Identification and Characterisation of Novel Bacteriocins Active Against *Clostridioides difficile*

Sarah May Olivia Chapman, BSc, MRes

Thesis Submitted to the University of Nottingham for the Degree of Doctor of Philosophy

October 2019

Declaration

Except where otherwise acknowledged the work presented in this thesis is my own. No part has been submitted for another degree at The University of Nottingham or any other institute of learning.

Sarah May Olivia Chapman

October 2019

Abstract

Clostridioides difficile is a Gram-positive, spore forming, anaerobic bacilli which is a leading causative agent of antibiotic associated diarrhoea and represents a significant burden to medical systems worldwide. The current treatment for *C. difficile* is antibiotics, including vancomycin, which are also a major risk factor for the disease. Due to this and the growing concern of the increasing prevalence of organisms with antibiotic resistance, including vancomycin resistance, a need for novel treatments has been repeatedly expressed. Bacteriocins are a group of antimicrobials undergoing increasingly extensive research for their applications as antimicrobials in the medical industry. In this study, putative bacteriocin harbouring organisms demonstrating antimicrobial activity towards clinically relevant strains of *C. difficile* were isolated from cow manure, a suggested reservoir of *C. difficile,* and characterised. In total six discrete organisms were positive for antimicrobial activity and *in silico* analysis was used to narrow this down to the three most probable candidates carrying novel bacteriocins. One, *Streptococcus* 30ANiii, was demonstrated to secrete an antimicrobial compound however this could not be attributed to the *in silico* candidates. Another organism, *Lactococcus lactis* 10, carried the well-known nisin bacteriocin which was shown to be the agent responsible for the observed antimicrobial activity. The last organism, *Bacillus pumilus* 1_16, solely harboured a putative novel bacteriocin which was shown to be of a proteinous nature through demonstrating antimicrobial activity after peptides were precipitated using 40% ammonium sulphate. Although the predicted bacteriocin structural gene, *bacP*, was not successfully expressed in an *E. coli* host, the putative bacteriocin locus did result in a preliminary demonstration of antimicrobial activity suggesting that this locus encodes a novel bacteriocin. This study presents a novel putative bacteriocin that

could be investigated as a future treatment for *C. difficile*, providing further evidence of the potential of bacteriocins for medical applications.

Acknowledgments

For my mamgu, Glenys Jones, who was always proud of me and was only ever a phone call away.

I would like to acknowledge my supervisors, Dr Chris Penfold, Dr Sarah Kuehne, and Dr Alan Cockayne for all their help and advice throughout the last four years. I would also like to thank Michelle Kelly for all her advice, time, and chats in the lab and office. Thank you to all the members of A02, B37, and B90 for providing me with entertainment, endless laughter, and friendship. Finally, I would like to thank my family and friends for all their constant support and belief. I would especially like to thank both my brothers and sister-in-law's who braved 'The North' to visit me, and my parents who were always there for me, whatever the time difference.

Abbreviations

%	Percentage
° C	Degrees Celsius
µl	Microlitres
µg	Micrograms
µm	Micrometer
µM	Micromolar
Ω	Ohm
ABC	ATP-binding cassette
Acryl-bis	Bis-acrylamide
APS	Ammonium persulfate
BM	Bacillus Medium
BHI	Brain Heart Infusion
BHIS	Brain Heart Infusion Supplemented
BLAST	Basic Local Alignment Search Tool
bp	Base Pair
$CaCl_2$	Calcium Chloride
CATH	Class, Architecture, Topology and Homology
CC	*Clostridium difficile* Selective Supplement
CDAD	*Clostridiodies difficile* Associated Disease
CDI	*Clostridiodies difficile* Infection
cfu	Colony Forming Units
cm	Centimetres
CO_2	Carbon Dioxide
CV	Coloumn Volumes
ddH_2O	Deionised Water
DNA	Deoxyribonucleic Acid
DTT	Dithiothreitol
EDTA	Ethylenediaminetetraacetic Acid
EtOH	Ethanol
E-Value	Expected Value
FDA	Food and Drug Administration
FMT	Faecal Microbiota Transplantation
g	Grams
g	Gravity
GC	Guanine-Cytosine
GM17	Glucose M17
GRAS	Generally Regarded As Safe
h	Hours
HCL	Hydrochloric Acid
ID	Identity
IMAC	Immobilized Metal Affinity Chromatography
InterPro	Integrated Resource of Protein domains and functional sites
IPTG	Isopropyl β-D-1-thiogalactopyranoside
kbp	Kilo Base Pair
kg	Kilograms
kV	Kilovolts
l	Litres
LAB	Lactic Acid Bacteria
LB	Lysogeny Broth

LiAc	Lithium Acetate
M	Mole
Mb	Mega Base Pair
MCS	Multiple Cloning Site
MEGA7	Molecular Evolutionary Genetics Analysis Seven
mg	Milligrams
MIC	Minimal Inhibitory Concentration
min	Minutes
MIRA4	Mimicking Intelligent Read Assembly Four
ml	Milliliters
mM	Millimolar
mm	Millimeters
mRNA	Messenger RNA
MRSA	Methicillin Resistant *Staphylococcus aureus*
NaAc	Sodium Acetate
NaCl	Sodium Chloride
NCBI	National Centre for Biotechnology Information
NEB	New England Biolabs
ng	Nanograms
NHLM	Nitrile Hydratase Leader Microcin
Nisin	Group N Streptococcus Inhibitory Substance IN
No	Number
NSR	Nisin-Resistance Protein
OD	Optical Density
ORF	Open Reading Frames
PaLoc	Pathogenicity Locus
PBS	Phosphate Buffered Saline
PCR	Polymerase Chain Reaction
Pham	Protein Families
pH	Potential Hydrogen
PHASTER	Phage Search Tool – Enhanced Release
PMC	Pseudomembranous Colitis
PTM	Posttranslational Modification
QC	Query Cover
RAST	Rapid Annotation using Subsystem Technology
RiPPS	Ribosomally synthesized and Post translationally modified Peptides
RNA	Ribonucleic Acid
rRNA	Ribosomal RNA
RPM	Rotations Per Minute
RT	PCR Ribotype
RT-PCR	Reverse Transcriptase PCR
s	Seconds
SBRC	Synthetic Biology Research Centre
SDS	Sodium Dodecyl Sulphate
SNP	Single Nucleotide Polymorphism
SOC	Super Optimal Broth
SOE	Splicing by Overlap Extension
TAE	Tris-acetate-EDTA
TBS-T	Tris-buffered Saline – Tween 20
TEMED	Tetramethylethylenediamine
TMB	Tetramethylbenzidine

Tris	Tris(hydroxymethyl)aminomethane
U	Units
UK	United Kingdom
UniProt	Universal Protein Resource
UV	Ultraviolet
V	Volts
vF	Velocity Factor
VRE	Vancomycin Resistant Enterococci

List of Figures

Figure 1- Diagram of the steps involved in pathogenesis of CDI. 20
Figure 2-Schematic of Gram-negative bacteriocin classes. 28
Figure 3-The different classes of gram-positive bacteriocins. 29
Figure 4- Diagram of the different types of overlays. 57
Figure 5- The multiple cloning site and available His-Tags in pET28b(+) vector. 65
Figure 6- Processing of pre-nisin A to the mature nisin. 85
Figure 7-The organization of operons in the nisin biosynthesis pathway. 87
Figure 8-Nisin interaction with lipid II and pore formation 90
Figure 9- Zones of clearance of L. lactis 10 concentrated cell free supernatant before and after catalase treatment against a lawn of *C. difficile* CD630. 93
Figure 10- Gram-stain of *L. lactis* 10. 94
Figure 11- Growth curve of *L. lactis* strain 10 grown in GM17 at 30°C with 10% CO2 including CFU/ml. - 94
Figure 12- Contig alignment of *L. lactis* 10 to reference sequences, generated by Mauve. 97
Figure 13- Circular Genome Sequence of *L. lactis* 10. 98
Figure 14 - Close up 16s rRNA phylogenetic tree of *L. lactis* 10 99
Figure 15- Amplification of partial nis operon to confirm presence of the operon, producing 518 bp product. 102
Figure 16- Amplification of *bacL* and *bacE,* producing a 428 bp and 747 bp product respectively. 102
Figure 17- Sequence alignment of BacE, EnlA, ZooA, LytM, lysostaphin and ALE-1. 106
Figure 18-The amino acid sequence of BacL. 108
Figure 19-The putative BacL operon. 108
Figure 20- Plasmid maps of pET28b(+)bacL and pET28b(+)bacE. 110
Figure 21-SDS-PAGE gel showing protein profile of E. coli pET28b(+)*bacE* cell lysate taken from induce and non-induced culture over time. 111
Figure 22- SDS-PAGE gel and western blot of the nickel protein purification of *BacE* 113
Figure 23- SDS-PAGE gel of the zinc protein purification of BacE. 113
Figure 24-Purified BacE separated under different conditions on a Native-page gel and re-run on SDS-PAGE gel 115
Figure 25-10% NativePage gel showing nickel purified BacE. 116
Figure 26-The zone of clearance observed from BacE on the overlaid half of the Native-PAGE gel corresponding to protein bands visualised on the strained half of the Native-PAGE gel. 116
Figure 27- Figure of the allelic exchange method designed to knock-out genes of interest in *L. lactis* 10. 119
Figure 28- Minimal inhibitory concentration of ampicillin and erythromycin against *L. lactis* 10 in GM17 media. 120
Figure 29- *nisIFEG* fragment, pMTL84123_*nisIFEG* plasmid map and conjugation PCR colony screen. 126
Figure 30- Agarose gel of RNA extracted from *C. difficile* wild-type R20291 and *C. difficile* R20291ΔpyrE_pMTL84123*nisIFEG* and 16s rRNA PCR using the RNA as template demonstrating a lack of DNA contamination. 129

Figure 31- RT-PCR of cDNA extracted from three biological replicates of *C. difficile* R20291ΔpyrE_pMTL84123*nisIFEG* expressing the nisin immunity genes, *nisIFEG,* compared to the wild-type *C. difficile* R20291. 129
Figure 32-Increased levels of nisin resistance of *C. difficile* 630ΔermΔpyrE_pMTL84123*nisIFEG* and R20291ΔpyrE_pMTL84123*nisIFEG* compared to wild-type and negative control strains. 131
Figure 33- Colony overlay depicting conferred resistance against *L. lactis* 10 of *C. difficile* strains carrying pMTL84123*nisIFEG*. 132
Figure 34- Protocol for purifying and identifying single bacterial isolates from a mixed culture. 150
Figure 35- Gram-stains of environmental samples, 1_29A, 8ANE, 1_16A, 41ANE and 30ANiii 151
Figure 36- Growth curve of *B. licheniformis* 41ANE grown in BHIS medium shaking at 37 °C. 156
Figure 37-The genetic organisation of the predicted lichencidin bacteriocin within *B. licheniformis* 41ANE. 160
Figure 38- The organisation of the bacteriocin cluster identified within *Streptococcus* 30ANiii 162
Figure 39- 16s rRNA phylogenetic tree of *Streptococcus* 30ANiii against other streptococci. 166
Figure 40- Growth curve and CFU/ml of *Streptococcus* 30ANiii. 167
Figure 41- *Streptococcus* 30ANiii spent supernatant assay against *C. difficile* CD630, DH1916, and R20291. 168
Figure 42- *Streptococcus* 30ANiii freeze-dried spent supernatant assay against *C. difficile* CD630. 169
Figure 43- Plasmid maps of pET28b(+)*bov255* and pET28b(+)*bacS*. 172
Figure 44- SDS-PAGE gels of cell lysates obtained from induced *E. coli* BL21 (DE3) harbouring the overexpression vectors pET28b(+)*bov255* and pET28b(+)*bacS*. 174
Figure 45-Subgroups of circular bacteriocins 188
Figure 46– Comparison of the structure of subgroup i and ii circular bacteriocins. . 189
Figure 47- Gram-stain of Sample 1_16. 193
Figure 48- Growth curve and CFU/ml of *B. pumilus* 1_16 over 72 h. 193
Figure 49-Section of the 16s rRNA Maximum Likelihood Phylogenetic Tree of *B. pumilus* 1_16 195
Figure 50- 16s rRNA phylogenetic tree of strain 1_16 against *B. pumilus* and *B. safensis* with available whole genome sequences. 195
Figure 51- *B. pumilus* 1_16 contig organisation aligned with the reference strain *B. pumilus* SAFR-032, generated by Mauve. 196
Figure 52-Circular genome of B. pumilus 1_16. 198
Figure 53- The organisation of operons in the putative BacP biosynthesis pathway. 201
Figure 54- Predicted tertiary structure of BacP compared to other circular bacteriocins 202
Figure 55- Phylogenetic tree of the putative *B. pumilus* 1_16 bacteriocin compared to the other circular bacteriocins 202
Figure 56- Amino acid sequence alignment of circular bacteriocins in subgroup i, and BacP. 203

Figure 57- Amino acid sequence alignment of MccF from *B. pumilus* 1_16 compared to four S66 peptidase members. 205
Figure 58- An overlay of *B. pumilus* 1_16 against *C. difficile* CD305 demonstrating a zone of clearance. 207
Figure 59- *B. pumilus* 1_16 spent supernatant assay against *C. difficile* CD630, R20291, DH196 and CD196. 210
Figure 60- A well overlay of protein precipitated from *B. pumilus* supernatant using different percentages of ammonium sulphate. 213
Figure 61- Inhibition results of *C. difficile* CD630 growth by 40% ammonium sulphate precipitated proteins at different dilutions. 214
Figure 62- A zone of clearance on an overlaid SDS-PAGE gel corresponding to a protein band on the silver stained part of the SDS-PAGE gel. 215
Figure 63- Plasmid map of pMTL84123_*bacP*. 220
Figure 64- Colony overlay of two biological repeats of *E. coli* pMTL84123_*bacP* against *C. difficile* CD630 compared to a negative control. 220
Figure 65- Primers for the assembly of pMTL84121_PtcdB_*BpMccF*. 222
Figure 66- Plasmid map of the self-immunity gene, *bpMccF*, expression vector 223

List of Tables

Table 1- Subgroups of Gram-positive Bacteriocins 31
Table 2- List of Media Used Throughout this Study 44
Table 3- List of Antibiotics Used Throughout this Study 45
Table 4- List of Buffers Used Throughout this Study 46
Table 5- Complete List of Bacteria Strains Used Throughout this Study 47
Table 6- Complete List of Plasmids Used Throughout this Study 48
Table 7- Complete List of Oligonucleotides Used in this Study 49
Table 8- Growth Conditions for the Organism Used in this Study 51
Table 9- The Requirements for a Positive Result of the Inhibition of *C. difficile* in Spent Supernatant 52
Table 10- The Chemicals Used to Make Up the Stacking and Resolving Gel in the Native-PAGE gels, and the Protocol to Make the Gels. 74
Table 11- Activity of *L. lactis* 10 Against Different *C. difficile* Strains. 92
Table 12- Putative Bacteriocins Predicted to be in the Genome of *L. lactis* 10 and their Corresponding BLASTp Match 101
Table 13- Sequence Alignment Results of Enterolysin A and BacE Against Proteins Showing Homology to Enterolysin A. 104
Table 14- Table of Trialled Electroporation and Recovery Conditions for Electrocompetent *L. lactis* 10 122
Table 15- The Antimicrobial Activity of Different Environmental Samples Against Three Strains of *C. difficile*. 149
Table 16- The Gram-Stain, Cell Morphology and 16s rRNA PCR Results for the Environmental Organisms Demonstrating Antimicrobial Activity Against *C. difficile*. 152
Table 17- The Gram-Stain, Cell Morphology and 16s rRNA PCR Results for the Different Colony Morphologies Identified in Culture 30AN and their Antimicrobial Activity Against *C. difficile* 152
Table 18- Antimicrobial Activity Observed from Concentrated or Stress Induced Cell Growth of the Isolated Environmental Isolates 154
Table 19-The results of the selection of environmental organisms for further investigation 166
Table 20- Locations of Putative Phage Regions within *B. pumilus* 1_16 Whole Genome Sequence Determined by PHASTER. 199
Table 21- The Sequence Similarity of MccF from *B. pumilus* 1_16 Compared to LdcA and MccF from Other Organisms. 206
Table 22- The Antimicrobial Activity of *B. pumilus* 1_16 Against Different Strains of *C. difficile*. 207

Contents

Identification and Characterisation of Novel Bacteriocins Active Against *Clostridioides difficile* 0
Declaration 1
Abstract 2
Acknowledgments 4
Abbreviations 5
List of Figures 8
List of Tables 11
Chapter 1: General Introduction 17
1.1 The *Clostridium* Genus 17
1.2 *Clostridioides difficile* 17
1.2.1 History and Phylogeny 17
1.2.2 *C. difficile* Infections 18
1.2.3 Current Treatment of *C. difficile* Infections 23
1.2.4 Alternative Treatments for *C. difficile* Infections 24
1.2.5 Future CDI Treatment in Ongoing Clinical Trials 25
1.3 Bacteriocins 26
1.3.1 History of Bacteriocins 26
1.3.2 Bacteriocin Classification 28
1.3.3 Bacteriocins Produced by LAB 33
1.3.4 Bacteriocins Produced by the *Bacillus* Genus 36
1.3.5 Bacteriocins as Alternatives to Antibiotics 37
1.4 Bacteriocins Active Against CDI 39
1.5 Objectives of this Study 42
Chapter 2: Materials and Methods 43
2.1 General Microbiology Techniques 43
2.1.1 Media 43
2.1.2 Antibiotics 43
2.1.3 Solutions 46
2.1.4 Bacterial Strains 47
2.1.5 Plasmids 48
2.1.5 Oligonucleotide Primers 49
2.1.6 Growth Conditions 50
2.1.7 Growth Curve and Colony Forming Units 51

2.1.8 Spent supernatant assay 51
2.1.9 Minimal Inhibitory Concentration 53
2.1.10 Competent Cells 53
2.1.11 Transformation Protocols 54
2.1.12 Conjugation into *C. difficile* Strains 55
2.1.13 Overlays 56
2.1.14 Isolation of Bacteriocin Producing Strains 57
2.1.15 Gram-Staining 58
2.2 General Molecular Techniques 58
2.2.1 Genomic DNA Extraction via Phenol Chloroform 58
2.2.2 RNA Extraction 59
2.2.3 Polymerase Chain Reactions 60
2.2.4 Reverse Transcription 61
2.2.5 Plasmid Extraction 61
2.2.6 DNA Purification 62
2.2.7 Gel Electrophoresis 62
2.2.8 DNA Extraction from Agarose Gel 63
2.2.9 Restriction Enzyme Digestion 63
2.2.10 De-phosphorylation of DNA 64
2.2.11 DNA Fragment Ligations 64
2.2.12 Construction of a Vector for Immunity Gene Expression in *C. difficile* 65
2.2.13 Construction of Over-Expression Vectors in *E. coli* 65
2.2.14 Nucleotide Sequencing 65
2.3 Protein techniques 65
2.3.1 Protein Overexpression and His-tag Protein Purification 65
2.3.2 Cell Lysis 66
2.3.3 Nickel Resin Protein Purification 67
2.3.4 Zinc Resin Protein Purification 68
2.3.5 Protein Precipitation Techniques 69
2.3.6 Protein Gels 71
2.4 Bioinformatic Tools 74
2.4.1 Genome Assembly and Analysis Tools 74
2.4.2 BLAST 77
2.4.3 Plasmid and PCR DNA Fragment Sequencing 77
2.4.4 Visualising Sequencing Data 78
2.4.5 Plasmid Maps 78

2.4.6 Oligonucleotide Design 78
2.4.7 Promoter and Terminator Sequence Identification 78
2.4.8 Data Analysis 78
Chapter 3: Characterisation of a *Lactococcus lactis* Strain with Antimicrobial Activity Against *C. difficile* 79
3.1 Introduction 79
3.1.1 Aims of this Project 79
3.1.2 *L. lactis* Background 79
3.1.3 Bacteriocins of *L. lactis* 80
3.2 Results 92
3.2.1 Discovery of a Bacteriocin Producing Strain, Strain Characterisation and Initial Activity Against *C. difficile* 92
3.2.2 *De novo* Whole Genome Sequencing 95
3.2.3 Bacteriocin Mining 99
3.2.4 Activity of Putative Bacteriocins 109
3.2.5 Alternative Methods of Determining Activity 117
3.3 Discussion 133
3.3.1 Characterisation of Anti-*C. difficile* Activity Demonstrated by *L. lactis* 10133
3.3.2 Inconclusive Evidence of Active BacE Peptide 134
3.3.3 Expression of BacL in an *E. coli* Host 135
3.3.4 Evidence that Nisin is Responsible for the Observed Antimicrobial Activity Against *C. difficile* 137
Chapter 4: Isolation of Environmental Organisms from Cow Manure with Antimicrobial activity against *C. difficile* 141
4.1 Introduction 141
4.1.1 Aims of this Project 141
4.1.2 Reservoirs of *C. difficile* 141
4.1.3 Cattle as a Reservoir for *C. difficile* 143
4.1.4 The Microflora of Cattle and its Potential as a Reservoir for Bacteriocin Harbouring Bacteria 147
4.2 Results 148
4.2.1 Isolation of Environmental Organism Active Against *C. difficile* 148
4.2.2 Characterization of the Isolated Environmental Organisms 149
4.2.3 Identification of Isolates Determined to Require Further Investigation 153
4.2.4 Further Characterization of *Streptococcus* 30ANiii 166

4.2.5 Determining Whether the Known Bacteriocin, Bov255 and/or the Putative Bacteriocin Candidate, BacS, Play a Role in the Observed Anti-*C. difficile* Activity 170
4.3 Discussion 174
4.3.1 Successful Isolation and Purification of Environmental Isolates with Antimicrobial Activity Against *C. difficile* 174
4.3.2 Rationale for Selection of Isolates for More Detailed Study 177
4.3.3 Successful Demonstration of Antimicrobial Activity from *Streptococcus* 30ANiii Against *C. difficile* 179
4.3.4 Overexpression of the *In Silico* Bacteriocin Candidates Identified within *Streptococcus* 30ANiii in an *E. coli* Host 180
4.3.5 Precipitation of Bovicin255 from *Streptococcus* 30ANiii 182
Chapter 5: Characterisation of a *Bacillus pumilus* Strain with Antimicrobial Activity Against *C. difficile* 184
5.1 Introduction 184
5.1.1 Objectives of this Project 184
5.1.2 Bacillus pumilus 184
5.1.3 Known Bacteriocins Produced by *Bacillus pumilus* 185
5.1.4 Circular Bacteriocins 186
5.2 Results 192
5.2.1 Strain Characterisation 192
5.2.2 *De novo* Whole Genome Sequencing 193
5.2.3 Bacteriocin Mining 197
5.2.4.2 Activity Demonstrated in the Supernatant of B. pumilus 1_16 208
5.2.5 Determining the Active Peptide Secreted by *B. pumilus* 1_16 211
5.2.6 Determining the Active Peptide through Expression in an *E. coli* Host 217
5.2.6.1 217
5.2.7 Identifying the Nature of the Antimicrobial Activity 221
5.3 Discussion 223
5.3.1 Characterisation of Antimicrobial Activity Observed from *B. pumilus* 1_16 Against *C. difficile* 223
5.3.2 Preliminary Evidence that BacP is a Putative Bacteriocin 229
Chapter 6: General Discussion 234
6.1 Addressing the Objectives of this study 234
6.2 Use of *In Silico* Mining in the Identification of Novel Bacteriocins 236
6.3 Comparing the Different Methods of Isolating Novel Bacteriocins 237
6.4 Future work 239

7 Appendix 242
Appendix 7.1 242
Appendix 7.2 243
Appendix 7.3 245
Appendix 7.4 – PIPS reflective statement 247
8 Bibliography 248

Chapter 1: General Introduction

1.1 The *Clostridium* Genus

The genus *Clostridium* contains Gram-positive, anaerobic, spore-forming bacilli, which are separated into distinct Clostridia clusters (Collins *et al.*, 1994). The genus was first proposed in 1880 by Prazmowski (Prazmowski, 1880), and its members are commonly found in a range of environments, from human and animal intestinal flora to sewage and soil (Wells and Wilkins, 1996). The most common clinically relevant species within this genus which cause a range of diseases including tetanus, (*Clostridium tetani),* botulism (*Clostridium botulinum*), gas gangrene (*Clostridium perfringens*) and diarrhoea and pseudomembranous colitis (*Clostridium difficile*) through the secretion of exotoxins (Wells and Wilkins, 1996; Actor, 2012).

1.2 *Clostridioides difficile*

1.2.1 History and Phylogeny

Originally described in 1935 by Hall and O'Toole, *Clostridioides difficile* was thought to be a normal component of the gastrointestinal flora of neonates (Hall and O'Toole, 1935). Although initially classified within the *Clostridium* genus, the similarity of *C. difficile* to the genus type strain, *Clostridium butyricum,* has been recently debated. It was proposed that the members of the *Clostridium* genus should be restricted to *C. butyricum* and its cluster I relatives only. Thus *C. difficile* was reclassified in 2013 to the *Peptostreptococcaceae* family, creating a new genus *Peptoclostridium* (Yutin and Galperin, 2013; Lawson and Rainey, 2016). The renaming of *C. difficile* to *P. difficile*, however, had cost implications to the health sector and was also confusing amongst the public, as well-known terms such as 'C. diff' and *Clostridium difficile* infection (CDI), would change to 'P. diff' and PDI (Lawson *et al.*, 2016). Due to these implications a new genus was coined, *Clostridioides*, in which *Clostridioides difficile* was

the type strain (Lawson *et al.*, 2016). In keeping with the new nomenclature, the organism in question will be referred to as *Clostridioides difficile* throughout this study.

There are different methods used to identify and classify strains of *C. difficile,* including multi-locus sequence typing (MLST), pulsed-field gel electrophoresis (PFGE) and PCR ribotyping (RT) (Knight *et al.*, 2015). The most widely accepted are PFGE and PCR ribotyping (Knight *et al.*, 2015); the former recognises different strain specific patterns of restriction digested chromosomal DNA whereas the latter uses variations in the intergenic spacer region (ISR) between the 16s and 23s rRNA genes to generate a strain specific pattern of PCR amplicons (Stubbs *et al.*, 1999; Knetsch *et al.*, 2013).

The most recent phylogenetic studies of *C. difficile* suggest that there are four distinct phylogenetic groups (PG), PG1-4, which encompass *C. difficile* isolated from the environment, humans and animals. Although there was evidence of ongoing speciation, it was determined that PG4 (Clade B) emerged prior to PG1-3, which are collectively known as 'Clade A' (Kumar *et al.*, 2019). Although there is a phylogenetic clustering of *C. difficile*, distinct clinically relevant strains were found throughout the phylogenetic groups and clades, for example the hypervirulent ribotypes RT078 and RT027 are in PG4 and PG2, respectively (Kamuju *et al.*, 2018; Kumar *et al.*, 2019).

1.2.2 *C. difficile* Infections

Pathogenic *C. difficile* produces toxins that mediate colonic disease in both animals and humans. Disease occurs when the natural barrier properties of the healthy gut microbiota are disrupted (Figure 1), often caused by broad spectrum antibiotics (Leffler and Lamont, 2015). Thus antibiotic therapy is a major risk factor for CDI as well as hospital stays, old age and underlying conditions such as inflammatory bowel disease (IBD), which would undermine the healthy gut microbiota or the immune

system (Goudarzi *et al.*, 2014). Diseases caused can range from mild self-limiting diarrhoea to pseudomembranous colitis and toxic megacolon, which can be fatal (George *et al.*, 1978). *C. difficile* strains can possess two major virulence factors, toxins TcdA and TcdB encoded by *tcdA* and *tcdB* respectively which are located in the LCT pathogenicity locus (PaLoc) (Hammond and Johnson, 1995; Genth *et al.*, 2008). Toxin B alone has been found to cause the symptoms of CDI however hyper-production of toxin A results in the same symptoms (Carter *et al.*, 2010; Kuehne *et al.*, 2010).There is also a third toxin a binary toxin termed *C. difficile* transferase (CDT) which is produced by some *C. difficile* strains and is thought to contribute to the virulence of these strains (Gerding *et al.*, 2014; Berry *et al.*, 2017; Oguri *et al.*, 2019). It is composed of two peptides (CdtA and CdtB) which induce diarrhoea through impairing the cytoskeleton in epithelial cells (Gerding *et al.*, 2014) however its clinical relevance has not been fully elucidated (Berry *et al.*, 2017).

Spores germinate into vegetative cells which adhere to intestinal epithelial cells after penetrating the mucus layer (Pérez-Cobas *et al.*, 2015). The exotoxins A and B bind to cell membrane receptors and enter the target cell through receptor mediated endocytosis (Na *et al.*, 2008; LaFrance *et al.*, 2015; Yuan *et al.*, 2015). The toxins undergo a conformational change which allows them to escape the endosome into the cytoplasm of the cell (Monaghan *et al.*, 2015). In the cell, toxins A and B inactivate Rho GTPases, which leads to various physiological effects on the cell including cell shrinking and rounding resulting in apoptosis (Voth and Ballard, 2005; Di Bella *et al.*, 2016). There are many cytoskeletal changes which lead to cell death, and TcdB is known to activate apoptosis directly (Qa'Dan *et al.*, 2002; Pérez-Cobas *et al.*, 2015). These alterations in the cell also lead to increased permeability in the colonic epithelium permitting the accumulation of inflammatory agents. It has been demonstrated that

TcdA stimulates cytokine production further increasing the inflammatory response; a response that has been linked to pseudo-membranous colitis (Voth and Ballard, 2005).

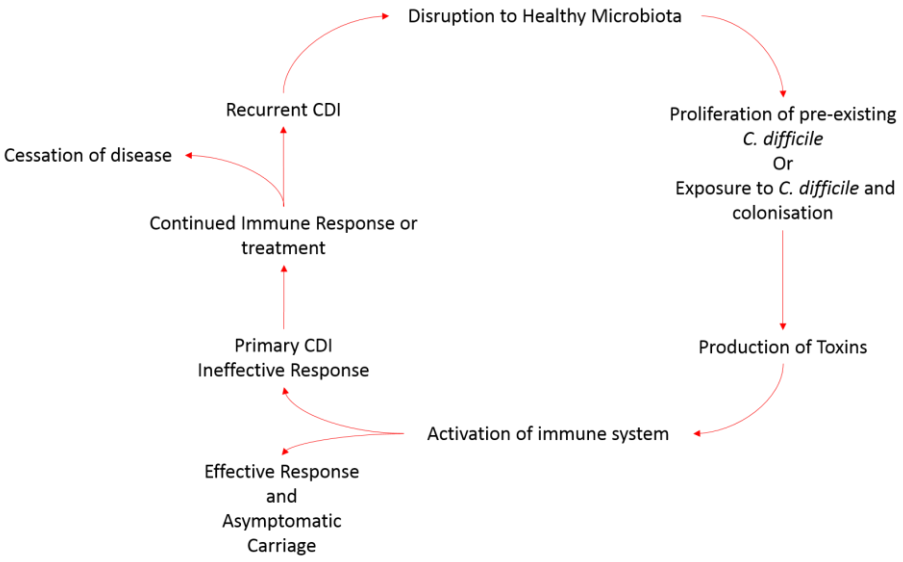


Figure 1- Diagram of the steps involved in pathogenesis of CDI. CDI, is initiated by disruption of the microbiota and the cessation of disease either occurs with an effective initial immune response, continued immune response or treatment. If disease is not successfully treated, RCDI can occur and the disease can persist.

An infection with this pathogen can clinically present in humans in four ways: short-term colonisation, acute diarrhoea, fulminant diarrhoea and recurrent infection (DuPont, 2011). Acute diarrhoea can range from mild to severe, however fulminant diarrhoea is linked to pseudo-membranous colitis and leukocytosis, which has a

mortality rate ranging between 3% and 15% (Pépin *et al.*, 2005; Abou Chakra *et al.*, 2014; Sheitoyan-Pesant *et al.*, 2015). CDI also has a high rate of recurrence ranging from 12% to 40% (Pépin *et al.*, 2005; Abou Chakra *et al.*, 2014; Sheitoyan-Pesant *et al.*, 2015) and recurrent CDI (RCDI) has been associated with higher mortality and economic burden (Garey *et al.*, 2008). RCDI is defined as either a 'relapse' from the original infection or 'reinfection' where a new or original strain is acquired from the environment (Shields *et al.*, 2015). It is unclear why there is such a high recurrence rate for CDI however it has been suggested to be caused by inadequate host immune response, persistence of spores within the gut after completion of antibiotic treatment, or that the intestinal-flora fails to be re-established (Figure 1) (Garey *et al.*, 2008; DuPont, 2011).

CDI is usually a hospital acquired disease with cases reaching an all-time high in 2007/8 in the UK at 108 cases per 100,000 population. This number has since substantially decreased to 26.3 cases per 100,000 population due to the implementation of improved procedures to limit the spread of disease including patient isolation, improved antibiotic prescribing, and improved hand hygiene (Office for National Statistics, 2013). However in the USA there is still a high number of CDI cases in patients who have not been to a hospital, a condition called community-acquired CDI (CA-CDI) (Lessa *et al.*, 2015). This highlights the prevalence of CDI, not only in the hospital setting but in the community as well (Shields *et al.*, 2015). The challenge is to be able to prevent and manage these infections, as they are an extreme burden to the medical system.

Primary CDI (PCDI) has been estimated to cost $5,243-$8,570 per case in Canada, the UK and Ireland. Higher costs, however, accumulate due to RCDI cases which are estimated per case to cost between $13,655 in Canada, the UK and Ireland

(Ghantoji *et al.*, 2010). With RCDI estimated in a high amount of PCDI cases (Aslam *et al.*, 2005; Bauer *et al.*, 2011; Lessa *et al.*, 2015), CDI is a significant economic burden to the healthcare systems in the US, Canada, the UK and Ireland with an annual associated cost between ~$400-$800 million in the US alone. Thus, the use of additional resources to prevent and control CDI is justified to reduce these costs (Ghantoji *et al.*, 2010). Another issue with CDI is the growing number of CA-CDI cases which have been increasing in the last 10-20 years (Khanna and Pardi, 2010). Over a 15 year period, 41% of CDI cases were CA-CDI (Gupta and Khanna, 2014) and in the USA there has been a significant increase in CA-CDI accounting for approximately half of hospitalised cases. These cases are particularly problematic as they are prevalent in patients previously considered low-risk, such as the younger generation and patients without prior exposure to antibiotics or gastric acid-suppressing medication.

Although the community-acquired cohort are less likely to experience severe infection, in the elderly, 61% of CA-CDI cases required hospitalisation (Khanna *et al.*, 2012). The large proportion of CA-CDI may suggest the presence of an undiscovered reservoir of infection, further discussed in Section 3.1, or may be simply due to higher clinician awareness of the disease resulting in more requests for *C. difficile* tests (Gupta and Khanna, 2014). One specific ribotype that is prevalent in CA-CDI are RT078 lineages which are hyper-virulent and have increased association with morbidity and mortality. It was previously thought this was due to a hyper-production of toxin A and B however this remains unclear as an increase in persistence due to higher spore formation and increased adherence has also been suggested (Warny *et al.*, 2005; Goorhuis *et al.*, 2008; Monaghan *et al.*, 2015; Kamuju *et al.*, 2018).

1.2.3 Current Treatment of *C. difficile* Infections

Current treatments for CDI are the antibiotics, fidaxomicin, and vancomycin, for more severe cases (McDonald *et al.*, 2018). Metronidazole is an inexpensive standard for non-severe cases, however almost all of the drug is absorbed in the gut resulting in very low concentrations reaching the colon, and there have been increasing reports of treatment failure (DuPont, 2011; Kelly, 2012). Due to this, the recommended use of metronidazole has been altered from first line treatment to non-severe initial cases when there is limited access to fidaxomicin and vancomycin (McDonald *et al.*, 2018). Although resistance to vancomycin is relatively low there are other resistant bacteria such as Vancomycin Resistant Enterococci (VRE), which have been found to coexist with *C. difficile* in more than 50% of CDI patients (Fujitani *et al.*, 2011; Spigaglia, 2016). This means there is an increased potential risk of gene transfer from VRE to *C. difficile* and given how quickly antimicrobial resistance evolves there is a need for multiple alternative treatments. There have also been reports of reduced susceptibility of certain *C. difficile* ribotypes to metronidazole and vancomycin (Baines *et al.*, 2008; Leeds *et al.*, 2014). New narrow spectrum treatments are required which will not severely affect the natural gut microflora of the host yet have a higher treatment success than metronidazole and a lower potential risk of development of antimicrobial resistance than vancomycin. One newer treatment is the antibiotic fidaxomicin, marketed as Dificid® and Dificlir®, and it has bactericidal activity against a range of Gram-positive bacteria including *C. difficile*. This antibiotic inhibits RNA synthesis and is non-inferior to vancomycin in the treatment of patients suffering mild to moderate CDI (Johnson and Wilcox, 2012). Fidaxomicin results in significantly fewer recurrences than vancomycin except in cases of NAP1/B1/027 strain infections (Petrella *et al.*, 2012). It is not absorbed in the gut and maintains levels up to 10,000-

fold higher than typical minimal inhibitory concentrations (MICs) for *C. difficile* strains, thus it has been speculated that resistance is unlikely to develop rapidly (Johnson and Wilcox, 2012). However, in the phase III clinical trial one strain showed significant increase in MIC although it was not above the threshold for resistance. Goldstein *et al*. (2011) indicated the need for resistance surveillance, especially in NAP1/B1/027 cases, and again expressed the need for alternative treatments. As fidaxomicin has similar clinical response rates to vancomycin and reduced RCDI rates, it is a recommended first line treatment despite fidaxomicin being significantly more expensive than vancomycin (Louie *et al.*, 2011; Cruz, 2012; Al Momani *et al.*, 2018).

Another, relatively new, treatment is bezlotoxumab (Zinplava™), a monoclonal antibody that binds to and neutralises the toxic effects of toxin B reducing RCDI rates in at-risk patients (Yang *et al.*, 2015; Wilcox *et al.*, 2017). This treatment was only approved for clinical use in 2017 and there are no current guidelines for its use in the UK (National Institute for Health and Care Excellence) however it is currently being reviewed for the next update in recommended treatment guidelines by the Infectious Disease Society of America (IDSA) (McDonald *et al.*, 2018). Its early clinical use will provide more data determining its true efficacy in at-risk patients and efficacy in reducing RCDI rates (Alonso and Mahoney, 2018).

1.2.4 Alternative Treatments for *C. difficile* Infections

CDI is a disease that occurs with the disruption of the healthy gut microbiome allowing *C. difficile* to proliferative and produce toxins. The reconstitution of the gut microbiome is a possible treatment for CDI, and this can be acquired through faecal microbiota transplantation (FMT). FMT is the transplantation of healthy faecal matter to the intestines of a CDI patient thus restoring a healthy gut microflora in the recipient (Smits *et al.*, 2013; Parmar and Glauser, 2019). It has been used to treat RCDI

successfully in 87-90% of cases in the literature, and shown few adverse side effects (Smits *et al.*, 2013; Gupta *et al.*, 2016). It has also been demonstrated that the treatment success rate increases with multiple administrations of faecal matter with up to 90% success after more than two administrations (Cammarota *et al.*, 2015; McDonald *et al.*, 2018). Recently a death occurred due to inadequate screening of donor faeces and the transmission of multidrug resistant *Escherichia coli* to immunocompromised patients (FDA, 2019; Giles *et al.*, 2019). Despite the FDAs warning for proper screening for multiple drug resistant organisms among donors (FDA, 2019) the recent death highlights the lack of research which is fundamental for ensuring patient safety through generating standardised regulations and elucidating long term effects of FMT (Smits *et al.*, 2013; Giles *et al.*, 2019). This lack of information is reflected in the current guidelines only recommending the use of FMT when there has been multiple RCDIs and antibiotic treatment has failed (McDonald *et al.*, 2018).

Another treatment option is the use of probiotics to restore a healthy gut microbiota. There is evidence that certain probiotics can have a protective effect and positively act against pathogenic *C. difficile* (Ollech *et al.*, 2015). Probiotics have been demonstrated to restore epithelium permeability caused by CDI and disrupt *C. difficile* TcdA and TcdB toxins binding their receptors (Kalliomäki and Walker, 2005; Banerjee *et al.*, 2009). Some probiotic strains produce antimicrobial peptides, bacteriocins, that have antimicrobial activity towards pathogenic *C.difficile* strains (Rea *et al.*, 2007).

1.2.5 Future CDI Treatment in Ongoing Clinical Trials

There are a number of antimicrobials currently in development with some in clinical trials. Those that show the most potential are Surotomycin, Fusidic acid, and Nitazoxanide which have undergone phase III clinical trials (Petrosillo *et al.*, 2018). Surotomycin and Fusidic acid show similar clinical response rates to the current CDI

treatment, however 55% of the *C. difficile* strains, after Fusidic acid treatment, showed resistance (Wullt and Odenholt, 2004; Norén *et al.*, 2006; Daley *et al.*, 2017). Despite these results, the low cost of Fusidic acid means it could still be used as future CDI treatment where current treatment is not suitable (Petrosillo *et al.*, 2018). In clinical trials Surotomycin demonstrated a lack of superiority over vancomycin implying it will not be a future treatment (Daley *et al.*, 2017; Petrosillo *et al.*, 2018). Nitazoxanide also showed similar clinical response rates and recurrence rates as the current treatment but was well tolerated and still could be a promising option for future treatment especially for RCDI (Musher *et al.*, 2006; Rafiullah *et al.*, 2011). Six other antimicrobials have undergone phase II clinical trials including Rifaximin, Cadazolid, Ridinilazole, LFF571, Ramoplanin, and Rifampin (Petrosillo *et al.*, 2018). Both Cadazolid and Ridinilazole are promising candidates that showed lower recurrence rates and higher sustained clinical response rates than vancomycin (Louie *et al.*, 2015; Petrosillo *et al.*, 2018). Vaccinating against *C. difficile* is another avenue that is being researched for the treatment of CDI. There are two vaccines currently in phase II clinical trials produced by Pfizer (Inc. P, 2017) and Valneva (Bézay *et al.*, 2016). Both these vaccines target the TcdA and TcdB toxins however the disadvantageous of this is that it results in asymptomatic carriers who could still transmit the disease, yet this is also a disadvantage of monoclonal therapies which are already in clinical use (Alonso and Mahoney, 2018).

1.3 Bacteriocins

1.3.1 History of Bacteriocins

Bacteriocins are small ribosomally synthesised antimicrobial peptides (AMP) which are produced by bacteria. These antimicrobial peptides can have narrow (same

species) or broad (across genera) spectrum activity against other bacteria, with the producer strain typically remaining unharmed due to specific immunity mechanisms (Cotter *et al.*, 2005). The first identified bacteriocins, later termed Colicins, were identified in 1925 when antimicrobial activity was observed from one *E*. *coli* strain against another *E. coli* strain (Gratia, 1925). Since then it has been hypothesised that between 30-99% of Bacteria and Archaea harbour at least one bacteriocin (Klaenhammer, 1988). These bacteriocins are thought to be produced as a part of a natural defence mechanism allowing the producer organism to compete for limiting resources (Nes and Holo, 2000; Chikindas *et al.*, 2018).

There is an increasing need for novel antimicrobials due to the emerging antibiotic resistance crisis, and bacteriocins are a possible solution. It has been suggested that bacteriocins have unknowingly been used in food production for over 8,000 years within fermented food (Cotter *et al.*, 2005). This is due to the identification of bacteriocins produced by well-known lactic acid bacteria (LAB) for example *Lactococcus lactis* which has been historically used in the dairy industry to produce products including cheese and yogurt (Song *et al.*, 2017). Due to its use in food production for centuries this strain has been given 'Generally Regarded as Safe' (GRAS) status, which has also been granted to many other LAB strains. This is significant as it means bacteriocins produced by GRAS strains or the producer organisms can be added to food products whereas antibiotics are not allowed to be used for this purpose (Chen and Hoover, 2003). Other advantages of bacteriocins are that they are typically heat and pH stable, meaning they can survive food process and they are usually protease sensitive resulting in low toxicity (Cotter *et al.*, 2013). Due to their ribosomal nature they can also be manipulated allowing for bioengineering (Cotter *et al.*, 2013). Despite their advantages there are currently only two commercially available bacteriocins and only a few have been approved for use in food (Dicks *et al.*, 2018). Nisin (Shin *et al.*,

2016) and pediocin PA-1 (Venema *et al.*, 1997) are successfully commercialised bacteriocins however there are always hurdles to overcome for each bacteriocin with respect to stability and solubility issues, potential cytotoxicity and production costs. Instead of successfully producing commercially viable peptides, bacteriocin-producing strains can also be commercialised such as BLIS K12™ and BLIS M18™ (Dicks *et al.*, 2018).

1.3.2 Bacteriocin Classification

There are a large number of bacteriocins, and they have been classified into different groups. These groups are separated based on the characteristics of the producer strain, protein characteristics, mode of action and more. The two main groups are Gram-negative (Fig 2) and Gram– positive (Fig 3) bacteriocins.

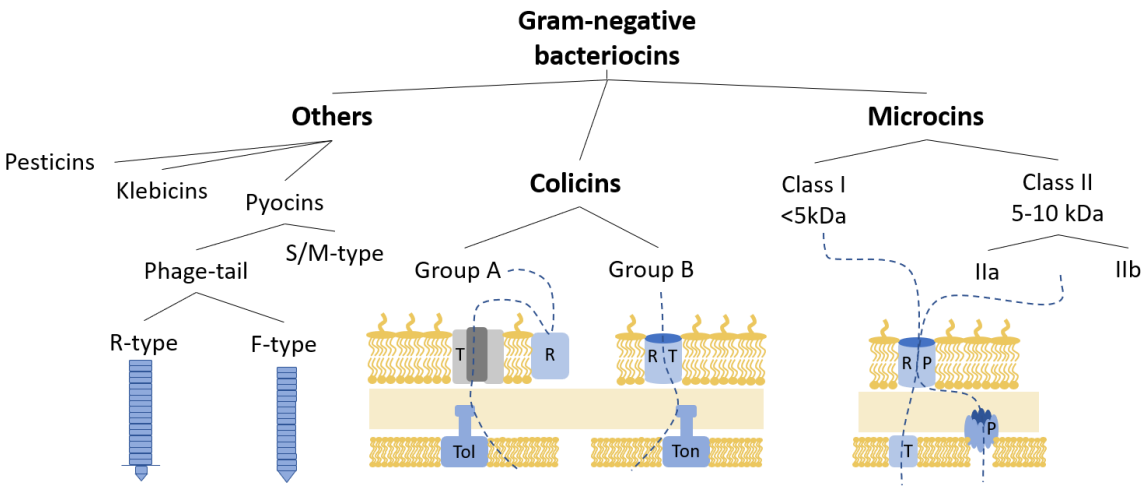


Figure 2-Schematic of Gram-negative bacteriocin classes. Gram-negative bacteriocins are classified into two main groups: colicins and microcins. Colicins are subgrouped into Group A which interact with at receptor (R) then a transporter (T) and utilise the Tol translocation system. Group B interact with a receptor that is also a transporter (RT) and utilises the TonB translocation system. Microcins are subgrouped into Class I and II, both enter the cell through a receptor or porin (RP). Another group with subgroups are pyocins split into colicin-like S/M-type and Phage-tail type groups.

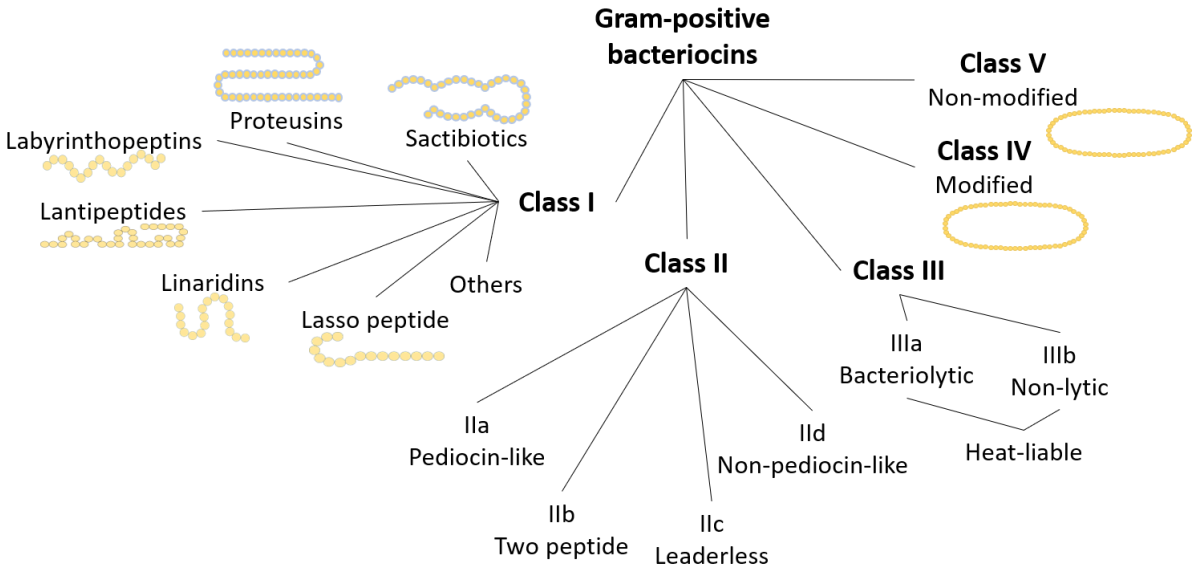


Figure 3-The different classes of gram-positive bacteriocins.The different subgroups of gram-positive bacteriocins separated through various protein features. A protein structure of an example peptide within subgroup demonstrated in table 1 is depicted for some classes.

1.3.2.1 Gram-negative Bacteriocins

Gram-negative bacteriocins are typically narrow spectrum and are produced by many different bacterial species such as *Pseudomonas aeruginosa* and *Escherichia coli* however the best characterised Gram-negative bacteriocins are produced by the latter (Rebuffat, 2011). *E. coli* bacteriocins are separated into two main subgroups, colicins which are the high molecular weight (20- 90 kDa) group and microcins are the low molecular weight peptides (<10 kDa) (Fig 2). The colicin subgroup are classified based on the machinery used to translocate the colicin to its target, Group A uses the Tol system and Group B the TonB system (Davies and Reeves, 1975b, 1975a). Typically, colicins are produced when the *E. coli* cells are under stress, and the expression of the colicin peptide is lethal to the producing cell due to the co-expression of a lysis protein essential for the release of the bacteriocin into the medium (van der Wal *et al.*, 1995; Braun *et al.*, 2002). There are typically three modes of action of colicins; the formation

of voltage dependent channels in the membrane, nuclease action, or peptidoglycan degradation (Barreteau *et al.*, 2010; Rebuffat, 2011).

The majority of microcins are expressed by *E. coli* typically under stress and are more similar to Gram-positive bacteriocins. Microcins are separated into classes by molecular weight; Class I consists of low molecular weight (<5 kDa) bacteriocins that undergo post-translation modification (PTM), whereas the molecular weight of class II is between 5-10 kDa. Class II is separated further by gene characteristics and protein structure, IIa microcins require three genes for the synthesis and assemble of a functional microcin whereas IIb are all linear microcins that can either undergo PTM or not (Yang *et al.*, 2014). Microcins act in a receptor-mediated manner with class I inhibiting enzymes, and class II targeting the inner membrane or membrane components (Rebuffat, 2011).

Other Gram-negative species are known to produce bacteriocins such as pyocins from *Pseudomonas* spp. (Michel-Briand and Baysse, 2002), pesticins from *Yersina pestis* (Patzer *et al.*, 2012) and klebicins produced by *Klebsiella pneumoniae* (Chavan *et al.*, 2005). These organisms typically produce colicin-like bacteriocins, for example the pyocins and pesticins, some of which show an evolutionary relationship with colicins (Michel-Briand and Baysse, 2002; Rebuffat, 2011). Microcins are also known to be produced by other Gram-negative species including *K. pneumoniae* which produces microcin E492 (Hetz *et al.*, 2002; Rebuffat, 2011).

1.3.2.2 Gram-Positive Bacteriocins

The subgroups of Gram-positive bacteriocins are more complex than Gram-negative bacteria with a greater variety of bacteriocins classified into different classes and examples are given in Table 1. Heat-stable bacteriocins can be classed in class I or

class II which are differentiated by molecular weight and PTM. Class I consists of low molecular weight bacteriocins that undergo PTM and can either be linear or globular bacteriocins. Class II bacteriocins consists of <10 kDa unmodified bacteriocins and both class I and II are then subdivided into multiple subclasses (Table 1) (Yang *et al.*, 2014). A large group of class I bacteriocins are the lantipeptide subgroup, which consists of peptides containing lanthionine and/or methyllanthionine residues. These bacteriocin have a variety of killing mechanisms with the majority of characterised lantipeptides disrupting membrane integrity or inhibiting the biosynthesis of the cell wall (Knerr and van der Donk, 2012). Lantipeptides are grouped into four classes based on their biosynthetic machinery, with class I requiring LanB dehydratase and LanC cyclase peptide, class II requiring LanM synthetase, class III requiring a LanC/LanM like trifunctional synthetase and class IV utilising a LanL synthetase peptide (Knerr and van der Donk, 2012). Class III of Gram-positive bacteriocins are larger proteins, usually >30 kDa and are heat liable bacteriocins. There are only two subclasses which are distinguished based on whether the bacteriocin is a bacteriolytic enzyme (Group A) or a non-lytic protein (Group B) (Table 1) (Yang *et al.*, 2014). There is a Class IV of complex bacteriocins that contain lipid and carbohydrate moieties in their chemical structure and are also cyclic peptide that have covalently linked N and C termini (Hammami *et al.*, 2013). Another newer class, Class V, differs from Class IV as the bacteriocins are unmodified cyclic peptides (Kemperman *et al.*, 2003).

Table 1- Subgroups of Gram-positive Bacteriocins

Subgroup I	Feature	Example	Reference
Lantipeptide	Contain lanthionine or β-methyl-lanthionine	Nisin	(Yang *et al.*, 2014)

Linaridins	Contain dehydrated amino acids and linear structure	Grisemycin	(Arnison *et al.*, 2013; Cotter *et al.*, 2013)
Lasso peptide	Lasso style structure	Lassomycin	
Linear azole-or azoline-containing peptide	Heterocyclic structure	MccB17	
Proteusins	Extensively modified and contain large leader region	Polytheonamide A	(Arnison *et al.*, 2013)
Sactibiotics	Cysteine Sulphur-α-carbon crosslink	Subtilosin A	(Arnison *et al.*, 2013; Yang *et al.*, 2014)
Cyanobactins	N-to-C macrocyclic	Anacyclamide	(Arnison *et al.*, 2013; Cotter *et al.*, 2013)
Bottromycins	Decarboxylated C-terminal thiazole and macrocylic amidine	Bottromycin A2	(Arnison *et al.*, 2013)
Glycocins	*S*-linked glycopeptides	Sublancin 168	(Cotter *et al.*, 2013)
Labyrinthopeptins	Labionin structure	Labyrinthipeptin A2	(Arnison *et al.*, 2013; Yang *et al.*, 2014)
Subgroup II	**Feature**	**Example**	**Reference**
IIa Pediocin-like	Flexible hinge separating N and C terminal	Pediocin PA-1	(Cotter *et al.*, 2013)
IIb Two peptide	Two peptides usually required for activity	Lactacin F	
IIc Leaderless	No N-terminal leader peptide and PTM on C-terminus	Enterocin L50	(Cintas *et al.*, 1998; Cotter *et al.*, 2013)

IId	Linear single peptide bacteriocins that are non-pediocin-like	Lactococcin A	(Cotter *et al.*, 2013)
Subgroup III	**Feature**	**Example**	**Reference**
IIIa	Bacteriolytic enzymes	Enterolysin A	(Nilsen *et al.*, 2003; Yang *et al.*, 2014)
IIIb	Non-lytic proteins	Caseicin	(Müller and Radler, 1993; Alvarez-Sieiro *et al.*, 2016)
Subgroup IV	**Feature**	**Example**	**Reference**
Class IV	Cyclic peptides with lipid or carbohydrate moieties	AS-48	(Gálvez *et al.*, 1986; Papagianni and Anastasiadou, 2009; Hammami *et al.*, 2013)
Subgroup V	**Feature**	**Examples**	**Reference**
Class V	Non-modified cyclic peptides	Circularin A	(Kemperman *et al.*, 2003)

1.3.3 Bacteriocins Produced by LAB

One group of bacteria that has been extensively researched, in relation to bacteriocins, are LAB. The term LAB is used for bacteria that produce energy from fermenting sugar and generated lactic acid which are the *Lactobacillales* order. This order encompasses multiple genera including *Pediococcus*, *Lactobacillus*, *Lactococcus* and *Streptococcus* (Dobbins, 2010), which the majority of these strains possessing GRAS status. However there are some known pathogens with the most notable being within the *Streptococcus* genus including *Streptococcus pyogenes, Streptococcus pneumoniae,* and *Streptococcus agalactiae* (Krzyściak *et al.*, 2013). LAB are known to produce a diverse variety of bacteriocins ranging across the different subgroups of Gram-positive

bacteriocins. Some of the best characterised bacteriocins produced by LAB are nisin (Class I) (Mulders *et al.*, 1991), lactococcins (Class II) (Nissen-Meyer *et al.*, 1992), lacticins (Class I) (Fujita *et al.*, 2007), pediocin PA1 (Class II) (Papagianni and Anastasiadou, 2009) and enterocins (Class II) (Cintas *et al.*, 1998). Nisin, produced by *L. lactis*, is a commercial bacteriocin (Chen and Hoover, 2003) which is the most characterised bacteriocin and is described in Section 3.1.3.4.

1.3.3.1 Streptococcal Bacteriocins

The majority of bacteriocins known to be produced by *Streptococcus* spp. are lantibiotics, although class II bacteriocins have also been characterised (Nes *et al.*, 2007). One of the first Streptococcal bacteriocins identified was the *Streptococcus salivarius* lantibiotic salivarcin A, which strongly inhibits *S. pyogenes* (Ross *et al.*, 1993). Since then bacteriocins have been found to be abundant across the *Streptococcus* genus, with mutacins produced by *Streptococcus mutans*, and nisin U and uberolysin produced by *Streptococcus uberis* (Wirawan *et al.*, 2006, 2007). One group of *Streptococcus* spp. that produce well known Streptococcal bacteriocins are the Group D *Streptococci* (GDS). Members of GDS were originally classified due to their Lancefield group D reactivity, and included strains of animal and human origin that displayed both non-hemolytic and α-hemolytic activity (Lancefield, 1933; Deibel, 1964). The GDS *Streptococcus bovis-Streptococcus equinus* complex contains *Streptococcus gallolyticus* subsp. *Macedonicus*, a food-grade GDS which produces macedocin, a lantibiotic with broad range activity against Gram-positive bacteria including other LAB (Georgalaki *et al.*, 2002). Another known bacteriocin produced by this organism is macedocin ST91KM which, in contrast, is a narrow range bacteriocin demonstrating activity against Gram-positive bacteria typically associated with mastitis (Pieterse *et al.*, 2008). Multiple bacteriocins have been characterized from *S. bovis*, and are known

as bovicins, and the most notable are bovicin HC5 and bovicin HJ50 (Whitford *et al.*, 2001; Mantovani and Russell, 2002; Xiao *et al.*, 2004). Bovicin 255 is known to be produced by a rumen strain of *S. bovis* and, in contrast to bovicin HC5 and HJ50, was classified as a class II bacteriocin due to its sequence similarity to lactococcin A and thermophillin A (Whitford *et al.*, 2001). Although *S. bovis* does not have GRAS status, it's bacteriocins do have potential applications within the food industry as bovicin HC5 and HJ50 have been shown to inhibit *Listeria monocytogenes, Clostridium aminophilum* and *Staphylococcus aureus* (Mantovani and Russell, 2002; Pimentel-Filho *et al.*, 2014).

1.3.3.2 Applications of LAB bacteriocins

The GRAS status of many LAB bacteriocins has resulted in a diverse range of applications. Although there are only currently two commercialised bacteriocins, nisin (Nisaplin, Danisco) and pediocin PA1 (Microgard™, ALTA 2431, Quest) other LAB bacteriocins have shown potential across many applications (Vijay Simha *et al.*, 2012). The most researched application is the use of LAB bacteriocins as food preservatives, as they are a natural additive meaning that less chemical preservatives are required (Maurício *et al.*, 2017). They have been studied for use within food and also through incorporation into food packaging as many have demonstrated antimicrobial activities against food pathogens (de Mello *et al.*, 2013; Sánchez-González *et al.*, 2013). Sakacin has been shown to reduce the growth of *L. monocytogenes* in salami maturation (de Souza Barbosa *et al.*, 2015). Lacticin 3147 has also been demonstrated to control *Listeria* and *Bacillus* in some dairy products, however the concentrations required were deemed impractical by the authors (Morgan *et al.*, 2001). One LAB bacteriocin was genetically engineered to create a hybrid of the LAB bacteriocin enterocin CRL35 and the microcin V, which demonstrated effective activity against *L. mononcytogenes*

and *E. coli* O157:57 in skim milk (Acuña *et al.*, 2015). . Another use of LAB bacteriocins is for food product reproducibility as they can be used in cheese and wine to control the quality and flavour of the final products (Oumer *et al.*, 2001; O'Sullivan *et al.*, 2002). LAB bacteriocins have been studied in relation to other applications such as veterinary medicine. The bacteriocins lacticin 3147 and macedoncin ST91KM have both demonstrated activity against bovine mastitis causing pathogens providing both prophylactic and treatment options against the disease (Klostermann *et al.*, 2010; Pieterse *et al.*, 2010). The LAB bacteriocin enterocin AS-48 has recently been shown to be non-toxic in *in vivo* models. This is pioneering work in the bacteriocin field as providing evidence for the safety of bacteriocins is important for moving towards preclinical trials. Although many bacteriocins have demonstrated potential applications in the biomedical field, there is a lack of research demonstrating their safety and efficacy in *in vivo* models which would strengthen the argument for beginning clinical trials (Cebrián *et al.*, 2019).

1.3.4 Bacteriocins Produced by the *Bacillus* Genus

Bacillus spp. are favourable species for industrial use due to their short fermentation cycles, high growth rate and the GRAS status of some species (Schallmey *et al.*, 2004). Bacilli are known to produce different classes of bacteriocins; subtilin is a class one bacteriocin which is a well characterised broad-spectrum lantibiotic produced by *B. subtilis* (Gross, 1973; Sahl and Bierbaum, 1998; Guder *et al.*, 2000). Another lantibiotic produced by *B. subtilis* is mersacidin which is active against methicillin- and vancomycin-resistant *S. aureus* (Chatterjee *et al.*, 1992). Haloduracin and lichenicidin are both two-peptide lantibiotics produced by *Bacillus halodurans* and *Bacillus licheniformis*, respectively. The former, a broad-spectrum bacteriocin, has activity against Gram-positive bacteria whereas lichenicidin has demonstrated activity against

methicillin-resistant *S. aureus* (MRSA) and vancomycin-resistant enterococcus (VRE) strains (Begley *et al.*, 2009). There are also known class II bacteriocins produced by *Bacillus* spp. including coagulin from *Bacillus coagulans*, thuricin-like peptides from *Bacillus thuringiensis*, lichenin from *B. licheniformis*, pumilicins from *Bacillus pumilus* and class III megacins from *Bacillus megaterium*.

1.3.4.1 Applications of Bacillus Bacteriocins

Due to the production of AMPs some *Bacillus* spp. are already on the market as probiotics including *B. licheniformis* which is used in animals as AlCare™, BioPlus® 2B, BioGrow® and Biostart® (Cutting, 2011). The spores of *Bacillus amyloliquefaciens* CECT 5940 are also marketed as Ecobiol® (Norel & Nature Nutrition) for reducing the effect of pathogenic bacteria such as *C. perfringens* and *E. coli* in poultry (Ahmed *et al.*, 2014). There are also potential biomedical applications of *Bacillus* bacteriocins. Some have demonstrated activity against antibiotic resistant organisms such as MRSA and VRE (Chatterjee *et al.*, 1992; Begley *et al.*, 2009) and potential applications need to be further researched There is also potential in using *Bacillus* spp. bacteriocin producing strains as probiotics for human use; *B. licheniformis* is already currently marketed as Biosporin, a licensed medicine in the Ukraine and Russia (Cutting, 2011). *Bacillius polyfermenticus* SCD is another commercial probiotic strain used to treat long term intestinal disorders and is active against *C. perfringens* (Lee *et al.*, 2001). Whereas *Bacillus clausii* strain O/C is utilised for its activity against *S. aureus, Enterococcus faecium* and *C. difficile* (Urdaci *et al.*, 2004).

1.3.5 Bacteriocins as Alternatives to Antibiotics

Currently antimicrobial resistance (AMR) is linked to an approximately 700,000 deaths annually worldwide, and this number has been estimated to increase to 10

million by 2050 (Ghosh *et al.*, 2019). There are many different alternatives being currently explored and bacteriocins are a valid alternative. Bacteriocins have a good reputation, a result of them being generally harmless to humans and the environment due to their sensitivity to proteolytic enzymes (Mills *et al.*, 2011). Many producer strains are already approved for use in food having GRAS status. This means that they can be added to food as bio-preservatives unlike antibiotics, whose use in food is illegal. Bacteriocins are not secondary metabolites, unlike antibiotics, meaning they can be easily genetically manipulated and they do not have undesirable side effects like newer generation antibiotics (Güllüce *et al.*, 2013). There are also commercially viable examples such as Nisin, a broad-spectrum bacteriocin produced by *L. lactis.* Nisin has been an approved preservative for use in foodstuff since 1983 and is used in many countries, including the USA and EU countries, (Guinane *et al.*, 2005; Field *et al.*, 2010) demonstrating the industrial potential of bacteriocins.

Although bacteriocins are a potential reservoir of new antimicrobials there are complications affecting their potential use that still need to be overcome. Different applications will each have their own hurdles to overcome such as storage conditions, shelf-life, chemical interactions and physical treatment (Balciunas *et al.*, 2013; Güllüce *et al.*, 2013). For the medical industry delivery of bacteriocins still needs to be optimised. Conventional drugs have relatively small molecular weights (MW) whereas the MW of bacteriocins range from under 5 kDa to above 60 kDa thus absorbance could be an issue for the larger MW bacteriocins (Cavera *et al.*, 2015). Proteolytic enzymes could also be a complication as protease sensitive bacteriocins could be quickly digested before a medical benefit is produced (Cavera *et al.*, 2015; Ghosh *et al.*, 2019). Possible immunogenicity of bacteriocins is also a potential problem, however safety and efficacy trials could highlight any toxic effects (De Almeida Vaucher *et al.*, 2011). Although there are areas requiring optimisation the proteinous nature of

bacteriocins mean they can be genetically altered to change features such as solubility and mode of action, thus increasing their suitability for specific applications (Güllüce *et al.*, 2013; Rea *et al.*, 2013). Synergisms can also be exploited to increase antimicrobial activity, however this could result in dual-resistant bacteria (Güllüce *et al.*, 2013). Antimicrobial resistance is another problem for bacteriocins as resistance has been previously demonstrated *in vitro,* however the resistance seen is usually spontaneous or induced through altering of the cell membrane or wall. Resistance is expected to be observed *in vivo*, and there is known inherent resistance to some bacteriocins such as nisin (Kramer *et al.*, 2006; Sun *et al.*, 2009; Zhou *et al.*, 2014). However it should be noted that although nisin has been used for decades there is no widespread resistance reported (Ghosh *et al.*, 2019). Close surveillance will be necessary to monitor any emerging resistance to bacteriocins and supervise strict regulations of use, so as to not repeat the same mistakes made with the overuse of antibiotics.

1.4 Bacteriocins Active Against CDI

There are promising bacteriocins with activity against *C. difficile* that are undergoing further research to determine their suitability as CDI treatment options. Avidocin-CDs are genetically manipulated diffocins (*C. difficile* produced bacteriocins) that consists of a scaffold structure and receptor binding protein (RBP) structure. The RBP was designed by genetically mining different *C. difficile* prophages that specifically targeted certain *C. difficile* strains. Through using this approach Av-CD291.2 was produced, which specifically targets the hyper-virulent ribotype 027 with antimicrobial activity against a few other ribotypes (Gebhart *et al.*, 2015). It has been shown to prevent the colonisation by *C. difficile*, after inoculation with spores, whilst having no activity against the normal mouse gut microbiota. This will hopefully reduce the

occurrence of severe CDI by limiting transmission and thus could by administered as a prophylaxis among at risk patients. Through genetically manipulating Avidocin-CDs it was determined that the RBP targeted a major component of the paracrystalline surface (S) – layer of *C. difficile*, SlpA (Kirk *et al.*, 2017). New variants of Aviodocin-CDs were generated to target the different S-layer variants, with intentions to develop a cocktail of Avidocin-CDs. This cocktail would target different *C. difficile* strains and limit resistance. The only resistance observed against Avidocin-CDs was through the loss of SlpA, however this also resulted in a loss of toxin production and virulence within hamsters (Kirk *et al.*, 2017). The highly specific targeting of Avidocin-CDs to *C. difficile*, the generation of a large scale production method and restricting the potential emergence of resistance, all highlight the potential of Avidocin-CDs for successfully clinical applications, generating strong evidence which will support pre-clinical trials (Kirk *et al.*, 2017) .

Another bacteriocin discovered with antimicrobial activity against *C. difficile* is thuricin CD. This bacteriocin was discovered using bioinformatics techniques, and is composed of two peptides, Trnα and Trnβ, which work in a synergistic manner to depolarise the cellular membrane (Mathur *et al.*, 2017). This is a narrow spectrum bacteriocin targeting *C. difficile* and other spore-forming bacteria, and the elucidation of its receptor could highlight a target that is likely to not exist in non-target species (Rea *et al.*, 2010; Mathur *et al.*, 2016, 2017). Thuricin CD also demonstrates activity against *C. difficile* biofilms where the use of vancomycin was ineffective against the majority of strains tested, and in combination with other antibiotics showed highly effective activity against stronger biofilm producers (Mathur *et al.*, 2014). This suggests it would be ideal in the treatment of CDI as it does not affect commensal non-spore forming bacteria, thereby maintaining the integrity of the normal microflora and reducing any potential recurrences (Rea *et al.*, 2010). To date, thuricin CD has not been

tested in clinical trials and its safety and efficacy against CDI is undetermined (Mathur *et al.*, 2016).

Another bacteriocin that demonstrated potential for use as a treatment for CDI was NVB302 (Novacta). This bacteriocin is produced by *Actinoplanes liguriae* NCIMB41362 and has a narrow spectrum of activity inhibiting peptidoglycan biosynthesis with limited effects on Gram-negative bacteria common in the gut microflora. *In vitro* human gut models, NVB302 showed non-inferiority to vancomycin treatment (Crowther *et al.*, 2013) and in 2011 it was the first lantibiotic to undergo phase I clinical trials (Boakes and Dawson, 2014). The results of the trial demonstrated that NVB302 was well tolerated with no serious adverse effects (Febbraro *et al.*, 2012), thus this is a promising treatment for CDI but any further work has currently stalled due to the liquidation of the company.

Although there are potential bacteriocins for CDI treatment already discovered, they are at such an early stage of research that there is still a need for continuation in bacteriocin discovery for CDI treatment. CDI results in such a burden on the health care system that multiple treatments need to be available that will not be detrimental to the health of patients as can be the case with current metronidazole and vancomycin treatments. With the expectation of increasing resistance and the obstacles that must be overcome, discovery and development of active bacteriocins that are commercially viable is a current challenge. This challenge needs to be approached with many potential bacteriocins and therefore continued research in discovery and characterising of bacteriocins against CDI is still relevant and very necessary.

1.5 Objectives of this Study

The overall objectives of this project are to screen environmental bacteria for antimicrobial activity with the aim to identify and characterise novel bacteriocins active against *C. difficile*. The first part of this study aimed to further characterise a strain of *L. lactis* that had previously been isolated from human sewage. The objectives for this strain was to characterise its genome to isolate *in silico* bacteriocin candidates that could have been responsible for the previously observed antimicrobial activity. Once bacteriocin candidates were highlighted the overall objective was to provide evidence linking an *in silico* candidate to the observed antimicrobial activity.

The aim of the second half of this study was to determine if cow manure is a reservoir for putative novel bacteriocins against *C. difficile*. The aim was to develop a screening protocol to identify bacterial candidates potentially harbour novel bacteriocins. Any isolated organism demonstrating antimicrobial activity against *C. difficile* will be characterised to determine if they are a prospective candidate for bacteriocin production, with the aim to discover and investigate the organism itself and the observed antimicrobial activity. The main objective was to identify a strain which was most likely harbouring a novel bacteriocin active against *C. difficile* and carry out an in-depth analysis of this strain aiming to provide evidence of an *in silico* genomic candidate linked to observed antimicrobial activity. With this is mind this study has been split into three sections: the first containing the results pertaining to the previously isolated *L. lactis* strain, the second providing results on the development of the environmental screening protocol and the characterisation of any initial strains of interest. The last section presents an organism, identified through the screening protocol, which underwent in-depth characterisation.

Chapter 2: Materials and Methods

2.1 General Microbiology Techniques

2.1.1 Media

All media used in this study are shown in Table 2 and all ingredients were manufactured by Oxoid™. These were made up with ddH_2O and autoclaved at 121 °C for 15 mins.

2.1.2 Antibiotics

The antibiotics used in this study were all supplied by Sigma-Aldrich® and are shown in Table 3. Stock solutions were all sterile filtered using a 0.22 µM syringe filter.

Table 2- List of Media Used Throughout this Study

	0.5% Soft-top Agar	BHIS Agar	BHIS Broth	LB Agar	LB Broth	Fructose Agar	GM17 Agar	GM17 Broth	GSM17 Agar	GSM17 Broth
C. difficile Agar Base	-	-	-	-	-	27.6 g	-	-	-	-
BHI agar	-	18.8 g	-	-	-	-	-	-	-	-
BHI	7.4 g	-	14.8 g	-	-	-	-	-	-	-
LB	-	-	-	10 g	10 g	-	-	-	-	-
Yeast Extract	1 g	2 g	2.0 g	-	-	2.0 g	-	-	-	-
L-cysteine	0.2 g	0.4 g	0.4 g	-	-	-	-	-	-	-
Bacteriological No.1 Agar	1 g	-	-	-	-	-	-	-	-	-
Bacteriological No.2 Agar	-	-	-	4.8 g	-	-	4.4 g	-	4.4 g	-
Deionised Water	200 ml	400 ml	400 ml	400 ml	400 ml	400 ml	380 ml	380 ml	280 ml	280 ml
Tryptone	-	-	-	-	-	-	2 g	2 g	2 g	2 g
Soya Peptone	-	-	-	-	-	-	2 g	2 g	2 g	2 g
Yeast Extract	-	-	-	-	-	-	1 g	1 g	1 g	1 g
Beef extract	-	-	-	-	-	-	2 g	2 g	2 g	2 g
L-ascorbic acid	-	-	-	-	-	-	0.2 g	0.2 g	0.2 g	0.2 g
β-disodium glycerophosphate	-	-	-	-	-	-	7.6 g	7.6 g	7.6 g	7.6 g
Magnesium Sulphate	-	-	-	-	-	-	0.1 g	0.1g	0.1 g	0.1 g
10% Glucose (after autoclave)							20 ml	20 ml	20 ml	20 ml
2M sucrose (after autoclave)							-	-	100 ml	100 ml

Table 3- List of Antibiotics Used Throughout this Study

Antibiotic	Solvent	Stock Concentration (mg/ml)	Working concentration *E. coli* (µg/ml)	Working concentration *C. difficile* (µg/ml)	Storage temperature (°C)
Kanamycin	Deionised water	50	50	-	-20
Ampicillin	Deionised water	50	100	-	-20
Chloramphenicol	100% Ethanol	25	12.5 (Broth) 25 (Agar)	-	4
Thiamphenicol	50% Ethanol	15	-	15	4
Erythromycin	50% Ethanol	50	500	10	4

2.1.3 Solutions

The buffers used in this study are shown in table 4 and all chemicals were supplied by Sigma-Aldrich® unless stated otherwise.

Table 4- List of Buffers Used Throughout this Study

Buffer	Composition
Protein Purification Binding buffer A1, pH 8.0	20 mM Sodium phosphate 500 mM NaCl 5 mM of imidazole pH with Hydrochloric acid (HCl)
Protein Purification wash buffer A2, pH 8.0	20 mM Sodium phosphate 500 mM NaCl 30 mM of imidazole pH with HCl
Protein Purification Elution Buffer B, pH 8.0	20 mM Sodium phosphate 500 mM NaCl 100 – 750 mM of imidazole pH with HCl
Native Page Running Buffer (1x)	190 mM Glycine 25 mM Trizma base
Native Page Loading Buffer (2x)	0.125M Tris (pH6.8) 20% sucrose 0.004% Bromophenol Blue
Tricine Gel Buffer (3x), pH 8.45	0.3% SDS 3 M Tris 1 M HCl
Cathode Buffer (10x), pH 8.25	1% SDS 1 M Tris 1 M Tricine
Anode Buffer (10x), pH 8.9	1 M Tris 0.225M HCl
Tricine SDS Reducing Buffer	12% SDS 6% β-mercaptoethanol 30% glycerol 0.05% coomassie blue G-250 150mM Tris/HCl (pH 7.0)
TAE Buffer (1x)	40mM Tris 20mM Glacial Acetic Acid 1mM EDTA
Tris Buffer, pH 7.5	50mM Tris 10mM NaCl

2.1.4 Bacterial Strains

All *Clostridioides difficile* strains used in this report, showed in the table below (Table 5), were provided by Synthetic Biology Research Centre (SBRC) which stored the strains in a 10% glycerol BHIS stock kept at -80 °C. All *C. difficile* strains were grown under anaerobic conditions (5% hydrogen, 5% carbon dioxide and 90% nitrogen) at 37 °C using BHIS media supplemented with *Clostridium difficile* selective supplement (CC) (Oxoid). All other strains (Table 5) were grown aerobically at either 37 °C or 30 °C using BHIS or LB media. Overnight cultures were incubated for approximately 16 h, in 5 ml of broth.

Table 5- Complete List of Bacteria Strains Used Throughout this Study

Strain	Description	Original Source
Lactococcus lactis 10	Environmental *Lactococcus lactis*	Human Sewage
Streptococcus 30ANiii	Environmental *Streptococcus* strain	Cow Manure
Bacillus pumilus 1_16	Environmental *Bacillus* stain	Cow Manure
Bacillus thermoamylovorans 8ANE	Environmental *Bacillus* strain	Cow Manure
Baciilus licheniformis 41ANE	Environmental *Bacillus* strain	Cow Manure
Oceanobacillus 1_29A	Environmental *Oceanobacillus* strain	Cow Manure
C. difficile R20291	PCR ribotype 027 (Stoke Mandeville, UK)	Anaerobe Reference Laboratory, Cardiff
C. difficile CD630 (NCTC 13307)	PCR ribotype 012	NCTC, Public Health England
C. difficile DH1916	PCR ribotype 027	Anaerobe Reference Laboratory, Cardiff
C. difficile M120	PCR ribotype 078	SBRC
C. difficile EK28	PCR ribotype 078	SBRC
C. difficile BI-1	PCR ribotype 027	SBRC
C. difficile TL178	PCR ribotype 002	SBRC
C. difficile CF5	PCR ribotype 017	SBRC
C. difficile TL176	PCR ribotype 014	SBRC
C. difficile CD305	PCR ribotype 023	SBRC
C. difficile CD196	PCR ribotype 027	SBRC
C. difficile LIV29	PCR ribotype 001	SBRC

C. difficile LIV22	PCR ribotype 106	SBRC
C. difficile DH478	PCR ribotype 027	Anaerobe Reference Laboratory, Cardiff
C. difficile BI-9	PCR ribotype 001	SBRC
C. difficile M68	PCR ribotype 017	SBRC
C. difficile R20291Δ*pyrE*	*C. difficile* R20291 with an in-frame deletion in *pyrE*	(Ng *et al.*, 2013)
C. difficile 630ΔermΔpyrE	*C. difficile* 630 with an in-frame deletion in *pyrE* and erythromycin sensitivity	(Ng *et al.*, 2013)
E. coli BL21 (DE3)	T7 over expression strain	SBRC
E. coli TOP10	Cloning strain	Invitrogen
E. coli DH5α	Cloning strain	NEB
E. coli C43 (DE3) pLysS	T7 overexpression strain	SBRC
E. coli CA434	Conjugation donor strain	(Purdy *et al.*, 2002)

2.1.5 Plasmids

All plasmids used in this study are shown in table 6 and were extracted using NEB monarch Plasmid extraction kit, and stored at -20 °C. The plasmids were also stored within a host strain such as *E. coli* Top10 or DH5α, which were stored at -80 °C.

Table 6- Complete List of Plasmids Used Throughout this Study

Plasmid	Marker	Description	Origin
pET28b(+)	Kanamycin	*E. coli* T7 expression vector	Invitrogen
pUC19	Ampicillin	Cloning vector	Invitrogen
pET28b(+)*bacE*	Kanamycin	Expression vector for bacE overexpression	This study
pET28b(+)*bacL*	Kanamycin	Expression vector for bacL overexpression	This study
pET28b(+)*bacSP1*	Kanamycin	Expression vector for bacSP1 overexpression	This study
pET28b(+)*bacSP2*	Kanamycin	Expression vector for bacSP2 overexpression	This study
pET28b(+)*bacBP*	Kanamycin	Expression vector for bacBP overexpression	This study
pMTL84123	Chloramphenicol/ Thiamphenicol	Modular vector containing p15a+*tra*, ferrodoxin promoter (P_{fdx}), and pCD6	Heap, *et al.* (2009)

pMTL84123_*nisIFEG*	Chloramphenicol/ Thiamphenicol	pMTL84123 backbone with Nisin immunity genes	This study
pMTL8amβ151	Chloramphenicol/ Thiamphenicol	Modular vector, ColE1+*tra*, with pAMβ origin of replicon	SBRC
pMTL8amβ251	Erythromycin	Modular vector, p15a+*tra*, with pAMβ origin of replicon	This study
pMTL84221	Erythromycin	Modular vector, with pCD6 and p15a+*tra*	Heap, *et al.* (2009)

2.1.5 Oligonucleotide Primers

All oligonucleotide primers were sourced from Sigma-Aldrich® and were designed for this study. The primers were used at a working concentration of 10 µM and are shown in the table below.

Table 7- Complete List of Oligonucleotides Used in this Study

Primer Name	Primer Sequence	Features
27F	AGAGTTTGATCCTGGCTCAG	16s rRNA primer
1492R	TACGGYTACCTTGTTACGACTT	16s rRNA primer
UP-1	GAAGTCATCATGACCGTTCTGCAYGCN GGNGGNAARTTYGA	DNA gyrase B
UP-2r	AGCAGGGTACGGATGTGCGAGCCRTC NACRTCNGCRTCNGTCAT	DNA gyrase B
pedBF	GGTGATTTTATGAATAAGACTAAGTCG	Pediocin primers
pedBR	CCCCTTTATCAGTACTATTGGCTAGGC	
BOV59F	GCATTAACCTCACAATCTG	Bovicin 255 primers
BOV302R	CGATTAGATCTGTTGTTGC	
mcdAF	ATGGAAAAAGAAACAACAAT	Macedocin primers
mcdAR	TTATGAGCAACATGTAGCAA	
suiA F	CAAACTGCAACTGATCAAGAAATTA	Suicin 3908
suiA R	AATTTTTGCACCCAGAGATGAATCC	
sslA F	AAACAACTCAGGAGCTTCAC	Suicin 90-1330
sslA R	CACAGGTCATCAAAATACCC	
sssA F	CTCCTGTAAGATGAACTTGT	Suicin 65
sssA R	CTTTATATGAGTCTGTCTGT	
LactGQ F	GAAAGAATTATCAGAAAAAG	Lactococcin G/Q
LactGQ R	CCACTTATCTTTATTTCCCTCT	
LactABM F	GAAGAGGCAATCAGTAGAG	Lactococcin A, B and M
LactB R	CCAGGATTTTCTTTGATTTACTTC	
LactM R	GTGTACTGGTCTAGCATAAG	
LactA R	GTGTTCTATTTATAGCTAATG	

Lcn972 F	TTGTAGCTCCTGCAGAAGGAACATGG	Lactococcin 972
Lcn927 R	GCCTTAGCTTTGAATTCTTACCAAAAG	
nis F	ATTATAAGGAGGCACTC	Nisin
nis R	CATGTCTGAACTAACAA	
bacE F	CATATGGCAAATGATTGGGGATG	*Nde*I restriction site
bacE R	CTCGAGCCTACTGACCTTATTATTTACC ATTCG	*Xho*I restriction site
bacP F	CATATGCGAAGGAGGTCATATTATGAC	*Nde*I restriction site
bacP R	CTCGAGTTTAGATTAAAAAGCAGCAGC	*Xho*I restriction site
bacP locus F	GAATTCATGACGAAAGCAACAGATTCT AAG	*Eco*RI restriction site
bacP locus R	CTCGAGCAAGAGTTCATGCTCTATTTAT TATTCA	*Xho*I restriction site
bov255 F	CATATGTTCAAATGAATACAAAAACTT	*Nde*I restriction site
bov255 R	CTCGAGCTACGGTTTTCCA	*Xho*I restriction site
bacS F	CATATGTTACATGAATACAAAAACGTTT C	*Nde*I restriction site
M13 F	TGTAAAACGACGGCCAGT	-
M13 R	CAGGAAACAGCTATGACC	-
T7 Terminator R	GCTAGTTATTGCTCAGCGG	-
T7 Promoter F	TAATACGACTCACTATAGGG	-
nisI F	CATATGAGAAAATATTTAATACTTATTG TGGCCTTA	*Nde*I restriction site
nisI R	TTACCTGCATCTAGTTTCCTACCTTCGTT GCA	Homologous overlapping region
nisFEG F	AGGAAACTAGATGCAGGTAAAAATTCA AAATC	Homologous overlapping region
nisFEG R	CTCGAGTTATCTAATCTTTTTTTTAGAT AATGCCACAAG	*XhoI* restriction site
nisI F RT-PCR	GGTGAGGTTTGACGAAGGAA	
nisI R RT-PCR	CTCGGGATACTTTTGCCACT	
nisF F RT-PCR	TTAATCGAGGCTCCAGCAAT	
nisF R RT-PCR	CATTCCTCCATGCACAATCTT	
nisE F RT-PCR	TATTGGTGGGAAGCTGGTTT	
nisE R RT-PCR	CAACTATGGAAAGTCCAACTGGT	
nisG F RT-PCR	TTAGAGCTTTTAACGGTTCCT	
nisG R RT-PCR	ACCCTACGGCAGCTATCAGA	

Restriction enzyme sites are shown in red.

2.1.6 Growth Conditions

Each organism was grown in the most optimal available conditions suited for their species. The different conditions were aerobic, anaerobic and 10% CO_2, at 37 °C or 30 °C. Aerobically cultures were either shaking or non-shaking however under 10%

CO_2 conditions only non-shaking 37 °C incubation was available. The anaerobic conditions used throughout were 80% Nitrogen, 10% CO_2 and 10% Hydrogen, which were maintained in a Don Whitley a95 workstation.

Table 8- Growth Conditions for the Organism Used in this Study

Organism	Media	Conditions	Temperature (°C)
E. coli	LB	Aerobic	37 or 30
Lactococcus lactis	GM17, GSM17, BHIS	10% CO_2	37
Bacillus pumilus	BHIS	Aerobic	37
Clostridioides difficile	BHIS CC	Anaerobic	37
Streptococcus infantarius	BHIS	Anaerobic	37

2.1.7 Growth Curve and Colony Forming Units

For a standard growth curve 50 ml of suitable pre-warmed broth was inoculated to an OD_{600nm} of 0.02, using a 16 h culture. The OD_{600} was then measured every hour until the stationary phase of growth was observed, a further OD_{600nm} reading was recorded at 24 h post inoculation. Triplicate readings from three separate biological replicates were taken for each time point. For each time point a range of dilutions were spotted (3 x 20 µL) onto either BHIS or GM17 agar plates and incubated for 16 h in order to determine the colony forming units per ml (CFU/ml). Dilutions ranged from 10^{-1} – 10^{-11}, and colonies were counted from spots that had between 2-50 colonies to determine the CFU/ml.

2.1.8 Spent supernatant assay

A 16 h culture of the sensitive *C. difficile* strain was used to inoculate 10 ml of fresh BHIS broth to an OD_{600nm} of 0.05. A 16 h culture of the strain of interest was also used to inoculate 10 ml of fresh BHIS to an OD_{600nm} of 0.02 or 0.05 depending on growth rate. Both cultures were incubated according to their optimal conditions (Table 8) until

mid-exponential phase (OD_{600nm} 0.4-0.6), at this point 1 ml of culture was placed on ice and the other 9 ml was centrifuged at 10,000 x *g* at 4 °C and sterile filtered using 0.22 µM filter. In a 96 well plate, wells were set up with spent supernatant from the strain of interest and inoculated with the chilled mid-exponential sensitive *C. difficile* strain to an OD_{600nm} of 0.05. This assay was repeated with three technical and three biological repeats if possible. To determine if there is any inhibition of *C. difficile* in the active strains spent supernatant, a positive control using fresh BHIS was always used. One potential cause of inhibition of growth is due to a lack of nutrients in the spent medium, however this was controlled for by observing normal growth of both strains in their own spent medium compared to fresh BHIS. A negative control of non-inoculated anaerobic and aerobic BHIS, and all spent medium was also used to ensure there was no contamination under either condition. Any aerobic growth was inoculated first and sealed with a plate seal before anaerobic cultures were inoculated within the anaerobic cabinet and a second plate seal was used to maintain anaerobic conditions. A positive inhibition result is shown through the combination of growth patterns shown in Table 9. If any of the controls failed the assay was repeated.

Table 9- The Requirements for a Positive Result of the Inhibition of *C. difficile* in Spent Supernatant

Growth outcome	Strain of Interest Growth	Suspected Sensitive *C. difficile* strain Growth
BHIS	✓	✓
Strain of Interest spent supernatant	✓	×
Own spent supernatant	✓	✓
Negative BHIS control	×	×
Result	*C. difficile* strain is sensitive to spent supernatant of queried strain	

2.1.9 Minimal Inhibitory Concentration

Pre-warmed media was aliquoted into a 24-well plate, with 3 ml of media in each well. A range of antibiotic concentrations were set up, in triplicate, with a positive control of no antibiotic. An 16 h culture was used to inoculate each well to an $OD6_{00nm}$ of 0.05, and readings were taken at 0, 4 or 6, 24 and 48 h, by agitating each well and aliquoting 200 µL into 96 well plate and read using a Glomax® Multi Detection Plate Reader (Promega). Depending on the known growth curve the 2nd time point to confirm normal growth in the positive control was taken at either 4 or 6 h. Three biological replicates were undertaken for each antibiotic concentration, and a non-inoculated control was used to ensure there was no contamination across the 48 h time period. The MIC was determined based on the lowest concentration of antibiotic that had no growth of bacteria after 48 h.

2.1.10 Competent Cells

2.1.10.1 Competent *L. lactis*

A 16 h culture of *L. lactis* was used to inoculate 100 ml of GM17 to an OD_{600nm} of 0.05. Once the cells reached an OD_{600nm} of 0.4-0.6 the cells were pelleted by centrifugation at 10,000 x *g* for 10 mins at 4 °C. All centrifugation steps were carried under these conditions. The supernatant was discarded, and the pellet was resuspended in 32 ml of 100 mM LiAc, 10 mM DTT, 0.6 M sucrose and 10 mM Tris-HCL, pH 7.5, and incubated at room temperature for 30 mins. The pelleted cells were washed using 5 different washes: 8 ml ddH_2O, 4 ml ddH_2O, 4 ml 50 mM EDTA, 4 ml ddH_2O, and 4 ml 0.3 M sucrose. After washing, the cells were resuspended in 1.2 ml 0.3 M sucrose and 10% glycerol. The cells were aliquoted out into 0.05 ml samples and immediately stored at -80 °C until further use.

2.1.10.2 Chemically competent *E. coli*

A single colony of *E. coli* TOP10, BL21(DE3), C43(DE3)pLysS or CA434 was inoculated into 5 ml of LB medium, and incubated for a minimum of 16 h (37°C 200 rpm shaking). 1 ml of this was then used to inoculate 100 ml of LB medium incubated at 37 °C, 200 rpm shaking until an OD_{600nm} of 0.4-0.7. The culture was then chilled on ice for 15 min before pelleting the cells at 3300 x *g* for 10 mins. The cell pellet was then resuspended in 40 ml of ice cold 0.1M $CaCl_2$ and incubated on ice for 30 min. The cells were pelleted using the same conditions as previous described and the cell pellet resuspended in 6 ml of ice cold 0.1 M $CaCl_2$ with 15% glycerol. The competent cells were stored in 50 µl aliquots at -80 ˚C

2.1.11 Transformation Protocols

2.1.11.1 Transformation of Electro-competent *E. coli*

Electro-competent cells (50 µL) were thawed on ice before the addition of ligation mixture (10 µL) and transferring into a chilled electroporation cuvette (0.2 cm gap, Bio-Rad). This was pulsed using a MicroPulser™ Electroporator (Bio-Rad) with pre-set conditions (2.5 kV, 200 Ω, 25 µF capacitance). Cells were immediately resuspended in 400 µL of SOC Medium (ThermoFisher Scientific) and transferred to a 1.5 ml microcentrifuge tube for a minimum of 1 h recovery at 30 °C or 37 °C with shaking (200 rpm).

After recovery, different volumes (10 µL, 100 µL) were spread plated across two agar plates. The remaining cells were pelleted for 30 seconds at 16000 x *g* and the pellet was re-suspended in 100 µL of supernatant and spread onto an agar plate. Plates were incubated at either 30 °C for two days or 37 °C for one day.

2.1.11.2 Transformation of Chemically competent *E. coli*

Chemically competent cells (50 µL) were thawed on ice for 10 mins before 10 µL of ligation mixture was added. The mixture was incubated on ice for 30 mins before heat shocking at 42 °C for 45 s and incubating on ice for a further 2 min. An 800 µL aliquot of room temperature SOC Medium (ThermoFisher Scientific) was added to the cells, which were recovered for at 30 °C or 37 °C with shaking (200 rpm) for a minimum of 1 h. After recovery, cells were spread plated as stated in 2.1.11.1.

2.1.11.3 Transformation of Electro-competent *L. lactis*

Electro-competent cells were thawed on ice before the addition of relevant concentration of plasmid. The mixture was then transferred to a chilled electroporation cuvette (0.2cm gap, Bio-Rad). This was pulsed using a MicroPulser™ Electroporator (Bio-Rad) with pre-set conditions (2.5 kV, 200 Ω, 25 µF capacitance). Cells were immediately resuspended in 2.5 ml of pre-warmed and reduced GM17 media supplemented with 1% sucrose, 20 mM $MgCl_2$ and 2 mM $CaCl_2$. The mixture was then left to recover for a minimum of 3 h recovery at 37°C with 10% CO_2. After recovery 10 µL, 100 µL was plated onto GM17 plates supplemented with antibiotic and 1% sucrose and the rest of the culture was pelleted at 16,000 x *g* and resuspended in 100 µL of its supernatant before spreading. The plates were incubated at 37 °C, 10% CO_2 for 24-48 h.

2.1.12 Conjugation into *C. difficile* Strains

A 5 ml 16 h culture of *E. coli* conjugal donor strain was harvested at 1,500 x *g* for 1 min, before discarding the supernatant and washing the cells 400 µL PBS. The cells were harvested again under the same conditions and the wash was repeated. The cells were pelleted again using the same conditions and the pellet was brought into the anaerobic cabinet. The conjugal donor pellet was resuspended in 200 µL of a 16 h

culture of the recipient *C. difficile* strain. This mixture was spotted onto non-selective BHIS agar plates, in 10 x 20 µL spots which were anaerobically incubated for 8 h, for *C. difficile* CD630, or 24 h, for *C. difficile* R20291. After incubation, any growth was harvested and suspended in 500 µL of anaerobic PBS, and was plated, in 100 µL aliquots onto 5 x BHIS agar plates supplemented with CC and selective antibiotic. The plates were incubated anaerobically for 48-72 h and any growth was re-streaked onto the same BHIS selective agar plates to ensure the growth was *C. difficile* transconjugants. Any potential transconjugants were checked using PCR and any positive strains were stored as glycerol stocks at -80 °C.

2.1.13 Overlays

For all overlays, 0.5% Soft Top Agar was melted and cooled to 50 °C before inoculation with a 16 h culture of *C. difficile* to an OD_{600nm} of 0.05.

2.1.13.1 Colony overlay

Bacteria of interest were grown (spotted or plated) for 12-16 h in its preferred growth conditions and media. 4 ml of soft-top agar was gently poured over the previous growth and incubated for 16 h according to the inoculum's growth conditions (Fig 4c). A zone of no growth of the inoculum around the colony, known as a zone of clearance, is a positive result.

2.1.13.2 Spot plating

For a spot plate or well overlay, 4 ml inoculated melted soft-top agar was poured over a BHIS agar plate and set aerobically before incubating for 1 hour anaerobically. Subsequently a well was made using a 0.8 cm cork borer and a 100 µL aliquot of liquid was added (Fig 4a) or 20 µL was spotted on to the top of the overlay

(Fig 4b). This was incubated for 16 h in the inoculums preferred growth conditions. No growth in the area of the spot, or well indicated a positive result.

2.1.13.3 SDS-page overlay

SDS-PAGE gels (Section 2.3.6.1) were washed with 0.5% Tween 80 (3X 40 mins), prior to rinsing with ddH_2O and transferring to an empty petri dish (100 mm). Inoculated melted soft-top agar (12 ml) was poured over the gel and left to set aerobically before incubating according to the growth requirement of the inoculum for 16 h (Fig 4d).

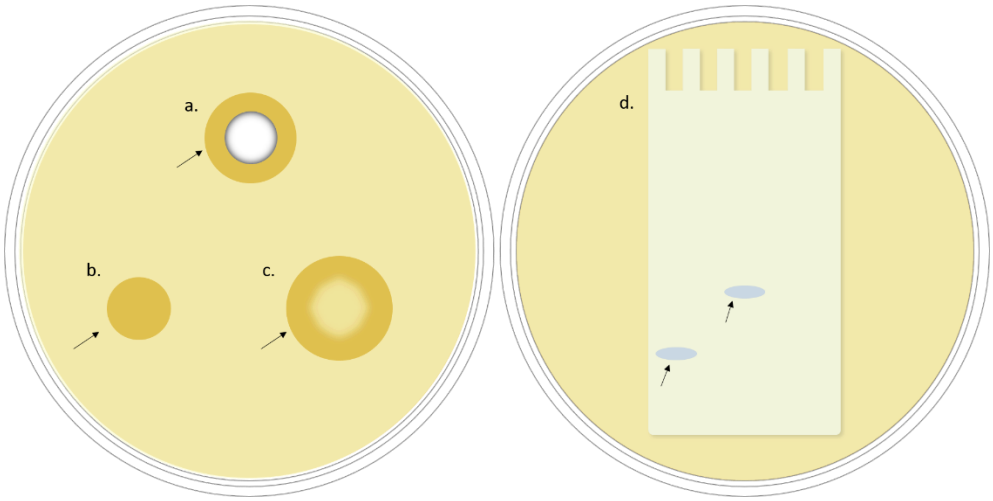


Figure 4- Diagram of the different types of overlays. All zones of clearance, where the sensitive strain cannot grow, are indicated with an arrow. (a). Well overlay, where a hole of agar is bored out of the plate after the overlay and liquid is added into the well, (b) spot overlay where liquid was spotted on to the of the overlay, (c) colony overlay with the colony in the center which has been overlaid after initial growth (d) SDS-page overlay, where the unstained gel is washed and overlaid.

2.1.14 Isolation of Bacteriocin Producing Strains

Bovine manure was homogenized, in Precellys Lysing tubes, with 200 µL PBS. To isolate spore-forming bacteria, 1 ml of 50% ethanol was used instead of PBS. The

mixture was pelleted (5 min at 5000 x *g*) and diluted to 10^{-4}. An aliquot (200 µL) was spread plated onto fructose agar and incubated under anaerobic or aerobic conditions for at least 16 h. Colonies were picked and replica plated onto two fructose agar plates and incubated under previously used conditions for 16 h. One plate was overlaid with *C. difficile* inoculated soft top agar to screen for zones of clearance whilst the other was used for further work on significant colonies.

2.1.15 Gram-Staining

A 10 µL loop was used to spread bacteria onto a glass slide and was fixed onto the slide by passing the slide over a flame. The fixed culture was flooded with crystal violet stain for 1 min before being washed off using ddH_2O. The culture was then flooded with Gram's iodine solution (Merck) for 1 min before washing off against with ddH_2O, and then acetone was added for 5 s before washing off with ddH_2O. The culture was then counterstained with Gram's safranin solution (Merck) for 1 min before being washed off with ddH_2O, and air-dried.

2.2 General Molecular Techniques

2.2.1 Genomic DNA Extraction via Phenol Chloroform

5 ml overnight cultures of the organism were centrifuged at 7,500 x *g* for 10 mins, and the cell pellet resuspended in 180 µl PBS with 10 g/ml lysozyme. This was then incubated at 37 °C for 30 mins, with occasional gentle agitation. 20 µl of RNase A was added and the mixture incubated at room temperature for 30 min. 25µl of proteinase K, 85µl ddH_2O, and 110 µl 10% w/v SDS was added and incubated at 65 °C for 30 min with occasional gentle agitation. The solution was transferred into a phase-lock tube and an equal volume of phenol:chloroform:isoamyl alcohol was added and mixed via inversion, then centrifuged at 16,000 x *g* for 5 min. The top layer was then

transferred into a fresh Eppendorf tube, 20 µl RNase A was added and the samples incubated at room temperature for 30 min. The mixture was then transferred to a fresh phase-lock tube containing phenol:chloroform:isoamyl alcohol, samples were mixed by inversion and centrifuged as described above. The top layer was transferred into a fresh phase lock tube containing phenol:chloroform:isoamyl alcohol and the centrifugation repeated. The top layer was transferred into an Eppendorf tube with a solution of 2x volume of 100% ethanol cooled to -80 °C and 3M NaAc. This was gently inverted before a melted glass pipette tip was used to move the DNA pellet into 1 ml of 70% ethanol. The ethanol was then pipetted off and the pellet allowed to air dry for 45 min, before eluting in 50 µl Elution Buffer. This was then stored at -20 °C.

2.2.2 RNA Extraction

Samples were prepared by harvesting at least 50 ml of culture at required time points through centrifuging at 10,000 x *g* for 10 mins at 4 °C. Cell pellets were resuspended in 2 ml RNAprotect® Cell Reagent (QIAGEN) and incubated at room temperature for 5 mins before being re-pelleted. At this stage the cell pellets can be stored at -80 °C. The cell pellets were resuspended in 1 ml of TRI Reagent® (Sigma-Aldrich) which is sufficient to lyse 5-10 x 10^7 cells. The samples were then passed a 21 gauge needle, x10, and then incubated at room temperature for 5 min. 0.1 ml of 1-bromo-3-chloroproane per 1 ml of Tri Reagent® was added before shaking vigorously for 15 s. After incubation at room temperature for 15 min the mixture was centrifuged at 12,000 x *g* for 15 min at 4 °C. The top colourless aqueous phase was transferred into a fresh RNA-free 1.5 ml tube containing 2-propanol (0.5 ml per 1 ml of Tri Reagent®). The samples were mixed and incubated at room temperature for 10 min before centrifuging at 12,000 x *g* for 10 min at 4 °C. The supernatant for each sample were removed and the pellets washed with 1 ml of 75% ethanol (made up with DEPC treated

water) per 1 ml TRI Reagent®. The samples were vortexed and centrifuged at 7,500 x *g* for 5 min at room temperature. The supernatant was removed, and the wash step repeated. After removing all the supernatant, the pellets were air dried for 10 min at room temperature, before being resuspended in 50 µL RNase free water. All samples were frozen in aliquots at -80 °C to prevent repeated freeze thawing. To check for contamination and degradation 10 µL of RNA sample was run on a 1% agarose gel and 1 µL was used in a control 16s rRNA PCR.

2.2.3 Polymerase Chain Reactions

All polymerase chain reactions were carried out using a Mastercycler® nexus thermocycler (Eppendorf) and using molecular grade nuclease free water (ThermoFisher Scientific). All primers (Section 2.1.5) used in the PCR reaction mix were at a stock concentration of 10 µM. All PCR were visualized on a 1% agarose gel, run for 45 minutes at 100 volts.

2.2.3.1 Colony Polymerase Chain Reactions

Unless specified otherwise, all colony PCRs were carried out with 10 µL DreamTaq Green PCR Master Mix (2x) (ThermoFisher Scientific), 1 µL of forward primer, 1 µL of Reverse primer and 1 µL of colony water. Colony water was generated through picking a colony with a 1 µL loop and resuspending it in 1 µL of nuclease free water. The PCR reaction mix made up to a final volume of 20 µL using nuclease free water. The cycle conditions for this polymerase used were, 95 °C for 10 minutes and then 30 cycles of 95 °C (30 s), annealing temperature for 30 s, 72 °C for extension time and then a final extension at 72 °C for 5 minutes. The annealing temperature for individual primer sets was calculated using the NEB Tm Calculator (https://tmcalculator.neb.com/#!/main), and the extension time was determined via manufactures guidelines of 1 min up to 2 kb and then an additional 1 min per kb.

2.2.3.2 Phusion Polymerase Chain Reaction

Unless specified otherwise, all PCRs were carried out using Phusion® High Fidelity enzyme. Briefly the reaction mix was made up of 4 µL HF Phusion Buffer (5x), 0.4 µL 10mM dNTP (NEB), 0.2 µL of Phusion polymerase, 1 µL of forward primer, 1 µL of reverse primer and ~50-100ng of template DNA. The reaction mix was made up to a final volume of 20 µL using nuclease free water. The cycle conditions for this polymerase used were, 98 °C for 30 s and then 30 cycles of 98 °C (10 s), annealing temperature for 30 s, 72 °C for extension time and then a final extension at 72 °C for 10 minutes. The annealing temperature for individual primer sets was calculated using the NEB Tm Calculator (https://tmcalculator.neb.com/#!/main), and the extension time was determined via manufactures guidelines of 30 s per kb.

2.2.4 Reverse Transcription

Reverse transcription was carried out using Omniscript® Reverse Transcription Kit (QIAGEN), which converts RNA into cDNA using a reverse transcriptase enzyme. Each reaction was made up to a final volume of 20 µL, with 2 µL 10x RT Buffer, 2 µL dNTP Mix, 2 µL of 10 µM Random Hexamer Primer (ThermoFisher Scientific), 1 µL of 10 units/µL RNase inhibitor, 1 µL Omniscript Reverse Transcriptase and 12 µL template RNA. The final mix was incubated at 37°C for 1 h, and was stored at -20 °C.

2.2.5 Plasmid Extraction

Plasmid extraction was carried out using NEB Monarch® Plasmid Miniprep Kit. Briefly, a 5 ml 16 hr culture was pelleted by centrifuging at 16,000 x *g* for 30 s. All centrifugation steps after this were carried out at 16,000 x *g* for 1 min unless otherwise stated. The pellet was resuspended in 200 µl of Plasmid resuspension buffer, before 200 µl of plasmid lysis buffer was added. The mix was inverted 5-6 times and incubated for 1 min at room temperature. 400 µl of Plasmid Neutralization buffer was then added

and mixed by again inverting the sample 5-6 times. The sample was incubated at room temperature for 1 min before centrifuging at 16,000 x *g* for 5 min. The supernatant was transferred into the supplied spin column and centrifuged, before discarding the flow through. The column was then washed with 200 µl plasmid wash buffer 1 before centrifuging and again discarding the flow through. The final wash was with 400 µl plasmid wash buffer 2, which was removed by centrifuging. The column was placed into a fresh 1.5 ml tube and 50 µl of ddH_2O, warmed to 50 °C, was applied to the column and incubated at room temperature for 10 min before centrifuging. The concentration and purity of plasmid DNA was checked using a NanoDrop™ Lite Spectrophotometer (ThermoFisher Scientific).

2.2.6 DNA Purification

DNA requiring purification, such as PCR products, were purified using NEB Monarch® Nucleic Acid Purification Kit. Briefly, a relative ratio (as specified by manufacturers' instructions) of binding buffer was added to the DNA and centrifuged at 16,000 x *g* for 1 min. All centrifugation steps after this were carried out at 16,000 x *g* for 1 min. The column was washed using 200 µl of wash buffer before centrifuging, and this step was repeated once more. After washing the DNA was eluted in 20 µl of ddH_2O water, warmed to 50 °C, and incubated for 5 min at room temperature, prior to centrifugation at 16,000 x *g* for 1 min. The concentration and purity of plasmid DNA was checked using a NanoDrop™ Lite Spectrophotometer (ThermoFisher Scientific).

2.2.7 Gel Electrophoresis

Gel electrophoresis was used to separate and visualize nucleic acids based upon their size. Agarose gels (1% w/v) were prepared using TAE Buffer (1X) and Agarose (ThermoFisher Scientific). Agarose gels were microwaved until the agarose was fully dissolved and were cooled (55 °C) before the addition of SYBR Safe

(Invitrogen), at 1:10000 dilution, which was used to visualize the nucleic acids. Cooled agarose gels were poured into trays with appropriate sample well comb and set at room temperature. The gel was removed from the setting tray and placed into the electrophoresis tank containing TAE buffer (1X). Samples run on the agarose gel were mixed with 6x Gel Loading Dye (NEB), if required, and loaded into the sample wells. 1 kb Plus DNA Ladder (NEB) was run for a size reference. All agarose gels were run at 100 volts for 45 min and visualised using the Molecular Imager® Gel Doc™ XR System using the SYBR™ Safe manufacturers protocol.

2.2.8 DNA Extraction from Agarose Gel

To extract and purify DNA from agarose gels the Monarch DNA Gel Extraction Kit was used following manufacturer's instructions. Briefly, the DNA band of interest was excised from the gel using a scalpel and placed in to a previously weighed 1.5 ml tube. The gel sliced was then weighed and x 4 volumes of Gel Dissolving Buffer was added to the tube. This was incubated at 55 °C for 10 min until the gel slice was completely dissolved. The mixture was then added to the spin column provided and centrifuged for 1 min at 16,000 x *g*, before discarding the flow through. The column was twice with 200 µL of wash buffer which was removed after centrifugation at 16,000 x *g* for 1 min each time. The column was then transferred to a fresh 1.5 ml tube and the DNA was eluted in 20 µl of ddH_2O water, warmed to 50 °C, and incubated for 5 min at room temperature, prior to centrifugation at 16,000 x *g* for 1 min. The concentration and purity of plasmid DNA was checked using a NanoDrop™ Lite Spectrophotometer (ThermoFisher Scientific).

2.2.9 Restriction Enzyme Digestion

All digestions were carried out using the same method. 1-0.5 µg of DNA was digested with 5 µL Cutsmart® Buffer (NEB) and either 1 µL of digestion enzyme for a

single digestion or 0.5 µL was used if it was a double digestion. This mix was made up to a total of 50 µL using nuclease-free water (ThermoFisher Scientific) and incubated between 4-17 h at 37 °C.

Digested DNA was mixed with 10 µL of Gel Loading Dye (6x) (NEB) and visualized on a 1% agarose gel. This mixture was then gel extracted using the Monarch® DNA Gel Extraction Kit previously described (Section 2.2.7). Digested fragments that could not be distinguished from the undigested fragment were purified using the Monarch® PCR & DNA Cleanup Kit described previously (Section 2.2.5).

2.2.10 De-phosphorylation of DNA

Following linearization of plasmids, Antarctic Phosphatase (NEB) was used to remove 5'-phophate group to prevent re-circularization. In a 20 µL DNA digestion reaction, 1 µL (5 units) of Antarctic phosphatase, and 2.3 µL of its buffer (10X) was added and incubated for a minimum of 30 mins at 37 °C

2.2.11 DNA Fragment Ligations

Ligations were carried out using concentration ratio of 3:1 (plasmid to insert). A vector only control was also used (2:0).

The concentration of vector used was always 50-100 ng, and the concentration of insert required was calculated using the New England BioLabs® Ligation Calculator (https://nebiocalculator.neb.com/#!/ligation). Both the vector and insert DNA was mixed with 2 µL of T4 Ligation Buffer, 1 µL of T4 Ligase and nuclease-free water which was added to bring the reaction total to 20 µL. This was either incubated at room temperature for 2 h or kept in ice water for a minimum of 16 h. Ligations were immediately transformed into competent cells (Section 2.1.10) or were stored at -20 °C.

2.2.12 Construction of a Vector for Immunity Gene Expression in *C. difficile*

The vector pMTL85153 was linearized with *Fse*I/*Sbf*I and the Gram-negative replicon and antibiotic marker were digested out of the vector pMTLMW4 also with *Fse*I/*Sbf*I. The fragment was ligated with the backbone to generate pMTL84123. This plasmid contained the low copy Gram-negative replicon p15a and the *C. difficile* replicon pCD6 and P*fdx* promoter.

2.2.13 Construction of Over-Expression Vectors in *E. coli*

The T7 expression vector pET28b(+), contains both a C and N terminal 6x Histidine tag, a thrombin site to cleave the N terminal tag, a T7 tag and multiple cloning site (MCS) for the addition of a gene of interest (Fig 5). All genes of interest in this study were cloned into the MCS using the restriction enzymes *Xho*I and *Nde*I, making sure they were still in frame with the start codon.

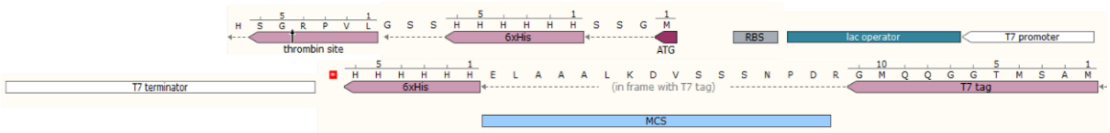


Figure 5- The multiple cloning site and available His-Tags in pET28b(+)

2.2.14 Nucleotide Sequencing

All PCR products and plasmids underwent sanger sequencing performed by Source Bioscience (UK) using provided oligonucleotide primers.

2.3 Protein techniques

2.3.1 Protein Overexpression and His-tag Protein Purification

Overexpression of desired proteins was generated using the T7 expression system. This is an inducible system which is repressed by lacI repressor, which is

present in the pET28b(+) plasmid, and the T7 lysozyme which is present in some *E. coli* expression strains such as C43 (DE3) pLysS. The addition of Isopropyl β-D-1-thiogalactopyranoside (IPTG) inhibits lacI repression and a T7 RNA polymerase is produced which transcribes the gene of interest located downstream of the T7 promoter. The T7 RNA polymerase is in T7 expression strains such as *E. coli* BL21 (DE3), C43 (DE3), Rosetta-gami-pRARE and others strains compatible with the T7 expression system. The level of expression variable across each strain and growth conditions which can be optimised through trialling different conditions such as growth temperature, time of induction initiation, length and temperature of induction and concentration of IPTG. These conditions were optimised for each strain and protein expressed; however, the core induction method remained the same for each. A 16 - 18 h culture of the expression strain was diluted into pre-warmed LB broth containing relevant antibiotics. This was grown shaking, at 200 rpm, at an optimised temperature, and once the required OD_{600nm} was reached IPTG was added and the culture was left shaking at an optimal temperature for a determined amount of time. Where necessary a negative control of the same expression strain carrying an empty vector was used.

2.3.2 Cell Lysis

After protein induction the cells were pelleted at 10,000 x *g* for 10 min at 4 °C. The cells were then lysed immediately or stored at -80 °C after snap freezing in liquid nitrogen. The cell pellets were lysed using two different methods, BugBuster®Master Mix or sonication. BugBuster®Master Mix was used to resuspend a cell pellet, at a ratio of 5 ml/g of wet cell pellet, and was left to shake on a shaking platform at room temperature for 10-20 min before pelleting at 16,000 x *g* for 20 min at 4 °C to remove the insoluble fraction. However, for Nickel resin His-tag purification the soluble fraction caused the resin to turn slightly brown and therefore a different lysis method

was used this purification. Sonication was used instead for cell lysis for nickel resin his-tag purification and for volumes of cell culture over 50 ml. The cell pellet was resuspended in 1:10 binding buffer (Section 2.1.3) and aliquoted into universal tubes which was placed into the chamber with the sterile probe inserted into the liquid and each sample was placed in ice to keep as cold as possible. For a typical sample which was 100 ml of culture, at an OD_{600nm} of 1, a sonication cycle 15 seconds on and 30 seconds off, repeated for 6 min was used. The insoluble fraction was pelleted by centrifuging at 10,000 x *g* for 20 min at 4 °C. The insoluble fraction could be stored at -20 °C and the soluble fraction, temporarily, at 4 °C.

2.3.3 Nickel Resin Protein Purification

The Nickel Resin used was Ni Sepharose™ 6 Fast Flow (GE Life Sciences) which was packed into an empty Pierce™ Spin Column (ThermoFisher Scientific). The resin is in a 20% ethanol mixture which was homogenized by gentle shaking, and 2 ml pipetted out into a 15 ml falcon tube. This was centrifuged at 500 x *g* for 5 min and the ethanol supernatant was pipetted off. The resin was washed in 5 ml water, left to shake gently for 3 min, and then centrifuged again at 500 x *g* for 5 min. The water was discarded, and the wash step was repeated with Buffer A2 (Section 2.1.3) which was again discarded. Fresh Buffer A2 was added to make 50% slurry with the resin which was added to the empty Pierce™ spin column (ThermoFisher Scientific) and left to stand for 30 min before the top layer of buffer was pipetted off.

Cell lysate was loaded onto the column 600 µL at a time. The sample was left to flow through via gravity, and this flow through was collected. The column was then washed twice with buffer A2, and the flow-through was collected each time. The protein was then eluted off the column with 200 µL of Buffer B (Section 2.1.3). The

concentration of imidazole was determined by trialling a range of concentrations from 100 mM – 750 mM to observe when the protein elutes off the resin. All collected flow-through was run on an SDS-PAGE to determine the success of the protein purification.

2.3.4 Zinc Resin Protein Purification

Uncharged resin was used to generate zinc resin, which was then packed into an empty Pierce™ Spin Column (ThermoFisher Scientific). Profinity™ immobilized metal affinity chromatography (IMAC) uncharged resin (BIO-RAD) was chelated with zinc ions through various washing and equilibrating steps. Each wash step included 5 min of gentle shaking and then centrifugation at 500 x *g* for 1 min before the next wash step. The order of washes was, 5x Column Volumes (CV) 50 mM Sodium acetate and 0.3 M NaCl (pH4.0), then 5x CV of 0.1M Zinc Sulphate which was repeated three times. Residual zinc was washed away with 5x CV of 50mM Sodium acetate and 0.3 M NaCl (pH 4.0), and 10x CV ddH_2O water before the resin was equilibrated with 5x CV of buffer A1 (Section 2.1.3). This was also centrifuged and discarded, before the resin resuspended in enough buffer A1 to make a 50% slurry which transferred to an empty Pierce™ spin column and left to stand for 10 mins. The separated top layer was discarded, and the resin was ready for protein purification.

Cell lysate was added to the column, 1x CV at a time, and pipetted up and down 5 times and the 1st CV would be left to stand for 5 mins before spinning at 1000 x *g* for 1 min. This procedure was repeated for the wash buffers as well, and all flow-through was collected to be run on an SDS-PAGE. The column was washed with 3x CV of buffer A1 and 3x CV of buffer A2 before any tagged protein was eluted off using 4x CV of buffer B (Section 2.1.3). The concentration of imidazole in buffer B was determined by trialling a range of concentrations from 100 mM – 750 mM during the first purification.

2.3.5 Protein Precipitation Techniques

Protein precipitation from supernatant was carried out using different methods to determine the best method to obtain proteins whilst maintaining antimicrobial activity. The methods trialled ranged from harsher methods such as solvent precipitation and adsorption/desorption to more gentle methods such as ammonium sulphate precipitation.

2.3.5.1 Solvent Protein Precipitation

Solvent precipitation uses different ratios of specific solvents to supernatant and causes the proteins to precipitate after incubation with the solvent typically overnight. Two different solvents were used, acetone and ethanol at different ratios of solvent to sample volume; acetone was used at 4:1 ratio and ethanol at a 9:1 ratio. The solvent was cooled to -20 °C, before added to sterile filtered supernatant. The supernatant was acquired from cultures grown to late exponential/early stationary phase which were centrifuged at 10,000 x *g* and sterile filtered using 0.22 µM filter. The solvent supernatant mix was vortexed to mix and incubated at -20 °C for a minimum of 60 min to 12 h. After incubation the precipitated proteins were pelleted by centrifuging for 10 min at 4 °C at 15,000 x *g*. The pelleted proteins were washed with the same solvent at a volume 1/10th of the original sample volume. The proteins were pelleted again for 10 min at 4 °C at 15,000 x *g*, and the solvent was discarded, and the pellet left to air dry to evaporate any residual solvent for 30 mins. The proteins were then resuspended in Tris buffer, pH 7.5.

2.3.5.2 Adsorption/Desorption Protein Precipitation

The adsorption/desorption protein precipitation method relies on proteins being absorbed onto the cell membrane at a slightly acidic pH such as 5 or 6, and then desorbed when exposed to extremely acidic conditions such as pH 1. Cultures are grown to a late exponential/early stationary phase and the pH was adjusted to either

6.5 or 5.5 with HCl. The cells and any proteins were pelleted at 10,000 x *g* for 15 min at 4 °C before being washed twice in 5 ml of 5 mM phosphate buffer at the same pH as the sample was previously adjusted to. The cells were harvested again using the previously stated conditions and resuspended in 1 ml of 100 mM NaCl pH 1.1, 1.5, 2.0 or 2.5. This was incubated, gently shaking, for 1 h at 4 °C and then centrifuged at 20,000 x *g* for 20 min at 4 °C, and the supernatant was dialysed into Tris Buffer (pH7.5) and the pelleted cells resuspended in 5 mM phosphate buffer pH 6.5 to check for residual activity.

2.3.5.3 Ammonium Sulphate Precipitation

Cultures were grown to a specified OD_{600nm}, and cells were pelleted at 10,000 x *g* and the supernatant sterile filtered using 0.22 μM filter. The supernatant was then altered to pH 5 using HCl for consistency, and cooled to 4 °C. To determine the percentage of ammonium sulphate required, solid ammonium sulphate was added in 20% increments from 20% - 80%. Once the ammonium sulphate was fully dissolved at each percentage it was left to precipitate, gently shaking, for 1 h. The precipitated proteins were then pelleted through centrifuging for a 4 °C 30 min at 20,000 x *g*. The protein pellet was then resuspended in Tris buffer (pH7.5) and dialysed against the same buffer but 10x the resuspension volume. The dialysis buffer was changed to fresh buffer after 2 h and left for at least approximately 16 h. A control of sterile BHIS medium was used to ensure the removal of all ammonium sulphate, at the dialysis stage, as any residual ammonium sulphate would result in a false positive and could interact with proteins running on SDS-pages gels.

2.3.6 Protein Gels

Protein gels were all carried out using cell lysate or purified protein. These methods were previously described 2.11 and 2.12 respectively. All protein gels were visualised and imaged using the Molecular Imager® Gel Doc™ XR System.

2.3.6.1 SDS-PAGE Gels

All SDS-PAGE gels were either 10% or 4 - 16% Invitrogen™ NuPage™ Bis-Tris gels (ThermoFisher Scientific) and were run using NuPage® MES SDS Running Buffer. 12 µL of sample was mixed with 3 µL NuPage® LDS Sample Buffer (4x) and 0.375 µL of 2M DTT. This mixture was heated to 70 °C for 10 min and 12 µL was loaded on to the gel. 10 µL of SeeBlue® Pre-stained Protein standard (New England BioLabs) or ColourPlus™ Prestained Broad Range (ThermoFisher Scientific) Protein Ladder was also applied to the gel. Gels were run under the following conditions 120 volts for approximately 90 min. Proteins were visualized using coomassie blue staining unless otherwise stated. The gels were stained with 0.1% Coomassie blue for 1 h and then de-stained for 1 h with a mixture of 50% water, 40% methanol and 10% acetic acid. Following de-staining the gels were incubated overnight in water to allow further de-staining.

2.3.6.2 Native-PAGE Gels

Native-PAGE gels were prepared using the BIO-RAD Mini-PROTEAN II Cell electrophoresis system and all the chemical were provided by Sigma-Aldrich. These gels were made with a resolving gel at either 12.5% or 10%. The resolving gel and stacking gel were prepared first without Ammonium Persulfate (AP) (Sigma-Aldrich) and TEMED (Sigma-Aldrich) (Table 10). Glass plates for the gel were assembled into the cassette, with separators between one large and one small plate which was placed into the cast correctly using the assembly stand. Once the cassettes were assembled and the gels ready to be poured the AP and TEMED were added into the resolving gel

mix and poured between the plates up to 4/5th of the total volume. Then 2-3 mm of water saturated butan-2-ol was poured on top ensure an even surface level. Once the resolving layer had set the saturated butan-2-ol was washed away with water and the well comb was then placed between the plates. The AP and TEMED was added into the stacking gel mix and poured in-between the plates until it reached the top of the comb. Once set, the comb was removed and the wells washed with water. In order to run samples on the gel, two gels were assembled into the BIO-RAD Mini-PROTEAN II Cell gel tank ensuring that the middle chamber was water-tight. The middle chamber between the two cassettes was filled with running buffer and the rest of the tank up was filled halfway with running buffer. Gels were run at 150 V for approximately 90mins, to ensure the tank did not overheat. Once run the gel was both stained and de-stained as previously described (Section 2.3.6.1 or Section 2.3.6.3), unless the gel was overlaid with 12 ml of 0.5% soft-top agar inoculated with 400 µL of *C. difficile* and left in the anaerobic cabinet overnight (Section 2.1.13.3).

2.3.6.3 Silver Staining

Silver staining was used on protein gels with proteins either in too low a concentration to be visualised with coomassie blue staining. Silver staining was carried out using Pierce Silver Staining Kit (ThermoFisher Scientific) according to manufactures conditions. Briefly, the gel was washed twice, for 5 mins each, in ultrapure water before the being fixed with two washes for 15 min in a mixture of 30% ethanol, 10% acetic acid and 60% ultrapure water. The gel was then washed twice, for 5 min each, in 10% ethanol before removing two more 5 min washes in ultrapure water to remove the ethanol. Next the gel was then exposed for 1 min to the Sensitizer Working Solution which was then washed away with two 1 min washes in ultrapure water. The gel was then stained for 30 min in the Stain Working Solution before being rapidly washed

twice with ultrapure water for 20 s each. The stain was then developed for 2 - 3 min using the Developer Working Solution and once successful development was stopped with a 10 min wash in 5% acetic acid. The gel could be temporally stored in water, before being imaged using the Molecular Imager® Gel Doc™ XR System.

2.3.6.4 Western Blotting

Western blotting was carried out using the Trans-Blot® Turbo™ Transfer System (BIO-RAD) using the Trans-Blot® Turbo™ Mini PVDF Transfer Blotting packs. The SDS-PAGE gel was placed between the top and bottom parts of the pre-assembled blotting paper packs, in the provided cassette, ensuring there were no bubbles. The lid of the cassette was tightly secured, and the proteins were transferred using the 'LOW MW' program which ran for 4 min at 25 V and 1.3 Amps. The blot was disassembled and the Whatman paper was discarded. The membrane blocked in TBS-T buffer plus 5% skimmed milk powder for 1 h at room temperature shaking. The blocking buffer was discarded, and the membrane was washed in TBS-T twice to remove any residual blocking buffer. The membrane was then incubated with 10 ml of TBS-T plus a 1:5,000 dilution of Anti-6X His-tag antibody for either 2 h at room temperature or 16 hrs at 4°C gently shaking. The antibody solution was then discarded before the membrane was washed twice in TBS-T for 5 min each time and the once for 5 min in TBS. Any bound anti-body was detected using a colour metric or chemiluminescence method. The colorimetric method was using 1 ml of 3,3',5,5'-Tetramethylbenzidine (TMB) to visualise the bands within 1- 5 min. The chemiluminescence detection used Western Lightening™ Plus-ECL (PerkinElmer), in a 1:1 ratio of enhanced luminol reagent and oxidizing reagent. 1 ml of the mix was incubated with the membrane for 1 min with gentle agitation before draining the liquid off and wrapping the membrane in cling-film. A Vilber Fusion FX imager was used to visualise any protein bands bound by the antibody.

Table 10- The Chemicals Used to Make Up the Stacking and Resolving Gel in the Native-PAGE gels, and the Protocol to Make the Gels.

	Stacking Gel	Resolving Gel	
		12.5%	10%
30% Acryl-bis	750 µl	4.2 ml	3.3 ml
3.0 M Tris-HCL (pH 8.8)	-	1.25 ml	1.25 ml
0.5 M Tris-HCL (pH 6.8)	1.5 ml	-	-
10% AP	150 µl	125 µL	125 µL
ddH_2O	3 ml	4.0 ml	4.8 ml
TEMED	6 µl	7.5 µL	7.5 µL

2.4 Bioinformatic Tools

2.4.1 Genome Assembly and Analysis Tools

2.4.1.1 Genomic Assembly Pipline

Genomic DNA extracted via the phenol-chloroform method (Section 2.2.1) was sent to MicrobesNG for Illumina HiSeq 2500 platform. The sequencing resulted in 2x250 bp pair-end sequence reads with a minimum coverage of 30X.

2.4.1.2 Quality Control of Sequencing Reads

After sequencing, MicrobesNG used a quality control pipeline to trim any low-quality bases and to remove adaptor sequences. The quality of the pair-end reads was confirmed using FastQC (Wingett and Andrews, 2018). Any low-quality sequences were flagged at this stage and determined if their quality could be improved for example a failure of 'Per base sequence quality' flags the need for further trimming of the reads to increase the quality. This software was also used to generate overall statistics of the reads such as GC% and total sequence length, which should correspond accordingly to the known data of the species.

2.4.1.3 Genome Assembly

The genomic assembler MIRA 4 (Chevreux *et al.*, 2004), was used to turn the quality checked Illumina HiSeq paired reads into contigs. The settings for MIRA were, accurate de novo mode '`job = genome,denovo,accurate`' defining the reads as pair ends '`readgroup = DataIlluminaPariedLib`' with '`autopairing`' stated which results in automatically pairing the reads without declaring the read orientation. The type of sequencing read is also defined using '`technology = solexa`'. The one parameter used was '`Parameter = SOLEXA_SETTINGS -CL quality_clip=on`' which allows MIRA to clip the reads if it determines the quality is too low. The output file '`projectname_LargeContigs.unpadded.fasta.qual`' was used for further downstream use, as this removes very small contigs and only contains large contigs which is more reliable for de novo assembles.

2.4.1.4 Contig Coverage

To determine the average number of reads covering a contig (contig coverage) the sequencing reads were mapped back to the assembled contigs using the software BWA (Li and Durbin, 2009). An index was created using the '`bwa index`' and the option '`-a is`' invoked an IS linear-time algorithm. The inputs were the MIRA 4 (Chevreux *et al.*, 2004) assembly output file and both sequencing read files. After this the function '`bwa mem`' was used with the number of threads option, '`-t`', specified to 4 and the output file specified to be in the SAM file format. This SAM file was then converted into a BAM file format using the software SAMtools (Li *et al.*, 2009b). This was done using the functions '`view`' option '`-b`' (bam file), '`sort`' specifying the format as '`out.s`', and index it as '`out.s.bam`'. This file was then opened using Qualimap (Garcia-Alcalde *et al.*,

2012), File>New Analysis>BAM QC. From here the contig coverage can be seen in the 'Chromosome stats' section, which displayed all the reads have been successfully mapped.

2.4.1.5 Genome Alignment

Aligning a genome to another assembled genome results in aligning of contigs, this allows for higher quality analysis as it provides a more informed genome structure. The top 3 largest contigs were uploaded to BLAST (Altschul *et al.*, 1990) to search for a reference strain using blastN. The top 5 closest full genome matches, in terms of the largest Query Coverage, Identity percentage and e-value for each were tallied and the 2 most frequent were used as a reference genome for contig alignment in Mauve (Darling *et al.*, 2004).

In Mauve (Darling *et al.*, 2004) the option Tools>Move Contigs was used to align the assembled contigs.fasta file to the reference genome. The contigs.fasta file in the highest number alignment output folder is the ordered contigs file.

2.4.1.6 Genome Annotation

Annotation of genes allows for easy analysis of the potential genes carried by the organism. The ordered contig.file from Mauve was uploaded to the Rapid Annotation using Subsystem Technology (RAST) server (Aziz *et al.*, 2008; Overbeek *et al.*, 2014). The taxonomy of the uploaded file was specified using the Taxonomic ID number. After annotation two file formats were downloaded, these were the amino acid file format and the merged Genbank file format, which can be opened in Artemis.

2.4.1.7 Identification of Bacteriocin Clusters and Prophage Regions within Bacterial Genomes

The identification of putative and known bacteriocins clusters within bacterial genomes was determined using BAGEL4 (van Heel *et al.*, 2018) and Bactibase (Hammami *et al.*, 2010), online bacteriocin mining tools. Another online tool, PHASTER (Arndt *et al.*, 2017), was used to identify the existence of prophage regions.

2.4.1.8 Phylogenetic Analysis of 16s rRNA sequences

Phylogenetic tress analysing 16s rRNA sequences was created using MEGA7 (Kumar *et al.*, 2016). The 16s rRNA NCBI database was mined for sequences with similarity to those taken from the whole genome sequences of isolated environmental strains. A pairwise comparison of the nucleotides sequences was generated using ClustalW and a phylogenetic analysis was carried out using the Maximum Likelihood method based on the Tamura-Nei model (Tamura and Nei, 1993). The first tree underwent bootstapping 500 times in order to generate the final phylogenetic tree.

2.4.2 BLAST

Mining of the nucleotide and sequence databases was performed using the Basic Local Alignment Search Tool (BLAST), using the nucleotide BLASTn and protein BLASTp algorithms (Altschul *et al.*, 1990) accessible at https://blast.ncbi.nlm.nih.gov/Blast.cgi.

2.4.3 Plasmid and PCR DNA Fragment Sequencing

Plasmid DNA and purified DNA fragments amplified from PCRs were sequenced using Source Bioscience Sanger Sequencing Service. The sequencing data was analysed using SnapGene® Viewer and any alignments of nucleotide and amino acid sequences was carried out using Clustal Omega (Sievers and Higgins, 2014) multiple sequence alignment tool.

2.4.4 Visualising Sequencing Data

All sequence data was viewed using both Artemis and SnapGene® Viewer software.

2.4.5 Plasmid Maps

SnapGene® Viewer was used to design and edit plasmid maps.

2.4.6 Oligonucleotide Design

Oligonucleotides used for PCR primers were manually designed using SnapGene® Viewer software. The primers were analysed for secondary structures, melting temperatures and self-dimerisation using Primer3Plus, an online tool accessible at http://www.bioinformatics.nl/cgi-bin/primer3plus/primer3plus.cgi.

2.4.7 Promoter and Terminator Sequence Identification

For the identification of putative promoter sequences within nucleotide sequences the software Softberry BPROM (Solovyev and Salamov, 2011) was used, which is accessible at http://www.softberry.com/berry.phtml?topic=bprom&group=programs&subgroup=gfindb.

To identify putative rho-independent terminator sequences within nucleotide sequences the software ARNold (Gautheret and Lambert, 2001; Macke *et al.*, 2001) was used, which is accessible at http://rssf.i2bc.paris-saclay.fr/toolbox/arnold/index.php.

2.4.8 Data Analysis

Generated data was visualised using Microsoft Excel and analysis and graphical representation of this data performed using both Microsoft Excel and GraphPad Prism.

Chapter 3: Characterisation of a *Lactococcus lactis* Strain with Antimicrobial Activity Against *C. difficile*

3.1 Introduction

3.1.1 Aims of this Project

Bacteriocins are hypothesised to be carried by 99% of all bacteria (Klaenhammer, 1988) and environmental strains are a potential source of novel bacteriocins. A *L. lactis* strain previously isolated from human sewage displayed antimicrobial activity against *C. difficile*, which is hypothesized to be due to a bacteriocin. This study aims to characterise this environmental strain, confirm the previously observed antimicrobial activity and elucidate the cause of this activity through *in silico* and *in vitro* analysis.

3.1.2 *L. lactis* Background

L. lactis, formerly *Streptococcus lactis*, is a Gram-positive, non-sporulating, facultatively anaerobic bacterium. There are three different subspecies of *L. lactis*, (i) *L. lactis* subsp *lactis*, (ii) *L. lactis* subsp *cremoris* and (iii) *L. lactis* subsp *hordniae*. *L. lactis* has GRAS status due to its historic use within the dairy industry where it has been used for centuries to ferment foods such as cheese and yogurt (Song *et al.*, 2017). *L. lactis* has become the model LAB strain for genetic modification widening its possible applications due to its small genome and abundance of genetic engineering tools (Song *et al.*, 2017). Its uses in the dairy industry include milk acidification, enhancing flavour via peptide and amino acid production, ripening, and preventing the growth of pathogenic bacteria or bacteria that induce food spoilage (Fernández *et al.*, 2011; Song *et al.*, 2017). Strains harbouring bacteriocins are useful for increasing the shelf-life of food products, and *L. lactis* strains have been shown to carry a range of bacteriocins including broad and very narrow spectrum bacteriocins. The main broad-spectrum

bacteriocins found in *L. lactis* strains are nisin, lacticin 3147, and lacticin 481 which are all class I lantibiotics. The narrow spectrum bacteriocins are a variety of lactococcins which are class II bacteriocins (Guinane *et al.*, 2005).

3.1.3 Bacteriocins of *L. lactis*

3.1.3.1 Lacticin 481

The lactococcus bacteriocin, lacticin, is produced by strain *L. lactis* subsp. *lactis* DPC5552. It has broad-spectrum activity which is limited to mostly LAB and *Clostridium tyrobutyricum*. It is used within the food industry to increase the shelf-life and standardise the flavour of some cheese. The lysis of sensitive cells reduces the rate limiting proteolytic step in some cheese maturation and also releases enzymes from the starter culture strains which improves flavour and quality (O'Sullivan *et al.*, 2002; Guinane *et al.*, 2005).

3.1.3.2 Lacticin 3147

Lacticin 3147 is a broad-spectrum two-component lantibiotic first isolated in 1996 (Ryan *et al.*, 1996). It is a pore-forming bacteriocin which causes cell death through undermining the integrity of the cell membrane. It is active against a range of Gram-positive bacteria including the pathogenic strains *Listeria monocytogenes* and *Enterococcus faecalis* (Ryan *et al.*, 1996; McAuliffe *et al.*, 1998). Although it was isolated from *L. lactis* DPC3147 this strain was associated with an off flavour. Therefore the genes, occurring naturally on a plasmid, were conjugated into more advantageous strains for cheese production (Guinane *et al.*, 2005). Another advantage of this plasmid is that it encodes bacteriophage resistance; phage attacks are a significant problem in the dairy industry (Coakley *et al.*, 1997).

3.1.3.3 Lactococcins

Lactococcins are lactococcus-specific bacteriocins produced by *L. lactis*. Two classes are known, class IId (lactococcin A, B, M and Z) and two-peptide class IIb

(lactococcin Q and G) lactococcins (Holo *et al.*, 1991; van Belkum *et al.*, 1992; Moll *et al.*, 1996; Zendo *et al.*, 2006; Diep *et al.*, 2007; Ishibashi *et al.*, 2015). Lactococcins A, B and M are encoded by separate genes on the same plasmid carried by *L. lactis* subsp. *cremoris* and show specific narrow spectrum activity only against other *L. lactis* strains (Holo *et al.*, 1991; Venema *et al.*, 1993). Lactococcin A and Z share significant sequence homology and use subunits from the mannose phosphotransferase system (Man-PTS) as a receptor on target cells. Lactococcin B is also known to utilise Man-PTS as a receptor however it lacks homology to lactococcin A and Z (Venema *et al.*, 1993; Diep *et al.*, 2007; Ishibashi *et al.*, 2015; Daba *et al.*, 2018). Lactococcin A, B, G, and Q all cause bactericidal activity through permeabilising the cell membrane causing leakage of intracellular substrates or ions (Holo *et al.*, 1991; Venema *et al.*, 1993). Uncharacteristically, and in spite of its homology to lactococcin A, lactococcin Z is thought to inhibit cell growth and been shown to be a non-pore forming bacteriocin (Daba *et al.*, 2018).

3.1.3.4 Nisin

3.1.3.4.1 History of Nisin and Food Applications

In 1928, the inhibitory effects of nisin were discovered in *Streptococcus lactis* (Rogers and Whittier, 1928). The name was devised from the previous Lancefield serological grouping of *L. lactis*, Group **N** *Streptococcus* **I**nhibitory **S**ubstance **IN (NISIN)**, by Mattick and Hirsch (1947). Nisin is a class Ia lantibiotic with, to date, eight established variants, four produced by *L. lactis* (A, Z, F, and Q) and four by *Streptococcus* strains (U, U2, P, and H). Nisin A was the first nisin to be characterised and is highly similar to the other lactococcal nisin peptides for example nisin Z and nisin A differ by only one amino acid, His27Asn (Mulders *et al.*, 1991). Nisin F has the same amino acid substitution plus Val30Ile (de Kwaadsteniet *et al.*, 2008), and nisin Q incorporates the previous two substitutions and Val15Ala and Leu21Met (Zendo *et al.*,

2003; Fukao *et al.*, 2008). The streptococcal variants U, U2 and P vary more from original nisin A. The most homologous streptococcal variant to nisin A is nisin H which is similar in molecular weight but contains amino acids changes observed in the *streptococcal* variants (Wirawan *et al.*, 2006; Zhang *et al.*, 2012; O'Connor *et al.*, 2015). Another more distinct variant isolated from *Blautai obeum* is nisin O (Hatziioanou *et al.*, 2017).

Nisin A is commercially available as Nisaplin™ (Danisco) and been used as a food preservative for over 60 years since its international acceptance as food additive E234 by the Joint Food and Agriculture Organisation/World Health Organisation (FAO/WHO) (WHO, 1969; Chen and Hoover, 2003; Khan and Oh, 2016). Nisin is used in a wide range of goods including, canned foods, baby food, meats, and dairy products. It is extensively used due to its broad spectrum activity against Gram-positive bacteria including pathogens such as *L. monocytogenes*, *Staphylococcus aureus*, *Mycobacterium* spp., *Bacillus* spp. and *Clostridium* spp. (Chen and Hoover, 2003). Gram-negative bacteria can be made sensitive by permeabilising the outer membrane for example by using chelating agents (Chen and Hoover, 2003).

Nisin is a fitting preservative because of its broad spectrum activity, high activity, protease sensitivity and its ability to survive the food pasteurisation and sterilisation processes due to heat stability (Santos *et al.*, 2018). An early use of nisin was the replacement of nitrates as a prevention method for the outgrowth of Clostridia spores in cheese and cheese spreads (Rayman *et al.*, 1981). By products of nitrates are known to be carcinogenic (Tricker and Preussmann, 1991); thus a natural preservatives, like nisin, are much more favourable; nevertheless, there are some limitations of nisin. Its amphipathic nature means it can interact with fat and some meat components, limiting its activity (Taylor *et al.*, 2007), thus it has been found to

be more effective in dairy products than meat (Deegan *et al.*, 2006). Although it is highly active in the acidic pH range, it loses activity above pH 7 limiting the range of food it can be added to. Studies have tried addressed these issues for example the incorporation of the peptide within packaging increases the length of activity due to the slow release of peptide (Chen and Hoover, 2003; Silva *et al.*, 2018). Its broad spectrum activity is another limiting factor as it cannot be used in food which requires LAB for fermentation processes (Silva *et al.*, 2018). Despite these limitations it is still widely used in the food industry.

Although nisin is currently only being used as a food bio-preservative it has the potential to be used in other industries such as the cosmetic. The use of nisin would lower the dose of chemical preservatives used, which are known to be associated with sensitivity and allergic reactions (Maurício *et al.*, 2017). Additionally, nisin has a wide ranging scope of potential applications within the biomedical industry ranging from anticancer, oral health, immunomodulatory, to antimicrobial applications (Shin *et al.*, 2016). There have been developments in the generation of genetic variants of nisin that enhance its properties such as increased activity against Gram-negatives, increased range of pH stability, and enhancing specificity and/or activity (Yuan *et al.*, 2004; Field *et al.*, 2012). Shin *et al* (2016), concluded that nisin could be an extremely useful therapeutic agent that needed to be validated with further *in vivo* studies.

3.1.3.4.2 Gene Organisation and Protein Structure

The production of nisin is coordinated by 11 different genes, in four different operons (Fig 7). The structural precursor peptide nisin A is a 57 residue peptide encoded by *nisA*. This peptide undergoes extensive modification to generate the mature active peptide (Buchman *et al.*, 1988). The nisin Z variant, encoded by *nisZ*, is structurally identical to nisin A and undergoes the same biosynthesis process (Mulders *et al.*, 1991). The biosynthesis pathway utilises four different genes, *nisC, nisB, nisT* and

nisP, which generate the final mature peptide (Fig 6). Firstly, the serine and threonine residues are dehydrated by a putative enzyme, NisB, into dehydroalanine and dehydrobutyrine, respectively (Lubelski *et al.*, 2008). NisC catalyses the reaction of these dehydrated amino acids with free cysteines, resulting in five thioether bridges called (β-methyl) lanthionines which are characteristic of lantibiotics (Fig 6)(Kuipers *et al.*, 2004). The leader peptide signals the transportation of the precursor nisin across the membrane using a dedicated, yet broad spectrum, transporter, NisT. NisT is only an ABC half-transporter, as it contains only two out of typical four modules, thus it is likely to form a homodimer in order to become an active unit (Lubelski *et al.*, 2008). To generate the active nisin molecule the leader peptide must be removed after transportation. NisP is thought to be anchored to the cell surface and is a subtilisin-like protease that proteolytically processes pre-nisin into the mature active peptide (van der Meer *et al.*, 1993; Schneewind *et al.*, 1995; Qiao *et al.*, 1996).

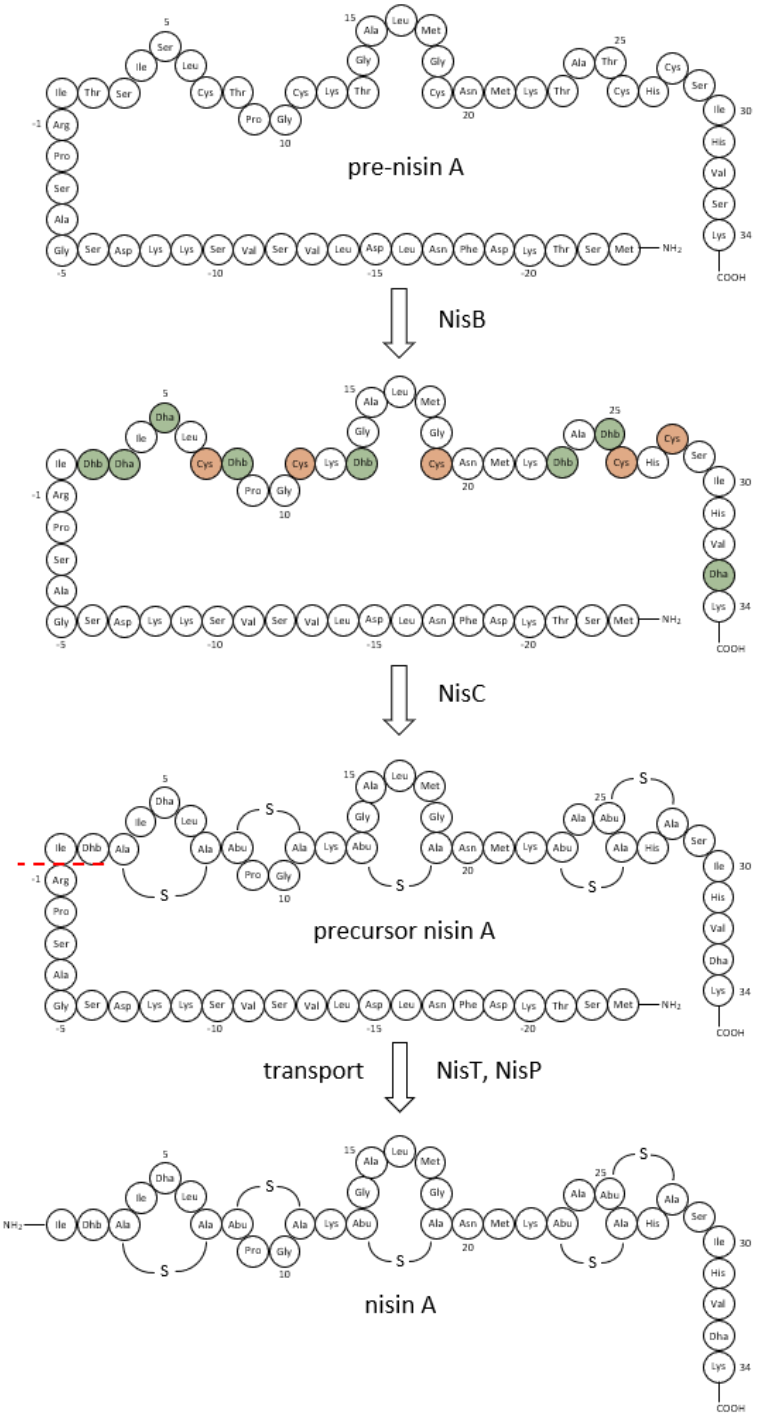


Figure 6- Processing of pre-nisin A to the mature nisin. Figure modified from Kuipers *et al.* (1995). Initially serine and threonine residues in pre-nisin A are dehydrated (green circles) by NisB. NisC then generates five thioether bridges (depicted by the arches) using the free cysteine residues (red circles) generating the precursor nisin A peptide. The precursor nisin A is transported across the cell membrane by NisT and NisP cleaves off the signal sequence (indicated by the red dotted line) resulting in the mature active nisin A peptide.

To protect the host from its own bacteriocins they can possess self-immunity genes. Two separate immunity operons, *nisI* and *nisFEG*, are present in the nisin biosynthesis pathway (Fig 7)(Siegers and Entian, 1995). These are two separate immunity systems which have a strong synergistic effect, resulting in the protection of the host from nisin. NisI is a lipoprotein which exists in two forms, a secreted lipid-free form and a membrane associated form, which is anchored by a lipid modification of the N-terminal cysteine residue (Stein *et al.*, 2003; Koponen *et al.*, 2004). Both forms function as a nisin intercepting molecule through binding the C-terminal of NisI (Kuipers *et al.*, 1993; Koponen *et al.*, 2004; Takala and Saris, 2006). The other immunity system is an ABC transporter complex made up of two integral membrane proteins, NisG and E, and an ATP binding protein, NisF (Siegers and Entian, 1995). The protein complex has been hypothesised to consist of two NisF molecules as typical ABC transporters and have two ATPases (Peschel and Gotz, 1996). Stein *et al*. (2003) demonstrated the function of this transporter complex was to pump nisin from the membrane into the extracellular space.

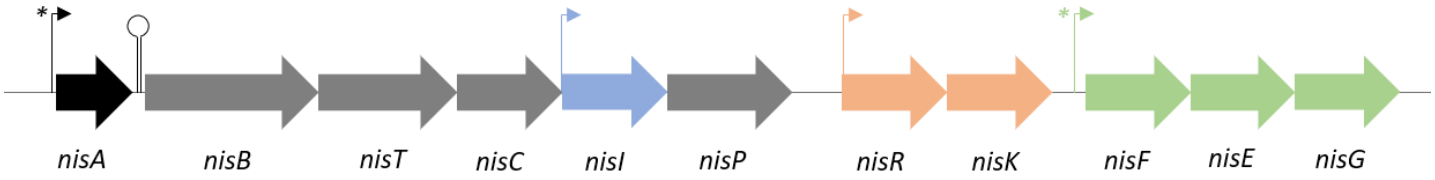


Figure 7-The organization of operons in the nisin biosynthesis pathway. The nisin biosynthesis pathway consist of 11 genes clustered into separate operons. The nisin structural gene, *nisA*, is shown in black with the processing and transportation genes shown in grey. The two immunity operons are indicated in blue and in green and the *nisRK* sensory kinases are shown in orange. The arrows represent promoters for each individual operon. The *nisBTC* operon does not have its own promoter but potential expression from read through of the *nisA* promoter. The promoters for *nisA* and *nisFEG* (indicated with a *) are known to be induced through the presence of nisin, whereas the *nisI* and *nisRK* have constitutive promoters. The loop represents an inverted repeat terminator.

3.1.2.4.3 Regulation of Nisin Production

Nisin production was determined to be autoregulated, meaning the presence of fully modified active nisin generates the production of more nisin (Engelke *et al.*, 1994). Four different promoters, P_{nisA}, P_{nisI}, P_{nisR}, and P_{nisF} generate the transcript of four operons, *nisABTCIPRK*, *nisI*, *nisRK*, and *nisFEG* respectively (Fig 7) (de Ruyter *et al.*, 1996). The gene, *nisI*, has its own weak constitutive internal promoter ensuring that the host is always protected against the presence of nisin (Li and O'Sullivan, 2006). It was determined that P_{nisA} and P_{nisF} were inducible promoters whereas P_{nisR} was a constitutive induction-independent promoter (de Ruyter *et al.*, 1996). NisRK act as a two component signal transduction system that is autoregulated by the extracellular modified nisin peptide (Kuipers *et al.*, 1995; de Ruyter *et al.*, 1996; Kleerebezem *et al.*, 1999). Nisin autophosphorylates NisK potentially via the interaction with the N-terminus of nisin. (van der Meer *et al.*, 1993; Engelke *et al.*, 1994; de Ruyter *et al.*, 1996). Transcription is induced from the subsequent transfer of phosphate to NisR, the transcriptional activator, inducing binding to P_{nisA} and P_{nisF} (Kuipers *et al.*, 1995; de Ruyter *et al.*, 1996; Kleerebezem, 2004). P_{nisA} generates two different mRNA transcripts, *nisA* and *nisABTCIPRK*, which is thought to be due to limited read-through caused by a downstream inverted repeat region (van der Meer *et al.*, 1993; Kuipers *et al.*, 1995; de Ruyter *et al.*, 1996). This terminator generates different levels of transcripts which is probably responsible for generating a balance of nisin precursor compared to modification, transportation and immunity proteins (Kuipers *et al.*, 1995). The gene organisation and protein homology is very similar to that of *B. subtilis* subtilin (Kleerebezem, 2004).

3.1.2.4.4 Mode of Action

Nisin bactericidal activity is the result of two different mechanisms. Nisin interacts with a central wall building block, lipid II, which is generated at the inner face of the membrane when the N-acetylmuramic acid moiety of lipid I is linked to undecaprenylphosphate-N-acetyl-glucosamine yielding lipid II (UDP-GlcNAc-MurNAcpentapeptide) (Heijenoort, 2001). This is then translocated to the outer cell membrane where it is incorporated into the growing peptidoglycan network (Heijenoort, 2001; Münch and Sahl, 2015). The binding of nisin to lipid II results in two different outcomes, firstly the permeabilization of the cell membrane and secondly the sequestering of lipid II away from its normal domains preventing cell wall synthesis (Brotz *et al.*, 1998; Breukink *et al.*, 1999, 2003).

The two thioether rings (A and B) of the N-terminal of nisin form a cage-like structure and bind to the pyrophosphate of lipid II (Hsu *et al.*, 2002). Once bound, nisin changes orientation to the membrane from parallel to perpendicular and is inserted into the bilayer. The N-terminus remains close to the lipid II headgroup, near the bilayer surface, and the centre of nisin is integrated in the middle of the phospholipid bilayer with the C-terminus near the phospholipid headgroup of the bilayer (Fig 8) (van Heusden *et al.*, 2002). Nisin forms a highly stable pore complex, with four lipid II and eight nisin molecules (Fig 8), embedded in the cell membrane permeabilising it, resulting in the loss of ion gradients and consequently cell death (Breukink and de Kruijff, 1999; Hasper *et al.*, 2004).

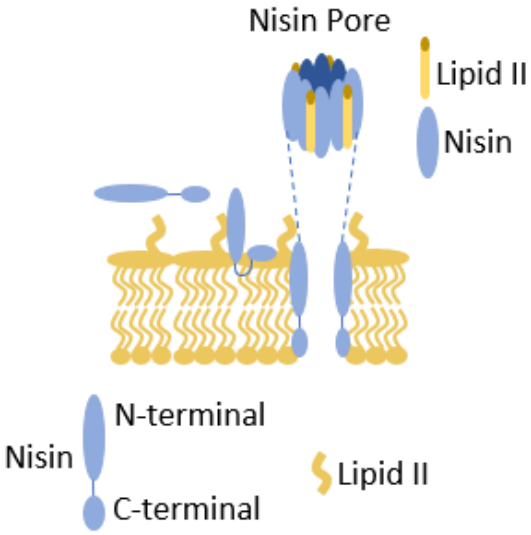


Figure 8-Nisin interaction with lipid II and pore formation Nisin molecules (blue) are parallel to the cell membrane before binding lipid II (yellow S) and inserting into the membrane generating a pore. The pore is made of eight nisin molecules (blue) bound to four lipid II molecules (yellow tube)

3.1.2.4.5 Resistance

As a result of the widespread use of nisin, within the food industry, the discovery of resistance was to be expected. Nisin resistance has been demonstrated through multiple different mechanisms ranging from a physical barrier, between nisin and the cell membrane, to preventing nisin binding to the membrane and enzymatic inactivation of nisin.

For nisin to bind its target, lipid II, it must pass through the negatively charged cell wall. One method of resistance is through the D-alanylation of teichoic acids which increases the positive charge of the cell wall, deterring nisin (Kramer *et al.*, 2006; Zhou *et al.*, 2014). After nisin binds to lipid II it generates a pore within the cell membrane however this can be inhibited through modifications to the phospholipid or fatty acid composition of the membrane. It has been shown that an increase in straight chain fatty acids or phosphatidylglycerol (PG) content of the membrane hinders the actions of nisin (Ming and Daeschel, 1995; Verheul *et al.*, 1997). Another form of resistance is the overexpression of ATP-binding cassette (ABC) transporters that pump nisin away from the cytoplasmic membrane (Kramer *et al.*, 2006). ABC transporters are a known

defence mechanism within many different bacteria against antimicrobial peptides. The expression of resistance peptides, such as nisinase and nisin-resistance protein (NSR), is another form of resistance which has been demonstrated in *Bacillus* spp., *Streptococcus thermophilus*, *Lactobacillus plantarum*, and *L. lactis* (Froseth and McKay, 1991; Sun *et al.*, 2009). Nisinase is a dehydropeptide reductase that enzymatically reduces the C-terminal dehydroalanyl-lysine in nisin, to alanyl-lysine (Alifax and Chevalier, 1962; Jarvis, 1970). NSR on the other hand is a nisin degrading protease which was discovered encoded on a plasmid in a strain of *L. lactis* which did not produce nisin. This is a significant resistance mechanism because the genes are harboured on a naturally transferrable element which allows for more rapid dissemination of the *nsr* genes (Froseth and McKay, 1991). NSR cleaves the C-terminus of nisin between the methyllanthionine^{28} and Ser^{29} residues thus decreasing its affinity for the cell membrane leading to a 100-fold decreased inhibitory activity against a highly nisin-sensitive *L. lactis* strain (Sun *et al.*, 2009). *nsr*-like genes have been found in a range of non-pathogenic species and are thought to confer resistance to nisin (Khosa *et al.*, 2013).

Despite the widespread use of nisin in the food industry for over 50 years, significant resistance outside of the laboratory setting has not been reported. However, with the potential broadening of its use into biomedical and other industries and the discovery of pathogenic strains harbouring resistance genes it is only natural to assume that resistance could be more commonplace (Draper *et al.*, 2015). Thus, increased monitoring of resistance and complementary strategies to limit or prevent resistance is crucial. One advantage of bacteriocins, over antibiotics, is that their proteinous nature allows for easier bioengineering. Protein bioengineering is a potential strategy against antimicrobial resistance. This was previously demonstrated for nisin when a bioactive NSR resistant variant was generated (Field *et al.*, 2019).

Another strategy to limit resistance is the use of nisin alongside other conventional drugs, such as antibiotics, as theoretically this should reduce the rate of resistance (Field *et al.*, 2019). Monitoring strategies need to be in place before the expansion of nisin use in order to ensure that it can continue to be used in the future.

3.2 Results

3.2.1 Discovery of a Bacteriocin Producing Strain, Strain Characterisation and Initial Activity Against *C. difficile*

Human sewage was previously collected from an East Midlands sewage plant and screened for bacteria with activity against *C. difficile* strain R20291. One sample, Strain 10, was isolated from non-heat treated raw sewage and demonstrated antimicrobial activity against the screening strain and a further four *C. difficile* strains, CD630, DH478, DH1916, and EK28 (Pickering, 2015). Additional screening, carried out in this study, showed strain 10 had activity against all the *C. difficile* strains tested (in duplicate), which were a range of ribotypes (Table 11).Previous work identified the active component was secreted into the medium as concentrated cell free supernatant (CFS) showed a zone of clearance. against *C. difficile* R20291 (Fig 9). The demonstrated activity was determined to not be an artefact of acidic pH of the spent supernatant through using catalase treatment to neutralise the hydrogen peroxide produced by the strain, however the zone was almost fully abolished (Chu Yuan Kee, 2015).

Table 11- Activity of *L. lactis* 10 Against Different *C. difficile* Strains.

C. difficile strain	PCR Ribotype	Activity
M120	078	++
TL178	002	+
CF5	017	+++
TL176	014	++
CD305	023	+++
CD196	027	++

LIV29	001	++
LIV22	106	+++
BI-9	001	+
M68	017	++
EK28	078	+
DH478	027	++
R20291	027	++
DH1916	027	+
CD630	012	+++

The sensitivity of the strain relates to the size of the diameter zone of clearance observed. 0.2 cm +, 0.3-0.4 cm ++, 0.4-0.6 cm +++



Figure 9- Zones of clearance of L. lactis 10 concentrated cell free supernatant before and after catalase treatment against a lawn of *C. difficile* CD630. Figure taken from Chu Yuan Kee, (2015). (a) Concentrated cell free supernatant after treatment with 100U of catalase, generating a zone of clearance (b) Concentrated cell free supernatant before catalase treatment, resulting in a more defined zone of clearance than observed in (a)

Previous work determined that Strain 10 was a Gram-positive coccus (Fig 10), and microscopy indicated an absence of spores therefore, this strain was allocated as a non-spore former. This was further confirmed by lack of growth after heat-treating a five-day culture at 65 °C for 30 mins. The initial 16s rRNA identified this strain as a potential *Lactococcus lactis*, and this was re-verified in this study using 16s rRNA primers 27F and 1492R and Sanger sequencing. The closest match using BLASTn in the 16s rRNA database was *Lactococcus lactis* with 96% Identification (ID) and 78% query

cover (QC), and an Expected-value (E-value) of 0. This strain was designated *L. lactis* 10, however the species and subspecies would need to be confirmed after whole genome sequencing due to the low ID and QC of the match (Rossi-Tamisier *et al.*, 2015). As bacteriocins are known to be produced at different phases of growth, a growth curve was carried out in this study (Fig 11). This growth curve demonstrates that the exponential phase of growth in GM17 medium at 10% CO_2, 30 °C and non-shaking was between 1 h – 4 h (Fig 11).

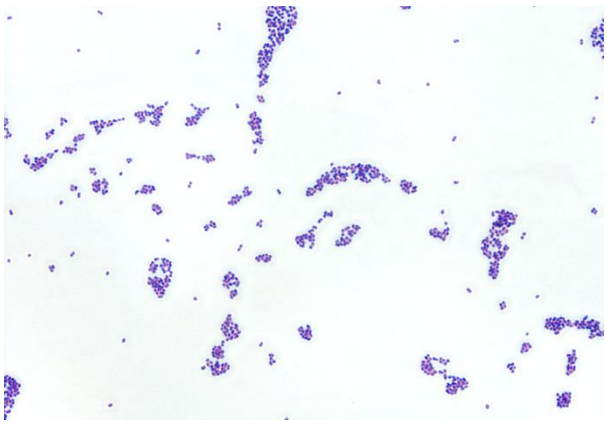


Figure 10- Gram-stain of *L. lactis* 10. Gram-staining of strain 10 at X1000 magnification shows Gram-positive cocci, in clusters or pairs. Pairs are more indicative of *L. lactis* arrangement.

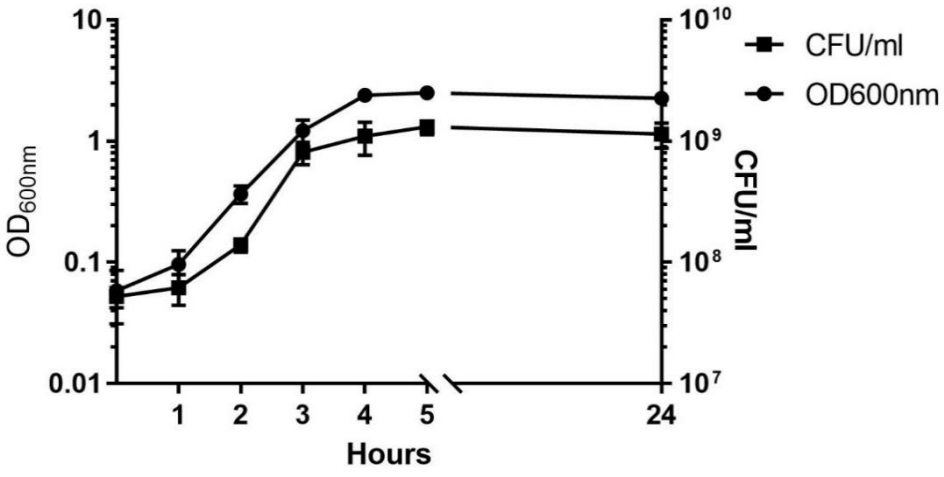


Figure 11- Growth curve of *L. lactis* strain 10 grown in GM17 at 30°C with 10% CO2 including CFU/ml. - OD_{600nm} shown on the left Y axis and CFU/ml shown on the right Y axis.

3.2.2 *De novo* Whole Genome Sequencing

L. lactis 10 was whole genome sequenced using the illumina MiSeq platform. Since this was an environmental strain with no reference genome, *de novo* assembly was used to generate a novel draft genome. *De novo* sequencing assembles sequence reads into contigs with the coverage quality dependent on the size and continuity of these contigs. *L. lactis* 10 sequencing resulted in 26 contigs, with a GC% of 34.83 and overall size of 2.54 Mb, which is consistent with the median *L. lactis* genomes on the NCBI database. As quality control, the paired end reads were mapped back to the *de novo* assembly which showed the mean mapping quality was 58.81, this means that there is a $10^{-5.881}$ probability that any given read is incorrectly aligned. There were also 0% unmapped reads, a low general error rate of 0.77% and an acceptable mean coverage of between 65X-120X (94.82 ± 28.21). The quality control showed the contigs to be reliable and therefore could be aligned to a known reference genome. This allows the contigs to be ordered based on similarity to known whole genome sequences and gives a more accurate representation of the genome *in vivo*.

Two possible reference strains, *L. lactis* KF147 and *L. lactis* subsp lactis NCD02118, had the highest sequence similarity (>98%) across the largest three contigs of *L. lactis* 10. Using Mauve Progressive Alignment (Darling *et al.*, 2004, 2010) the genomes of the references were aligned to *L. lactis* 10 demonstrating three large differences (red arrows) between *L. lactis* KF147 and *L. lactis* 10 compared to only one in *L. lactis* subsp lactis NCDO2118 (Fig 12) and thus the latter strain was used as the final alignment reference. The contigs were shuffled using Mauve (Darling *et al.*, 2004, 2010) according to the structure of *L. lactis* subsp lactis NCDO2118, which resulted in the draft whole genome sequence of *L. lactis* 10 (Fig 13). This sequence was annotated

using the RAST pipeline (Aziz *et al.*, 2008; Overbeek *et al.*, 2014; Brettin *et al.*, 2015), which identifies and assigns functions to protein-encoding genes based on homology within its database.

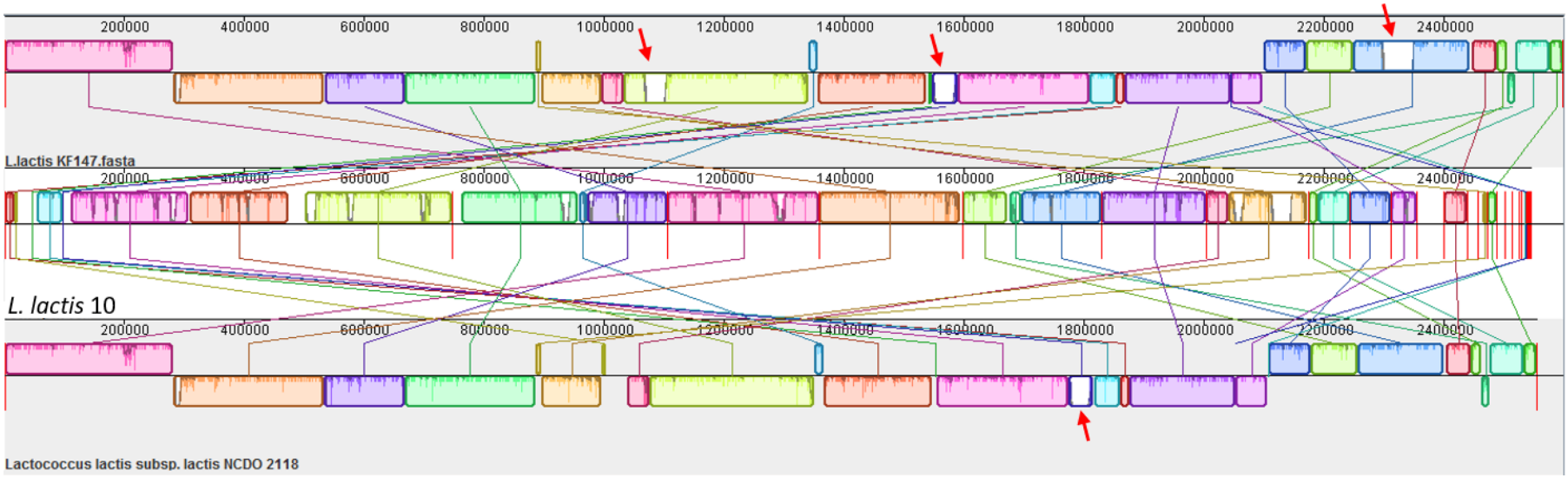


Figure 12- Contig alignment of *L. lactis* 10 to reference sequences, generated by Mauve. The two reference strains are *L. lactis* KF147 (Top strand) and *L. lactis* NCDO 2118 (bottom strand) whilst the unknown strain, *L. lactis* 10 is the central strand. The different blocks represent blocks of homology and the lines and colours indicate corresponding homology within the other genomes. The red arrows indicate non-homologous regions compared to *L. lactis* 10.

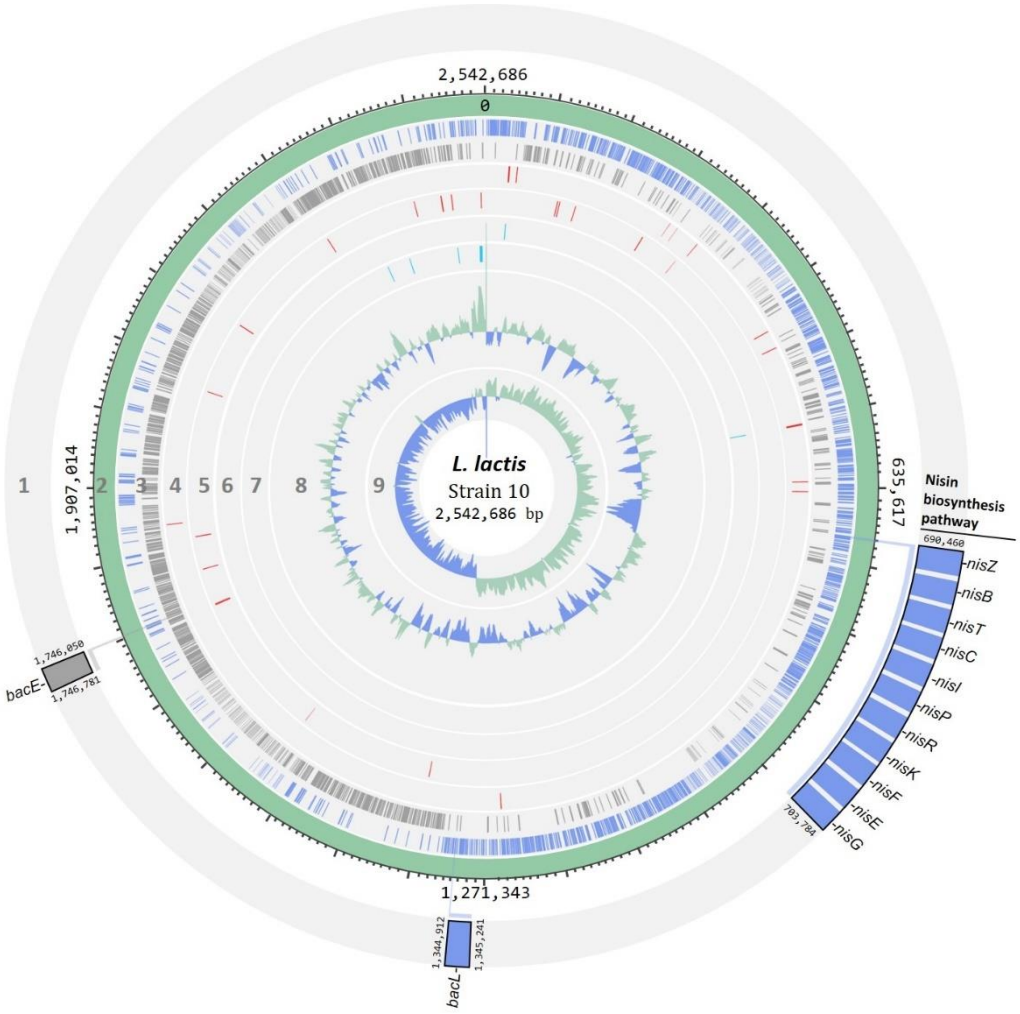


Figure 13- Circular Genome Sequence of *L. lactis* 10. Track 1: Genome features of interest; the *bacL* and *bacE* genes, and the nisin biosynthesis pathway; Track 2: source sequence; Track 3: Coding sequences on the positive (blue) and negative (green) strand; Track 4: positive strand tRNA CDS; Track 5: negative strand tRNA CDS; Track 6: positive strand rRNA CDS; Track 7: negative strand rRNA CDS; Tracks 8 and 9: G+C content and G+C skew, (blue above average GC, green below average GC).

The identification of the strain was confirmed using the DNA sequences of primers 27F and 1492R to locate the 16s rRNA *in silico*. This sequence was compared against the NCBI 16s rRNA database, and the top 94 matches were used to generate a maximum-likelihood phylogenetic tree (Fig 14, Appendix 7.1). This was created to confirm the closest species and subspecies of *L. lactis* 10, and this tree was bootstrapped 500 times to ensure enough accuracy in the prediction. According to the phylogenetic tree the

closest relatives are two *L. lactis* subsp. lactis, NCDO604 and NBRC100933 and two *L. lactis* subsp. hordniae, NCDO2181 and NBRC100931 (Fig 14, Appendix 7.1).

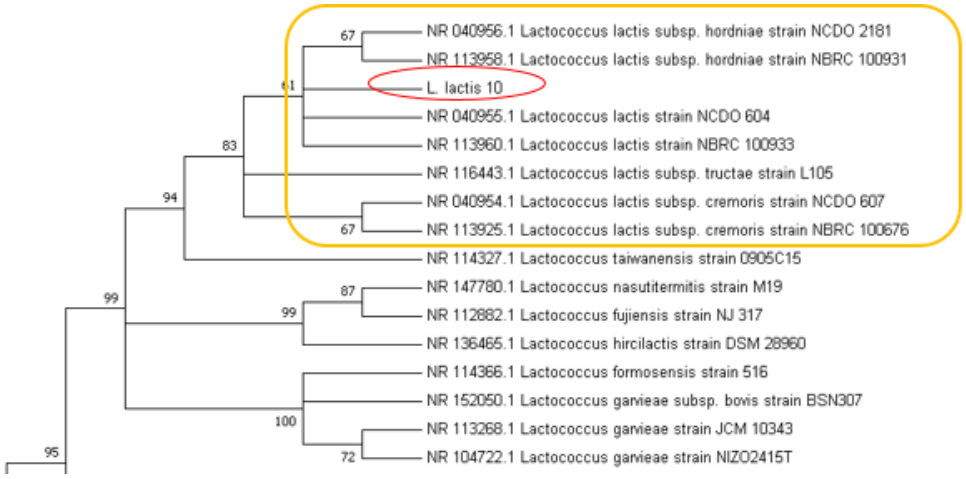


Figure 14 - Close up 16s rRNA phylogenetic tree of *L. lactis* 10. The maximum likelihood tree was generated by MEGA7 (Kumar *et al.*, 2016), and shows the relationship between *L. lactis* 10 16s rRNA and the top 94 matches against the NCBI 16s database. This phylogenetic tree used *Streptococcus* sp. as an outgroup (not shown) and was bootstrapped 500 times. The yellow square shows the *L. lactis* cluster with *L. lactis* 10 circled in red. The full phylogenetic tree is shown in appendix 7.1.

3.2.3 Bacteriocin Mining

The dedicated bacteriocin databases, BAGEL4 (van Heel *et al.*, 2018) and Bactibase (Hammami *et al.*, 2010), were used to mine the draft genome sequence of *L. lactis* 10 for known or putative bacteriocins. In total, one known bacteriocin, nisin Z, and four putative bacteriocins were annotated, BacE, BacL, enterolysin #2 and sactipepetide (Fig 13). The presence of nisin Z was confirmed in the genome using PCR with the primers *nis* F/*nis* R, which conveyed the possession of the *nisZ* structural gene and part of the downstream *nisB* gene (Fig 15). To determine the reliability of the BAGEL4 annotations for the putative bacteriocins, the sequences were also run against the NCBI database using BLASTp (Table 12). Two of the identified candidates,

enterolysin #2 and sactipeptide were further identified as a phage-related peptide and non-antimicrobial protein, respectively (Table 12) and therefore were not deemed to be likely putative bacteriocins. The other two candidates, BacE and BacL were confirmed to have homology to known bacteriocins (Table 12), enterolysin (40% similarity) and lactococcin B (37.5% similarity), and therefore considered putative bacteriocins. The presence of the encoding genes was confirmed *in vivo* using PCR, validating the BAGEL4 identification. The primers *bacE* F/*bacE* R and *bacL* F/*bacL* R were used to amplify *bacE* and *bacL*, respectively (Fig 16) and the products were Sanger sequenced to confirm their sequence. As the encoding genes were confirmed to be present for the two putative bacteriocin candidates, BacE and BacL, as well as the nisin Z structural gene, *nisZ*, further *in silico* analysis was undertaken on all three genes to investigate their potential role in the observed antimicrobial activity against *C. difficile*.

Table 12- Putative Bacteriocins Predicted to be in the Genome of *L. lactis* 10 and their Corresponding BLASTp Match

Putative bacteriocin	Location	Identity (%)	E-value	BlastP Result	Query Cover (%)	Identity (%)	E-value*	Gene name
Enterolysin A #1	1736241-1756775	46.17	2e-50	M23 family metallopeptidase	100	99.59	2e-176	*bacE*
Enterolysin A #2	625634-646078	48.8	5e-46	Phage-related lysozyme	100	99.74	0	-
Nisin Z	680462-705619	100	1e-37	Gallidermin/Nisin Family Lantibiotic	100	100	2e-32	*nisZ*
Sactipeptides	2374616-2396734	24.20	7e-19	Radical SAM protein	99	100	0	-
Lactococcin B	345104-1345232	42	6e-07	Hypothetical protein	100	100	1e-72	*bacL*

*The expected value (E-value) is a statistical likelihood the result would have been a false positive match due to the large size of the database rather than a significant match. The closer to zero the value the more significant the match is, and the less likely it is a false positive.

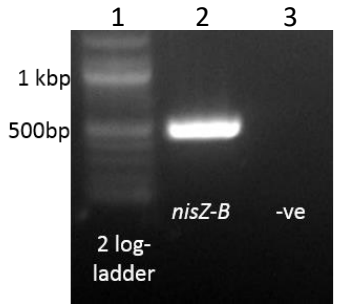


Figure 15- Amplification of partial nis operon to confirm presence of the operon, producing 518 bp product. Lane 1: 2-log ladder, lane 2 *L. lactis* 10 showing product of approximately 500 bp, lane 3: No template negative control

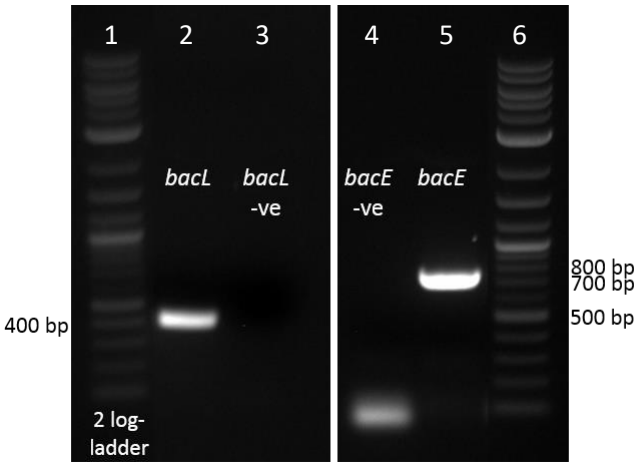


Figure 16- Amplification of *bacL* and *bacE*, producing a 428 bp and 747 bp product respectively. Lane 1 & 6: 2-log ladder, lane 2 & 4 *L. lactis* 10 with *bacL* and *bacE* primers respectively, lane 3 & 5: No template negative control with *bacL* F/ *bacL* R and *bacE* F/ *bacE* R primers respectively.

3.2.3.1 *In silico* Analysis of BacE

Using multiple protein sequence analysis software (HMMER, Pfam, InterPro, and UniProt) the amino acid sequence of BacE was allocated to the Pfam peptidase M23 family (Bateman *et al.*, 1999; Apweiler *et al.*, 2000; Finn *et al.*, 2016; UniProt Consortium, 2018). The peptidase M23 family encompasses zinc metallopeptidases, however, using the CATH classification (Orengo *et al.*, 1997) the protein falls into the outer membrane antigenic lipoprotein B functional family instead

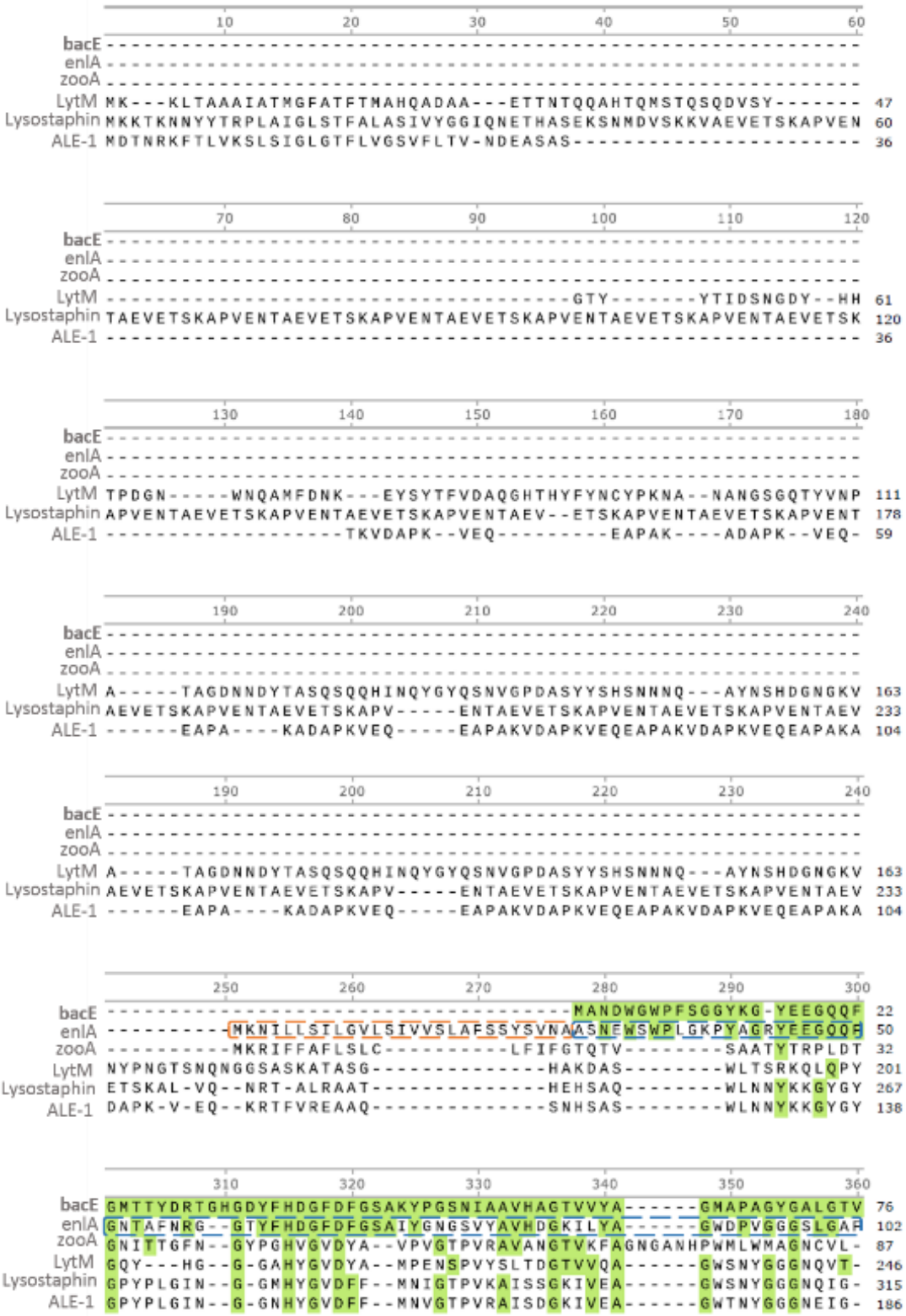
of the M23 peptidase functional family. The bacteriocin enterolysin A (EnlA) falls within the glycyl-glycine endopeptidase and the lysostaphin functional family within the same superfamily. Enterolysin A is also within this family and has 40% similarity to BacE, however BacE is smaller with only 253 amino acids and an estimated molecular weight of 26 kDa compared to enterolysin A's 316 amino acids and predicted molecular weight of 34.5 kDa. Enterolysin A is split into two regions, the N-terminal catalytic domain and the C-terminal substrate recognition domain. The N terminal domain was 54% identical to BacE, whereas the C-terminal domain was not a significant match to BacE. The N terminal domain of enterolysin A has previously been shown to be homologous to cell wall-degrading proteins also within the M23 metallopeptidase family (Table 13). These proteins are zoocin A (zooA), ALE-1, lysostaphin and LytM, which all show conserved His residues which have been suggested to function as a Zn^{2+} ligand (Fig 17) (Sugai *et al.*, 1997; Nilsen *et al.*, 2003). BacE also contains these conserved His residues at positions 120 and 122 (Fig 17), which suggests this protein could also have catalytic activity. The C-terminal domain of EnlA shares homology to a lysin from *Lactobacillus casei* bacteriophage A2 and an *N*-acetylmuramoyl-$_L$-alanine amidase from *L. casei* bacteriophage PL-1 (Fig 17). Although there is little sequence homology between enterolysin A and BacE within this C-terminal domain, there is significant homology to a 1, 4-*N*-acetylmuramidase lysozyme from *Enterococcus faecalis* ATCC6055 (WOU_02507) (Table 13, Fig 14). This suggests that the substrate recognised by this protein is different to enterolysin A's substrate, which is also indicated through the drastically different isoelectric point (pI); enterolysin A having a pI of 9.24 whilst BacE has a predicted pI of 5.27. The cytoplasm is usually at pH 7 therefore both proteins can sit comfortably within this environment although cytoplasmic proteins tend to have a pI around 5 or 6. Integral membrane proteins usually have a pI of approximately 8.5-

9.0 (Schwartz *et al.*, 2001) as the membrane and cell wall are negatively charged, indicating that enterolysin A and not BacE might target these areas.

A sequence search of UniProt for proteins similar to BacE, showed two highly similar matches from *L. lactis*. The top match with 87.6% ID, across 95.4% query coverage (QC), is a protein annotated as enterolysin A, a M23 peptidase and from *L. lactis* subsp. *cremoris*. The second match is to a Nitrile Hydratase Leader Microcin (NHLM) bacteriocin system secretion protein from *L. lactis* subsp. *lactis* with 72.8% ID across 98.8% QC. This indicates that other *L. lactis* strains encode this protein and supports a bacteriocin functionality.

Table 13- Sequence Alignment Results of Enterolysin A and BacE Against Proteins Showing Homology to Enterolysin A.

Protein	vs EnlA			vs BacE		
	Query cover (%)	Identity (%)	E-value	Query cover (%)	Identity (%)	E-value
Enterolysin A	-	-	-	58	49.2	1e-52
Zoocin	55	27.3	9e-07	67	30.2	1e-04
LytM	26	33.7	8e-08	51	27.6	0.02
Lysostaphin	26	29.9	3e-07	17	29.5	0.012
ALE-1	23	30.4	1e-07	20	27.2	8e-04
A2	43	66.9	5e-71	4	42.9	1.7
PL-1	49	58.9	9e-69	29	30.3	0.011
WOU_02507	5	42.1%	0.79	45	63.8	2e-31


10 20 30 40 50 60
bacE
enlA
zooA
LytM MK---KLTAAAIATMGFATFTMAHQADAA---ETTNTQQAHTQNSTQSQDVSY------- 47
Lysostaphin MKKTKNNYYTRPLAIGLSTFALASIVYGGIQNETHASEKSNMDVSKKVAEVETSKAPVEN 60
ALE-1 MDTNRKFTLVKSLSIGLGTFLVGSVFLTV-NDEASAS---------- 36
70 80 90 100 110 120
bacE
enlA
zooA
LytM GTY-------YTIDSNGDY--HH 61
Lysostaphin TAEVETSKAPVENTAEVETSKAPVENTAEVETSKAPVENTAEVETSKAPVENTAEVETSK 120
ALE-1 36
130 140 150 160 170 180
bacE
enlA
zooA
LytM TPDGN-----WNQAMFDNK---EYSYTFVDAQGHTHYFYNCYPKNA--NANGSGQTYVNP 111
Lysostaphin APVENTAEVETSKAPVENTAEVETSKAPVENTAEV--ETSKAPVENTAEVETSKAPVENT 178
ALE-1 ----------TKVDAPK--VEQ---------EAPAK----ADAPK--VEQ- 59
190 200 210 220 230 240
bacE
enlA
zooA
LytM A-----TAGDNNDYTASQSQQHINQYGYQSNVGPDASYYSHSNNNQ---AYNSHDGNGKV 163
Lysostaphin AEVETSKAPVENTAEVETSKAPV-----ENTAEVETSKAPVENTAEVETSKAPVENTAEV 233
ALE-1 ------EAPA----KADAPKVEQ-----EAPAKVDAPKVEQEAPAKVDAPKVEQEAPAKA 104
190 200 210 220 230 240
bacE
enlA
zooA
LytM A-----TAGDNNDYTASQSQQHINQYGYQSNVGPDASYYSHSNNNQ---AYNSHDGNGKV 163
Lysostaphin AEVETSKAPVENTAEVETSKAPV-----ENTAEVETSKAPVENTAEVETSKAPVENTAEV 233
ALE-1 ------EAPA----KADAPKVEQ-----EAPAKVDAPKVEQEAPAKVDAPKVEQEAPAKA 104
250 260 270 280 290 300
bacE ----------MANDWGWPFSGGYKG-YEEGQQF 22
enlA ----------MKNILLSILGVLSIVVSLAFSSYSVNAASNEWSWPLGKPYAGRYEEGQQH 50
zooA ----------MKRIFFAFLSLC----------LFIFGTQTV--------SAATYTRPLDT 32
LytM NYPNGTSNQNGGSASKATASG----------HAKDAS--------WLTSRKQLQPY 201
Lysostaphin ETSKAL-VQ--NRT-ALRAAT----------HEHSAQ--------WLNNYKKGYGY 267
ALE-1 DAPK-V-EQ--KRTFVREAAQ----------SNHSAS--------WLNNYKKGYGY 138
310 320 330 340 350 360
bacE GMTTYDRTGHGDYFHDGFDFGSAKYPGSNIAAVHAGTVVYA------GMAPAGYGALGTV 76
enlA GNTAFNRG--GTYFHDGFDFGSAIYGNGSVYAVHDGKILYA------GWDPVGGGSLGAF 102
zooA GNITTGFN--GYPGHVGVDYA--VPVGTPVRAVANGTVKFAGNGANHPWMLWMAGNCVL- 87
LytM GQY---HG--G-GAHYGVDYA--MPENSPVYSLTDGTVVQA------GWSNYGGGNQVT- 246
Lysostaphin GPYPLGIN--G-GMHYGVDFF--MNIGTPVKAISSGKIVEA------GWSNYGGGNQIG- 315
ALE-1 GPYPLGIN--G-GNHYGVDFF--MNVGTPVRAISDGKIVEA------GWTNYGGGNEIG- 186

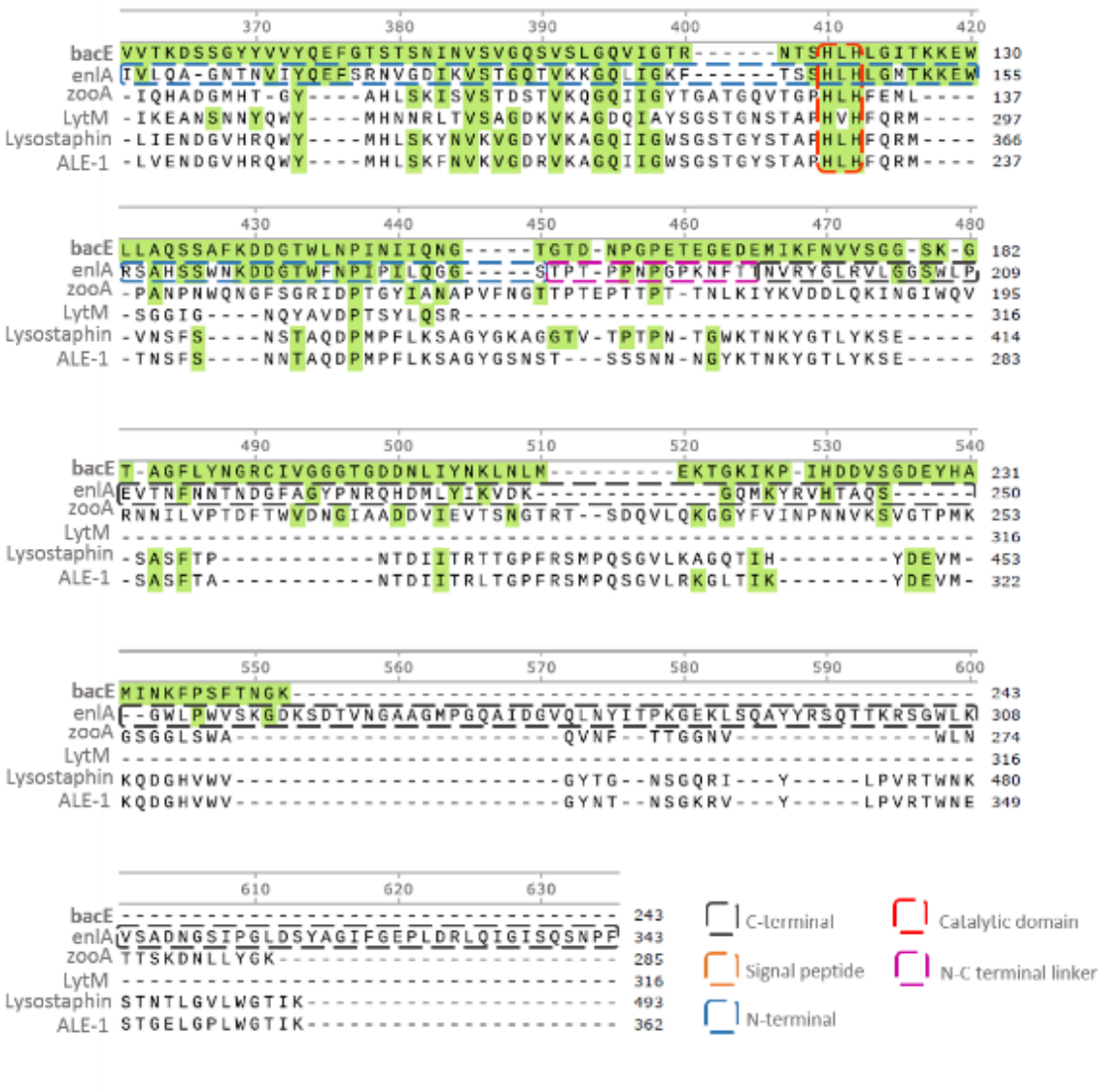


Figure 17- Sequence alignment of BacE, EnlA, ZooA, LytM, lysostaphin and ALE-1. BacE is highlighted as the reference in green with homologous residues from other bacteriocins also shown in green underneath. The black, orange, blue, red and purple dotted boxes show the C-terminus, signal peptide, N-terminus, catalytic domain and N-C terminal linker of EnlA respectively.

3.2.3.2 *In silico* Analysis of BacL

BacL was considered a putative bacteriocin candidate within the genome of *L. lactis* 10 due to homology to lactococcin B (LcnB), a known bacteriocin. There is a sequence similarity of 37.5% between BacL and LcnB (QC 58%, E-value 8e-11). The protein sequence also shows a lactococcin family domain architecture across the first 60 residues, further indicating it as a putative lactococcin bacteriocin (Fig 18). Lactococcin A, B and M are all processed at a glycine doublet, and although there is a signal peptide predicted in BacL between Ala28 and Asp29 there is a downstream glycine doublet present (Fig 18) which would result in a processed peptide similar to the size of LcnA, B and M (van Belkum *et al.*, 1992). Mirroring the operon layout of lactococcin B (*lcnB*), a putative self-immunity gene (*baiL*) is located downstream of *bacL* (Fig 19). This gene encodes a protein with a Pfam Enterocin A immunity family domain indicating its potential role in host immunity. The operon of *lcnB*, consists of a single promoter upstream of *lcnB* with a self-immunity gene (*lciB*) and terminator downstream. In the case of the *bacL* region there is a possible RBS site upstream of both genes and predicted -10 and -35 promoter binding sites upstream of *bacL*. The distance between the promoter binding sites and RBS is a similar distance as observed in the *lcnB* promoter. There is also a predicted terminator downstream of *baiL*, which again reflects the structure of the operon of *lcnB* (van Belkum *et al.*, 1992). The *in silico* analysis therefore suggests the potential for *L. lactis* 10 to express a lactococcin bacteriocin and also confer host protection with an immunity protein.

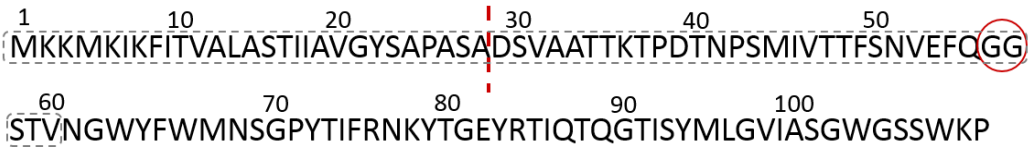


Figure 18-The amino acid sequence of BacL. The lactococcin family domain match is indicated with the blue dotted box which spans across the first 60 residues. The predicted signal peptide is shown between Ala28 and Asp29 with a red dotted line however the downstream glycine doublet is indicated with a red circle.



Figure 19-The putative BacL operon. The BacL structural gene (blue), *bacL,* is located upstream of the predicted self-immunity gene, *baiL* (green). The predicted RBS site (Red) is downstream of the predicted -35 (grey) and -10 (yellow) promoter binding sites. A terminator is also predicted downstream of *baiL* (loop).

3.2.3.3 *In silico* Analysis of the Nisin Operon

All the genes from the nisin biosynthesis pathway (Section 3.1.2.4, Fig 5) are present in *L. lactis* 10 with the closest match to *L. lactis* UC11 and UC08 with 99% identity over 100% QC. All the promoters and terminator in the pathway are intact with the differences occurring in the NisB, NisT, NisP and NisG protein sequence with the most variation occurring between NisB and NisP, which have three amino acid changes in the structural proteins. The three amino acid substitutions in NisB are not in areas previously demonstrated to be essential for activity (Khusainov *et al.*, 2015). There are also amino acid changes in NisT, NisP, and NisG compared to the known protein sequences which could affect the protein function. Due to the lack of annotation of the significant residues in NisT, P and G it is unknown if these changes would have any effect on nisin biosynthesis and immunity. It is important to note that

the nisin structural gene in *L. lactis* 10 is the variant nisin Z which has a higher diffusion rate compared to NisA (de Vos *et al.*, 1993).

3.2.4 Activity of Putative Bacteriocins

Two putative bacteriocins, BacE and BacL were identified within the genome of *L. lactis* 10 which may be responsible for the observed antimicrobial activity against *C. difficile* strains. To test whether these putative bacteriocins were active against *C. difficile* we sought to clone the genes into an *E. coli* T7 over-expression system and test the purified protein for biological activity against *C. difficile*.

3.2.4.1 Creation of Overexpression Vectors pET28b(+)*bacE* and pET28b(+)*bacL*

The identified putative bacteriocin genes, *bacE* and *bacL* were amplified from the *L. lactis* 10 genome using the primers *bacE* F/*bacE* R and *bacL* F/*bacL* R, respectively (Table 7, Section 2.1.5). These primers also incorporated the restriction enzyme sites *Xho*I and *Nde*I into the 5' and 3' end of both genes respectively. A T7 expression vector pET28b (+) was selected to be used due to its high expression level and an N or C terminal six histidine tag (His-tag). In this instance the N terminal tag was selected due to the presence of a thrombin site to cleave off the His-tag, if necessary. The vector was linearised using *Xho*I and *Nde*I restriction sites, in the multiple cloning site (MCS), and the DNA backbone was gel purified. The PCR gene products were digested with the same enzymes and gel purified. The vector backbone and insert were ligated together to create pET28b(+)*bacE* (Fig 20a) and pET28b(+)*bacL* (Fig 20b) which were then transformed into *E. coli* Top10 and the sequence confirmed with Sanger sequencing using the T7 Terminator and T7 promoter primers (Table 7, Section 2.1.5). The plasmids were then transformed into the *E. coli* expression strain BL21 (DE3),

which was selected due to its compatibility with the T7 expression vector, its high transformation efficiency and reduced expression of proteases.

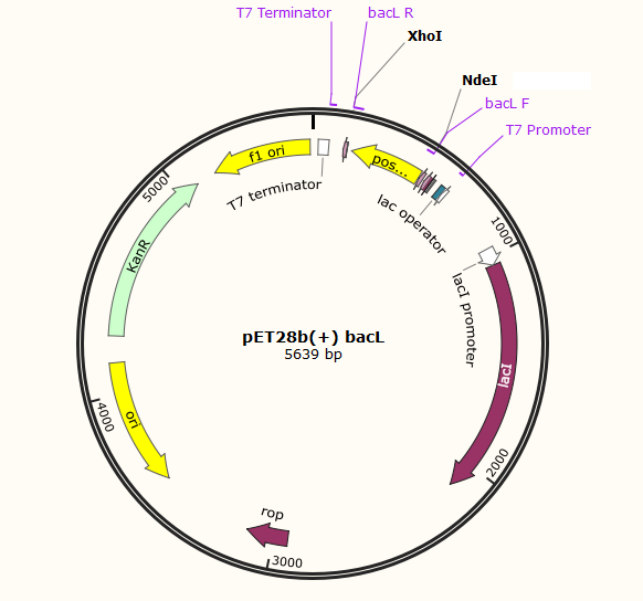
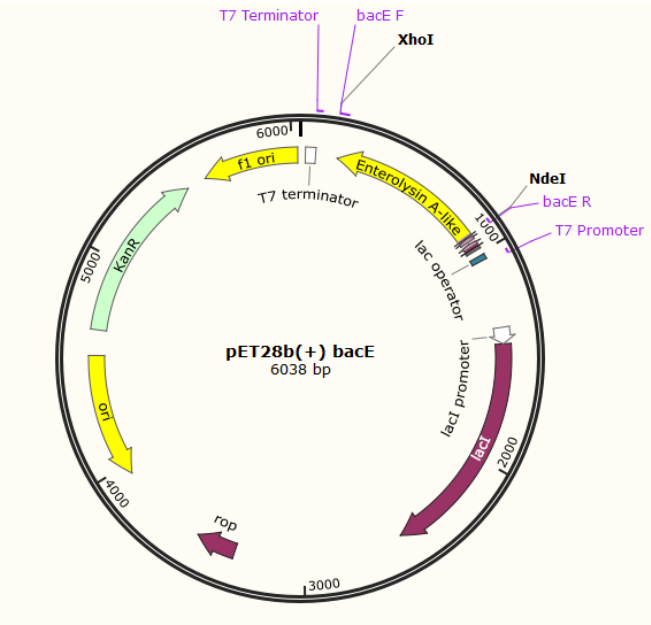


Figure 20- Plasmid maps of pET28b(+)bacL and pET28b(+)bacE. (a) Plasmid map of the expression vector pET28b(+) with the *bacL* gene cloned into the MCS using the *Xho*I/*Nde*I restriction sites. The gene amplification primers (*bacL* R and *bacL* F) and the screening primers (T7 Terminator and T7 promoter) are indicated in purple.



(b) Vector map of the pET28b(+)*bacE* with *bacE* cloned into the pET28b(+) vector MSC using *Xho*I/*Nde*I restriction sites. The gene amplification primers (*bacE* R and *bacE* F) and the screening primers (T7 Terminator and T7 promoter) are indicated in purple.

3.2.4.2 Overexpression of *bacE*

Overexpression of the *bacE* gene was carried out at room temperature for 10 h using a single *E. coli* overexpression strain, *E. coli* BL21(DE3). The strain carrying BacE and a second carrying the empty pET28b(+) expression vector, as a negative control, were cultured to an OD_{600nm} of 0.4-0.6 before being induced with 1mM IPTG for 4 h at 37 °C. To determine if there was overexpression of BacE, samples of culture were taken every hour for 4 hours and run on an SDS-PAGE. The sample was pelleted, and the cells lysed using BugBuster® Master Mix and the soluble fraction was run on a 4-12% Bis-Tris NuPage SDS-PAGE gel. A difference in profile can be seen between non-induced and induced cultures (Fig 21). The predicted molecular weight of BacE with the His-tag is 28 kDa and in the SDS-PAGE profile (Fig 21) an increase in the intensity of two polypeptides can be seen over the first 4 h of induction, one above 28 kDa and one below. However, there is also a slight increase in the same size proteins in the non-induced control (Fig 18); this is probably due to the induction system leaking and expressing the protein despite the lack of inducer. Consequently the negative control was changed from non-induced culture to *E. coli* BL21 (DE3) carrying the empty pET28b(+) vector and 1 mM IPTG was added to this culture to mirror the induction.

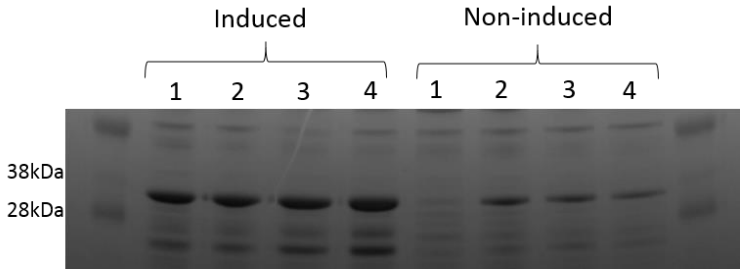


Figure 21-SDS-PAGE gel showing protein profile of E. coli pET28b(+)*bacE* cell lysate taken from induce and non-induced culture over time. The lanes correspond to hours of induction for each condition. Two proteins between 28kDa-38kDa show an increase in production of protein over time and are more concentrated compared to the non-induced, however there is also an increase in these two proteins in the non-induced.

As there was good overexpression of protein, His-tag purification was used to isolate the protein from the cell lysate. Nickel resin was used for the purification and showed three high intensity proteins compared to the negative control in the final eluate (Fig 22). There was some loss of protein in the flow-through which did not bind to the resin and some protein was washed out in the 1st and 2nd washes (Fig 22 W1 and W2 respectively) however, there was still a high yield of protein in the eluate. The eluate was very contaminated with other proteins and the purification process needed optimisation for further downstream processes. The eluted samples were tested in a Western blot, which showed two proteins in the eluate that contained the His-tag, their sizes are between 25-37 kDa and 20-25 kDa (Fig 22b) suggesting the correct protein is the higher protein as it is within range of the predicted molecular weight of BacE. To aid in the removal of the potential contamination in the elute fraction, zinc resin was used instead of nickel resin as it would better suit the BacE protein as it is a zinc metalloprotease. Although there was less protein binding to the resin and more protein loss in the washes, there was also less contamination in the eluate resulting in higher purity (Fig 23). There is one other very low molecular weight protein visible, however this is probably degradation (Fig 23). Smaller proteins evident on the western blot could be the non- his tagged C-terminal sequestered away from the His-tag after protein degradation (Fig 23).

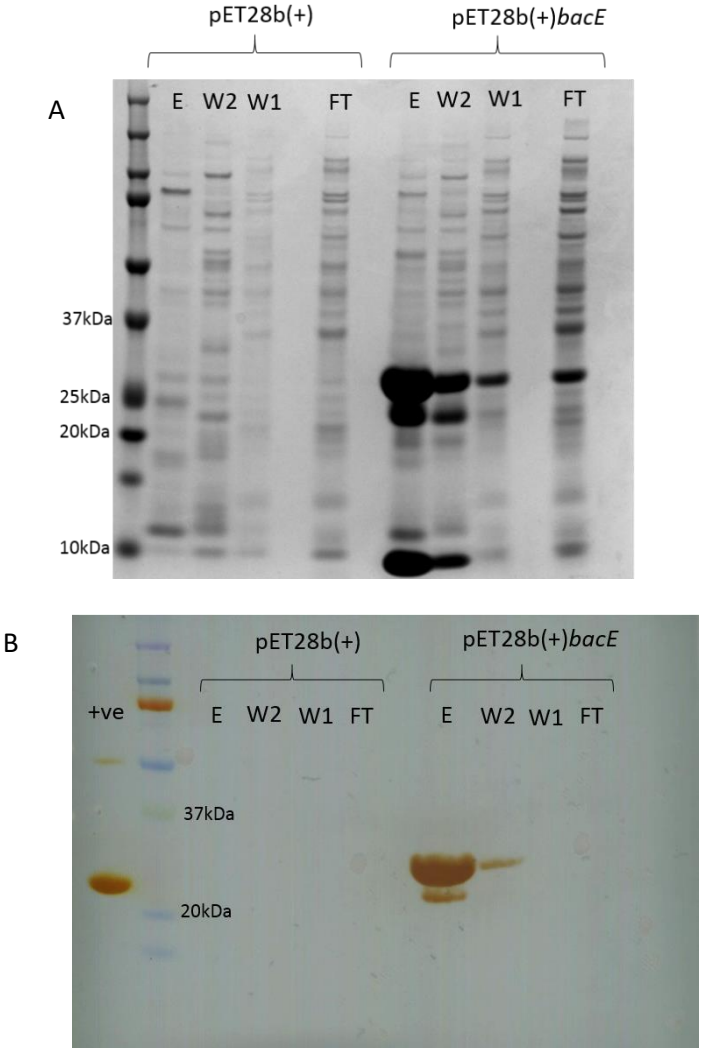


Figure 22- SDS-PAGE gel and western blot of the nickel protein purification of *BacE* (a) Coomassie stained NuPage Bis-tris 4-12% SDS-PAGE of nickel resin His-tag purification of *E. coli* pET28b(+) and pET28b(+)*bacE*. The left hand lane contains the molecular maker Precision Plus Protein™Kaledioscope™ (Bio-Rad). From right to left the lanes show proteins in the flow through (FT), 1st column wash (W1), 2nd column wash (W2) and the elute (E) fractions for pET(+)*bacE* and the empty pET28(+) vector respectively. (b) A western blot of the same samples using an anti-6XHis-peroxidase antibody (Roche) was carried out on (a) in the previous order with a positive western blot control in the left land lane.

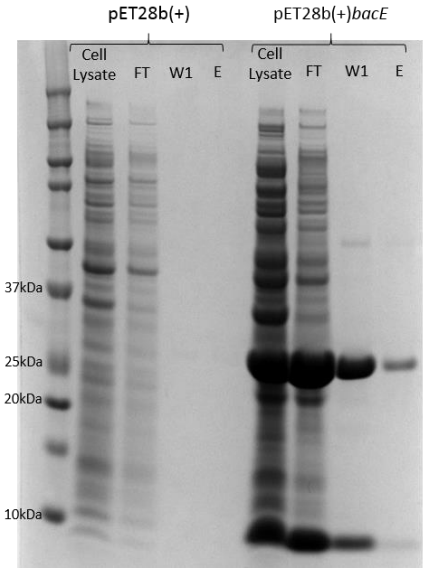


Figure 23- SDS-PAGE gel of the zinc protein purification of BacE. Coomassie stained NuPage Bis-tris 4-12% SDS-PAGE of zinc resin His-tag purification of *E. coli* pET28b(+) and pET28b(+)*bacE*. From right to left the lanes show proteins in the cell lysate, flow through (FT), 1st column wash (W1), and the eluate (E) for pET(+)*bacE* and the empty pET28(+) vector respectively. The left hand lane contains the molecular maker Precision Plus Protein™Kaledioscope™ (Bio-Rad).

3.2.4.3 Activity of BacE Against C. difficile

The zinc purified BacE protein was tested for activity against *C. difficile,* using both a well overlay and the SDS-PAGE overlay method (Section 2.1.13). Smaller molecular weight proteins can be separated through on a normal denaturing SDS-PAGE gel and renatured by washing the SDS off and the addition of 0.1% v/v Tween80 to washes. However, the larger size of BacE (~25kDa) could result in a more complex protein structure and therefore lead to more difficulty renaturing. To overcome this, a Native-PAGE gel was used without any detergent or boiling of the sample prior to running on the gel. Unfortunately, the protein in this sample was not able to run to its predicted molecular weight (Data not shown), however this is not unusual for Native-PAGE gels. The addition of different detergents (Tween 80, Triton X-100 and Chaps) and altering the pH of the elution buffer to trial a range of pHs between 5.3-9.3, however this did not result in any change in the profile of the banding pattern (Fig 24a). The two bands with the highest intensity were sliced out of the gel and boiled in more SDS-PAGE reducing buffer and the extract loaded on to an SDS-PAGE gel (Fig 24b). Both bands ran to the same MW between 28kDa and 38kDa, although there was a much lower concentration of the lower molecular weight band (Fig 24a, blue arrow) thus resulting in an extremely faint band (Fig 24b, blue arrow). This shows that the protein is running into the non-reducing gel but not to the same size as on an SDS-PAGE, which could be a result of aggregation, protein folding, and/or charge effects.

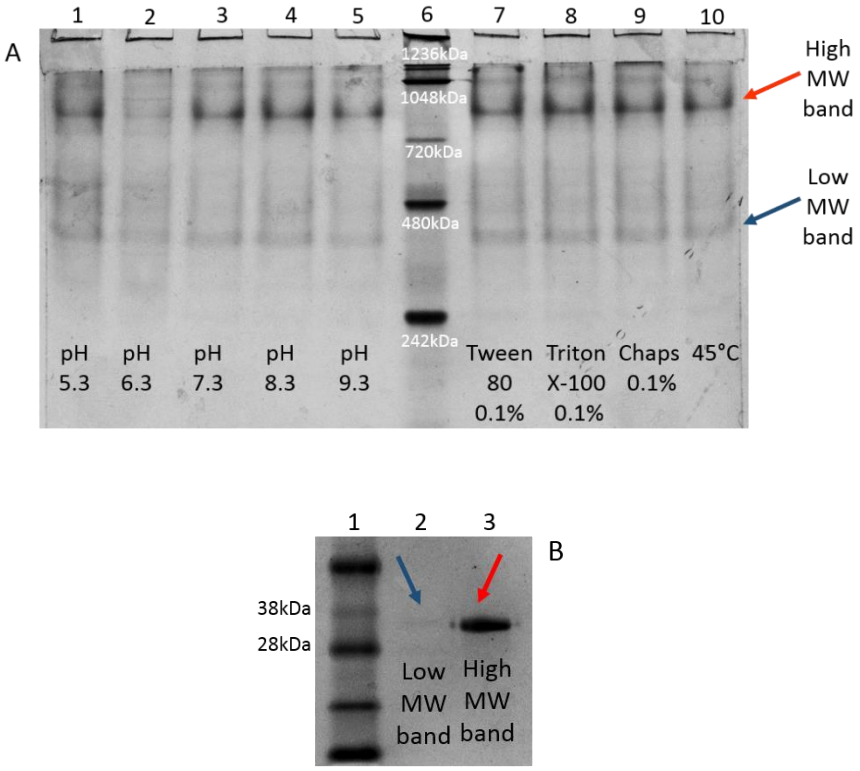


Figure 24-Purified BacE separated under different conditions on a Native-page gel and re-run on SDS-PAGE gel (a) Nickel resin purified BacE eluted in different pH elution buffer shown in lane 1-5 corresponding to pH 5.3-9.3. Lane 6 contains protein standard ladder NativeMark Unstained Protein Standard (Thermo Fisher Scientific). Lane 7-9 shows BacE in sample buffer with 0.1% of different detergents added and lane 10 shows BacE heated in sample buffer to 45°C for 30 mins. (b) shows extracts of the lower MW band in lane 2 (blue arrow) and high MW band in lane 3 (red arrow) from native-page gel shown in a. re-run on a 4-12% Bis-Tris NuPage SDS-PAGE gel. Lane 1 is the SeeBlue Pre-stained Protein Standard (Thermo Fisher Scientific).

Activity could be seen with nickel purified BacE, after running the Native-PAGE gel (Fig 25) with two mirror repeats. One half was overlaid with *C. difficile* CD630 (Fig 26a) and the other half was coomassie stained (Fig 26b). The zone of clearance can be seen in all four lanes and not in the empty lane, however the lane containing BacE alone has a zone further down the gel than the other three wells (Fig 26). This could be due to the lack of detergent meaning the protein ran faster through the gel than the other samples. This was repeated using a protein sample with the cell lysate obtained via freeze-thawing instead of using BugBuster® Master Mix (data not shown) however the detection of activity could not be repeated using zinc purified protein.

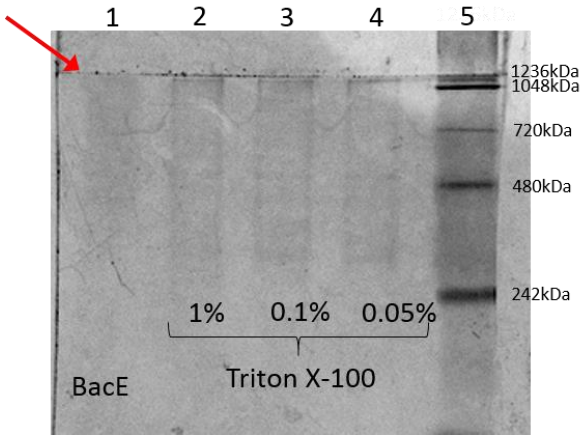


Figure 25-10% NativePage gel showing nickel purified BacE. Lane 1: BacE showing a smear of protein starting from the red arrow, Lane 2-4: BacE with different percentages of Triton X-100, which shows the same smear as lane 1. Lane 5: NativeMark Unstained Protein Standard (Thermo Fisher Scientific)

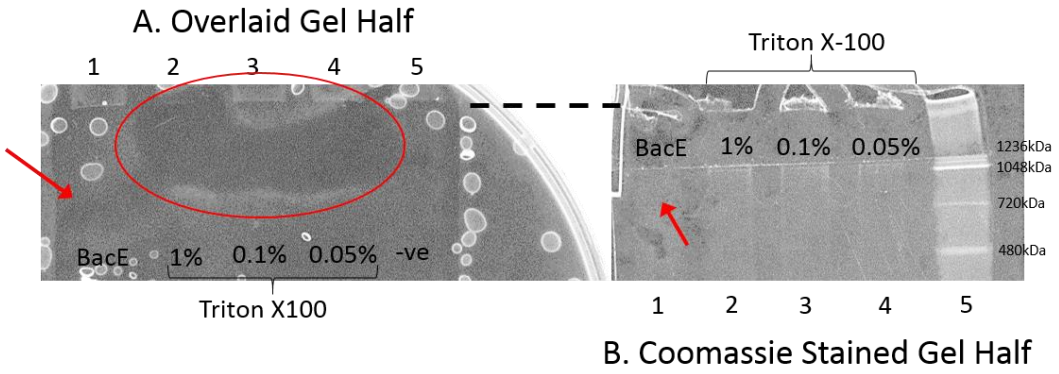


Figure 26-The zone of clearance observed from BacE on the overlaid half of the Native-PAGE gel corresponding to protein bands visualised on the strained half of the Native-PAGE gel. (a) Native-page gel overlaid with 0.5% soft-top agar inoculated with *C. difficile* CD630 which can be seen as a lawn of growth in light grey. A zone of clearance can be seen across lanes 1-4 (red arrow and circle) with an empty well as a control in lane 5. The zone in lane 1 (red arrow) is lower than the zone in lanes 2-4 (red circle) (b) The Native-PAGE gel as seen in figure 22, imaged next to the overlaid halfin order to match up the peptide bands (red arrow) to the zone of clearance. The top of the overlaid half and stained half are aligned as shown by the black dotted line. This is the coomassie stained half of the gel in shown in a. which contains BacE samples with different concentration of Triton X-100. The alignment of the stained and overlaid halves of the Native-PAGE gel shows the zone partially overlaps the protein smear present in lanes 1-4 (red arrow).

3.2.4.4 Overexpression of Lactococcin B-like Protein BacL

Overexpression of *bacL* was carried out in two different *E. coli* expression strains. The first strain *E. coli* BL21 (DE3), was cultured to an OD_{600nm} of 0.4-0.6 prior to being cooled to room temperature and induced with 1mM IPTG for 10 h. The cells were then pelleted and lysed using BugBuster® Master Mix, the insoluble and soluble

fraction were run on a 16% Tricine SDS-PAGE which were silver stained and western blotted to locate the presence of any overexpressed protein. There was no observable difference in protein pattern compared to the negative control, and the western blot was negative (data not shown), thus it was determined that the protein was not being overexpressed. A second expression strain, *E. coli* C43(DE3)pLysS was used to try and achieve protein expression; however, there was still no observable protein expression compared to the negative control (data not shown).

3.2.5 Alternative Methods of Determining Activity

Since the over-expression of the two putative bacteriocins in an *E. coli* hosts did not produce conclusive or repeatable results of activity, a different approach was required. Creating knock-outs of each putative bacteriocin within the *L. lactis* 10 strain could identify the causative agent of the observed antimicrobial activity. This approach would also determine if there were any potential synergistic effects between the putative bacteriocins and nisin. However, since this is an environmental strain of *L. lactis* and not a known laboratory strain its ability to be genetically manipulated is unknown and therefore this was investigated.

3.2.5.1 Strategy 1: Creating Bacteriocin Knock-Outs In *L. lactis* 10

3.2.5.1.1 Designing a Knock-Out Plasmid

Three different *L. lactis* knock-out strains were designed, each strain would have one putative bacteriocin and Nisin Z knocked-out and one strain would have all three candidates knocked-out. In order to do this, three different knock-out constructs were developed based an adaptation of the allelic exchange method (Heap *et al.*, 2012). This technique uses sequences of homology flanking the gene targeted for deletion to incorporate an antibiotic resistance cassette in the place of the targeted gene following two distinct homologous recombination events. A plasmid is generated containing the antibiotic resistance cassette with a left and right homology arm (LHA

and RHA, respectively) over 1 kb in length, these arms include the first and last 9 nucleotides of the targeted gene. As described in figure 24 the first homologous recombination results in a single cross over with the plasmid incorporated into the chromosome of the organism. A second homologous recombination generates a double cross over in which the antibiotic cassette is incorporated into the chromosome and the plasmid backbone is excised and subsequently lost. For this study a suicide vector with this knock-out cassette was designed in order to knock-out the targeted gene leaving behind a truncated gene consisting of three amino acids from the 5' and 3' prime ends of the targeted gene (Fig 27). The *E. coli* cloning vector pUC19 was chosen as the backbone vector because it does not have a lactococcal origin of replication and thus would act as a suicide vector. At the 5' and 3' prime ends of the antibiotic selection marker is a *lox* site, these are specific recognition sites for a *cre* recombinase. When the *lox* sites are in the same orientation, they recombine through transient expression of *cre* recombinase, and the addition of the *cre*-helper plasmid, and cause the excision of any DNA between the sites including the loss of one of the *lox* sites themselves. This would result in the loss of the antibiotic resistance cassette generating a clean knock-out mutant (Fig 27).

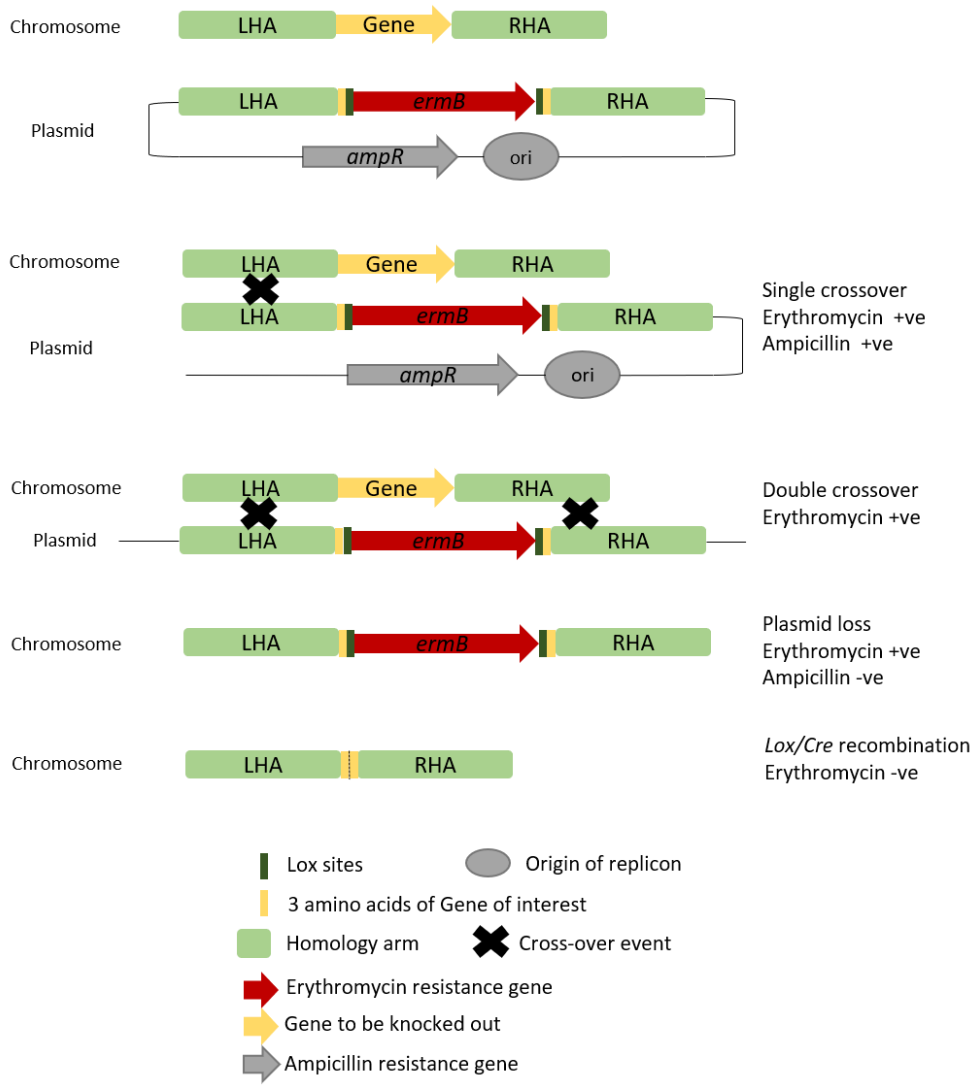


Figure 27- Figure of the allelic exchange method designed to knock-out genes of interest in *L. lactis* 10. Left and right homology arms (LHA and RHA respectively, green boxes) are designed to be homologous to the chromosomal region either side of the gene of interest. In between the homology arms on the plasmid is an erythromycin resistance cassette (red box) which would be incorporated into the chromosome instead of the gene of interest as an antibiotic selection marker (red box). Once homologous recombination has occurred generating a double cross over, plasmid loss will be selected for through the loss of ampicillin resistance. Once a mutant has been generated the erythromycin resistance cassette could be lost using the cre/lox system as lox sites (dark green box) were incorporated at both the 5' and 3' prime ends of the erythromycin resistance gene. This would then result in a clean knock-out mutant which is once against sensitive to erythromycin.

3.2.5.1.2 Minimal Inhibitory Concentration

To select a suitable antibiotic selection cassette and compatibility to the pUC19 vector, the antibiotic resistance profile of *L. lactis* 10 was determined. Antibiotic sensitivity was tested over a 48 h period with time points at 0, 4, 24, and 48 h. The lack

of growth at 48 h (Fig 28) showed that this strain was sensitive to both antibiotics and the minimal inhibitory concentrations were 2.5 µg/ml for ampicillin (Fig 28a) and 0.5 µg/ml for erythromycin (Fig 28b). The working concentrations were set at 5 µg/ml for ampicillin and 1 µg/ml for erythromycin.

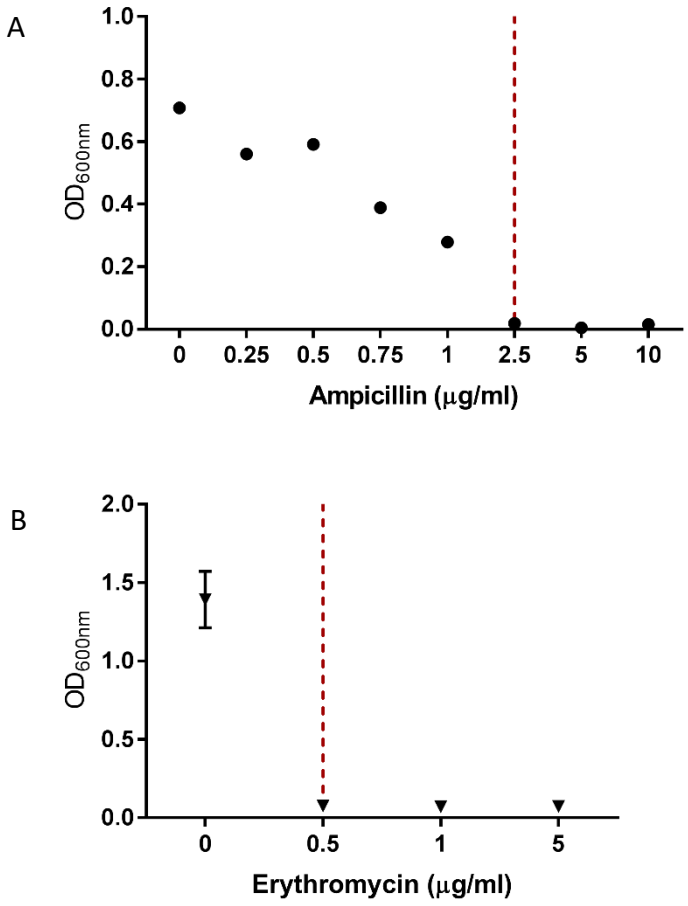


Figure 28- Minimal inhibitory concentration of ampicillin and erythromycin against *L. lactis* 10 in GM17 media. The OD_{600nm} for *L. lactis* 10 grown in GM17 media at 37°C at 10% CO_2 with different concentrations of (A) ampicillin or (B) erythromycin and the MIC value indicated by the red dotted line.

3.2.5.1.3 Trialing Electrocompetent L. lactis 10

There are multiple different methods for preparing electro-competent *L. lactis*. In order to determine if this environmental strain of *L. lactis* would be able to uptake foreign DNA, cells were prepared using a protocol adapted from Holo and Nes (1995), Geertsma and Poolman (2007), Papagianni, Avramidis and Filioussis (2007).

The cells were plated after -80 °C storage, and after electroporation, without plasmid, to confirm their viability and survival of electroporation. Once confirmed, the conditions trialled were kilovolts for electroporation, length of recovery, recovery broth and volume, and plasmid concentration. All transformations were carried out with a lactococcal compatible positive control plasmid pMTL8amβ251. This plasmid was generated through digesting pMTL8amβ151 with restriction enzymes *Pme*I/*Fse*I and ligating in antibiotic selection marker *ermB* digested out from pMTL84221 using the same enzymes.

The electro-competent cells generated only six colonies, which is not deemed to be a high enough efficiency for generation of knock-out mutants (Table 14). Therefore, this protocol will need to be optimised to improve the efficiency of transformation, this could be done by growing the cells in the presence of a cell wall weakening chemicals, such as threonine or glycine, which have both been tried before with *L. lactis* (Holo and Nes, 1995; Papagianni *et al.*, 2007). Also, *in silico* analysis did identify the presence of an endonuclease encoding gene, *LlaFI*, (Section 3.2.2, Fig 10) which plays a role in the type III restriction system previously shown to be active in some *L. lactis* strains. Although it is known to incompletely digest foreign DNA it could also be a cause of the low transformation efficiency (Su *et al.*, 1999). Before optimising this protocol, a second strategy (Section 3.2.5.2) was explored to establish if nisin was solely responsible for the observed antimicrobial activity thus ascertaining the requirement of knock-out mutants.

Table 14- Table of Trialled Electroporation and Recovery Conditions for Electrocompetent *L. lactis* 10

Plasmid concentration	Electroporation	Recovery broth	Recovery volume	Recovery Length	Colonies
1ug	2.5kV, 200Ω, 25vF	GSM17 MgCl CaCl	5 ml	1 hr	No
1ug	2.5kV, 200Ω, 25vF	GSM17	5 ml	1 hr	No
1ug	2.5kV, 200Ω, 25vF	GSM17 MgCl CaCl	5 ml	3 hrs	No
100 ng	2.5kV, 200Ω, 25vF	GSM17 MgCl CaCl	2.5 ml	Ice 5 mins 37°C 3 hrs	6
100 ng	1.3kV, 200Ω, 25vF	GSM17 MgCl CaCl	2.5 ml	Ice 5 mins 37°C 3 hrs	No
200 ng	2.5kV, 200Ω, 25vF	GSM17 MgCl CaCl	2.5 ml	Ice 5 mins 37°C 3 hrs	No
200 ng	1.3kV, 200Ω, 25vF	GSM17 MgCl CaCl	2.5 ml	Ice 5 mins 37°C 3 hrs	No

3.2.5.2 Strategy 2: Generation of Nisin Resistant *C. difficile* 630Δ*ermΔpyrE* and *C. difficile* R20291Δ*pyrE*

A second strategy was designed to generate nisin resistant strains of previously sensitive *C. difficile* which would ascertain if the observed antimicrobial activity was solely caused by nisin. Due to the known activity of nisin against *C. difficile* and the *in silico* analysis showing the presence of the intact operon it was hypothesised that this operon was active and, in all probability, responsible for the majority of the observed activity. This approach would be to confer resistance to *C. difficile* strains through the expression of the nisin immunity genes. Previously a *Bacillus subtilis* strain was used as a surrogate host to determine the role of four genes in nisin immunity, resulting in the host becoming resistant up to 8 µg of nisin (Stein *et al.*, 2003). By engineering the same genes expressed in *C. difficile*, a nisin resistant strain would also be generated. In order to generate this mutant, four genes would need to be expressed, one nisin-intercepting gene (*nisI*) and three transportation genes (*nisFEG*) which export nisin from the cytoplasmic membrane back into the environment. The expression of all four genes were shown to confer the greatest resistance against Nisin (Stein *et al.*, 2003) and therefore should all be simultaneously expressed in the *C. difficile* strains.

3.2.5.2.1 Generation of nisIFEG Expression Vector

Following the method of Stein *et al.* (2003), a vector was designed to incorporate the genes identified in conferring resistance to nisin (*nisIFEG)* and under control of the *C. perfringens* ferredoxin promoter, P_{fdx}, compatible with *C. difficile* (Takamizawa *et al.*, 2004). The plasmid generated to express these genes in *C. difficile* was pMTL84123, this contains a *C. difficile* compatible replicon (pCD6), a low copy *E. coli* replicon (p15a), and a chloramphenicol resistance marker (*catP*) compatible in

both strains (Heap *et al.*, 2009). In *L. lactis* the genes *nisFEG* are under control of the same promoter and *nisI* is part of a different transcriptional unit, therefore all four genes were fused together and put under one *C. difficile* compatible promoter. Using the primers *nisI*F and *nisI*R, the *nisI* gene was amplified with an *Nde*I restriction site introduced at the 3' prime end and with 10 bp of overlapping sequence from *nisF* at the 5' prime end. The primers *nisFEG* F and *nisFEG* R were used to introduce an *Xho*I restriction site at the 5' prime end and also to add 10bp of overlapping *nisI* sequence at the 3' prime end. The overlapping sequences added allowed the primers *nisI* F and *nisFEG* R to be used to generate the 2.7 kbp *nisIFEG* fragment using SOE PCR to amplify across both the fragments fusing them together (Fig 29a). This PCR resulted in multiple DNA fragments and therefore the correct size fragment was excised and purified from an agarose gel (Fig 29b). This fragment was then digested with the enzymes *Xho*I/*Nde*I and ligated into pMTL84123, linearised with the same enzymes, generating pMTL84123*nisIFEG* (Fig 29c). The ligated plasmid was chemically transformed into *E. coli* Top10, and the colonies were screened using colony PCR with *nisG R* RT-PCR/M13 R primers (Fig 29d). The plasmid was extracted from two positive colonies and digested with *XhoI/NdeI* restriction enzymes to determine if the correct size fragment was present. The insert was gel purified and Sanger sequenced using internal primers for each individual gene; *nisI* F RT-PCR/*nisI* R RT-PCR, *nisF* F RT-PCR/*nisF* R RT-PCR, *nisE* F RT-PCR/*nisE* R RT-PCR and, *nisG* F RT-PCR/*nisG* R RT-PCR. This confirmed plasmid was transformed into the conjugal donor strain, *E. coli* NEB SExpress (Woods *et al.*, 2019), before being conjugated into both *C. difficile* 630Δ*ermΔpyrE* and *C. difficile* R20291Δ*pyrE*. These two strains were selected as they are suitable for genes to be expressed chromosomally if this was necessary for future work rather than expressing the immunity genes from a plasmid. Putative plasmid carrying colonies from each strain were screened using M13 R/ *nisG* R RT-PCR primers which span across the

plasmid and *nisIFEG* region, producing a PCR amplicon of the expected size of ~2950 bp (Fig 29c). Thus, three individual plasmid carrying colonies from each strain were selected for further study.

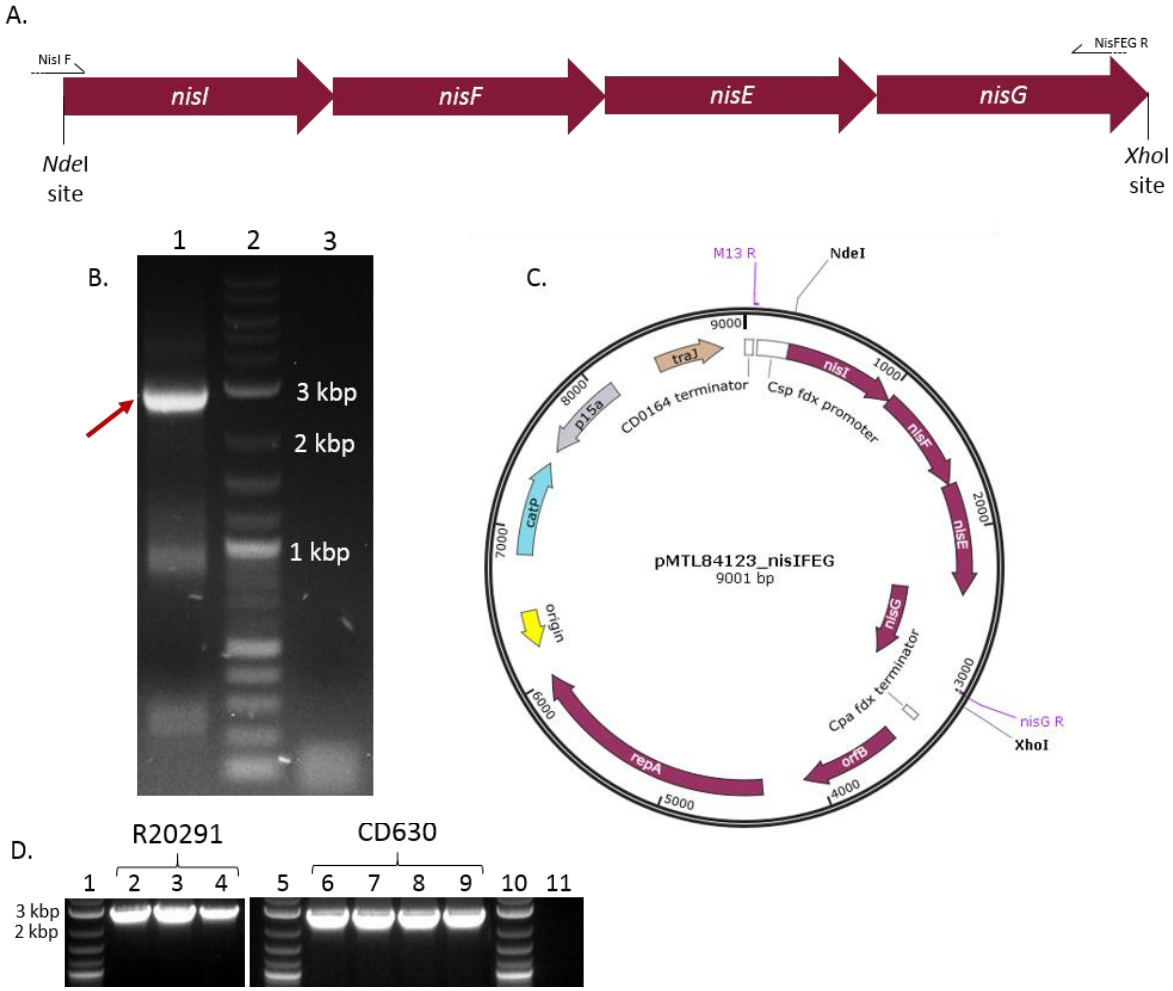


Figure 29- *nisIFEG* fragment, pMTL84123_*nisIFEG* plasmid map and conjugation PCR colony screen. A. the final *nisIFEG* product resulted from fusing *nisI* to *nisFEG* with overlapping primers nisI F and nisFEG R leaving restriction sites at either end for introduction into the pMTL84123 vector. B. Agarose gel picture of the *nisIFEG* SOE PCR using primers *nisI* F and *nisFEG* R, which resulted in multiple fragments shown in lane 1, with the negative no template control in lane 3 and a 2-log ladder in lane 2. The correct size fragment, approximately 2.7 kbp (red arrow) was excised from this gel and gel purified. C. pMTL84123_*nisIFEG* vector map, with the screening primers M13R and nisG R RT-PCR shown in purple and the restriction sites, *Nde*I/*Xho*I, used to incorporate *nisIFEG* in bold. D. Agarose gel picture of colony PCR screening for *C. difficile* R20291Δ*pyrE* and CD630Δ*ermΔpyrE* positive for pMTL84123_*nisIFEG*. Labelled lanes 2, 3, 4, 6, 7, 8, 9 represent individual colonies for each strain and the negative control (lane 11) is a no template control. A fragment at the expected size ~2950bp can be seen indicated by the 2-log ladder in lanes 1, 5, and 10.

Confirmation of the expression of all four genes was conducted using RT-PCR to amplify the individual genes expressed in the transcriptional mRNA unit in *C. difficile* R20291Δ*pyrE*_pMTL84123*nisIFEG*. Extracted RNA was confirmed through agarose gel visualising for RNA from wild-type *C. difficile* R20291 and three biological replicates of *C. difficile* R20291Δ*pyrE*_pMTL84123*nisIFEG* (Fig 30a). There were two RNA fragments observed on the agarose gel which do not correspond to the correct 23S and 16S rRNA sizes however this is probably due to the incorrect approximation of fragment size as a 2-log DNA ladder was used rather than an RNA. A lack of amplified product from a 16s rRNA PCR determined the lack of DNA contamination within the RNA samples (Fig 30) and the RNA was converted into cDNA. Primers amplifying the individual genes, *nisI* F RT-PCR/*nisI* R RT-PCR, *nisF* F RT-PCR/*nisF* R RT-PCR, *nisE* F RT-PCR/*nisE* R RT-PCR and *nisG* F RT-PCR/*nisG* R RT-PCR, were used to confirm the expression of all the genes within the mRNA unit (Fig 30a, data not shown for *nisEG*). The gDNA positive control (Fig 30b) showed the PCR should have successfully amplified the genes, with the no template control demonstrating this result was not a false positive (Fig 30b). However successful amplification of all the genes, *nisIFEG* was not observed across the three RNA samples. As one *C. difficile* R20291Δ*pyrE*_pMTL84123*nisIFEG* biological repeat possessed both *nisI* and *nisF* and the other two repeats possessed *nisF* (Fig 31a) all three repeats should be positive for all four genes as they are transcribed as one mRNA unit, however the PCR results did not reflect this. In order to demonstrate the nisin resistant strain expresses all the nisin resistant genes the extracted RNA was used as the template for an alternative RT-PCR to amplify across the whole mRNA transcription unit, using the primers *nisI* F RT-PCR/*nisG* R RT-PCR, instead of amplifying the individual genes (Fig 31e). This PCR successfully generated a product of the correct size, ~2.5 kbp, with the wild-type control showing no product and the no template control displaying

no PCR contamination (Fig 31e). This indicates that the *nisIFEG* operon is being expressed within the *C. difficile* strain carrying the plasmid.

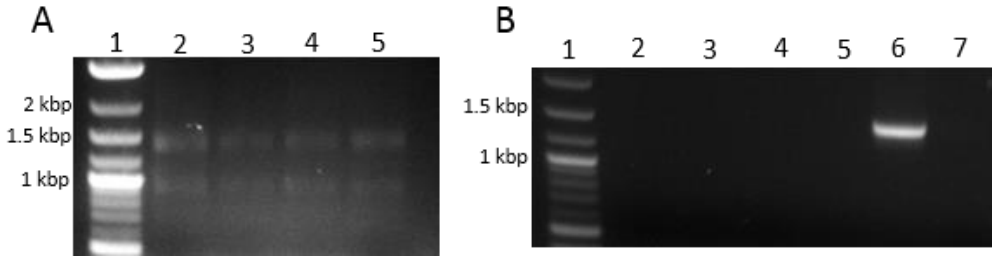


Figure 30- Agarose gel of RNA extracted from *C. difficile* wild-type R20291 and *C. difficile* R20291ΔpyrE_pMTL84123*nisIFEG* and 16s rRNA PCR using the RNA as template demonstrating a lack of DNA contamination. (a) Agarose gel of RNA extracted from wild-type R20291 (lane 2) and three biological replicates of *C. difficile* R20291Δ*pyrE*_pMTL84123*nisIFEG,* #1 (lane 3), #2 (lane 4) & #3 (lane 5). A 2-log ladder is in lane 1. (b) Agarose gel of 16s rRNA PCR using the RNA as a template. Lane 1: 2-log ladder, lane 2: RNA from wild-type R20291, lane 3-5: RNA from *C. difficile* R20291Δ*pyrE*_pMTL84123*nisIFEG* #1, #2, #3 respectively, lane 6: *B. pumilus* 1_16 gDNA, lane 7: no template control.

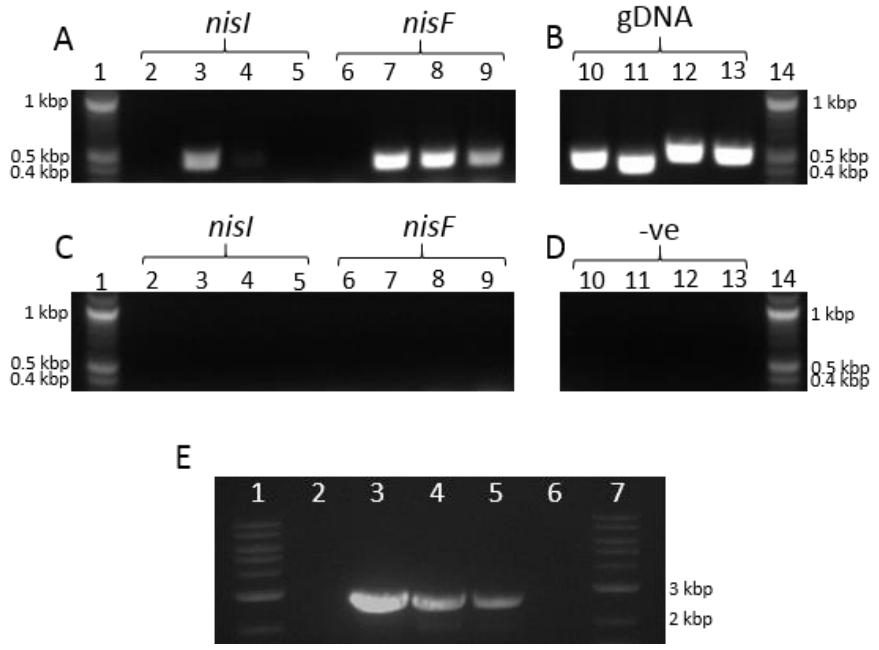


Figure 31- RT-PCR of cDNA extracted from three biological replicates of *C. difficile* R20291ΔpyrE_pMTL84123*nisIFEG* expressing the nisin immunity genes, *nisIFEG*, compared to the wild-type *C. difficile* R20291. (a) RT-PCR using cDNA template from wild-type R20291 (lanes 2 & 6) and three *C. difficile* R20291Δ*pyrE*_pMTL84123*nisIFEG* biological replicates #1 (lanes 3 & 7), #2 (lanes 4 & 8) and #3 (lanes 5 & 9). Lanes 2, 3, 4 & 5 show the PCR of *nisI*, with a successful amplification in lane 3. Lanes 6, 7, 8, & 9 show the amplification of *nisF* which is successful in lanes 7, 8, & 9. (b) Agarose gel of PCR, using *B. pumilus* 1_16 gDNA template, of genes *nisI, nisF, nisE,* and*, nisG* shown in lanes 10, 11, 12, and 13 respectively. There is a product in each lane ranging between 400-600 bp (c) Agarose gel of negative RT-PCR using RNA as a template from wild-type R20291 (lanes 2 & 6) and three *C. difficile* R20291Δ*pyrE*_pMTL84123*nisIFEG* biological replicates #1 (lanes 3 & 7), #2 (lanes 4 & 8) and #3 (lanes 5 & 9). Lanes 2, 3, 4 & 5 show the PCR of *nisI* and lanes 6, 7, 8, & 9 show the amplification of *nisF*, no product is observed in any lanes. (d) Agarose gel showing the no DNA template negative control for PCRs shown in figure A-C. (e) Agarose gel showing the RT-PCR of cDNA from wild-type R20291 (lane 2) *C. difficile* R20291Δ*pyrE*_pMTL84123*nisIFEG* #1 (lane 3), #2 (lane 4), #3 (lane5). There is successful amplification of product ~2.5 kbp in lanes 3, 4, and 5. A 2-log ladder is used in all agarose gels pictured in lanes (a) 1 (b) 14, (c) 1 (d) 14, and (e) 1 & 7.

3.2.5.2.2 Nisin Resistance of C. difficile Strains

Nisin resistance was examined using the agar well diffusion method, as shown in Stein et al. (2003). Different concentrations of nisin were tested against *C. difficile* 630Δ*ermΔpyrE*_pMTL84123*nisIFEG, C. difficile* R20291Δ*pyrE*_pMTL84123*nisIFEG* or the same strain carrying the empty plasmid pMTL84123 as a control. A zone of clearance around the well depicted sensitivity to the specific concentration of nisin, and the addition of *nisIFEG* resulted in an increase of resistance from 2 μg/ml of nisin to 8 μg/ml for both strains (Fig 32). The zones of clearance for both resistant strains were also smaller than the negative control and wild-type suggesting some level of protection at higher concentrations (Fig 32).

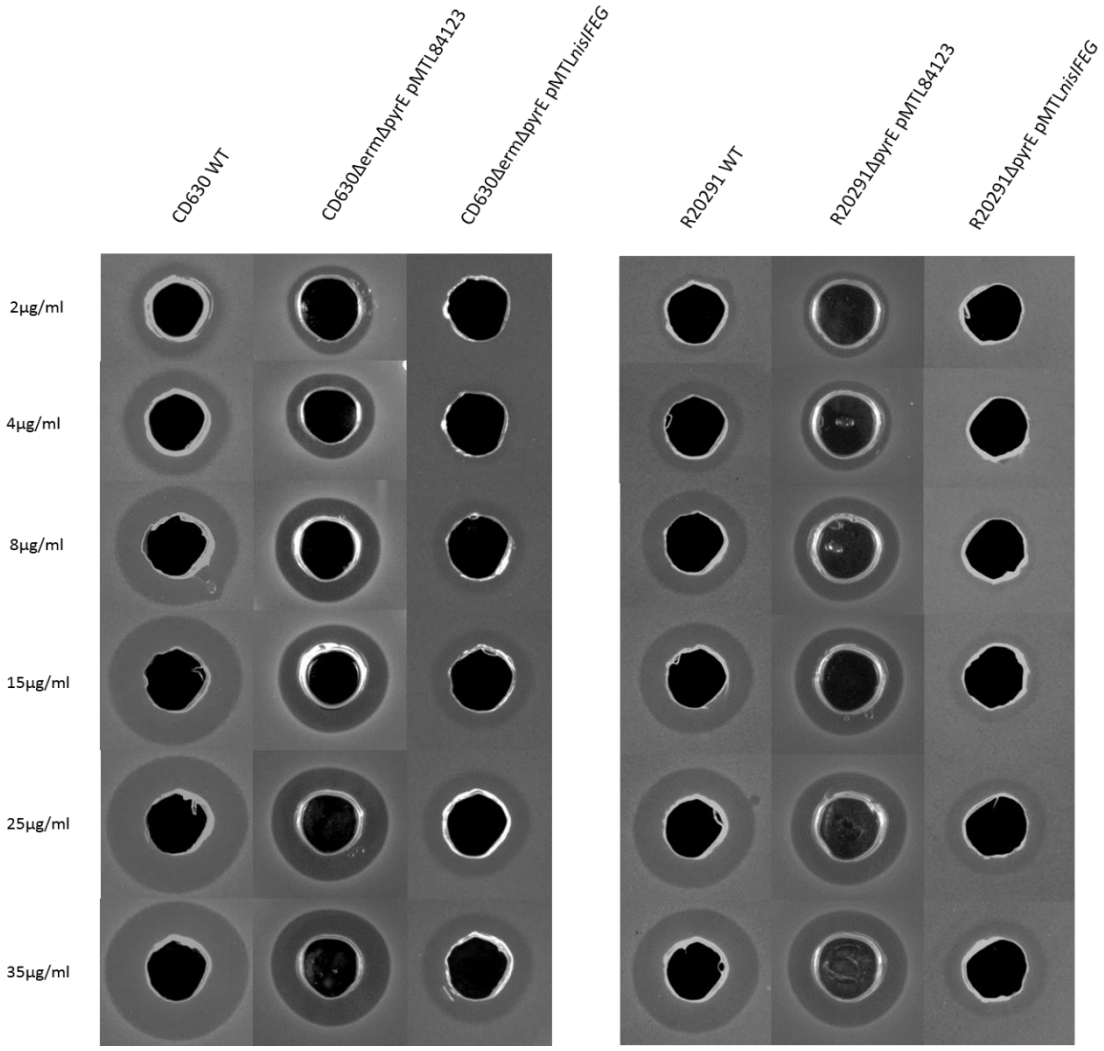


Figure 32-Increased levels of nisin resistance of *C. difficile* 630ΔermΔpyrE_pMTL84123*nisIFEG* and R20291ΔpyrE_pMTL84123*nisIFEG* compared to wild-type and negative control strains. A range of nisin concentrations, 2 µg/ml- 35 µg/ml were tested using well overlays against wild-type *C. difficile* CD630 and R20291 which demonstrated comparable sensitivity to the negative control strains *C. difficile* 630Δ*ermΔpyrE* and R20291Δ*pyrE*. *C. difficile* 630Δ*ermΔpyrE*_pMTL84123*nisIFEG* and R20291Δ*pyrE*_pMTL84123*nisIFEG* which were expressing the nisin immunity genes, *nisIFEG*, demonstrated higher levels of resistance against the same range of nisin concentrations. A zone of clearance for both resistant strains was only observed starting from 8 µg/ml whereas the wild-type and negative controls demonstrated sensitivity from 2 µg/ml. The observed zones of clearance for the two resistant strains were also smaller than that of the wild-type and negative control strains.

To determine whether nisin caused the zone of clearance and activity of *L. lactis* strain 10 against *C. difficile*, the two resistant strains were tested in an overlay. *L. lactis* strain 10 showed no inhibitory activity against the Nisin resistant strain, whereas the negative control harbouring the plasmid backbone was sensitive to *L. lactis* 10 (Fig 33). This suggests the activity was cause solely by the presence of nisin and not due to the other putative bacteriocins.

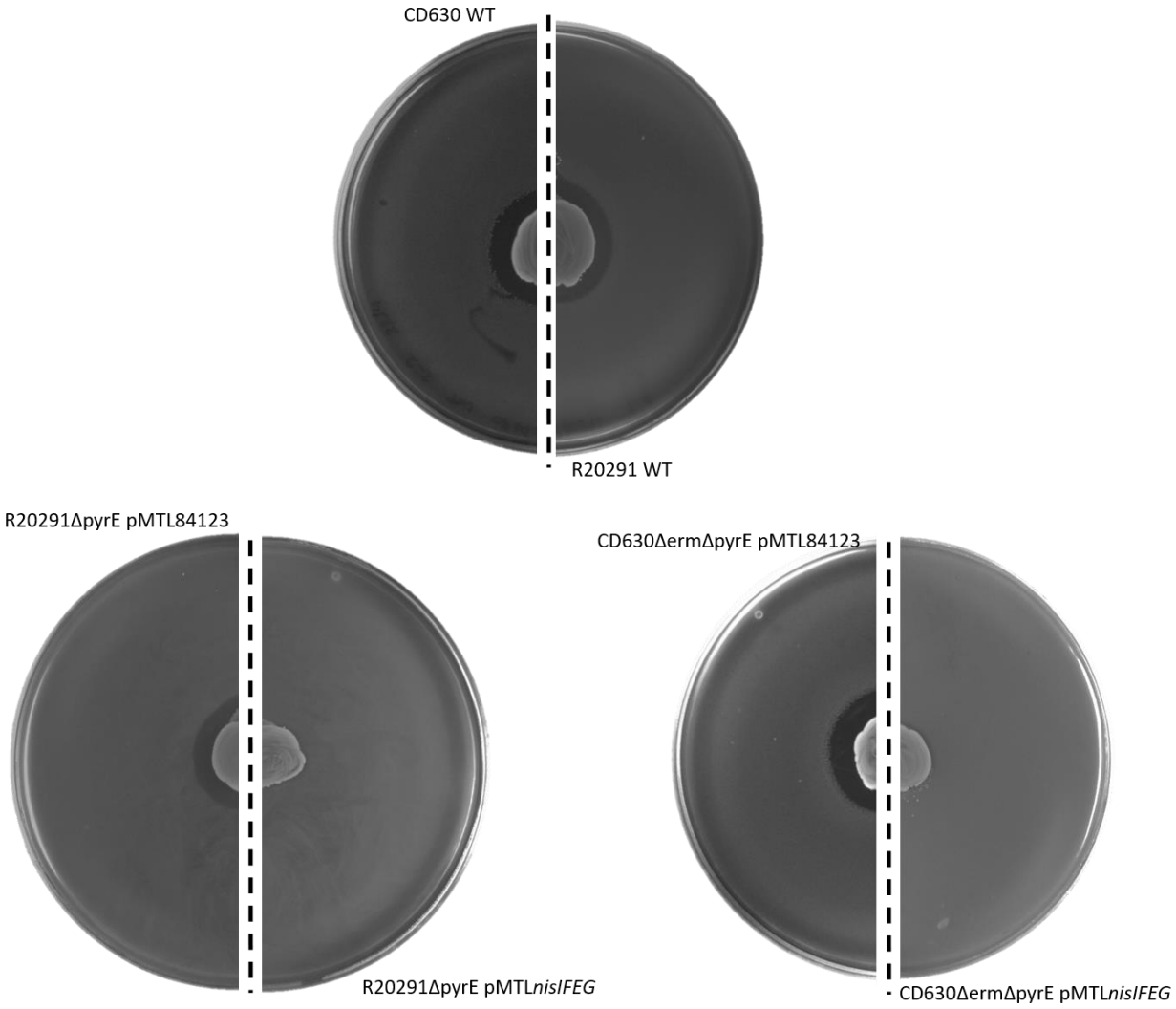


Figure 33- Colony overlay depicting conferred resistance against *L. lactis* 10 of *C. difficile* strains carrying pMTL84123*nisIFEG*. Comparative zone of clearances can be seen around the controls carrying the backbone pMTL84123 vector compared to the parental wild-type strain. No zone of clearance is observed in the two strains harbouring pMTL84123*nisIFEG*.

3.3 Discussion

3.3.1 Characterisation of Anti-*C. difficile* Activity Demonstrated by *L. lactis* 10

L. lactis strain 10 was previously isolated and showed activity against a few *C. difficile* strains with clinically relevant PCR ribotypes. After further testing *L. lactis* 10 demonstrated inhibition activity against all 15 strains of *C. difficile* tested which spanned a range of ten different PCR ribotypes including 012, 027, 078, and 017. As bacteriocins are secreted into the medium, the concentrated cell free supernatant was tested for activity, however the supernatant of *L. lactis* has a low pH and contains hydrogen peroxidase. Although the majority of inhibition activity was deemed to be due to the presence of hydrogen peroxidase, some inhibition remained evident and therefore the whole genome was sequenced to identify any putative bacteriocin candidates. Although two putative bacteriocins, BacE and BacL, were highlighted, the intact nisin biosynthesis pathway was also identified. Nisin is a well-known bacteriocin isolated from *L. lactis* and highly active against *C. difficile* (Le Lay *et al.*, 2016), thus it is the most likely cause of the observed antimicrobial activity. The presence of nisin could have been discovered earlier in the screening process, after 16s rRNA identification, if PCR had been used to search for previously characterised bacteriocins known to be produced by the same species. Although screening for known bacteriocins doesn't rule out the possibility of multiple bacteriocins present within the genome and without a positive control for each bacteriocin the PCR could give false negative results, it is a potentially useful tool to narrow down the screening strains and increase the chances of isolating novel bacteriocins. The activity seen against *C. difficile* is most likely due to nisin however the two other putative bacteriocins could be having a synergistic effect or be contributing to a second mechanism of *C. difficile* inhibition.

3.3.2 Inconclusive Evidence of Active BacE Peptide

In silico analysis of BacE showed homology to enterolysin A, which is a bacteriocin produced by *Enterococcus faecalis* LMG 2333. Enterolysin A is a cell wall-degrading bacteriocin which is active against a range of species including *Bacillus, Listeria, Staphylococcus, Lactococcus and Lactobacillus.* As this bacteriocin is typically produced by *E. faecalis* and is known to be active against strains of *L. lactis* it was unusual to identify it within the genome of *L. lactis* 10 however the limited homology suggest BacE is significantly different to enterolysin A (Khan *et al.*, 2013). The activity of Enterolysin A toward *C. difficile* remains unknown however it did not demonstrate activity towards other *Clostridium* strains (Nilsen *et al.*, 2003). . BacE shows significant homology to the catalytic N-terminus of enterolysin A (Nilsen *et al.*, 2003), including the conserved histidine residues which are thought to be the Zn^{2+} ligand binding residues. This, along with its categorisation as an M23 peptidase, suggested that BacE could have catalytic activity. BacE, however, significantly differs from enterolysin A at its C-terminal recognition domain, which is shorter than enterolysin A, suggesting a different target than the cell wall. The typical pI for cell wall targeting proteins is above 7, and the predicted pI of BacE is 5.3 (Schwartz *et al.*, 2001). This again suggests a different target for its putative catalytic activity which does not rule out its potential as a novel bacteriocin.

Overexpression of *bacE* was demonstrated in an *E. coli* host and was initially purified using nickle ion resin. However subsequent switching to zinc ion resulted in a less contaminated product. Zinc ion resin should have been the initial choice in resin due to the zinc metalloprotease nature of the peptide which was previously identified *in silico*. The purification of the BacE peptide, did not result in active peptide and showed unreliable activity on a native-PAGE gel. This could have been due to the smearing of the protein on the native-PAGE gel, a focused band of peptide would have

resulted in a more concentrated zone of peptide. The smear of peptide did generate a zone of clearance suggesting that the peptide was active against *C. difficile* however this result could not be repeated and was deemed unreliable, despite no such zone in the negative control. The usual cause of false positives for a protein gel overlay is due to the premature solidification of the soft top agar preventing the growth of *C. difficile* however this was not observed in this case. The zone of clearance does not fully match the protein bands visualised on the gel and extends up to the wells of the native-PAGE gel whereas no protein is observed here, however if silver staining or western blot were used instead it could have more accurately displayed the peptide banding. The negative control used was an empty well and whilst this demonstrates the native-PAGE gel does not contain any components that inhibit *C. difficile* growth, a superior negative control would have been the His-tag elution buffer. The unreliable results could be due to protein degradation or inactivation during the purification process and although steps such as using a non-denaturing conditions were used to prevent inactivation; the result could not be repeated reliably.

3.3.3 Expression of BacL in an *E. coli* Host

In silico analysis of BacL identified a lactococcin domain and showed homology with LcnB, a lactococci bacteriocin known for its narrow host range against other lactococci. The bacteriocin cluster also contained a self-immunity gene (*baiL*) downstream of *bacL* reflecting organisation homology to the previously described LcnB operon (van Belkum *et al.*, 1992), suggesting *L. lactis* 10 has the potential to express and self-protect against this putative lactococcin bacteriocin. The presence of a lactococcin bacteriocin within the genome of *L. lactis* 10 is quite typical as multiple lactococcins have been characterised from *L. lactis* strains (Holo *et al.*, 1991; van

Belkum *et al.*, 1992; Moll *et al.*, 1996; Zendo *et al.*, 2006; Diep *et al.*, 2007; Ishibashi *et al.*, 2015).

E. coli host expression was a valid option to express BacL as there have been previous bacteriocin structural genes successfully expressed, however there are known difficulties such as discrepancies in expression between host strains and toxicity towards the host (Mesa-Pereira *et al.*, 2018). The activity of some bacteriocins, including nisin, is dependent on the functioning of additional gene products which are involved in transportation or peptide modification. In the case of BacL, the encoding operon contains only one extra gene, *baiL*, the putative self-immunity gene suggesting host machinery is utilised for its processing and secretion. Although *bacL* was successfully cloned into a *E. coli* expression system there was no evidence of over-expression in both the soluble and insoluble fractions. Although different induction conditions were trialled, these could have been carried out in a less time-consuming manner through probing the cell lysates using dot blotting which could more rapidly identify the presence of any His-tagged peptide across more samples. Although there is a possibility of false positives it would narrow down the different conditions that would need more in-depth examination. It is known that expression levels vary between *E. coli* expression hosts (Mesa-Pereira *et al.*, 2018) therefore trialling multiple hosts increases the chance of obtaining overexpressed protein, and although both hosts used in this study (C43 (DE3) pLysS and Rosetta-gami 2) did not result in protein expression additional hosts could have been used. Trialling multiple hosts with multiple growth and induction conditions is an extremely time-consuming process, therefore using a more rapid probing system such as dot blotting would be extremely beneficial. Previous studies have reported successful expression of bacteriocins from only their structural genes, such as LcnB, (Kawai *et al.*, 2003; Jasniewski *et al.*, 2008; Chen *et al.*, 2012; Sandiford and Upton, 2012), however other lactococcins, such as

lactococcin A, are known to require transporter and modification peptides(Stoddard *et al.*, 1992). With this is mind the ORFs surrounding *bacL* could be investigated for any role using both *in silico* and *in vitro* techniques.

Although the role of BacL was not ascertained in this study it the *in silico* evidence highlights it an an unlikely candidate for the antimicrobial activity observed against *C. difficile.* However, this does not rule out its potential as a novel lactococcal bacteriocin, which could be tested for using nisin resistant lactococcal strains. Any lactococcal strain expressing nisin are inherently resistant to nisin activity due to the co-expression of the immunity genes. Therefore, any inhibition caused by *L. lactis* 10 could be shown to originate from a second bacteriocin. As the aim of this study was to discover novel bacteriocins against *C. difficile*, no further work was carried out on BacL however generating a *L. lactis* 10 strain with nisin and BacE knocked out would provide a platform for further elucidation of its role as a putative lactococcin.

3.3.4 Evidence that Nisin is Responsible for the Observed Antimicrobial Activity Against *C. difficile*

In silico analysis shows that the intact nisin biosynthesis pathway is present in the genome of *L. lactis* 10. Several amino acid changes are present within NisT, NisP and NisG, whose effects could be investigated further however the biosynthesis process of nisin has been elucidated. Further investigations might not result in any significant discoveries. Since historical evidence of nisin activity has been repeatedly demonstrated (Le Lay *et al.*, 2016) it is reasonable to hypothesise that the majority, if not all, the activity of *L. lactis* 10 observed against *C. difficile* is due to nisin Z.

The most conclusive way to determine if nisin Z was solely responsible for the observed activity against *C. difficile* was to generate a knock-out mutant. *L. lactis* has been historically a good model for genetic engineering, in part, due to its readiness to

accept foreign DNA, however electro-transformation of *L. lactis* 10 resulted in too low a transformation efficiency for genetic manipulation. This could be due to the presence of a type III restriction modification system, as the endonuclease encoding gene *LlaFI* was identified *in silico* (Su *et al.*, 1999). Previously this gene was shown to incompletely digest foreign DNA, however knocking out this gene could increase the transformation efficiency of *L. lactis* 10. Optimising the electro-competent cells and transformation protocol could improve the transformation efficiency for example using cell wall weakening agents (Holo and Nes, 1995; Papagianni *et al.*, 2007) or optimising cell density and growth medium (Gerber and Solioz, 2007). Optimisation of the protocol was not undertaken due to the positive results generated through expressing the nisin immunity genes in *C. difficile*, however in order to provide evidence for the inactivity of BacE and BacL the knock-out mutants should be generated. As well as knocking out the putative restriction system in *L. lactis* 10,

The strategy that successfully showed that nisin was the responsible candidate for the antimicrobial activity of *L. lactis* 10 against *C. difficile* was expressing the nisin immunity genes, *nisIFEG*, in two strains of *C. difficile,* CD630 and R20291. RT-PCR demonstrated the expresson of *nisIFEG*, which conferred increased resistance to nisin compared with the wild-type and empty plasmid control. This increased resistance, up to 8 µg/ml, corresponds with previous observations from Stein *et al.* (2003), who expressed *nisIFEG* in a *B. subtilis* host. An overlay of the *C. difficile* resistant strains against *L. lactis* 10 demonstrated no zone of clearance, whereas there was a large zone of clearance against the wild-type and the negative controls, suggesting that nisin was solely responsible for the observed antimicrobial activity. It is interesting to note that there was a difference in sensitivity between wild-type *C. difficile* CD630 and R20291 with the latter demonstrating higher resistance to nisin. Further work to confirm this difference through liquid assays could increase our understanding of the difference in

nisin sensitivity between *C. difficile* strains. From the data obtained in this study there is not a readily observable pattern of nisin sensitivity associated with specific PCR ribotypes, however more strains need to be screened in order to conclusively demonstrate this or to observe a pattern of resistance. As there are multiple causes of nisin resistance it would be difficult to identify a single cause however screening a large number of strains combined with *in silico* comparisons could help elucidate any conserved factors (Zhou *et al.*, 2014). There has been recent work in using bioinformatic tools to study microevolution in *C. difficile* pertaining to transmission pathways between animals and humans, and this resulted in the development of core genomes for certain PCR ribotypes of *C. difficile* (Knight and Riley, 2019). If these strains were screened for nisin sensitivity, then *in silico* single nucleotide variant analysis could provide insight into the phenotypic differences of nisin resistance across different *C. difficile* strains. As nisin has previously been suggested as a candidate for use against *C. difficile* (Le Lay *et al.*, 2016) investigating this inherent sensitivity difference could provide insight into significant treatment factors such as strain dependent dosing or potentially using nisin in conjunction with other antimicrobials to minimise failure due to inherent resistance.

LAB strains are highly recommended as bacteriocin production strains due to their GRAS status and positive publicity; however, many do contain known bacteriocins such as nisin. In order to discover novel bacteriocins, screening strains that are resistant to known bacteriocins would be extremely useful. In order to investigate the role of four genes in nisin resistances, Stein et al. (2003) expressed those genes in a previously sensitive *B. subtilis* strain thus generating a nisin resistant strain. This study demonstrated the ability to generate two nisin resistant *C. difficile* strains through the same principle, expressing the genes on a plasmid. Incorporating these genes into the chromosome would generate stable nisin resistant strains which could be used as

screening strains for novel bacteriocins against *C. difficile*. As there is a need for new antimicrobial treatments against *C. difficile* and bacteriocins are a viable option discovering novel bacteriocins is important. Therefore, by increasing the speed at which novel bacteriocins are discovered, through ruling out known bacteriocins early in the screening process, more novel bacteriocins could be discovered.

Take home points

- Nisin is a known bacteriocin with effective antimicrobial activity against *C. difficile* and its known immunity genes can be successfully expressed in *C. difficile* hosts providing resistance to nisin.
- Utilising the bacteriocin immunity genes and expressing them in sensitive hosts is a viable approach to determine the bacteriocins role in antimicrobial activity

Chapter 4: Isolation of Environmental Organisms from Cow Manure with Antimicrobial activity against *C. difficile*

4.1 Introduction

4.1.1 Aims of this Project

The objectives of the work presented in this chapter was to develop a screening method for isolating environmental organisms with antimicrobial activity against *C. difficile*. This was based on the hypothesis that cow manure is a potential reservoir for environmental organisms harboring novel bacteriocins with activity against *C. difficile.* Any strain successfully isolated would be further characterized by genetic, phenotypic and *in silico* methods for the discovery of putative novel bacteriocin candidates.

4.1.2 Reservoirs of *C. difficile*

As *C. difficile* is known to inhabit a range of environments such as soil, water, food, and animals, there are multiple potential reservoirs for community acquired *C. difficile* infections (CA-CDI). There are specific ribotypes of *C. difficile* that are associated with different hosts e.g. humans (RT002 and RT014), pigs (RT078, RT150 and RT045) and cattle (RT033), however some strains are known to have multiple hosts (Janezic *et al.*, 2014; Rodriguez Diaz *et al.*, 2018). Previously it was hypothesised that CDI should be considered a zoonotic or anthroponotic disease, as its presence in animals plays a role in the continued maintenance of the infection in nature. However, the difference in prevalence of binary-toxin positive strains suggested that animals and humans had distinct populations of *C. difficile*. Over time the prevalence of binary-toxin strains in humans has increased and is often associated with both HA- and CA-

CDI (Rupnik *et al.*, 2003; Terhes *et al.*, 2004; Barbut *et al.*, 2005; Rupnik, 2007). Further evidence of the zoonotic nature of *C. difficile* was observed in the hypervirulent RT027 strains, a ribotype originally found in animals and in human CA-CDI cases but rapidly disseminated within hospital systems across North America and Europe (Rodriguez-Palacios *et al.*, 2006; He *et al.*, 2013; Knight and Riley, 2019). Another hypervirulent *C. difficile* ribotype, RT078, has also been shown to be capable of transmission between animals and humans (Knetsch *et al.*, 2014, 2018). RT078 has previously been isolated from both animals and humans and is emerging as an important human pathogen as a cause of both HA- and CA-CDI. A potential reservoir for this ribotype is pigs, as genotypically and phenotypically similar isolates have previously been found to be carried by both pig farmers and their pigs (Keessen *et al.*, 2013). The hypervirulent RT078 lineages were found to be predominately on pig farms in multiple studies (Keel *et al.*, 2007; Goorhuis *et al.*, 2008; Debast *et al.*, 2009; Keessen *et al.*, 2013; Wu *et al.*, 2016). A gene substitution within RT078 strains that conferred moxifloxacin resistance in pig isolates was also noted in human isolates strains further suggesting a transmission between humans and pigs (Wu *et al.*, 2016). More recent studies have confirmed the genetic similarities between animal and human RT078 strains and using, whole genome phylogeny, provided evidence of the frequent transmission of RT078s between animals and humans on a global scale (Knetsch *et al.*, 2018). There is also evidence of strains closely related to RT078s, having zoonotic potential as well (Knight *et al.*, 2019). The use of *in silico* techniques to decipher lineages of *C. difficile* has recently shown that, across a variety of ribotypes, the core genomes of *C. difficile* strains isolated from humans and animals did not form distinct populations. Indicating that the disease is passing between the two and is therefore zoonotic (Knight and Riley, 2019).

4.1.3 Cattle as a Reservoir for *C. difficile*

C. difficile is found within the gastrointestinal tract of cattle, with the first report in 1974 (Hafiz, 1974), and have been studied due to their potential as a reservoir for CA-CDI. Multiple PCR ribotypes of *C. difficile* have been isolated from cattle, including ribotypes found in human CDI such as RT003, RT002 and RT078; the latter being commonly detected across different European countries (Pirs *et al.*, 2008; Avbersek *et al.*, 2009; Hoffer *et al.*, 2010; Koene *et al.*, 2012; Rodriguez *et al.*, 2012, 2013; Romano *et al.*, 2012; Zidaric *et al.*, 2012; Schmid *et al.*, 2013; Schneeberg *et al.*, 2013). Cattle have therefore been hypothesised to be a source of CA-CDI, through contact with contaminated animals and their products including meat, carcasses and manure (Rodriguez-Palacios *et al.*, 2007; Rupnik, 2007; Rodriguez *et al.*, 2013). Evidence of cross-contamination between livestock and humans has been previously demonstrated with an increased carriage rate (25-14%) in people with direct contact with piglets (Keessen *et al.*, 2013), livestock known for its high prevalence of *C. difficile*, compared to the normal <5% carriage rate of non-hospitalised adults (Loo *et al.*, 2011). One study showed evidence against cross-contamination between cattle and cattle farmers however the study was limited by its small sample population (n = 5). The authors suggested that a larger scale study, throughout the year, was more appropriate for the intermittent asymptomatic excretion of *C. difficile* in humans (Rodriguez *et al.*, 2017). Although there is evidence of a persistent *C. difficile* reservoir in cattle, there are differences amongst the different breeds of cattle and type of farming practice. Breeding and rearing Simmental cattle were five times more likely to shed *C. difficile* than Holstein-Friesian cattle (Bandelj *et al.*, 2016). In one study, which controlled for farm practice and cattle maintenance, it was noted that one farm accounted for over 50% of the *C. difficile* strains isolated, and it was hypothesised that this prevalence could be attributed to the Limousin breed, which are a non-docile

constantly stressed breed (Rodriguez *et al.*, 2017). Chronic stress in piglets can lead to physiological changes which favour *C. difficile* colonisation including alterations in the intestinal barrier and receptor expression (Li *et al.*, 2017). Thus, a similar stress induced increase of *C. difficile* could be associated with the known anxiousness of the Limousin breed. This association of *C. difficile* prevalence and stress has previously been observed, as calves were twice as likely to be positive for *C. difficile* one week after arriving in the feedlots compared to later weeks (Costa *et al.*, 2011). At this early stage calves have to undergo many stresses inducing changes such as separation from their mother, dehorning, earring implantation, vaccinations, transportation and feed changes (Rodriguez *et al.*, 2017).

Another factor affecting the prevalence of *C. difficile* in cattle is age which has repeatedly been observed amongst pigs and cattle (Rodriguez-Palacios *et al.*, 2006; Weese *et al.*, 2010; Costa *et al.*, 2011; Rodriguez *et al.*, 2012; Bandelj *et al.*, 2016; Grześkowiak *et al.*, 2016). Veal calves have a higher incidence of *C. difficile* on arrival to farms compared with the prevalence at slaughter (Costa *et al.*, 2011; Knight *et al.*, 2013). In one study calves tested within 24 h of birth were positive for *C. difficile* and demonstrated shedding in their faeces (Bandelj *et al.*, 2018). These new-born calves had the highest prevalence, which decreased over time (Bandelj *et al.*, 2018) in agreement with Costa *et al*. (2011) where recent arrivals had a higher prevalence (32%) of *C. difficile,* when the calves were between 2-10 days old, compared to a 2% positive rate after 21 weeks. This age-related effect has also been demonstrated in both humans (Rupnik *et al.*, 2009) and piglets with the latter demonstrating a high prevalence at day 2 of age, with 74% of piglets colonised, compared to 56% at day 7, 40% at day 30 and 3.7% at day 62 (Weese *et al.*, 2010). The reasons and source of colonisation remain unknown (Rodriguez *et al.*, 2017) as although perinatal sows have been suggested to carry *C. difficile* it is at a very low level, yet there is evidence of

colonisation of neonatal piglets, suggesting an extremely short window for colonisation (Grześkowiak *et al.*, 2016). Although there is some limited evidence of an environmental spread in manure or via vermin there is currently no clear widespread evidence for this route of environmental spread or of a maternal origin (Weese *et al.*, 2010; Burt *et al.*, 2012; Grześkowiak *et al.*, 2016; Usui *et al.*, 2017). It has been suggested that the higher prevalence in neonatal piglets is due to the underdeveloped intestinal microflora which becomes more colonisation resistant as it develops with age (Rodriguez-Palacios *et al.*, 2006). This could also be the case for calves, in addition to the stress due to birth, different feeding and management systems or a naïve immune response (Bandelj *et al.*, 2018). The age-related effect of *C. difficile* colonisation is an important risk-factor as, although in North America slaughter of veal calves can be up to 21 weeks, in Australia it occurs at a much younger age (<7 days of age) and thus the higher prevalence of *C. difficile* could be a risk for foodborne contamination (Knight *et al.*, 2013). There is also the possibility of transmitting *C. difficile* between farms through the trading of colonised calves. Trading calves older than three weeks has been suggested to limit the spread of *C. difficile* to other farms and thus possible dissemination to the wider community (Costa *et al.*, 2012). However some young animals colonised with *C. difficile* do not show symptoms of disease, which has previously been observed in human neonates (Lees *et al.*, 2016). It has been hypothesised that a lack of disease is due to the lack of *C. difficile* toxin receptors in the gastrointestinal tract, protection from breast milk, or protection from the infant gut microflora (Pant *et al.*, 2013; Lees *et al.*, 2016); these could also be possible reasons for the lack of *C. difficile* disease shown in some animal neonates.

The management of livestock has previously been suggested to be associated with the occurrence of *C. difficile* in farm animals. In humans a risk-factor of CDI is the use of antibiotics, which impair the normal gut microflora allowing *C. difficile* to

colonise. Although well known in humans, there is a lack of information for an association between the use of antimicrobials and shedding of *C. difficile* in cattle, despite the routine use of antibiotics within the farming industry. Two studies have demonstrated a lack of association between antimicrobial use and *C. difficile* shedding in cattle, however both suggested the need for further studies linking the two (Rodriguez-Palacios *et al.*, 2011; Costa *et al.*, 2012). Another management factor is the scale of the farming practice. Multiple studies have been carried out on cattle on large scale intensive farms, which is only one type of industrial cattle farming. Bandelj *et al*. (2016) studied a smaller mid-size family farm, more representative of European agriculture, and determined that 39.8% of cattle (calves and cows) were positive, throughout the year, for *C. difficile.* Although the prevalence was higher than reported for large scale intensive farming, the screening method was more sensitive and there was a difference in ribotype (Bandelj *et al.*, 2016). On large scale cattle farms, one of the most prevalent ribotypes is RT078 (Keel *et al.*, 2007; Costa *et al.*, 2011; Rodriguez-Palacios *et al.*, 2011; Zidaric *et al.*, 2012) where as Bandelj *et al*. (2016) reported most prevalent ribotype as RT033, (Zidaric *et al.*, 2012). The difference in ribotypes could be dynamic, changing with time, as studies in 2017 have shown different patterns of ribotypes not observed in studies from 2012 and 2013 (Rodriguez *et al.*, 2012, 2013, 2017). It was hypothesized that the source of the bacterium in cattle was the environment and that continued epidemiological studies need to monitor new emerging *C. difficile* types (Rodriguez *et al.*, 2017). This could also be significant as there is increasing evidence of frequent transmission of *C. difficile* between animals and humans (Knight and Riley, 2019), implying that the monitoring of cattle farms could provide insights into emerging CA-CDI strains that could disseminate into hospital systems. This monitoring could be trialled with the RT033 which is known to be prevalent in cattle. Increasing reports have linked RT033 with human CDI, indicating

its potential as an emerging ribotype (Grandesso *et al.*, 2016; McGovern *et al.*, 2017; Krutova *et al.*, 2018). Although there are discrepancies in prevalence and ribotypes reported on cattle farms it has repeatedly been demonstrated that there is a persistent *C. difficile* reservoir in cattle.

4.1.4 The Microflora of Cattle and its Potential as a Reservoir for Bacteriocin Harbouring Bacteria

The gastrointestinal microbiota of bovines is made up of a diverse mix of protozoa, archaea, bacteria, and fungi which is reflected in the faecal microflora. The faecal microflora of dairy cattle contains a number of different phyla, with the majority of microflora (91%) belonging to the Firmicutes and Bacteroidetes phyla (Hagey *et al.*, 2019). It has been reported that *Clostridium, Porphyromonas, Bacteroides, Ruminococcus, Alistipes, Lachnospira, Prevotella,* and *Streptococcus* are all highly abundant genera in dairy cattle faeces (Dowd *et al.*, 2008; Shanks *et al.*, 2011). Ruminant bacteria have been shown to commonly produce bacteriocin like inhibitory substances that confer both a narrow and broad spectrum of activity (Kalmokoff *et al.*, 1996). It has been suggested that the bacterial population within the bovine gastrointestinal environment are partially regulated through the production of antimicrobials including bacteriocins (Kalmokoff *et al.*, 1996, 2003; Russell and Mantovani, 2002). Bacteriocins could enhance the survival of the producer organisms, inhibiting bacteria competing for the same resources. Since the bovine microflora is such a diverse environment, with a plethora of different bacterial species, it could be an abundant reservoir of bacteriocins with a diverse range of activity (Cookson *et al.*, 2004). . Currently multiple ruminant bacterial species have been shown to produce bacteriocins such as *Streptococcus gallolyticus* (Pieterse *et al.*, 2008)*, S. bovis* (Whitford *et al.*, 2001)*, Butyrivibrio fibrisolvens* (Kalmokoff *et al.*, 1996)*, Ruminococcus albus* (Odenyo *et al.*, 1994) and *Enterococcus faecium* (Morovský *et al.*, 2001) which indicates

that cattle are a reservoir of bacteriocins (Cookson *et al.*, 2004). As there is evidence that cattle are a reservoir of *C. difficile* it is reasonable to hypothesise that cattle might also be a reservoir for organisms producing bacteriocins with activity against *C. difficile*.

4.2 Results

4.2.1 Isolation of Environmental Organism Active Against *C. difficile*

Fresh bovine faeces was collected from a mid-scale fully automated dairy farm in the East Midlands, UK. All faecal matter was collected from either the barn floor or the slurry tank which also contains infected milk or milk contaminated with antibiotics.

Samples were homogenized and the suspension was plated on to rich medium (Fructose agar) grown at 37 ^{o}C both aerobically and anaerobically to broaden the variety of species screened. Ethanol shock step, to kill all vegetative cells, selecting for spore-forming anaerobic bacteria was also used (Rea *et al.*, 2010). A minimum of 50 colonies were picked from each condition and were screened for activity against *C. difficile* CD630 via overlaying with inoculated soft-top agar. Any colonies which showed a zone of clearance against *C. difficile,* after 12 h incubation, were picked for further screening against two more *C. difficile* strains, DH1916 and R20291, to confirm antimicrobial activity against other clinically relevant strains. Isolates active against at least two *C. difficile* strains were investigated further. In total, out of 200 screened isolates, 36 were re-tested for activity and 16 were repeatedly active against *C. difficile* CD630. 11 of these isolates showed activity against at least two out of the three tested *C. difficile* strains (Table 15).

Table 15- The Antimicrobial Activity of Different Environmental Samples Against Three Strains of *C. difficile*.

Sample name	*C. difficile* strain		
	CD630	DH1916	R20291
30AN[1]	+	+	-
1_16A[2]	+	+	-
1_29A	+	+	-
2_16A	+	+	-
2_36A	+	+	-
2_46A	+	+	-
2_49A	+	+	-
8ANE[3]	+	+	-
33ANE	+	-	+
41ANE	+	+	+
42ANE	+	+	+

[1]AN-anaerobic growth, [2]A- aerobic growth, [3]ANE- anaerobic growth after a 50% ethanol wash

4.2.2 Characterization of the Isolated Environmental Organisms

As each bacterial sample was originally obtained from an environmental source, they underwent a process of purification (Fig 34) to isolate the bacteria responsible for the observed antimicrobial activity. This process involved streaking the sample to isolate single colony morphologies, verified using Gram-staining (Fig 34). Once the Gram-stain showed a single cellular morphology the isolate underwent 16s rRNA PCR for further identification (Fig 35). 16s rRNA PCR screening indicated most isolates samples were *Bacillus* spp*.* (Table 16) and one isolate of each unique species (30ANiii, 1_16A, 8ANE, 1_29A, 41ANE) underwent further investigation (indicated in green, Table 16). Culture 30AN was found to be a mix of four different colony morphologies which were each purified and identified, three of which were *Proteus* sp. and one *Streptococcus* sp. (Table 17). The *Proteus* sp. showed variable cellular morphology under Gram-stain suggesting it might have been still contaminated with the *Streptococcus* sp.. Only the *Streptococcus* sp. (30ANiii) demonstrated antimicrobial activity against *C. difficile*, however it is interesting to note that as a pure culture its antimicrobial range increased from not only *C. difficile* CD630 and DH1916 but also R20291 (Table 17). The *Streptococcus* sp. was also selected for further downstream investigation.

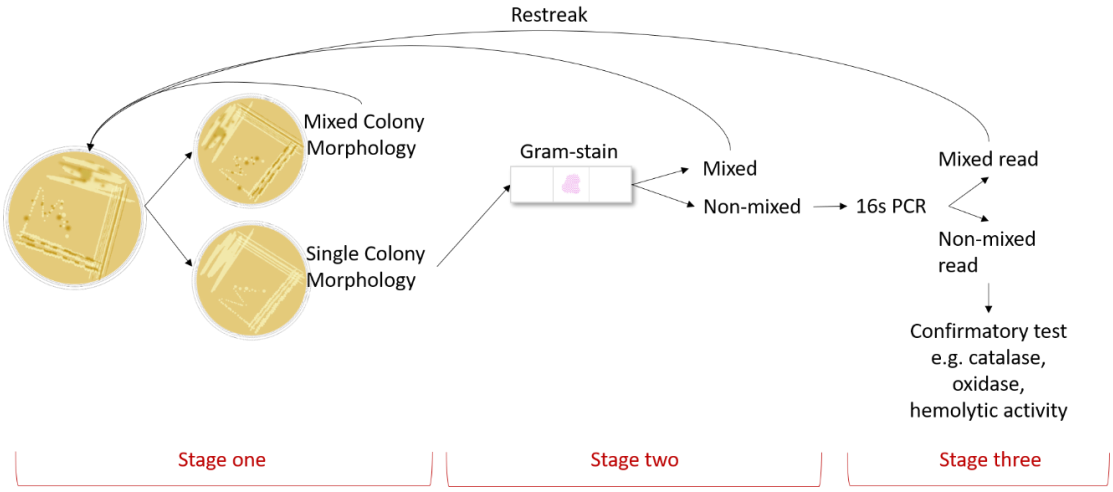


Figure 34- Protocol for purifying and identifying single bacterial isolates from a mixed culture. In the first stage of the purification process of potential mixed culture was to streak on to separate agar plates to obtain single colonies. The second stage involved is Gram-staining the potentially pure culture to verify a single cell morphology. This then underwent the third stage of the process, 16s rRNA PCR to attempt an identification of the bacterial species. The identification of the isolate was confirmed using secondary confirmatory tests if necessary. If the culture appeared mixed at any stage then the process was repeated from stage one.

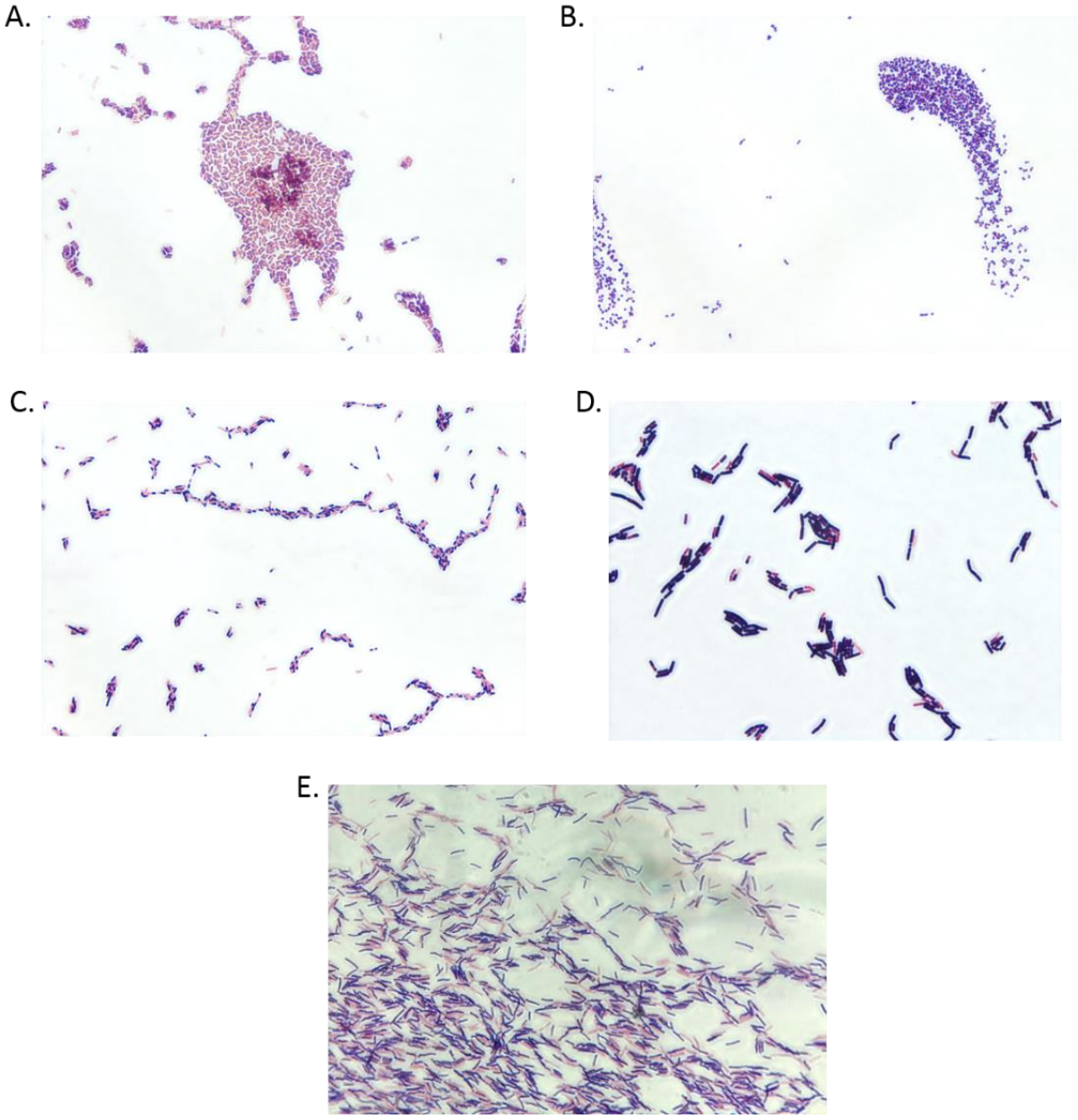


Figure 35- Gram-stains of environmental samples, 1_29A, 8ANE, 1_16A, 41ANE and 30ANiii. X1000 magnification of Gram stain of cultures (A) 1_16A Gram-positive bacilli (B) 30ANiii Gram-positive cocci, (C) 41ANE Gram-positive bacilli, (D) 1_29A Gram-positive bacilli and (E) 8ANE Gram-positive bacilli.

Table 16- The Gram-Stain, Cell Morphology and 16s rRNA PCR Results for the Environmental Organisms Demonstrating Antimicrobial Activity Against *C. difficile*.

Culture name	Gram-stain	Morphology	Initial 16s rRNA identification
1_16A	+	Bacilli	*Bacillus pumilus*
1_29A	+	Bacilli	*Oceanobacillus sp.*
2_16A	+	Bacilli	*Bacillus pumilus*
2_36A	+	Bacilli	*Bacillus pumilus*
2_46A	+	Bacilli	*Bacillus pumilus*
2_49A	+	Bacilli	*Bacillus pumilus.*
8ANE	+	Bacilli	*Bacillus thermoamylovorans*
33ANE	Variable	Bacilli	*Bacillus licheniformis*
41ANE	Variable	Bacilli	*Bacillus licheniformis*
42ANE	Variable	Bacilli	*Bacillus licheniformis*

Table 17- The Gram-Stain, Cell Morphology and 16s rRNA PCR Results for the Different Colony Morphologies Identified in Culture 30AN and their Antimicrobial Activity Against *C. difficile*.

Culture 30AN	Gram stain	Morphology	Initial 16s rRNA identification
i	-	Cocci	*Proteus* sp
ii	Variable	Bacilli	*Proteus* sp
iii	+	Cocci	*Streptococcus* sp
iv	-	Cocci	*Proteus* sp
Culture 30AN	Activity		
	CD630	DH1916	R20291
i	-	-	-
ii	-	-	-
iii	+	+	+
iv	-	-	-

4.2.3 Identification of Isolates Determined to Require Further Investigation

Five isolates, *B. pumilus* 1_16A, *Oceanobacillus* 1_29A, *B. thermoamylovorans* 8ANE, *B. licheniformis* 41ANE, and *Streptococcus* 30ANiii were identified from the screening process as having antimicrobial activity against multiple strains of *C. difficile*. *B. pumilus* 1_16 underwent further characterization and investigation discussed in Chapter 5. The other four isolates underwent additional characterization of their antimicrobial activity as well as *in silico* whole genome analysis to identify potential bacteriocin candidate(s) responsible for their activity. These results of these investigations highlighted the isolates warranting further investigation.

4.2.3.1 Further Characterization of Antimicrobial Activity

In order to determine if the observed antimicrobial activity was due to a secreted peptide(s), the supernatant of the environmental strains was tested for activity. Due to the unknown nature of the produced antimicrobial activity, several growth conditions were tested which included using different media and stress inducers positive stimuli to increase the likelihood of inducing antimicrobial activity (Table 18).

Table 18- Antimicrobial Activity Observed from Concentrated or Stress Induced Cell Growth of the Isolated Environmental Isolates

Test		Observed Antimicrobial Activity			
		B. thermoamylovorans 8ANE	*B. licheniformis* 41ANE	*Oceanobacillus* spp. 1_29	*Streptococcus* 30ANiii
< 10 kDa Supernatant	BHIS	-	-	-	-
	TSB	-	-	NT	NT
	BM	Not Tested (NT)	+	NT	NT
	MM	NT	NT	-	NT
> 10 kDa Supernatant	BHIS	-	-	-	-
	TSB	-	-	NT	NT
	BM	NT	NT	NT	NT
	MM	NT	NT	-	NT
Mitomycin C		-	-	-	NT
UV irradiation	16 h	-	NT	-	-
	1:50 dilution	-	NT	-	-
Plaque Assay		NT	-	NT	NT

4.2.3.1.1 B. thermoamylovorans 8ANE

A zone of clearance was previously observed when a colony of *B. thermoamylovorans* 8ANE, grown at 37 °C on BHIS medium, was overlaid with *C. difficile*. Therefore the supernatant was tested for antimicrobial activity after growth at 37 °C and not at the optimal 50 °C. In order to observe any antimicrobial activity from the supernatant of *B. thermoamylovorans* 8ANE, 5 ml of sterile filtered supernatant from a 16 h culture grown in two different types of media, BHIS and TSB, with and without stress inducers, mitomycin C and UV. The supernatant was concentrated using a 10 kDa spin column and both fractions, >10 kDa and <10 kDa, were then spot plate overlaid. However, no activity against *C. difficile* CD630 was observed (Table 18).

4.2.3.1.2 B. licheniformis 41ANE

Antimicrobial activity was observed from this isolate from a colony overlay after growth on BHIS medium at 37 °C. Cell free supernatant was obtained from a 16 h culture in BHIS medium and was concentrated using a 10 kDa spin column. Both the >10 kDa and <10 kDa fractions were negative for antimicrobial activity. The isolate was then grown in two different types of media, BM and TSB for 16 h and supernatants tested for antimicrobial activity after concentration. There was a faint zone of clearance in the <10 kDa fraction from cells grown in BM however none was observed in supernatant of cells grown in the TSB medium (Table 18) (data not shown). To induce further activity mitomycin C was added at either early exponential (0.2 OD_{600nm}) or mid-exponential (0.4 OD_{600nm}) (Fig 36) and was incubated overnight or for 2 h before the cells were pelleted and resuspended in fresh media. However, the cells appeared to have lysed after the addition of mitomycin C and therefore the supernatant was not tested for antimicrobial activity. As mitomycin C is also used to induce phage lysis from

cells, a phage plaque assay was carried out to determine if the observed inhibition of *C. difficile* was due to lytic phage infection. Cell free supernatant was obtained from a 16 h culture grown in BHIS that was induced with mitomycin C until complete cell lysis. Cell free supernatant was also collected from a 16 h culture of *B. licheniformis* 41ANE grown in BM which appeared to have undergone cell lysis as previously observed. These were both tested for plaque formation against *C. difficile* however there were no observable plaques (data not shown) which suggest there is no phage or no phage infection of *C. difficile*. Any further experimental work on *B. licheniformis* 41ANE was hindered due to a lack of cell viability during non-stressed growth.

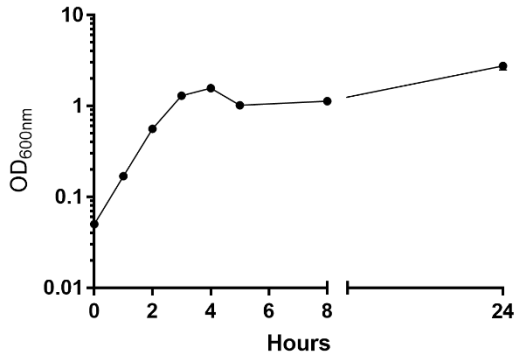


Figure 36- Growth curve of *B. licheniformis* 41ANE grown in BHIS medium shaking at 37 °C. This data reflects three technical repeats of three independent biological repeats.

4.2.3.1.3 Oceanobacillus 1_29A

Antimicrobial activity against *C. difficile* was observed from *Oceanobacillus* 1_29A after a colony was grown aerobically at 37 °C on BHIS medium. BHIS broth was therefore used to grow *Oceanobacillus* 1_29A to obtain cell free supernatant from a 16 h culture. This supernatant was concentrated using a 10 kDa spin column and both the >10 kDa and <10 kDa fractions were negative for antimicrobial activity (Table 18). As growth in BHIS did not result in observed antimicrobial activity the optimal medium for *Oceanobacillus* spp., Marine Medium (MM), was investigated. 3h after inoculation

the isolate was stressed through the addition of mitomycin C for either a 16 h incubation or 2 h incubation after which the cells were resuspended into fresh medium and incubated for 16 h. Another form of stress induction was UV irradiation after which cells were inoculated into fresh MM at a 1:50 dilution and both the original UV irradiated culture and inoculated culture were incubated for 16 h at 37°C. Both >10 kDa and <10 kDa concentrated cell free supernatant fractions from all stress conditions demonstrated no antimicrobial activity (Table 18).

4.2.3.1.4 Streptococcus 30ANiii

Streptococcus 30ANiii demonstrated antimicrobial activity against *C. difficile* when it was grown anaerobically at 37°C on BHIS medium. As *Streptococcus* spp. are facultative anaerobes both aerobic and anaerobic growth conditions were used to obtain cell free BHIS supernatant. The supernatant was concentrated using a 10 kDa spin column but both the >10 kDa and <10 kDa fractions demonstrated no antimicrobial activity in either condition (data not shown). The isolate was then stressed using UV irradiation as described in 4.2.3.1.3 and the concentrated supernatant demonstrated no antimicrobial activity (Table 18) (data not shown)

4.2.3.2 *In silico* Mining for Known and Putative Bacteriocins within the Genomes of the Environmental Isolates

Each environmental isolate underwent *de novo* whole genome sequencing using the illumina HiSeq 2500 platform. The generated paired reads from each strain were assembled into contigs using Mira, which were then annotated for genes using the RAST platform (Brettin *et al.*, 2015). The genomes were then mined for known and putative bacteriocins using BAGEL4 (van Heel *et al.*, 2018) and bactibase (Hammami *et al.*, 2010) which highlighted significant regions.

4.2.3.2.1 B. thermoamylovorans 8ANE and Oceanobacillus 1_29A

Bacteriocin mining of the *Oceanobacillus* 1_29A and *B. thermoamylovorans* 8ANE genomes revealed no known or putative bacteriocins present. It is a possibility that there is no homology to known bacteriocins and that these strains still carry putative novel bacteriocins that remain unidentified and not in either database. One method that has been utilized to identify novel lantibiotics, other than BAGEL4, is to screen for sequence homology using the more conserved peptides within the different types of lantibiotic classes (Begley *et al.*, 2009; Singh and Sareen, 2014). Due to the lack of putative bacteriocins identified in the genomes of *Oceanobacillus* 1_29A and *B. thermoamylovorans*, other mechanisms for the antimicrobial activity were investigated. In order to identify any secondary metabolites or phage that could be responsible, the databases antiSMASH and PHASTER were used, however neither highlighted any putative cause for the observed antimicrobial activity.

4.2.3.2.2 B. licheniformis 41ANE

In silico analysis of the genome of *B. licheniformis* 41ANE revealed the presence of a known two-component lantibiotic, lichenicidin VK, bacteriocin locus present in *B. licheniformis* ATCC 14580 which consisted of modification, processing, transportation and structural bacteriocin genes. This locus was also detected in *B. licheniformis* VK21 and DSM13 as part of a conserved region. The same region is also present in *B. licheniformis* 41ANE which contains both the structural genes, *lchA1* and *lchA2*, in the typical orientation each with their downstream *lchM1* and *lchM2* genes (Fig 37). The amino acid sequences for LchA1 and LchA2 in *B. licheniformis* 41ANE are identical to the prepeptides produced by *B. licheniformis* VK21, which are responsible for the mature peptides Lchα and Lchβ, respectively. The rest of the locus is also present within the *B. licheniformis* 41ANE genome with 99% identity to the lichenicidin locus within *B. licheniformis* ATCC 14580 with all the same genes present (Fig 37). The

presence of this intact region suggests that this lantibiotic is a likely candidate for the antimicrobial activity against *C. difficile*. Although there is no published evidence that this lantibiotic has been tested on *C. difficile*, it is known to be active against Gram-positive bacteria.

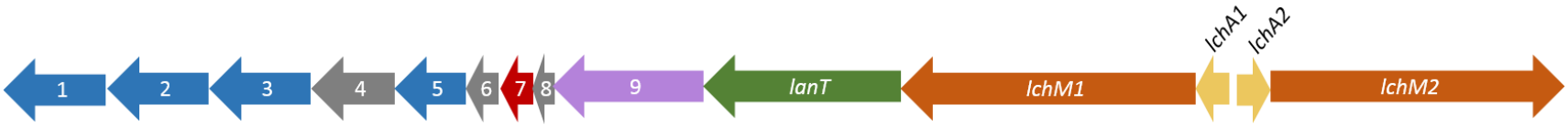


Figure 37-The genetic organisation of the predicted lichencidin bacteriocin within *B. licheniformis* 41ANE. Genes present which relate to ABC immunity and transport (blue), transcriptional regulation (red), hypothetical proteins (grey), proteases (purple), transportation (green), peptide modification (orange) and bacteriocin prepeptides (yellow). The predicted function of the encoded peptides based on sequence homology to *B. licheniformis* ATCC 14580 are 1. Permease, 2. ATPase, 3. Bacteriocin immunity, 4. No predicted function, 5. ATPase, 6. No predicted function, 7. Transcriptional regulator, 8. No predicted function, 9. Serine protease. *lanT* (green) is a ABC-type bacteriocin exporter, *lchM1* and *lchM2* (orange) encode LanM type modification enzymes, and *lchA1* and *lchA2* (yellow) encode the prepeptides for Lchα and Lchβ, respectively.

4.2.3.2.3 Streptococcus 30ANiii

In silico mining of the *Streptococcus* 30ANiii genome for known and putative bacteriocins using BAGEL4 highlighted one significant region within the whole genome sequence. The highlighted region spans over 20 kbp and within this region there are genes identified with similarities to bacteriocins or bacteriocin related genes. These genes are, one known and one putative bacteriocin (bov255 and BacS, respectively), three putative immunity genes (*caax1*, *caax2*, and *bai255*), one putative transportation gene (*lagD*) and one processing gene (*lcnD*) (Fig 38). *bacS* encodes a putative bacteriocin with a predicted ComC family domain. ComC is a family of competence proteins which also includes peptides that form bacteriocin-like propeptides that typically have a glycine-glycine cleavage site. The two genes downstream of *bacS*, *caax1* and *caax2* contain CPBP intramembrane metalloprotease domains (PF02517) and were annotated by RAST (Brettin *et al.*, 2015) as CAAX-self immunity metalloproteases (Abi). There is little known about this family of peptides in prokaryotes however Abi genes are known to be found in bacteriocin loci and have been demonstrated to be involved in bacteriocin self-immunity (Kjos *et al.*, 2010). The conserved Abi motifs within the self-immunity peptides were also found to be essential for the observed immunity function which suggested immunity via a proteolytic function (Kjos *et al.*, 2010).



Figure 38- The organisation of the bacteriocin cluster identified within *Streptococcus* 30ANiii. The bacteriocin locus consists of two bacteriocin structural genes (green), one putative, *bacS*, and one known, *bov255*. Three bacteriocin self-immunity genes were identified (orange), two Abi genes (*caax1* and *caax2*) and one EntA domain containing gene (*baiI255*). *lcnD* and *lagD* (yellow) are in the opposite orientation to the rest of the genes and are two known essential bacteriocin genes, encoding a secretory accessory peptide and ABC transporter, respectively. A single open reading frame was identified upstream of *caax2* however it encodes a hypothetical peptide (grey) with no known homology. Two terminators (loops) were predicted using ARNold (Naville *et al.*, 2011), after the hypothetical peptide (grey) and the *bai255* immunity gene.

Downstream of the putative self-immunity genes is a gene annotated as encoding a hypothetical protein with no significant homology to any known proteins or any identified sequence domains therefore, the role of this gene, if expressed, is currently unknown. More is known about the downstream *lcnD* gene which is a part of the 'Secretion protein MesE/LcnD' interpro family (IPR005696) and has over 98% similarity to a bacteriocin secretion accessory protein from *S. equinus.*. LcnD is a known secretion accessory protein found in an operon with an ABC transporter, LcnC, and upstream of the leucocin A bacteriocin structural gene (LcnA). It was demonstrated that LcnD and LcnC form a type I sec independent secretion system which is essential for the secretion and maturation of LcnA (Holo *et al.*, 1991). The function of other members of the MesE/LcnD secretion accessory family remains unclear in Gram-positive bacteria, however they are known to be required for bacteriocin secretion (Varcamonti *et al.*, 2001). Downstream of *lcnD* is *lagD*, which has over 98% similarity to an ABC-type bacteriocin transporter protein from *S. equinus*. The encoded peptide contains three domains, two ABC transporter domains (ABC membrane and ABC transporter domain) and a third peptidase domain, peptidase C39 (PF03412). The ABC membrane domain (IPR011527) is the transmembrane structural unit of the transporter and the ABC transport domain (IPR003439) is the ATP-binding region of the ABC transporter. ABC transporters are known to be the maturation proteases for some bacteriocins and a peptidase C39 domain is typically found within these peptides (Rawlings and Barrett, 1995). However the catalytic residues are not always conserved as come peptides have be reclassified as non-peptidase homologues due to this (Havarstein *et al.*, 1995).

Bovicin 255 has some similarity to Lactococcin A (27.7% ID over 86% of the amino acid sequence), which is a LAB non-pediocin like single peptide bacteriocin (Holo *et al.*, 1991). Additionally bovicin 255 was suggested to be a class II bacteriocin due to

it being a small heat stable cationic peptide which is processed at a glycine doublet (Whitford *et al.*, 2001). The biosynthesis of Lactococcin A requires four genes, *lcnC*, *lcnD*, *lcnA*, and *lciA* which encode the ABC transporter, secretion accessory proteins, the structural peptide and immunity gene, respectively. These four genes are all encoded in the same orientation (Stoddard *et al.*, 1992). The layout of the bacteriocin cluster within *S. infantarius* 30ANiii reflects the lactococcin A biosynthesis pathway. *lcnC* (lagD) and *lcnD* are downstream of *bov255* and a putative immunity gene, *bai255*. In contrast to the lactococcin A pathways *lcnC* and *lcnD* are encoded in the opposite orientation. The putative immunity gene, *bai255* carries an *entA* domain, a class IIc enterocin A immunity protein, which is also carried by other class II bacteriocins that possess immunity peptides such as LciA (Venema *et al.*, 1994). The EntA family of immunity peptides typically use their C-terminal domain to recognize and interact with the bacteriocin altering its orientation within the bacterial cell membrane, closing the pore that is the usual mode of action for their associated bacteriocin (Soliman *et al.*, 2007).

4.2.3.3 Selection of Organism for Further Investigation

Although there was evidence of antimicrobial activity from *B. thermoamylovorans* 8ANE and *Oceanobacillus* 1_29A, no putative bacteriocin candidates were identified from the *in silico* analysis. There was also no historical evidence of bacteriocins produced from these organisms inferring that it is unlikely that a bacteriocin was responsible for the activity. Although *in silico* analysis does not rule out the possibility of a bacteriocin, and protein precipitation and purification from the supernatant could identify a peptide candidate. The preliminary results of the spent supernatant experiments suggest that the antimicrobial compound is not secreted into the supernatant at a concentration that could be readily purified.

Therefore, these two isolates, *B. thermoamylovorans* 8ANE and *Oceanobacillus* 1_29A, were not selected for further study (Table 19).

Although *B. licheniformis* 41ANE did demonstrate evidence of an active antimicrobial compound within its supernatant (data not shown) the preliminary *in silico* analysis only highlighted a known bacteriocin, lichenicidin VK21, as the probable candidate. Although activity against *C. difficile* has not been previously demonstrated, there is evidence that lichenicidin VK21 has broad-spectrum activity against Gram-positive bacteria. This isolate was therefore not further investigated due to the lack of novelty (Table 19).

In silico analysis of *Streptococcus* 30ANiii also identified a known bacteriocin, Bov255, as being the most likely candidate for the observed antimicrobial activity, however it also highlighted a putative uncharacterised bacteriocin, BacS. Bov255 has not been well characterised and previously demonstrated no activity against two clostridial organisms, *C. perfringens* and *C. clostridiforme*. This, along with the presence of a second putative bacteriocin candidate, indicated that this strain could carry a novel bacteriocin active against *C. difficile*. Although preliminary results showed no evidence of antimicrobial activity demonstrated in the supernatant, the *in silico* evidence of a bacteriocin locus was robust enough to warrant further investigation of *Streptococcus* 30ANiii (Table 19).

Table 19-The results of the selection of environmental organisms for further investigation

Organism	Zone of clearance	Supernatant activity	Bacteriocin mining	Selection for further investigation	Other reasons for outcome
B. thermoamylovorans 8ANE	✓	X	X	X	
B. licheniformis 41ANE	✓	✓	✓	X	Lack of novel bacteriocin candidate and repeated cell lysis
Oceanobacillus 1_29A	✓	X	X	X	
Streptococcus 30ANiii	✓	X	✓	✓	

4.2.4 Further Characterization of *Streptococcus* 30ANiii

In order to characterize *Streptococcus* 30ANiii and determine its most recent ancestor, the 16s rRNA sequence from the whole genome sequence was aligned against the 16s rRNA nucleotide NCBI database and a phylogenetic tree was generated using the 16s rRNA sequences of the matched strains (Fig 39). This phylogenetic tree demonstrated that the 16s rRNA of *Streptococcus* 30ANiii clusters within the *S. equinus*/*S. bovis* complex (Fig 39, appendix 7.2). As group D streptococci are known to be non-hemolytic, *Streptococcus* 30ANiii was streaked onto horse blood agar plates to confirm whether it belonged to this streptococcal group. No hemolysis on blood agar was observed (data not shown).

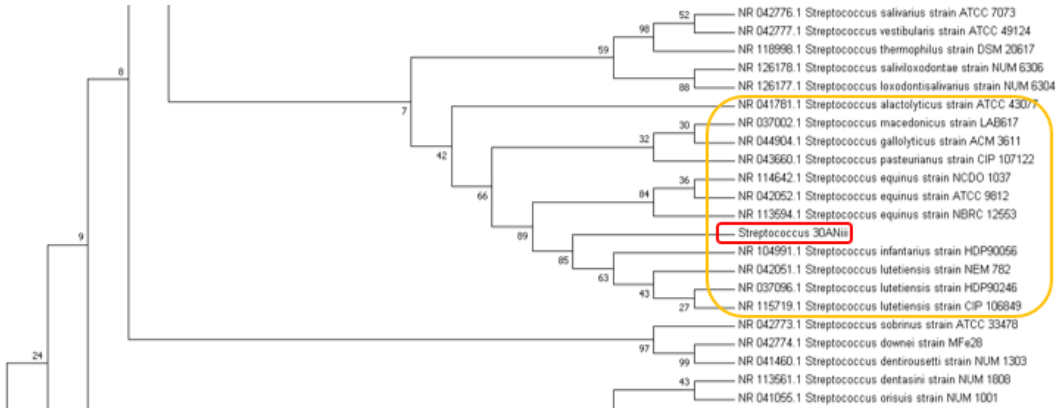


Figure 39- 16s rRNA phylogenetic tree of *Streptococcus* 30ANiii against other streptococci. Maximum likelihood tree was generated by MEGA7 (Kumar *et al.*, 2016) showing the 16s rRNA sequence from *Streptococcus* 30ANiii and the 89 closest matches using the 16s rRNA nucleotide NCBI database. Phylogenetic analysis was bootstrapped 500 times and is rooted using an *Enterococcus* sp. outgroup (not shown). The yellow square indicates the *S. bovis/S.equinus* complex with *Streptococcus* 30ANiii falling within this group highlighted in the red box. The full tree is shown in appendix 7.2.

A growth curve and CFU/ml was carried out to determine the exponential phase of growth in BHIS media incubated at 37 °C anaerobically (Fig 40). To determine if there was an active component secreted into the culture supernatant, cell free supernatant from mid-exponential phase *Streptococcus* 30ANiii (OD_{600nm} 0.4) was inoculated with *C. difficile* and cultures incubated for 24 h. Using mid-exponential phase supernatant resulted in no inhibition of *C. difficile* growth across the three strains tested, *C. difficile* CD630 (Fig 41a), DH1916 (Fig 41b) and R20291 (Fig 41c).

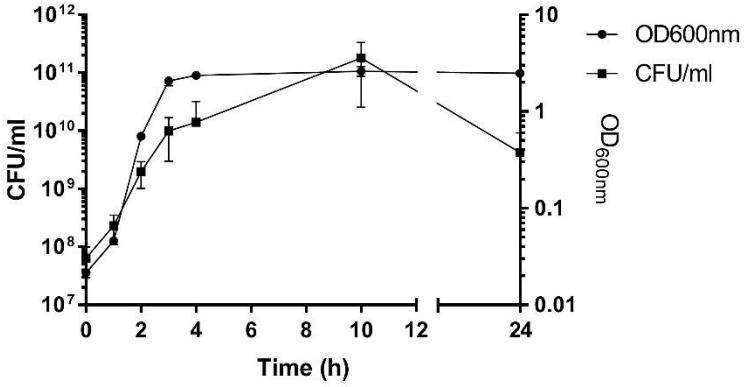


Figure 40- Growth curve and CFU/ml of *Streptococcus* 30ANiii. The growth curve and CFU/ml count were carried out at 37°C under anaerobic non-shaking conditions in BHIS medium over 24 h. The OD_{600nm} is aligned to the right Y axis and the CFU/ml count is on the left Y axis. The data shows the mean and error bars represent standard deviation from three technical repeats for each of the three biological repeats

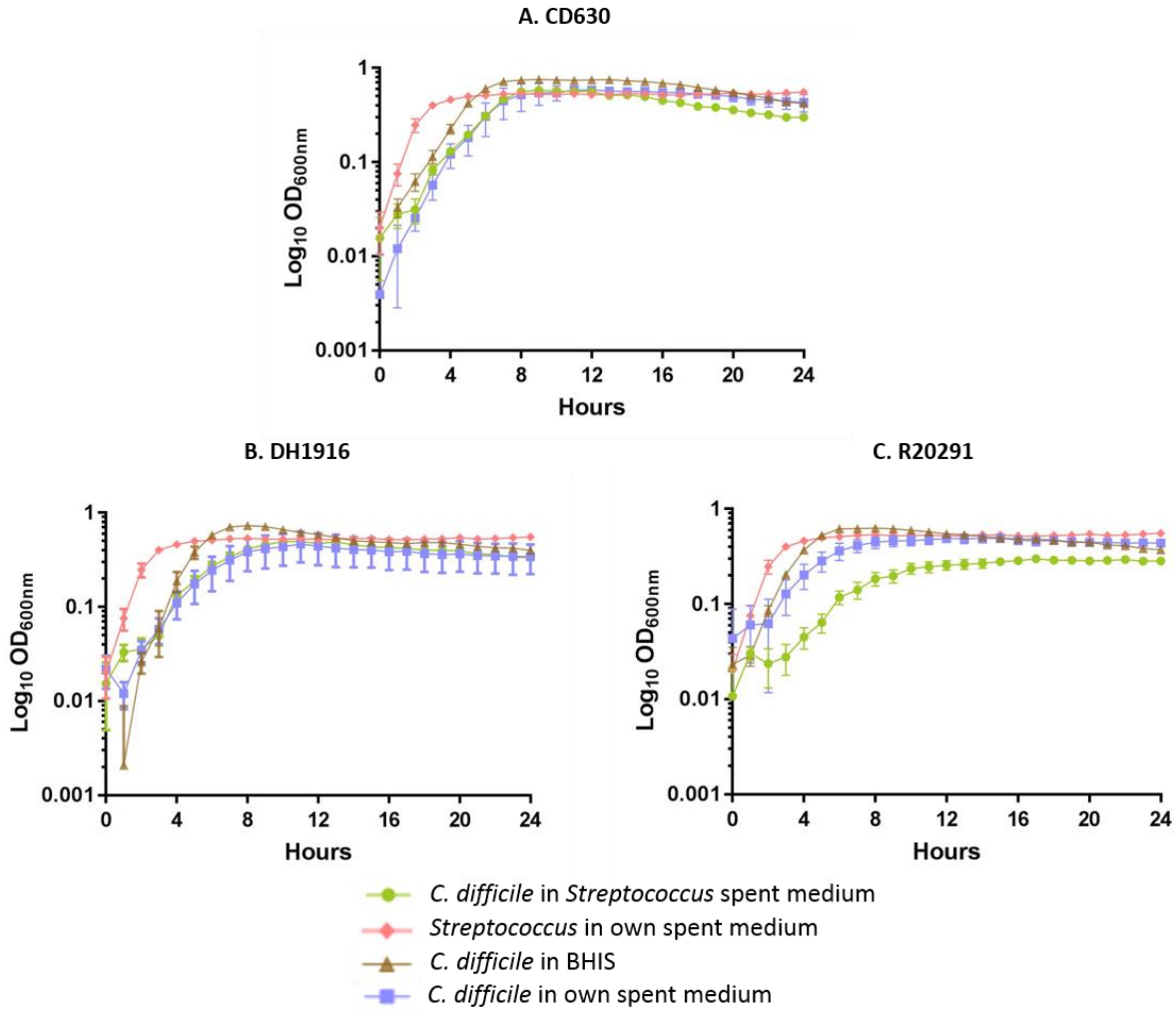


Figure 41- *Streptococcus* 30ANiii spent supernatant assay against *C. difficile* CD630, DH1916, and R20291. The lack of inhibition of *C. difficile* CD630 (A), DH1916 (B), R20291 (C) in *Streptococcus* 30AN mid exponential phase spent supernatant (green) can be seen compared to the non-inhibited growth of the controls *C.difficile* cultures grown in fresh BHIS, their its own spent medium and *Streptococcus* 30ANiii grown it its own spent medium (brown, blue and pink respectively). The OD_{600nm} readings were carried out at 37°C anaerobically, shaking before each reading on a 96-well microplate reader over a 24 h period. The data shows the mean and errors bars represent the standard deviation taken from 9 repeats which include 3 independent biological replicates each with 3 technical repeats.

As bovicin 255 was previously purified from cultures after overnight growth, the growth phase at which it was produced and secreted was not identified (Whitford *et al.*, 2001). Thus, the negative results (observed in Fig 41) could simply be due to the lack of bacteriocin expression at that growth phase. To test this hypothesis supernatant was collected from a later stage of growth however a lack of nutrients in the medium prevented normal growth of the controls in the subsequent assays (data not shown). Early stationary phase *C. difficile* CD630 and *Streptococcus* 30ANiii

supernatant (OD_{600nm} 2-2.5) were therefore lyophilized and rehydrated resulting in 2-fold concentrated spent supernatant. This, concentrated and nutrient rich, spent supernatant was used in place of the mid-exponential phase spent supernatant previously used. The concentrated spent supernatant resulted in slower growth of *C. difficile* in its own spent supernatant, compared to fresh BHIS medium, however by the 24 and 48 h time point it was at comparable levels whereas there was a complete inhibition of growth of *C. difficile* in the *Streptococcus* 30ANiii spent supernatant (Fig 42). This demonstrates that there is a compound secreted by *Streptococcus* 30ANiii that inhibits the growth of *C. difficile* which is either expressed starting from the early stationary growth phase or was concentrated sufficiently to be inhibitory after freeze-drying and resuspending in half the original volume.

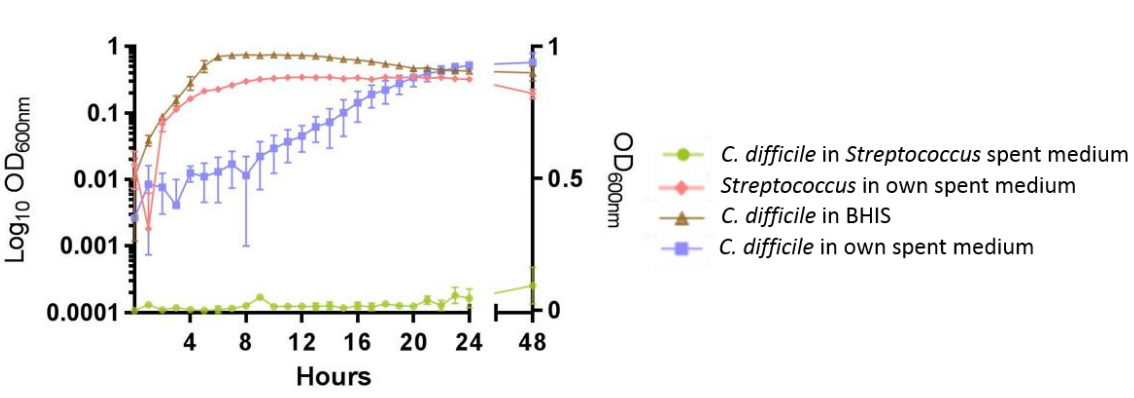


Figure 42- *Streptococcus* 30ANiii freeze-dried spent supernatant assay against *C. difficile* CD630. The lack of *C. difficile* CD630 growth in concentrated early stationary phase *Streptococcus* 30ANiii spent supernatant is shown in green. The controls, *C. difficile* grown in fresh BHIS, its own spent medium and *Streptococcus* 30ANiii grown it its own spent medium are shown in brown, blue and pink, respectively. The OD_{600nm} readings were carried out at 37°C anaerobically, shaking before each reading on a 96-well microplate reader over a 24 h period. The data shows the mean and error bars represent the standard deviation taken from six repeats which include two biological replicates each with three technical repeats.

4.2.5 Determining Whether the Known Bacteriocin, Bov255 and/or the Putative Bacteriocin Candidate, BacS, Play a Role in the Observed Anti-*C. difficile* Activity

In order to determine whether Bov255 and/or BacS play a role in the previously observed antimicrobial activity against *C. difficile* it was pertinent to obtain partially purified peptide which could be tested for activity. There are multiple different ways to purify bacteriocins, one involves protein precipitation of the native peptide from culture supernatant using ammonium sulphate and another one is plasmid based overexpression of the recombinant peptide in *E. coli*

4.2.5.1 Protein Precipitation of the Active Compound from the Supernatant of *Streptococcus* 30ANiii

As there is a known bacteriocin, Bov255, within the genome of *Streptococcus* 30ANiii that has been previously successfully purified from supernatant, the same protocol was followed in this study (Whitford *et al.*, 2001). *Streptococcus* 30ANiii was grown for 16 h at 37 °C anaerobically in BHIS broth before the bacteriocin was desorbed from the cells via resuspension in 50 mM sodium phosphate (pH 2.0) for 2 h. The supernatant was collected, and the peptides precipitated using 40% w/v ammonium sulphate. Whitford *et al*. (2001) noted a floating pellet after the ammonium sulphate precipitation, which was not observed in this study. The peptides in the resolubilized protein pellet were then tested for antimicrobial activity via a well overlay against *C. difficile* CD630. However, the precipitated peptides demonstrated no antimicrobial activity (data not shown). It is possible that Bov255 was not successfully precipitated out of the supernatant as compared to Whitford *et al*. (2001) some experimental conditions differ from the previously described ones, such as the medium and the growth temperature. It is also possible that Bov255 was successfully precipitated and is not responsible for the observed antimicrobial activity against *C. difficile*. An alternative approach to obtain purified Bov255 was to overexpress the

peptide using an *E. coli* host. This method was also to be used to obtain purified BacS to determine if that peptide plays a role in the antimicrobial activity against *C. difficile*.

4.2.5.2 E. coli Host Expression of the Putative Bacteriocin Candidates

4.2.5.2.1 Creation of the Overexpression Vectors

The creation of the Bov255 and BacS overexpression vectors was carried out in the same manner as previously described (Section 3.2.4.1). The identified putative bacteriocin genes, *bov255* and *bacS* were amplified from the *Streptococcus* 30ANiii genome using the primers *bov255* F/*bov255* R and *bacS* F/*bacS* R respectively (Table 7, Section 2.1.5). These primers also incorporated the restriction enzyme sites *Xho*I and *Nde*I into the 5' and 3' ends of both genes respectively. The T7 expression vector pET28b (+) was linearised in the multiple cloning site (MCS) with restriction enzymes *Xho*I and *Nde*I and the DNA backbone gel purified. The amplified PCR gene products were digested with the same enzymes and gel purified. The vector backbone and insert were ligated together to create pET28b (+)*bov255* and pET28b(+)*bacS* (Fig 43) which were transformed into *E. coli* Top10 and confirmed by Sanger sequencing using the T7 Terminator R and T7 promoter F primers. The confirmed plasmids were then transformed into the *E. coli* expression strain, BL21 (DE3) and the plasmid sequence was re-confirmed by Sanger sequencing.

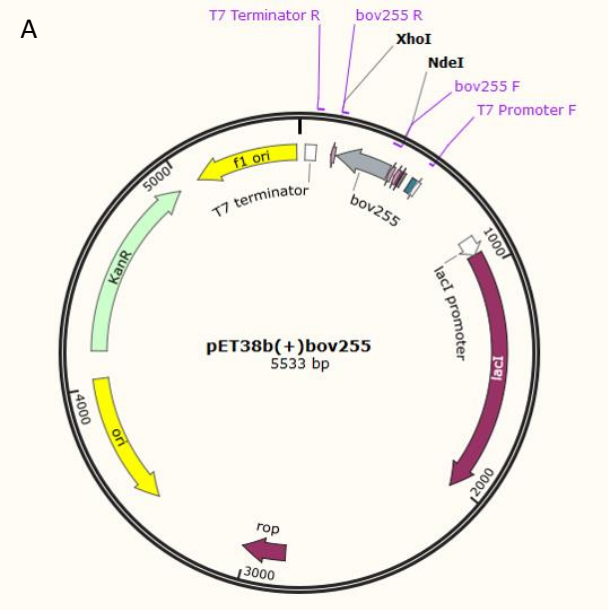
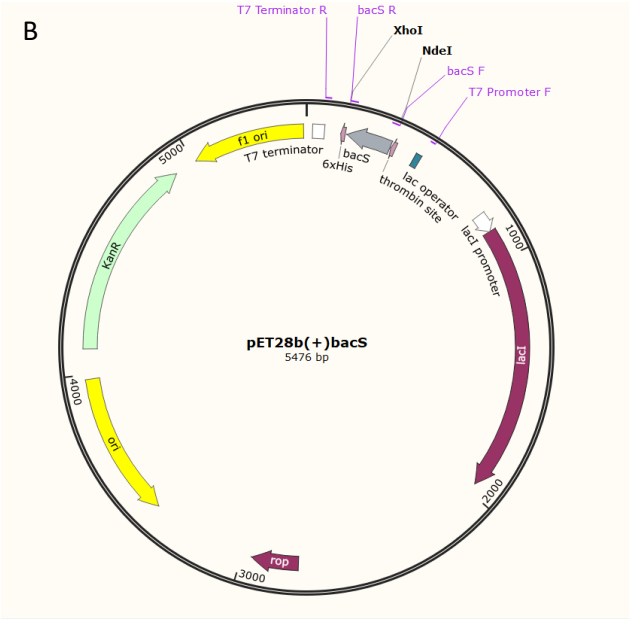


Figure 43- Plasmid maps of pET28b(+)*bov255* and pET28b(+)*bacS*. (a) Plasmid map of the expression vector pET28b(+) with the *bov255* gene cloned into the MCS using the *Xho*I/*Nde*I restriction sites, shown in black. The gene amplification primers (*bov255* F and *bov255* R) and the screening primers (T7 Terminator F and T7 R promoter) are indicated in purple.



(b) Vector map of the pET28b(+)*bac*S with *bacS* cloned into the pET28b(+) vector MSC using *Xho*I/*Nde*I restriction sites indicated in black. The gene amplification primers (*bacS* F and *bacS* R) and the screening primers (T7 Terminator R and T7 promoter F) are indicated in purple.

4.2.5.2.2 Analysis of the Overexpression of Bov255 and BacS

The *E. coli* BL21(DE3) strains harbouring the overexpression vectors pET28b(+)*bov255* and pET28b(+)*bacS* were grown with shaking to mid-exponential phase (OD_{600nm} 0.3-0.4) at 37 °C and induced with 1 mM IPTG for 3-4 h. *E. coli* cells were then harvested and lysed using BugBuster® MasterMix and evidence for overexpressed proteins was determined using a 17.5% Bis-tris SDS-page (Fig 44a). However, no overexpressed proteins were observed. As bacteriocins can be small peptides, and bovicin 255 is less than 6.5 kDa, the cell lysate samples underwent size exclusion using a 50 kDa spin column to filter out larger peptides that might inhibit the migration of smaller peptides through acrylamide gel. Samples from the <50kDa cell lysate fraction was separated on a 16% tricine SDS-PAGE gel to facilitate visualizing smaller peptides however no overexpressed peptides were observed and no peptides significantly smaller than 10 kDa were visualized (Fig 44b). Silver staining was then used to visualize peptides smaller than 10 kDa however there was still no evidence of any overexpressed peptides (data not shown).

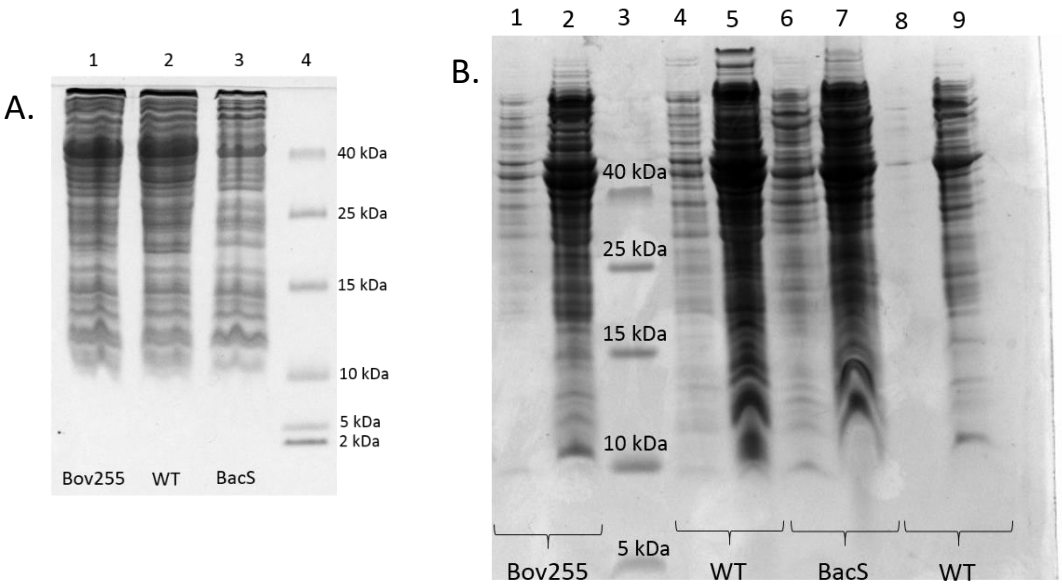


Figure 44- SDS-PAGE gels of cell lysates obtained from induced *E. coli* BL21 (DE3) harbouring the overexpression vectors pET28b(+)*bov255* and pET28b(+)*bacS*. (A) 17.5% Bis-Tris NuPage gel visualising the whole cell lysates from induced *E. coli* BL21 (DE3) pET28(+)*bov255* (lane 1), *E. coli* BL21(DE3) (lane 2) and *E. coli* BL21(DE3) pET28b(+)*bacS* (lane 3). (B) 16% Tricine NuPage gel visualising the same cell lysates as in Figure XA. Lane 1 & 2 contain different loading volumes of *E. coli* BL21 (DE3) pET28(+)*bov255*. Lane 6 & 7 contain *E. coli* BL21 (DE3) pET28(+)*bacS*. Lane 5 & 9 contain different loading volumes of *E. coli* BL21 (DE3). Samples in Lanes 1, 4, 6 & 8 had all undergone size exclusion using a 50 kDa spin column. Lane 4 (A) and Lane 3 (B) contains the Spectra™ multicolour low range protein ladder (ThermoScientific).

4.3 Discussion

4.3.1 Successful Isolation and Purification of Environmental Isolates with Antimicrobial Activity Against *C. difficile*

Samples of cow manure, collected from a dairy farm, were grown under different conditions to isolate organisms with antimicrobial activity against *C. difficile*. Out of 200 cultures screened, 11 were positive for activity against *C. difficile*; with the majority of these were isolated under aerobic growth conditions. Four isolates were identified from the ethanol treated samples that were grown anaerobically whereas only one culture, showing activity, was discovered under anaerobic growth condition without ethanol treatment. Out of these 11 positive cultures there were five unique organisms, four of which are in the *Bacillaceae* family. It was difficult to determine if different cultures containing the same species were different strains within the species

or the same strain. Due to this, only one isolate of each organism was selected for further investigation which limited the selection to five out of 11 cultures thus eliminating several *Bacillus* cultures. *Bacillus* spp. are extremely abundant in diverse environments such as gastrointestinal tracts of animals, soil and food; therefore it is not surprising to have isolated so many from cow manure (Nicholson, 2002). Although multiple cultures containing the seemingly same *Bacillus* sp. were isolated, they are known for their production of diverse antimicrobial substances (Stein, 2005). However, 16s rRNA sequencing is not discriminatory enough to distinguish *Bacillus* spp. down to the subspecies level (Wang *et al.*, 2007). In order to prevent the loss of potential bacteriocin producing strains early in the screening process, further characterization of the organisms down to the subspecies level could have been carried out. For *Bacillus* spp. this could be done through the use of DNA gyrase B sequencing, as it is more discriminating down to the subspecies level, (Wang *et al.*, 2007) or whole genome sequencing. DNA gyrase B sequencing was not carried out here the five organisms selected for further study were deemed a large enough workload to manage within the time constraints of the study.

Another way to prevent the repeated isolation of the same species of bacteria, is to diversify the environmental samples and growth conditions used for isolating cultures. The growth medium could have been varied in order to select for different types of bacteria, such as the use of MRS agar (De Man *et al.*, 1960) to encourage the growth of LAB strains which are known for their carriage of bacteriocins and GRAS status (Alvarez-Sieiro *et al.*, 2016). Also sampling from different areas of the dairy farm could help boost variety, such as collecting samples from within the slurry tank and specifically from calves, whose high prevalence of *C. difficile* (Rodriguez Diaz *et al.*, 2018) could lead to a reservoir of bacteriocins. The discovery of multiple organisms with antimicrobial activity against *C. difficile* on initial screening, demonstrates that

cow manure is a potential reservoir for novel bacteriocins against this organism. It would be interesting to determine the prevalence of *C. difficile* on the dairy farm used in the present study, in comparison with the number of organisms isolated with activity against this bacterium to demonstrate if there is an observable association.

It is interesting to note that the organisms most frequently isolated here, *Bacillus* spp., were not previously found to be the most abundant within the faecal microflora of dairy cattle. The Firmicutes phylum is a highly abundant phylum in dairy cattle faeces however the *Bacillus* genus was not among the common genera identified (Hagey *et al.*, 2019). *B. thermoamylovorans* is known in the dairy industry, however it is known for its role as a food spoilage agent. *Oceanobacillus* spp. are found in a diverse range of environments including the human gastrointestinal tract and fermented foods and *B. licheniformis* is predominantly found in the soil. The selection for bacteriocin producing organisms could have biased the isolation of *Bacillus* spp. however it is possible that these organisms are contaminants from the environment, especially *B. licheniformis* which is common in soil. This highlights a sample collection issue which could have introduced contaminant organisms into the collected faecal matter. The screened samples may not have contained organisms that interact with *C. difficile* in the gastrointestinal tract which could have been a screening limitation. However, the identification of the *Streptococcus* spp. which is a common genus in the microflora of dairy cattle faeces (Dowd *et al.*, 2008; Shanks *et al.*, 2011) does indicate that some of the common genera found in the microflora were screened. To limit contamination with environmental organisms fresh faeces could be collected which has not contacted the barn floor or any potential areas of contamination for example where soil and food might also be present.

4.3.2 Rationale for Selection of Isolates for More Detailed Study

Both *in vitro* and *in silico* techniques, in combination with current literature, were employed to narrow the selection of environmental organism to the candidates most likely to harbor novel bacteriocins. The literature and *in silico* analysis both highlighted the same candidates, *B. licheniformis* 41ANE and *Streptococcus* 30ANiii. *B. licheniformis* strains have been previously found to produce a wide range of bacteriocins active against a range of Gram-positive bacteria (Cladera-Olivera *et al.*, 2004; Begley *et al.*, 2009; Beric *et al.*, 2014; Kayalvizhi *et al.*, 2016; Vadakedath and Halami, 2019). The *in silico* investigation of both organisms generated data in agreement with the current literature with a known bacteriocin cluster was identified within the genome of *B. licheniformis* 41ANE, however there was no putative or known bacteriocin present in the genome of *B. thermoamylovorans* 8ANE. General agreement between the literature-based evidence and data from the *in silico* investigation of known bacteriocins was also apparent for the two other organisms isolated, *Oceanobacillus* sp. and *Streptococcus* sp. Although *Oceanobacillus* sp. have previously appeared in *in silico* searches mining for putative novel bacteriocins, to date, there are no known bacteriocins described (Dirix *et al.*, 2004). *In silico* mining in this study did not identify any putative bacteriocin candidates. *Streptococcus* sp. are known to produce a diverse range of bacteriocins, with the most prevalent being type A lantibiotics, however class II two-peptide bacteriocin have also been identified (Nes *et al.*, 2007). Some *Streptococcus* spp. are known to produce multiple lantibiotics such as *S. mutans* which can produce three types of mutacin (Nes *et al.*, 2007). The GDS are also known to produce a range of bacteriocins with bovicin 255, a class II non-lantibiotic, and bovicin HJ50, a lantibiotic, both being produced from *S. bovis* strains (Whitford *et al.*, 2001; Xiao *et al.*, 2004). The *in silico* investigation of the whole genome

of *Streptococcus* 30ANiii identified both a known bacteriocin cluster, bovicin 255, and a putative known bacteriocin within the same locus.

Although the literature and *in silico* investigation highlighted the same candidates, it is pertinent to note that the bacteriocin database, BAGEL4, was generated from known bacteriocins. BAGEL4 searches genomes for similarities to known bacteriocin and analyses the clustered genes for any transport, immunity, modification or processing functions. To date there are no characterized bacteriocins from *Oceanobacillus* sp. and *B. thermoamylovorans* and therefore homology to known bacteriocins might be too low for the mining tool to identify novel bacteriocins. Screening methods using more conserved peptides, such as the ABC transporter peptides, have been successfully used to identify novel bacteriocin clusters (Marsh *et al.*, 2010; Singh and Sareen, 2014) which could be used to confirm the lack of putative bacteriocin clusters with the genome of *Oceanobacillus* 1_29A and *B. thermoamylovorans* 8ANE. Although other *in silico* tools were used to explain the initial antimicrobial activity observed from *Oceanobacillus* 1_29A and *B. thermoamylovorans* 8ANE, there was no putative responsible candidate identified. Further work on these two strains to identify likely causes of this activity could be carried out through determining if the compound could be an active peptide secreted into the supernatant, via undertaking spent supernatant assays and subsequent peptide precipitation. First, however, re-sequencing the entire genome of both organisms would increase the coverage and quality of the genome sequence, thus making the *in silico* results more reliable.

Further characterization of the initial activity observed from all the environmental organisms resulted in no activity demonstrated from the supernatant, apart from *B. thermoamylovorans* 41ANE. Although different conditions, such as UV

induction, mitomycin C induction and different growth media were used to try and induce activity in the supernatant, spot plating or well overlays never demonstrated activity. The protocol for demonstrating activity in the supernatant was never successfully optimized and activity was only demonstrated through the inhibition of *C. difficile* in the freeze-dried spent supernatant of the organism. This assay was extremely time consuming and the inability to demonstrate activity in the supernatant was a limitation within the strain selection screening process. Optimization of this step in the screening process is essential to be able to select suitable candidates for in depth analysis. A more systematic approach to trial the different conditions that could affect bacteriocin expression could have been employed to streamline the screening process. The approach would first employ the utilization of literature to determine the different conditions that needed to be tested such as growth conditions, growth phase, peptide concentration and induction. This would then be related to the individual producer organisms; for example, *B. thermoamylovorans* growth temperature range varies greatly from that of *B. licheniformis* and the various growth media that could have been trialed would also be different for each organism. This would ensure a more logical approach to demonstrate antimicrobial activity in the supernatant. However, if activity is not observed through plate overlays, the more labor intensive, spent supernatant assays could still be used to demonstrate activity within the supernatant.

4.3.3 Successful Demonstration of Antimicrobial Activity from *Streptococcus* 30ANiii Against *C. difficile*

Although characterization of the initial activity from *Streptococcus* 30ANiii against *C. difficile* yielded no observed activity from the supernatant, the *in silico* analysis highlighted *Streptococcus* 30ANiii as a promising candidate. As the *in silico* analysis identified both a putative and known bacteriocin within the *Streptococcus* 30ANiii genome, further effort was made to demonstrate antimicrobial activity within

the supernatant. Concentrated spent supernatant resulted in the inhibition of *C. difficile* CD630 however the control culture of *C. difficile* growing in its own spent supernatant was also slower than the controls. The growth of the *C. difficile* spent supernatant control was comparable to normal growth by the 24 h time point and a 48 h time point was also taken to ensure there was still no growth of *C. difficile* within the spent supernatant of *Streptococcus* 30ANiii at that time. A possibility is that the salt concentration or another compound within the medium was too high in the lyophilizate and caused an increase in *C. difficile* lag phase but did not inhibit its growth completely. Despite the abnormal growth of the control, there is evidence that an active compound is secreted into the supernatant of *Streptococcus* 30ANiii with antimicrobial activity against *C. difficile* CD630. Further independent repeats of this assay should be carried out to determine the cause for the slower growth of *C. difficile* within its own concentrated spent supernatant. This assay also needs to be extended to include other strains of *C. difficile* such as DH1916, R20291, and CD196, to determine if the same pattern of an increased lag phase and the inhibition of *C. difficile* is observed.

4.3.4 Overexpression of the *In Silico* Bacteriocin Candidates Identified within *Streptococcus* 30ANiii in an *E. coli* Host

One known and one putative bacteriocin candidate (Bov255 and BacS, respectively) were identified in the genome of *Streptococcus* 30ANiii and cloned into a T7 overexpression vector. This vector was transformed into different *E. coli* overexpression hosts, however no overexpressed peptides were observed. To optimize the expression of Bov255 and BacS different growth temperatures, induction concentration and induction lengths could have been trialled in a more methodical manner. Alternatively, the reason for the lack of observed overexpressed peptide could be due to a low concentration of the peptide. To determine if this was the case,

the incorporated His-tag could have been utilized to either probe the cell lysates using the more sensitive western blot assay or to concentrate the peptide through His-tag purification. If it was determined that either peptide was produced, then anti-*C. difficile* activity could be tested. This could be done using the individual peptides and trialing a combination of both peptides at different concentrations to observe any synergistically activity or determine if they act as a two-peptide bacteriocin.

Another possible strategy to utilize *E. coli* as an expression host would be to express the whole locus or parts of the locus. Bov255 is similar to lactococcin A which requires the expression of three other genes to be produced, *lcnC, lcnD* and *lciA*. LcnC and LcnD are the ABC transporter and secretion accessory peptides where as LciA is the immunity peptide (Stoddard *et al.*, 1992). The Bov255/BacS locus has multiple genes however there is a clustering of four genes (*bov255, bai255, lagD,* and *lcnD*) which resembles the layout of the lactococcin A locus. This locus could be expressed in *E. coli* and tested for anti-*C. difficile* activity. It is important, however, to test the bacteriocin against previously confirmed sensitive strains as well as *C. difficile* to confirm the peptide is functional. Also, it would be interesting to express the entire locus to determine if all the genes are necessary for the successful expression of a mature bacteriocin active against *C. difficile*. It could be possible to isolate the specific active peptide using SDS-PAGE gels and identify it using LC-MS/MS. Another possibility is to generate two vectors, to be carried within the same host, one with the locus without the structural gene and another vector with the structural gene which incorporates a protein tag such as a His-tag. This would allow the active peptide to be successfully generated whilst also providing a downstream method for purification. The benefit of this method is that the essential genes could be determined through generating a range of different constructs with combinations of the genes.

4.3.5 Precipitation of Bovicin255 from *Streptococcus* 30ANiii

Bov255 was originally isolated through a mixture of two protein purification techniques; firstly absorption/desorption was partially used to allow any peptides bound to the cell membrane to be released into the supernatant through incubating the cells at pH 2.0. Then ammonium sulphate was used to precipitate peptides from the supernatant which resulted in a fraction of peptides that demonstrated activity against the indicator *S. bovis* strain (Whitford *et al.*, 2001). This process was repeated using *Streptococcus* 30ANiii as it was shown to harbor the *bov255* structural gene, however, the protocol did not result in the floating pellicle as described by Whitford *et al*. (2001) and there was no observed antimicrobial activity. The previously purified Bov255 was not active against two clostridia, *C. perfringens* and *C. clostridiforme,* therefore the negative result could be because it is inactive against *C. difficile* rather than the bacteriocin not being precipitated. One method to determine it if has been precipitated would be to observe if it is active against previously sensitive strains or to conduct LC-MS/MS on the precipitated peptides to determine if the peptide is present. LC-MS/MS, however, is highly sensitive and thus it might not accurately show all the peptide sequences present if there is a high concentration of contaminants thus further downstream purification would most likely be required. Bov255 could also simply not have been expressed by the *Streptococcus* 30ANiii strain as the growth medium and the temperature used in the present study did differ slightly from that used by Whitford *et al.* (2001). Gene transcription could be determined through RT-PCR however this approach would only confirm that the peptide is expressed and not if it is secreted. The generation of *bov255* and *bacS* knock-out mutants is another potential method however there could be natural barriers preventing easy genetic manipulation.

If Bov255 is not responsible, or solely responsible, for the previously observed anti-*C. difficile* activity, then other protein precipitation techniques could be used to try and purify the active compound. There is no standard precipitation method for bacteriocins, even within the same classes due to the structural diversity of the peptides. Previous successful precipitation techniques used include salting out via ammonium sulphate (Contreras *et al.*, 1997; Flynn *et al.*, 2002; Georgalaki *et al.*, 2002; Todorov *et al.*, 2004; Bauer *et al.*, 2005; Deraz *et al.*, 2005), absorption-desorption (Elegado *et al.*, 1997; Atrih *et al.*, 2001), and solvent precipitation (Venema *et al.*, 1997; Fujita *et al.*, 2007), however the optimisation of these techniques along with the possible expression conditions are very labour intensive. As ammonium sulphate precipitation is known to be a mild form of protein precipitation, this should be trialled first using a sequential increase in concentration to fractionate the peptides within the supernatant and checking for anti-*C. difficile* activity in each fraction.

Take Home Points

- Bacteriocins producing strains with antimicrobial activity against *C. difficile* were successfully isolated from dairy cow manure
- A screening approach utilizing both *in silico* and *in vitro* techniques highlighted the most likely candidates carrying novel bacteriocins

Chapter 5: Characterisation of a *Bacillus pumilus* Strain with Antimicrobial Activity Against *C. difficile*

5.1 Introduction

5.1.1 Objectives of this Project

A bacterial strain, isolated from cow manure in Nottingham, demonstrated antimicrobial activity against *C. difficile* (Section 4.2.1 and 4.2.2). This study aims to characterise this strain and identify the potential causes of this activity *in silico*. It also aims to demonstrate if the observed activity is due to a secreted peptide and to characterise and identify this peptide to determine if it is a putative bacteriocin with activity against *C. difficile*.

5.1.2 Bacillus pumilus

B. pumilus was one of the four original members of the *B. subtilis* group, within the *Bacillus* genus (Zeigler and Perkins, 2008). *B. pumilus* has a history of use in industrial processes, as it produces bioactive substances such as vanillin, keratinase and xylanase (Nwokoro and Dibua, 2014; Saranya *et al.*, 2014). This species has also been found to promote plant growth in healthy plant tissue (Myresiotis *et al.*, 2014) however, some strains are known plant pathogens causing disease in mango trees, ginger and Asian pear trees (Galal *et al.*, 2006; Li *et al.*, 2009a; Yuan and Gao, 2015). Although not as common as other *Bacillus* spp. there have been reports of *B. pumilus* causing rare cases of septicaemia, cutaneous infection and food poisoning (Bentur *et al.*, 2007; From *et al.*, 2007; Tena *et al.*, 2007; Kimouli *et al.*, 2012). Some *B. pumilus* strains produce pumilacidins A, B, C, D, E, F and G, a heat-stable cyclic lipopeptide similar to other *Bacillus* spp. lipopeptides, surfactin and lichenysin (Peypoux *et al.*, 1999). Previous studies have shown these lipopeptides lead to the destruction of cell membranes (Sheppard *et al.*, 1991; Peltola *et al.*, 2001) and a pumilacidin producing

strain has been implicated in a small outbreak of food poisoning (From *et al.*, 2007). It is unknown why there are so few cases of reported *B. pumilus* food poisonings (Logan, 2012). Interestingly, although there are reported cases of human pathogenesis, a probiotic strain of *B. pumilus* is also currently on the market as Biosubtyl in Southeast Asia (Biophar Co. Ltd., Nha Trang, Vietnam) (Green *et al.*, 1999; Duc *et al.*, 2004).

5.1.3 Known Bacteriocins Produced by *Bacillus pumilus*

The *B. subtilis* group are known to produce a wide range of active compounds including bacteriocins (Rawlings and Salvesen, 2013). There are two characterised bacteriocins produced by *B. pumilus*, pumilicin 4 and pumilarin (Aunpad and Na-Bangchang, 2007; van Heel *et al.*, 2017); the former was the first bacteriocin characterised from *B. pumilus*. It is produced by *B. pumilus* WAPB4 and was identified for its potential use against methicillin-resistant *Staphylococcus aureus* (MRSA) and vancomycin-resistant *Enterococcus faecalis* (VRE) (Aunpad and Na-Bangchang, 2007). Pumilicin 4 was determined to be heat-stable, stable across a range of pHs, hydrophobic and bacteriocidal. It was found to be active against Gram-positive bacteria including *B. cereus, B. megaterium, E. faecium,* MRSA, and VRE but not against Gram-negative bacteria which is typical for Gram-positive bacteriocins (Gray *et al.*, 2006; Aunpad and Na-Bangchang, 2007). *B. pumilus* WAPB4, was further characterised and found to be carrying two putative holin-like peptides, XhlB and BhlA, the latter of which was identified as bacteriocin related (Aunpad and Panbangred, 2012). Holins are used by bacteriophage to generate non-specific lesions in the cytoplasmic membrane which allow access to the peptidoglycan and results in cell lysis (Young and Blasi, 1995). It was suggested that BhlA was a bacteriocin-related holin-like peptide as it resulted in the death of the *E. coli* expressing host due to cell membrane disruption (Aunpad and Panbangred, 2012).

The second characterised bacteriocin produced by *B. pumilus* was classified as a head-to-tail cyclised bacteriocin, the first known in *B. pumilus*. Genome mining identified a widespread occurrence of cyclised bacteriocin clusters within *Bacillus* spp. and highlighted the high diversity of cyclised bacteriocins within this genus (van Heel *et al.*, 2017). Pumilarin is located in a cluster of five genes including the structural gene, *pumA*, encoding a 7 kDa peptide, but no self-immunity gene. The sequence of the prepeptide showed limited homology to other circular bacteriocins however there were several strongly conserved regions within the leader sequence and the core peptide sequence. The prepeptide is processed, in a similar manner to known circular bacteriocins, removing the leader peptide. The N- and C- termini are fused resulting in a mature peptide which constitutes the active bacteriocin. Pumilarin does not show as strong an antimicrobial pattern of activity as enterocin AS-48 from *Enterococcus faecalis* (Gálvez *et al.*, 1989; Burgos *et al.*, 2014), the prototype for circular bacteriocins. However it does show strong activity against other *Bacillus* spp. Interestingly, it was also active against *E. coli* which is not typical of Gram-positive bacteriocins (Gray *et al.*, 2006; Aunpad and Na-Bangchang, 2007). The amino acid sequence of pumilarin could therefore give insight into possibilities for protein engineering to develop circular bacteriocins active against Gram-negative bacteria (van Heel *et al.*, 2017).

5.1.4 Circular Bacteriocins

Circular bacteriocins are a class of bacteriocins that are characterised by fused N- and C-termini generating a cyclic peptide. There has been some debate as to what class they fall under, previously making up class IIc, although newer classes IV or V have been proposed as well as class Ib, as modified peptides termed 'head-to-tail cyclized peptides' (Kemperman *et al.*, 2003; Heng *et al.*, 2007; Maqueda *et al.*, 2008; Alvarez-

Sieiro *et al.*, 2016). Due to their compact and circular nature, these bacteriocins are known for being resistant to many proteolytic enzymes as well as being heat and pH stable (Maqueda *et al.*, 2008). These characteristics, along with their typical broad-spectrum antimicrobial activity, make them favourable for industrial and medical applications (Towle and Vederas, 2017). Circular bacteriocins are typically produced by Gram-positive bacteria, mainly LAB strains and a number of *Bacilli* and *Clostridia* (Gabrielsen *et al.*, 2014a). To date there are over a dozen identified circular bacteriocins which can be categorised into two subgroups based on homology and physiochemical characteristics (Fig 45). Subgroup i is made up of circular bacteriocins with a cationic surface and high isoelectric points (pI) with some conserved homology within the group (Gabrielsen *et al.*, 2014a). Some examples of subgroup i are enterocin AS-48 (AS-48) (Martinez-Bueno *et al.*, 1994), carnocyclin (Martin-Visscher *et al.*, 2008), circularin A (Kawai *et al.*, 2004), uberolysin (Wirawan *et al.*, 2007), and garvicin ML (Borrero *et al.*, 2011). Lactocyclicin Q (Sawa *et al.*, 2009) and leucocyclicin Q (Masuda *et al.*, 2011) are also members, however they are outliers as they are evolutionarily more closely related to each other than the other members of the subgroup. Subgroup ii have low pIs and an overall anionic surface. This group is not as well characterised as subgroup i, with only two members, gassericin A (Kawai *et al.*, 2004) and butyrivibriocin AR10 being defined so far (Kalmokoff *et al.*, 2003). Subgroup ii members also do not conform to the Saposin-like fold but rather are alpha helical bundles (Acedo *et al.*, 2015). One exception to these subgroups is subtilisin A which, in contrast to the other circular bacteriocins, contains cysteines and therefore does not conform, in composition, to either subgroup and thus it was suggested for reclassification in a new sactibiotic class (Kawulka *et al.*, 2003; Alvarez-Sieiro *et al.*, 2016).

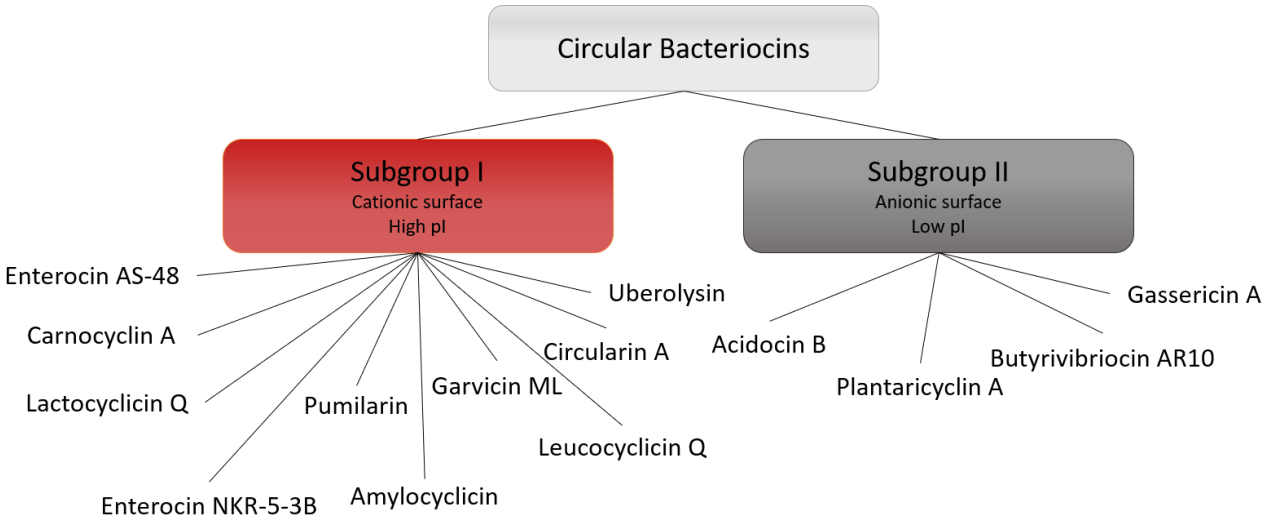


Figure 45-Subgroups of circular bacteriocins The two different subgroups of circular bacteriocins, subgroup I shown in red and subgroup II shown in black, with their corresponding members.

5.1.4.1 Genetic and Structural Features

Although only a few of the circular bacteriocins have been characterised there are some shared traits between them genetically and structurally. The gene clusters of circular bacteriocins usually contains 5-7 genes, which includes the structural gene, an immunity gene, ATPase, putative membrane protein(s) and an integral membrane protein. There are observed differences between the two subgroups, with subgroup i typically containing a gene encoding a large protein with predicted transmembrane topology but no known functional domain or homology; whereas subgroup ii usually contains a gene encoding a smaller membrane protein with homology to ABC-2 membrane transporters (Gabrielsen *et al.*, 2014a).

The biosynthesis process of circular bacteriocins remains largely unknown with only three major steps known: the cleavage of the leader sequence, circularisation, and exportation out of the cell. The leader sequence ranges from 2-48 residues and is highly variable across the different bacteriocins (Gabrielsen *et al.*, 2014a). It has been previously demonstrated that the removal of the leader sequence is the first step in the biosynthesis pathway and not required for circularisation or

exportation of the peptide out of the cell (Gabrielsen *et al.*, 2014b). There is little known about the circularisation process of these bacteriocins however it has been demonstrated that the first and last residues of the mature peptide affects the process and that hydrophobic residues, in the N- and C- termini, are important (Cebrian *et al.*, 2010; Gabrielsen *et al.*, 2014a). The mature peptides all contain a conserved structural motif of a Saposin like fold or α-helical bundle, with the former typical of subgroup i and the latter typical of subgroup ii (Towle and Vederas, 2017). Both motifs typically consist of four to five α-helices however the Saposin like fold is a more specific formation of two α-helices forming a 'v-shape' with an α-helix positioned almost perpendicular to this 'v-shape' (Fig 46). The difference between the Saposin fold and the Saposin like fold is the lack of di-sulphide bonds stabilising the structure in the latter fold, which is due to absence of cysteines (Towle and Vederas, 2017). It has been postulated that this structural motif is conserved as it might be a key feature for the interaction of these bacteriocin with lipids (Towle and Vederas, 2017).

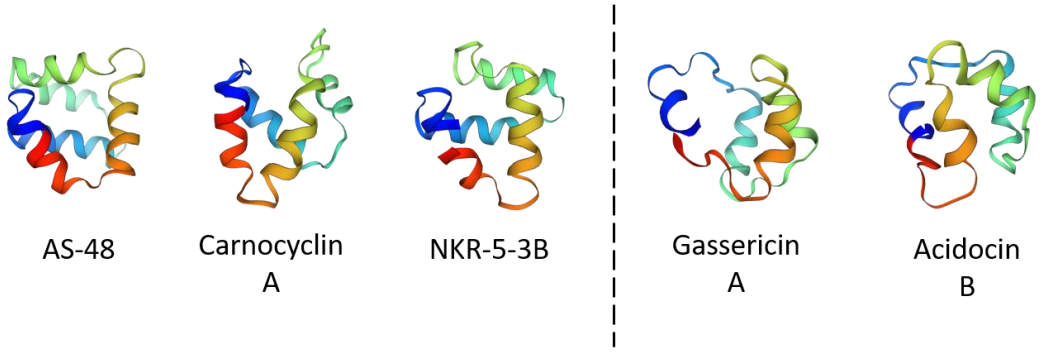


Figure 46– Comparison of the structure of subgroup i and ii circular bacteriocins. Subgroup i (left) circular bacteriocins show the typical Saposin like fold, with two α-helices forming a 'v-shape' and another α-helix almost perpendicular to this. These members include enterocin AS-48 (Q47765), carnocyclin A (B2MVM5) and NKR-5-3B (A0A0P0YL94). The subgroup ii circular bacteriocins (right) show a simple α-helical bundle motif lacking the 'v-shape'. The members of subgroup ii shown are gassericin A (O24790) and acidocin B (Q48499).

5.1.4.2 Mode of Action

Circular bacteriocins are known to have a broad spectrum of antimicrobial activity against Gram-positive bacteria. Activity has also been demonstrated against Gram-negative bacteria however usually at much higher concentrations or when the cell membrane is disrupted (Vasilchenko and Valyshev, 2019). The conserved structural motifs of this class of bacteriocins suggests a similar mode of action, however there are significant differences. Although typically these bacteriocins form pores the observed pores are different, either being non-selective or ion selective (Vasilchenko and Valyshev, 2019). The two best studied pore-formers in this class are AS-48 and carnocyclin A. AS-48 was demonstrated to form a homodimer at physiological pH, undergoing a conformational change when it came into contact with the cell membrane. This resulted in the dimer being embedded into the membrane, forming a non-selective pore, leading to a loss of ions and membrane potential (Gálvez *et al.*, 1991; Sanchez-Barrena *et al.*, 2003). Carnocyclin A, in contrast, forms highly stable voltage dependent anion-selective pores and gassericin A results in the leakage of potassium ions (Kawai *et al.*, 2004; Gong *et al.*, 2009). Although these bacteriocins form different types of channels they all result in cell death through undermining the integrity of the cell membrane.

It has been previously assumed that circular bacteriocins work in a receptor-independent manner, as studies have established that gassericin A, carnocyclin A and AS-48 do not need a receptor to cause cell death (Gálvez *et al.*, 1991; Kawai *et al.*, 2004; Gong *et al.*, 2009). It has also been demonstrated that enterocin NKR-5-3B non-selectively binds to the cell membrane, solely due to the membranes negative charge (Himeno *et al.*, 2015). However more recently Garvicin ML was found to be able to use the maltose ABC transporter as a receptor; yet at higher concentrations the activity was shown to be receptor-independent (Gabrielsen *et al.*, 2012). Thus it has now been

recognised that some circular bacteriocins have a concentration dependent mode of action, where higher concentrations result in receptor independent activity and lower concentrations in receptor dependent activity (Gabrielsen *et al.*, 2014a; Vasilchenko and Valyshev, 2019). A dual mode of action, due to concentration difference, has been previously demonstrated in other bacteriocins such as nisin, where it causes cell death by pore-formation at high concentrations and though inhibition of cell wall synthesis at low concentrations (Breukink *et al.*, 1999).

5.1.4.3 Circularin A

Circularin A is a circular bacteriocins produced by *Clostridium beijerinckii* ATCC25752, which was isolated due to its antimicrobial activity against the cheese-spoilage bacterium, *Clostridium tyrobutyricum* NIZOB570. It was demonstrated that circularin A was partially resistant to proteinase digestion despite containing digestion sites, which is a known result of the circular bacteriocins compact structure. It was also resistant to protein denaturing conditions which allowed antimicrobial activity to be observed on an overlaid SDS-PAGE gel, indicating a molecular weight of 2.2 kDa. Circularin A was partially purified using 50% ammonium sulphate precipitation (which resulted in a crude purification), before being further purified using HPLC. The result was a purified head-to-tail cyclised protein with the leader sequence removed and a new peptide bond formed between the N-terminal Val^{4} and the C-terminal Tyr^{72}. Circularin A, a 72-residue peptide with a pI of 10.6, was found to be encoded chromosomally by *cirA* (AY164463) (Kemperman *et al.*, 2003). The amino acid sequence of circularin A has 30% identity to AS-48, however circularin A only has a leader sequence of 3 amino acids whereas AS-48 has an extended leader sequence (Kemperman *et al.*, 2003; Kawai *et al.*, 2004). This difference is typical of this bacteriocin class, which are known to possess highly diverse leader sequences (Gabrielsen *et al.*, 2014a).

5.2 Results

5.2.1 Strain Characterisation

Sample 1_16 was isolated under aerobic conditions on rich media from cow manure (Section 4.2.2). The organism(s) present in this sample demonstrated activity against *C. difficile* DH1916, and CD630 but not R20291. To determine what organism(s) was present in this sample it was cultured to purity (Section 4.2.2 Fig 34). Once purified into single colonies its antimicrobial activity was confirmed and it now also demonstrated activity against *C. difficile* R20291 which could have been an artefact of the colony size that was previously tested or its activity was masked as it was not a pure isolate. Initial characterisation showed the strain to be a Gram-positive bacillus with potential spores observed (Fig 47), and the 16s rRNA PCR identified it as a possible *B. pumilus* (Section 4.2.2 Table 16) therefore the strain was named *B. pumilus* 1_16. The growth of the strain aerobically, shaking in BHIS medium at 37 °C, showed mid-exponential phase at three hours, with early stationary phase at five to six hours of growth (Fig 48). The viable cell count also follows the pattern of the OD_{600nm}, however there is a steep drop in cell numbers between four and six hours of growth (Fig 48) which could be a change in growth characteristics brought about by sporulation as endospores were observed at the six-hour time point by microscopy (data not shown).

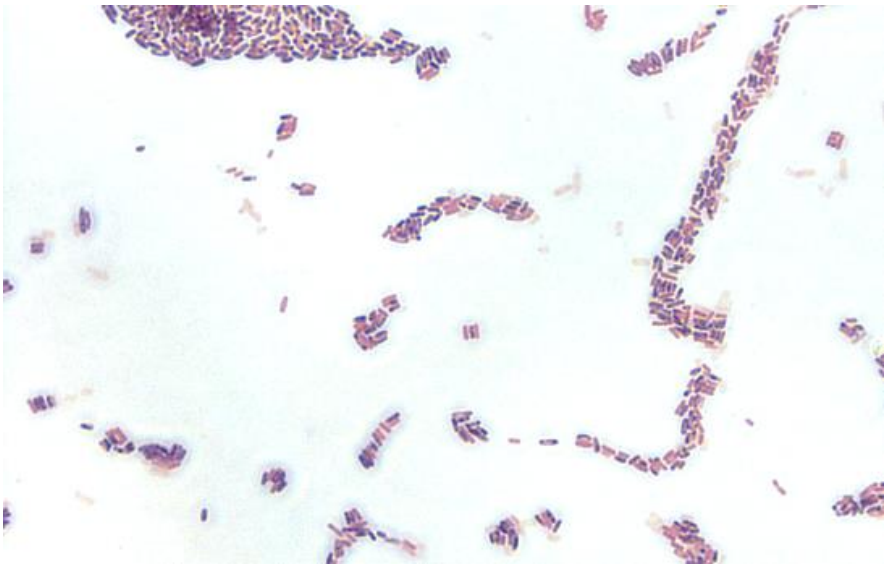


Figure 47- Gram-stain of Sample 1_16. Gram-positive Bacilli, with potential endospores. X1000 magnification

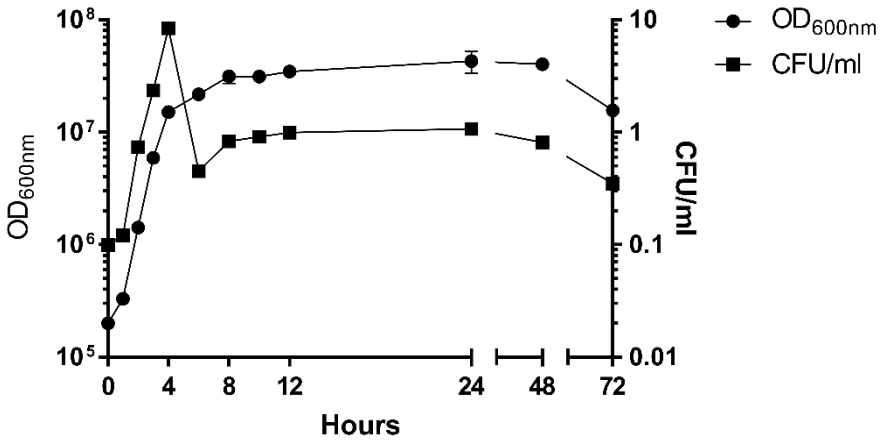


Figure 48- Growth curve and CFU/ml of *B. pumilus* 1_16 over 72 h. *B. pumilus* 1_16 was grown aerobically shaking at 37°C in BHIS medium and plated onto BHIS agar for the CFU/ml count. OD_{600nm} readings are shown on the left Y axis and CFU/ml is on the right Y axis. The error bars represent the standard deviation across the three technical repeats for the three independent biological repeats. The steep drop after the 4 h time point in viable cells is most likely due to sporulation.

5.2.2 *De novo* Whole Genome Sequencing

To identify possible bacteriocins, *B. pumilus* 1_16 was genome sequenced using the illumina HiSeq 2500 platform and was assembled using a *de novo* method, due to its environmental nature. *De novo* sequencing assembles sequence reads into

contigs with the coverage quality dependent on the size and continuity of these contigs. *B. pumilus* sequencing produced a 3.8 Mb size genome, with a 41.5% GC content which is consistent with the median *B. pumilus* genomes on the NCBI database. As quality control, the paired end reads were mapped back to the *de novo* assembly which showed the 99.65% of reads were mapped to the genome with a mean mapping quality score of 33.44. Also, the average coverage was 52X (±15.4). The quality control showed the contigs to be reliable, although a higher coverage would have been better for *de novo* sequencing as it would have resulted in fewer more reliable contigs (Desai *et al.*, 2013).

The 16s rRNA from the assembled whole genome was used to find a closest reference. First a phylogenetic tree was generated to confirm its categorisation as *B. pumilus*, and it was confirmed that its closest relatives were *B. pumilus* and *B. safensis* (Fig 49). Since a whole genome reference was required for ordering the generated contigs into the most likely order, a second phylogenetic tree was generated using the 16s rRNA of all available *B. pumilus* and *B. safensis* completed whole genomes (Fig 50). The closest relative here was *B. pumilus* SAFR-032 (NC_009848.4) which was the reference for the final alignment using Mauve Progressive Alignment (Darling *et al.*, 2004, 2010). The contigs were shuffled to reflect the closest possible organisational layout of *B. pumilus* SAFR-032, which resulted in 22 locally collinear blocks (LCB) of conserved regions between the two strains (Fig 51). The newly shuffled *B. pumilus* 1_16 contigs reflected the organisational layout of *B. pumilus* SAFR-032 with two noticeable differences, a gap and a different arrangement of a block of homology (Fig 51 indicated with arrows). The ordered contigs were then uploaded to the RAST pipeline (Aziz *et al.*, 2008; Overbeek *et al.*, 2014; Brettin *et al.*, 2015), which identifies and annotates open reading frames (ORF).

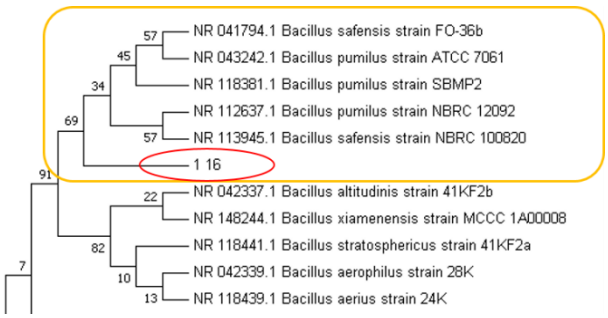


Figure 49-Section of the 16s rRNA Maximum Likelihood Phylogenetic Tree of *B. pumilus* 1_16. The phylogenetic tree shows the relationship of *B. pumilus* 1_16 to close relatives identified through NCBI 16s rRNA database. The phylogenetic tree was generated by MEGA7 (Kumar *et al.*, 2016) and bootstrapped 500 times. The *B. pumilus* and *B. safenesis* branch is indicated through the yellow rectangle and strain 1_16 by the red circle. The out group used was *Escherichia coli* K-12 (not shown) and the full tree is shown in appendix 7.3.

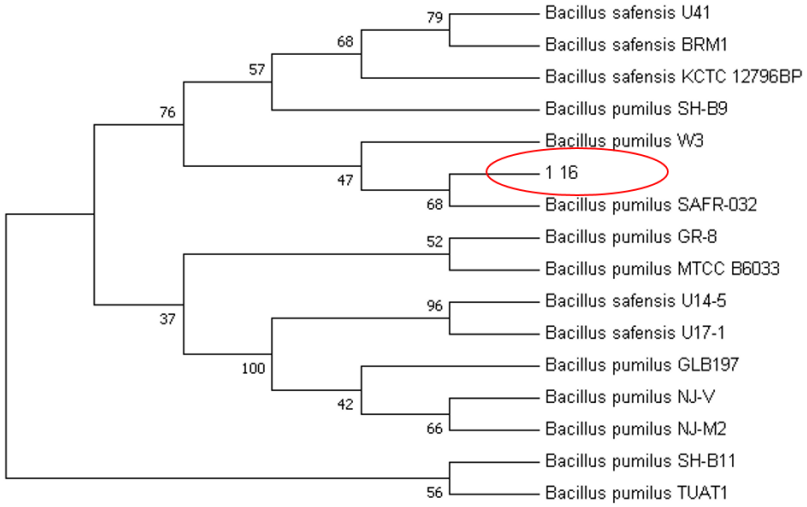


Figure 50- 16s rRNA phylogenetic tree of strain 1_16 against *B. pumilus* and *B. safensis* with available whole genome sequences. *B. pumilus* and *B. safenesis* were determined to be the closest relatives to strain 1_16. Any strains with available whole genome sequences were aligned with strain 1_16 16s rRNA sequence in a maximum likelihood phylogenetic tree which was bootstrapped 500 times by MEGA7 (Kumar *et al.*, 2016). The closest relative is *B. pumilus* SAFR-032 on the same branch as strain 1_16 which is circled in red.

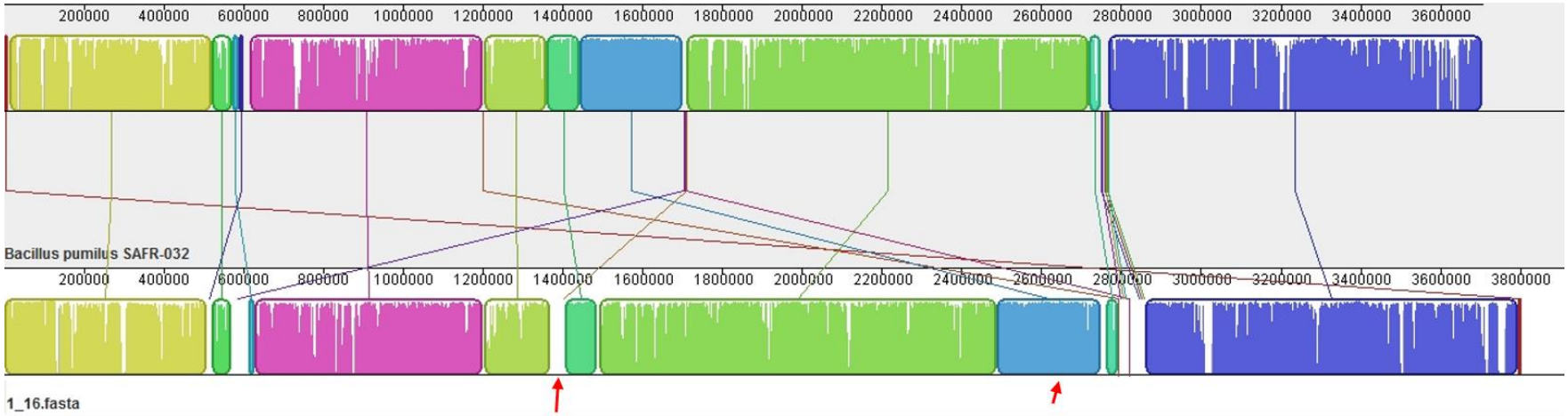


Figure 51- *B. pumilus* 1_16 contig organisation aligned with the reference strain *B. pumilus* SAFR-032, generated by Mauve. The reference strain is the top stand with *B. pumilus* 1_16 on the bottom. The different blocks represent local collinear blocks of homology and the lines indicates the location of the homology within the other genome. The red arrows indicate two noticeable differences in the organisation of both genomes.

5.2.3 Bacteriocin Mining

The draft genome of *B. pumilus* 1_16 was mined for any potential cause of antibacterial activity against *C. difficile*. Two areas of interest were identified using BAGEL4, one of which was a phage gene and another a potential amylocyclicin like class IIc bacteriocin, termed *bacP*. Due to the identification of this phage gene, the genome was further searched for potential phage regions using PHASTER, and two intact phage genomes were identified along with two potential phage regions and one incomplete region (Fig 52).

5.2.3.1 *In Silico* Analysis of Phage Regions

The phage database PHASTER identified two intact phage genomes, PM1 and Osiris within the genome of *B. pumilus* 1_16, and three other areas containing phage genes (Table 20). Although there are two intact phage regions present in the genome, *C. difficile* phage typically have very narrow host range usually originating from other *C. difficile* strains (Hargreaves and Clokie, 2014) and need to be induced to leave the genome. Therefore, the two phage present within the genome are most likely not responsible for the activity seen as it is across a variety of different *C. difficile* strains and was detected without induction by mitomycin C.

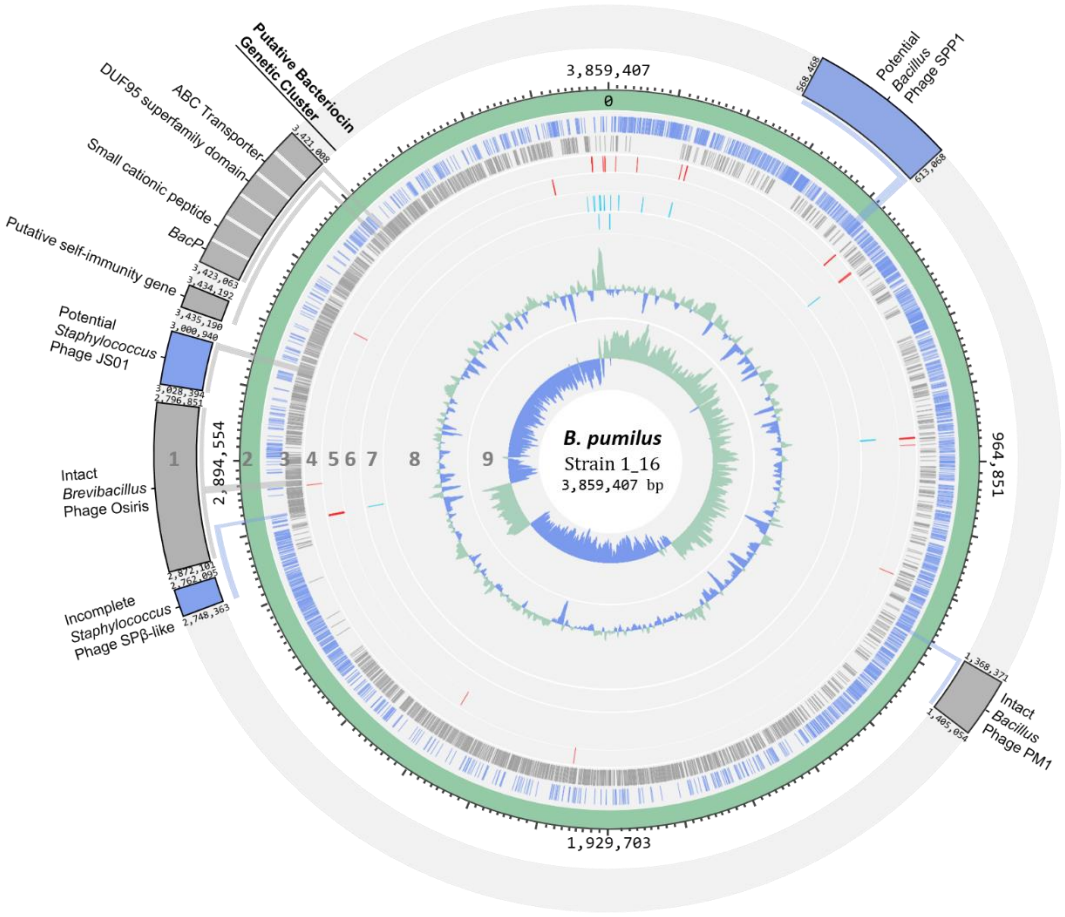


Figure 52-Circular genome of B. pumilus 1_16. Track 1: Genome features of interest; the *bacP* locus and two intact phage (Phage Osiris and Phage PM1), and three phage regions (light blue); Track 2: source sequence; Track 3: Coding sequences on the positive (blue) and negative (green) strand; Track 4: positive strand tRNA CDS; Track 5: negative strand tRNA CDS; Track 6: positive strand rRNA CDS; Track 7: negative strand rRNA CDS; Tracks 8 and 9: G+C content and G+C skew, (blue above average GC, green below average GC).

Table 20- Locations of Putative Phage Regions within *B. pumilus* 1_16 Whole Genome Sequence Determined by PHASTER.

Region Length (Kb)	Completeness	Region position	Phage with the highest homology
44.6	Potential	568468-613068	Bacillus phage SPP1 NC_004166
36.6	Intact	1368371-1405054	Bacillus phage PM1 NC_020883
13.7	Incomplete	2748363-2762095	Staphylococcus phage SPβ-like NC_029119
75.2	Intact	2796851-2872101	Brevibacillus phage Osiris NC_028969
27.4	Potential	3000940-3028394	Staphylococcus phage JS01 NC_021773

5.2.3.2 Genetic *In Silico* Analysis of *bacP* Region

The putative bacteriocin encoding gene identified by BAGEL4, seemed to be a solo gene as the database did not highlight any other genes of interest in the region. However previous studies comparing the genetic clusters of circular bacteriocins illustrated the need for peptide modification and transportation genes, typically between five and seven genes in a bacteriocin locus. There are five genes located upstream of *bacP* (Fig 53), which were annotated by RAST as all hypothetical proteins and one as an ABC transporter gene. *In silico* analysis suggests hypothetical protein 3 has a DUF59 domain which conforms to genes located in other circular bacteriocin clusters. Four genes are separated from *bacP* by a terminator, and there is another predicted terminator after the fifth gene (Fig 53). Typically, a self-immunity gene would be in this cluster, however none is identified by BAGEL4 and a putative self-immunity gene is located further downstream which is further analysed in section 5.2.3.4. The hypothetical protein 1 (Fig 53) satisfies all the requirements of being a short (estimated 5.9 kDa) hydrophobic(GRAVY score of 0.815) and cationic peptide (theoretical pI of 10.13) that is typically found in the subgroup i circular bacteriocin gene cluster. Hypothetical protein 3 (Fig 53) has a predicted DUF95 superfamily domain which contains the Stage II sporulation protein M domain (InterPro IPR002798) identified within the sequence by PHMMER (Potter *et al.*, 2018). This domain encompasses predicted integral membrane proteins and is typically found within the cluster of genes required for circular bacteriocins biosynthesis. Hypothetical protein 2 and 4 show no homology to known proteins and contain no known domains. Overall, this cluster of genes does contain some of the known necessary genes for circular bacteriocin biosynthesis however the lack of known domains in hypothetical protein 2 and 4 hinders the classification of this cluster. The lack of information known

about circular bacteriocin biosynthesis and associated genes should be noted, hence this genetic cluster's involvement with BacP cannot be ruled out.



Figure 53- The organisation of operons in the putative BacP biosynthesis pathway. *bacP* is shown as the gene arrow. The grey arrows depict hypothetical proteins 1-4, and the orange arrow is the putative ABC transporter. Predicated terminators are depicted by loops.

5.2.3.3 *In Silico* Analysis of BacP

Two identical matches can be found in the newly identified *B. pumilus* strain for the amino acid sequence of BacP; however, both are predicted peptides for other *B. pumilus* strains. Two domains were identified by BlastP, circularin A/uberolysin family (Accession TIGR03651) and bacteriocin class IId (pfam09221). Both families incorporate membrane-interacting peptides such as the bacteriocin AS-48 with broad spectrum anti-microbial activity against Gram-positive and Gram-negative bacteria. Circular bacteriocins are split into two subgroups, i and ii, however both subgroups of circular peptides are structurally similar; BacP does not possess any cysteines complying with the Saposin-like fold typical of subgroup i. Also, the predicted model of the peptide shows the correct formation of alpha helices for this fold when compared with other members of this family (Fig 54). This model covers residues 33-96, which represents ~60% of the full amino acid sequence and uses bacteriocin NKR-5-3B as the reference (Fig 54), as its sequence has a higher similarity (39% ID) to BacP when compared with other potential references. This generated a model with a Global Model Quality Estimation of 0.55, and a Qualitative Model Energy Analysis of -1.39 which indicates the fidelity of the model (Fig 54). A phylogenetic tree including

members of both circular bacteriocin subgroups, proposes circularin A as the closest relative (Fig 55), which is in subgroup i, with a sequence homology of 33% across a QC of 86% and an E-value of 6e-05. Sequence alignment of members of subgroup i has previously highlighted conserved residues across this group (Gabrielsen *et al.*, 2014a) and present in BacP (Fig 56) which indicates significant homology between BacP and subgroup i members, placing it within this subgroup and not within subgroup ii.

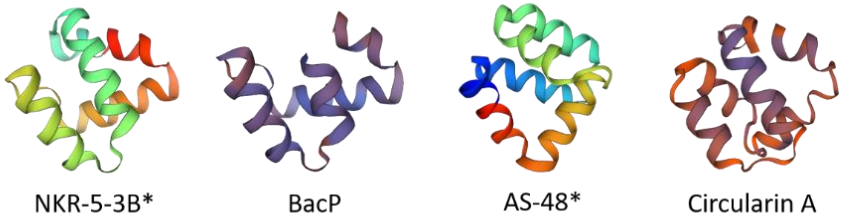


Figure 54- Predicted tertiary structure of BacP compared to other circular bacteriocins. The two reference structures are indicated with an asterisk (*) which are NKR-5-3B (2mp8.1 .A) and AS-48 (1083.1.A) deposited in the SWISS-MODEL template library. BacP and Circularin A were modelled on the templates respectively, to generate a predicted tertiary structure using SWISS-MODEL (Waterhouse *et al.*, 2018).

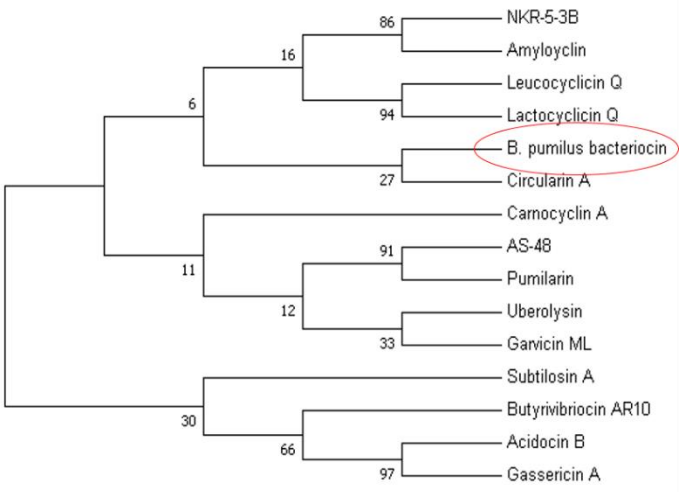


Figure 55- Phylogenetic tree of the putative *B. pumilus* 1_16 bacteriocin compared to the other circular bacteriocins. This phylogenetic tree was generated by MEGA7 (Kumar *et al.*, 2016) using ClustalW sequence alignment and maximum likelihood tree which has been bootstrapped 500 times. BacP is circled in red with its closest relative being Circularin A.



Figure 56- Amino acid sequence alignment of circular bacteriocins in subgroup i, and BacP. This ClustalW alignment demonstrates conserved resides above 50% conservation highlighted in green. A consensus sequence is indicated in the top row in bold, with the level of residue conservation demonstrated through bars with increasing conservation from blue to red. BacP sequence is shown in red containing all, except two residues, of the consensus sequence.

5.2.3.4 *In Silico* Analysis of a Putative Immunity Gene

Upstream of *bacP* is a gene potentially involved in bacteriocin self-immunity. This gene, *BpMccF*, is classified as a peptidase in the S66 family which has two well characterised members, MccF and LD-carboxypeptidases (LdcA), the latter of which are involved in peptidoglycan recycling through amide bond cleavage. BpMccF contains domains characteristic of both LdcA and MccF. MccF is a self-immunity protein for the class I *E. coli* bacteriocin microcin C (McC) which it detoxifies through cleaving the carboxamide bond between the McC peptide and its nucleotide moiety (Tikhonov *et al.*, 2010). MccF homologs have been found to form a separate clade,

within the S66 peptidase family, and have also been identified with the same detoxifying action in other bacteria, such as *Bacillus anthracis* (Nocek *et al.*, 2012). The *B. anthracis* MccF (BaMccF) has been compared to two LdcA peptides from *Pseudomonas aeruginosa* and *Novosphingobium aromaticivorans* and the original MccF peptide. It was observed that they all contain a catalytic triad of residues however BaMccF has a 27 residue insert that forms a long loop, which can also be seen in MccF (Fig 57). BpMccF also contains the catalytic triad and retains homology to BaMccF over this 27 residues insert (Fig 57). This, along with higher homology to both BaMccF and MccF (63% and 52% ID, respectively) (Table 21), suggests it belongs in the MccF clade rather than the LdcA clade.

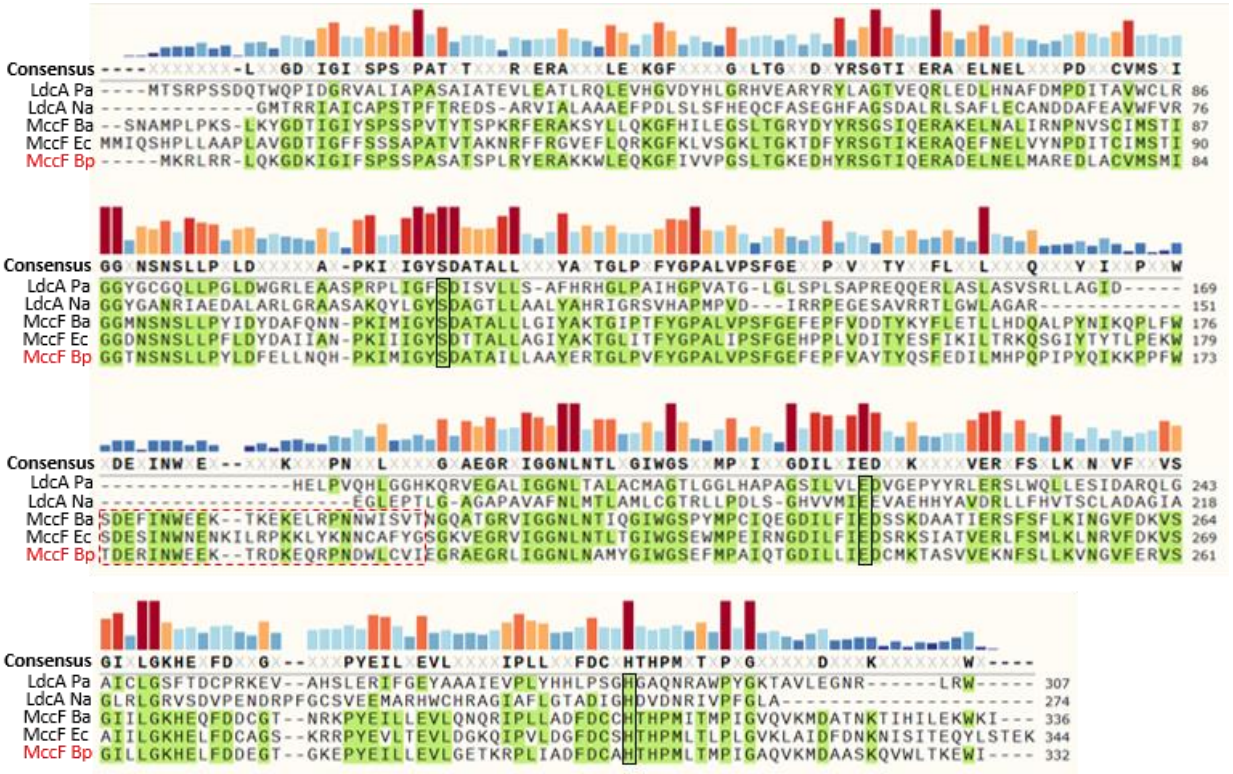


Figure 57- Amino acid sequence alignment of MccF from *B. pumilus* 1_16 compared to four S66 peptidase members. This ClustalW alignment demonstrates conserved resides above 50% conservation highlighted in green. A consensus sequence is indicated in the top row in bold, with the level of residue conservation demonstrated through bars with increasing conservation from blue to red. BpMccF sequence is shown in red containing the 27 residue insert indicated by the red dotted box and also the catalytic residues highlighted by the black boxes.

Table 21- The Sequence Similarity of MccF from *B. pumilus* 1_16 Compared to LdcA and MccF from Other Organisms.

B. pumilus MccF			
Versus	QC (%)	E-value	ID (%)
P. aeruginosa LdcA	78	1e-17	28.17
N. aromaticivorans LdcA	31	0.041	48
B. anthracis MccF	98	1e-166	63
E. coli MccF	98	9e-127	52

5.2.4 Initial activity

5.2.4.1 Overlays

Having identified a putative bacteriocin candidate within the genome of *B. pumilus* 1_16 the extent of the initial activity from *B. pumilus* 1_16 was then tested. Colony overlays were carried out against a range of *C. difficile* strains and antimicrobial activity was observed against all the *C. difficile* strains tested despite their differing PCR ribotypes, and the zone of clearance was measured to determine the range of activity observed (Fig 58, Table 22). The only strain which showed some inconsistent sensitivity was *C. difficile* CF5. It is also pertinent to note that the size of the colony overlaid made a difference in the activity seen against *C. difficile* R20291, with a small colony (<5 mm) not resulting in a zone of clearance.

Table 22- The Antimicrobial Activity of *B. pumilus* 1_16 Against Different Strains of *C. difficile*.

C. difficile strains	PCR Ribotype	Activity
CD305	023	+++
CD196	027	++
LIV29	001	++
LIV22	106	++
EK28	078	+++
CF5	017	+/-
R20291	027	+
CD630	012	++
DH1916	027	++
TL178	002	++
CD196	027	+
BI-9	001	++
BI-1	027	++
M68	017	+
M120	078	+++
TL176	014	++

The sensitivity of the strain relates to the size of the zone of clearance observed which were independently repeated three times. The measurement refers to the difference between the radius of the zone and radius of the inner colony, 0.2 cm +, 0.3-0.4 cm ++, 0.4-0.6 cm +++, no zone -.

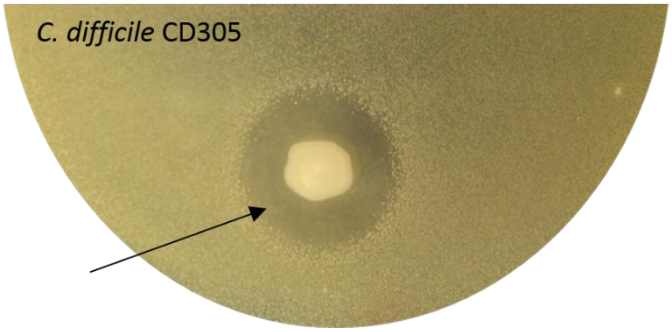


Figure 58- An overlay of *B. pumilus* 1_16 against *C. difficile* CD305 demonstrating a zone of clearance. A late exponential culture of the sensitive indicator *C. difficile* CD305 was inoculated to an OD_{600nm} 0.05 in soft-top agar which was overlaid against the producer strain which was an aerobically grown 16 hr *B. pumilus* 1_16 colony. A large zone of clearance can be observed (black arrow).

5.2.4.2 Activity Demonstrated in the Supernatant of B. pumilus 1_16

5.2.4.2.1 Demonstrating Antimicrobial Activity through Concentrating the Supernatant and Inducing Activity

To demonstrate activity in the supernatant, sterile filtered supernatant was concentrated using 10 kDa spin columns. This was carried out on ~16 hr growth of *B. pumilus* 1_16 grown in BHIS, TSB and Bacillus medium, however no zone was seen in either of the above or below 10 kDa fractions (data not shown). To try and induce bacteriocin production mitomycin C and UV were used to stress the growing cells. Mitomycin C, at a concentration of 3 µg/ml, was added to an exponentially growing culture and incubated for ~16 hrs or for 2 hr and resuspended into fresh medium and left incubating for ~16 hrs, however there was still no observed activity in the concentrated supernatant. Exposing exponentially growing cells to UV irradiation and lysed *C. difficile* CD630 cells also did not result in activity in the concentrated supernatant (data not shown).

5.2.4.2.2 Spent Supernatant Assays

Another method of demonstrating that the supernatant contains the component responsible for the activity was to try and grow the sensitive organism in the spent supernatant of the *B. pumilus* 1_16 strain. The sensitive strain was inoculated into the sterile filtered spent supernatant of the strain of interest to determine if there is any inhibition. The same medium, BHIS, was used for both strains and the strains were always inoculated back into their own spent supernatant to demonstrate a lack of growth is not due to nutrient deficiency. A positive control using fresh medium is also included to demonstrate a normal growth curve for that strain. *C. difficile* CD630, R20291, DH195, and CD196 were all inoculated into the spent supernatant of *B. pumilus* 1_16 and the growth of *C. difficile* was inhibited for all strains apart from R20291 (Fig 59). The growth of R20291 matched that of the positive control

which also demonstrates that there are sufficient nutrients left in the medium and the inhibition of growth in not due to nutrient deficiency (Fig 59). As previously mentioned it was observed that the size of the *B. pumilus* 1_16 colony on a plate overlay did affect the presence of a zone against R20291, therefore it might be possible that there is not a high enough concentration of the active compound to successfully inhibit the growth of R20291 in this assay. The inhibition of three out of four of the *C. difficile* strains tested shows that a component secreted into the supernatant is responsible for the activity observed against *C. difficile*.

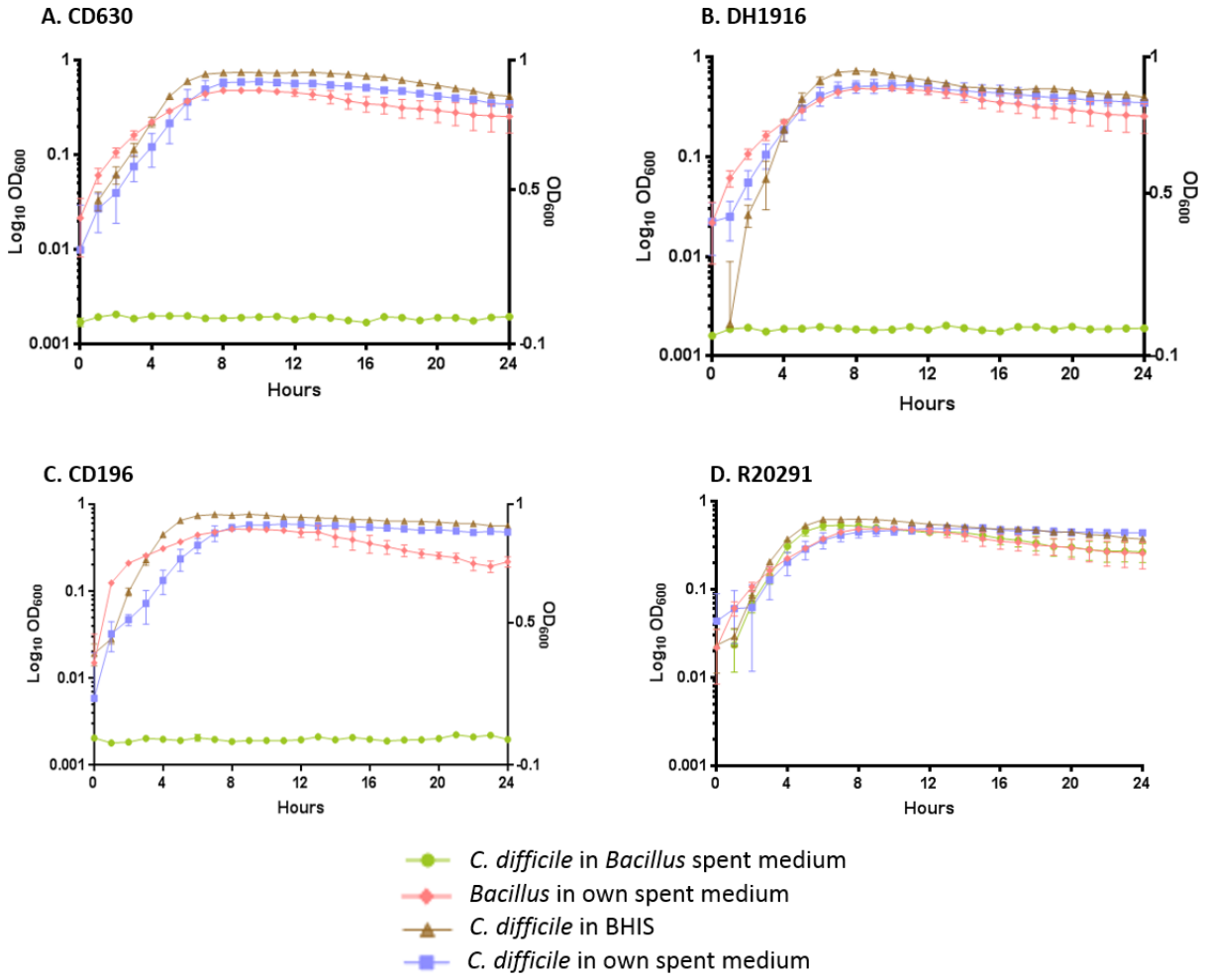


Figure 59- *B. pumilus* 1_16 spent supernatant assay against *C. difficile* CD630, R20291, DH196 and CD196. This microplate assay was incubated for 24 h at 37°C with shaking prior to OD_{600nm} readings. The inhibition of *C. difficile* CD630 (A), DH1916 (B), CD196 (C) in *B. pumilus* 1_16 spent supernatant (green) can be seen against a linear right Y axis, compared to the non-inhibited growth of the controls *C. difficile* grown in fresh BHIS, its own spent medium and *B. pumilus* 1_16 grown it its own spent medium (brown, blue and red respectively) on the logarithmic left Y axis. There is no inhibition of R20291 (D) as the growth in the spent supernatant (green) matches that of the controls (red, brown and blue). The error bars represent the standard deviation across nine repeats, three technical repeats for each independent biological repeat.

5.2.5 Determining the Active Peptide Secreted by *B. pumilus* 1_16

5.2.5.1 Whole Cell Lysate Overlay

A ~16 h culture of *B. pumilus* 1_16 was lysed using three different methods: sonication, Bugbuster® Master Mix and freeze-thaw. The three different methods were used to determine the most effective and gentle method of cell lysis that optimised release of undamaged protein. All the lysates were run in duplicate on a Bis-Tris 10% SDS-PAGE gel, with one half of the gel coomassie stained and the other overlaid with *C. difficile* CD630 or R20291 to observe any activity. However, no zone of clearance was seen. This could be due to the type of SDS-PAGE gel which only resolved proteins down to 5 kDa whereas some classes of bacteriocins are known to be smaller than 5 kDa. This is especially pertinent as a likely bacteriocin candidate from the *in silico* analysis suggested a class of bacteriocins known to be as small as 3 kDa. The abundance of the protein within the cell could also be too low to visualise and generate a zone of clearance, as bacteriocins are usually secreted from the cell to act against other bacteria. Therefore, concentrating the proteins from the supernatant would be required which can be carried out using multiple techniques; however, it is well known that bacteriocin isolation from media is difficult (Saavedra and Sesma, 2011).

5.2.5.2 Protein Precipitation from the Supernatant

5.2.5.2.1 Negatively Charged Peptide Purification

A characterised head-to-tail cyclised bacteriocin produced from *B. pumilus* was previously isolated using a resin that binds negatively charged peptides (van Heel *et al.*, 2017). This CM Sepharex® C-25 resin was rehydrated and incubated with the spent supernatant of *B. pumilus* 1_16 (grown to late exponential phase) for 1 h. The resin was then packed into an empty column and the bound negatively charged peptides were eluted from the resin using a high salt concentrated elution buffer. The

eluent was then desalted using PD-10 columns. The resultant desalted eluent was checked for antimicrobial activity using a well overlay but no antimicrobial activity against *C. difficile* was observed (data not shown).

5.2.5.2.2 Acid Protein Precipitation

Acid protein precipitation has previously (Elegado *et al.*, 1997; Vera Pingitore *et al.*, 2007) been used to purify bacteriocins as it has been determined that they typically fully absorb to the producer at a range of pH from 5-7. In order to determine if this technique could be used so *B. pumilus* 1_16 two different absorbtion pH's were tested, pH 5.5 and pH 6.5 to allow peptides to be absorbed on to the cell surface. Any absorbed proteins where desorbed for 1 h in acidic pH's ranging from 1.1 to 2.5. Samples were dialysed into Tris buffer pH 7.5 to prevent false positives due to the acidic pH and the samples were tested for activity using well overlays, however no activity was observed (data not shown).

5.2.5.2.3 Solvent Precipitation

Another protein precipitation method investigated was the use of the solvents acetone and ethanol to precipitate out any proteins within the supernatant. Ice cold acetone and ethanol were mixed with cell free supernatant, at a ratio of 1:4 and 1:9 respectively and incubated overnight at -20 °C to precipitate out any proteins. After venting the sample to completely remove any trace of solvents samples were resuspended in Tris buffer. The samples were tested for activity against *C. difficile* CD630 using a well overlay. No activity was observed (data not shown).

5.2.5.2.4 Salt Precipitation

A common precipitation method which is known to be less damaging to proteins is salt precipitation using ammonium sulphate. This has been previously used for purifications of many bacteriocins (Saavedra and Sesma, 2011). Cell free supernatant of *B. pumilus* 1_16 was obtained at early stationary/late exponential

phase of growth and proteins were precipitated out at 80% w/v. Precipitation was carried out overnight and the proteins were dialysed against Tris buffer, pH7.5 to remove any ammonium sulphate to prevent false positives and the distortion of proteins running on an SDS-PAGE gel. The precipitated proteins were tested for activity against *C. difficile* CD630 using a well overlay, however no zone of clearance was observed at 80%, therefore the method was optimised to determine if the active peptide was precipitated out at a different percentage of ammonium sulphate. A zone of clearance was observed at only 40% ammonium sulphate (w/v) (Fig 60), and not at 20%, 60% (data not shown) or 80% (Fig 60). This confirms that a peptide that precipitated out at 40% w/v ammonium sulphate is secreted from *B. pumilus* 1_16 has antimicrobial activity towards *C. difficile*. The precipitated protein was tested at two different dilutions for inhibition of *C. difficile* CD630 growth in liquid BHIS medium. Whilst the 1/100 dilution (Fig 61b) showed no inhibition of growth, the 1/10 dilution (Fig 61a) completely inhibited the growth of *C. difficile* CD630 over the complete time course of 24 hr (Fig 61).

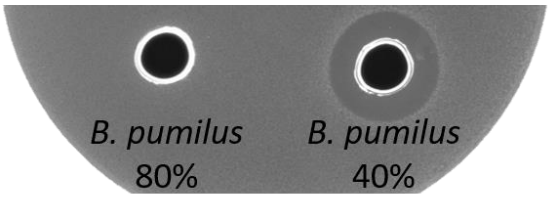


Figure 60- A well overlay of protein precipitated from *B. pumilus* supernatant using different percentages of ammonium sulphate. A zone of clearance is observed in the 40% fraction and not the 80% fraction against *C. difficile* CD630, demonstrating a protein active against *C. difficile* is present in this 40% fraction.

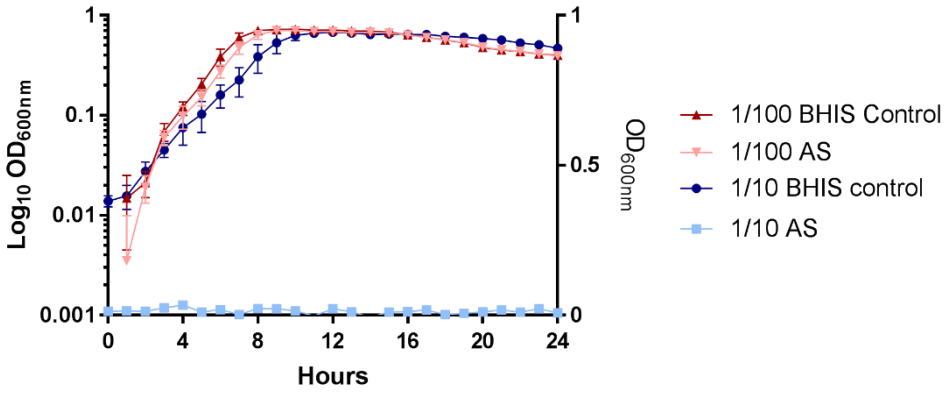


Figure 61- Inhibition results of *C. difficile* CD630 growth by 40% ammonium sulphate precipitated proteins at different dilutions. The growth of *C. difficile* CD630 is inhibited by the precipitated proteins (light blue) on the right linear Y axis diluted 1/10 into BHIS compared to the growth of the control (dark blue) on the left Y axis in Log_{10}. There is a lack of inhibition by the same precipitated proteins diluted by 1/100 (pink) as compared to the growth of the control (dark red).

To isolate the protein of interest from within the 40% fraction, the precipitated proteins were run on a 10% Tris-tricine SDS-PAGE gel in triplicate. One third of the gel was coomassie stained, the other was silver stained, and the last third was overlaid with *C. difficile* CD630 (Fig 62). A zone of clearance was observed on the overlaid third of the gel, demonstrating antimicrobial activity against *C. difficile*. The adjacent lanes, of 20% ammonium sulphate precipitated proteins and Tris buffer negative control show no zone of clearance. This demonstrates that the ammonium sulphate is not causing a false positive and neither is the resuspension buffer. The zone of clearance does not correspond to a band in the coomassie stained third of the gel (Fig 62). It does, however, correspond with the same shape and size of a peptide band visualised through silver staining, which is a peptide less than 5 kDa. The band was excised and sent for MS/MS sequencing however it was too contaminated to give a definitive sequence or match the sequence of the *in silico* bacteriocin candidate (Data not shown).

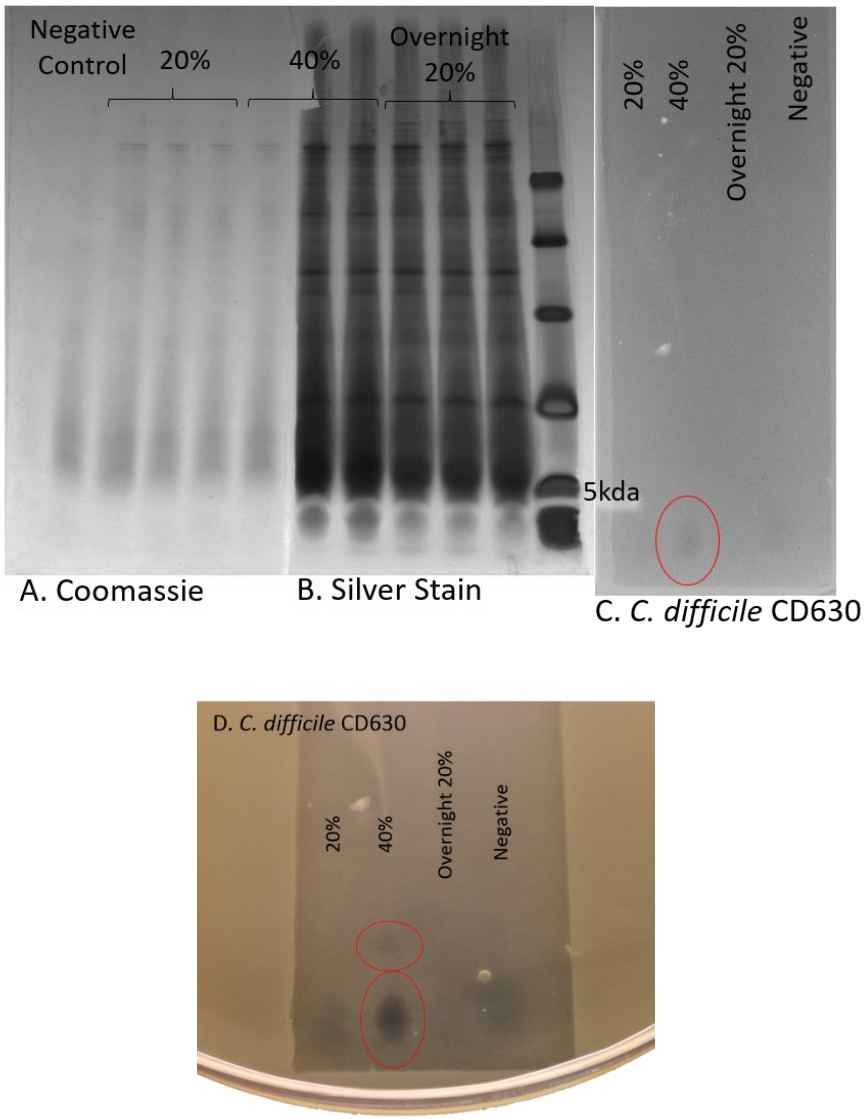


Figure 62- A zone of clearance on an overlaid SDS-PAGE gel corresponding to a protein band on the silver stained part of the SDS-PAGE gel. A. is the coomassie stained third of the SDS-PAGE gel with the negative control, 20% fraction and 40% fraction which does not show a protein band corresponding to the zone of clearance seen. B. the silver stained third of the SDS-PAGE gel showing fractions 40% and 20% precipitated overnight. C. the overlaid third of the SDS-PAGE gel, showing a zone of clearance matching to a protein band in the silver stained gel (red circle). It also shows a lack of zone in both 20% fractions and the negative control. D. the overlaid third again zoomed in on the section of the gel with an observable zone of clearance. There is no zone in the negative control, which is ammonium sulphate precipitated non-inoculated BHIS medium, however there is a blue tinge which is a result of the SDS-PAGE sample reducing buffer.

5.2.5.3 Summary of Results Generated from Supernatant Precipitation Experiments

The ammonium sulphate precipitation method successfully demonstrated a peptide under 5 kDa caused antimicrobial activity against *C. difficile*. However, the precipitation of the active peptide proved unreliable as antimicrobial activity could not be demonstrated after every precipitation attempt. Although ammonium sulphate precipitation is known to be unreliable, different precipitation conditions were controlled for in order to increase the reliability of the method. The pH of the supernatant was controlled, the percentage of ammonium sulphate, the temperature of precipitation, the length of precipitation and all buffers for resuspension and dialysis were all controlled. The active peptide could have been lost in the process through binding to the filter paper, or the dialysis tubing however these two were ruled out as the filter paper showed no antimicrobial activity and the dialysis tubing was switched to PD-10 desalting columns and the eluent also showed no activity.

To overcome the unreliability of the ammonium sulphate protein precipitation method, the gene encoding the putative bacteriocin identified through *in silico* analysis, *bacP*, was cloned into an *E. coli* expression host. This approach was to determine if *bacP* could be cloned into *E. coli* and the protein successfully overexpressed. If the protein was overexpressed, then it could be checked for antimicrobial activity against *C. difficile*, confirming that this putative bacteriocin is responsible for the observed activity.

5.2.6 Determining the Active Peptide through Expression in an *E. coli* Host

5.2.6.1 Creation of a BacP Overexpression Vector

The creation of the BacP overexpression vector was carried out in the same manner as previously described (Section 3.2.4.1). The identified candidate gene, *bacP* was amplified from the *B. pumilus* 1_16 genome using the primers *bacP* F/*bacP* R, which incorporated the restriction enzyme sites *Xho*I and *Nde*I into the 5' and 3' ends respectively. The T7 expression vector pET28b (+) was linearised in the multiple cloning site (MCS) with restriction enzymes *Xho*I and *Nde*I and the DNA backbone gel purified. The amplified PCR gene product was digested with the same enzymes and gel purified. The vector backbone and insert were ligated together to create pET28b (+)*bacP* which was transformed into *E. coli* Top10 and confirmed with Sanger sequencing using the T7 Terminator and T7 promoter primers. The confirmed plasmid was then transformed into the first *E. coli* expression strain, C43 (DE3) pLysS, and re-confirmed with Sanger sequencing. Due to the circular nature of the bacteriocin either an N- or C- terminal His-tag could interfere with the bacteriocins structure. The N-terminal His-tag was chosen as it could be cleaved off using the adjacent thrombin site. However due to the typical biosynthesis process of circular bacteriocins the His-tag could be lost with the cleaving of the leader sequence meaning His-tag purification would not be effective.

5.2.6.2 Overexpression of BacP

Overexpression of BacP*,* was trialled in several *E. coli* expression strains to try and demonstrate antimicrobial activity against *C. difficile*. The only observed activity was in *E. coli* C43 (DE3) pLysS using the His-tag to purify the protein. This strain carrying the BacP overexpression vector was grown in duplicate until an OD_{600nm} ~1 and one culture was induced with 1 mM IPTG and the other was inoculated with 0.2% glucose as a non-induced negative control. Both cultures were incubated at 37 °C for 3.5 h

before being pelleted and lysed by sonication. The cell lysate was run on a nickel resin column to determine if any His-tag protein was present and demonstrating activity. Any resin bound proteins where eluted off the resin using 500 mM imidazole and the fractions were run on a 16% Tris-tricine SDS-PAGE gel with the induced and non-induced elute run in duplicate on the same SDS-PAGE gel. It had previously been determined that 500 mM imidazole did not inhibit the growth and cause a zone of clearing against *C. difficile* CD630 (data not shown). The duplicated elutes on one half were overlaid with *C. difficile* CD630 and showed a zone of clearance, however this did not match any band on the silver stained half of the SDS-PAGE gel (data not shown). This could be due to the low abundance of the protein therefore western-blots were carried out using antibodies against the His-tag, and visualised using both colorimetric and chemiluminescence, however although the positive control worked no other band was observed (data not shown). The purification was repeated which again demonstrated a zone of clearance, however a zone was also shown in the non-induced control (data not shown) which could have been due to leakiness in the expression system or an indication something else was causing a false positive. Therefore, it is inconclusive if BacP was overexpressed and resulted in antimicrobial activity.

Due to the potential for the loss of the N-terminal His-tag, late exponential *E. coli* C43(DE3) pLysS pET28b (+)*bacP* cells were induced for 4 h at 30 °C, harvested and sonicated. The sterile filtered soluble fraction was incubated shaking with CM-Sephadex® C-25 for 1 h, and any cationic peptides present in the soluble fraction, were eluted off the resin using a high concentration of salt. The salt was removed from the eluent using a PD-10 desalting column however no activity was observed in the eluent (data not shown). This could be due to incorrect processing of BacP by the *E. coli* host, or due to protease degradation or insufficient protein overexpressing.

5.2.6.3 Preliminary Results from the Expression of the Entire BacP Locus in an *E. coli* Host

The *in silico* analysis combined with the current literature on circular bacteriocins highlighted a bacteriocin cluster. This putative bacteriocin locus includes the BacP structural gene (*bacP*), an ABC transporter, and four hypothetical genes. This cluster of genes, including the putative terminators, was amplified using *bacP* locus F/*bacP* locus R primers, which included the restriction sites *Eco*RI and *Xho*I onto the 5' and 3' prime ends respectively. This was then cloned into an *E. coli* compatible vector, pMTL84123, under the control of P_{fdx} which is known to function in *E. coli*. The resulting plasmid, pMTL84123_*bacP* (Fig 63), was then transformed into *E. coli* Top10 and transformants screened using M13 F and M13 R primers. Two confirmed transformants were colony overlaid with *C. difficile* CD630 to observe any antimicrobial activity compared to an empty plasmid negative control. Preliminary results demonstrate antimicrobial activity against *C. difficile* CD630 compared to the control as a zone of clearance can be observed for both independent biological repeats (Fig 6). In conjunction with the *in silico* data, this suggests that the *bacP* gene encodes a putative bacteriocin with activity against *C. difficile* however it requires the expression of the entire locus to be active. This is only preliminary data as it has not been fully repeated due to time constraints, although there are two independent biological repeats and a negative control suggesting the results are reliable.

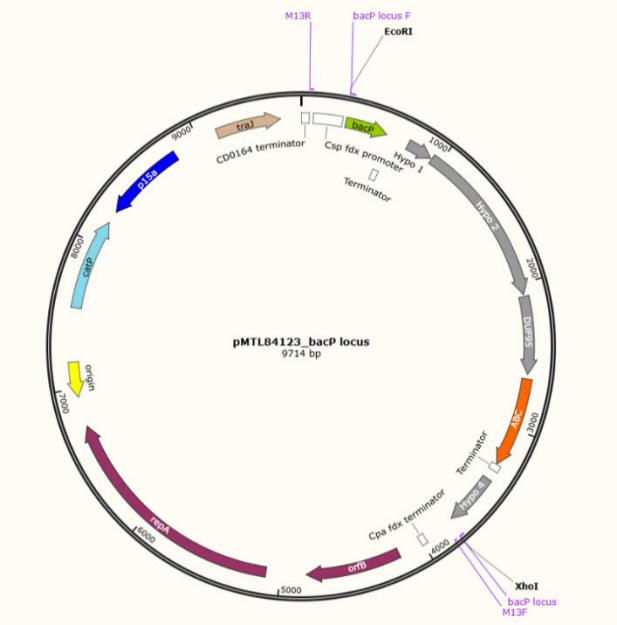


Figure 63- Plasmid map of pMTL84123_*bacP*. Plasmid map of the vector pMTL84123 with the *bacP* locus cloned into the MCS using the *Eco*RI/*Xho*I restriction sites. The cluster amplification primers (*bacP* locus F and *bacP* locus R) and the screening primers (M13 F and M13 R promoter) are indicated in purple. The *bacP* structural gene is shown in green, with the hypothetical genes 1-4 are in grey and the ABC transporter gene in orange.

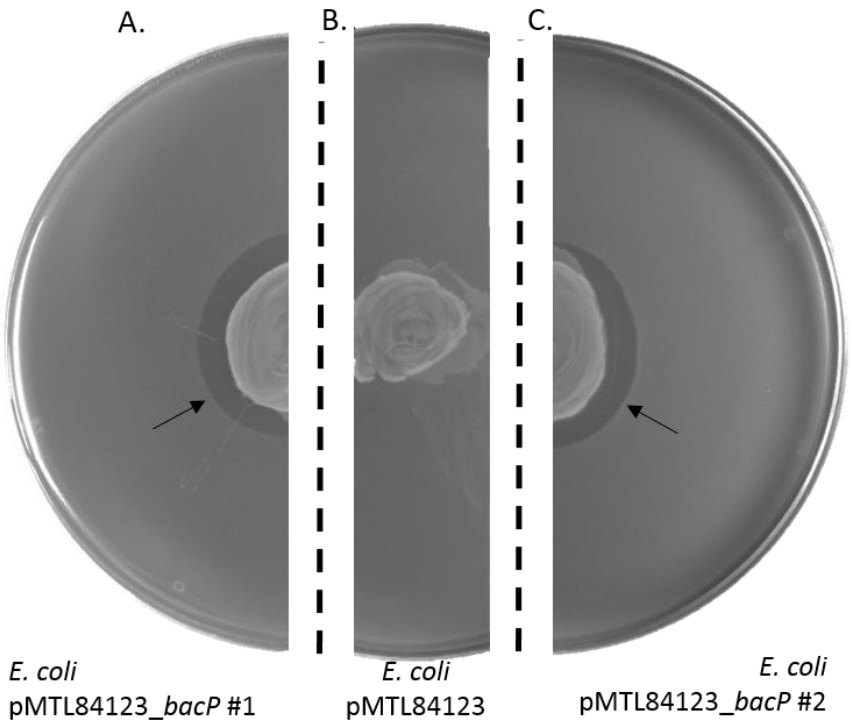


Figure 64- Colony overlay of two biological repeats of *E. coli* pMTL84123_*bacP* against *C. difficile* CD630 compared to a negative control. (A & C) A colony overlay of two independent biological repeats of *E. coli* pMTL84123_*bacP* (#1 and #2 respectively) which demonstrates a zone of clearance (arrow) against a lawn of the indicator strain *C. difficile* CD630. (B) A colony overlay of the *E. coli* negative control harbouring the empty plasmid pMTL84123 demonstrating no zone of clearance against a lawn of *C. difficile* CD630.

5.2.7 Identifying the Nature of the Antimicrobial Activity

An alternative approach to determine if the antimicrobial activity observed from *B. pumilus* 1_16 against *C. difficile* was caused by a bacteriocin was to create a resistant *C. difficile*. Circular bacteriocins are known to have self-immunity genes which protect the host from the action of a produced bacteriocin. *In silico* analysis identified a putative self-immunity gene, *bpMccF,* in the genome of *B. pumilus* 1_16. To investigate the role of this gene and the nature of the observed antimicrobial activity, *bpMccF* could be expressed in sensitive *C. difficile* strains. Any resulting insensitivity to the previously observed activity would confirm the activity was due to a putative bacteriocin and the self-immunity role of *bpMccF*.

5.2.7.1 Creating the Self-Immunity Gene Expression Vector

Through *in silico* analysis, BpMccF was highlighted as a putative bacteriocin self-immunity gene. Previous work in section 3.2.5.2 demonstrated that known self-immunity genes can be expressed in *C. difficile* strains resulting in bacteriocin resistant *C. difficile*. This concept was trialled for BpMccF, using the same cloning method described in Section 3.2.5.2.1, and it was incorporated into pMTL84123 and transformed into *E. coli* Top10 and DH5α. No positive colonies were identified when screening which could be due to the nature of BpMccF, a protease which placed under the control of P_{fdx} could be expressed in *E. coli* and be toxic to the cells. Therefore, a second vector was generated which contained the *tcdB* promoter, a promoter for the *C. difficile* toxin B gene, which functions in *C. difficile* but not in *E. coli.* The pMTL84121_PtcdB vector was generated by removing the P_{fdx} from pMTL84123 using the restriction enzymes *Not*I and *Nde*I. The P_{tcdB} promoter was then amplified using primers PtcdB F and PtcdB R which incorporated an overhang of 25 bp at the 5' and 3' ends, respectively. Both these overhang regions were homologous to the pMTL84123 linear vector enabling the new promoter to be incorporated using Gibson Assembly.

The ligated plasmid was transformed into *E. coli* DH5α. Positive transformants were confirmed by Sanger sequencing and the confirmed plasmid was linearised using restriction enzyme *Nde*I and gel purified. The BpMccF fragment was amplified to incorporate regions of homology to pMTL84121_P_{tcdB} at both 5' and 3' prime ends using the primers PtcdB_MccF F (Fig 65a) and pMTL_MccF R (Fig 65b). The amplified BpMccF product and the linear pMTL84121_PtcdB vector were ligated together generating pMTL84121_P_{tcdB}_BpMccF (Fig 66), using the Gibson assembly method and transformed into *E. coli* DH5α. Any transformants were screened for the BpMccF insert using the M13 F and M13 R screening primers however none of the transformants were positive. The ligated plasmid was also transformed into *E. coli* Top10 and *E. coli* CA434 however no positive transformants were identified.



Figure 65- Primers for the assembly of pMTL84121_PtcdB_*BpMccF*. The primers used for the cloning of the *bpMccF* self-immunity gene into the pMTL84121_PtcdB expression vector are indicated in the purple boxes (PtcdB_BpMccF F (A) and pMTL_BpMccF R (B)) which overlap, by 25 bp (capital letters), across the P_{tcdB} (A) and the plasmid backbone (B) respectively.

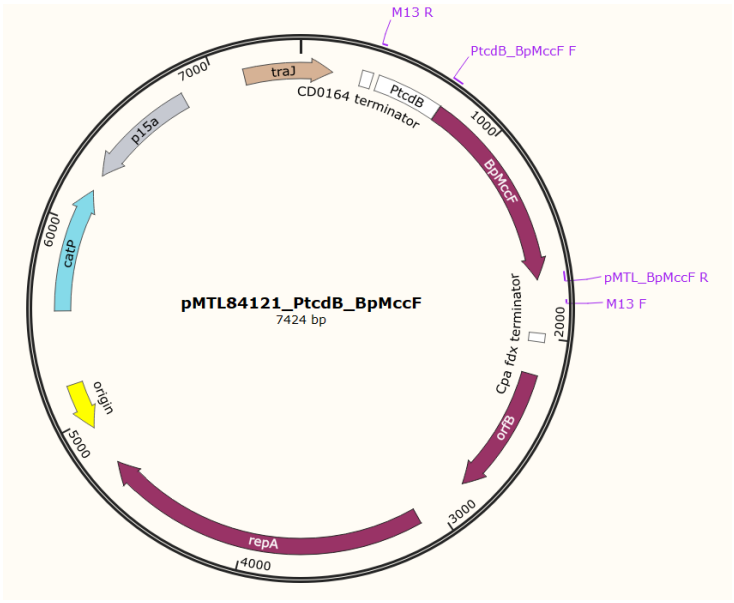


Figure 66- Plasmid map of the self-immunity gene, *bpMccF*, expression vector. Vector pMTL84121 with the *bpMccF* gene cloned into the MCS via Gibson assembly using the primers PtcdB_BpMccF F and pMTL_BpMccF R (indicated in purple). The screening primers (M13 F and M13 R promoter) are indicated in purple.

5.3 Discussion

5.3.1 Characterisation of Antimicrobial Activity Observed from *B. pumilus* 1_16 Against *C. difficile*

5.3.1.1 Evidence that the Activity is Due to a Peptide

Initial antimicrobial activity was demonstrated from a Gram-positive bacillus against three different strains of *C. difficile*. This strain was identified through 16s rRNA and whole genome sequencing as *B. pumilus* 1_16. Further testing demonstrated antimicrobial activity against a wide range of *C. difficile* strains across different PCR ribotypes. The strength of activity also ranged across the different *C. difficile* strains with CD305 being very sensitive and CF5 demonstrating inconsistent sensitivity. Although the overlaid soft-top agar was inoculated to the same OD_{600nm} each time (0.05), the *B. pumilus* 1_16 colony size was not well controlled. Thus overlaying different size colonies could determine if the concentration of active compound plays

a role in the inconsistency of sensitivity observed from CF5. The activity observed was most likely not due to the intact phage within the genome of *B. pumilus* 1_16 as no activity was observed in the supernatant after mitomycin C induction. *C. difficile* phage are known to have narrow host range and be carried by other *C. difficile* strains (Hargreaves and Clokie, 2014), and since activity was observed against a wide range of *C. difficile* this also counteracts the known information on phage activity suggesting the phage are not responsible for the observed activity.

The antimicrobial activity against *C. difficile* was also demonstrated in liquid through the inhibition of *C. difficile* growth in the spent supernatant of *B. pumilus* 1_16. There was clear inhibition of *C. difficile* strains over a 24 h period. It is interesting to note that, although there was inhibition of *C. difficile* CD196, DH1916 and CD630, there was no inhibition of *C. difficile* R20291. A zone of clearance was repeatedly observed from a colony overlay against *C. difficile* R20291, however it was observed that the size of the zone was dependent on the size of the *B. pumilus* 1_16 colony overlaid. This could suggest that a larger concentration of active compound is required for inhibition of *C. difficile* R20291 which was not observed for other *C. difficile* strains, apart from the inconsistent sensitivity of *C. difficile* CF5. The lack of inhibition of *C. difficile* R20291 does suggest that the inhibition is not a side effect from nutrient deficiency from using medium that another strain has grown in. The inhibition of *C. difficile* growth also indicates that the active component was produced and secreted into the BHIS medium by mid-exponential phase, as the spent supernatant was harvested at this growth phase. The active compound was also still present in cell free supernatant suggesting it is not lost during filtering with 0.22 µM filter, as some peptides can bind filters, resulting in a loss of peptide and activity.

As the *in silico* analysis suggests a putative circular bacteriocin candidate, different bacteriocin purification strategies were trialled in order to purify the active component in the *B. pumilus* 1_16 supernatant. Circular bacteriocins have previously been successfully precipitated from supernatant using multiple different conventional techniques such as salting out and cation-exchange (Gálvez *et al.*, 1986; Kemperman *et al.*, 2003). As there has previously been a head-to-tail cyclised bacteriocin purified from a strain of *B. pumilus*, using cation-exchange this was first trialled for *B. pumilus* 1_16 (van Heel *et al.*, 2017). CM Sephadex® C-25 resin binds cationic peptides and was incubated with sterile filtered supernatant from *B. pumilus* grown to late exponential phase. The eluent however demonstrated no activity despite undergoing desalting which can reduced bacteriocin (Saavedra and Sesma, 2011; van Heel *et al.*, 2017). To determine if the active compound was lost during the process, the flow through, wash fractions and the resin itself were checked for activity but no zone of clearance was observed. In hindsight these fractions were not desalted and therefore the compound could have been present but the high salt concentrations prevented observable activity. The compound could also not be a cationic peptide and therefore may not bind the CM Sephadex® C-25 resin, or the process required more optimisation to successfully purify the active compound. Conditions that could have been optimised are, the growth phase in which the resin was added, the pH of desalting buffers and the concentration of salt in the elution buffer.

Acid purification has also been successfully described for bacteriocins such as pediocin AcH, nisin, and sakacin A (Yang *et al.*, 1992; Saavedra and Sesma, 2011). It was demonstrated that a range of pH 5-7 was required for the absorption of bacteriocins onto the cell membrane of the producer and then lowering the pH allowed the bacteriocin to be desorbed. Two different absorption pH, 5.5 and 6.5, were used and the desorption pH's were 1.1-2.5 with subsequent dialysis into a neutral pH

solution using 1 kDa MWCO. Unfortunately, no antimicrobial activity was observed, however further absorption pH's could be trialled to determine if the lack of activity is due to lack of method optimisation. Also acid precipitation is known not to be suitable for all bacteriocins. However it has been successfully used as the purification strategy for lactocyclicin Q, a subgroup i circular bacteriocin (Sawa *et al.*, 2009),. suggest it is suitable for circular bacteriocins. It should be noted that lactocyclicin Q is an outlier within this group and less homologous to the putative bacteriocin candidate, BacP (Section 5.2.3.3, Fig 55, 56). Due to the bias of CM Sephadex® C-25 resin (cationic peptides only) and acid purification (cell membrane targeting peptides) toward bacteriocins more general techniques were also tested.

Solvent precipitation methods are a more general protein precipitation technique that have also been previously described for the precipitation of bacteriocins. The limitation of this technique is that the solvents could inactivate active peptides, however it has been successfully used for active peptides such as bacteriocins previously (Marciset *et al.*, 1997). Two different solvents, ethanol, and acetone, were used to precipitate the active component from the supernatant of *B. pumilus* 1_16 however these both resulted in no activity. As the active compound is secreted during exponential phase, as shown previously, it should have been present however it could have been too low a concentration or inactivated by the solvents. A widely used technique for precipitating active peptides is salt precipitation using ammonium sulphate, however it is known to be an unreliable technique which is why this strategy was not tested first.

Salt precipitation using ammonium sulphate has previously been demonstrated to work for numerous bacteriocins such as plantaricin NC8 (Maldonado *et al.*, 2003), enterocin 1071 (Balla *et al.*, 2000), lactococcin G (Nissen-Meyer *et al.*,

1992) and circular bacteriocins including butyrivibriocin AR10 (Kalmokoff *et al.*, 2003), uberolysin (Wirawan *et al.*, 2007) and circularin A (Kawai *et al.*, 2004). This technique was deemed appropriate due to peptides homology to circularin A and non-denaturing nature of it. A sequential increase of ammonium sulphate was used to precipitate proteins out of cell free supernatant from late exponential *B. pumilus* 1_16 growth and the residue salt was removed through dialysis. There was observed antimicrobial activity in the 40% ammonium sulphate fraction and no observed activity in the 20%, 60% and 80% fractions and in the negative controls corresponding to the same fraction. The antimicrobial activity within the 40% fraction was confirmed through the addition of different dilutions to medium inoculated with *C. difficile*. The inhibition of growth spanning 24 h period at a 1/10 dilution provides further evidence that the active peptide was precipitated out of the supernatant. This demonstrated that a secreted peptide precipitated out of the supernatant using 40% ammonium sulphate causes the previously demonstrated antimicrobial activity.

To characterise the precipitated active peptide, the 40% fraction was visualised on an SDS-PAGE gel which demonstrated that the activity was due to a low molecular weight peptide, less than 5 kDa. This is in agreement with the known low molecular weight of circular bacteriocins with circularin A being 2.2 kDa. It must be noted that the peptide band observed on the silver stained SDS-PAGE gel was a smear and thus it is difficult to precisely determine the size. It is also evident that there is a second smaller zone of clearance on the overlaid section of the SDS-PAGE gel which could be a dimerization of the smaller peptide or a completely different active peptide. Although this band was excised for MS/MS the band was too contaminated to get an accurate sequence. This is reported to be common for this type of precipitation as it usually results in a crude purification and could have been hindered by the rich growth

medium. Typically, bacteriocins undergo further purification before MS/MS sequencing to prevent this type of contamination.

Further work with this precipitated peptide was planned however the precipitation was extremely unreliable and the active peptide was not precipitated out successfully during most attempts. In order to optimise the precipitation different conditions were controlled, the pH of the supernatant was always standardised to a pH of 5, the duration and temperature of precipitation was also controlled. The dialysis tubing was also changed from a MWCO of 3.5 to 1, to ensure reduced loss of peptide however the method still proved unreliable therefore the peptide was not characterised further. However, if more peptide was obtained the planned characterisation would be to conduct peptidase experiments to determine its resistance pattern, also heat stability would be investigated. Further purification using cation-exchange would also be optimised to obtain enough pure peptide for MS/MS sequencing to match the sequence with the *in silico* putative bacteriocin candidate, BacP. As the salt protein purification technique could not be further optimised different approaches were explored to investigate the role of BacP in the observed antimicrobial activity.

5.3.1.2 Bacteriocin Nature of Activity

To investigate the nature of the antimicrobial activity demonstrated by *B. pumilus* 1_16, one approach was to generate a resistant *C. difficile* strain. This approach has been previously demonstrated and described in section 3.2.5.2. A putative bacteriocin self-immunity gene was identified through *in silico* analysis displaying homology to MccF, a microcin C self-immunity peptide. A previous study demonstrated that standalone MccF homologous genes have serine peptidase activity, cleaving carboxyamide bonds (Nocek *et al.*, 2012). As there are asparagine residues, thus the potential for carboxyamide bond formation, within BacP it could be a

potential target for BpMccF and should therefore be considered as a possible self-immunity gene. Therefore, if BpMccF was a putative bacteriocin self-immunity gene and its expression in a previously sensitive *C. difficile* strain resulted in a resistant strain then the previously observed antimicrobial activity was most likely due to bacteriocin activity. However, the self-immunity gene candidate, BpMccF, could not be cloned into an expression vector. This was initially thought to be due *E. coli* conjugation strain not tolerating the potential low-level expression from the *fdx* promoter however, the vector could also not be generated when the promoter was switched to an *E. coli* incompatible promoter. To generate this vector and thus test the hypothesis further attempts to generate the non-leaky vector through screening more *E. coli* colonies could be carried out. The gene could also be cloned into the vector under the control of no promoter to determine if any read through is causing problems, if so the vector could be designed to prevent this by placing the gene in the opposite direction. Another method would be to use an inducible system such as a riboswitch system, RiboCas, which is used to express lethal genes in *Clostridium* spp. as the expression of the genes are under tight regulation and only translated when induced (Cañadas *et al.*, 2019).

5.3.2 Preliminary Evidence that BacP is a Putative Bacteriocin

5.3.2.1 *In Silico* Evidence

Mining the genome of *B. pumilus* 1_16 using a bacteriocin database, BAGEL4, highlighted a putative bacteriocin, BacP. The amino acid sequence of BacP contains a circularin A/uberolysin domain placing it within the head-to-tail cyclical class of bacteriocins. The lack of cysteine residues and predicted Saposin-like tertiary structure of is typical of the characteristics of circular bacteriocins (Towle and Vederas, 2017). The amino acid sequence alignment of subgroup i members indicated a limited consensus sequence of ten conserved residues. BacP contains eight of these conserved

residues which equally as many as circularin A, the bacteriocin which has the highest homology to BacP. The initial *in silico* analysis suggests that BacP conforms to the typical characteristics of circular bacteriocins in subgroup i.

The genetic organisation of characterised circular bacteriocin gene clusters have been previously investigated leading to limited characterisation of typical motifs (Gabrielsen *et al.*, 2014a). The gene clusters around the putative bacteriocin encoding gene, *bacP*, were analysed and compared to the known motifs of circular bacteriocins. Typically, circular bacteriocins have between five to seven essential genes which are not well characterised. The five genes in the vicinity of *bacP* mostly encode hypothetical proteins with no known homology, however there are three genes that fit the required categories. There is a small cationic peptide, a DUF59 domain containing gene, and an ABC transporter gene however an ATPase and a putative membrane protein with unknown function could not be identified within this cluster. This could be due to the lack of homology and limited characterisation of known circular bacteriocins or the region might not conform to the known typical circular bacteriocin cluster. Another possibility is the region is incomplete due to an error in the genome assembly process or it is not a circular bacteriocin gene cluster. The area in question could be re-sequenced to rule out any missing genes due to an artefact of the genome assembly pipeline. Due to the limited characterisation of the genetic organisation of circular bacteriocin clusters it is extremely difficult to determine if this fits within the class, however it should not be ruled out as a potential region encoding a bacteriocin as it contains some characteristic motifs.

5.3.2.2 *E. coli* Expression of BacP and the BacP Locus

In order to investigate if the BacP peptide was responsible for the observed antimicrobial activity against *C. difficile* it was cloned into an *E. coli* expression host. Previous studies have demonstrated successful expression of bacteriocins by cloning the encoding gene into an overexpression vector and trialling different *E. coli* expression strains. Although successful expression could not be confirmed through SDS-PAGE or western blot for all *E. coli* expression strains used, there was a zone of clearance observed on an SDS-PAGE overlaid with *C. difficile* CD630 whilst using *E. coli* (DE3) C43 pLysS. This result could not be repeated reliably and therefore the role of BacP remains inconclusive. Although different *E. coli* expression strain were trialled, no overexpression could be confirmed, this could be due to the loss of the His-tag. If the leader sequence of the bacteriocin was cleaved off the His-tag would have been lost and therefore a His-tag purification and western blot would not have been successful. However, if the His-tag was lost then theoretically there should have been antimicrobial activity observed in the flow-through and wash fractions of the His-tag purification which was not seen.

Further experiments to express the entire predicted genetic region of BacP within *E. coli* did result in antimicrobial activity against *C. difficile* CD630. Further work needs to be undertaken to ensure that the results are reliable through repeating them and screening for activity against other strains of sensitive *C. difficile*. To confirm that the active peptide is BacP, rather than any other expressed peptides, running the cell lysate on an SDS-PAGE gel and overlaying it with the sensitive *C. difficile* CD630 strain should highlight an active peptide band. This band of peptide could be analysed through LC-MS/MS to elucidate the peptide sequence. The peptide could also be further purified from the cell lysate such as through ammonium sulphate precipitation or cationic CM Sepharex® C-25 resin to remove any contaminants. Once the active

peptide is purified its properties could be further characterised such as pH and heat stability, protease resistance, and its structure could be elucidated through crystallography to determine if it does conform to the typical circular bacteriocin structure that was predicted *in silico* (Towle and Vederas, 2017). The antimicrobial activity itself could be further analysed determining its spectrum of activity through testing a range of bacterial species, both Gram-negative and Gram-positive. Typically circular bacteriocins have a broad-spectrum range against Gram-positive bacteria however the use of additional agents such as EDTA does sensitise some Gram-negative bacteria (Vasilchenko and Valyshev, 2019). It would be interesting to discover if the putative bacteriocin BacP conforms to this. Classically circular bacteriocins are pore-formers they have been shown to work in a receptor-independent or dependent manner with some utilising both options. It has been previously observed and suggested that some circular bacteriocins have a dual mode of actions which is concentration dependent (Gabrielsen *et al.*, 2014a). It would be interesting to determine if the putative bacteriocin, BacP, binds a specific receptor or works in an independent manner such as enterocin NKR-5-3B which is attracted to the negative charge of the cell membrane (Himeno *et al.*, 2015). One way that has been previously used to determine if there is a putative receptor is to isolate bacteriocin resistant colonies and use *in silico* techniques to identify receptor candidates. This method was used to identify the maltose ABC transporter as the receptor for the circular bacteriocin garvicin ML (Gabrielsen *et al.*, 2012).

The successfully demonstration of activity from the expression of the putative BacP locus suggests that the *E. coli* host does not have the necessary machinery to process the mature active peptide. The essentiality of all the expressed genes could be determined through expressing different combinations of the genes and observing if antimicrobial activity is still present. This would demonstrate if BacP conforms to the

typical circular group i bacteriocin biosynthetic cluster, providing further evidence for the requirement of a minimum of five to seven genes (Perez *et al.*, 2018). The presence of a predicted terminator after the BacP structural gene has previously been observed in the circularin A biosynthesis pathway and the layout of the biosynthesis pathway is similar to known circular bacteriocins pumarlin, uberolysin, enterosin NKR-5-3B and amylocyclicin (Perez *et al.*, 2018). Although Pumarlin is another *B. pumilus* group i circular bacteriocin it does not have an immunity gene within its biosynthesis pathway, differing to BacP which has a predicted peptide fitting the characteristics of a self-immunity gene. Expressing this gene in *C. difficile* could elucidate its role by potentially confering immunity to any antimicrobial activity. Alternatively the expression of the entire locus, minus the putative self-immunity gene, has previously been used for lactococcin B to demonstrate the protective role of its immunity peptide (van Belkum *et al.*, 1992).

Take home points

- Antimicrobial activity is due to secreted protein that could be a bacteriocin
- Preliminary results suggest the BacP locus contains a putative bacteriocin active against *C. difficile* and *in silico* analysis highlights BacP as the most likely candidate

Chapter 6: General Discussion

6.1 Addressing the Objectives of this study

The objectives of this study were:

- To screen environmental bacteria for antimicrobial activity against *C. difficile* with the intention of isolating novel bacteriocins.
- To determine if cow manure was a reservoir for putative novel bacteriocins against *C. difficile*
- To develop a screening protocol to identify bacterial candidate that may harbour novel bacteriocins
- Characterise any organisms demonstrating putative antimicrobial activity against *C. difficile* with the aim to investigate the origin of any antimicrobial activity to determine its origin focusing on any putative novel bacteriocin candidates.

During this study over 200 bacterial colonies isolated from cow manure were screened resulting in the identification of 11 bacterial strains, a high number of strains for this scale of project. One previous study only identified 34 strains out of 1000, screened and from that 34 only one isolate was investigated and found to carry a novel bacteriocin (Aunpad and Na-Bangchang, 2007). Another study identified eight strains from 155 LAB strains screened (Ndlovu *et al.*, 2015) whereas other studies have reported higher results such as 40 positive isolates out of 125 screened LAB strains

(Bali *et al.*, 2011). In this study two putative bacteriocins were determined amongst the genomes of the five strains investigated. This suggests that a robust screening protocol was developed which could be applied to the discovery of novel bacteriocins against a range of specified organisms including antimicrobial resistant bacteria. The protocol developed also has adaptability to look for specific types of bacteriocins. For example to select for bacteriocins with broad spectrum or narrow spectrum activity could be carried out early on in the process through tailoring the initial screening by including multiple organisms both resistant and sensitive (Bali *et al.*, 2011). In hindsight a superior screening protocol for novel bacteriocins against CDI would be to screen for narrow spectrum activity identifying organisms with activity against *C. difficile* which lack activity against typical organisms in the gastrointestinal microflora.

The isolation of organisms active against *C. difficile*, in this study, indicates that cow manure could be a reservoir of putative bacteriocin harbouring organisms. Whether the majority of those bacteriocins are known or novel needs to be further determined through additional screening. As it has been predicted that up to 99% of bacteria can produce at least one bacteriocin, we can hypothesise that there are further novel bacteriocins to be identified (Riley and Wertz, 2002; Yang *et al.*, 2014). At least two putative bacteriocin candidates have been identified in this study, providing evidence for the presence of novel bacteriocins within this reservoir of cow manure.

Five bacterial isolates from cow manure, and one from human sewage, underwent further characterisation in this study to determine if their observed antimicrobial activity against *C. difficile* could be due to the production of a bacteriocin. Three bacterial isolates underwent additional characterisation due to *in silico* evidence of putative novel bacteriocin candidates within their genome sequence. The

approaches utilised were both protein precipitation techniques and expression host to best characterise the antimicrobial activity observed. Out of the characterisation of these three isolates, two resulted in the evidential attribution of activity to a protein encoding gene or cluster of genes. The characterisation of *L. lactis* 10 resulted in the demonstration of the known bacteriocin, nisin, being solely responsible for the observed antimicrobial activity. The characterisation of *B. pumilus* 1_16 resulted in preliminary evidence for the antimicrobial activity to be attributed to a cluster of genes, one of which was a predicted bacteriocin encoding gene. This demonstrates that the methodology of further characterisation utilised in this study was successful in identifying responsible candidates of antimicrobial activity. However, there were two isolates where no responsible candidate was determined for any antimicrobial activity which conveys the limitation of the methodologies utilised.

6.2 Use of *In Silico* Mining in the Identification of Novel Bacteriocins

In silico mining for the screening of bacteriocins has become a common place approach to identify novel bacteriocin. This method has been utilised due to the advancement of genome analysis tools, bacteriocin databases, and the plethora of available genomes. While this tool is extremely useful in high-throughput screening systems, it does have drawbacks. For example, it does not allow for the search of bacteriocins with specific activity such as against certain organisms or with a desired spectrum of activity. This tool, however, can be utilised after physical screening methods, as demonstrated throughout this study. The use of *in silico* mining in this study conveys the effectiveness of the online tools in narrowing the screening for novel bacteriocin as it highlighted the one organism solely carrying a potential putative bacteriocin. This allowed the majority of research power to be focused on the isolated *B. pumilus* 1_16 strain to increase the probability of determining the responsible

candidate for the observed activity. Narrowing the screening strains early on is extremely important as it allows for more in depth analysis of more likely candidates. A limitation of *in silico* mining was, however, noted in this study as no putative candidates for activity were identified in the two strains. This could either be due to the lack of putative bacteriocins present, the poor quality of genome sequence, or a limitation of the *in silico* tools utilised. For both strains, *B. thermoamylovorans* and *Oceanobacillus* spp., there are, to date, no current characterised bacteriocins produced by these organisms. Therefore, it is not surprising that the *in silico* tools did not find putative bacteriocin candidates as it might be too novel compared to their database of known bacteriocins. This is where protein precipitation techniques are best utilised to discover completely novel bacteriocins or peptides responsible for antimicrobial activity when *in silico* tools have fail.

6.3 Comparing the Different Methods of Isolating Novel Bacteriocins

Bacteriocin purification using protein precipitation techniques is a common first step in the purification of bacteriocins. This step is a limitation in the identification of novel bacteriocins as it is time consuming and initial precipitation usually results in a partially purified crude bacteriocin which requires further purification. Downstream purification can result in an overall expensive method of purification that is not suitable for upscaling which is required for potential commercialisation. On the other hand, pure bacteriocin preparations allows for further characterisation of the peptide leading to the elucidation of novel drug targets. Although there are many bacteriocins successfully purified using protein precipitation (Gálvez *et al.*, 1989; Ross *et al.*, 1993; Cintas *et al.*, 1998; McAuliffe *et al.*, 1998; Kemperman *et al.*, 2003; Kawai *et al.*, 2004; Todorov *et al.*, 2004; Wirawan *et al.*, 2007; Martin-Visscher *et al.*, 2008; Sawa *et al.*, 2009), the trial and failure of numerous methods through this study shows the difficult

nature of this approach. There are multiple variables that can affect bacteriocin expression such as growth phase, growth medium, and temperature (Saavedra and Sesma, 2011). After an initial precipitation method is deemed successful the optimisation of downstream protocol such as cation exchange needs to then be carried out (Saavedra and Sesma, 2011). As the protein precipitation attempts made in this study proved unreliable or unsuccessful alternative methods were undertaken to determine if an alternative method was better suited for novel bacteriocins. Another method is utilising host organisms for expression, which can result in more cost-effective yield of peptide as hosts such as *E. coli* rapidly grow on inexpensive media (Mesa-Pereira *et al.*, 2018). The most common host organism is *E. coli* in part due to the available variety of genetic tools and expression systems (Mergulhão and Monteiro, 2004; Rosano and Ceccarelli, 2014; Jia and Jeon, 2016; Mesa-Pereira *et al.*, 2018), with multiple successful examples of bacteriocin expression (Kawai *et al.*, 2003; Caetano *et al.*, 2011b, 2011a; Shi *et al.*, 2011; Chen *et al.*, 2012; Himes *et al.*, 2016; Mesa-Pereira *et al.*, 2017). However utilising *E. coli* as an overexpression host requires consideration, time consuming trial and error experiments, which can be labour intensive. The genes required for successful expression can be different for each bacteriocin, with some only requiring the structural gene (Kawai *et al.*, 2003; Chen *et al.*, 2012) and other requiring transporter genes (Mesa-Pereira *et al.*, 2017), and others requiring post-translation modification genes (Caetano *et al.*, 2011b, 2011a; Shi *et al.*, 2011; Himes *et al.*, 2016), which might not be readily identifiable for novel bacteriocins. Other things to consider are the potential toxic effects towards the host, the presence and choice of affinity tags and fusion protein partners, and the choice of vector. Trialling different combinations to generate successful expression is a time-consuming process. Once the vector is designed then there is a choice of host strains, which is another significant decision as one study noted successfully expression when

they switched from *E. coli* BL21(DE3) to *E. coli* C43(DE3) (Wang *et al.*, 2016). Once a production strain is selected there is then the optimisation of growth conditions including temperature and media, as well as induction conditions such as the induction concentration and length. Overall this method needs optimisation for each individual bacteriocin due to magnitude of variables that can effect bacteriocin production (Mesa-Pereira *et al.*, 2018). Throughout this study the limitations of *E. coli* host expression approach has been evident, as multiple bacteriocin candidates were trialled in different *E. coli* strains however the results were only partially reliable for what was a time-consuming approach. The cause of the lack of successfully expression is also hard to identify as there are so many variables that can affect expression. Thus, the results of this study demonstrate the difficult nature of this approach. However, this study also conveys the simplicity of the approach and the powerful results when it is successful. The preliminary antimicrobial activity observed from the expression of the BacP locus in an *E. coli* host demonstrates the immediate nomination of a small number of genes as putative bacteriocin encoding candidates, linking the *in silico* data with *in vitro* results.

6.4 Future work

The priority of any future work would be to further investigate the BacP locus to demonstrate that BacP is a novel bacteriocin active against *C. difficile*. This could be carried out through running any cell lysate or concentrated supernatant of the *E. coli* overexpression host on an SDS-PAGE gel and overlaying it with *C. difficile*. Any peptide that shows activity could be sent for LC-MS/MS to elucidate the sequence and link it to one of the expressed genes. Another method would be to express the locus with different combinations of the genes which could also illuminate the roles of the other genes at the same time.

Further characterisation of the antimicrobial activity should also be prioritised as the reliable expression of the putative bacteriocin provides sufficient peptide for extensive characterisation. Characterisation such as activity spectrum, peptidase sensitivity, pH and heat stability, would determine if this peptide would be suitable for any future applications and the scope of those applications. It would also be interesting to determine if this activity is a result of typical circular bacteriocin pore-forming mode of action (Gabrielsen *et al.*, 2014a) as this would further the circular bacteriocin field of research. This could be determined through using microscopy to observe whether the peptide enters the cell or builds up on the surface. Model membranes can be utilised to determine if there is integration into the membrane or attachment, and also depolarisation of lipid bilayers could also be monitored (Brogden, 2005). Since many bacteriocin producing strains have applications as probiotics both the peptide and the producing strain could undergo efficacy and safety testing through the observation of any cytotoxic effects on mammalian cell lines and through *in vivo* studies administering the peptide to hamsters infected with *C. difficile*. This could result in evidence supporting the continued research into this strain and the putative bacteriocin.

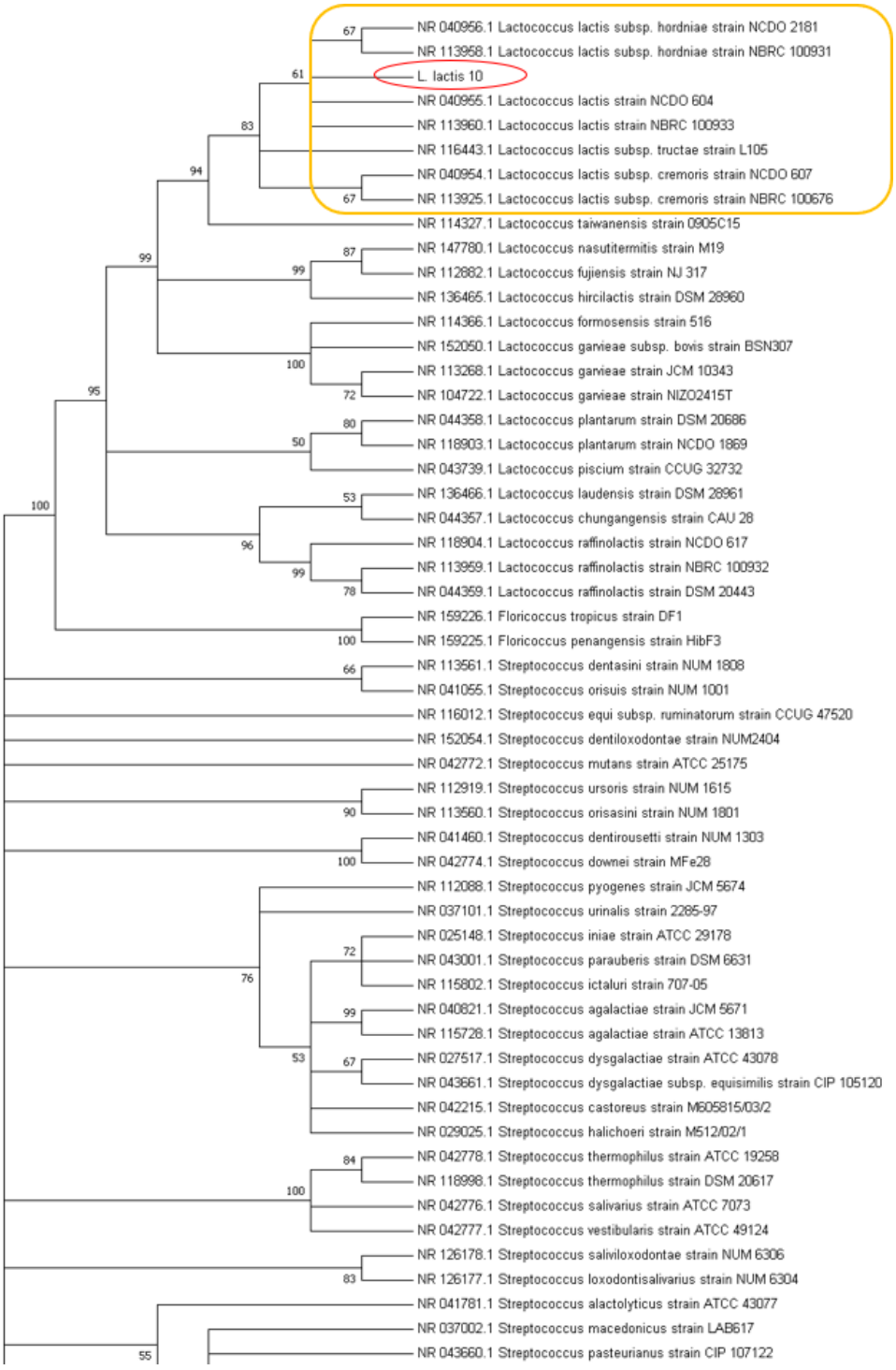
Investigating the putative bacteriocin would result in providing more insight in to the bacteriocin field as well as benefitting *C. difficile*. Although the antimicrobial peptide produced by *B. pumilus* 1_16 could result in a future novel treatment for CDI directly, future work could lead to furthering the field of work indirectly. Determining the peptide target or receptor, if any, could result in a novel putative target for antimicrobials against *C. difficile*. Through the utilisation of resistant mutants and genome sequencing, *C. difficile* phage and bacteriocins targeting *C. difficile* have been shown to utilise the SlpA peptide within the S-layer (Kirk *et al.*, 2017). This could also be applied to the putative bacteriocin identified in this body of work, and if a different target is determined then it provides a possible alternative method of killing *C. difficile*.

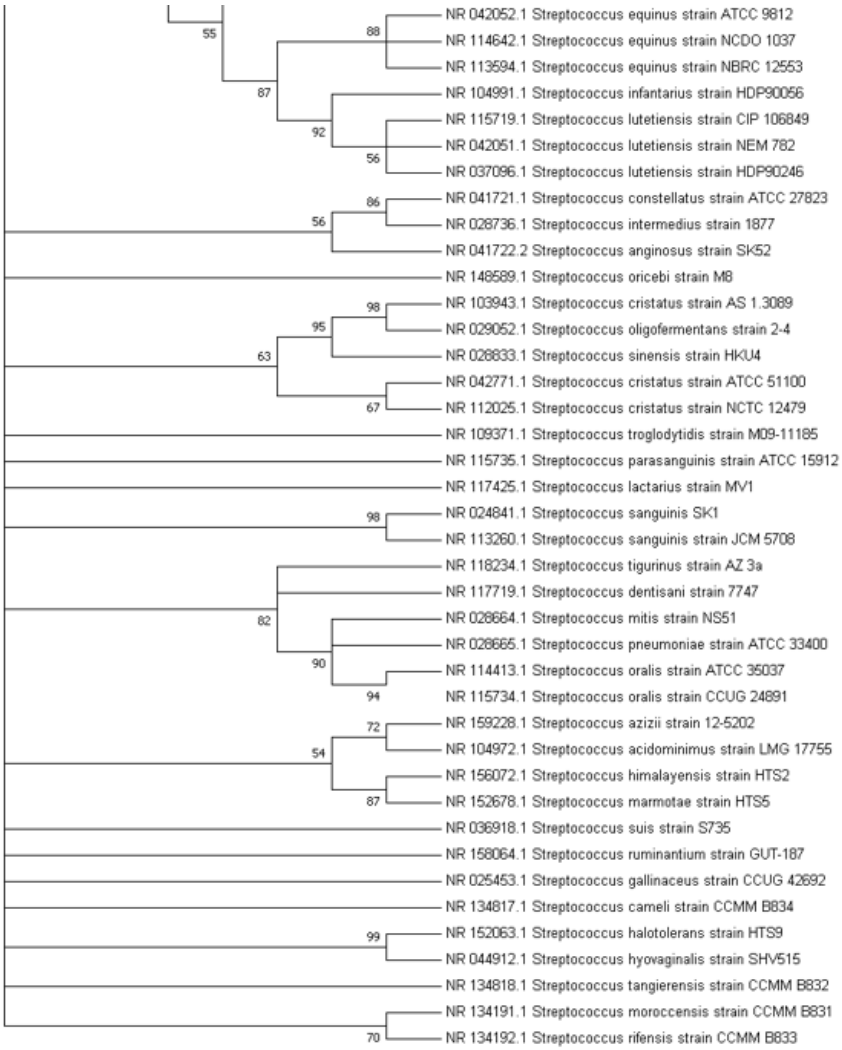
If multiple targets are available and utilised together it limits the likelihood of generating resistant mutants. This would also be the case if multiple alternative treatments were available that utilised different modes of actions.

7 Appendix

Appendix 7.1

The maximum likelihood tree was generated by MEGA7 (Kumar *et al.*, 2016), and shows the relationship between *L. lactis* 10 16s rRNA and the top 94 matches against the NCBI 16s database. This phylogenetic tree used *Streptococcus* sp. as an outgroup and was bootstrapped 500 times. The yellow square shows the *L. lactis* cluster with *L. lactis* 10 circled in red.

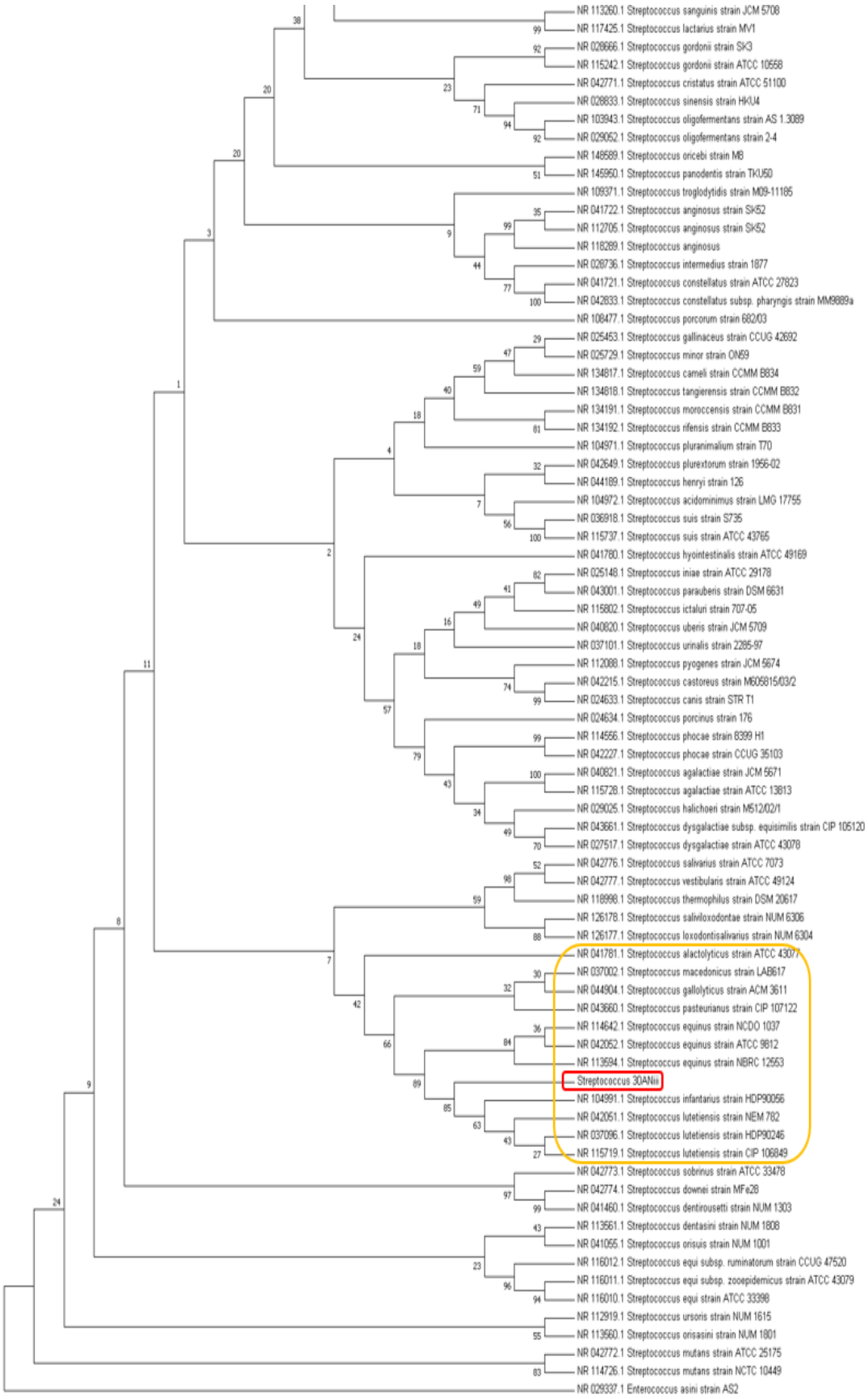




Appendix 7.2

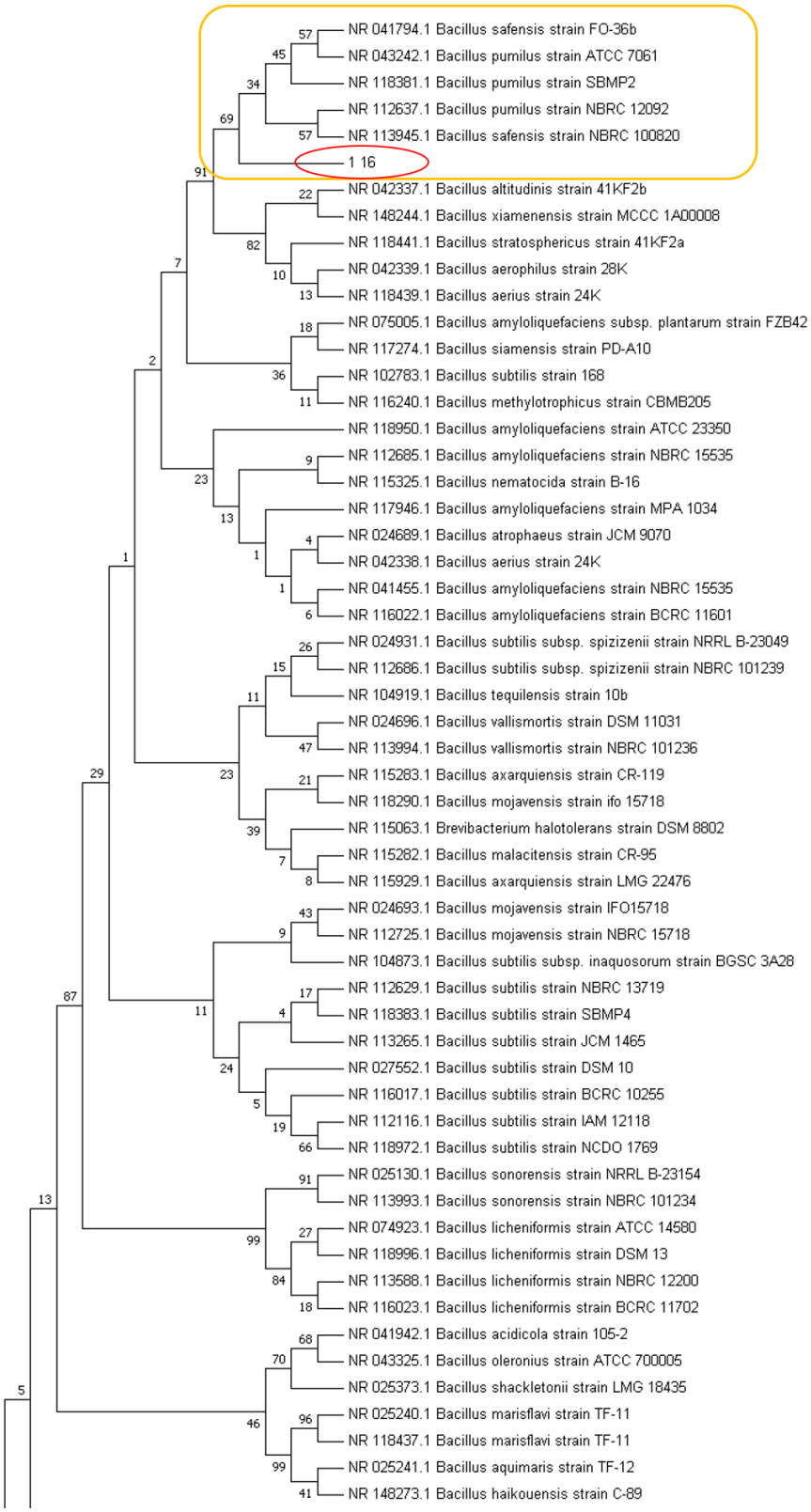
Maximum likelihood tree was generated by MEGA7 (Kumar *et al.*, 2016) showing the 16s rRNA sequence from *Streptococcus* 30ANiii and the 89 closest matches using the 16s rRNA nucleotide NCBI database. Phylogenetic analysis was bootstrapped 500 times and is rooted using an *Enterococcus* sp. outgroup. The yellow square indicates the *S. bovis/S.equinus* complex with *Streptococcus* 30ANiii falling within this group highlighted in the red box.

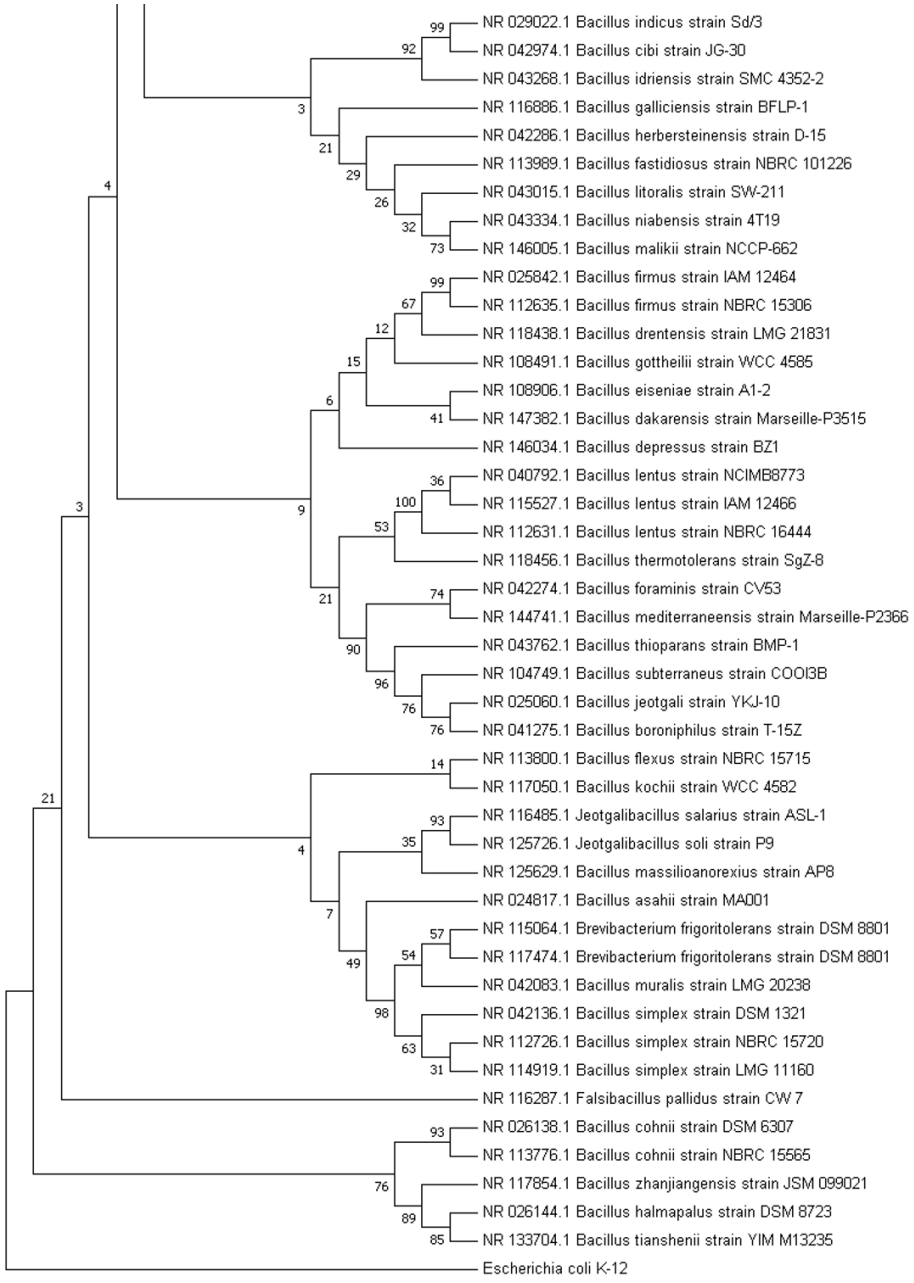



NR 113260.1 Streptococcus sanguinis strain JCM 5708
NR 117425.1 Streptococcus lactarius strain MV1
NR 028666.1 Streptococcus gordonii strain SK3
NR 115242.1 Streptococcus gordonii strain ATCC 10558
NR 042771.1 Streptococcus cristatus strain ATCC 51100
NR 028833.1 Streptococcus sinensis strain HKU4
NR 103943.1 Streptococcus oligofermentans strain AS 1.3089
NR 029052.1 Streptococcus oligofermentans strain 2-4
NR 148589.1 Streptococcus oricebi strain M8
NR 145960.1 Streptococcus panodentis strain TKU50
NR 109371.1 Streptococcus troglodytidis strain M09-11185
NR 041722.1 Streptococcus anginosus strain SK52
NR 112705.1 Streptococcus anginosus strain SK52
NR 118289.1 Streptococcus anginosus
NR 028736.1 Streptococcus intermedius strain 1877
NR 041721.1 Streptococcus constellatus strain ATCC 27823
NR 042833.1 Streptococcus constellatus subsp. pharyngis strain MM9889a
NR 108477.1 Streptococcus porcorum strain 682/03
NR 025453.1 Streptococcus gallinaceus strain CCUG 42692
NR 025729.1 Streptococcus minor strain ON59
NR 134817.1 Streptococcus cameli strain CCMM B834
NR 134818.1 Streptococcus tangierensis strain CCMM B832
NR 134191.1 Streptococcus moroccensis strain CCMM B831
NR 134192.1 Streptococcus rifensis strain CCMM B833
NR 104971.1 Streptococcus pluranimalium strain T70
NR 042649.1 Streptococcus plurextorum strain 1956-02
NR 044189.1 Streptococcus henryi strain 126
NR 104972.1 Streptococcus acidominimus strain LMG 17755
NR 036918.1 Streptococcus suis strain S735
NR 115737.1 Streptococcus suis strain ATCC 43765
NR 041780.1 Streptococcus hyointestinalis strain ATCC 49169
NR 025148.1 Streptococcus iniae strain ATCC 29178
NR 043001.1 Streptococcus parauberis strain DSM 6631
NR 115802.1 Streptococcus ictaluri strain 707-05
NR 040820.1 Streptococcus uberis strain JCM 5709
NR 037101.1 Streptococcus urinalis strain 2285-97
NR 112088.1 Streptococcus pyogenes strain JCM 5674
NR 042215.1 Streptococcus castoreus strain M605815/03/2
NR 024633.1 Streptococcus canis strain STR T1
NR 024634.1 Streptococcus porcinus strain 176
NR 114566.1 Streptococcus phocae strain 8399 H1
NR 042227.1 Streptococcus phocae strain CCUG 35103
NR 040821.1 Streptococcus agalactiae strain JCM 5671
NR 115728.1 Streptococcus agalactiae strain ATCC 13813
NR 029025.1 Streptococcus halichoeri strain M512/02/1
NR 043661.1 Streptococcus dysgalactiae subsp. equisimilis strain CIP 105120
NR 027517.1 Streptococcus dysgalactiae strain ATCC 43078
NR 042776.1 Streptococcus salivarius strain ATCC 7073
NR 042777.1 Streptococcus vestibularis strain ATCC 49124
NR 118998.1 Streptococcus thermophilus strain DSM 20617
NR 126178.1 Streptococcus salivloxodontae strain NUM 6306
NR 126177.1 Streptococcus loxodontisalivarius strain NUM 6304
NR 041781.1 Streptococcus alactolyticus strain ATCC 43077
NR 037002.1 Streptococcus macedonicus strain LAB617
NR 044904.1 Streptococcus gallolyticus strain ACM 3611
NR 043660.1 Streptococcus pasteurianus strain CIP 107122
NR 114642.1 Streptococcus equinus strain NCDO 1037
NR 042052.1 Streptococcus equinus strain ATCC 9812
NR 113694.1 Streptococcus equinus strain NBRC 12553
Streptococcus 30ANii
NR 104991.1 Streptococcus infantarius strain HDP90056
NR 042051.1 Streptococcus lutetiensis strain NEM 782
NR 037096.1 Streptococcus lutetiensis strain HDP90246
NR 115719.1 Streptococcus lutetiensis strain CIP 106849
NR 042773.1 Streptococcus sobrinus strain ATCC 33478
NR 042774.1 Streptococcus downei strain MFe28
NR 041460.1 Streptococcus dentirousetti strain NUM 1303
NR 113561.1 Streptococcus dentasini strain NUM 1808
NR 041055.1 Streptococcus orisuis strain NUM 1001
NR 116012.1 Streptococcus equi subsp. ruminatorum strain CCUG 47520
NR 116011.1 Streptococcus equi subsp. zooepidemicus strain ATCC 43079
NR 116010.1 Streptococcus equi strain ATCC 33398
NR 112919.1 Streptococcus ursoris strain NUM 1615
NR 113560.1 Streptococcus orisasini strain NUM 1801
NR 042772.1 Streptococcus mutans strain ATCC 25175
NR 114726.1 Streptococcus mutans strain NCTC 10449
NR 029337.1 Enterococcus asini strain AS2

Appendix 7.3

A phylogenetic tree of the relationship of *B. pumilus* 1_16 compared to close relatives identified through NCBI 16s rRNA database. The phylogenetic tree was generated by MEGA7 (Kumar *et al.*, 2016) and bootstrapped 500 times. The *B. pumilus* and *B. safenesis* branch is indicated through the yellow rectangle and strain 1_16 by the red circle. The out group used was *Escherichia coli* K-12.




NR 029022.1 Bacillus indicus strain Sd/3
NR 042974.1 Bacillus cibi strain JG-30
NR 043268.1 Bacillus idriensis strain SMC 4352-2
NR 116886.1 Bacillus galliciensis strain BFLP-1
NR 042286.1 Bacillus herbersteinensis strain D-15
NR 113989.1 Bacillus fastidiosus strain NBRC 101226
NR 043015.1 Bacillus litoralis strain SW-211
NR 043334.1 Bacillus niabensis strain 4T19
NR 146005.1 Bacillus malikii strain NCCP-662
NR 025842.1 Bacillus firmus strain IAM 12464
NR 112635.1 Bacillus firmus strain NBRC 15306
NR 118438.1 Bacillus drentensis strain LMG 21831
NR 108491.1 Bacillus gottheilii strain WCC 4585
NR 108906.1 Bacillus eiseniae strain A1-2
NR 147382.1 Bacillus dakarensis strain Marseille-P3515
NR 146034.1 Bacillus depressus strain BZ1
NR 040792.1 Bacillus lentus strain NCIMB8773
NR 115527.1 Bacillus lentus strain IAM 12466
NR 112631.1 Bacillus lentus strain NBRC 16444
NR 118456.1 Bacillus thermotolerans strain SgZ-8
NR 042274.1 Bacillus foraminis strain CV53
NR 144741.1 Bacillus mediterraneensis strain Marseille-P2366
NR 043762.1 Bacillus thioparans strain BMP-1
NR 104749.1 Bacillus subterraneus strain COOI3B
NR 025060.1 Bacillus jeotgali strain YKJ-10
NR 041275.1 Bacillus boroniphilus strain T-15Z
NR 113800.1 Bacillus flexus strain NBRC 15715
NR 117050.1 Bacillus kochii strain WCC 4582
NR 116485.1 Jeotgalibacillus salarius strain ASL-1
NR 125726.1 Jeotgalibacillus soli strain P9
NR 125629.1 Bacillus massilioanorexius strain AP8
NR 024817.1 Bacillus asahii strain MA001
NR 115064.1 Brevibacterium frigoritolerans strain DSM 8801
NR 117474.1 Brevibacterium frigoritolerans strain DSM 8801
NR 042083.1 Bacillus muralis strain LMG 20238
NR 042136.1 Bacillus simplex strain DSM 1321
NR 112726.1 Bacillus simplex strain NBRC 15720
NR 114919.1 Bacillus simplex strain LMG 11160
NR 116287.1 Falsibacillus pallidus strain CW 7
NR 026138.1 Bacillus cohnii strain DSM 6307
NR 113776.1 Bacillus cohnii strain NBRC 15565
NR 117854.1 Bacillus zhanjiangensis strain JSM 099021
NR 026144.1 Bacillus halmapalus strain DSM 8723
NR 133704.1 Bacillus tianshenii strain YIM M13235
Escherichia coli K-12

Appendix 7.4 – PIPS reflective statement

Note to examiners:

This statement is included as an appendix to the thesis in order that the thesis accurately captures the PhD training experienced by the candidate as a BBSRC Doctoral Training Partnership student.

The Professional Internship for PhD Students is a compulsory 3-month placement which must be undertaken by DTP students. It is usually centred on a specific project and must not be related to the PhD project. This reflective statement is designed to capture the skills development which has taken place during the student's placement and the impact on their career plans it has had.

PIPS Reflective Statement

During my third year of my PhD between the months of April and July, I understood a PIP placement at metaLinear Ltd in BioCity Alderly Edge. During this placement I worked as scientist in the laboratory setting assisting with current research projects developing my laboratory knowledge and skills. I also helped contribute to the planning of future projects through generating and analysing data.

This placement allowed me to gain insight into industrial biological businesses especially small start-up companies. It taught me the advantageous and disadvantageous of working for a small company providing crucial experience for my future career. Through this placement I have furthered my scientific skills which was useful of my PhD and I have also developed contacts within industry which could help me in my career.

8 Bibliography

Abou Chakra, C. N., Pepin, J., Sirard, S. and Valiquette, L. (2014) 'Risk Factors for Recurrence, Complications and Mortality in *Clostridium difficile* Infection: A Systematic Review', *PLOS ONE*. Public Library of Science, 9(6), p. e98400. Available at: https://doi.org/10.1371/journal.pone.0098400.

Acedo, J. Z., van Belkum, M. J., Lohans, C. T., McKay, R. T., Miskolzie, M. and Vederas, J. C. (2015) 'Solution structure of acidocin B, a circular bacteriocin produced by *Lactobacillus acidophilus* M46.', *Applied and environmental microbiology*. United States, 81(8), pp. 2910–2918. doi: 10.1128/AEM.04265-14.

Actor, J. K. (2012) '12 - Clinical Bacteriology', in Actor, J. K. B. T.-E. I. R. I. and M. (Second E. (ed.). Philadelphia: W.B. Saunders, pp. 105–120. doi: https://doi.org/10.1016/B978-0-323-07447-6.00012-0.

Acuña, L., Corbalan, N. S., Fernandez-No, I. C., Morero, R. D., Barros-Velazquez, J. and Bellomio, A. (2015) 'Inhibitory Effect of the Hybrid Bacteriocin Ent35-MccV on the Growth of *Escherichia coli* and *Listeria monocytogenes* in Model and Food Systems', *Food and Bioprocess Technology*, 8(5), pp. 1063–1075. doi: 10.1007/s11947-015-1469-0.

Ahmed, S. T., Islam, M. M., Mun, H.-S., Sim, H.-J., Kim, Y.-J. and Yang, C.-J. (2014) 'Effects of *Bacillus amyloliquefaciens* as a probiotic strain on growth performance, cecal microflora, and fecal noxious gas emissions of broiler chickens', *Poultry Science*, 93(8), pp. 1963–1971. doi: 10.3382/ps.2013-03718.

Alifax, P. R. and Chevalier, R. (1962) 'Etude de la nisinase produite par *Streptococcus thermophilus*', *Journal of Dairy Research*. 2009/06/01. Cambridge University Press, 29(3), pp. 233–240. doi: DOI: 10.1017/S0022029900011043.

De Almeida Vaucher, R., De Campos Velho Gewehr, C., Folmer Correa, A. P., Sant'Anna, V., Ferreira, J. and Brandelli, A. (2011) 'Evaluation of the immunogenicity and in vivo toxicity of the antimicrobial peptide P34', *International Journal of Pharmaceutics*, 421(1), pp. 94–98. doi: 10.1016/j.ijpharm.2011.09.020.

Alonso, C. D. and Mahoney, M. V (2018) 'Bezlotoxumab for the prevention of *Clostridium difficile* infection: a review of current evidence and safety profile', *Infection and drug resistance*. Dove Medical Press, 12, pp. 1–9. doi: 10.2147/IDR.S159957.

Altschul, S. F., Gish, W., Miller, W., Myers, E. W. and Lipman, D. J. (1990) 'Basic local alignment search tool.', *Journal of molecular biology*. England, 215(3), pp. 403–410. doi: 10.1016/S0022-2836(05)80360-2.

Alvarez-Sieiro, P., Montalbán-López, M., Mu, D. and Kuipers, O. P. (2016) 'Bacteriocins Of Lactic Acid Bacteria: Extending The Family', *Applied Microbiology and Biotechnology*, 100(7 LB-Alvarez-Sieiro2016), pp. 2939–2951. doi: 10.1007/s00253-016-7343-9.

Apweiler, R. *et al.* (2000) 'InterPro--an integrated documentation resource for protein families, domains and functional sites.', *Bioinformatics (Oxford, England)*. England, 16(12), pp. 1145–1150. doi: 10.1093/bioinformatics/16.12.1145.

Arndt, D., Marcu, A., Liang, Y. and Wishart, D. S. (2017) 'PHAST, PHASTER and PHASTEST: Tools for finding prophage in bacterial genomes.', *Briefings in bioinformatics*. England. doi: 10.1093/bib/bbx121.

Arnison, P. G. *et al.* (2013) 'Ribosomally Synthesized and Post-Translationally Modified Peptide Natural Products: Overview and Recommendations for a Universal Nomenclature', *Nat Prod Rep*. 2012/11/21. Prairie Plant Systems Inc, Botanical Alternatives Inc, Suite 176, 8B-3110 8th Street E, Saskatoon, SK, S7H 0W2, Canada., 30(1), pp. 108–160. doi: 10.1039/c2np20085f.

Aslam, S., Hamill, R. J. and Musher, D. M. (2005) 'Treatment of *Clostridium difficile*-Associated Disease: Old Therapies and New Strategies', *The Lancet Infectious Diseases*, 5, pp. 549–557.

Atrih, A., Rekhif, N., Moir, A. J., Lebrihi, A. and Lefebvre, G. (2001) 'Mode of action, purification and amino acid sequence of plantaricin C19, an anti-Listeria bacteriocin produced by *Lactobacillus plantarum* C19.', *International journal of food microbiology*. Netherlands, 68(1–2), pp. 93–104. doi: 10.1016/s0168-1605(01)00482-2.

Aunpad, R. and Na-Bangchang (2007) 'Pumilicin 4, A Novel Bacteriocin with Anti-MRSA and Anti-VRE Activity Produced by Newly Isolated Bacteria *Bacillus pumilus* Strain WAPB4', *Current Microbiology*, 55, pp. 308–313.

Aunpad, R. and Panbangred, W. (2012) 'Evidence for Two Putative Holin-Like Peptides Encoding Genes of *Bacillus pumilus* Strain WAPB4', *Current Microbiology*, 64, pp. 343–348.

Avbersek, J., Janezic, S., Pate, M., Rupnik, M., Zidaric, V., Logar, K., Vengust, M., Zemljic, M., Pirs, T. and Ocepek, M. (2009) 'Diversity of *Clostridium difficile* in pigs and other animals in Slovenia.', *Anaerobe*. England, 15(6), pp. 252–255. doi: 10.1016/j.anaerobe.2009.07.004.

Aziz, R. K. *et al.* (2008) 'The RAST Server: rapid annotations using subsystems technology', *BMC Genomics*. 2008/02/12. Fellowship for Interpretation of Genomes, Burr Ridge, IL 60527, USA. ramy.aziz@gmail.com, 9, p. 75. doi: 10.1186/1471-2164-9-75.

Baines, S. D., O'Connor, R., Freeman, J., Fawley, W. N., Harmanus, C., Mastrantonio, P., Kuijper, E. J. and Wilcox, M. H. (2008) 'Emergence of Reduced Susceptibility to Metronidazole in *Clostridium difficile*', *Journal of Antimicrobial Chemotherapy*, 62(5), pp. 1046–1052.

Balciunas, E. M., Castillo Martinez, F. A., Todorov, S. D., Franco, B. D. G. M., Converti, A. and Oliveira, R. P. S. (2013) 'Novel biotechnological applications of bacteriocins: A review', *Food Control*, 32(1), pp. 134–142. doi: http://dx.doi.org/10.1016/j.foodcont.2012.11.025.

Bali, V., Panesar, P. S. and Bera, M. B. (2011) 'Isolation, Screening and Evaluation of Antimicrobial Activity of Potential Bacteriocin Producing Lactic Acid Bacteria Isolate', *Microbiology Journal*. Science Alert, 1(3), pp. 113–119. doi: 10.3923/mj.2011.113.119.

Balla, E., Dicks, L. M., Du Toit, M., Van Der Merwe, M. J. and Holzapfel, W. H. (2000) 'Characterization and cloning of the genes encoding enterocin 1071A and enterocin 1071B, two antimicrobial peptides produced by *Enterococcus faecalis* BFE 1071.',

Applied and environmental microbiology. United States, 66(4), pp. 1298–1304. doi: 10.1128/aem.66.4.1298-1304.2000.

Bandelj, P., Blagus, R., Briski, F., Frlic, O., Vergles Rataj, A., Rupnik, M., Ocepek, M. and Vengust, M. (2016) 'Identification of risk factors influencing *Clostridium difficile* prevalence in middle-size dairy farms', *Veterinary research*. BioMed Central, 47, p. 41. doi: 10.1186/s13567-016-0326-0.

Bandelj, P., Harmanus, C., Blagus, R., Cotman, M., Kuijper, E. J., Ocepek, M. and Vengust, M. (2018) 'Quantification of *Clostridioides (Clostridium) difficile* in feces of calves of different age and determination of predominant Clostridioides difficile ribotype 033 relatedness and transmission between family dairy farms using multilocus variable-number ta', *BMC Veterinary Research*, 14(1), p. 298. doi: 10.1186/s12917-018-1616-8.

Banerjee, P., Merkel, G. J. and Bhunia, A. K. (2009) 'Lactobacillus delbrueckii ssp. bulgaricus B-30892 can Inhibit Cytotoxic Effects and Adhesion of Pathogenic *Clostridium difficile* to Caco-2 Cells', *Gut Pathogens*, 1(1 LB-Banerjee2009), pp. 1–11. doi: 10.1186/1757-4749-1-8.

Barbut, F. *et al.* (2005) 'Clinical features of *Clostridium difficile*-associated diarrhoea due to binary toxin (actin-specific ADP-ribosyltransferase)-producing strains.', *Journal of medical microbiology*. England, 54(Pt 2), pp. 181–185. doi: 10.1099/jmm.0.45804-0.

Barreteau, H., Bouhss, A., Gérard, F., Duché, D., Boussaid, B., Blanot, D., Lloubès, R., Mengin-Lecreulx, D. and Touzé, T. (2010) 'Deciphering the catalytic domain of colicin M, a peptidoglycan lipid II-degrading enzyme.', *The Journal of biological chemistry*, 285(16), pp. 12378–89. doi: 10.1074/jbc.M109.093583.

Bateman, A., Birney, E., Durbin, R., Eddy, S. R., Finn, R. D. and Sonnhammer, E. L. (1999) 'Pfam 3.1: 1313 multiple alignments and profile HMMs match the majority of proteins.', *Nucleic acids research*. England, 27(1), pp. 260–262. doi: 10.1093/nar/27.1.260.

Bauer, M. P., Notermans, D. W., van Benthem, B. H. G., Brazier, J. S., Wilcox, M. H., Rupnik, M., Monnet, D. L., van Dissel, J. T. and Kuijper, E. J. (2011) '*Clostridium difficile* Infection in Europe: A Hospital-Based Survey', *Lancet*, 377(9759), pp. 63–73.

Bauer, R., Chikindas, M. L. and Dicks, L. M. T. (2005) 'Purification, partial amino acid sequence and mode of action of pediocin PD-1, a bacteriocin produced by *Pediococcus damnosus* NCFB 1832.', *International journal of food microbiology*. Netherlands, 101(1), pp. 17–27. doi: 10.1016/j.ijfoodmicro.2004.10.040.

Begley, M., Cotter, P. D., Hill, C. and Ross, R. P. (2009) 'Identification of a Novel Two-Peptide Lantibiotic, Lichenicidin, following Rational Genome Mining for LanM Proteins', *Applied and Environmental Microbiology*, 75(17), pp. 5451–5460.

van Belkum, M. J., Kok, J. and Venema, G. (1992) 'Cloning, sequencing, and expression in *Escherichia coli* of lcnB, a third bacteriocin determinant from the lactococcal bacteriocin plasmid p9B4-6.', *Applied and Environmental Microbiology*, 58(2), pp. 572 LP – 577.

Di Bella, S., Ascenzi, P., Siarakas, S., Petrosillo, N. and di Masi, A. (2016) '*Clostridium difficile* toxins A and B: Insights into pathogenic properties and extraintestinal effects', *Toxins*. MDPI AG. doi: 10.3390/toxins8050134.

Bentur, H. N., Dalzell, A. M. and Riordan, F. A. I. (2007) 'Central venous catheter infection with *Bacillus pumilus* in an immunocompetent child: a case report', *Annals of clinical microbiology and antimicrobials*. BioMed Central, 6, p. 12. doi: 10.1186/1476-0711-6-12.

Beric, T., Stankovic, S., Draganic, V., Kojic, M., Lozo, J. and Fira, D. (2014) 'Novel antilisterial bacteriocin licheniocin 50.2 from *Bacillus licheniformis* VPS50.2 isolated from soil sample.', *Journal of applied microbiology*. England, 116(3), pp. 502–510. doi: 10.1111/jam.12393.

Berry, C. E., Davies, K. A., Owens, D. W. and Wilcox, M. H. (2017) 'Is there a relationship between the presence of the binary toxin genes in *Clostridium difficile* strains and the severity of *C. difficile* infection (CDI)?', *European Journal of Clinical Microbiology and Infectious Diseases*. Springer Verlag, 36(12), pp. 2405–2415. doi: 10.1007/s10096-017-3075-8.

Bézay, N., Ayad, A., Dubischar, K., Firbas, C., Hochreiter, R., Kiermayr, S., Kiss, I., Pinl, F., Jilma, B. and Westritschnig, K. (2016) 'Safety, immunogenicity and dose response of VLA84, a new vaccine candidate against *Clostridium difficile*, in healthy volunteers', *Vaccine*, 34(23), pp. 2585–2592. doi: https://doi.org/10.1016/j.vaccine.2016.03.098.

Boakes, S. and Dawson, M. J. (2014) 'Discovery and Development of NVB302, a Semisynthetic Antibiotic for Treatment of *Clostridium difficile* Infection', in *Natural Products: Discourse, Diversity, and Design*. Wiley Blackwell, pp. 455–468. doi: 10.1002/9781118794623.ch24.

Borrero, J., Brede, D. A., Skaugen, M., Diep, D. B., Herranz, C., Nes, I. F., Cintas, L. M. and Hernandez, P. E. (2011) 'Characterization of garvicin ML, a novel circular bacteriocin produced by *Lactococcus garvieae* DCC43, isolated from mallard ducks (Anas platyrhynchos).', *Applied and environmental microbiology*. United States, 77(1), pp. 369–373. doi: 10.1128/AEM.01173-10.

Braun, V., Patzer, S. I. and Hantke, K. (2002) 'Ton-dependent colicins and microcins: modular design and evolution.', *Biochimie*, 84(5–6), pp. 365–80. doi: 10.1016/s0300-9084(02)01427-x.

Brettin, T. *et al.* (2015) 'RASTtk: a modular and extensible implementation of the RAST algorithm for building custom annotation pipelines and annotating batches of genomes', *Scientific Reports*, 5.

Breukink, E., van Heusden, H. E., Vollmerhaus, P. J., Swiezewska, E., Brunner, L., Walker, S., Heck, A. J. R. and de Kruijff, B. (2003) 'Lipid II is an intrinsic component of the pore induced by nisin in bacterial membranes.', *The Journal of biological chemistry*. United States, 278(22), pp. 19898–19903. doi: 10.1074/jbc.M301463200.

Breukink, E. and de Kruijff, B. (1999) 'The Lantibiotic Nisin, a Special Case or Not?', *Biochimica et Biophysica Acta (BBA)- Biomembranes*, 1462(1–2), pp. 223–234.

Breukink, E., Wiedemann, I., van Kraaij, C., Kuipers, O. P., Sahl, H. G. and de Kruijff, B. (1999) 'Use of the cell wall precursor lipid II by a pore-forming peptide antibiotic.', *Science (New York, N.Y.)*. United States, 286(5448), pp. 2361–2364. doi: 10.1126/science.286.5448.2361.

Brogden, K. A. (2005) 'Antimicrobial peptides: Pore formers or metabolic inhibitors in bacteria?', *Nature Reviews Microbiology*, pp. 238–250. doi: 10.1038/nrmicro1098.

Brotz, H., Josten, M., Wiedemann, I., Schneider, U., Gotz, F., Bierbaum, G. and Sahl, H. G. (1998) 'Role of lipid-bound peptidoglycan precursors in the formation of pores by nisin, epidermin and other lantibiotics.', *Molecular microbiology*. England, 30(2), pp. 317–327.

Buchman, G. W., Banerjee, S. and Hansen, J. N. (1988) 'Structure, Expression, and Evolution of a Gene Encoding The Precursor of Nisin, a Small Protein Antibiotic', *The Journal of Biological Chemistry*. 1988/11/05. Department of Chemistry and Biochemistry, University of Maryland, College Park 20742., 263(31), pp. 16260–16266.

Burgos, M. J. G., Pulido, R. P., Aguayo, M. del C. L., Gálvez, A. and Lucas, R. (2014) 'The cyclic antibacterial peptide enterocin AS-48: Isolation, mode of action, and possible food applications', *International Journal of Molecular Sciences*. MDPI AG, pp. 22706–22727. doi: 10.3390/ijms151222706.

Burt, S. A., Siemeling, L., Kuijper, E. J. and Lipman, L. J. A. (2012) 'Vermin on pig farms are vectors for *Clostridium difficile* PCR ribotypes 078 and 045', *Veterinary Microbiology*, 160(1), pp. 256–258. doi: https://doi.org/10.1016/j.vetmic.2012.05.014.

Caetano, T., Krawczyk, J. M., Mösker, E., Süssmuth, R. D. and Mendo, S. (2011a) 'Heterologous expression, biosynthesis, and mutagenesis of type II lantibiotics from *Bacillus licheniformis* in *Escherichia coli*.', *Chemistry & biology*, 18(1), pp. 90–100. doi: 10.1016/j.chembiol.2010.11.010.

Caetano, T., Krawczyk, J. M., Mösker, E., Süssmuth, R. D. and Mendo, S. (2011b) 'Lichenicidin biosynthesis in *Escherichia coli*: licFGEHI immunity genes are not essential for lantibiotic production or self-protection.', *Applied and environmental microbiology*, 77(14), pp. 5023–6. doi: 10.1128/AEM.00270-11.

Cammarota, G., Masucci, L., Ianiro, G., Bibbò, S., Dinoi, G., Costamagna, G., Sanguinetti, M. and Gasbarrini, A. (2015) 'Randomised clinical trial: faecal microbiota transplantation by colonoscopy vs. vancomycin for the treatment of recurrent *Clostridium difficile* infection', *Alimentary Pharmacology & Therapeutics*. John Wiley & Sons, Ltd (10.1111), 41(9), pp. 835–843. doi: 10.1111/apt.13144.

Cañadas, I. C., Groothuis, D., Zygouropoulou, M., Rodrigues, R. and Minton, N. P. (2019) 'RiboCas: A Universal CRISPR-Based Editing Tool for Clostridium', *ACS Synthetic Biology*. American Chemical Society, 8(6), pp. 1379–1390. doi: 10.1021/acssynbio.9b00075.

Carter, G. P., Rood, J. I. and Lyras, D. (2010) 'The Role of Toxin A and Toxin B in *Clostridium difficile*-Associated Disease', *Gut Microbed*, 1(1), pp. 58–64.

Cavera, V. L., Arthur, T. D., Kashtanov, D. and Chikindas, M. L. (2015) 'Bacteriocins and their Position in the Next Wave of Conventional Antibiotics', *International Journal of Antimicrobial Agents*, 46, pp. 494–501.

Cebrián, R. *et al.* (2019) 'Preclinical studies of toxicity and safety of the AS-48 bacteriocin', *Journal of Advanced Research*. Elsevier B.V., 20, pp. 129–139. doi: 10.1016/j.jare.2019.06.003.

Cebrian, R., Maqueda, M., Neira, J. L., Valdivia, E., Martinez-Bueno, M. and Montalban-Lopez, M. (2010) 'Insights into the functionality of the putative residues involved in enterocin AS-48 maturation.', *Applied and environmental microbiology*.

United States, 76(21), pp. 7268–7276. doi: 10.1128/AEM.01154-10.

Chatterjee, S., Chatterjee, S., Lad, S. J., Phansalkar, M. S., Rupp, R. H., Ganguli, B. N., Fehlhaber, H.-W. and Kogler, H. (1992) 'Mersacidin, A New Antibiotic From Bacillus: Fermentation, Isolation, Purification and Chemical Characterization', *The Journal of Antibiotics*, 45(6), pp. 832–838.

Chavan, M., Rafi, H., Wertz, J., Goldstone, C. and Riley, M. A. (2005) 'Phage associated bacteriocins reveal a novel mechanism for bacteriocin diversification in Klebsiella.', *Journal of molecular evolution*, 60(4), pp. 546–56. doi: 10.1007/s00239-004-0263-9.

Chen, H. and Hoover, D. G. (2003) 'Bacteriocins and their Food Applications', *Comprehensive Reviews in Food Science and Food Safety*. John Wiley & Sons, Ltd (10.1111), 2(3), pp. 82–100. doi: 10.1111/j.1541-4337.2003.tb00016.x.

Chen, H., Tian, F., Li, S., Xie, Y., Zhang, H. and Chen, W. (2012) 'Cloning and heterologous expression of a bacteriocin sakacin P from *Lactobacillus sakei* in *Escherichia coli*.', *Applied microbiology and biotechnology*. Germany, 94(4), pp. 1061–1068. doi: 10.1007/s00253-012-3872-z.

Chevreux, B., Pfisterer, T., Drescher, B., Driesel, A. J., Muller, W. E. G., Wetter, T. and Suhai, S. (2004) 'Using the miraEST assembler for reliable and automated mRNA transcript assembly and SNP detection in sequenced ESTs.', *Genome research*. United States, 14(6), pp. 1147–1159. doi: 10.1101/gr.1917404.

Chikindas, M. L., Weeks, R., Drider, D., Chistyakov, V. A. and Dicks, L. M. (2018) 'Functions and emerging applications of bacteriocins', *Current opinion in biotechnology*. 2017/08/05, 49, pp. 23–28. doi: 10.1016/j.copbio.2017.07.011.

Chu Yuan Kee, M.-J. (2015) *Characterisation of Bacteriocins Active Against Clinically Relevant Strains of Clostridium difficile*. University of Nottingham.

Cintas, L. M., Casaus, P., Holo, H., Hernandez, P. E., Nes, I. F. and Håvarstein, L. S. (1998) 'Enterocins L50A and L50B, Two Novel Bacteriocins from *Enterococcus faecium* L50, are Related to Staphylococcal Hemolysins', *Journal of Bacteriology*, 180, pp. 1988–1994.

Cladera-Olivera, F., Caron, G. R. and Brandelli, A. (2004) 'Bacteriocin-like Substance Production by *Bacillus licheniformis* Strani P40', *Letters in Applied Microbiology*, 38, pp. 251–256.

Coakley, M., Fitzgerald, G. and Ros, R. P. (1997) 'Application and evaluation of the phage resistance- and bacteriocin-encoding plasmid pMRC01 for the improvement of dairy starter cultures.', *Applied and environmental microbiology*. United States, 63(4), pp. 1434–1440.

Collins, M. D., Lawson, P. A., Willems, A., Cordoba,', J. J., Fernandez-Garayzabal, J., Garcia, ' P, Cai, J., Hippe, ' H and Farrow', J. A. E. (1994) *The Phylogeny of the Genus Clostridium: Proposal of Five New Genera and Eleven New Species Combinations*.

Contreras, B. G., De Vuyst, L., Devreese, B., Busanyova, K., Raymaeckers, J., Bosman, F., Sablon, E. and Vandamme, E. J. (1997) 'Isolation, purification, and amino acid sequence of lactobin A, one of the two bacteriocins produced by *Lactobacillus amylovorus* LMG P-13139.', *Applied and environmental microbiology*. United States, 63(1), pp. 13–20.

Cookson, A. L., Noel, S. J., Kelly, W. J. and Attwood, G. T. (2004) 'The Use of PCR for

the Identification and Characterisation of Bacteriocin Genes from Bacterial Strains Isolated from Rumen or Caecal Contents of Cattle and Sheep', *FEMS Microbiology Ecology*, 48, pp. 199–207.

Costa, M. C., Reid-Smith, R., Gow, S., Hannon, S. J., Booker, C., Rousseau, J., Benedict, K. M., Morley, P. S. and Weese, J. S. (2012) 'Prevalence and molecular characterization of *Clostridium difficile* isolated from feedlot beef cattle upon arrival and mid-feeding period', *BMC Veterinary Research*, 8(1), p. 38. doi: 10.1186/1746-6148-8-38.

Costa, M. C., Stampfli, H. R., Arroyo, L. G., Pearl, D. L. and Weese, J. S. (2011) 'Epidemiology of *Clostridium difficile* on a veal farm: prevalence, molecular characterization and tetracycline resistance.', *Veterinary microbiology*. Netherlands, 152(3–4), pp. 379–384. doi: 10.1016/j.vetmic.2011.05.014.

Cotter, P. D., Hill, C. and Ross, R. P. (2005) 'Bacteriocins: developing innate immunity for food', *Nature Reviews Microbiology*, 3(10), pp. 777–788. doi: 10.1038/nrmicro1273.

Cotter, P., Ross, R. P. and Hill, C. (2013) 'Bacteriocins-A Viable Alternative To Antibiotics?', *Nature Reviews*, 11, pp. 95–105.

Crowther, G. S., Baines, S. D., Todhunter, S. L., Freeman, J., Chilton, C. H. and Wilcox, M. H. (2013) 'Evaluation of NVB302 versus vancomycin activity in an in vitro human gut model of *Clostridium difficile* infection.', *The Journal of antimicrobial chemotherapy*. England, 68(1), pp. 168–176. doi: 10.1093/jac/dks359.

Cruz, M. P. (2012) 'Fidaxomicin (Dificid), a Novel Oral Macrocyclic Antibacterial Agent For the Treatment of Clostridium difficile-Associated Diarrhea in Adults', *P & T : a peer-reviewed journal for formulary management*. MediMedia USA, Inc., 37(5), pp. 278–281. Available at: https://www.ncbi.nlm.nih.gov/pubmed/22876085.

Cutting, S. M. (2011) 'Bacillus probiotics.', *Food microbiology*. England, 28(2), pp. 214–220. doi: 10.1016/j.fm.2010.03.007.

Daba, G. M., Ishibashi, N., Gong, X., Taki, H., Yamashiro, K., Lim, Y. Y., Zendo, T. and Sonomoto, K. (2018) 'Characterisation of the action mechanism of a Lactococcus-specific bacteriocin, lactococcin Z', *Journal of Bioscience and Bioengineering*, 126(5), pp. 603–610. doi: https://doi.org/10.1016/j.jbiosc.2018.05.018.

Daley, P., Louie, T., Lutz, J. E., Khanna, S., Stoutenburgh, U., Jin, M., Adedoyin, A., Chesnel, L., Guris, D., Larson, K. B. and Murata, Y. (2017) 'Surotomycin versus vancomycin in adults with *Clostridium difficile* infection: primary clinical outcomes from the second pivotal, randomized, double-blind, Phase 3 trial', *Journal of Antimicrobial Chemotherapy*, 72(12), pp. 3462–3470. doi: 10.1093/jac/dkx299.

Darling, A. E., Mau, B., Blattner, F. R. and Perna, N. T. (2004) 'Mauve: Multiple Alignment of Conserved Genomic Sequence With Rearrangements', *Genome Research*, 14, pp. 1394–1403. doi: 10.1101.

Darling, A. E., Mau, B. and Perna, N. T. (2010) 'progressiveMauve: Multiple Genome Alignment with Gene Gain, Loss and Rearrangement', *PLOS ONE*, 5(6), p. e11147. doi: 10.1371/journal.pone.0011147.

Davies, J. K. and Reeves, P. (1975a) 'Genetics of Resistance to Colicins in *Escherichia coli* K-12: Cross-Resistances Among Colicins of Group A', *Journal of Bacteriology*,

123(1), pp. 102–117.

Davies, J. K. and Reeves, P. (1975b) 'Genetics of Resistance to Colicins in Escherichia coli K-14: Cross-Resistance Among Colicins of Group B', *Journal of Bacteriology*, 123(1), pp. 96–101.

Debast, S. B., van Leengoes, L. A. M. G., Goorhuis, A., Harmanus, C., Kuijper, E. J. and Bergwerff, A. A. (2009) '*Clostridium difficile* PCR Ribotype 078 Toxinotype V Found in Diarrhoeal Pigs Identical to Isolates from Affected Humans', *Environmental Microbiology*, 11(2), pp. 505–511.

Deegan, H. L., Cotter, P., Hill, C. and Ross, P. (2006) *Bacteriocins: Biological tools for bio-preservation and shelf-life extension*, *International Dairy Journal*. doi: 10.1016/j.idairyj.2005.10.026.

Deibel, R. H. (1964) 'The Group D Streptococci', *Bacteriological Reviews*, 28(3), pp. 330 LP – 366. Available at: http://mmbr.asm.org/content/28/3/330.abstract.

Deraz, S. F., Karlsson, E. N., Hedstrom, M., Andersson, M. M. and Mattiasson, B. (2005) 'Purification and characterisation of acidocin D20079, a bacteriocin produced by *Lactobacillus acidophilus* DSM 20079.', *Journal of biotechnology*. Netherlands, 117(4), pp. 343–354. doi: 10.1016/j.jbiotec.2005.02.005.

Desai, A., Marwah, V. S., Yadav, A., Jha, V., Dhaygude, K., Bangar, U., Kulkarni, V. and Jere, A. (2013) 'Identification of optimum sequencing depth especially for de novo genome assembly of small genomes using next generation sequencing data.', *PloS one*, 8(4), p. e60204. doi: 10.1371/journal.pone.0060204.

Dicks, L. M. T., Dreyer, L., Smith, C. and van Staden, A. D. (2018) 'A Review: The Fate of Bacteriocins in the Human Gastro-Intestinal Tract: Do They Cross the Gut–Blood Barrier? ', *Frontiers in Microbiology* , p. 2297. Available at: https://www.frontiersin.org/article/10.3389/fmicb.2018.02297.

Diep, D. B., Skaugen, M., Salehian, Z., Holo, H. and Nes, I. F. (2007) 'Common mechanisms of target cell recognition and immunity for class II bacteriocins', *Proceedings of the National Academy of Sciences*, 104(7), pp. 2384 LP – 2389. doi: 10.1073/pnas.0608775104.

Dirix, G., Monsieurs, P., Dombrecht, B., Daniels, R., Marchal, K., Vanderleyden, J. and Michiels, J. (2004) 'Peptide signal molecules and bacteriocins in Gram-negative bacteria: a genome-wide in silico screening for peptides containing a double-glycine leader sequence and their cognate transporters.', *Peptides*. United States, 25(9), pp. 1425–1440. doi: 10.1016/j.peptides.2003.10.028.

Dobbins, J. J. (2010) 'Prescott's Microbiology, Eighth Edition', *Journal of Microbiology & Biology Education : JMBE*. American Society of Microbiology, 11(1), pp. 64–65. doi: 10.1128/jmbe.v11.i1.154.

Dowd, S. E., Callaway, T. R., Wolcott, R. D., Sun, Y., McKeehan, T., Hagevoort, R. G. and Edrington, T. S. (2008) 'Evaluation of the bacterial diversity in the feces of cattle using 16S rDNA bacterial tag-encoded FLX amplicon pyrosequencing (bTEFAP)', *BMC microbiology*. BioMed Central, 8, p. 125. doi: 10.1186/1471-2180-8-125.

Draper, L. A., Cotter, P. D., Hill, C. and Ross, R. P. (2015) 'Lantibiotic Resistance', *Microbiology and Molecular Biology Reviews*, 79(2), pp. 171 LP – 191. doi: 10.1128/MMBR.00051-14.

Duc, L. H., Hong, H. A., Barbosa, T. M., Henriques, A. O. and Cutting, S. M. (2004) 'Characterization of Bacillus Probiotics Available for Human Use', *Applied and Environmental Microbiology*, 70(4), pp. 2161–2171.

DuPont, H. L. (2011) 'The Search for Effective Treatment of *Clostridium difficile* Infection', *The New England Journal of Medicine*, 364(5), pp. 473–475.

Elegado, F. B., Kim, W. J. and Kwon, D. Y. (1997) 'Rapid purification, partial characterization, and antimicrobial spectrum of the bacteriocin, Pediocin AcM, from *Pediococcus acidilactici* M.', *International journal of food microbiology*. Netherlands, 37(1), pp. 1–11. doi: 10.1016/s0168-1605(97)00037-8.

Engelke, G., Gutowski-Eckel, Z., Kiesau, P., Siegers, K., Hammelmann, M. and Entian, K. D. (1994) 'Regulation of nisin biosynthesis and immunity in *Lactococcus lactis* 6F3.', *Applied and environmental microbiology*. United States, 60(3), pp. 814–825.

FDA (2019) *Important Safety Alert Regarding Use of Fecal Microbiota for Transplantation and Risk of Serious Adverse Reactions Due to Transmission of Multi-Drug Resistant Organisms*, *fda.gov*. Available at: https://www.fda.gov/vaccines-blood-biologics/safety-availability-biologics/important-safety-alert-regarding-use-fecal-microbiota-transplantation-and-risk-serious-adverse (Accessed: 13 October 2019).

Febbraro, S., Hancock, A., Boyd, A. and Dawson, M. J. (2012) 'A Phase I, Double-Blind, Randomised, Placebo-Controlled, Dose Escalating Study to Assess the Safety, Tolerability, and Pharmacokinetics of Single and Multiple Doses of NVB302 Administered Orally to Healthy Volunteers', in *52nd Interscience Conference on Antimicrobial Agents and Chemotherapy*. San Francisco.

Fernández, E., Alegría, A., Delgado, S., Martín, M. C. and Mayo, B. (2011) 'Comparative phenotypic and molecular genetic profiling of wild *Lactococcus lactis* subsp. *lactis* strains of the *L. lactis* subsp. *lactis* and *L. lactis* subsp. *cremoris* genotypes, isolated from starter-free cheeses made of raw milk', *Applied and environmental microbiology*. American Society for Microbiology, 77(15), pp. 5324–5335. doi: 10.1128/AEM.02991-10.

Field, D., Begley, M., O'Connor, P. M., Daly, K. M., Hugenholtz, F., Cotter, P. D., Hill, C. and Ross, R. P. (2012) 'Bioengineered nisin A derivatives with enhanced activity against both Gram positive and Gram negative pathogens', *PloS one*. Public Library of Science, 7(10), pp. e46884–e46884. doi: 10.1371/journal.pone.0046884.

Field, D., Blake, T., Mathur, H., O' Connor, P. M., Cotter, P. D., Paul Ross, R. and Hill, C. (2019) 'Bioengineering nisin to overcome the nisin resistance protein', *Molecular Microbiology*. John Wiley & Sons, Ltd (10.1111), 111(3), pp. 717–731. doi: 10.1111/mmi.14183.

Field, D., Quigley, L., O'Connor, P. M., Rea, M. C., Daly, K., Cotter, P. D., Hill, C. and Ross, R. P. (2010) 'Studies With Bioengineered Nisin Peptides Highlight The Broad-spectrum Potency of Nisin V', *Microbial Biotechnology*, 3(4), pp. 473–486.

Finn, R. D. *et al.* (2016) 'The Pfam protein families database: towards a more sustainable future', *Nucleic Acids Research*, 44(D1), pp. D279–D285. doi: 10.1093/nar/gkv1344.

Flynn, S., van Sinderen, D., Thornton, G. M., Holo, H., Nes, I. F. and Collins, J. K. (2002) 'Characterization of the genetic locus responsible for the production of ABP-118, a

novel bacteriocin produced by the probiotic bacterium *Lactobacillus salivarius* subsp. *salivarius* UCC118.', *Microbiology (Reading, England)*. England, 148(Pt 4), pp. 973–984. doi: 10.1099/00221287-148-4-973.

From, C., Hormazabal, V. and Granum, P. E. (2007) 'Food poisoning associated with pumilacidin-producing *Bacillus pumilus* in rice.', *International journal of food microbiology*. Netherlands, 115(3), pp. 319–324. doi: 10.1016/j.ijfoodmicro.2006.11.005.

Froseth, B. R. and McKay, L. L. (1991) 'Molecular characterization of the nisin resistance region of *Lactococcus lactis* subsp. *lactis* biovar diacetylactis DRC3.', *Applied and Environmental Microbiology*, 57(3), pp. 804 LP – 811.

Fujita, K., Ichimasa, S., Zendo, T., Koga, S., Yoneyama, F., Nakayama, J. and Sonomoto, K. (2007) 'Structural analysis and characterization of lacticin Q, a novel bacteriocin belonging to a new family of unmodified bacteriocins of gram-positive bacteria.', *Applied and environmental microbiology*. United States, 73(9), pp. 2871–2877. doi: 10.1128/AEM.02286-06.

Fujitani, S., George, W. L., Morgan, M. A., Nichols, S. and Murthy, A. R. (2011) 'Implications for Vancomycin-Resistance Enterococcus Colonisation Associated with *Clostridium difficile* Infections', *American Journal of Infection Control*, 39(9), pp. 188–193.

Fukao, M., Obita, T., Yoneyama, F., Kohda, D., Zendo, T., Nakayama, J. and Sonomoto, K. (2008) 'Complete covalent structure of nisin Q, new natural nisin variant, containing post-translationally modified amino acids.', *Bioscience, biotechnology, and biochemistry*. England, 72(7), pp. 1750–1755. doi: 10.1271/bbb.80066.

Gabrielsen, C., Brede, D. A., Hernandez, P. E., Nes, I. F. and Diep, D. B. (2012) 'The maltose ABC transporter in *Lactococcus lactis* facilitates high-level sensitivity to the circular bacteriocin garvicin ML.', *Antimicrobial agents and chemotherapy*. United States, 56(6), pp. 2908–2915. doi: 10.1128/AAC.00314-12.

Gabrielsen, C., Brede, D. A., Nes, I. F. and Diep, D. B. (2014a) 'Circular bacteriocins: biosynthesis and mode of action.', *Applied and environmental microbiology*. United States, 80(22), pp. 6854–6862. doi: 10.1128/AEM.02284-14.

Gabrielsen, C., Brede, D. A., Salehian, Z., Nes, I. F. and Diep, D. B. (2014b) 'Functional genetic analysis of the GarML gene cluster in *Lactococcus garvieae* DCC43 gives new insights into circular bacteriocin biosynthesis.', *Journal of bacteriology*. United States, 196(5), pp. 911–919. doi: 10.1128/JB.01115-13.

Galal, A., A. El-Bana, A. and Janse, J. (2006) '*Bacillus pumilus*, A New Pathogen on Mango Plants', *Egypt J Phytopathol*, 34.

Gálvez, A., Gimenez-Gallego, G., Maqueda, M. and Valdivia, E. (1989) 'Purification and amino acid composition of peptide antibiotic AS-48 produced by *Streptococcus (Enterococcus) faecalis* subsp. *liquefaciens* S-48', *Antimicrobial Agents and Chemotherapy*, 33(4), pp. 437–441. doi: 10.1128/AAC.33.4.437.

Gálvez, A., Maqueda, M., Martinez-Bueno, M. and Valdivia, E. (1991) 'Permeation of bacterial cells, permeation of cytoplasmic and artificial membrane vesicles, and channel formation on lipid bilayers by peptide antibiotic AS-48.', *Journal of bacteriology*. United States, 173(2), pp. 886–892. doi: 10.1128/jb.173.2.886-892.1991.

Gálvez, A., Maqueda, M., Valdivia, E., Quesada, A. and Montoya, E. (1986) 'Charaterisation and Partial Purification of a Broad Spectrum Antibiotic AS-48 Produced by *Streptococcus faecalis*', *Canadian Journal of Microbiology*, 32(10), pp. 765–771.

Garcia-Alcalde, F., Okonechnikov, K., Carbonell, J., Cruz, L. M., Gotz, S., Tarazona, S., Dopazo, J., Meyer, T. F. and Conesa, A. (2012) 'Qualimap: evaluating next-generation sequencing alignment data.', *Bioinformatics (Oxford, England)*. England, 28(20), pp. 2678–2679. doi: 10.1093/bioinformatics/bts503.

Garey, K. W., Sethi, S., Yadav, Y. and DuPont, H. L. (2008) 'Meta-Analysis to Assess Risk Factors for Recurrent *Clostridium difficile* Infection', *Journal of Hospital Infection*, 70(4), pp. 298–304.

Gautheret, D. and Lambert, A. (2001) 'Direct RNA motif definition and identification from multiple sequence alignments using secondary structure profiles.', *Journal of molecular biology*. England, 313(5), pp. 1003–1011. doi: 10.1006/jmbi.2001.5102.

Gebhart, D., Lok, S., Clare, S., Tomas, M., Stares, M., Scholl, D., Donskey, C. J., Lawley, T. D. and Govoni, G. R. (2015) 'A Modified R-Type Bacteriocin Specifically Targeting *Clostridium difficile* Prevents Colonization of Mice without Affecting Gut Microbiota Diversity', *Mbio*. [Gebhart, Dana; Lok, Stephen; Scholl, Dean; Govoni, Gregory R.] AvidBiotics Corp, San Francisco, CA 94080 USA. [Clare, Simon] Wellcome Trust Sanger Inst, Microbial Pathogenesis Lab, Hinxton, England. [Tomas, Myreen; Donskey, Curtis J.] Louis Stokes Clevel, 6(2). doi: 10.1128/mBio.02368-14.

Geertsma, E. and Poolman, B. (2007) 'High-throughput Cloning and Expression in *Lactococcus lactis*', *Nature Methods*. doi: 10.1038/nprot.2007.290.

Genth, H., Dreger, S. C., Huelsenbeck, J. and Just, I. (2008) '*Clostridium difficile* Toxins: More Than Mere Inhibitors of Rho Proteins', *The International Journal of Biochemistry & Cell Biology*, 40(4), pp. 592–597.

Georgalaki, M. D., Van den Berghe, E., Kritikos, D., Devreese, B., Van Beeumen, J., Kalantzopoulos, G., De Vuyst, L. and Tsakalidou, E. (2002) 'Macedocin, a Food-Grade Lantibioitc Produced by *Streptococcus macedonicus* ACA-DC 198', *Applied and Environmental Microbiology*, 68(12), pp. 5891–5903.

George, R. H., Symonds, J. M., Dimock, F., Brown, J. D., Arabi, Y., Shinagawa, N., Alexander-Williams, J. and Burdon, D. W. (1978) 'Identification of *Clostridium difficile* as a Cause of Pseudomembranous Colitis', *British Medical Journal*, 1(6114), p. 695.

Gerber, S. D. and Solioz, M. (2007) 'Efficient transformation of *Lactococcus lactis* IL1403 and generation of knock-out mutants by homologous recombination.', *Journal of basic microbiology*. Germany, 47(3), pp. 281–286. doi: 10.1002/jobm.200610297.

Gerding, D. N., Johnson, S., Rupnik, M. and Aktories, K. (2014) '*Clostridium difficile* binary toxin CDT: mechanism, epidemiology, and potential clinical importance.', *Gut microbes*, 5(1), pp. 15–27. doi: 10.4161/gmic.26854.

Ghantoji, S. S., Sail, K., Lairson, D. R., DuPont, H. L. and Garey, K. W. (2010) 'Economic Healthcare Costs of *Clostridium difficile* Infection: A Systematic Review', *Journal of Hospital Infection*, 74(4), pp. 309–318.

Ghosh, C., Sarkar, P., Issa, R. and Haldar, J. (2019) 'Alternatives to Conventional Antibiotics in the Era of Antimicrobial Resistance', *Trends in Microbiology*, 27(4), pp.

323–338. doi: https://doi.org/10.1016/j.tim.2018.12.010.

Giles, E. M., D'Adamo, G. L. and Forster, S. C. (2019) 'The future of faecal transplants', *Nature Reviews Microbiology*. doi: 10.1038/s41579-019-0271-9.

Gong, X., Martin-Visscher, L. A., Nahirney, D., Vederas, J. C. and Duszyk, M. (2009) 'The circular bacteriocin, carnocyclin A, forms anion-selective channels in lipid bilayers', *Biochimica et Biophysica Acta (BBA) - Biomembranes*, 1788(9), pp. 1797–1803. doi: https://doi.org/10.1016/j.bbamem.2009.05.008.

Goorhuis, A., Bakker, D., Corver, J., Debast, S. B., Harmanus, C., Notermans, D. W., Bergwerff, A. A., Dekker, F. W. and Kuijper, E. J. (2008) 'Emergence of *Clostridium difficile* Infection Due to a New Hypervirulent Strain, Polymerase Chain Reaction Ribotype 078', *Clinical Infectious Disease*, 47(9), pp. 1162–1170.

Goudarzi, M., Seyedjavadi, S. S., Goudarzi, H., Aghdam, E. M. and Nazeri, S. (2014) '*Clostridium difficile* Infection: Epidemiology, Pathogenesis, Risk Factors, and Therapeutic Options', *Scientifica*, 2014, pp. 1–9.

Grandesso, S., Arena, F., Eseme, F. E., Panese, S., Henrici De Angelis, L., Spigaglia, P., Barbanti, F. and Rossolini, G. M. (2016) '*Clostridium difficile* ribotype 033 colitis in a patient following broad-spectrum antibiotic treatment for KPC producing *Klebsiella pneumoniae* infection, Italy.', *The new microbiologica*, 39(3), pp. 235–236. Available at: http://www.ncbi.nlm.nih.gov/pubmed/27284989 (Accessed: 21 October 2019).

Gratia, A. (1925) 'Sur un remarquable exemple d'antagonisme entre deux souches de coilbacille', *C. R. Seances Soc. Biol. Fil.*, 93, pp. 1040–1041. Available at: https://ci.nii.ac.jp/naid/10027104803/en/.

Gray, E. J., Lee, K. D., Souleimanov, A. M., Di Falco, M. R., Zhou, X., Ly, A., Charles, T. C., Driscoll, B. T. and Smith, D. L. (2006) 'A novel bacteriocin, thuricin 17, produced by plant growth promoting rhizobacteria strain *Bacillus thuringiensis* NEB17: isolation and classification', *Journal of Applied Microbiology*. John Wiley & Sons, Ltd (10.1111), 100(3), pp. 545–554. doi: 10.1111/j.1365-2672.2006.02822.x.

Green, D. H., Wakeley, P. R., Page, A., Barnes, A., Baccigalupi, L., Ricca, E. and Cutting, S. M. (1999) 'Characterization of two Bacillus probiotics', *Applied and environmental microbiology*. American Society for Microbiology, 65(9), pp. 4288–4291. Available at: https://www.ncbi.nlm.nih.gov/pubmed/10473456.

Gross, E. (1973) 'Subtilin VI : Die struktur des subtilins', *Hoppe-Seyler's Z. Physiol. Chem.*, 354, pp. 810–812. Available at: https://ci.nii.ac.jp/naid/10011440128/en/.

Grześkowiak, Ł., Zentek, J. and Vahjen, W. (2016) 'Determination of the extent of *Clostridium difficile* colonisation and toxin accumulation in sows and neonatal piglets', *Anaerobe*, 40, pp. 5–9. doi: https://doi.org/10.1016/j.anaerobe.2016.04.012.

Guder, A., Wiedemann, I. and Sahl, H.-G. (2000) 'Posttranslationally modified bacteriocins—the lantibiotics', *Peptide Science*. John Wiley & Sons, Ltd, 55(1), pp. 62–73. doi: 10.1002/1097-0282(2000)55:1<62::AID-BIP60>3.0.CO;2-Y.

Guinane, C. M., Cotter, P. D., Hill, C. and Ross, R. P. (2005) 'Microbial Solutions to Microbial Problems; Lactococcal Bacteriocins for the Control of Undesirable Biota in Food', *Journal of Applied Microbiology*, 98(6), pp. 1316–1325.

Güllüce, M., Karadayi, M. and Barış, Ö. (2013) 'Bacteriocins: Promising Natural Antimicrobials', in University, A. (ed.) *Microbial Pathogens and Strategies for*

Combating Them: Science, Technology and Education. Turkey: Formatex, pp. 1016–1027.

Gupta, A. and Khanna, S. (2014) 'Community-Acquired *Clostridium difficile* Infection: An Increasing Public Health Threat', *Infection and Drug Resistance*, 7, pp. 63–72.

Gupta, S., Allen-Vercoe, E. and Petrof, E. O. (2016) 'Fecal Mircrobiota Transplanatation: In Perspective', *Therapeutic Advances in Gastroenterology*, 9(2), pp. 229–239.

Hafiz, S. (1974) *Clostridium difficile and its toxins*. University of Leeds. Available at: http://europepmc.org/theses/ETH/490383.

Hagey, J. V, Bhatnagar, S., Heguy, J. M., Karle, B. M., Price, P. L., Meyer, D. and Maga, E. A. (2019) 'Fecal Microbial Communities in a Large Representative Cohort of California Dairy Cows', *Frontiers in Microbiology* , p. 1093. Available at: https://www.frontiersin.org/article/10.3389/fmicb.2019.01093.

Hall, I. C. and O'Toole, E. (1935) 'Intestinal Flora in New-Born Infants: With a Description of a New Pathogenic Anaerobe, Bacillus Difficilis', *JAMA Pediatrics*, 49(2), pp. 390–402. doi: 10.1001/archpedi.1935.01970020105010.

Hammami, R., Fernandez, B., Lacroix, C. and Fliss, I. (2013) 'Anti-infective Properties of Bacteriocins: An Update', *Cellular and Molecular Life Sciences*, 70, pp. 2947–2967.

Hammami, R., Zouhir, A., Le Lay, C., Ben Hamida, J. and Fliss, I. (2010) 'BACTIBASE second release: a database and tool platform for bacteriocin characterization', *BMC Microbiol*. 2010/01/29. STELA Dairy Research Center, Nutraceuticals and Functional Foods Institute, Universite Laval, G1K 7P4 Quebec, QC, Canada., 10, p. 22. doi: 10.1186/1471-2180-10-22.

Hammond, G. A. and Johnson, J. L. (1995) 'The toxigenic element of *Clostridium difficile* strain VPI 10463', *Microbial Pathogenesis*, 19(4), pp. 203–213. doi: 10.1016/S0882-4010(95)90263-5.

Hargreaves, K. R. and Clokie, M. R. J. (2014) '*Clostridium difficile* phages: still difficult? ', *Frontiers in Microbiology* , p. 184. Available at: https://www.frontiersin.org/article/10.3389/fmicb.2014.00184.

Hasper, H. E., de Kruijff, B. and Breukink, E. (2004) 'Assembly and stability of nisin-lipid II pores.', *Biochemistry*. United States, 43(36), pp. 11567–11575. doi: 10.1021/bi049476b.

Hatziioanou, D., Gherghisan-Filip, C., Saalbach, G., Horn, N., Wegmann, U., Duncan, S. H., Flint, H. J., Mayer, M. J. and Narbad, A. (2017) 'Discovery of a novel lantibiotic nisin O from *Blautia obeum* A2-162, isolated from the human gastrointestinal tract.', *Microbiology (Reading, England)*. England, 163(9), pp. 1292–1305. doi: 10.1099/mic.0.000515.

Havarstein, L. S., Diep, D. B. and Nes, I. F. (1995) 'A family of bacteriocin ABC transporters carry out proteolytic processing of their substrates concomitant with export.', *Molecular microbiology*. England, 16(2), pp. 229–240. doi: 10.1111/j.1365-2958.1995.tb02295.x.

He, M. *et al.* (2013) 'Emergence and global spread of epidemic healthcare-associated *Clostridium difficile*.', *Nature genetics*. United States, 45(1), pp. 109–113. doi: 10.1038/ng.2478.

Heap, J. T., Ehsaan, M., Cooksley, C. M., Ng, Y.-K., Cartman, S. T., Winzer, K. and Minton, N. P. (2012) 'Integration of DNA into bacterial chromosomes from plasmids without a counter-selection marker', *Nucleic Acids Research*, 40(8), pp. e59–e59. doi: 10.1093/nar/gkr1321.

Heap, J. T., Pennington, O. J., Cartman, S. T. and Minton, N. P. (2009) 'A modular system for Clostridium shuttle plasmids', *Journal of Microbiological Methods*, 78(1), pp. 79–85. doi: https://doi.org/10.1016/j.mimet.2009.05.004.

van Heel, A. J., de Jong, A., Song, C., Viel, J. H., Kok, J. and Kuipers, O. P. (2018) 'BAGEL4: a user-friendly web server to thoroughly mine RiPPs and bacteriocins', *Nucleic Acids Res*. 2018/05/23. Molecular Genetics, GBB, University of Groningen, Groningen, 9747AG, the Netherlands., 46(W1), pp. W278-w281. doi: 10.1093/nar/gky383.

van Heel, A. J., Montalban-Lopez, M., Oliveau, Q. and Kuipers, O. P. (2017) 'Genome-guided identification of novel head-to-tail cyclized antimicrobial peptides, exemplified by the discovery of pumilarin', *Microbial genomics*. Microbiology Society, 3(10), pp. e000134–e000134. doi: 10.1099/mgen.0.000134.

Heijenoort, J. van (2001) 'Recent advances in the formation of the bacterial peptidoglycan monomer unit', *Natural Product Reports*. The Royal Society of Chemistry, 18(5), pp. 503–519. doi: 10.1039/A804532A.

Heng, N. C. K., Wescombe, P. A., Burton, J. P., Jack, R. W. and Tagg, J. R. (2007) 'The Diversity of Bacteriocins in Gram-Positive Bacteria BT - Bacteriocins: Ecology and Evolution', in Riley, M. A. and Chavan, M. A. (eds). Berlin, Heidelberg: Springer Berlin Heidelberg, pp. 45–92. doi: 10.1007/978-3-540-36604-1_4.

Hetz, C., Bono, M. R., Barros, L. F. and Lagos, R. (2002) 'Microcin E492, a channel-forming bacteriocin from *Klebsiella pneumoniae*, induces apoptosis in some human cell lines', *Proceedings of the National Academy of Sciences of the United States of America*, 99(5), pp. 2696–2701. doi: 10.1073/pnas.052709699.

van Heusden, H. E., de Kruijff, B. and Breukink, E. (2002) 'Lipid II induces a transmembrane orientation of the pore-forming peptide lantibiotic nisin.', *Biochemistry*. United States, 41(40), pp. 12171–12178. doi: 10.1021/bi026090x.

Himeno, K. *et al.* (2015) 'Identification, Characterization, and Three-Dimensional Structure of the Novel Circular Bacteriocin, Enterocin NKR-5-3B, from *Enterococcus faecium*.', *Biochemistry*. United States, 54(31), pp. 4863–4876. doi: 10.1021/acs.biochem.5b00196.

Himes, P. M., Allen, S. E., Hwang, S. and Bowers, A. A. (2016) 'Production of Sactipeptides in Escherichia coli: Probing the Substrate Promiscuity of Subtilosin A Biosynthesis.', *ACS chemical biology*, 11(6), pp. 1737–44. doi: 10.1021/acschembio.6b00042.

Hoffer, E., Haechler, H., Frei, R. and Stephan, R. (2010) 'Low occurrence of *Clostridium difficile* in fecal samples of healthy calves and pigs at slaughter and in minced meat in Switzerland.', *Journal of food protection*. United States, 73(5), pp. 973–975. doi: 10.4315/0362-028x-73.5.973.

Holo, H. and Nes, I. F. (1995) 'Transformation of Lactococcus by Electroporation', *Methods Mol Biol*. 1995/01/01. Laboratory of Microbial Gene Technology, Agricultural University of Norway, As., 47, pp. 195–199. doi: 10.1385/0-89603-310-

4:195.

Holo, H., Nilssen, Ø. and Nes, I. F. (1991) 'Lactococcin A, a New Bacteriocin for *Lactococcus lactis* subsp. *cremoris*: Isolation and Characterization of the Protein and Its Gene', *American Society for Microbiology*, 173(12).

Hsu, S.-T., Breukink, E., de Kruijff, B., Kaptein, R., Bonvin, A. M. J. J. and van Nuland, N. A. J. (2002) 'Mapping the targeted membrane pore formation mechanism by solution NMR: the nisin Z and lipid II interaction in SDS micelles.', *Biochemistry*. United States, 41(24), pp. 7670–7676. doi: 10.1021/bi025679t.

Inc. P (2017) 'A Phase 3, Placebo-Controlled, Randomized, Observer-Blinded Study To Evaluate The Efficacy, Safety, And Tolerability Of A Clostridium Difficile Vaccine In Adults 50 Years Of Age And Older'.

Ishibashi, N., Seto, H., Koga, S., Zendo, T. and Sonomoto, K. (2015) 'Identification of Lactococcus-Specific Bacteriocins Produced by Lactococcal Isolates, and the Discovery of a Novel Bacteriocin, Lactococcin Z', *Probiotics and Antimicrobial Proteins*, 7(3), pp. 222–231. doi: 10.1007/s12602-015-9196-4.

Janezic, S. *et al.* (2014) 'International *Clostridium difficile* animal strain collection and large diversity of animal associated strains', *BMC Microbiology*, 14(1), p. 173. doi: 10.1186/1471-2180-14-173.

Jarvis, B. (1970) 'Enzymic reduction of the C-terminal dehydroalanyl-lysine sequence in nisin', *Biochemical Journal*, 119(5), pp. 56P LP-56P. doi: 10.1042/bj1190056P.

Jasniewski, J., Cailliez-Grimal, C., Gelhaye, E. and Revol-Junelles, A.-M. (2008) 'Optimization of the production and purification processes of carnobacteriocins Cbn BM1 and Cbn B2 from *Carnobacterium maltaromaticum* CP5 by heterologous expression in *Escherichia coli*.', *Journal of microbiological methods*. Netherlands, 73(1), pp. 41–48. doi: 10.1016/j.mimet.2008.01.008.

Jia, B. and Jeon, C. O. (2016) 'High-throughput recombinant protein expression in *Escherichia coli*: Current status and future perspectives', *Open Biology*. Royal Society of London. doi: 10.1098/rsob.160196.

Johnson, A. P. and Wilcox, M. H. (2012) 'Fidaxomicin: A New Option for the Treatment of *Clostridium difficile* Infection', *Journal of Antimicrobial Chemotherapy*, 67, pp. 2788–2792.

Kalliomäki, M. A. and Walker, W. A. (2005) 'Physiologic and Pathologic Interactions of Bacteria with Gastrointestinal Epithelium', *Gastroenterology Clinics of North America*, 34(3), pp. 383–399.

Kalmokoff, M. L., Bartlett, F. and Teather, R. M. (1996) 'Are Ruminal Bacteria Armed with Bacteriocins?', *Journal of Dairy Science*, 79(12), pp. 2297–2306. doi: https://doi.org/10.3168/jds.S0022-0302(96)76608-0.

Kalmokoff, M. L., Cyr, T. D., Hefford, M. A., Whitford, M. F. and Teather, R. M. (2003) 'Butyrivibriocin AR10, a new cyclic bacteriocin produced by the ruminal anaerobe *Butyrivibrio fibrisolvens* AR10: characterization of the gene and peptide.', *Canadian journal of microbiology*. Canada, 49(12), pp. 763–773. doi: 10.1139/w03-101.

Kamuju, V., Kumar, S., Khan, W. H. and Vivekanandan, P. (2018) 'Hypervirulent *Clostridium difficile* ribotypes are CpG depleted', *Virulence*. Taylor & Francis, 9(1), pp. 1422–1425. doi: 10.1080/21505594.2018.1509669.

Kawai, Y., Arakawa, K., Itoh, A., Saitoh, B., Ishii, Y., Nishimura, J., Kitazawa, H., Itoh, T. and Saito, T. (2003) 'Heterologous expression of gassericin A, a bacteriocin produced by *Lactobacillus gasseri* LA39', *Animal Science Journal*. John Wiley & Sons, Ltd (10.1111), 74(1), pp. 45–51. doi: 10.1046/j.1344-3941.2003.00085.x.

Kawai, Y., Kemperman, R., Kok, J. and Saito, T. (2004) 'The circular bacteriocins gassericin A and circularin A.', *Current protein & peptide science*. United Arab Emirates, 5(5), pp. 393–398.

Kawulka, K., Sprules, T., McKay, R. T., Mercier, P., Diaper, C. M., Zuber, P. and Vederas, J. C. (2003) 'Structure of subtilosin A, an antimicrobial peptide from *Bacillus subtilis* with unusual posttranslational modifications linking cysteine sulfurs to alpha-carbons of phenylalanine and threonine.', *Journal of the American Chemical Society*. United States, 125(16), pp. 4726–4727. doi: 10.1021/ja029654t.

Kayalvizhi, N., Rameshkumar, N. and Gunasekaran, P. (2016) 'Cloning and characterization of mersacidin like bacteriocin from *Bacillus licheniformis* MKU3 in *Escherichia coli*', *Journal of food science and technology*. 2016/05/27. Springer India, 53(5), pp. 2298–2306. doi: 10.1007/s13197-016-2195-y.

Keel, K., Brazier, J. S., Post, K. W., Weese, S. and Songer, J. G. (2007) 'Prevalence of PCR Ribotypes Among *Clostridium difficile* Isolated from Pigs, Calves, and Other Species', *Journal of Clinical Microbiology*, 45(6), pp. 1963–1964.

Keessen, E. C., Harmanus, C., Dohmen, W., Kuijper, E. J. and Lipman, L. J. A. (2013) '*Clostridium difficile* Infection Associated with Pig Farms', *Emerging Infectious Diseases*, 19(6), pp. 1032–1034.

Kelly, C. P. (2012) 'Can We Identify Patients at High Risk of Recurrent *Clostridium difficile* Infection?', *Clinical Microbiology and Infection*, 18(6), pp. 21–27.

Kemperman, R., Kuipers, A., Karsens, H., Nauta, A., Kuipers, O. and Kok, J. (2003) 'Identification and Characterisation of Two Novel Clostridial Bacteriocins, Circularin A and Closticin 574', *Applied and Environmental Microbiology*, 69(3), pp. 1589–1597.

Khan, H., Flint, S. H. and Yu, P.-L. (2013) 'Determination of the mode of action of enterolysin A, produced by *Enterococcus faecalis* B9510.', *Journal of applied microbiology*. England, 115(2), pp. 484–494. doi: 10.1111/jam.12240.

Khan, I. and Oh, D.-H. (2016) 'Integration of nisin into nanoparticles for application in foods', *Innovative Food Science & Emerging Technologies*, 34, pp. 376–384. doi: https://doi.org/10.1016/j.ifset.2015.12.013.

Khanna, S. and Pardi, D. S. (2010) 'The growing incidence and severity of *Clostridium difficile* infection in inpatient and outpatient settings', *Expert Review of Gastroenterology & Hepatology*. Taylor & Francis, 4(4), pp. 409–416. doi: 10.1586/egh.10.48.

Khanna, S., Pardi, D. S., Aronson, S. L., Kammer, P. P., Orenstein, R., St Sauver, J. L., Harmsen, W. S. and Zinsmeister, A. R. (2012) 'The Epidemiology of Community-Acquired *Clostridium difficile* Infection: A Population-Based Study', *The American Journal of Gastroenterology*, 107, pp. 89–95.

Khosa, S., AlKhatib, Z. and Smits, S. H. J. (2013) 'NSR from *Streptococcus agalactiae* confers resistance against nisin and is encoded by a conserved nsr operon.', *Biological chemistry*. Germany, 394(11), pp. 1543–1549. doi: 10.1515/hsz-2013-0167.

Khusainov, R., van Heel, A. J., Lubelski, J., Moll, G. N. and Kuipers, O. P. (2015) 'Identification Of Essential Amino Acid Residues In The Nisin Dehydratase NisB', *Frontiers in Microbiology*.

Kimouli, M., Vrioni, G., Papadopoulou, M., Koumaki, V., Petropoulou, D., Gounaris, A., Friedrich, A. W. and Tsakris, A. (2012) 'Two Cases of Severe Sepsis Cause by *Bacillus pumilus* in Neonatal Infants', *Journal of Medical Microbiology*, 61, pp. 596–599.

Kirk, J. A., Gebhart, D., Buckley, A. M., Lok, S., Scholl, D., Douce, G. R., Govoni, G. R. and Fagan, R. P. (2017) 'New class of precision antimicrobials redefines role of *Clostridium difficile* S-layer in virulence and viability', *Science Translational Medicine*. American Association for the Advancement of Science, 9(406). doi: 10.1126/scitranslmed.aah6813.

Kjos, M., Snipen, L., Salehian, Z., Nes, I. F. and Diep, D. B. (2010) 'The Abi Proteins and Their Involvement in Bacteriocin Self-Immunity', *Journal of Bacteriology*, 192(8), pp. 2068–2076.

Klaenhammer, T. R. (1988) 'Bacteriocins of Lactic Acid Bacteria', *Biochimie*, 70(3), pp. 337–349.

Kleerebezem, M. (2004) 'Quorum sensing control of lantibiotic production; nisin and subtilin autoregulate their own biosynthesis', *Peptides*, 25(9), pp. 1405–1414. doi: https://doi.org/10.1016/j.peptides.2003.10.021.

Kleerebezem, M., de Vos, W. M. and Kuipers, O. P. (1999) 'The lantibiotics nisin and subtilin act as extracellular regulators of their own biosynthesis', in Dunny, G. M. and Winans, S. C. (eds) *Cell-cell signalling in bacteria*. Washington D.C.: ASM Press, pp. 159–174.

Klostermann, K., Crispie, F., Flynn, J., Meaney, W. J., Paul Ross, R. and Hill, C. (2010) 'Efficacy of a teat dip containing the bacteriocin lacticin 3147 to eliminate Gram-positive pathogens associated with bovine mastitis', *Journal of Dairy Research*. 2009/09/29. Cambridge University Press, 77(2), pp. 231–238. doi: DOI: 10.1017/S0022029909990239.

Knerr, P. J. and van der Donk, W. A. (2012) 'Discovery, Biosynthesis, and Engineering of Lantipeptides', *Annual Review of Biochemistry*. Annual Reviews, 81(1), pp. 479–505. doi: 10.1146/annurev-biochem-060110-113521.

Knetsch, C. W. *et al.* (2014) 'Whole genome sequencing reveals potential spread of *Clostridium difficile* between humans and farm animals in the Netherlands, 2002 to 2011', *Euro surveillance : bulletin Europeen sur les maladies transmissibles = European communicable disease bulletin*, 19(45), p. 20954. doi: 10.2807/1560-7917.es2014.19.45.20954.

Knetsch, C. W. *et al.* (2018) 'Zoonotic Transfer of *Clostridium difficile* Harboring Antimicrobial Resistance between Farm Animals and Humans', *Journal of Clinical Microbiology*. Edited by B. Fenwick, 56(3), pp. e01384-17. doi: 10.1128/JCM.01384-17.

Knetsch, C. W., Lawley, T. D., Hensgens, M. P., Corver, J., Wilcox, M. W. and Kuijper, E. J. (2013) 'Current application and future perspectives of molecular typing methods to study *Clostridium difficile* infections', *Eurosurveillance*, 18(4). doi: https://doi.org/10.2807/ese.18.04.20381-en.

Knight, D. R., Elliott, B., Chang, B. J., Perkins, T. T. and Riley, T. V (2015) 'Diversity and Evolution in the Genome of *Clostridium difficile*', *Clinical Microbiology Reviews*, 28(3), pp. 721 LP – 741. doi: 10.1128/CMR.00127-14.

Knight, D. R., Kullin, B., Androga, G. O., Barbut, F., Eckert, C., Johnson, S., Spigaglia, P., Tateda, K., Tsai, P.-J. and Riley, T. V (2019) 'Evolutionary and Genomic Insights into *Clostridioides difficile* Sequence Type 11: a Diverse Zoonotic and Antimicrobial-Resistant Lineage of Global One Health Importance', *mBio*. Edited by A. B. P. Onderdonk Robin Smits, Wiep Klaas, 10(2), pp. e00446-19. doi: 10.1128/mBio.00446-19.

Knight, D. R. and Riley, T. V (2019) 'Genomic Delineation of Zoonotic Origins of *Clostridium difficile*', *Frontiers in Public Health* , p. 164. Available at: https://www.frontiersin.org/article/10.3389/fpubh.2019.00164.

Knight, D. R., Thean, S., Putsathit, P., Fenwick, S. and Riley, T. V (2013) 'Cross-Sectional Study Reveals High Prevalence of *Clostridium difficile* Non-PCR Ribotype 078 Strains in Australian Veal Calves at Slaughter', *Applied and Environmental Microbiology*, 79(8), pp. 2630 LP – 2635. doi: 10.1128/AEM.03951-12.

Koene, M. G. J., Mevius, D., Wagenaar, J. A., Harmanus, C., Hensgens, M. P. M., Meetsma, A. M., Putirulan, F. F., van Bergen, M. A. P. and Kuijper, E. J. (2012) '*Clostridium difficile* in Dutch animals: their presence, characteristics and similarities with human isolates.', *Clinical microbiology and infection : the official publication of the European Society of Clinical Microbiology and Infectious Diseases*. England, 18(8), pp. 778–784. doi: 10.1111/j.1469-0691.2011.03651.x.

Koponen, O., Takala, T. M., Saarela, U., Qiao, M. and Saris, P. E. J. (2004) 'Distribution of the NisI immunity protein and enhancement of nisin activity by the lipid-free NisI.', *FEMS microbiology letters*. England, 231(1), pp. 85–90. doi: 10.1016/S0378-1097(03)00934-0.

Kramer, N. E., van Hijum, S. A. F. T., Knol, J., Kok, J. and Kuipers, O. P. (2006) 'Transcriptome Analysis Reveals Mechanisms by Which *Lactococcus lactis* Acquires Nisin Resistance', *Antimicrobial Agents and Chemotherapy*, 50(5), pp. 1753 LP – 1761. doi: 10.1128/AAC.50.5.1753-1761.2006.

Krutova, M., Wilcox, M. H., Kuijper, E. J. and Allerberger, F. (2018) 'The pitfalls of laboratory diagnostics of *Clostridium difficile* infection', *Clinical Microbiology and Infection*, 24, pp. 682–683. doi: 10.1016/j.cmi.2018.02.026.

Krzyściak, W., Pluskwa, K. K., Jurczak, A. and Kościelniak, D. (2013) 'The pathogenicity of the Streptococcus genus', *European journal of clinical microbiology & infectious diseases : official publication of the European Society of Clinical Microbiology*. 2013/07/03. Springer Berlin Heidelberg, 32(11), pp. 1361–1376. doi: 10.1007/s10096-013-1914-9.

Kuehne, S. A., Cartman, S. T., Heap, J. T., Kelly, M. L., Cockayne, A. and Minton, N. P. (2010) 'The Role of Toxin A and Toxin B in *Clostridium difficile* Infection', *Nature*, 467(7316), pp. 711–713.

Kuipers, A., de Boef, E., Rink, R., Fekken, S., Kluskens, L. D., Driessen, A. J. M., Leenhouts, K., Kuipers, O. P. and Moll, G. N. (2004) 'NisT, the transporter of the lantibiotic nisin, can transport fully modified, dehydrated, and unmodified prenisin and fusions of the leader peptide with non-lantibiotic peptides.', *The Journal of biological chemistry*. United States, 279(21), pp. 22176–22182. doi:

10.1074/jbc.M312789200.

Kuipers, O. P., Beerthuyzen, M. M., de Ruyter, P. G. G. A., Luesink, E. J. and de Vos, W. M. (1995) 'Autoregulation of Nisin Biosynethesis in *Lactococcus lactis* by Signal Transduction', *The Journal of Biological Chemistry*, 270(45), pp. 27299–27304.

Kuipers, O. P., Beerthuyzen, M. M., Siezen, R. J. and De Vos, W. M. (1993) 'Characterization Of The Nisin Gene Cluster nisABTCIPR Of *Lactococcus lactis*. Requirement Of Expression Of The nisA And nisI Genes For Development Of Immunity', *European Journal of Biochemistry*, 216(1), pp. 281–291. Available at: internal-pdf://189.27.88.186/KUIPERS 1993 Nisin.pdf.

Kumar, N. *et al.* (2019) 'Adaptation of host transmission cycle during *Clostridium difficile* speciation', *Nature Genetics*, 51(9), pp. 1315–1320. doi: 10.1038/s41588-019-0478-8.

Kumar, S., Stecher, G. and Tamura, K. (2016) 'MEGA7: Molecular Evolutionary Genetics Analysis Version 7.0 for Bigger Datasets.', *Molecular biology and evolution*. United States, 33(7), pp. 1870–1874. doi: 10.1093/molbev/msw054.

de Kwaadsteniet, M., Ten Doeschate, K. and Dicks, L. M. T. (2008) 'Characterization of the structural gene encoding nisin F, a new lantibiotic produced by a *Lactococcus lactis* subsp. *lactis* isolate from freshwater catfish (Clarias gariepinus).', *Applied and environmental microbiology*. United States, 74(2), pp. 547–549. doi: 10.1128/AEM.01862-07.

LaFrance, M. E., Farrow, M. A., Chandrasekaran, R., Sheng, J., Rubin, D. H. and Lacy, D. B. (2015) 'Identification of an Epithelial Cell Receptor Responsible for *Clostridium difficile* TcdB-Induced Cytotoxicity', *Proc Natl Acad Sci U S A*, 112(22), pp. 7073–7078.

Lancefield, R. C. (1933) 'A Serological Differentiation of Human and Other Groups of Hemolytic Streptococci', *The Journal of Experimental Medicine*, 57(4), pp. 571 LP – 595. doi: 10.1084/jem.57.4.571.

Lawson, P. A., Citron, D. M., Tyrrell, K. L. and Finegold, S. M. (2016) 'Reclassification of *Clostridium difficile* as *Clostridioides difficile* (Hall and O'Toole 1935) Prevot 1938.', *Anaerobe*. England, 40, pp. 95–99. doi: 10.1016/j.anaerobe.2016.06.008.

Lawson, P. A. and Rainey, F. A. (2016) 'Proposal to restrict the genus Clostridium Prazmowski to *Clostridium butyricum* and related species', *International Journal of Systematic and Evolutionary Microbiology*. Microbiology Society, 66(2), pp. 1009–1016. doi: https://doi.org/10.1099/ijsem.0.000824.

Le Lay, C., Dridi, L., Bergeron, M. G., Ouellette, M. and Fliss, I. L. (2016) 'Nisin is an effective inhibitor of *Clostridium difficile* vegetative cells and spore germination', *J Med Microbiol*. 65(2), pp. 169–175. doi: 10.1099/jmm.0.000202.

Lee, K.-H., Jun, K.-D., Kim, W.-S. and Paik, H.-D. (2001) 'Partial characterization of polyfermenticin SCD, a newly identified bacteriocin of *Bacillus polyfermenticus*', *Letters in Applied Microbiology*. John Wiley & Sons, Ltd (10.1111), 32(3), pp. 146–151. doi: 10.1046/j.1472-765x.2001.00876.x.

Leeds, J. A., Sachdeva, M., Mullin, S., Barnes, S. W. and Ruzin, A. (2014) 'In Vitro Selection, via Serial Passage, of *Clostridium difficile* Mutants with Reduced Susceptibility to Fidaxomicin or Vancomycin', *Journal of Antimicrobial Chemotherapy*, 69(1), pp. 41–44.

Lees, E. A., Miyajima, F., Pirmohamed, M. and Carrol, E. D. (2016) 'The role of *Clostridium difficile* in the paediatric and neonatal gut — a narrative review', *European Journal of Clinical Microbiology & Infectious Diseases*, 35(7), pp. 1047–1057. doi: 10.1007/s10096-016-2639-3.

Leffler, D. A. and Lamont, J. T. (2015) '*Clostridium difficile* Infection', *The New England Journal of Medicine*, 372, pp. 1539–1548.

Lessa, F. C. *et al.* (2015) 'Burden of *Clostridium difficile* Infection in the United States', *The New England Journal of Medicine*, 372(9), pp. 825–834.

Li, B., Qiu, W., Tan, Q. M., Su, T., Fang, Y. and Xie, G.-L. (2009a) 'Association of a bacillus species with leaf and twig dieback of Asian pear (*Pyrus Pyrifolia*) in China', *Journal of Plant Pathology*, 91. doi: 10.4454/jpp.v91i3.565.

Li, H. and Durbin, R. (2009) 'Fast and accurate short read alignment with Burrows-Wheeler transform.', *Bioinformatics (Oxford, England)*. England, 25(14), pp. 1754–1760. doi: 10.1093/bioinformatics/btp324.

Li, H., Handsaker, B., Wysoker, A., Fennell, T., Ruan, J., Homer, N., Marth, G., Abecasis, G. and Durbin, R. (2009b) 'The Sequence Alignment/Map format and SAMtools.', *Bioinformatics (Oxford, England)*. England, 25(16), pp. 2078–2079. doi: 10.1093/bioinformatics/btp352.

Li, H. and O'Sullivan, D. J. (2006) 'Identification of a nisI promoter within the nisABCTIP operon that may enable establishment of nisin immunity prior to induction of the operon via signal transduction.', *Journal of bacteriology*. United States, 188(24), pp. 8496–8503. doi: 10.1128/JB.00946-06.

Li, Y., Song, Z., Kerr, K. A. and Moeser, A. J. (2017) 'Chronic social stress in pigs impairs intestinal barrier and nutrient transporter function, and alters neuro-immune mediator and receptor expression', *PLOS ONE*. Public Library of Science, 12(2), p. e0171617. Available at: https://doi.org/10.1371/journal.pone.0171617.

Logan, N. A. (2012) 'Bacillus and relatives in foodborne illness', *Journal of Applied Microbiology*. John Wiley & Sons, Ltd (10.1111), 112(3), pp. 417–429. doi: 10.1111/j.1365-2672.2011.05204.x.

Loo, V. G. *et al.* (2011) 'Host and pathogen factors for *Clostridium difficile* infection and colonization.', *The New England journal of medicine*. United States, 365(18), pp. 1693–1703. doi: 10.1056/NEJMoa1012413.

Louie, T. J., Miller, M. A., Mullane, K. M., Weiss, K., Lentnek, A., Golan, Y., Gorbach, S., Sears, P. and Shue, Y.-K. (2011) 'Fidaxomicin versus Vancomycin for *Clostridium difficile* Infection', *New England Journal of Medicine*. Massachusetts Medical Society, 364(5), pp. 422–431. doi: 10.1056/NEJMoa0910812.

Louie, T. J., Nord, C. E., Talbot, G. H., Wilcox, M., Gerding, D. N., Buitrago, M., Kracker, H., Charef, P. and Cornely, O. A. (2015) 'Multicenter, Double-Blind, Randomized, Phase 2 Study Evaluating the Novel Antibiotic Cadazolid in Patients with *Clostridium difficile* Infection', *Antimicrobial Agents and Chemotherapy*, 59(10), pp. 6266 LP – 6273. doi: 10.1128/AAC.00504-15.

Lubelski, J., Rink, R., Khusainov, R., Moll, G. N. and Kuipers, O. P. (2008) 'Biosynthesis, immunity, regulation, mode of action and engineering of the model lantibiotic nisin.', *Cellular and molecular life sciences : CMLS*. Switzerland, 65(3), pp. 455–476. doi:

10.1007/s00018-007-7171-2.

Macke, T. J., Ecker, D. J., Gutell, R. R., Gautheret, D., Case, D. A. and Sampath, R. (2001) 'RNAMotif, an RNA secondary structure definition and search algorithm.', *Nucleic acids research*. England, 29(22), pp. 4724–4735. doi: 10.1093/nar/29.22.4724.

Maldonado, A., Ruiz-Barba, J. L. and Jiménez-Díaz, R. (2003) 'Purification and Genetic Characterization of Plantaricin NC8, a Novel Coculture-Inducible Two-Peptide Bacteriocin from *Lactobacillus plantarum* NC8', *Applied and Environmental Microbiology*, 69(1), pp. 383 LP – 389. doi: 10.1128/AEM.69.1.383-389.2003.

De Man, J. C., Rogosa, M. and Sharpe, M. E. (1960) 'A Medium for the Cultivation of Lactobacilli', *Journal of Applied Bacteriology*. John Wiley & Sons, Ltd (10.1111), 23(1), pp. 130–135. doi: 10.1111/j.1365-2672.1960.tb00188.x.

Mantovani, H. C. and Russell, J. B. (2002) 'The Ability of a Bacteriocin of *Streptococcus bovis* HC5 (bovicin HC5) to Inhibit *Clostridium aminophilum*, An Obligate Amino Acid Fermenting Bacterium from the Rumen', *Anaerobe*, 8(5), pp. 247–252. doi: https://doi.org/10.1006/anae.2002.0437.

Maqueda, M., Sánchez-Hidalgo, M., Fernández, M., Montalbán-López, M., Valdivia, E. and Martínez-Bueno, M. (2008) 'Genetic Features of Circular Bacteriocins Produced by Gram-Positive Bacteria', *FEMS Microbiology Reviews*, (32), pp. 2–22.

Marciset, O., Jeronimus, M., Mollet, B. and Poolman, B. (1997) 'Thermophilin 13, a Nontypical Antilisterial Poration Complex Bacteriocin, That Functions without a Receptor', *The Journal of biological chemistry*, 272, pp. 14277–14284. doi: 10.1074/jbc.272.22.14277.

Marsh, A. J., O'Sullivan, O., Ross, R. P., Cotter, P. D. and Hill, C. (2010) '*In silico* analysis highlights the frequency and diversity of type 1 lantibiotic gene clusters in genome sequenced bacteria.', *BMC genomics*. England, 11, p. 679. doi: 10.1186/1471-2164-11-679.

Martin-Visscher, L. A., van Belkum, M. J., Garneau-Tsodikova, S., Whittal, R. M., Zheng, J., McMullen, L. M. and Vederas, J. C. (2008) 'Isolation and characterization of carnocyclin a, a novel circular bacteriocin produced by *Carnobacterium maltaromaticum* UAL307.', *Applied and environmental microbiology*. United States, 74(15), pp. 4756–4763. doi: 10.1128/AEM.00817-08.

Martinez-Bueno, M., Maqueda, M., Galvez, A., Samyn, B., Van Beeumen, J., Coyette, J. and Valdivia, E. (1994) 'Determination of the gene sequence and the molecular structure of the enterococcal peptide antibiotic AS-48.', *Journal of bacteriology*. United States, 176(20), pp. 6334–6339. doi: 10.1128/jb.176.20.6334-6339.1994.

Masuda, Y., Ono, H., Kitagawa, H., Ito, H., Mu, F., Sawa, N., Zendo, T. and Sonomoto, K. (2011) 'Identification and characterization of leucocyclicin Q, a novel cyclic bacteriocin produced by *Leuconostoc mesenteroides* TK41401.', *Applied and environmental microbiology*. United States, 77(22), pp. 8164–8170. doi: 10.1128/AEM.06348-11.

Mathur, H., Fallico, V., O'Connor, P. M., Rea, M. C., Cotter, P. D., Hill, C. and Ross, R. P. (2017) 'Insights into the mode of action of the sactibiotic thuricin CD', *Frontiers in Microbiology*. Frontiers Research Foundation, 8(APR). doi: 10.3389/fmicb.2017.00696.

Mathur, H., O'Connor, P. M., Cotter, P. D., Hill, C. and Ross, R. P. (2014) 'Heterologous Expression of Thuricin CD Immunity Genes in *Listeria monocytogenes*', *Antimicrobial Agents and Chemotherapy*, 58(6), pp. 3421–3428.

Mathur, H., Rea, M. C., Cotter, P., Hill, C. and Ross, R. P. (2016) 'The Efficacy of Thuricin CD, Tigecycline, Vancomycin, Teicoplanin, Rifampicin and Nitazoxanide, Independently and in Paired Combinations Against *Clostridium difficile* Biofilms and Planktonic Cells', *Gut Pathogens*, 8(20).

Maurício, E., Rosado, C., Duarte, P. M., Verissimo, J., Bom, S. and Vasconcelos, L. (2017) 'Efficiency of Nisin as Preservative in Cosmetics and Topical Products', *Cosmetics* . doi: 10.3390/cosmetics4040041.

McAuliffe, O., Ryan, M. P., Ross, R. P., Hill, C., Breeuwer, P. and Abee, T. (1998) 'Lacticin 3147, a broad-spectrum bacteriocin which selectively dissipates the membrane potential.', *Applied and environmental microbiology*. United States, 64(2), pp. 439–445.

McDonald, L. C. *et al.* (2018) 'Clinical Practice Guidelines for Clostridium difficile Infection in Adults and Children: 2017 Update by the Infectious Diseases Society of America (IDSA) and Society for Healthcare Epidemiology of America (SHEA)', *Clinical Infectious Diseases*, 66(7), pp. e1–e48. doi: 10.1093/cid/cix1085.

McGovern, A. M., Androga, G. O., Knight, D. R., Watson, M. W., Elliott, B., Foster, N. F., Chang, B. J. and Riley, T. V. (2017) 'Prevalence of binary toxin positive *Clostridium difficile* in diarrhoeal humans in the absence of epidemic ribotype 027', *PLoS ONE*. Public Library of Science, 12(11). doi: 10.1371/journal.pone.0187658.

van der Meer, J. R., Polman, J., Beerthuyzen, M. M., Siezen, R. J., Kuipers, O. P. and De Vos, W. M. (1993) 'Characterization Of The Lactococcus lactis Nisin A Operon Genes nisP, Encoding A Subtilisin-like Serine Protease Involved In Precursor Processing, and nisR, Encoding A Regulatory Protein Involved In Nisin Biosynthesis', *Journal of bacteriology*, 175(9), pp. 2578–2588.

de Mello, M., Malheiros, P., Brandelli, A., Silveira, N., Jantzen, M. and Motta, A. (2013) 'Characterization and Antilisterial Effect of Phosphatidylcholine Nanovesicles Containing the Antimicrobial Peptide Pediocin', *Probiotics and Antimicrobial Proteins*, 5. doi: 10.1007/s12602-013-9125-3.

Mergulhão, F. and Monteiro, G. A. (2004) *Design of bacterial vector systems for the production of recombinant proteins in Escherichia coli Engineered MRI nanoprobes based on superparamagnetic iron oxide nanoparticles View project Drug delivery using electrospun fibers View project*, *Article in Journal of Microbiology and Biotechnology*. Available at: https://www.researchgate.net/publication/273132208 (Accessed: 14 October 2019).

Mesa-Pereira, B., O'Connor, P. M., Rea, M. C., Cotter, P. D., Hill, C. and Ross, R. P. (2017) 'Controlled functional expression of the bacteriocins pediocin PA-1 and bactofencin A in *Escherichia coli*.', *Scientific reports*, 7(1), p. 3069. doi: 10.1038/s41598-017-02868-w.

Mesa-Pereira, B., Rea, M. C., Cotter, P. D., Hill, C. and Ross, R. P. (2018) 'Heterologous Expression of Biopreservative Bacteriocins With a View to Low Cost Production', *Frontiers in microbiology*. Frontiers Media S.A., 9, p. 1654. doi: 10.3389/fmicb.2018.01654.

Michel-Briand, Y. and Baysse, C. (2002) 'The pyocins of *Pseudomonas aeruginosa*', *Biochimie*, pp. 499–510. doi: 10.1016/S0300-9084(02)01422-0.

Mills, S., Stanton, C., Hill, C. and Ross, R. P. (2011) 'New Developments and Applications of Bacteriocins and Peptides in Foods', *Annual Review of Food Science and Technology*. Annual Reviews, 2(1), pp. 299–329. doi: 10.1146/annurev-food-022510-133721.

Ming, X. and Daeschel, M. A. (1995) 'Correlation of Cellular Phospholipid Content with Nisin Resistance of *Listeria monocytogenes* Scott A', *Journal of Food Protection*. International Association for Food Protection, 58(4), pp. 416–420. doi: 10.4315/0362-028X-58.4.416.

Moll, G., Ubbink-Kok, T., Hildeng-Hauge, H., Nissen-Meyer, J., Nes, I. F., Konings, W. N. and Driessen, A. J. (1996) 'Lactococcin G is a potassium ion-conducting, two-component bacteriocin.', *Journal of Bacteriology*, 178(3), pp. 600 LP – 605. doi: 10.1128/jb.178.3.600-605.1996.

Al Momani, L. A., Abughanimeh, O., Boonpheng, B., Gabriel, J. G. and Young, M. (2018) 'Fidaxomicin vs Vancomycin for the Treatment of a First Episode of *Clostridium difficile* Infection: A Meta-analysis and Systematic Review', *Cureus*. Cureus, 10(6), pp. e2778–e2778. doi: 10.7759/cureus.2778.

Monaghan, T. M., Cockayne, A. and Mahida, Y. R. (2015) 'Pathogenesis of *Clostridium difficile* Infection and Its Potential Role in Inflammatory Bowel Disease', *Inflammatory Bowel Diseases*, 21(8), pp. 1957–1966.

Morgan, S. M., Galvin, M., Ross, R. P. and Hill, C. (2001) 'Evaluation of a spray-dried lacticin 3147 powder for the control of Listeria monocytogenes and *Bacillus cereus* in a range of food systems', *Letters in Applied Microbiology*. John Wiley & Sons, Ltd (10.1111), 33(5), pp. 387–391. doi: 10.1046/j.1472-765X.2001.01016.x.

Morovský, M., Pristaš, P., Javorský, P., Nes, I. F. and Holo, H. (2001) 'Isolation and characterization of enterocin BC25 and occurrence of the ent A gene among ruminal Gram-positive cocci', *Microbiological Research*, 156(2), pp. 133–138. doi: https://doi.org/10.1078/0944-5013-00090.

Mulders, J. W., Boerrigter, I. J., Rollema, H. S., Siezen, R. J. and de Vos, W. M. (1991) 'Identification and Characterization of The Lantibiotic Nisin Z, A Natural Nisin Variant', *European Journal of Biochemistry*. 1991/11/01. Department of Biophysical Chemistry, Netherlands Institute for Dairy Research (NIZO), Ede., 201(3), pp. 581–584.

Müller, E. and Radler, F. (1993) 'Caseicin, a Bacteriocin from *Lactobacillus casei*', *Folia Microbiologica*, 38(6), pp. 441–446.

Münch, D. and Sahl, H.-G. (2015) 'Structural variations of the cell wall precursor lipid II in Gram-positive bacteria — Impact on binding and efficacy of antimicrobial peptides', *Biochimica et Biophysica Acta (BBA) - Biomembranes*, 1848(11, Part B), pp. 3062–3071. doi: https://doi.org/10.1016/j.bbamem.2015.04.014.

Musher, D. M., Logan, N., Hamill, R. J., DuPont, H. L., Lentnek, A., Gupta, A. and Rossignol, J.-F. (2006) 'Nitazoxanide for the Treatment of *Clostridium difficile* Colitis', *Clinical Infectious Diseases*, 43(4), pp. 421–427. doi: 10.1086/506351.

Myresiotis, C. K., Vryzas, Z. and Papadopoulou-Mourkidou, E. (2014) 'Enhanced root

uptake of acibenzolar-S-methyl (ASM) by tomato plants inoculated with selected Bacillus plant growth-promoting rhizobacteria (PGPR)', *Applied Soil Ecology*, 77, pp. 26–33. doi: https://doi.org/10.1016/j.apsoil.2014.01.005.

Na, X., Kim, H. U., Moyer, M. P., Pothoulakis, C. and LaMont, J. T. (2008) 'gp96 is a Human Colonocyte Plasma Membrane Binding Protein for *Clostridium difficile*Toxin A', *Infection and Immunity*, 76(7), pp. 2862–2871.

Naville, M., Ghuillot-Gaudeffroy, A., Marchais, A. and Gautheret, D. (2011) 'ARNold: A web tool for the prediction of Rho-independent transcription terminators', *RNA Biology*. Taylor & Francis, 8(1), pp. 11–13. doi: 10.4161/rna.8.1.13346.

Ndlovu, B., Schoeman, H., Franz, C. M. A. P. and Du Toit, M. (2015) 'Screening, identification and characterization of bacteriocins produced by wine-isolated LAB strains'. doi: 10.1111/jam.12752.

Nes, I. F., Diep, D. B. and Holo, H. (2007) 'Bacteriocin Diversity in Streptococcus and Enterococcus', *Journal of Bacteriology*, 189(4), pp. 1189–1198.

Nes, I. F. and Holo, H. (2000) 'Class II Antimicrobial peptides from lactic acid bacteria', *Biopolymers*, 55, pp. 50–61. doi: 10.1002/1097-0282(2000)55:1<50::AID-BIP50>3.0.CO;2-3.

Ng, Y. K., Ehsaan, M., Philip, S., Collery, M. M., Janoir, C., Collignon, A., Cartman, S. T. and Minton, N. P. (2013) 'Expanding the repertoire of gene tools for precise manipulation of the *Clostridium difficile* genome: allelic exchange using pyrE alleles', *PLoS One*, 8(2), p. e56051. doi: 10.1371/journal.pone.0056051.

Nicholson, W. L. (2002) 'Roles of Bacillus endospores in the environment.', *Cellular and molecular life sciences : CMLS*. Switzerland, 59(3), pp. 410–416.

Nilsen, T., Nes, I. F. and Holo, H. (2003) 'Enterolysin A, a cell wall-degrading bacteriocin from *Enterococcus faecalis* LMG 2333', *Applied and Environmental Microbiology*, 69(5), pp. 2975–2984. Available at: internal-pdf://244.176.220.81/Nilsen EnterolysinA.pdf.

Nissen-Meyer, J., Holo, H., Havarstein, L. S., Sletten, K. and Nes, I. F. (1992) 'A novel lactococcal bacteriocin whose activity depends on the complementary action of two peptides.', *Journal of bacteriology*. United States, 174(17), pp. 5686–5692. doi: 10.1128/jb.174.17.5686-5692.1992.

Nocek, B. *et al.* (2012) 'Structural and Functional Characterisation of Microcin C Resistance Peptidase MccF from *Bacillus anthracis*', *Journal of Molecular Biology*, 420, pp. 366–383.

Norén, T., Wullt, M., Åkerlund, T., Bäck, E., Odenholt, I. and Burman, L. G. (2006) 'Frequent Emergence of Resistance in Clostridium difficile during Treatment of C. difficile-Associated Diarrhea with Fusidic Acid', *Antimicrobial Agents and Chemotherapy*, 50(9), pp. 3028 LP – 3032. doi: 10.1128/AAC.00019-06.

Nwokoro, O. and Dibua, M. E. U. (2014) 'Degradation of soil cyanide by single and mixed cultures of *Pseudomonas stutzeri* and *Bacillus subtilis*', *Archives of Industrial Hygiene and Toxicology*, 65(1), pp. 113–119. doi: https://doi.org/10.2478/10004-1254-65-2014-2449.

O'Connor, P. M., O'Shea, E. F., Guinane, C. M., O'Sullivan, O., Cotter, P. D., Ross, R. P. and Hill, C. (2015) 'Nisin H Is a New Nisin Variant Produced by the Gut-Derived Strain

Streptococcus hyointestinalis DPC6484.', *Applied and environmental microbiology*. United States, 81(12), pp. 3953–3960. doi: 10.1128/AEM.00212-15.

O'Sullivan, L., Morgan, S. M., Ross, R. P. and Hill, C. (2002) 'Elevated enzyme release from lactococcal starter cultures on exposure to the lantibiotic lacticin 481, produced by *Lactococcus lactis* DPC5552.', *Journal of dairy science*. United States, 85(9), pp. 2130–2140. doi: 10.3168/jds.S0022-0302(02)74291-4.

Odenyo, A. A., Mackie, R. I., Stahl, D. A. and White, B. A. (1994) 'The use of 16S rRNA-targeted oligonucleotide probes to study competition between ruminal fibrolytic bacteria: development of probes for Ruminococcus species and evidence for bacteriocin production.', *Applied and Environmental Microbiology*, 60(10), pp. 3688 LP – 3696. Available at: http://aem.asm.org/content/60/10/3688.abstract.

Office for National Statistics (2013) *Deaths Involving Clostridium difficile, England and Wales - Office for National Statistics*. Available at: https://www.ons.gov.uk/peoplepopulationandcommunity/birthsdeathsandmarriages/deaths/bulletins/deathsinvolvingclostridiumdifficileenglandandwales/2013-08-22 (Accessed: 21 October 2019).

Oguri, N. *et al.* (2019) 'Community-acquired fulminant colitis caused by binary toxin-producing *Clostridium difficile* in Japan', *Clinical Journal of Gastroenterology*. Springer Tokyo. doi: 10.1007/s12328-019-00949-z.

Ollech, J. E., Shen, N. T., Crawford, C. V and Ringel, Y. (2015) 'Use of Probiotics in Prevention and Treatment of Patients with *Clostridium difficile* Infection', *Best Practice & Research Clinical Gastroenterology*, 30, pp. 111–118.

Orengo, C. A., Michie, A. D., Jones, S., Jones, D. T., Swindells, M. B. and Thornton, J. M. (1997) 'CATH--a hierarchic classification of protein domain structures.', *Structure (London, England : 1993)*. United States, 5(8), pp. 1093–1108.

Oumer, A., Garde, S., Gaya, P., Medina, M. and Nunez, M. (2001) 'The Effects of Cultivating Lactic Starter Cultures with Bacteriocin-Producing Lactic Acid Bacteria', *Journal of Food Protection*. International Association for Food Protection, 64(1), pp. 81–86. doi: 10.4315/0362-028X-64.1.81.

Overbeek, R. *et al.* (2014) 'The SEED and the Rapid Annotation of microbial genomes using Subsystems Technology (RAST)', *Nucleic Acids Research*, 42(Database Issue), pp. D206–D214.

Pant, C., Deshpande, A., Altaf, M. A., Minocha, A. and Sferra, T. J. (2013) '*Clostridium difficile* infection in children: a comprehensive review.', *Current medical research and opinion*. England, 29(8), pp. 967–984. doi: 10.1185/03007995.2013.803058.

Papagianni, M. and Anastasiadou, S. (2009) 'Pediocins: The Bacteriocins of Pediococci. Sources, Production, Properties and Applications', *Microbial Cell Factories*, 8(3).

Papagianni, M., Avramidis, N. and Filioussis, G. (2007) 'High Efficiency Electrotransformation of *Lactococcus lactis* spp. *lactis* Cells Pretreated with Lithium Acetate and Dithiothreitol', *BMC Biotechnology*, 7(15).

Parmar, N. V and Glauser, J. (2019) 'Systematic Review of Current Treatment and Prevention Strategies for *Clostridium difficile*', *Current Emergency and Hospital Medicine Reports*, 7(3), pp. 66–73. doi: 10.1007/s40138-019-00186-1.

Patzer, S. I., Albrecht, R., Braun, V. and Zeth, K. (2012) 'Structural and mechanistic studies of pesticin, a bacterial homolog of phage lysozymes', *Journal of Biological Chemistry*, 287(28), pp. 23381–23396. doi: 10.1074/jbc.M112.362913.

Peltola, J., Andersson, M. A., Haahtela, T., Mussalo-Rauhamaa, H., Rainey, F. A., Kroppenstedt, R. M., Samson, R. A. and Salkinoja-Salonen, M. S. (2001) 'Toxic-Metabolite-Producing Bacteria and Fungus in an Indoor Environment', *Applied and Environmental Microbiology*, 67(7), pp. 3269 LP – 3274. doi: 10.1128/AEM.67.7.3269-3274.2001.

Pépin, J., Alary, M.-E., Valiquette, L., Raiche, E., Ruel, J., Fulop, K., Godin, D. and Bourassa, C. (2005) 'Increasing Risk of Relapse after Treatment of *Clostridium difficile* Colitis in Quebec, Canada', *Clinical Infectious Diseases*, 40(11), pp. 1591–1597. doi: 10.1086/430315.

Pérez-Cobas, A. E., Moya, A., Gosalbes, M. J., Latorre, A. and Chang, Y.-F. (2015) 'Colonisation Resistance of Gut Microbiota Against *Clostridium difficile*', *Antibiotics*, 4(3), pp. 337–357.

Perez, R. H., Zendo, T. and Sonomoto, K. (2018) 'Circular and Leaderless Bacteriocins: Biosynthesis, Mode of Action, Applications, and Prospects', *Frontiers in microbiology*. Frontiers Media S.A., 9, p. 2085. doi: 10.3389/fmicb.2018.02085.

Peschel, A. and Gotz, F. (1996) 'Analysis of the *Staphylococcus epidermidis* genes epiF, -E, and -G involved in epidermin immunity.', *Journal of bacteriology*. United States, 178(2), pp. 531–536. doi: 10.1128/jb.178.2.531-536.1996.

Petrella, L. A., Sambol, S. P., Cheknis, A., Nagaro, K., Kean, Y., Sears, P. S., Babakhani, F., Johnson, S. and Gerding, D. N. (2012) 'Decreased Cure and Increased Recurrence Rate for *Clostridium difficile* Infections Caused by the Epidemic *C. difficile* BI Strain', *Clinical Infectious Disease*, 55(3), pp. 351–357.

Petrosillo, N., Granata, G. and Cataldo, M. A. (2018) 'Novel Antimicrobials for the Treatment of *Clostridium difficile* Infection', *Frontiers in Medicine* , p. 96. Available at: https://www.frontiersin.org/article/10.3389/fmed.2018.00096.

Peypoux, F., Bonmatin, J. M. and Wallach, J. (1999) 'Recent trends in the biochemistry of surfactin', *Applied Microbiology and Biotechnology*, 51(5), pp. 553–563. doi: 10.1007/s002530051432.

Pickering, D. S. (2015) *Characterisation of Bacteriocins Active Against Clinically Relevant Strains of Clostridium difficile*. University of Nottingham.

Pieterse, R., Todorov, S. D. and Dicks, L. M. T. (2008) 'Bacteriocin ST91KM, Produced by *Streptococcus gallolyticus* subsp. *macedonicus* ST91KM, is a Narrow-Spectrum Peptide Active Against Bacteria Associated with Mastitis in Dairy Cattle', *Canadian Journal of Microbiology*, 54(7), pp. 525–531.

Pieterse, R., Todorov, S. D. and Leon M T, D. (2010) 'Mode of action and In Vitro susceptibility of mastitis pathogens to macedocin ST91KM and preparation of a teat seal containing the bacteriocin', *Brazilian journal of microbiology : [publication of the Brazilian Society for Microbiology]*. 2010/03/01. Sociedade Brasileira de Microbiologia, 41(1), pp. 133–145. doi: 10.1590/S1517-838220100001000020.

Pimentel-Filho, N. de J., Mantovani, H. C., de Carvalho, A. F., Dias, R. S. and Vanetti, M. C. D. (2014) 'Efficacy of bovicin HC5 and nisin combination against *Listeria*

monocytogenes and *Staphylococcus aureus* in fresh cheese', *International Journal of Food Science & Technology*. John Wiley & Sons, Ltd (10.1111), 49(2), pp. 416–422. doi: 10.1111/ijfs.12316.

Pirs, T., Ocepek, M. and Rupnik, M. (2008) 'Isolation of *Clostridium difficile* from food animals in Slovenia.', *Journal of medical microbiology*. England, 57(Pt 6), pp. 790–792. doi: 10.1099/jmm.0.47669-0.

Potter, S. C., Luciani, A., Eddy, S. R., Park, Y., Lopez, R. and Finn, R. D. (2018) 'HMMER web server: 2018 update', *Nucleic Acids Research*, 46(W1), pp. W200–W204. doi: 10.1093/nar/gky448.

Prazmowski, A. (1880) 'Untersuchung über die Entwickelungsgeschichte und Fermentwirking einiger Bacterien-Arten Hugo Voigt, Leipzig'.

Purdy, D., O'Keeffe, T. A. T., Elmore, M., Herbert, M., McLeod, A., Bokori-Brown, M., Ostrowski, A. and Minton, N. P. (2002) 'Conjugative transfer of clostridial shuttle vectors from *Escherichia coli* to *Clostridium difficile* through circumvention of the restriction barrier.', *Molecular microbiology*. England, 46(2), pp. 439–452.

Qa'Dan, M., Ramsey, M., Daniel, J., Spyres, L. M., Safiejko-Mroczka, B., Ortiz-Leduc, W. and Ballard, J. D. (2002) '*Clostridium difficile* Toxin B Activates Dual Caspase-Dependent and Caspase-Independent Apoptosis in Intoxicated Cells', *Cellular Microbiology*, 4(7), pp. 425–434.

Qiao, M., Ye, S., Koponen, O., Ra, R., Usabiaga, M., Immonen, T. and Saris, P. E. (1996) 'Regulation of the nisin operons in *Lactococcus lactis* N8.', *The Journal of applied bacteriology*. England, 80(6), pp. 626–634.

Rafiullah, F., Kanwal, S., Majeed, U. M., Korsten, M. A., Cheema, F. H., Luthra, M. and Sohail, M. R. (2011) 'Successful use of nitazoxanide in the treatment of recurrent *Clostridium difficile* infection', *BMJ Case Reports*, 2011, p. bcr0420114123. doi: 10.1136/bcr.04.2011.4123.

Rawlings, N. D. and Barrett, A. J. (1995) 'Evolutionary families of metallopeptidases', *Methods in enzymology*. United States, 248, pp. 183–228. doi: 10.1016/0076-6879(95)48015-3.

Rawlings, N. D. and Salvesen, G. (2013) 'Lysostaphin', in *Handbook of Proteolytic Enzymes*. London: Elsevier, pp. 1558–1560.

Rayman, M. K., Aris, B. and Hurst, A. (1981) 'Nisin: a possible alternative or adjunct to nitrite in the preservation of meats', *Applied and environmental microbiology*, 41(2), pp. 375–380. Available at: https://www.ncbi.nlm.nih.gov/pubmed/7195188.

Rea, M. C., Alemayehu, D., Ross, R. P. and Hill, C. (2013) 'Gut Solutions To A Gut Problem: Bacteriocins, Probiotics and Bacteriophage For Control Of *Clostridium difficile* Infection', *Journal of Medical Microbiology*, 62, pp. 1369–1378.

Rea, M. C., Clayton, E., O'Connor, P. M., Shanahan, F., Kiely, B., Ross, R. P. and Hill, C. (2007) 'Antimicrobial Activity of Lacticin 3,147 Against Clinical *Clostridium difficile* Strains', *Journal of Medical Microbiology*, 56(Pt 7), pp. 940–946.

Rea, M. C., Sit, C. S., Clayton, E., O'Connor, P. M., Whittal, R. M., Zheng, J., Vederas, J. C., Ross, R. P. and Hill, C. (2010) 'Thuricin CD, A Posttranslationally Modified Bacteriocin With A Narrow Spectrum Of Activity Against *Clostridium difficile*', *Proceedings of The National Academy of Sciences of The United States of America*,

107(20), pp. 9352–9357.

Rebuffat, S. (2011) 'Bacteriocins from Gram-Negative Bacteria: A Classification?', in *Prokaryotic Antimicrobial Peptides*. Springer New York, pp. 55–72. doi: 10.1007/978-1-4419-7692-5_4.

Riley, M. A. and Wertz, J. E. (2002) 'Bacteriocin Diversity: Ecological and Evolutionary Perspectives', *Biochimie*, 84(5–6), pp. 357–364.

Rodriguez-Palacios, A., Pickworth, C., Loerch, S. and LeJeune, J. T. (2011) 'Transient fecal shedding and limited animal-to-animal transmission of *Clostridium difficile* by naturally infected finishing feedlot cattle', *Applied and environmental microbiology*, 77(10), pp. 3391–3397. doi: 10.1128/AEM.02736-10.

Rodriguez-Palacios, A., Staempfli, H. R., Duffield, T. and Weese, S. J. (2007) *'Clostridium difficile* in Retail Ground Meat, Canada', *Emerging Infectious Diseases*, 13(3), pp. 485–487.

Rodriguez-Palacios, A., Stampfli, H. R., Duffield, T., Peregrine, A. S., Trotz-Williams, L. A., Arroyo, L. G., Brazier, J. S. and Weese, J. S. (2006) '*Clostridium difficile* PCR ribotypes in calves, Canada.', *Emerging infectious diseases*, 12(11), pp. 1730–1736. doi: 10.3201/eid1211.051581.

Rodriguez, C., Avesani, V., Van Broeck, J., Taminiau, B., Delmee, M. and Daube, G. (2013) 'Presence of *Clostridium difficile* in pigs and cattle intestinal contents and carcass contamination at the slaughterhouse in Belgium.', *International journal of food microbiology*. Netherlands, 166(2), pp. 256–262. doi: 10.1016/j.ijfoodmicro.2013.07.017.

Rodriguez, C., Hakimi, D.-E., Vanleyssem, R., Taminiau, B., Van Broeck, J., Delmée, M., Korsak, N. and Daube, G. (2017) '*Clostridium difficile* in beef cattle farms, farmers and their environment: Assessing the spread of the bacterium', *Veterinary Microbiology*, 210, pp. 183–187. doi: https://doi.org/10.1016/j.vetmic.2017.09.010.

Rodriguez, C., Taminiau, B., Van Broeck, J., Avesani, V., Delmee, M. and Daube, G. (2012) '*Clostridium difficile* in young farm animals and slaughter animals in Belgium.', *Anaerobe*. England, 18(6), pp. 621–625. doi: 10.1016/j.anaerobe.2012.09.008.

Rodriguez Diaz, C., Seyboldt, C. and Rupnik, M. (2018) 'Non-human *C. difficile* Reservoirs and Sources: Animals, Food, Environment.', *Advances in experimental medicine and biology*. United States, 1050, pp. 227–243. doi: 10.1007/978-3-319-72799-8_13.

Rogers, L. A. and Whittier, E. O. (1928) 'Limiting Factors in the Lactis Fermentation', *Journal of bacteriology*. United States, 16(4), pp. 211–229.

Romano, V., Albanese, F., Dumontet, S., Krovacek, K., Petrini, O. and Pasquale, V. (2012) 'Prevalence and genotypic characterization of *Clostridium difficile* from ruminants in Switzerland.', *Zoonoses and public health*. Germany, 59(8), pp. 545–548. doi: 10.1111/j.1863-2378.2012.01540.x.

Rosano, G. L. and Ceccarelli, E. A. (2014) 'Recombinant protein expression in *Escherichia coli*: Advances and challenges', *Frontiers in Microbiology*. Frontiers Research Foundation. doi: 10.3389/fmicb.2014.00172.

Ross, K. F., Ronson, C. W. and Tagg, J. R. (1993) 'Isolation and characterization of the lantibiotic salivaricin A and its structural gene salA from *Streptococcus salivarius*

20P3.', *Applied and environmental microbiology*. United States, 59(7), pp. 2014–2021.

Rossi-Tamisier, M., Benamar, S., Raoult, D. and Fournier, P.-E. (2015) 'Cautionary tale of using 16S rRNA gene sequence similarity values in identification of human-associated bacterial species', *International journal of systematic and evolutionary microbiology*, 65. doi: 10.1099/ijs.0.000161.

Rupnik, M. (2007) 'Is *Clostridium difficile*-associated infection a potentially zoonotic and foodborne disease?', *Clinical Microbiology and Infection*. John Wiley & Sons, Ltd (10.1111), 13(5), pp. 457–459. doi: 10.1111/j.1469-0691.2007.01687.x.

Rupnik, M., Grabnar, M. and Geric, B. (2003) 'Binary toxin producing *Clostridium difficile* strains.', *Anaerobe*. England, 9(6), pp. 289–294. doi: 10.1016/j.anaerobe.2003.09.002.

Rupnik, M., Wilcox, M. H. and Gerding, D. N. (2009) '*Clostridium difficile* infection: new developments in epidemiology and pathogenesis.', *Nature reviews. Microbiology*. England, 7(7), pp. 526–536. doi: 10.1038/nrmicro2164.

Russell, J. B. and Mantovani, H. C. (2002) 'The bacteriocins of ruminal bacteria and their potential as an alternative to antibiotics.', *Journal of molecular microbiology and biotechnology*. Switzerland, 4(4), pp. 347–355.

de Ruyter, P. G., Kuipers, O. P., Beerthuyzen, M. M., van Alen-Boerrigter, I. and de Vos, W. M. (1996) 'Functional analysis of promoters in the nisin gene cluster of *Lactococcus lactis*', *Journal of bacteriology*, 178(12), pp. 3434–3439. doi: 10.1128/jb.178.12.3434-3439.1996.

Ryan, M. P., Rea, M. C., Hill, C. and Ross, R. P. (1996) 'An Application in Cheddar Cheese Manufacture for a Strain of *Lactoccus lactis* Producing a Novel Broad-Spectrum Bacteriocin, Lacticin 3147', *Applied and Environmental Microbiology*, 62(2), pp. 612–619.

Saavedra, L. and Sesma, F. (2011) 'Purification Techniques of Bacteriocins from Lactic Acid Bacteria and Other Gram-Positive Bacteria', in Drider, D. and Rebuffat, S. (eds) *Prokaryotic Antimicrobial Peptides: From Genes to Applications*. Springer, pp. 99–113. doi: 10.1007/978-1-4419-7692-5_7.

Sahl, H.-G. and Bierbaum, G. (1998) 'LANTIBIOTICS: Biosynthesis and Biological Activities of Uniquely Modified Peptides from Gram-Positive Bacteria', *Annual Review of Microbiology*. Annual Reviews, 52(1), pp. 41–79. doi: 10.1146/annurev.micro.52.1.41.

Sanchez-Barrena, M. J., Martinez-Ripoll, M., Galvez, A., Valdivia, E., Maqueda, M., Cruz, V. and Albert, A. (2003) 'Structure of bacteriocin AS-48: from soluble state to membrane bound state.', *Journal of molecular biology*. England, 334(3), pp. 541–549. doi: 10.1016/j.jmb.2003.09.060.

Sánchez-González, L., Quintero Saavedra, J. I. and Chiralt, A. (2013) 'Physical properties and antilisterial activity of bioactive edible films containing Lactobacillus plantarum', *Food Hydrocolloids*, 33(1), pp. 92–98. doi: https://doi.org/10.1016/j.foodhyd.2013.02.011.

Sandiford, S. and Upton, M. (2012) 'Identification, Characterization, and Recombinant Expression of Epidermicin NI01, a Novel Unmodified Bacteriocin

Produced by *Staphylococcus epidermidis*', *Antimicrobial Agents and Chemotherapy*, 56(3), pp. 1539 LP – 1547. doi: 10.1128/AAC.05397-11.

Santos, J. C. P., Sousa, R. C. S., Otoni, C. G., Moraes, A. R. F., Souza, V. G. L., Medeiros, E. A. A., Espitia, P. J. P., Pires, A. C. S., Coimbra, J. S. R. and Soares, N. F. F. (2018) 'Nisin and other antimicrobial peptides: Production, mechanisms of action, and application in active food packaging', *Innovative Food Science & Emerging Technologies*, 48, pp. 179–194. doi: https://doi.org/10.1016/j.ifset.2018.06.008.

Saranya, P., Sukanya Kumari, H., Prasad Rao, B. and Sekaran, G. (2014) 'Lipase production from a novel thermo-tolerant and extreme acidophile *Bacillus pumilus* using palm oil as the substrate and treatment of palm oil-containing wastewater', *Environmental Science and Pollution Research*, 21(5), pp. 3907–3919. doi: 10.1007/s11356-013-2354-x.

Sawa, N., Zendo, T., Kiyofuji, J., Fujita, K., Himeno, K., Nakayama, J. and Sonomoto, K. (2009) 'Identification and characterization of lactocyclicin Q, a novel cyclic bacteriocin produced by *Lactococcus* sp. strain QU 12.', *Applied and environmental microbiology*. United States, 75(6), pp. 1552–1558. doi: 10.1128/AEM.02299-08.

Schallmey, M., Singh, A. and Ward, O. P. (2004) 'Developments in the use of *Bacillus* species for industrial production', *Canadian Journal of Microbiology*. NRC Research Press, 50(1), pp. 1–17. doi: 10.1139/w03-076.

Schmid, A., Messelhausser, U., Hormansdorfer, S., Sauter-Louis, C. and Mansfeld, R. (2013) 'Occurrence of zoonotic clostridia and *Yersinia* in healthy cattle.', *Journal of food protection*. United States, 76(10), pp. 1697–1703. doi: 10.4315/0362-028X.JFP-13-151.

Schneeberg, A., Neubauer, H., Schmoock, G., Grossmann, E. and Seyboldt, C. (2013) 'Presence of *Clostridium difficile* PCR ribotype clusters related to 033, 078 and 045 in diarrhoeic calves in Germany.', *Journal of medical microbiology*. England, 62(Pt 8), pp. 1190–1198. doi: 10.1099/jmm.0.056473-0.

Schneewind, O., Fowler, A. and Faull, K. F. (1995) 'Structure of the cell wall anchor of surface proteins in *Staphylococcus aureus*.', *Science (New York, N.Y.)*. United States, 268(5207), pp. 103–106.

Schwartz, R., Ting, C. S. and King, J. (2001) 'Whole proteome pI values correlate with subcellular localizations of proteins for organisms within the three domains of life.', *Genome research*. United States, 11(5), pp. 703–709. doi: 10.1101/gr.158701.

Shanks, O. C., Kelty, C. A., Archibeque, S., Jenkins, M., Newton, R. J., McLellan, S. L., Huse, S. M. and Sogin, M. L. (2011) 'Community Structures of Fecal Bacteria in Cattle from Different Animal Feeding Operations', *Applied and Environmental Microbiology*, 77(9), pp. 2992 LP – 3001. doi: 10.1128/AEM.02988-10.

Sheitoyan-Pesant, C., Abou Chakra, C. N., Pépin, J., Marcil-Héguy, A., Nault, V. and Valiquette, L. (2015) 'Clinical and Healthcare Burden of Multiple Recurrences of *Clostridium difficile* Infection', *Clinical Infectious Diseases*, 62(5), pp. 574–580. doi: 10.1093/cid/civ958.

Sheppard, J. D., Jumarie, C., Cooper, D. G. and Laprade, R. (1991) 'Ionic channels induced by surfactin in planar lipid bilayer membranes', *Biochimica et Biophysica Acta (BBA) - Biomembranes*, 1064(1), pp. 13–23. doi: https://doi.org/10.1016/0005-2736(91)90406-X.

Shi, Y., Yang, X., Garg, N. and Van Der Donk, W. A. (2011) 'Production of lantipeptides in *Escherichia coli*', *Journal of the American Chemical Society*, 133(8), pp. 2338–2341. doi: 10.1021/ja109044r.

Shields, K., Araujo-Castillo, R. V, Theethira, T. G., Alonso, C. D. and Kelly, C. P. (2015) 'Recurrent *Clostridium difficile* Infection: From Colonization To Cure', *Anaerobe*, 34, pp. 59–73.

Shin, J. M., Gwak, J. W., Kamarajan, P., Fenno, J. C., Rickard, A. H. and Kapila, Y. L. (2016) 'Biomedical applications of nisin', *Journal of applied microbiology*. 2016/02/12, 120(6), pp. 1449–1465. doi: 10.1111/jam.13033.

Siegers, K. and Entian, K. D. (1995) 'Genes involved in immunity to the lantibiotic nisin produced by *Lactococcus lactis* 6F3.', *Applied and environmental microbiology*. United States, 61(3), pp. 1082–1089.

Sievers, F. and Higgins, D. G. (2014) 'Clustal Omega, accurate alignment of very large numbers of sequences.', *Methods in molecular biology (Clifton, N.J.)*. United States, 1079, pp. 105–116. doi: 10.1007/978-1-62703-646-7_6.

Silva, C. C. G., Silva, S. P. M. and Ribeiro, S. C. (2018) 'Application of Bacteriocins and Protective Cultures in Dairy Food Preservation', *Frontiers in microbiology*. Frontiers Media S.A., 9, p. 594. doi: 10.3389/fmicb.2018.00594.

Singh, M. and Sareen, D. (2014) 'Novel LanT associated lantibiotic clusters identified by genome database mining.', *PloS one*. United States, 9(3), p. e91352. doi: 10.1371/journal.pone.0091352.

Smits, L. P., Bouter, K. E. C., De Vos, W. M., Borody, T. J. and Nieuwdorp, M. (2013) 'Therapeutic Potential of Fecal Microbiota Transplantation', *Gastroenterology*, 145(5), pp. 946–953.

Soliman, W., Bhattacharjee, S. and Kaur, K. (2007) 'Molecular dynamics simulation study of interaction between a class IIa bacteriocin and its immunity protein.', *Biochimica et biophysica acta*. Netherlands, 1774(8), pp. 1002–1013. doi: 10.1016/j.bbapap.2007.05.009.

Solovyev, V. and Salamov, A. (2011) 'Automatic Annotation of Microbial Genomes and Metagenomic Sequences', in Li, R. W. (ed.) *Metagenomics and its Applications in Agriculture, Biomedicine and Environmental Studies*. Nova Science Publishers, pp. 61–78.

Song, A. A.-L., In, L. L. A., Lim, S. H. E. and Rahim, R. A. (2017) 'A review on *Lactococcus lactis*: from food to factory', *Microbial cell factories*. BioMed Central, 16(1), p. 55. doi: 10.1186/s12934-017-0669-x.

de Souza Barbosa, M., Todorov, S. D., Ivanova, I., Chobert, J.-M., Haertlé, T. and de Melo Franco, B. D. G. (2015) 'Improving safety of salami by application of bacteriocins produced by an autochthonous *Lactobacillus curvatus* isolate', *Food Microbiology*, 46, pp. 254–262. doi: https://doi.org/10.1016/j.fm.2014.08.004.

Spigaglia, P. (2016) 'Recent Advances in the Understanding of Antibiotic Resistance in *Clostridium difficile* Infection', *Therapeutic Advances in Infectious Disease*, 3(1), pp. 23–42.

Stein, T. (2005) '*Bacillus subtilis* antibiotics: structures, syntheses and specific functions.', *Molecular microbiology*. England, 56(4), pp. 845–857. doi:

10.1111/j.1365-2958.2005.04587.x.

Stein, T., Heinzmann, S., Solovieva, I. and Entian, K. D. (2003) 'Function of *Lactococcus lactis* Nisin Immunity Genes nisI and nisFEG after Coordinate Expression in the Surrogate Host *Bacillus subtilis*', *The Journal of Biology Chemistry*, 278(1), pp. 89–94.

Stoddard, G. W., Petzel, J. P., van Belkum, M. J., Kok, J. and McKay, L. L. (1992) 'Molecular analyses of the lactococcin A gene cluster from *Lactococcus lactis* subsp. *lactis* biovar diacetylactis WM4.', *Applied and environmental microbiology*. United States, 58(6), pp. 1952–1961.

Stubbs, S. L. J., Brazier, J. S., O'Neill, G. L. and Duerden, B. I. (1999) 'PCR Targeted to the 16S-23S rRNA Gene Intergenic Spacer Region of *Clostridium difficile* and Construction of a Library Consisting of 116 Different PCR Ribotypes', *Journal of Clinical Microbiology*, 37(2), pp. 461 LP – 463. Available at: http://jcm.asm.org/content/37/2/461.abstract.

Su, P., Im, H., Hsieh, H., Kang'A, S. and Dunn, N. W. (1999) 'LlaFI, a Type III Restriction and Modification System in *Lactococcus lactis*', *Applied and Environmental Microbiology*, 65(2), pp. 686–693. Available at: http://www.ncbi.nlm.nih.gov/pmc/articles/PMC91080/.

Sugai, M., Fujiwara, T., Akiyama, T., Ohara, M., Komatsuzawa, H., Inoue, S. and Suginaka, H. (1997) 'Purification and molecular characterization of glycylglycine endopeptidase produced by *Staphylococcus capitis* EPK1.', *Journal of bacteriology*. United States, 179(4), pp. 1193–1202. doi: 10.1128/jb.179.4.1193-1202.1997.

Sun, Z., Zhong, J., Liang, X., Liu, J., Chen, X. and Huan, L. (2009) 'Novel mechanism for nisin resistance via proteolytic degradation of nisin by the nisin resistance protein NSR.', *Antimicrobial agents and chemotherapy*. United States, 53(5), pp. 1964–1973. doi: 10.1128/AAC.01382-08.

Takala, T. M. and Saris, P. E. J. (2006) 'C-Terminus of NisI Provides Specificity to Nisin', *Microbiology*, 152, pp. 3543–3549. doi: 10.1099/mic.0.29083-0.

Takamizawa, A., Miyata, S., Matsushita, O., Kaji, M., Taniguchi, Y., Tamai, E., Shimamoto, S. and Okabe, A. (2004) 'High-level expression of clostridial sialidase using a ferredoxin gene promoter-based plasmid', *Protein Expression and Purification*, 36(1), pp. 70–75. doi: https://doi.org/10.1016/j.pep.2004.03.004.

Tamura, K. and Nei, M. (1993) 'Estimation of the number of nucleotide substitutions in the control region of mitochondrial DNA in humans and chimpanzees.', *Molecular biology and evolution*. United States, 10(3), pp. 512–526. doi: 10.1093/oxfordjournals.molbev.a040023.

Taylor, T. M., Davidson, P. M. and Zhong, Q. (2007) 'Extraction of nisin from a 2.5% commercial nisin product using methanol and ethanol solutions.', *Journal of food protection*. United States, 70(5), pp. 1272–1276.

Tena, D., Martinez-Torres, J. A., Perez-Pomata, M. T., Saez-Nieto, J. A., Rubio, V. and Bisquert, J. (2007) 'Cutaneous infection due to *Bacillus pumilus*: report of 3 cases.', *Clinical infectious diseases : an official publication of the Infectious Diseases Society of America*. United States, 44(4), pp. e40-2. doi: 10.1086/511077.

Terhes, G., Urban, E., Soki, J., Hamid, K. A. and Nagy, E. (2004) 'Community-acquired *Clostridium difficile* diarrhea caused by binary toxin, toxin A, and toxin B gene-

positive isolates in Hungary.', *Journal of clinical microbiology*. United States, 42(9), pp. 4316–4318. doi: 10.1128/JCM.42.9.4316-4318.2004.

Tikhonov, A., Kazakov, T., Semenova, E., Serebryakova, M., Vondenhoff, G., Van Aerschot, A., Reader, J. S., Govorun, V. M. and Severinov, K. (2010) 'The mechanism of microcin C resistance provided by the MccF peptidase.', *The Journal of biological chemistry*. United States, 285(49), pp. 37944–37952. doi: 10.1074/jbc.M110.179135.

Todorov, S. D., van Reenen, C. A. and Dicks, L. M. T. (2004) 'Optimization of bacteriocin production by *Lactobacillus plantarum* ST13BR, a strain isolated from barley beer.', *The Journal of general and applied microbiology*. Japan, 50(3), pp. 149–157.

Towle, K. M. and Vederas, J. C. (2017) 'Structural Features of Many Circular and Leaderless Bacteriocins are Similar to those in Saposins and Saposin-like Peptides', *Medicinal Chemical Communications*, 8(276).

Tricker, A. R. and Preussmann, R. (1991) 'Carcinogenic N-nitrosamines in the diet: occurrence, formation, mechanisms and carcinogenic potential.', *Mutation research*. Netherlands, 259(3–4), pp. 277–289.

UniProt Consortium, T. (2018) 'UniProt: the universal protein knowledgebase', *Nucleic Acids Research*, 46(5), p. 2699. doi: 10.1093/nar/gky092.

Urdaci, M. C., Bressollier, P. and Pinchuk, I. (2004) '*Bacillus clausii* Probiotic Strains: Antimicrobial and Immunomodulatory Activities', *Journal of Clinical Gastroenterology*, 38. Available at: https://journals.lww.com/jcge/Fulltext/2004/07002/Bacillus_clausii_Probiotic_Strains__Antimicrobial.10.aspx.

Usui, M., Kawakura, M., Yoshizawa, N., San, L. L., Nakajima, C., Suzuki, Y. and Tamura, Y. (2017) 'Survival and prevalence of *Clostridium difficile* in manure compost derived from pigs', *Anaerobe*, 43, pp. 15–20. doi: https://doi.org/10.1016/j.anaerobe.2016.11.004.

Vadakedath, N. and Halami, P. M. (2019) 'Characterization and mode of action of a potent bio-preservative from food-grade *Bacillus licheniformis* MCC 2016', *Preparative Biochemistry and Biotechnology*. Taylor & Francis, 49(4), pp. 334–343. doi: 10.1080/10826068.2019.1566141.

Varcamonti, M., Nicastro, G., Venema, G. and Kok, J. (2001) 'Proteins of the lactococcin A secretion system: lcnD encodes two in-frame proteins', *FEMS Microbiology Letters*, 204(2), pp. 259–263. doi: 10.1111/j.1574-6968.2001.tb10894.x.

Vasilchenko, A. S. and Valyshev, A. V (2019) 'Pore-forming bacteriocins: structural–functional relationships', *Archives of Microbiology*, 201(2), pp. 147–154. doi: 10.1007/s00203-018-1610-3.

Venema, K., Abee, T., Haandrikman, A. J., Leenhouts, K. J., Kok, J., Konings, W. N. and Venema, G. (1993) 'Mode of Action of Lactococcin B, a Thiol-Activated Bacteriocin from *Lactococcus lactis*.', *Applied and environmental microbiology*. United States, 59(4), pp. 1041–1048.

Venema, K., Chikindas, M. L., Seegers, J., Haandrikman, A. J., Leenhouts, K. J., Venema, G. and Kok, J. (1997) 'Rapid and Efficient Purification Method for Small, Hydrophobic, Cationic Bacteriocins: Purification of Lactococcin B and Pediocin PA-1.',

Applied and environmental microbiology. United States, 63(1), pp. 305–309.

Venema, K., Haverkort, R. E., Abee, T., Haandrikman, A. J., Leenhouts, K. J., de Leij, L., Venema, G. and Kok, J. (1994) 'Mode of action of LciA, the lactococcin A immunity protein.', *Molecular microbiology*. England, 14(3), pp. 521–532. doi: 10.1111/j.1365-2958.1994.tb02186.x.

Vera Pingitore, E., Salvucci, E., Sesma, F. and Nader-Macías, M. (2007) 'Different strategies for purification of antimicrobial peptides from Lactic Acid Bacteria (LAB)', *Microbiology book series*, 2.

Verheul, A., Russell, N. J., Van'T Hof, R., Rombouts, F. M. and Abee, T. (1997) 'Modifications of membrane phospholipid composition in nisin-resistant *Listeria monocytogenes* Scott A', *Applied and environmental microbiology*, 63(9), pp. 3451–3457. Available at: https://www.ncbi.nlm.nih.gov/pubmed/9292996.

Vijay Simha, B., Sood, S. K., Kumariya, R. and Garsa, A. K. (2012) 'Simple and rapid purification of pediocin PA-1 from *Pediococcus pentosaceous* NCDC 273 suitable for industrial application', *Microbiological Research*, 167(9), pp. 544–549. doi: https://doi.org/10.1016/j.micres.2012.01.001.

de Vos, W. M., Mulders, J. W., Siezen, R. J., Hugenholtz, J. and Kuipers, O. P. (1993) 'Properties of nisin Z and distribution of its gene, nisZ, in *Lactococcus lactis*', *Applied and environmental microbiology*, 59(1), pp. 213–218. Available at: https://www.ncbi.nlm.nih.gov/pubmed/8439149.

Voth, D. E. and Ballard, J. D. (2005) '*Clostridium difficile* Toxins: Mechanism of Action and Role in Disease', *Clinical Microbiology Reviews*, 18(2), pp. 247–263.

van der Wal, F. J., Luirink, J. and Oudega, B. (1995) 'Bacteriocin release proteins: mode of action, structure, and biotechnological application.', *FEMS microbiology reviews*, 17(4), pp. 381–99. doi: 10.1111/j.1574-6976.1995.tb00221.x.

Wang, J., Ge, X., Zhang, L., Teng, K. and Zhong, J. (2016) 'One-pot synthesis of class II lanthipeptide bovicin HJ50 via an engineered lanthipeptide synthetase.', *Scientific reports*, 6, p. 38630. doi: 10.1038/srep38630.

Wang, L.-T., Lee, F.-L. and Kasai, H. (2007) 'Comparison of gyrB Gene Sequences, 16S rRNA Gene Sequences and DNA-DNA Hybridization in the *Bacillus subtilis* Group', *International Journal of Systematic and Evolutionary Microbiology*, 57, pp. 1846–1850.

Warny, M., Pepin, J., Fang, A., Killgore, G., Thompson, A., Brazier, J. S., Frost, E. and McDonald, L. C. (2005) 'Toxin Production by an Emerging Strain of *Clostridium difficile* Associated with Outbreaks of Severe Disease in North America and Europe', *Lancet*, 366, pp. 1079–1084.

Waterhouse, A. *et al.* (2018) 'SWISS-MODEL: homology modelling of protein structures and complexes', *Nucleic Acids Research*, 46(W1), pp. W296–W303. doi: 10.1093/nar/gky427.

Weese, J. S., Wakeford, T., Reid-Smith, R., Rousseau, J. and Friendship, R. (2010) 'Longitudinal investigation of *Clostridium difficile* shedding in piglets', *Anaerobe*, 16(5), pp. 501–504. doi: https://doi.org/10.1016/j.anaerobe.2010.08.001.

Wells, C. L. and Wilkins, T. D. (1996) 'Clostridia: Sporeforming Anaerobic Bacilli', in Baron, S. (ed.). Galveston (TX).

Whitford, M. F., McPherson, M. A., Forster, R. J. and Teather, R. M. (2001) 'Identification of Bacteriocin-Like Inhibitors from Rumen *Streptococcus* spp. and Isolation and Characterization of Bovicin 255', *Applied and Environmental Microbiology*, 67(2), pp. 569–574.

WHO, W. H. O. (1969) *Specifications for identity and purity of some antibiotics*, *WHO/Food Additives*.

Wilcox, M. H. *et al.* (2017) 'Bezlotoxumab for Prevention of Recurrent *Clostridium difficile* Infection', *New England Journal of Medicine*. Massachusetts Medical Society, 376(4), pp. 305–317. doi: 10.1056/NEJMoa1602615.

Wingett, S. W. and Andrews, S. (2018) 'FastQ Screen: A tool for multi-genome mapping and quality control.', *F1000Research*. England, 7, p. 1338. doi: 10.12688/f1000research.15931.2.

Wirawan, R. E., Klesse, N. A., Jack, R. W. and Tagg, J. R. (2006) 'Molecular and genetic characterization of a novel nisin variant produced by *Streptococcus uberis*.', *Applied and environmental microbiology*. United States, 72(2), pp. 1148–1156. doi: 10.1128/AEM.72.2.1148-1156.2006.

Wirawan, R. E., Swanson, K. M., Kleffmann, T., Jack, R. W. and Tagg, J. R. (2007) 'Uberolysin: a novel cyclic bacteriocin produced by *Streptococcus uberis*.', *Microbiology (Reading, England)*. England, 153(Pt 5), pp. 1619–1630. doi: 10.1099/mic.0.2006/005967-0.

Woods, C., Humphreys, C. M., Rodrigues, R. M., Ingle, P., Rowe, P., Henstra, A., Winzer, K., Köpke, M., Simpson, S. D. and Minton, N. P. (2019) *A Novel Conjugal Donor Strain for Improved DNA transfer into Clostridium spp.*, *Anaerobe*.

Wu, Y. C., Lee, J. J., Tsai, B.-Y., Liu, Y. F., Chen, C.-M., Tien, N., Tsai, P.-J. and Chen, T.-H. (2016) 'Potentially Hypervirulent *Clostridium difficile* PCR Ribotype 078 Lineage Isolates in Pigs and Possible Implactions for Humans in Taiwan', *International Journal of Medical Microbiology*, 306, pp. 115–122.

Wullt, M. and Odenholt, I. (2004) 'A double-blind randomized controlled trial of fusidic acid and metronidazole for treatment of an initial episode of *Clostridium difficile*-associated diarrhoea', *Journal of Antimicrobial Chemotherapy*, 54(1), pp. 211–216. doi: 10.1093/jac/dkh278.

Xiao, H., Chen, X., Chen, M., Tang, S. and Zhao, X. (2004) 'Bovicin HJ50, A Novel Lantibiotic Produced by *Streptococcus bovis* HJ50', *Microbiology*, 150, pp. 103–108.

Yang, R., Johnson, M. C. and Ray, B. (1992) 'Novel method to extract large amounts of bacteriocins from lactic acid bacteria', *Applied and environmental microbiology*, 58(10), pp. 3355–3359. Available at: https://www.ncbi.nlm.nih.gov/pubmed/1444369.

Yang, S.-C., Lin, C.-H., Sung, C. T. and Fang, J.-Y. (2014) 'Antibacterial Activities of Bacteriocins: Application in Food and Pharmaceuticals', *fronitiers in Microbiology*, 5, pp. 1–10.

Yang, Z. *et al.* (2015) 'Mechanisms of protection against *Clostridium difficile* infection by the monoclonal antitoxin antibodies actoxumab and bezlotoxumab', *Infection and immunity*. 2014/12/08. American Society for Microbiology, 83(2), pp. 822–831. doi: 10.1128/IAI.02897-14.

Young, R. and Blasi, U. (1995) 'Holins: form and function in bacteriophage lysis.', *FEMS microbiology reviews*. England, 17(1–2), pp. 191–205. doi: 10.1111/j.1574-6976.1995.tb00202.x.

Yuan, J., Zhang, Z.-Z., Chen, X.-Z., Yang, W. and Huan, L.-D. (2004) 'Site-directed mutagenesis of the hinge region of nisinZ and properties of nisinZ mutants.', *Applied microbiology and biotechnology*. Germany, 64(6), pp. 806–815. doi: 10.1007/s00253-004-1599-1.

Yuan, P. *et al.* (2015) 'Chondroitin Sulfate Proteoglycan 4 Functions as the Cellular Receptor for *Clostridium difficile* Toxin B', *Cell Research*, 25(2), pp. 157–168. doi: 10.1038/cr.2014.169.

Yuan, Y. and Gao, M. (2015) 'Genomic Analysis Of A Ginger Pathogen *Bacillus pumilus* Providing The Understanding to The Pathogenesis and The Novel Control Strategy', *Nature Scientific Reports*, 5.

Yutin, N. and Galperin, M. Y. (2013) 'A genomic update on clostridial phylogeny: Gram-negative spore formers and other misplaced clostridia.', *Environmental microbiology*. England, 15(10), pp. 2631–2641. doi: 10.1111/1462-2920.12173.

Zeigler, D. and Perkins, J. (2008) 'The Genus Bacillus', in, pp. 309–337. doi: 10.1201/9781420009330.ch24.

Zendo, T., Fukao, M., Ueda, K., Higuchi, T., Nakayama, J. and Sonomoto, K. (2003) 'Identification of the lantibiotic nisin Q, a new natural nisin variant produced by *Lactococcus lactis* 61-14 isolated from a river in Japan.', *Bioscience, biotechnology, and biochemistry*. England, 67(7), pp. 1616–1619. doi: 10.1271/bbb.67.1616.

Zendo, T., Koga, S., Shigeri, Y., Nakayama, J. and Sonomoto, K. (2006) 'Lactococcin Q, a Novel Two-Peptide Bacteriocin Produced by *Lactococcus lactis* QU 4', *Applied and Environmental Microbiology*, 72(5), pp. 3383 LP – 3389. doi: 10.1128/AEM.72.5.3383-3389.2006.

Zhang, Q., Yu, Y., Vélasquez, J. E. and van der Donk, W. A. (2012) 'Evolution of lanthipeptide synthetases', *Proceedings of the National Academy of Sciences*, 109(45), pp. 18361 LP – 18366. doi: 10.1073/pnas.1210393109.

Zhou, H., Fang, J., Tian, Y. and Lu, X. Y. (2014) 'Mechanisms of nisin resistance in Gram-positive bacteria', *Annals of Microbiology*, 64(2), pp. 413–420. doi: 10.1007/s13213-013-0679-9.

Zidaric, V., Pardon, B., Dos Vultos, T., Deprez, P., Brouwer, M. S. M., Roberts, A. P., Henriques, A. O. and Rupnik, M. (2012) 'Different antibiotic resistance and sporulation properties within multiclonal *Clostridium difficile* PCR ribotypes 078, 126, and 033 in a single calf farm', *Applied and environmental microbiology*. 2012/09/21. American Society for Microbiology, 78(24), pp. 8515–8522. doi: 10.1128/AEM.02185-12.