1.33		
D-Ala5	7.99	4.15	1.32		
C_{15}					H1 = 4.40, H2 = 5.41, H4 = 1.72, H5 = 2.10, H6 = 2.06, H7 = 5.15, H9 = 1.65, H10 = 2.01, H11 = 2.03, H12 = 5.10, H14 = 1.65, H15 = 1.57

Table 7.20 ^{1}H chemical shift assignments for Z,Z-C_{15} LII-DAP in 6oxoDPC$_{d36}$

Residue	NH	Hα	Hβ	Hγ	Others
GlcNAc	8.36				H1 = 4.53, H2 = 3.70, H3 = 3.53, H4 = 3.38, H5 = 3.69, CH_2 = 3.68 & 3.84, Ac = 2.00
MurNac	8.21				H1 = 5.45, H2 = 3.98, H3 = 3.70, H4 = 3.87, H5 =3.38, CH_2 = 3.89 & 3.71, Ac = 2.00, OCH = 4.41, CH_3 = 1.36
Ala1	8.24	4.31	1.38		
D-Glu2	8.13	4.18	2.35	2.10, 1.91	
DAP3	8.24	4.17	1.87, 1.77	1.42	δ = 1.72, ε = 3.69, NH_2 = 8.26
D-Ala4	8.36	4.27	1.34		
D-Ala5	8.03	4.18	1.33		
C_{15}					H1 = 4.40, H2 = 5.43, H4 = 1.71, H5 = 2.11, H6 = 2.07, H7 = 5.15, H9 = 1.66, H10 = 2.02, H11 = 2.02, H12 = 5.11, H14 = 1.64, H15 = 1.59

Table 7.21 ^{1}H chemical shift assignments for Z,Z-C_{15} LII-DAP in 10oxoDPC$_{d36}$

Residue	NH	Hα	Hβ	Hγ	Others
GlcNAc	8.36				H1 = 4.53, H2 = 3.70, H3 = 3.53, H4 = 3.38, H5 = 3.69, CH_2 = 3.68 & 3.84, Ac = 2.00
MurNac	8.21				H1 = 5.45, H2 = 3.98, H3 = 3.70, H4 = 3.87, H5 = 3.38, CH_2 = 3.89 & 3.71, Ac = 2.00, OCH = 4.41, CH_3 = 1.36
Ala1	8.23	4.31	1.38		
D-Glu2	8.10	4.17	2.34	2.09, 1.90	
DAP3	8.23	4.17	1.88, 1.76	1.43	δ = 1.72, ε = 3.70
D-Ala4	8.35	4.28	1.33		
D-Ala5	8.00	4.17	1.33		
C_{15}					H1 = 4.41, H2 = 5.42, H4 = 1.72, H5 = 2.08, H6 = 2.07, H7 = 5.15, H9 = 1.66, H10 = 2.01, H11 = 2.02, H12 = 5.10, H14 = 1.64, H15 = 1.57

7.5.4 Chemical shift assignments of Oct-TriA$_1$ and *Z,Z*-C$_{15}$ LII-DAP

Table 7.22 ^{1}H chemical shift assignments for Oct-TriA$_1$ with Z,Z-C$_{15}$ LII-DAP in 6oxoDPC$_{d36}$

Residue	NH	Hα	Hβ	Hγ	Others
lipid		2.35, 2.26	1.55, 1.58	1.25	1.23, 1.21, 1.19, 0.82
D-Val1	7.81	4.27	2.1	0.9	
D-Dab2	8.67	4.81	2.18	2.96	
Gly3	8.24	4.08, 3.81			
D-Ser4	8.18	4.78	3.72		
D-Trp5	8.48	-	3.20		E1 = 10.52, D1 = 7.26, Z2 = 7.40, H2 = 7.00, Z3 = 6.93, E3 = 7.50
Ser6	8.56	4.06	3.76, 3.58		
Dab7	8.52	4.34	2.19	3.01	
D-Dab8	8.05	-	2.06, 1.89	2.82	
Phe9	8.55	4.81	3.09		D1, D2, Z = 7.21, E1, E2 = 7.13
Glu10	8.3	4.57	1.91, 1.86	2.29	
Val11	8.21	4.58	1.91	0.92	
D-*a*Ile12	8.15	4.46	1.97	1.42	γCH$_3$ = 1.18, δ = 0.90
Ala13	8.20	4.17	1.34		

Table 7.23 ^{1}H chemical shift assignments for Z,Z-C$_{15}$ LII-DAP with Oct-TriA$_1$ in 6oxoDPC$_{d36}$

Residue	NH	Hα	Hβ	Hγ	Others
GlcNAc	8.34				H1 = 4.52, H2 = 3.71, H3 = 3.53, H4 = 3.39, H5 = 3.69 , CH$_2$ = 3.68 & 3.85, Ac = 2.00
MurNac	8.13				H1 = 5.47, H2 = 3.97, H3 = 3.72, H4 = 3.91, H5 = 3.38, CH$_2$ = 3.90 & 3.71, Ac = 2.01, OCH = 4.41, CH$_3$ = 1.36
Ala1	8.25	4.33	1.4		
D-Glu2	8.15	4.19	2.11, 1.91	2.36	
DAP3	8.24	4.18	1.81, 1.88	1.41	
D-Ala4	8.36	4.28	1.34		
D-Ala5	8.04	4.2	1.34		
C$_{15}$					H1 = 4.43, H2 = 5.42, H4 = 1.72, H5 = 2.00, H6 = 2.05, H7 = 5.14, H9 = 1.65, H10 = 2.00, H11 = 2.00, H12 = 5.09, H14 = 1.62, H15 = 1.55

Table 7.24 ^{1}H chemical shift assignments for Oct-TriA$_1$ with Z,Z-C$_{15}$ LII-DAP in 10oxoDPC$_{d36}$

Residue	NH	Hα	Hβ	Hγ	Others
lipid		2.36, 2.26	1.56	1.25	1.24, 1.22, 0.83
D-Val1	7.75	4.27	2.1	0.90	
D-Dab2	8.64	4.8	2.17, 2.02	2.95	
Gly3	8.43	4.05, 3.84			
D-Ser4	8.16	4.78	3.68		
D-Trp5	8.48	-	3.18		E1 = 10.52, D1 = 7.27, Z2 = 7.42, H2 = 7.03, Z3 = 6.92, E3 = 7.50
Ser6	8.59	4.02	3.76, 3.46		
Dab7	8.53	4.33	2.18, 1.99	3.01	
D-Dab8	8.01	4.45		2.85	
Phe9	8.57	4.81	3.08		D1, D2, Z = 7.21, E1, E2 = 7.15
Glu10	8.3	4.57	1.91, 1.86	2.29	
Val11	8.21	4.58	1.89	0.92	
D-*a*Ile12	8.11	4.46	1.98	1.44	γCH$_3$ = 1.18, δ = 0.90
Ala13	8.2	4.17	1.34		

Table 7.25 ^{1}H chemical shift assignments for Z,Z-C$_{15}$ LII-DAP with Oct-TriA$_1$ in 10oxoDPC$_{d36}$

Residue	NH	Hα	Hβ	Hγ	Others
GlcNAc	8.37				H1 = 4.54, H2 = 3.70, H3 = 3.54, H4 = 3.39, H5 = 3.69 , CH$_2$ = 3.71 & 3.85, Ac = 2.00
MurNac	8.23				H1 = 5.45, H2 = 3.98, H3 = 3.71, H4 = 3.89, H5 = 3.38, CH$_2$ = 3.90 & 3.71, Ac = 2.01, OCH = 4.41, CH$_3$ = 1.36
Ala1	8.26	4.31	1.39		
D-Glu2	8.10	4.16	2.34	2.10, 1.91	
DAP3	8.24	4.18	1.79, 1.89	1.43	δ = 1.73, ε = 3.71, NH$_2$ = 8.16
D-Ala4	8.35	4.29	1.33		
D-Ala5	8.01	4.17	1.33		
C$_{15}$					H1 = 4.43, H2 = 5.43, H4 = 1.71, H5 = 2.08, H6 = 2.04, H7 = 5.15, H9 = 1.65, H10 = 2.00, H11 = 2.00, H12 = 5.10, H14 = 1.63, H15 = 1.56

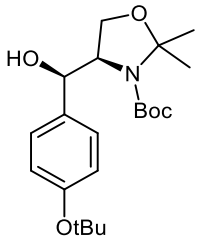
7.6 Experimental procedures for microcin J25

7.6.1 Determination of disulfide insertion location

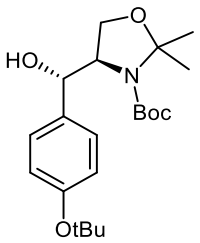
The structure coordinate file of MccJ25 (PDB: 1Q71) was uploaded into PyMOL software. Distance measurements were taken between residue Glu_8 and the terminal residues Phe_{19} or Tyr_{20}. It was determined that the distance between the β-carbon of Glu_8 and the β-carbon of Tyr_{20}, at 4.8 Å would be ideal. Additionally, it was determined both thiols would need to be in the *R* configuration as to form a disulfide bond without major changes to the native lasso structure of MccJ25.

7.6.2 Synthesis of β-thiol Tyr

***tert*-butyl (4*R*)-4-((4-(*tert*-butoxy)phenyl)(hydroxy)methyl)-2,2-dimethyloxazolidine-3-carboxylate (61)**



61a



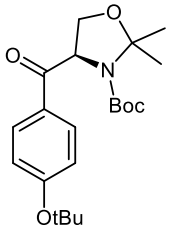
61b

This new compound was synthesized using literature precedent.[150] To a dry 3-neck flask, a stir bar, magnesium granules (230 mg, 9.4 mmol) and iodine crystals (a small spatula tip) were added. The flask was fitted to a reflux condenser and THF (2 mL) was added and stirred for 25 min at rt. A solution of 1-bromo-4-*tert*-butoxybenzene (1.12 g, 4.94 mmol) in THF (3 mL) was added dropwise over 3 min and the mixture was heated to 40 ^{o}C for 1 h. The reaction was cooled to 0 ^{o}C for 5 min and a solution

of (*R*)-Garner's aldehyde **60** (844 mg, 3.68 mmol) in THF (5 mL) was added dropwise and the reaction was stirred at 0 ^{o}C for 20 min. Next, the mixture was warmed to rt and stirred for an additional 1 h. The reaction was quenched with saturated NH_4Cl (50 mL) and extracted with EtOAc (3 × 25 mL). Combined organic layers were dried ($MgSO_4$), filtered, and concentrated *in vacuo*. The crude product was purified by flash column chromatography (1:4 EtOAc:Hex) as a clear oil (1:1 *syn:anti* diastereomeric ratio*) which hardened upon cooling (1.33 g, 95%). R_f = 0.39 (1:4 EtOAc:Hex); $[\alpha]_D^{25}$ = +76.1 (*c* = 1.10, DCM); IR (cast film, DCM): 3457, 2977, 1699, 1506, 1391, 1366, 1173, 1070 cm^{-1}; ^{1}H NMR (700 MHz, $CDCl_3$, rotamers) *anti*: δ 7.28-7.23 (m, 2H, ArH), 6.99-6.94 (m, 2H, ArH), 5.23-4.91 (m, 1H, benzylic CH), 4.76 (app s, 1H, CH_α), 4.32 (s, 1H, OH), 4.10-3.94 (m, 2H, CH_2), 1.73-1.28 (m, 24H, Boc, tBu, & diCH_3); *syn*: δ 7.28-7.23 (m, 2H, ArH), 6.99-6.94 (m, 2H, ArH), 5.30 (br s, 1H, OH), 4.73 (app s, 1H, benzylic CH), 4.18 (m, 1H, CH_α), 3.84-3.60 (m, 2H, CH_2), 1.73-1.28 (m, 24H, Boc, tBu, & diCH_3); ^{13}C NMR (175 MHz, $CDCl_3$, diastereomers, rotamers) δ 155.6, 155.1, 154.6, 154.2, 152.0, 136.6, 135.9, 127.7, 126.9, 126.2, 124.0, 123.8, 94.5, 81.0, 78.4, 78.3, 75.0, 73.2, 64.6, 63.5, 63.2, 62.9, 28.9, 28.8, 28.7, 28.3, 27.0, 26.4, 25.8, 24.3, 22.7; HRMS (ESI) calcd for $C_{21}H_{33}NO_5NA$ $[M + Na]^+$ 402.2251, found 402.2250.

*rotamers resolved by high temperature ^{1}H and ^{13}C NMR (80 ^{o}C, toluene$_{d8}$); *syn* **61a** proceeded with oxidation step below, *anti* **61b** proceeded with S_N2 using KSCN.

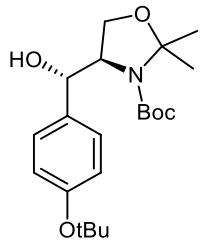
***tert*-butyl (*R*)-4-(4-(*tert*-butoxy)benzoyl)-2,2-dimethyloxazolidine-3-carboxylate (62)**



62

This new compound was synthesized using literature precedent.[150] To a solution of the undesired *syn* isomer *tert*-butyl (4*R*)-4-((*S*)-4-(*tert*-butoxy)phenyl)(hydroxy)methyl)-2,2-dimethyloxazolidine-3-carboxylate **61a** (700 mg, 1.85 mmol) in DCM (42.6 mL), freshly made and dried manganese(IV) oxide (1.6 g, 18.5 mmol) was added. The mixture was stirred at rt overnight and then passed through a cotton plug. The filtrate was concentrated *in vacuo* and crude product was purified by flash chromatography (1:4 EtOAc:Hex) to yield a while solid (550 mg, 79%). R_f = 0.51 (1:4 EtOAc:Hex); $[\alpha]_D^{25}$ = +30.9 (*c* = 0.78, DCM); IR (cast film, DCM): 3069, 2987, 1700, 1687, 1598, 1393, 1365, 1230, 1163, 1093, 1052 cm^{-1}; ^1H NMR (700 MHz, $CDCl_3$, rotamers) *major*: δ 7.84-7.86 (m, 2H, ArH), 7.05-7.01 (m, 2H, ArH), 5.33-5.31 (m, 1H, CH_α), 4.31-4.28 (m, 1H, CH_{2a}), 3.92 (dd, 1H, *J* = 3.1, 9.1 Hz, CH_{2b}), 1.75-1.24 (m, 24H, Boc, tBu, & diCH_3); *minor*: δ 7.84-7.86 (m, 2H, ArH), 7.05-7.01 (m, 2H, ArH), 5.44-5.43 (m, 1H, CH_α), 4.31-4.28 (m, 1H, CH_{2a}), 3.96-3.95 (dd, 1H, *J* = 3.9, 9.1 Hz, CH_{2b}), 1.75-1.24 (m, 24H, Boc, tBu, & diCH_3); ^{13}C NMR (175 MHz, $CDCl_3$, rotamers) *major*: δ 194.7, 160.9, 152.1, 130.0, 129.4, 95.2, 80.2, 29.8, 66.3, 61.3, 31.6, 28.2, 25.5, 24.8, *minor*: δ193.7, 160.8, 151.3, 129.7, 129.5, 94.6, 80.7, 79.9, 65.9, 61.6, 29.0, 28.4, 25.9, 24.7; HRMS (ESI) calcd for $C_{21}H_{31}NO_5$ $[M + Na]^+$ 400.2094, found 400.2094.

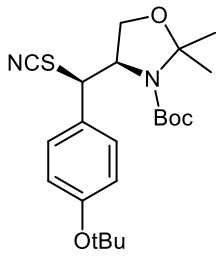
***tert*-butyl (*R*)-4-((*S*)-(4-(*tert*-butoxy)phenyl)(hydroxy)methyl)-2,2-dimethyloxazolidine-3-carboxylate (61b)**



61b

This new compound was synthesized following a literature protocol.[150] To a solution of the oxidation intermediate **62** (70 mg, 0.185 mmol), in dry THF (5.5 mL) at 0 ^{o}C, DIBAlH (1 M in Hex, 393 μL, 0.393 mmol) was added dropwise over 20 min. The reaction was stirred at 0 ^{o}C for 35 min and then quenched with cold MeOH (10 mL), followed by dropwise addition of potassium sodium tartrate (1 M, 10 mL). The mixture was allowed to stir vigorously for 5 min, then poured into cold 1 M HCl (100 mL) and extracted with EtOAc (3 × 20 mL). Combined organic layers were dried with brine (20 mL) and over $MgSO_4$ for 5 min, filtered, and concentrated *in vacuo*. The crude was purified by flash column chromatography (1:4 EtOAc:Hex) as a white solid (58 mg, 83%). R_f = 0.39 (1:4 EtOAc:Hex); $[\alpha]_D^{25}$ = +76.1 (*c* = 1.10, DCM); IR (cast film, DCM): 3457, 2977, 1699, 1506, 1391, 1366, 1173, 1070 cm^{-1}; ^{1}H NMR (700 MHz, $CDCl_3$, rotamers) *anti*: δ 7.28-7.23 (m, 2H, ArH), 6.99-6.94 (m, 2H, ArH), 5.23-4.91 (m, 1H, benzylic CH), 4.76 (app s, 1H, CH_α), 4.32 (s, 1H, OH), 4.10-3.94 (m, 2H, CH_2), 1.73-1.28 (m, 24H, Boc, tBu, & diCH_3); ^{13}C NMR (175 MHz, $CDCl_3$, rotamers) δ 155.6, 154.2, 152.0, 136.6, 135.9, 127.7, 126.2, 123.8, 94.5, 81.0, 78.4, 73.2, 63.2, 62.9, 28.8, 27.0, 24.3, 22.7; HRMS (ESI) calcd for $C_{21}H_{33}NO_5NA$ [M + Na]$^+$ 402.2251, found 402.2252.

tert-butyl (*R*)-4-((*R*)-(4-(*tert*-butoxy)phenyl)(thiocyanato)methyl)-2,2 dimethyloxazolidine-3-carboxylate (63)

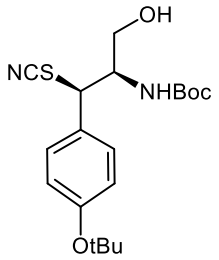


63

This new compound was synthesized using literature precedent.[150] *tert*-butyl (*R*)-4-(4-(*tert*-butoxy)benzoyl)-2,2-dimethyloxazolidine-3-carboxylate **61b** (600 mg, 1.58 mmol) and DCM (7 mL) were added to a dry flask placed in a 0 ^{o}C ice bath. Next, Et_3N (183.5 µL, 1.90 mmol) was added dropwise followed by the dropwise addition of methasulfonyl chloride (321.9 µL, 2.37 mmol)). The reaction was monitored by TLC and stirred at 0 ^{o}C until starting material was consumed (~3 h). The solution was diluted with DCM (25 mL) and poured into a separatory funnel containing saturated NH_4Cl (40 mL). The organic layer was washed with water (40 mL), brine (40 mL), and dried over $MgSO_4$. The mixture was filtered, concentrated *in vacuo* to a crude yellow oil, which was carried onto the next step without further purification. Dry ACN (17 mL) and crude mesylate intermediate (710 mg) were added to a 3-neck flask containing a stir bar. Oven-dried KSCN (770 mg, 7.91 mmol) was added and the solution was stirred at rt for 2 d. The solution was concentrated *in vacuo* and reconstituted in DCM (20 mL) and poured into a separatory funnel containing water (40 mL). Using DCM (3 × 20 mL), the aqueous layer was extracted and the combined organic layers were dried with brine (20 mL) and over $MgSO_4$. Mixture was filtered, concentrated *in vacuo*, and purified using flash chromatography (1:9 to 1:4

EtOAc:Hex) to yield a yellow oil (156.1 mg, 23%). $[\alpha]_D^{25}$ = +82.5 (*c* = 1.10 , DCM); IR (cast film, DCM): 2979, 2936, 2882, 2155 (SCN), 1797, 1607, 1507, 1377, 1206, 1165 cm^{-1}; ^1H NMR (600 MHz, $CDCl_3$, rotamers) δ 7.32-7.28 (m, 2H, ArH), 7.01-6.97 (m, 2H, ArH), 5.20-5.04 (m, 1H, benzylic CH), 4.40-4.26 (m, 1H, CH_α), 4.12-4.10 (app d, *J* = 17.8 Hz, CH_{2a}) 3.94-3.89 (m, 1H, CH_{2b}), 1.61-1.35 (m, 24H, Boc, tBu, & diCH_3); ^{13}C NMR (175 MHz, $CDCl_3$, rotamers) δ 171.2, 156.0, 152.7, 130.6, 128.8, 128.5, 124.2, 124.1, 94.8, 81.4, 78.8, 64.0, 61.8, 61.5, 56.3, 55.4, 28.9, 28.6, 28.4, 26.6, 24.1; HRMS (ESI) calcd for $C_{22}H_{32}N_2O_4S$ $[M + Na]^+$ 443.1975, found 443.1975.

tert-butyl ((1*R*,2*R*)-1-(4-(_tert_-butoxy)phenyl)-3-hydroxy-1-thiocyanatopropan-2-yl)carbamate (64)

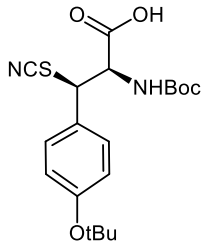


64

This new compound was synthesized using literature precedent.[150] To a flask containing dry 1,4-dioxane (3.8 mL), tert-butyl (*R*)-4-((*R*)-(4-(*tert*-butoxy)phenyl)(thiocyanato)methyl)-2,2 dimethyloxazolidine-3-carboxylate **63** (97.0 mg, 0.23 mmol) and *para*-toluenesulfonic acid (57.9 mg, 0.53 mmol) were added and the reaction was stirred for 3 h at rt. The reaction was quenched with saturated $NaHCO_3$ (10 mL) and the mixture was extracted with EtOAc (3 × 10 mL). The

combined organic layers were dried over $MgSO_4$, filtered, and concentrated *in vacuo*. The resulting crude material was purified using flash chromatography (30% to 50% EtOAc in Hex) to yield a white solid (45.9 mg, 0.12 mmol, 52%). $[\alpha]_D^{25}$ = +68.6 (*c* = 1.00, DCM); IR (cast film, DCM): 3315, 2978, 2934, 2880, 2153 (SCN), 2069, 1744, 1509,1475, 1419, 1392, 1367, 1241, 1162, 1050, 897, 675 cm^{-1}; ^1H NMR ($CDCl_3$, 600 MHz) δ 7.29-7.23 (m, 2H, ArH), 6.98-6.95 (m, 2H, ArH), 5.15 (br s, 1H, OH), 4.71 (d, *J* = 9.3 Hz, 1H, benzylic CH), 4.12 (d, *J* = 6.4 Hz, 1H, CH_α), 3.57 (dd, *J* = 11.3, 3.4 Hz, 1H, CH_{2a}), 3.36 (d, *J* = 10.9 Hz, 1H, CH_{2b}), 1.46-1.39 (m, 9H, Boc), 1.33 (s, 9H, tBu); ^{13}C NMR ($CDCl_3$, 125 MHz) δ 171.4, 156.3, 131.0, 128.9, 128.7, 126.7, 124.3, 112.0, 80.6, 79.8, 62.0, 60.5, 56.1, 28.9, 28.3; HRMS (ESI) calcd for $C_{22}H_{32}N_2O_4S$ $[M + Na]^+$ 403.1662, found 403.1661.

(2*R*,3*R*)-3-(4-(tert-butoxy)phenyl)-2-((tert-butoxycarbonyl)amino)-3-thiocyanatopropanoic acid (65)



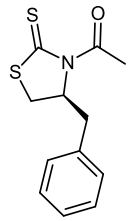
65

This new compound was synthesized using literature precedent.[150] To a RBF containing *tert*-butyl ((1*R*,2*R*)-1-(4-(*tert*-butoxy)phenyl)-3-hydroxy-1-thiocyanatopropan-2-yl)carbamate **64** (27.1 mg, 71.2 μmol), DMF (1.4 mL) and pyridinium dichromate (269 mg, 71.2 μmol) were added. The reaction was stirred at

rt; reaction completion was monitored by TLC and citric acid (10%, 5 mL) was added once all starting material had been consumed (~20 h). The mixture was extracted with EtOAc (3 × 10 mL) and combined organic layers were washed with brine (10 mL) and dried over $MgSO_4$. Flash column chromatography (5:95 MeOH:DCM) was used to purify compound **65** as pale yellow solid (15.2 mg, 39.9 μmol, 56%). $[\alpha]_D^{25}$ = +63.12 (*c* = 1.9, $CHCl_3$); IR ($CHCl_3$, cast film) 3115, 2970, 2935, 2157 (SCN), 1762, 1740, 1512, 1470, 1426, 1381, 1238, 1164, 1051, 899, 673 cm^{-1}; ^1H NMR ($CDCl_3$, 500 MHz) δ 7.25 (d, *J* = 7.5 Hz, 2H, ArH), 6.84 (d, *J* = 7.5 Hz, 2H, ArH), 4.94 (d, *J* = 9.3 Hz, 1H, benzylic CH), 4.86 (d, *J* = 6.4 Hz, 1H, CH_α), 1.36-1.28 (m, 18H, Boc & tBu); ^{13}C NMR ($CDCl_3$, 125 MHz) δ 170.1, 155.8, 155.6, 133.2, 128.4, 118.5, 114.4, 86.2, 79.6, 69.6, 44.3, 28.4, 27.2; HRMS (ESI) Calcd for $C_{19}H_{26}N_2NaO_5S$ $[M+Na]^+$ 417.1561, found 417.1563.

7.6.3 Synthesis of β-thiol Glu

(*S*)-1-(4-benzyl-2-thioxothiazolidin-3-yl)ethan-1-one (69)

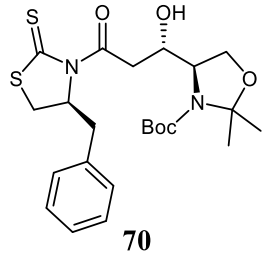


69

This known compound was made following a literature protocol.[152] To a round bottom flask equipped with a stir bar, (*S*)-4-benzylthiazolidine-2-thione (10.0 g, 47.77 mmol) and pyridine (29.1 mL, 286.63 mmol) were added and the mixture was stirred at rt until dissolved. Acetic anhydride (22.6 mL, 238.86 mmol) was added and

solution turned a bright yellow. The reaction was refluxed at 135 ^{o}C for 1.5 h, during which the mixture turned dark brown, and then concentrated *in vacuo*. The crude was dissolved in EtOAc (250 mL) and washed with 10% K_2CO_3 (3 × 100 mL), 1 M HCl (3 x 100 mL), and brine (1 × 100 mL). The organic layer was dried over $MgSO_4$ and concentrated *in vacuo*. The crude yellow-brown solid was purified by flash chromatography (1:9 EtOAc:Hex) to yield bright yellow crystals (11.1 g, 93% yield). R_f = 0.38 (1:4 EtOAc: Hex); $[\alpha]_D^{25}$ = +248.02 (*c* = 0.58, DCM); IR (cast film, DCM): 3026, 2927, 2856, 2695, 1953(C=S), 1888(C=S), 1694, 1454, 1366, 1341, 1275, 1209, 1022, 848, 747 cm^{-1}; ^{1}H NMR (700 MHz, $CDCl_3$) δ 7.36-7.34 (m, 2H, ArH), 7.30-7.27 (m, 3H, ArH), 5.38 (ddd, *J* = 10.5, 7.0, 3.5 Hz, 1H, CH), 3.38 (dd, *J* = 11.2, 7.2 Hz, 1H, $PhCH_2$) 3.22 (dd, *J* = 13.0, 3.9 Hz, 1H, SCH_2), 3.04 (dd, *J* = 13.0, 10.9 Hz, 1H, SCH_2), 2.89 (d, *J* = 11.9 Hz, 1H, $PhCH_2$), 2.80 (s, 3H, CH_3); ^{13}C NMR (175 MHz, CDCl) δ 201.6, 179.8, 136.6, 129.5, 129.0, 127.3, 68.3, 36.8, 31.9, 27.1; HRMS (EI) calcd for $C_{12}H_{13}NOS_2$ $[M]^{+\cdot}$ 251.0439, found 251.0437.

tert-butyl (*R*)-4-((*S*)-3-((*S*)-4-benzyl-2-thioxothiazolidin-3-yl)-1-hydroxy-3-oxopropyl)-2,2-dimethyloxazolidine-3-carboxylate (70)



70

This new compound was synthesized using a method with literature precedent.[150] To a 250 mL dried RBF under argon equipped with a stir bar, (*S*)-1-(4-benzyl-2-

thioxothiazolidin-3-yl)ethan-1-one **69** (500 mg, 1.99 mmol) and DCM (82.9 mL) were added. The solution was placed in a -78 ^{o}C IPA bath and titanium tetrachloride (2.19 mL, 2.19 mmol, 1 M in Hex) was added. A solution of (–)-sparteine (1.00 mL, 4.38 mmol) in DCM (6 mL) was added dropwise over 15 min, the solution turned red, and was stirred for 2 h at -78 ^{o}C. Garner's aldehyde (415 μL, 1.99 mmol) in DCM (6 mL) was added dropwise and the solution was stirred for 1 h at -78 ^{o}C. Citric acid (10%, 30 mL) was added to quench the reaction and the reaction flask was warmed to rt. The solution was extracted with EtOAc (3 × 10 mL) and combined organic layers were dried with brine (20 mL), $MgSO_4$, and solvent was removed *in vacuo*. The resulting crude was purified using flash chromatography (1:4 to 2:3 EtOAc:Hex) to yield both isomers (quant) as yellow solids; the desired *S*-hydroxyl isomer was isolated (670 mg, 1.39 mmol, 70% yield). R_f = 0.58 (1:4 EtOAc:Hex); $[\alpha]_D^{25}$ = +212.12 (*c* = 0.36, $CDCl_3$); IR (cast film, $CDCl_3$): 3459, 2979, 2934, 2885, 1694, 1455, 1366, 1258, 1167, 1988, 1066, 852, 755 cm^{-1}; ^{1}H NMR (600 MHz, $CDCl_3$, rotomers) δ 7.35-7.28 (m, 5H, ArH), 5.39-5.36 (m, 1H, R_2NCH), 4.53-4.49 (m, 1H), 4.16-3.97 (m, 3H, rotomers), 3.57-3.54 (m, 1H), 3.41 (dd, *J* = 11.0, 7.2 Hz,1H, SCH_{2a}), 3.26 (d, *J* = 12.4 Hz, 1H, $PhCH_{2a}$), 3.05 (app t, *J* = 11.8 Hz, 1H, $PhCH_{2a}$), 2.89 (d, *J* = 11.4 Hz, 1H, SCH_{2a}), 1.49 (s, 15H, Boc & $diCH_3$); ^{13}C NMR (172 MHz, $CDCl_3$) δ 201.1, 171.3, 169.2, 136.7, 136.0, 129.5, 129.0, 127.3, 99.3, 81.2, 68.5, 68.0, 65.0, 47.1, 41.7, 36.8, 32.0, 29.8, 28.8, 18.8; HRMS (ESI) calcd for $C_{22}H_{32}N_2NaO_5S_2$ $[M + Na]^+$ 503.1645, found 503.1643.

References

1. van Leeuwenhoek A. Observations, communicated to the publisher by Mr. Antoni van Leewenhoeck, in a Dutch letter of the 9th Octob. 1676. here English'd: concerning little animals by him observed in rain-well-sea- and snow water; as also in water wherein pepper had lain infused. *Phil Trans R Soc.* **1677**, *12*, 821-831.
2. (a) Messaoudi M, Lalonde R, Violle N, Javelot H, Desor D, Nejdi A, Bisson JF, Rougeot C, Pichelin M, Cazaubiel M, Cazaubiel J. Assessment of psychotropic-like properties of a probiotic formulation (*Lactobacillus helveticus* R0052 and *Bifidobacterium longum* R0175) in rats and human subjects. *Brit J Nutrition*. **2011**, *105*, 755-764. (b) O'Mahony L, McCarthy J, Kelly P, Hurley G, Luo F, Chen K, O'Sullivan GC, Kiely B, Collins JK, Shanahan F, Quigley EM. *Lactobacillus* and *bifidobacterium* in irritable bowel syndrome: symptom responses and relationship to cytokine profiles. *Gastroenterology*. **2005**, *128*, 541-551. (c) Nguyen PW, Botyanszki Z, Tay PK, Joshi NS. Programmable biofilm-based materials from engineered curli nanofibers. *Nat Commun.* **2014**, *5*, 4945.
3. Pham VJ, Yilma MA, Feliz A, Majid MT, Maffetone N, Walker JR, Kim E, Cho HJ, Reynolds JM, Song MC, Park SR, Yoon YJ. A Review of the microbial production of bioactive natural products and biologics. *Front Microbiol.* **2019**, 10, 1404.
4. Lei J, Sun L, Huang S, Zhu C, Li P, He J, Mackey V, Coy CH, He Q. The antimicrobial peptides and their potential clinical applications. *Am J Transl Res*. **2019**, *11*, 3919-3931.

5. (a) Bérdy, J. Thoughts and facts about antibiotics: where we are now and where we are heading. *J. Antibiot. (Tokyo)* **2012**, *65*, 385-395. (b) Wellcome Trust Review on Antimicrobial Resistance. Tackling drug-resistant infections globally: final report and recommendations. **2016**. (c) World Health Organization. Antimicrobial resistance: global report on surveillance. Geneva, Switzerland: WHO **2014**.

6. Fleming A. Classics in infectious diseases: on the antibacterial action of cultures of a penicillium, with special reference to their use in the isolation of *B. influenzae*. *Br. J. Exp. Pathol.* **1929**, *10*, 226-236.
7. Vetola CL. The antibiotic resistance crisis. *Pharmacy & Therapueitcs.* **2015**, *40*, 277-283.
8. Mahlapuu M, Håkansson J, Ringstand L, Björn C. Antimicrobial peptides: an emerging category of therapeutic agents. *Front Cell Infect Microbiol.* **2019**, *6*, 194.
9. Hassan M, Kjos M, Nes IF, Diep DB, Lotfipour F. Natural antimicrobial peptides from bacteria: characteristics and potential applications to fight against antibiotic resistance. *J Appl Microbiol.* **2012**, *113*, 723-736.
10. Haspser HE, de Kruijff B, Breukink E. Assembly and stability of nisin-lipid II pores. *Biochemistry*. **2004,** *43*, 11567-11575.
11. Mukhopadhyay J, Sineva E, Knight J, Levy RM, Ebright RH. Antibacterial peptide microcin J25 inhibits transcription by binding within and obstructing the RNA polymerase secondary channel. *Mol Cell*. **2004,** *14*, 739-751.

12. Cotter PD, Hill C, Ross P. Bacteriocins: developing innate immunity for food. *Nat Rev Micro*. **2005**, *3*, 777-788.
13. Arnison PG *et al.* Ribosomally synthesised and post-translationally modified peptides natural products: overview and recommendations for a universal nomenclature. *Nat Prod Rep*. **2013**, *30*, 108-160.
14. Chiorean S, Vederas JC, van Belkum MJ. Identification and heterologous expression of the *sec*-dependent bacteriocin faerocin MK from *Enterococcus faecium* M3K31. *Probiotics Antimicrob Proteins*. **2018**, *10*, 142-147.
15. Voet D, Voet J. Biochemistry. *4th ed*. John Wiley & Sons, Inc.: Hoboken, NY. **2010**.
16. (a) Worobo RW, van Belkum MJ, Sailer M, Roy KL, Vederas JC, Stiles ME. A signal peptide secretion-dependent bacteriocin from *Carnbacterium divergens. J bacterial.* **1995**, *177*, 3143-3149. (b) Kelly WL, Pan L, Li C. Thiostrepton biosynthesis: prototype for a new family of bacteriocins. *J Am Chem Soc.* **2009**, *131*, 4327-4334.
17. Omar TJ, van der Bonk WA. Follow the leader: the use of leader peptides to guide natural product biosynthesis. *Nat Chem Biol.* **2010***, 6*, 9-18.
18. Tsirigotaki A, De Geyter J, Šoštaric, N. Economou A, Karamanou S. Protein export through the bacterial *Sec* pathway. *Nat Rev Microbiol.* **2017**, *15*, 21-36.
19. Van Belkum MJ, Worobo RW, Stiles, ME. Double-glycine-type leader peptides direct secretion of bacteriocins by ABC transporters: colicin V secretion in *Lactococcus lactis*. *Mol Microbiol.* **1997**, *23***,** 1293-1301.

20. Schneider E, Hunke S. ATP-binding-cassette (ABC) transport systems: functional and structural aspects of the ATP-hydrolyzing subunits/domains. *FEMS Microbiol Rev.* **1998***, 22*, 1-20.

21. Håvarstein LS, Diep DB, Nes IF. A family of bacteriocin ABC transporters carry out proteolytic processing of their substrates concomitant with export. *Mol Microbiol.* **1995***, 16*, 229-240.

22. Acedo JZ, Chiorean S, Vederas JC, van Belkum MJ, The expanding structural variety among bacteriocins from Gram-positive bacteria. *FEMS Microbiol Rev*. **2018***, 42*, 805-828.

23. Ge P, Scholl D, Leiman PG, Yu X, Miller JF, Zhou ZH. Atomic structures of a bactericidal contractile nanotube in its pre- and postcontraction states. *Nat Struct Mol Biol.* **2015**, *22*, 377-382.

24. Hammami R, Zouhir A, Le Lay C, Ben Hamida J, Fliss, I. BACTIBASE second release: a database and tool platform for bacteriocin characterization. *BMC Microbiol.* **2010**, *10*, 22.

25. (a) Duquesne S, Destoumieux-Garzón D, Peduzzi J, Rebuffat S. Microcins, gene-encoded antibacterial peptides from enterobacteria. *Nat Prod Rep.* **2007**, *24*, 708-734. (b) Baquero F, Lanza VF, Baquero MR, Del Campo R, Bravo-Vázquez DA. Microcins in *Enterobacteriaceae*: peptide antimicrobials in the eco-active intestinal chemosphere. *Front Microbiol.* **2019**, *10*, 2261.

26. (a) Pugsley AP. The ins and outs of colicins. Part I: production, and translocation across membranes. *Microbiol Sci.* **1984**, *1*, 168-175. (b) Pugsley AP. The ins and

outs of colicins. Part II. Lethal action, immunity and ecological implications. *Microbiol Sci.* **1984**, *1*, 203-205.

27. Michel-Briand Y, Baysse C. The pyocins of *Pseudomonas aeruginosa*. *Biochimie*. **2002**, *84*, 499-510.
28. Mach B, Reich E, Tatum EL. Separation of the biosynthesis of the antibiotic polypeptide tyrocidine from protein biosynthesis. *Proc Natl Acad Sci USA*. **1963,** *50*, 175-181.
29. Süssmuth RD, Mainz A. Nonribosomal peptide synthesis – principles and prospects. *Angew Chem Int Ed.* **2017**, *56*, 3770-3821.
30. Ziemert N, Alanjary M, Weber T. The evolution of genome mining in microbes – a review. *Nat Prod Rep.* **2016**, *33*, 988-1005.
31. Hur GH, Vickery CR, Burkart MD. Explorations of catalytic domains in non-ribosomal peptide synthetase enzymology. *Nat Prod Rep*. **2012**, *29*, 1074-1098.
32. Walsh CT, Chen H, Keating TA, Hubbard BK, Losey HC, Luo L, Marshall CG, Miller DA, Patel HM. Tailoring enzymes that modify nonribosomal peptides during and after chain elongation on NRPS assembly lines. *Curr Opin Chem Biol*. **2001**, *5*, 525-534.
33. Schmartz PC, Zerbe K, Abou-Hadeed K, Robinson JA. Bis-chlorination of a heptapeptide-PCP conjugate by halogenase involved in vancomycin biosynthesis. *Org Biomol Chem*. **2014**, *12*, 5574-5577.
34. Finking R, Marahiel MA. Biosynthesis of nonribosomal peptides. *Annu Rev Microbiol.* **2004**, *58*, 453-488.

35. (a) Gang DH, Kim DQ, Park, HS. Cyclic peptides: promising scaffolds for biopharmaceuticals. *Genes (Basel)*. **2018**, *9*, 557. (b) Driggers EM, Hale SP, Lee J, Terrett NK. The exploration of macrocycles for drug discovery--an underexploited structural class. *Nat Rev Drug Discov*. **2008**, 7, 608-624.
36. Cochrane SA, Vederas JC. Lipopeptides from Bacillus and Paenibacillus spp.: a gold mine of antibiotic candidates. *Med Res Rev.* **2016**, *36*, 4-31.
37. Williamson R, Hakenbeck R, Tomasz A. *In vivo* interaction of beta-lactam antibiotics with the penicillin-binding proteins of *Streptococcus pneumoniae*. *Antimicrob Agents Chemother*. **1980,** *18,* 629-637.
38. (a) Bednarska NG, Wren BW, Willcocks SJ. The importance of the glycosylation of antimicrobial peptides: natural and synthetic approaches. *Drug Discov Today.* **2017,** *22*, 919-926. (b) Křen V, Řezanka T. Sweet antibiotics – the role of glycosidic residues in antibiotic and antitumor activity and their randomization. *FEMS Microbiol Rev.* **2008**, *32*, 858-889.
39. Leclère V, Weber T, Jacques P, Pupin M. Bioinformatics tools for the discovery of new nonribosomal peptides. *Methods Mol Biol*. **2016**, *1401*, 209-232.
40. Staley JT, Konopka A. Measurements of in situ activities of nonphotosynthetic microorganisms in aquatic and terrestrial habitats. *Ann Rev Microbiol*. **1985**, *39***,** 321-346.
41. Connon SA, Giovannoni SJ. High-throughput methods for culturing microorganisms in very-low-nutrient media yield diverse new marine isolates. *Appl Environ Microbiol*. **2002**, *68*, 3878-3885.

42. Katz M, Hover BM, Brady SF. Culture-independent discovery of natural products from soil metagenomes. *J Ind Microbiol Biotechnol*. **2016**, *43*, 129-141. (b) Stevenson LJ, Owen JG, Ackerly DF. Metagenome driven discovery of nonribosomal peptides. *ACS Chem. Biol.* **2019**, *14*, 2115-2126.

43. (a) Owen JG, Charlop-Powers Z, Smith AG, Ternei MA, Calle PY, Reddy BVB, Montiel D, Brady SF. Multiplexed metagenome mining using short DNA sequence tags facilitates targeted discovery of epoxyketone proteasome inhibitors. *Proc Natl Acad Sci USA*. **2015**, *112*, 4221-4226. (b) Hover BM, Kim SH, Katz M, Charlop-Powers Z, Owen JG, Ternei MA, Maniko J, Estrela AB, Molina H, Park S, Perlin DS, Brady SF. Culture-independent discovery of the malacidins as calcium-dependent antibiotics with activity against multidrug-resistant Gram-positive pathogens. *Nat Microbiol.* **2018**, *3*, 415-422. (c) Wu C, Shang Z, Lemetre C, Ternei MA, Brady SF. Cadasides, Calcium-dependent acidic lipopeptides from the soil metagenome that are active against multidrug-resistant bacteria. *J Am Chem Soc.* **2019**, *141*, 3910-3919.

44. Nichols D, Cahoon N, Trakhtenberg EM, Pham L, Mehta A, Belanger A, Kanigan T, Lewis K, Epstein SS. Use of ichip for high-throughput *in situ* cultivation of "uncultivable" microbial species. *App. Environ Microbiol.* **2010**, *76*, 2445-2450.

45. Ling LL, Schneider T, Peoples AJ, Spoering AL, Engels I, Conlon BP, Mueller A, Schäberle TF, Hughes DE, Epstein S, Jones M, Lazarides L, Steadman VA, Cohen DR, Felix CR, Fetterman KA, Millett WP, Nitti AG, Zullo AM, Chen C, Lewis K. A new antibiotic kills pathogens without detectable resistance. *Nature*. **2015**, *517*, 455-459.

46. Henninot A, Collins JC, Nuss JM. The current state of peptide drug discovery: back to the future? *J Med Chem.* **2018**, *61*, 1382-1414.

47. Munita JM, Arias CA. Mechanisms of antibiotic resistance. *Microbiol Spectr*. **2016**, *4*, 1-37.

48. Ramirez MS, Tolmasky ME. Aminoglycoside modifying enzymes. *Drug Resist Updat*. **2010**, *13*, 151-171.

49. Abraham EP, Chain E. An enzyme from bacteria able to destroy penicillin 1940. *Rev Infect Dis.* **1988**, *10*, 677-678.

50. Pagès JM, James CE, Winterhalter M. The porin and the permeating antibiotic: a selective diffusion barrier in Gram-negative bacteria. *Nat Rev Microbiol.* **2008**, *6*, 893-903.

51. McMurry L, Petrucci RE Jr, Levy SB. Active efflux of tetracycline encoded by four genetically different tetracycline resistance determinants in *Escherichia coli. Proc Natl Acad Sci USA*. **1980**, *77*, 3974-3977.

52. (a) Poole K. Efflux-mediated antimicrobial resistance. *J Antimicrob Chemother*. **2005**, *56*, 20-51. (b) Piddock LJ. Clinically relevant chromosomally encoded multidrug resistance efflux pumps in bacteria. *Clin Microbiol Rev*. **2006**, *19*, 382-402.

53. Zhou H, Fan J, Tian Y, Lu ZY. Mechanisms of nisin resistance in Gram-positive bacteria. *Ann Microbiol.* **2014**, *64*, 413-420.

54. Hooper DC. Fluoroquinolone resistance among Gram-positive cocci. *Lancet Infect Dis*. **2002**, *2*, 530-538.

55. Long KS, Vester B. Resistance to linezolid caused by modifications at its binding site on the ribosome. *Antimicrob Agents Chemother*. **2012**, *56*, 603-612.

56. Golkar T, Zieliński M, Berghuis AM. Look and outlook on enzyme-mediated macrolide resistance. *Front Microbiol.* **2018**, *9*, 1942.

57. Hiramatsu K, Ito T, Tsubakishita S, Sasaki T, Takeuchi F, Morimoto Y, Katayama Y, Matsuo M, Kuwahara-Arai K, Hishinuma T, Baba T. Genomic basis for methicillin resistance in *Staphylococcus aureus. Infect Chemother*. **2013**, *45*, 117-136.

58. Arthur M, Courvalin P. Genetics and mechanisms of glycopeptide resistance in enterococci. *Antimicrob Agents Chemother*. **1993**, *37*, 1563-1571.

59. Livermore DM. Mechanisms of resistance to cephalosporin antibiotics. *Drugs*. **1987**, *34*, S64-88.

60. Hooper DC, Wolfson JS, Ng EY, Swartz MN. Mechanisms of action of and resistance to ciprofloxacin. *Am J Med*. **1987,** *82*, 12-20.

61. Chopra I, Roberts M. Tetracycline antibiotics: mode of action, applications, molecular biology, and epidemiology of bacterial resistance. *Microbiol Mol Biol Rev*. **2001**, *65*, 232-260.

62. Fosgerau K, Hoffmann T. Peptide therapeutics: current status and future directions. *Drug Discov Today.* **2015**, *20*, 122-128.

63. Craik DJ, Fairlie DP, Liras S, Price D. The future of peptide-based drugs. *Chem Biol Drug Des*. **2013**, *81*, 136-147.

64. Statistics Market Research Consulting Pvt Ltd.: Peptide Therapeutics - Global Market Outlook (2019-2027), June **2020**, ID: 5116326.

65. Vlieghe P, Lisowski V, Martinez J, Khrestchatisky M. Synthetic therapeutic peptides: science and market. *Drug Discov Today.* **2010**, *15*, 40-56.

66. Fjell C, Hiss J, Hancock R, Schneider G. Designing antimicrobial peptides: form follows function. *Nat Rev Drug Discov.* **2012**, *11***,** 37-51.

67. Sutherland R, Croydon EA, Rolinson GN. Flucloxacillin, a new isoxazolyl penicillin, compared with oxacillin, cloxacillin, and dicloxacillin. *Br Med J*. **1970**, *4*, 455-460.

68. Liu J, Smith PA, Barrios Steed D, Romesberg F. Efforts toward broadening the spectrum of arylomycin antibiotic activity. *Bioorg Med Chem Lett.* **2013**, *23*, 5654-5659.

69. Wenzler E, Rodvold KA, Telavancin: the long and winding road from discovery to food and drug administration approvals and future directions. *Clin Infect Dis*. **2015**, *61*, S38-S47.

70. Weiss GA, Watanabe CK, Zhong A, Goddard A, Sidhu SS. Rapid mapping of protein functional epitopes by combinatorial alanine scanning. *Proc Natl Acad Sci.* **2000**, *97*, 8950-8954.

71. Goodwin D, Simerska P, Toth I. Peptides as therapeutics with enhanced bioactivity. *Curr Med Chem*. **2012**, *19,* 4451-4461.

72. Werner HM, Cabalteja CC, Horne WS. Peptide backbone composition and protease susceptibility: impact of modification type, position, and tandem substitution. *ChemBioChem.* **2016**, *17*, 712-718.

73. Di L. Strategic approaches to optimizing peptide ADME properties. *AAPS J.* **2015**, *17*, 134-143.

74. Frey V, Viaud J, Subra G, Cauquil N, Guichou JF, Casara P, Grassy GR, Chavanieu A. Structure-activity relationships of Bak derived peptides: affinity and specificity modulations by amino acid replacement. *Eur J Med Chem.* **2008**, *43*, 966-972.

75. Drider D, Fimland G, Héchard Y, McMullen LM, Prévost H. The continuing story of class IIa bacteriocins. *Microbiol Mol Biol Rev.* **2006**, *70*, 564-582.

76. Arbulu S, Frantzen C, Lohans CT, Cintas LM, Herranz C, Holo H, Diep DB, Vederas JC, Hernández PE. Draft genome sequence of the bacteriocin-producing strain *Enterococcus faecium* M3K31, isolated from griffon vultures (*Gyps fulvus* subsp. *fulvus*). *Genome Announc.* **2016**, *4,* e00055-e00016.

77. Svetoch EA, Stern NJ, Eruslanov BV, Kovalev YN, Volodina LI, Perelygin VV, Mitsevich EN, Mitsevich IP, Pokhilenko VD, Borzenkov VN, Levchuk VP, Svetoch OE, Kudriavtseva TY. Isolation of *Bacillus circulans* and *Paenibacillus polymyxa* strains inhibitory to *Campylobacter jejuni* and characterization of associate bacteriocins. *J Food Prot*. **2005**, *68*, 11-17.

78. Lohans CT, van Belkum MJ, Li J, Vederas JC. Characterization of bacterial antimicrobial peptides active against *Campylobacter jejuni. Can J Chem* **2015**, *93*, 381-388.

79. van Belkum MJ, Lohans CT, Vederas JC. Draft genome sequences of *Paenibacillus polymyxa* NRRL B-30509 and *Paenibacillus terrae* NRRL B-30644, strains from a poultry environment that produce tridecaptin A and paenicidins. *Genome Announc.* **2015**, *3*, e00372-15.

80. Lohans CT, Huang Z, van Belkum MJ, Giroud M, Sit CS, Steels EM, Zheng J, Whittal RM, McMullen LM, Vederas JC. Structural characterization of the highly cyclized lantibiotic paenicidin A via a partial desulfurization/reduction strategy. *J Am Chem Soc.* **2012**, *134*, 19540-19543.

81. Gravesen A, Ramnath M, Rechinger KB, Andersen N, Jänsch L, Héchard Y, Hastings JW, Knøchel S. High-level resistance to class IIa bacteriocins is associated with one general mechanism in Listeria monocytogenes. *Microbiology.* **2002**, *148*, 2361-2369.

82. Sievers F, Higgins DG. Clustal omega. *Curr Protoc Bioinformatics.* **2014***, 48*, 3.13.1-3.13.16.

83. Wang Y, Henz M, Gallagher NL, Chai S, Gibbs A, Yan LZ, Stiles M, Wishart D, Vederas JC. Solution structure of carnobacteriocin B2 and implications for structure-activity relationships among type IIa bacteriocins from lactic acid bacteria. *Biochemistry*. **1999**, *38*, 15438-15448.

84. Du L, Somkuti GA, Renye JA Jr. Molecular analysis of the bacteriocin-encoding plasmid pDGL1 from *Enterococcus durans* and genetic characterization of the durancin GL locus. *Microbiology*. **2012**, *158*, 1523-1532.

85. Sánchez J, Diep DB, Herranz C, Nes IF, Cintas LM, Hernández PE. Amino acid and nucleotide sequence, adjacent genes, and heterologous expression of hiracin JM79, a *sec*-dependent bacteriocin produced by *Enterococcus hirae* DCH5, isolated from Mallard ducks (*Anas platyrhynchos*). *FEMS Microbiol Lett.* **2007**, *270*, 227-236.

86. Derksen DJ, Boudreau MA, Vederas JC. Hydrophobic interactions as substitutes for a conserved disulfide linkage in the type IIa bacterioicns, leucocinA and pediocin PA-1. *ChemBioChem.* **2008**, *9*, 1898-1901.

87. Greenfield NJ, Fasman GD. Computed circular dichroism spectra for the evaluation of protein conformation. *Biochemistry*. **1969**, *8*, 4108-4116.

88. Fregeau Gallagher NL, Sailer M, Niemczura WP, Nakashima TT, Stiles ME, Vederas JC. Three-dimensional structure of leucocin A in trifluoroethanol and dodecylphosphocholine micelles: spatial location of residues critical for biological activity in type IIa bacteriocins from lactic acid bacteria. *Biochemistry*. **1997**, *36*, 15062-15072.

89. Buck M. Trifluoroethanol and colleagues: cosolvents come of age. Recent studies with peptides and proteins. *Q Rev Biophys*. **1998,** *31*, 297-355.

90. Reiersen H, Rees AR. Trifluoroethanol may form a solvent matrix for assisted hydrophobic interactions between peptide side chains. *Protein Eng*. **2000**, *13*, 739-743.

91. Morrow JA, Segall ML, Lund-Katz S, Phillips MC, Knapp M, Rupp B, Weisgraber KH. Differences in stability among the human apolipoprotein E isoforms determined by the amino-terminal domain. *Biochemistry.* **2000**, *39*, 11657-11666.

92. Ben Belgacem Z, Rehaiem A, Fajardo Bernárdez P, Manai M, Pastrana Castro L. Interactive effects of pH and temperature on the bacteriocin stability by response surface analysis. *Microbiology.* **2012**, *8*, 195-200.

93. Muskens MW, Vissers AP, Mol JN, Kooter JM. Role of inverted DNA repeats in transcriptional and post-transcriptional gene silencing. *Plant Mol Biol.* **2000**, *43*, 243-260.

94. Butt TR, Edavettal SC, Hall JP, Mattern MR. SUMO fusion technology for difficult-to-express proteins. *Protein Expr Purif.* **2005**, *43*, 1-9.

95. Gunjal VB, Thakare R, Chopra S, Reddy DS. Teixobactin: a paving stone toward a new class of antibiotics? *J Med Chem*. **2020**, *63*, 12171-12195.

96. Guo C, Mandalapu D, Ji X, Gao J, Zhang Q. Chemistry and biology of teixobactin. *Chemistry*. **2018**, *24*, 5406-5422.

97. Mulani MS, Kamble EE, Kumkar SN, Tawre MS, Pardesi KR. Emerging strategies to combat ESKAPE pathogens in the era of antimicrobial resistance: a review. *Front. Microbiol.* **2019**, *10*, 539.

98. Vollmer W, Blanot D, de Pedro MA. Peptidoglycan structure and architecture. *FEMS Microbiol. Rev.* **2008**, *32*, 149-167.

99. Bouhss A, Trunkfield AE, Bugg TD, Mengin-Lecreulx D. The biosynthesis of peptidoglycan lipid-linked intermediates. *FEMS Microbiol Rev.* **2008**, *32*, 208-233.

100. Typas A, Banzhaf M, Gross C, Vollmer W. From the regulation of peptidoglycan synthesis to bacterial growth and morphology. *Nat Rev Microbiol.* **2012***, 10***,** 123-136.

101. Karas JA, Chen F, Schneider-Futschik EK, Kang Z, Hussein M, Swarbrick J, Hoyer D, Giltrap AM, Payne RJ, Li J, Velkov T. Synthesis and structure-activity relationships of teixobactin. *Ann NY Acad Sci.* **2020**, *1459*, 86-105.

102. Giltrap AM, Dowman LJ, Nagalingam G, Ochoa JS, Linington RG, Britton WJ, Payne, RJ. Total synthesis of teixobactin. *Org Lett.* **2016**, *18*, 2778-2791.

103. Jad YE, Acosta GA, Naicker T, Ramtahal M, El-Faham A, Govender T, Kruger HG, de la Torre BG, Albericio F. Synthesis and biological evaluation of a teixobactin analogue. *Org Lett.* **2015**, *17*, 6182-6185.

104. Yang H, Pishenko AV, Li X, Nowick JS. Design, synthesis, and study of lactam and ring-expanded analogues of teixobactin. *J Org Chem*. **2020**, *85*, 1331-1339.

105. Velkov T, Swarbrick JD, Hussein MH, Schneider-Futschik EK, Hoyer D, Li J, Karas JA. The impact of backbone *N*-methylation on the structure-activity relationship of Leu_{10}-teixobactin. *J Pept Sci.* **2019**, *25*, e3206.

106. Shukla R, Medeiros-Silva J, Parmar A, Vermeulen BJA, Das S, Paioni AL, Jekhmane S, Lorent J, Bonvin AMJJ, Baldus M, Lelli M, Veldhuizen EJA, Breukink E, Singh I, Weingarth M. Mode of action of teixobactins in cellular membranes. *Nat Commun.* **2020**, *11*, 2848.

107. VanNieuwenhze MS, Mauldin SC, Zia-Ebrahimi M, Winger BE, Hornback WJ, Saha SL, Aikins JA, Blaszczak LC. The first total synthesis of lipid II: the final monomeric intermediate in bacterial cell wall biosynthesis. *J Am Chem Soc*. **2002**, *124*, 3656-3660.

108. Foster MP, McElroy CA, Amero CD. Solution NMR of large molecules and assemblies. *Biochemistry*. **2007**, *46*, 331-340.

109. Yu JS, Kleckley TS, Wiemer DF. Synthesis of farnesol isomers via a modified Wittig procedure. *Org Lett.* **2005**, *7*, 4803-4806.

110. López SN, Lopes AA, Batista JM Jr, Flausino O Jr, Bolzani Vda S, Kato MJ, Furlan M. Geranylation of benzoic acid derivatives by enzymatic extracts from *Piper crassinervium* (Piperaceae). *Bioresour Technol.* **2010**, *101*, 4251-4260.

111. Schenk DJ, Starks CM, Manna KR, Chappell J, Noel JP, Coates RM. Stereochemistry and deuterium isotope effects associated with the cyclization-rearrangements catalyzed by tobacco epiaristolochene and hyoscyamus premnaspirodiene synthases, and the chimeric CH4 hybrid cyclase. *Arch Biochem Biophys*. **2006**, *448*, 31-44.

112. Li L, Woodward R, Han W, Qu J, Song J, Ma C, Wang PG. Chemoenzymatic synthesis of the bacterial polysaccharide repeating unit undecaprenyl pyrophosphate and its analogs. *Nat Protoc.* **2016**, *11***,** 1280-1298.

113. Duff MR Jr, Grubbs J, Howell EE. Isothermal titration calorimetry for measuring macromolecule-ligand affinity. *J Vis Exp.* **2011**, *55*, e2796.

114. Cochrane SA, Findlay B, Bakhtiary A, Acedo JZ, Rodriguez-Lopez EM, Mercier P, Vederas JC. Antimicrobial lipopeptide tridecaptin A1 selectively binds to Gram-negative lipid II. *Proc Natl Acad Sci USA*. **2016**, *113*, 11561-11566.

115. Hsu ST, Breukink E, Tischenko E, Lutters MA, de Kruijff B, Kaptein R, Bonvin AM, van Nuland NA. The nisin-lipid II complex reveals a pyrophosphate cage that provides a blueprint for novel antibiotics. *Nat Struct Mol Biol.* **2004**, *11*, 963-967.

116. Abdel Monaim SAH, Jad YE, Ramchuran EJ, El-Faham A, Govender T, Kruger HG, de la Torre BG, Albericio F. Lysine scanning of Arg_{10}-teixobactin:

deciphering the role of hydrophobic and hydrophilic residues. *ACS Omega*. **2016**, *1*, 1262-1265.

117. Erger A, Linderstrom-Lang K. Deuterium exchange of poly-DL-alanine in aqueous solution. *Arch Biochem Biophys*. **1957**, *69*, 106-118.

118. Öster C, Walkowiak GP, Hughes DE, Spoering AL, Peoples AJ, Catherwood AC, Tod JA, Lloyd AJ, Herrmann T, Lewis K, Dowson CG, Lewandowski JR. Structural studies suggest aggregation as one of the modes of action for teixobactin. *Chem Sci*. **2018**, *9*, 8850-8859.

119. Yang H, Wierzbicki M, Du Bois DR, Nowick JS. X-ray crystallographic structure of a teixobactin derivative reveals amyloid-like assembly. *J Am Chem Soc.* **2018,** *140,* 14028-14032.

120. Cochrane SA, Vederas JC. Unacylated tridecaptin A_1 acts as an effective sensitiser of Gram-negative bacteria to other antibiotics. *Int J Antimicrob Agents*. **2014**, *44*, 493-499.

121. Vaara M. Polymyxin derivatives that sensitize gram-negative bacteria to other antibiotics. *Molecules*. **2019**, *24*, 249.

122. Vaara M, Viljanen P. Binding of polymyxin B nonapeptide to gram-negative bacteria. *Antimicrob Agents Chemother*. **1985**, *27*, 548-554.

123. Zhang Y, Carney D, Henninot A, Srinivasan K. Novel high-throughput strategy for the aqueous solubility assessment of peptides and proteins exhibiting a propensity for gelation: application to the discovery of novel antibacterial teixobactin analogues. *Mol Pharm*. **2021**, *18*, 469-474.

124. (a) Shoji J, Hinoo H, Sakazaki R, Kato T, Wakisaka Y, Mayama M, Matsuura S, Miwa H. Isolation of tridecaptins A, B and C. (Studies on antibiotics from the genus Bacillus. XXIII). *J. Antibiot (Tokyo)*. **1978**, *31*, 646-651. (b) Kato T, Hinoo H, Shoji J. The structure of tridecaptin A (studies on antibiotics from the genus Bacillus. XXIV). *J Antibiot (Tokyo).* **1978**, *31*, 652-661.

125. Lohans CT, van Belkum MJ, Cochrane SA, Huang Z, Sit CS, McMullen LM, Vederas JC. Biochemical, structural, and genetic characterization of tridecaptin A_1, an antagonist of *Campylobacter jejuni*. *Chembiochem*. **2014**, *15*, 243-249.

126. Schwarzer D, Mootz HD, Linne U, Marahiel MA. Regeneration of misprimed nonribosomal peptide synthetases by type II thioesterases. *Proc Natl Acad Sci USA*. **2002**, *99*, 14083-14088.

127. Cochrane SA, Lohans CT, Brandelli JR, Mulvey G, Armstrong GD, Vederas JC. Synthesis and structure-activity relationship studies of N-terminal analogues of the antimicrobial peptide tridecaptin A_1. *J Med Chem*. **2014**, *57*, 1127-1131.

128. Cochrane SA, Findlay B, Vederas JC, Ratemi ES. Key residues in octyl-tridecaptin A1 analogues linked to stable secondary structures in the membrane. *ChemBioChem*. **2014**, *15*, 1295-1299.

129. Ito N, Watahiki T, Maesawa T, Maegawa T, Sajiki H. Synergistic effect of a palladium-on-carbon/platinum-on-carbon mixed catalyst in hydrogen/deuterium exchange reactions of alkyl-substituted aromatic compounds. *Adv Synth.* **2006**, *348*, 1025-1028.

130. Maksimov MO, Pan SJ, Link JA. Lasso peptides: structure, function, biosynthesis, and engineering. *Nat Prod Rep*. **2012**, *29*, 996-1006.

131. Hegemann JD, Zimmermann M, Xie X, Marahiel MA. Lasso peptides: an intriguing class of bacterial natural products. *Acc Chem Res*. **2015**, *48*, 1909-1919.

132. Salomón RA, Farías RN. Microcin 25, a novel antimicrobial peptide produced by *Escherichia coli*. *J Bacteriol.* **1992**, *174*, 7428-7435.

133. Delgado MA, Rintoul MR, Farías RN, Salomón RA. *Escherichia coli* RNA polymerase is the target of the cyclopeptide antibiotic microcin J25. *J Bacteriol*. **2001**, *183*, 4543-4550.

134. Rintoul MR, de Arcuri BF, Salomón RA, Farías RN, Morero RD. The antibacterial action of microcin J25: evidence for disruption of cytoplasmic membrane energization in *Salmonella newport*. *FEMS Microbiol Lett*. **2001**, *204*, 265-270.

135. Bellomio A, Vincent PA, de Arcuri BF, Farías RN, Morero RD. Microcin J25 has dual and independent mechanisms of action in *Escherichia coli*: RNA polymerase inhibition and increased superoxide production. *J Bacteriol*. **2007**, *189*, 4180-4186.

136. Blond A, Péduzzi J, Goulard C, Chiuchiolo MJ, Barthélémy M, Prigent Y, Salomón RA, Farías RN, Moreno F, Rebuffat S. The cyclic structure of microcin J25, a 21-residue peptide antibiotic from *Escherichia coli*. *Eur J Biochem.* **1999**, *259*, 747-755.

137. Rosengren KJ, Clark RJ, Daly NL, Göransson U, Jones A, Craik DJ. Microcin J25 has a threaded sidechain-to-backbone ring structure and not a head-to-tail cyclized backbone. *J Am Chem Soc*. **2003**, *125*, 12464-12474.

138. Solbiati JO, Ciaccio M, Farías RN, González-Pastor JE, Moreno F, Salomón RA. Sequence analysis of the four plasmid genes required to produce the circular peptide antibiotic microcin J25. *J Bacteriol*. **1999**, *181*, 2659-2662.

139. Duquesne S, Destoumieux-Garzón D, Zirah S, Goulard C, Peduzzi J, Rebuffat S. Two enzymes catalyze the maturation of a lasso peptide in *Escherichia coli*. *Chem Biol*. **2007**, *14*, 793-803.

140. Tietz JI, Schwalen CJ, Patel PS, Maxson T, Blair PM, Tai HC, Zakai UI, Mitchell DA. A new genome-mining tool redefines the lasso peptide biosynthetic landscape. *Nat Chem Biol*. **2017**, *13*, 470-478.

141. Cheung-Lee WL, Link AJ. Genome mining for lasso peptides: past, present, and future. *J Ind Microbiol Biotechnol.* **2019**, *46*, 1371-1379.

142. Pavlova O, Mukhopadhyay J, Sineva E, Ebright RH, Severinov K. Systematic structure-activity analysis of microcin J25. *J Biol Chem*. **2008**, *283*, 25589-25595.

143. Knappe TA, Manzenrieder F, Mas-Moruno C, Linne U, Sasse F, Kessler H, Xie X, Marahiel MA. Introducing lasso peptides as molecular scaffolds for drug design: engineering of an integrin antagonist. *Angew Chem Int Ed*. **2011**, *50*, 8714-8717.

144. Zong C, Maksimov MO, Link AJ. Construction of lasso peptide fusion proteins. *ACS Chem Biol*. **2016**, *11*, 61-68.

145. Cheng C, Hua Z-C. Lasso peptides: heterologous production and potential medical application. *Front Bioeng Biotechnol*. **2020**, *8*, 571165.

146. Lear S, Munshi T, Hudson AS, Hatton C, Clardy J, Mosely JA, Bull TJ, Sit CS, Cobb SL. Total chemical synthesis of lassomycin and lassomycin-amide. *Org Biomol Chem*. **2016**, *14*, 4534-4541.

147. Rosengren KJ, Blond A, Afonso C, Tabet JC, Rebuffat S, Craik DJ. Structure of thermolysin cleaved microcin J25: extreme stability of a two-chain antimicrobial peptide devoid of covalent links. *Biochemistry*. **2004**, *43*, 4696-4702.

148. Yan LZ, Dawson PE. Synthesis of peptides and proteins without cysteine residues by native chemical ligation combined with desulfurization. *J Am Chem Soc*. **2001**, *123*, 526-533.

149. Lide DR. CRC Handbook of Chemistry and Physics, *82nd ed*. CRC Pres: Boca Raton, FL. **2001**.

150. Malins L, Giltrap AM, Dowman LJ, Payne RJ. Synthesis of β-thiol phenylalanine for applications in one-pot ligation-desulfurization chemistry. *Org Lett*. **2015**, *17*, 2070-2073.

151. Isidro-Llobet A, Álvarez M, Albericio F. Amino acid-protecting groups. *Chem Rev*. **2009**, *109*, 2455-2504.

152. Yang S-C, Wang H-M, Chen L-C. A mild and efficient method for selective acetylation of amines. *J Chin Chem Soc*. **1995**, *42*, 585-587.

153. Zimmermann M, Hegemann JD, Xie X, Marahiel MA. Characterization of caulonodin lasso peptides revealed unprecedented N-terminal residues and a precursor motif essential for peptide maturation. *Chem Sci*. **2014**, *5*, 4032-4043.

154. Perrin DDA, Amarego WLF. Purification of laboratory chemicals, *3rd ed*. Pergamon Press: New York City, NY. **1993**.

155. Hwang TL, Shaka AJ. Water suppression that works. Excitation sculpting using arbitrary wave-forms and pulsed-field gradients. *J Magn Reson A*. **1995**, *112*, 275-279.

156. Methods for dilution antimicrobial susceptibility tests for bacteria that grow aerobically. Clinical and Laboratory Standards Institute: Wayne, PA. **2011**.

157. Gasteiger E, Hoogland C, Gattiker A, Duvaud S, Wilkins MR, Appel RD, Bairoch A. Protein identification and analysis tools on the ExPASy server. In The Proteomics Protocols Handbook, *Walker JM Ed*. Humana Press: New York City, NY. **2005**, 571-607.

158. DeLano WL. The PyMol Molecular Graphics System. DeLano Scientific. San Carlos, CA. **2002**.

159. Johnson BA. Using NMRView to visualize and analyze the NMR spectra of macromolecules. *Methods Mol Biol*. **2004**, *278*, 313-352.

160. Delaglio F, Grzesiek S, Vuister GW, Zhu G, Pfeifer J, Bax A. NMRPipe: a multidimensional spectral processing system based on Unix pipes. *J Biomol NMR*. **1995**, *6*, 277-293.

161. (a) Rodger A. Far UV protein circular dichroism. Encyclopedia of Biophysics, *Roberts GCK Ed*. Springer: Heidelberg, Berlin. **2013**. (b) Nordén B, Rodger A, Dafforn TR. Linear dichroism and circular dichroism: a textbook on polarized spectroscopy. Royal Society of Chemistry: Cambridge, UK. **2010**.

162. Dai Y, Whittal RM, Li L. Two-layer sample preparation: a method for MALDI-MS analysis of complex peptide and protein mixtures. *Anal Chem*. **1999**, *71*, 1087-1091.

163. Ma B, Zhang K, Hendrie C, Liang C, Li M, Doherty-Kirby A, Lajoie G. PEAKS: powerful software for peptide de novo sequencing by tandem mass spectrometry. *Rapid Commun Mass Spectrom*. **2003**, *17*, 2337-2342.

164. Kaitlyn, MT. Structural studies of peptides that influence the pathogenicity of bacterial infections, and investigation of structure-activity-relationships of antimicrobial peptides with application to cancer therapy. *University of Alberta Theses and Dissertations*. **2017**.

165. Wuthrich K. NMR of proteins and nucleic acids, *1st ed*. John Wiley & Sons: New York City, NY. **1986**.

166. van de Guchte M, van der Vossen JM, Kok J, Venema G. Construction of a lactococcal expression vector: expression of hen egg white lysozyme in *Lactococcus lactis* subsp. *lactis*. *Appl Environ Microbiol*. **1989**, *55*, 224-248.

Appendix: Grassroots initiatives to foster equity, diversity, and inclusivity

*This work was previously published in the Canadian Journal of Chemistry special issue In Honour of Dr. Margaret-Ann Armour as "*Chiorean S, Oakley MS, Sinclair J. Starting grassroots initiatives to foster equity, diversity, and inclusivity in the Chemistry Department at the University of Alberta. *Can J Chem*. **2021**, *99*, 679-684." *and a version is reproduced here with permission.*

A.1 Motivations

Less than four decades ago, there were zero women faculty and about 70% of the student body was male in the Department of Chemistry at the University of Alberta.[A1] After Dr. Margaret-Ann Armour joined the Department, she spearheaded the formation of the Women in Scholarship, Engineering, Science, and Technology[A2] (WISEST) organization in 1982 to increase the number of female students in the Faculty of Science. Currently, WISEST is focused on increasing engagement of high school students from underrepresented groups to consider careers in science by hosting summer research opportunities. With WISEST focused on introducing young women to science, Women in Science, Engineering & Research[A3] (WiSER) was designed to promote the retention and advancement of women in the STEM workforce. As Associate Dean of Diversity in the Faculty of Science at the University of Alberta, Dr. Armour directed Project Catalyst, a committee formed in 2005 aimed at understanding hiring biases towards female candidates (family life assumptions, technical skill assessments) and overcoming the structural barriers (maternity leave,

timeline to tenure, lack of mentorship) in order to increase the appointment and retainment of women in the Faculty of Science. Even after 11 years of work, the overall percentage of female faculty has only increased to a total of 15% at the University of Alberta.[A4]

Although steps have been taken to rectify gender disparity in chemistry, such as the initiatives outlined above, only 28% of assistant professors in chemistry departments throughout Canada are women, and significantly less (10%) are full professors.[A5] These numbers are also reflected at the University of Alberta, where faculty and staff are disproportionate in their representation compared to the student body. The lack of mentorship and representation levy a heavy burden on graduate students, who may be first generation academics, far from their support networks, or struggle to find their place in the chemistry community for other reasons.[A6] This is reflected in the disproportionate attrition of underrepresented individuals at the juncture between graduate school and post-graduate careers in science.[A7,8]

Many of the equity-focused groups at the University of Alberta target the recruitment of high school students into the sciences (WISEST), or those who are entering or are already in the workforce as career scientists (WiSER, Project Catalyst). While these projects focus on STEM-wide issues, a group serving a small sub-section of STEM may be able to enact changes much more quickly at the Department level. In the spring of 2017, the University of Alberta's Women in Chemistry initiative was founded to bridge this gap and provide support to underrepresented groups in the trainee stages of their careers in chemistry.

The drive behind the foundation of this group was varied, with each graduate student bringing different perspectives based on their personal past experiences. Many members recounted instances of discrimination, implicit bias, microaggressions, gaslighting, imposter syndrome, and sexual harassment. These experiences fuelled our desire to create a more inclusive environment where diversity is celebrated and equity is fought for. The group, led by a graduate student executive team with support from a Faculty Advisor and occasional undergraduate student volunteers, has a vision of engaging the Department of Chemistry in a cultural shift towards the active retention of underrepresented groups in chemistry. Since its initiation, the group has grown, both in terms of membership and in addressing the needs of our peers. Recognizing that this group should serve all persons in the community, and after a series of discussions and departmental feedback, the name was officially changed to University of Alberta Working for Inclusivity in Chemistry (UAWIC) to better reflect the true nature of our membership. The UAWIC group set out to build an equitable, diverse, and inclusive environment through the following goals:

1. To build community among members of the Chemistry Department by breaking down the barriers between hierarchical academic roles, allowing for a greater understanding amongst groups (faculty, students, staff) to develop. The sharing of lived experiences can allow for open discussion and will help Department members gain insight into other perspectives in chemistry.

2. To retain the diverse graduate student population in chemistry by giving graduate students tools and training to enrich their career development, with a specific focus on highlighting diverse scientists and career paths. We hope to

provide the space for individuals to be supported in their pursuit of science at the graduate level and to place the inclusion of historically excluded groups at the forefront.

The events and initiatives described below were organized to execute our vision of creating a more inclusive environment at the University of Alberta. They are presented here to act as a template for further action and support for equity, diversity, and inclusion (EDI) initiatives.

A.2 Goal 1: Community building

UAWIC set out to achieve our first goal by organizing department-wide community building through promoting visibility of underrepresented groups and allies, and hosting social events, but also facilitating peer-to-peer community building by networking at meet & greets with departmental invited speakers.

Department-wide community building: visibility and food

Most of the social events organized by UAWIC utilize food as a way to bring people together. World Food Day invites all members of the Department to contribute food to a potluck-style lunch and gives people the opportunity to share some of their culture. This event has been particularly successful in accommodating people with food restrictions, as participants are asked to indicate the name and ingredients of their dish. Pi Day is another potluck event where Department members contribute

both sweet and savoury pies for others to enjoy. The event spans the entire Department as Chairpersons, administrative staff, professors and others volunteer to be pie'd in the face. The event helps break down the hierarchy of the Department, as graduate students pie Faculty, service staff pie administrators, and undergraduates pie their lab teaching assistants. These events facilitate community building and camaraderie through shared experience and good food, which in turn supports our first goal.

Visible allyship

As the scope of our group expanded, we wanted to outwardly show our support for different intersectional identities within both EDI and STEM. Visibility of allyship is one major area that has been identified to help make workspaces in the physical sciences more welcoming to underrepresented groups.[A9] Flying a rainbow flag or posting Black Lives Matter signs, for example, can promote more inclusive spaces and increase visibility of allies. Additionally, these symbols can generate conversation and can help one locate safe spaces to talk about issues they may be facing. In an attempt to increase visibility of allyship at conferences, we designed a set of enamel pins that people can wear to signify that they are potential allies. Current pins advocate for women in chemistry, diversity, and support for those who have invisible identities. These pins, seen in Figure A.1, also serve as ongoing fundraising for our group and have been sold internationally for chemists to identify themselves as allies at work and conferences.

Figure A.1 Enamel pins designed by UAWIC to promote visible allyship

In an effort to engage the entire Department in visibility initiatives, we participated in the international LGBTQ+ in STEM day[A10] through a door decorating contest. Research groups in the Chemistry Department volunteered to decorate their door and UAWIC hosted a tour to view all the participating doors. Approximately 30% of the research groups in the department participated, and many of the decorations were left up throughout the year to indicate allies to the LGBTQ+ community.

The LGBTQ+ in STEM event and the annual IUPAC Global Women's Breakfast[A11] are two events which, though hosted internationally, we were able to participate in locally. These events bring together communities around the world to learn from each other. The global recognition of groups working towards diversity and inclusion is important for the ongoing success of groups like UAWIC, as they allow the sharing of event ideas and prevent burnout via mutual support. These events act to highlight diverse chemists internationally, as well as within our own Department, and aid in creating an inclusive environment which will have a lasting effect on the community, supporting the ongoing goals of UAWIC.

Peer-to-peer community building: Meet & Greets

As a fledgling group, UAWIC sought to highlight diverse chemists who visited the Department to give research seminars. UAWIC offers to host visiting speakers who identify as members of underrepresented groups in chemistry or who follow non-traditional careers in a casual coffee hour open to the entire Department. Being an informal activity, these meetings give students and postdoctoral fellows time to ask questions about careers, barriers to advancement in chemistry, and work-life balance in a safe setting. Given the gender disparity and profound lack of diversity in faculty membership across Canada,[A5] these events allow access to a larger pool of representative academics and professionals. The mentorship by these visiting speakers, however brief, provided invaluable insight into work environments and challenges upon the completion of graduate school. This event series regularly draws groups of 10-15 attendees. Although this represents only 5% of the graduate student body, these attendees hail from different groups with different interests, depending on the speaker. For example, new female professors tended to draw postdoctoral researchers and graduate students interested in pursuing academia, while pharmacological industry representatives drew an entirely different cohort of student attendees. Together, the reach for these events would be approximately 25% of the total graduate student population. The Meet & Greet hours function within the framework of both of our goals by facilitating conversation and connections among individuals with similar interests as well as highlighting diverse scientists in the wider chemistry community.

A.3 Goal 2: Retention of diverse graduate student chemists

Multiple factors can affect the desire of a student to continue in academia and pursue advanced careers in chemistry, and our group strives to give graduate students tools to navigate their chosen career path. We provide professional development tools through workshops (LOGIC, highlighting diverse career paths) and also train students in potential barriers they or their colleagues may face along the way (LOGIC, Margaret-Ann Armour Lecture Series, Diversity in STEMinars).

Leaders Overcoming Gender Inequality in Chemistry Retreat

In 2018, UAWIC organized the 2nd annual Leaders Overcoming Gender Inequality in Chemistry (LOGIC) retreat in Edmonton, AB. This annual event began in 2017 as an initiative hosted by the graduate student members of the Women in Chemistry (now known as Working towards Inclusivity in Chemistry) group at the University of Toronto. The inaugural event was inspired by the Puget Sound Women Chemists retreat and was held as a satellite event to the Canadian Chemistry Conference and Exhibition (CCCE) to reach a wider audience of Canadian chemists. LOGIC includes seminars, workshops, panel discussions, and networking sessions focused on a central theme. In past years, themes have included "Becoming a Confident Chemist and Future Leader" (Toronto, 2017) and "Paving a Path to a Career in Chemistry" (Edmonton, 2018). These retreats focused on building professional development skills for graduate students and early career researchers and providing resources for identifying mentors and sponsors for your career path. In 2019 and 2020, the themes

were “Women Leading and Excelling” (Quebec, 2019) and “Beyond the Visible Spectrum” (Winnipeg/Virtual, 2020) where the focus was predominantly on educating and training individuals on equity, diversity and inclusivity. This retreat is based on learning uncomfortable topics in an environment which encourages questions and meaningful discussion. Mentorships often begin within LOGIC, as some attendees are new to these topics and how they relate to STEM, while others are highly educated in EDI issues and can help guide beginners.

The breakdown of attendee statistics, collected by the respective organizing committee of that year, over the last four years can be seen in Figure A.2. Graduate students make up the majority of the attendee roster with an average of 53%. The remaining half is distributed among postdoctoral researchers, academics (assistant professors, associate professors, full professors, technical staff), industry, and government chemists. Since the first LOGIC retreat, the number of participants has increased, with a major increase in 2020 due to the retreat being moved online and the alleviation of registration fees. Additionally, the breakdown of attendee gender for LOGIC shows that there are a much greater number of female attendees than any other gender. This may be due to the original scope of the conference (women in chemistry) and we are working to incorporate components in the future that will bring people of all genders to the discussion.

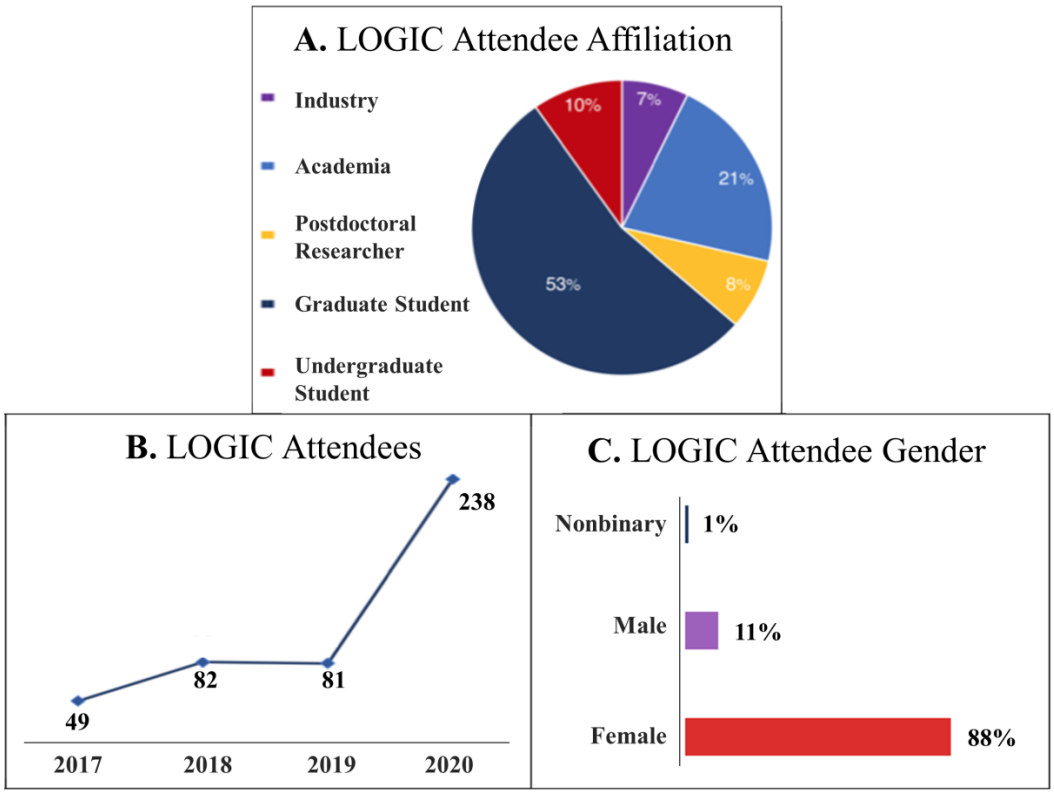


Figure A.2: Attendee statistics based on data over four years of the LOGIC retreats: (A) attendee affiliation averaged over data from 2017 to 2020 (B) growth in annual number of attendees (C) self-reported attendee gender percentages averaged over data from 2017 to 2020.

Anonymous feedback from previous LOGIC retreats provides insight to the most valuable aspects of the conference as identified by attendees. Participants explained that they left the conference feeling invigorated and inspired to tackle EDI challenges:

"It was important and empowering to learn about what issues women in science face when heading into the workforce as PIs or in other chemistry related fields, and I feel more prepared heading into the next stage in my career being aware of these issues. I will be proactive in fighting for equity in my place of work, and have been inspired by

so many amazing women who have already paved their way in their respective fields." – participant 1

"Pursue what you love to do no matter what barriers or challenges you encounter, and don't be afraid to seek out mentors and colleagues in the chemistry community for support." – participant 2

Feedback also showed that attendees learned about aspects of intersectionality and how to better support underrepresented groups in STEM:

"Being a white woman is not ideal in science, but it's a helluva lot more privileged than being a woman of colour, and I need to recognize that advantage as such, and make sure that I work to increase diversity of all its forms, and not just women. That is important." – participant 3

"It gave me tools and language to express myself and help me become a more effective and active ally." – participant 4

Attendees have reported that bringing these discussions to their own department would be valuable and would like to set up avenues for dissemination of EDI topics:

"You've inspired me to encourage students in my department to come together and discuss equity, diversity and inclusion since as it stands, we do not have any such organization or platform for these conversations." – participant 5

"I think all Department Heads should be attending as it really opens one's eyes to the challenges faced by so many." – participant 6

LOGIC 2020 was adjusted to a virtual format due to COVID-19 restrictions. The online platform boosted attendee numbers three-fold from the previous year and, since travel and accommodations were no longer required, there were a large number of international registrants from countries like USA, Australia, England, Belgium, Scotland, and Switzerland. Additionally, this allowed us to offer more inclusive and flexible programming such as pre-recorded and close-captioned seminars. Discussions on the Slack platform allowed attendees to be part of the conference without excluding those with career or caretaking commitments or language barriers.

Over time, LOGIC has evolved into a community-driven event organized by an overarching network, Canadians Working for Inclusivity in Chemical Sciences, Engineering, and Technology[A12] (CWIC). Representatives from other inclusivity groups across Canada are now involved and a community is built through working towards a common goal.

Margaret-Ann Armour Lecture Series

In honour of Dr. Margaret-Ann Armour's contribution to scientific research and advancement of diversity at the University of Alberta, UAWIC established the annual student-invited Margaret-Ann Armour Lecture Series in 2019.[A13] This lecture series showcases that chemists can be both outstanding in their research and promote diversity and inclusion within chemistry.

Potential speakers for this series are short-listed by the graduate student members of the UAWIC executive and an anonymous voting ballot is electronically sent to all graduate students and postdoctoral fellows in the Department for final selection. The UAWIC Co-Chairs will reach out to the chosen speaker and invite them to visit our campus for a two-part lecture series. One of the lectures will be a research presentation and the second lecture will be dedicated to showcasing the career path and personal involvement of our visitor in EDI initiatives. Once they have accepted, members of the UAWIC executive team will begin planning and executing all aspects of the speaker's visit, from flights to accommodations to meals to campus room and projector booking to advertising. To ensure both students and faculty have a chance to meet our speaker, we arrange for a student pizza lunch open to all graduate students and postdoctoral fellows, a dinner with the entire UAWIC executive team, as well as individual one-on-one faculty meetings between the visitor and professors of our Chemistry Department. An endeavour of this type does require significant planning and we begin this process, initiated by the student vote, about a year in advance of the anticipated visit. Furthermore, to successfully deliver this lecture series as a free event to all members of our university community interested in attending, we actively fundraise through the year by hosting food-centered events, promoting pin sales, and writing grants for additional support.

Our inaugural speaker was Dr. Geraldine Richmond, Professor of Chemistry at the University of Oregon and founder of COACh (Committee On the Advancement of Women Chemists), an organization focusing on increasing the presence of women in science by providing resources on building successful careers.[A14-16] Dr. Richmond

spoke about her research on nanoemulsions, as well as highlighting the importance of groups such as COACh. Our invited speaker for 2020 was Dr. Molly Shoichet, Professor of Chemistry at the University of Toronto, who spoke to us about her work on developing hydrogels and her career path. In light of COVID-19, our 2021 Armour Lecture Series took place virtually over a two-day period. Our invited speaker was Dr. Jennifer Heemstra, Professor of Chemistry at Emory University and a columnist for Chemical & Engineering News, targeting early career chemists with monthly advice and guidance. Dr. Heemstra shared her work on nucleic acid molecular recognition and assembly, her struggles with imposter syndrome throughout a career in academia, and engaged in lively discussions with our executive team, students, and professors.

Department members were exposed to strong role models that excel both in research and embracing diversity. Meeting leaders that exemplify our goals is a valuable network building opportunity that has been well received by the Department (65 attendees or more). This seminar series highlights EDI work as crucial to the success of science, rather than a pursuit external to it. The speakers are lifelong champions for inclusivity in their own communities and provide valuable insight into what still needs to be done in this area.

Diversity in STEMinars

The Diversity in STEMinar lecture series was founded in 2019 to bring more diverse perspectives into the Chemistry Department and foster discussion on topics outside of

the expertise of the UAWIC executive committee. We focus on identifying local speakers with research interests under the STEM umbrella, but who also make diversity and inclusivity a focus in their career. Our first speaker, Dr. Kim TallBear, addressed the colonial past and present of science, in particular in genomic and biological studies.[A17,18] For our second STEMinar, we leveraged a mandatory professional development program within the Faculty of Graduate Studies and Research at the University of Alberta to increase attendance and broaden the pool of attendees to the entire Faculty of Science. Reaching over 80 students, faculty, and staff, Dr. Lisa Willis discussed the statistics of diverse representation in STEM academia,[A19] the fallacy of improvement in this area,[A20] and the benefits to problem solving and creative idea production that result from assembling diverse teams.[A21-23] These seminars will continue to operate as a way to educate the community on the intersections of diversity and scientific research, and highlight local scientists who are experts and passionate about this work. This broad education on the status of EDI movements in STEM supports both of our goals by fostering discussion within the Department and highlighting areas in which we can improve support for underrepresented groups during their time at the University of Alberta.

Highlighting diverse career paths

Chemistry-degree holders pursue a vast range of career trajectories after graduation[A24] and exposure to non-academic jobs is extremely limited in university settings. UAWIC regularly uses its network of professional chemists to showcase possible career opportunities available through panel discussions. Speakers include

individuals from all walks of life, ranging from government scientists, to process chemists, to scientific editors. Attendees engaged in discussions about navigating a career in chemistry and were exposed to different employment opportunities available to them. These events are specifically tailored to support our second goal. By including a diverse range of speakers, open discussions, and sharing experiences, we equip attendees with in-depth knowledge of what to expect when entering the chemistry workforce.

Canadians Working for Inclusivity in Chemical Sciences, Engineering, and Technology Network

Through many of these events, we met other graduate students and groups working on EDI issues in chemistry. In collaboration with the University of Toronto Working for Inclusivity in Chemistry[A25] and the University of Saskatchewan Women in Chemistry[A26] groups, the Canadian Women in Chemical Sciences, Engineering, and Technology (CWIC) Network was formed. This team is focused on providing guidance and resources to up and coming CWIC chapters across Canada, as well as creating a network for these chapters to share resources and experiences with each other. Since 2018, the number of affiliated groups has grown from three to fifteen. The CWIC Network is affiliated with the Chemical Institute of Canada (CIC) as a member resource group and is the first of its kind in the Canadian chemistry community. CWIC shares similar goals to the UAWIC group, but has a far greater reach into the academic and industrial spheres. For example, CWIC will organize an EDI symposium at the 104th Canadian Chemistry Conference and Exhibition / 48th

World Chemistry Congress / 51st IUPAC General Assembly focused on allyship in the chemistry community.

A.4 Conclusions and future perspectives

As EDI initiatives grow in size and number in the international community, UAWIC has worked to be at the forefront of the movement within graduate chemistry communities in Canada. Due to the mentorship and sponsorship of Dr. Margaret-Ann Armour during the inception of the group, we have been able to make noticeable changes within our Department. Professional development activities such as career panels, the LOGIC conferences, and seminars (Margaret-Ann Armour Lecture Series, Diversity in STEMinars) provide attendees with the tools to navigate academic and industrial systems and identify inequities so they can be addressed. Visibility through social events, door decorating initiatives, and lapel pins have helped to recognize allies in our community and help those of all backgrounds and identities feel welcome. Looking forward, we aim to continue advocacy efforts within the University of Alberta. The Department of Chemistry has recently formed an Equity, Diversity, and Inclusion Committee, on which a graduate student representative from UAWIC is present. The prerogatives of the Departmental EDI Committee[A27] are heavily based on the EDI components of NSERC funding applications, upon which UAWIC was previously consulted. Involvement in these higher-level Departmental Committees allows for direct input on policy development, resource allocation, and other crucial, concrete changes to the social and professional environment at the

University. Connections within the CWIC Network will also enable a continued push for the re-examination of systemic bias and inequitable practices in the wider Canadian chemistry community. The LOGIC conferences, as well as the hosting of EDI sessions within the CCCE meetings, will enable the ongoing support and growth of these initiatives, and continue to promote diversity and inclusion in the Canadian chemistry community through visibility and education. However, as UAWIC is a local group serving the community of underrepresented minorities in chemistry at the University of Alberta, our primary focus will continue to be the development of a positive and inclusive environment which strives for equity, diversity, and inclusivity. We hope Dr. Armour would be proud.

The aims of UAWIC will continue to evolve as new individuals become part of the executive team and share their stories and perspectives. UAWIC has flourished due to its varied membership, representing diverse identities and experiences of those who wish to join us. It is vital to acknowledge that, although we are working towards the same goals of inclusion and diversity, the ways in which these initiatives manifest in each department and in each group may be different.

As a last thought, we encourage *everyone* to take a stand and be part of the conversation around equity, diversity, and inclusivity in STEM, but *especially* those who are not part of underrepresented or marginalized groups, as they tend to disproportionately carry most diversity initiatives.[A28] Take responsibility to educate yourself and others about systemic racism and biases that are ingrained in academia and Canadian culture, and help to improve the quality of life and safety of others around you who may be less privileged.

A.5 Acknowledgements

The authors would like to acknowledge members of Canadian chemistry community for their continued support as well as the generous funding sources that enabled us to make our initiatives possible. We would like to extend a heartfelt thank you to our incredible executive team members for dedicating their time, effort, and ideas throughout the last few years.

Thank you especially to Dr. Florence Williams for setting this in motion by laying the groundwork, providing valuable guidance for our foundation, and offering unwavering support in structuring of UAWIC's initial activities as our first Faculty Advisor. A sincere thank you to our fellow co-founders, Dr. Regina Sinelnikov and Dr. Rochelin Dalangin, for diving into this adventure with us and helping make LOGIC 2018 a reality. Since then, UAWIC has grown to include Karen Hoy, Michelle Ha, Hansol Park, Kleinberg Fernandez, William Medroa Del Pino, Brittany Reib, Sarah Milliken, Caley Craven, Iris Chan, Alexandra Burnett, Laura Matchett, Kay McCallum, Santiago Tíjaro Bulla, and Keija Yan, who have all helped shape UAWIC into what it is today. Thank you to the University of Alberta Chemistry Department, with particular thanks to Dr. Julianne Gibbs for serving as our current Faculty Advisor, Dr. Jillian Buriak for moral support, and of course Dr. Margaret-Ann Armour for supporting us since our very first event. We are also grateful to the other CWIC founders for sharing this journey with us, in particular Dr. Anika Tarasewicz (who reviewed this manuscript), Dr. Elaheh Khozeimeh, Dr. Nimrat Obhi, and Dr. Bryony McAllister. We attribute the success of UAWIC to all of you!

A.6 References

A1. Armour MA. Academic chemistry disciplines through a gender diversity lens. *Making Chemistry Inclusive*. Macmillan Learning Curriculum Solutions. **2016**, 5-18.

A2. Women in Scholarship, Engineering, Science, and Technology. ualberta.ca/services/wisest/index.html (accessed Nov 29, 2020).

A3. WiSER. wiseredmonton.ca (accessed Nov 29, 2020).

A4. Faculty of Science Diversity Report. https://www.ualberta.ca/science/about-us/contours/2016-fall-contours/2016/october/faculty-of-science-diversity-report.html (accessed Nov 29, 2020).

A5. Canadian Association of University Teachers. CAUT Almanac of Post-Secondary Education in Canada. *Canadian Association of University Teachers*, **2014**. https://www.caut.ca/docs/default-source/almanac/almanac_2013-2014_print_finalE20A5E5CA0EA6529968D1CAF.pdf?sfvrsn=2 (accessed July 23, 2020).

A6. Ferreira MM. Gender issues related to graduate student attrition in two science departments. *Int J Sci Educ*. **2003**, *25*, 969-989.

A7. National Science Foundation. Women, Minorities, and Persons with Disabilities in Science and Engineering. **2019**. https://ncses.nsf.gov/pubs/nsf19304/digest/ (accessed Nov 13, 2020).

A8. Royal Society of Chemistry. Diversity Landscape of the Chemical Sciences. **2018**. https://www.rsc.org/globalassets/02-about-us/our-strategy/inclusion-

diversity/cm-044-17_a4-diversity-landscape-of-the-chemical-sciences-report_web-2.pdf (accessed Nov 13, 2020).

A9. Royal Society of Chemistry. Exploring the workplace for LGBTQ+ physical scientists. **2019**. https://www.rsc.org/globalassets/04-campaigning-outreach/campaigning/lgbt-report/lgbt-report_web.pdf (accessed July 15, 2020).

A10. Pride in STEM. LGBTQ+ In Stem Day 2019. https://prideinstem.org/lgbtstem-day-2019-press-release (accessed July 15, 2020)

A11. International Union of Pure and Applied Chemistry. IUPAC Global Women's Breakfast. https://iupac.org/100/global-breakfast/ (accessed Nov 29, 2020).

A12. Canadians Working for Inclusivity in Chemical Sciences, Engineering, and Technology. www.cwicnetwork.ca (accessed Nov 29, 2020).

A13. Univewrsity of Alberta Chemistry Department. Chemistry Lecture Series. https://www.ualberta.ca/chemistry/news-and-events/ lecture-series/index.html (accessed Nov 29, 2020).

A14. Stockard J, Greene J, Richmond G, Lewis P. Is the gender climate in chemistry still chilly? Changes in the last decade and the long-term impact of COACh-sponsored workshops. *J Chem Educ*. **2018**, *95*, 1492-1499.

A15. Stockard J, Greene J, Lewis P, Richmond G. Promoting gender equity in academic departments: a study of department heads in top-ranked chemistry departments. *J Women Minor Scien Eng.* **2008**, *14*, 1-27.

A16. Stockard J, Greene J, Lewis P, Richmond G. Promoting mentoring among and for women in chemistry: The experiences of COACh. In mentoring strategies to

facilitate the advancement of women faculty. *American Chemical Society.* **2010**, 153-163.

A17. TallBear, K. Native American DNA - tribal belonging and the false promise of genetic science. University of Minnesota Press: Minneapolis, **2013**.

A18. Tallbear, K. Caretaking Relations, Not American Dreaming. *Kalfous.* **2019**, 6, 24-41.

A19. Natural Science and Engineering Research Council. Women in Science and Engineering in Canada, **2017**. https://www.nserc-crsng.gc.ca/_doc/Reports-Rapports/WISE2017_e.pdf (accessed July 15, 2020)

A20. Eaton AA, Saunders JF, Jacobson RK, West K. How gender and race stereotypes impact the advancement of scholars in STEM: Professors' biased evaluations of physics and biology post-doctoral Ccandidates. *Sex Roles.* **2020**, *82*, 127-141.

A21. Woolley AW, Chabris CF, Pentland A, Hashmi N, Malone TW. Evidence for a collective intelligence factor in the performance of human groups. *Science.* **2010**, *330*, 686-688.

A22. Freeman RB, Huang W. Collaboration: strength in diversity. *Nature.* **2014**, *513*, 305.

A23. Levine SS, Apfelbaum EP, Bernard M, Bartelt VL, Zajac EJ, Stark D. Ethnic diversity deflates price bubbles. *Proc Natl Acad Sci USA.* **2014**, *111*, 18524-18529.

A24. American Chemical Society. ACS Salary Report 2015. https://acs.org/content/acs/en/careers/salaries/surveys/salary-report-2015.html (accessed July 15, 2020).

A25. Working towards Inclusivity in Chemistry Toronto. http://croftw.chem.utoronto.ca/wp/womeninchemto/ (accessed Nov 29, 2020).

A26. University of Saskatchewan Women in Chemistry. https://usask-wic.com/ (accessed Nov 29, 2020).

A27. University of Alberta Department of Chemistry EDI Committee. Our Commitment to Equity, Diversity, and Inclusion. **2020**. https://www.ualberta.ca/chemistry/news-and-events/news/2020/august/edi_post.html (accessed Nov 9, 2020).

A28. Jimenez MF, Laverty TM, Bombaci SP, Wilkins K, Bennett DE, Pejchar L. Underrepresented faculty play a disproportionate role in advancing diversity and inclusion. *Nat Ecol Evol.* **2019**, *3*, 1030-1033.