



Ajeigbe, Mayokun (2023) Genetic analysis and Manipulation of Bacterial Microcin Peptides Biosynthesis. Doctoral thesis, The University of Sunderland.

Downloaded from: <http://sure.sunderland.ac.uk/id/eprint/17023/>

Usage guidelines

Please refer to the usage guidelines at <http://sure.sunderland.ac.uk/policies.html> or alternatively contact sure@sunderland.ac.uk.



**University of
Sunderland**

Genetic Analysis and Manipulation of Bacterial Microcin Peptides Biosynthesis

Mayokun Ajeigbe

This thesis was submitted in partial fulfilment of the requirements of the University of Sunderland, for the degree of Doctor of Philosophy.

Supervisor: Dr Lewis Bingle

Co-supervisor: Dr Timothy Paget

Co-supervisor: Dr Stephen Childs

June 2023.

Dedication

This PhD thesis is dedicated to my beloved daughter, Iseoluware Joanna, born during this thesis writeup.

Acknowledgments

I am filled with joy that this journey was a success. During my PhD, many people supported me in one way or another, and it is worthwhile to appreciate them.

My profound gratitude goes to my supervisor Dr. Lewis Bingle for his support, guidance, and kind gestures from the beginning of my PhD. I thank my co-supervisors, Dr. Timothy Paget and Dr. Stephen Childs, for their supervision and encouragement.

I cannot forget the prayer, encouragement, and financial support of my mother, Mrs. Olufunmilola Ajeigbe.

My family, Mr. Tope & Mrs. Jumoke Bamgbose, Mr. Michael & Mrs. Itunu Situ, Mrs. Abike Situ, Prof. Tunde & Mrs. Wumi Oladiran, Mr. Olubunmi & Mrs. Titilope Sobodu, Dr. Fiyinfolu & Mrs. Toyin Oladiran, Dr. Olumide & Mrs. Ibukun Aluko, have all encouraged me at one point and I am grateful. To my nephews Ayomide, Nifemi, Ethan and Elliot who provided the much-needed comic relief when required. Thank you all so much.

I also want to thank Dr. Callum Cooper for the insights and encouragement he gave from time to time.

I am grateful for the tireless support from the technicians in the University of Sunderland microbiology lab. and analytical chemistry lab. Linda Henry, Kayleigh Ironside, Paul Stronach, Maria Barry, and Julie McLaren for their constant help and assistance around the lab.

Special acknowledgement has to be given to my wife, Dr. Ore-Ofe Ajeigbe for her understanding, prayers and ceaseless support. I love you, my Sunshine.

Lastly, I cannot begin to show my appreciation to my heavenly Father. God, I thank You for giving me strength and grace throughout my lab work and write-up. Without You, all this would have been impossible.

Abstract

Researchers are actively searching for new antibiotics or alternatives to antibiotics to combat the global challenge of microbial resistance. One group of alternative antimicrobials that have been intensively studied are the microcins. Microcin B17 (MccB17) is a ribosomally synthesized and post-translationally modified peptide (RiPP) that targets DNA gyrase subunit B. MccB17 is encoded by 7 genes arranged in an operon (*mcb* operon) carried on a large conjugative plasmid, pMccB17, that was originally found in *E. coli* LP17.

This work provided complete DNA sequence and in-depth bioinformatic analysis of the 69 kb circular pMccB17 plasmid and confirmed its previous assignment to the IncFII family. Replicon typing using the FAB formula placed pMccB17 into F2:A-:B- and it was assigned to the MOB_{F12} group. pMccB17 has an average GC content of 51% and 81 predicted coding sequences, 30 of which encode conjugation proteins and as previously suggested, it has no antibiotic resistance genes. Plasmid-encoded Pbf protein, homologous to proteins involved in plasmid or chromosome partitioning, was of particular interest due to its conservation in other IncF plasmids and its unknown function. Sequence variations in *mcbB*, *mcbD* and *mcbF* between the historic sequences of *mcb* operon (GenBank accession numbers M24253 and X07875) in comparison to the *mcb* operon from this work were noted which account for 2 amino acids changes in the encoded proteins of McbB and McbD, and 17 amino acids changes in the encoded protein of McbF.

Homologues of the *mcb* operon genes were discovered in the genome of the plant pathogen *Erwinia persicina* NBRC 102418. In previous work, the microcin operon from *E. persicina* (*emo* operon) has been PCR amplified with and without an extra ORF found downstream of its operon, in addition to the predicted upstream promoter region, and cloned into the *E. coli* pUC 18 and pUC19 plasmid vectors. When transformed into *E. coli* DH5 α for expression, the pUC19-derived constructs showed a slightly higher level of microcin production (10%) which could be due to the orientation of the *lac* promoter in the vectors. The presence of the extra ORF was found not to affect microcin production. Bioassay experiments carried out to characterise the *Erwinia persicina* microcin (EMB17) showed that it is maximally produced during the stationary phase of growth. EMB17 has a narrow activity range being able to inhibit the growth of some proteobacteria such as *Pseudomonas aeruginosa* and *Klebsiella pneumonia*, but not the growth of any Gram-positive bacteria tested. Two peaks having the same mass of 1088.9 m/z were picked up by mass spectrometer

during the analysis of EMB17, purification parameters were set around this mass using preparative HPLC-MS. When tested, it was confirmed that the purified product retained its antimicrobial activity.

The *mcbA* and *mcbB* genes from *emo* and *mcb* operons were individually deleted from these operons and cloned into inducible expression vectors for expression *in trans*. The single cloned gene was then paired with the remainder of the operon complementarily to make up cognate and non-cognate microcin operons *in trans*. Comparison of bioassay results from cognate and non-cognate reassembled operon indicated that the protein-protein interaction between the promicrocin and the synthase is not species-specific. The resulting microcin was active against the sensitive *E. coli* BL21 strain. *Cis* operon arrangement favours more microcin production in *mcb* operon, while the *trans* arrangement favours microcin production in *emo* operon. This molecular promiscuity may facilitate the engineering of mutant microcins with altered chemistry and potentially improved or novel antimicrobial properties.

Table of Contents

<i>Dedication</i>	2
<i>Acknowledgments</i>	3
<i>Abstract</i>	4
<i>List of figures</i>	10
<i>List of tables</i>	14
<i>List of abbreviations</i>	15
<i>Chapter one</i>	16
INTRODUCTION	16
1.0 Overview	16
1.1 Antibiotics	17
1.1.1 Spectrum of activity of antibiotics	17
1.1.2 Antimicrobial resistance	18
1.2 Bacteriocins	19
1.2.1 Classification of bacteriocins	20
1.2.1.1 Bacteriocins from Archaea	20
1.2.1.2 Bacteriocins from Gram-positive bacteria	21
1.2.1.3 Bacteriocins from Gram-negative bacteria	22
1.2.2 Mode of actions of bacteriocins	23
1.2.2.1 Cell wall/peptidoglycan disruption.....	24
1.2.2.2 Bacterial membrane disruption	24
1.2.2.3 Inhibition of protein biosynthesis.....	25
1.2.2.4 Inhibition of DNA replication and transcription	25
1.2.2.5 Inhibition of folate biosynthesis	25
1.2.3 Current and Future Applications of bacteriocins	25
1.3 Microcin B17	26
1.3.1 Microcin B17 Biosynthesis	27
1.4 Ribosomally synthesized and post-translationally modified peptide (RiPP)	30
1.4.1 Classes of RiPPs	31
1.5 Taxonomy of the Enterobacterales with a focus on <i>E. coli</i> and <i>E. persicina</i>	33
1.5.1 Gammaproteobacteria	33
1.5.1.1 Enterobacterales	35
1.6 Plasmids	37
1.6.1 Plasmid replication systems	37
1.6.1.1 Theta Replication	38
1.6.1.2 Strand displacement replication	38
1.6.1.3 Rolling circle replication	38
1.6.2 Classification of Plasmids	39
1.6.2.1 Incompatibility group.....	39

1.6.2.2	Conjugative functions	39
1.6.2.3	Host range.....	40
1.6.3	Plasmid copy number.....	40
1.6.4	Plasmid encoded maintenance systems.....	41
1.6.4.1	Multimer resolution system.....	41
1.6.4.2	Active partitioning systems	41
1.6.4.3	Post segregation killing systems	43
1.7	Aims and objectives of this work.....	44
	Chapter two	45
	Materials and methods.....	45
2.1	Materials.....	45
2.1.1	Buffers	45
2.1.2	Media	45
2.1.3	Enzymes.....	45
2.1.4	Oligonucleotides.....	46
2.1.5	Bacterial strains and plasmids.....	48
2.2	Methods	51
2.2.1	Isolation of plasmid DNA	51
2.2.2	Illumina sequencing of pMccB17 DNA.....	51
2.2.3	Annotation and Visualization of pMccB17.....	51
2.2.4	Plasmid replicon and conjugation system typing.....	52
2.2.5	Phylogenetic tree building	52
2.2.6	Identification of Resistance Genes, Insertion Sequences and Virulence factors.....	53
2.2.7	PCR amplification of DNA.....	54
2.2.8	Agarose gel electrophoresis	54
2.2.9	Purification of PCR products	54
2.2.10	Restriction enzyme digestion	55
2.2.11	Alkaline dephosphorylation	55
2.2.12	Phosphorylation	55
2.2.13	Ligation.....	55
2.2.14	Transformation	55
2.2.16	In vivo determination of microcin activity (bioassay).....	56
2.2.17	Microcin purification steps	56
2.2.18	Liquid chromatography–mass spectrometry analysis.....	57
2.2.19	Preparative High pressure liquid chromatography analysis.....	57
	Chapter Three	59

<i>pMccB17</i>	59
3.0 Overview	59
3.1 Introduction.....	59
3.2 Results and discussion	59
3.2.1 Purification and sequencing of plasmid <i>pMccB17</i>	59
3.2.2 Plasmid sequence assembly	60
3.2.3 PCR and plasmid gap closing	63
3.2.4 Annotation of <i>pMccB17</i> sequence and submission to GenBank	63
3.2.5 Visualization.....	67
3.2.6 Direct repeats.....	68
3.2.7 Plasmid replicon and conjugation system typing.....	70
3.2.8 <i>pMccB17</i> overview	70
3.3 The microcin B17 operon of <i>pMccB17</i>	71
3.3.1 Sequence variation	74
3.3.2 Confirming the variations.....	74
3.4 ParB-fusion protein (<i>Pbf</i>).....	79
Chapter four.....	91
Cloning and analysis of <i>Erwinia persicina</i> microcin operon.....	91
4.0 Overview	91
4.1 Discovery of <i>Erwinia</i> microcin operon (<i>emo</i> operon)	91
4.1.1 Previous background work	92
4.2 An extra ORF included in the <i>E. persicina</i> microcin operon	92
4.3 Comparison of <i>E. coli</i> and <i>E. persicina</i> microcin operon	94
4.3.1 Comparison of <i>McbA</i> peptides.....	95
4.3.2 Comparison of <i>McbBs</i> proteins	98
4.4 Cloning of <i>Erwinia persicina</i> microcin operon	100
4.5 Next generation sequencing of <i>pMBB-3</i>	103
4.6 Bioassay of antimicrobial activity of the <i>EMB17</i>	104
4.7 Characterization of <i>Erwinia</i> microcin B17 (<i>EMB17</i>)	107
4.7.1 Effects of IPTG on microcin production	107
4.7.2 Time course optimization experiment	108
4.7.3 Examining mechanism of action via bioassay of deletion mutant strains	112
4.7.3.1 Study with Δ <i>tldD</i> producer cells	112
4.7.3.2 Study with Δ <i>ompF</i> and Δ <i>sbmA</i> strains.....	113
4.7.4 Spectrum of action of <i>EMB17</i>	114
4.8 Analysis and Purification of <i>EMB17</i>	116

4.8.1	<i>Antimicrobial activity of DMSO on E. coli BL21[pUC19]</i>	117
4.8.2	<i>Mass spectrometry analysis of EMB17</i>	118
4.8.3	<i>Preparative HPLC-MS analysis</i>	124
<i>Chapter five</i>		132
<i>Genetic analysis of promicrocin – synthase interactions of E. persicina and E. coli microcin gene clusters</i>		132
5.0	<i>Overview</i>	132
5.1	<i>Strategy for genetic analysis of modified microcin gene clusters</i>	132
5.1.2	<i>Cloning of mcbA into SEVA234 plasmid vector</i>	144
5.1.3	<i>Deletion of mcbB from E. persicina and E. coli microcin operon</i>	151
5.1.4	<i>Cloning of mcbB into SEVA234 plasmid vector</i>	159
5.2	<i>Pairing of complementary gene(s) in the microcin operon (mix and match)</i>	163
5.2.1	<i>Deriving pBR322 from pPY113</i>	165
5.2.2	<i>Bioassay analysis of the paired microcin genes</i>	167
<i>Chapter six</i>		172
<i>Discussion and conclusions</i>		172
6.1	<i>Plasmid pMccB17</i>	172
6.2	<i>The Erwinia microcin, EMB17</i>	174
<i>References</i>		177

List of figures

Figure 1.1	Targets of both antibiotics and bacteriocins, their general location, and examples of each capable of inhibition of these targets	24
Figure 1.2	Amino acid sequence of MccB17 showing the location of the post-translational modifications	28
Figure 1.3	Biosynthesis of MccB17 and mechanism of action in sensitive cells.....	29
Figure 1.4	Phylogenetic tree of the old Enterobacteriaceae family	34
Figure 3.1	Relationship between the first and second draft sequence of pMccB17.....	62
Figure 3.2	Circular Map of pMccB17	67
Figure 3.3	Genomic view of the regions between the direct repeats in pMccB17 showing the location of the 149 bp repeats as well as the genes present between them.....	69
Figure 3.4	A chromatogram (from Sanger sequencing of pMccB17 PCR amplicons) and a sequence view in Artemis of the original <i>mcb</i> operon sequence (M24253) showing the single base substitution errors (GC swaps) in the original <i>mcb</i> operon.....	77
Figure 3.5	Deletion in <i>mcbF</i> of the X07875 sequence, which resulted into frameshift.....	78
Figure 3.6	A BLASTn result of <i>mcbF</i> ON989342 sequence showing other homologues with 100% identity in the nr/nt database	78
Figure 3.7	A BLASTn result of <i>mcbF</i> X07875 sequence showing homologous sequences in the nr/nt database	79
Figure 3.8	Location of known conserved domains in the ParB-fusion protein.....	80
Figure 3.9	Prediction of transmembrane segment in selected Pbf proteins	82
Figure 3.10	MSA of ParB protein family members, including the fused-ParB-like protein	84
Figure 3.11	An overview of the fused-ParB-like protein	85
Figure 3.12	Phylogenetic analysis of ParB family proteins	87
Figure 3.13	Conservation of neighbouring genes of <i>pbf</i>	89

Figure 4.1	Overview of <i>Erwinia</i> microcin operon (<i>emo</i> operon) showing the location of the <i>eeg</i>	93
Figure 4.2	Alignment of McbA from <i>E. coli</i> and <i>E. persicina</i> amino acid sequences.....	95
Figure 4.3	Amino acid composition of the types of rings (heterocycles) formed during cyclization of <i>Erwinia</i> and <i>E. coli</i> McbA	97
Figure 4.4	Alignment of McbB from <i>E. coli</i> and <i>E. persicina</i> amino acid sequences.....	99
Figure 4.5	Maps of the cloned <i>emo</i> operon showing complete operon plasmid constructs.....	102
Figure 4.6	pMBB-3 showing the microcin genes (<i>mcbA – mcbG</i>) and the <i>eeg</i> cloned into pUC19 vector	103
Figure 4.7	<i>In vivo</i> activity of microcin production using <i>E. coli</i> BL21[pUC19] as the indicator strain	105
Figure 4.8	Bioassay of microcin production by cloned <i>emo</i> operon	106
Figure 4.9	Bioassay experiment to study the effect of IPTG on microcin production for the <i>emo</i> operon clones	107
Figure 4.10	Bacterial growth curve for <i>E. coli</i> carrying pMBB-3, pPY113 and pUC19 in LB medium	109
Figure 4.11	Bacterial growth curve in M63 medium	110
Figure 4.12	Time course study of microcin production assessed by bioassay in LB and M63 media	111
Figure 4.13	Time-course study comparing <i>in vivo</i> activity of pMBB-3 microcin production with time in LB and M63 medium	112
Figure 4.14	Effects of <i>tldD</i> on microcin production	113
Figure 4.15	Effects of <i>ompF</i> and <i>sbmA</i> on the transport of EMB17 into sensitive strains.....	114
Figure 4.16	Chromatogram showing the peaks of EMB17	122
Figure 4.17	The mass spectrum of the molecular ion of EMB17 from the peak at 29th minute showing the m/z 1088.9	123
Figure 4.18	Chromatogram showing time-based collection of EMB17 sample.....	125

Figure 4.19	The mass spectrum of the selected peaks from the preparative HPLC analysis of EMB17	128
Figure 4.20	Image of Time-based collection bioassay	130
Figure 5.1	Schematic illustration of how plasmid constructs were made and used for the mix and match pairing using <i>mcbA</i> as an example	134
Figure 5.2	The deletion pattern of <i>mcbA</i> and the residue that remained after deletion.....	136
Figure 5.3	Agarose gel electrophoresis of PCR amplicons after the deletion of <i>mcbA</i> from the microcin operons	137
Figure 5.4	Agarose gel electrophoresis to confirm deletion of <i>mcbA</i> by restriction enzyme digest of pMBB-3 Δ <i>mcbA</i> and pMBB-3	139
Figure 5.5	Agarose gel electrophoresis to confirm deletion of <i>mcbA</i> by restriction enzyme digest of pPY113 Δ <i>mcbA</i> and pPY113	140
Figure 5.6	Plasmid maps of the cloned microcin operons showing deletion of <i>mcbA</i> and incomplete microcin operon	141
Figure 5.7	Plasmid maps of cloned microcin operon deletion mutants of <i>mcbA</i>	143
Figure 5.8	Agarose gel electrophoresis of PCR amplicons, EcoRI – Sall <i>mcbA</i> cassette (~200bp) from <i>E. coli</i> and <i>Erwinia</i> microcin operons	145
Figure 5.9	Plasmid maps of pSEVA234 derivatives with <i>mcbA</i> insert	146
Figure 5.10	Confirmation of cloning of <i>mcbA</i> into pSEVA234	147
Figure 5.11	Plasmid maps of cloned <i>mcbA</i> into pSEVA234	150
Figure 5.12	Agarose gel electrophoresis of PCR amplicons after the deletion of <i>mcbB</i> from the microcin operons	152
Figure 5.13	Plasmid maps of the cloned microcin operons showing deletion of <i>mcbB</i> and incomplete microcin operon	153
Figure 5.14	Plasmid maps of cloned microcin operon deletion mutants of <i>mcbB</i>	156
Figure 5.15	Agarose gel electrophoresis to confirm deletion of <i>mcbB</i> by restriction enzyme digest of pMBB-3 Δ <i>mcbB</i> and pMBB-3	157
Figure 5.16	Agarose gel electrophoresis to confirm deletion of <i>mcbB</i> by restriction enzyme digest of pPY113 Δ <i>mcbB</i> and pPY113	158
Figure 5.17	Plasmid maps of pSEVA234 derivatives with <i>mcbB</i> insert	160

Figure 5.18	Plasmid maps of cloned <i>mcbB</i> into pSEVA234	163
Figure 5.19	Digestion of pPY113 with BamHI and EcoRI to get pBR322	166
Figure 5.20	<i>In vivo</i> activity of microcin production by the cognate and non-cognate microcin operons. <i>E. coli</i> BL21(SUC19) was used as the indicator strain	169
Figure 5.21	Alignment of McbC from <i>E. coli</i> and <i>E. persicina</i> amino acid sequences.....	170
Figure 5.22	Alignment of McbD from <i>E. coli</i> and <i>E. persicina</i> amino acid sequences.....	171

List of tables

Table 2.1	List of PCR primers	46
Table 2.2	List of organisms	48
Table 2.3	List of plasmids	50
Table 3.1	CDS identified in pMccB17 and origin of annotation	63
Table 3.2	Comparison of the genes in <i>mcb</i> operon in pMccB17 (ON989342) with some published sequences of <i>mcb</i> operon	72
Table 3.3	Comparison of the amino acid in MccB17 operon in pMccB17 (ON989342) with some published sequences of MccB17 operon ...	73
Table 3.4	Details of the sequence variation between the ON989342 sequence and the published sequences	75
Table 3.5	Neighbouring genes common to Pbf (<i>yffB</i>)	86
Table 4.1	BLASTp search of <i>eeg</i> located next to the <i>E. persicina</i> microcin operon.....	93
Table 4.2	Functional description and comparison of proteins encoded by microcin operons of <i>E. coli</i> and <i>E. persicina</i>	94
Table 4.3	Assigned names of cloned vectors with and without the <i>eeg</i>	100
Table 4.4	Activity spectrum of EMB17	116
Table 4.5	Analytical run order of the pooled QC EMB17 and MccB17 sample.....	119
Table 4.6	Bioassay result of partial purified microcin	120
Table 4.7	Bioassay for HPLC time-based collection	129
Table 5.1	Pairing selection of microcin biosynthesis gene(s) used in the mix and match experiment	165

List of abbreviations

Abbreviations	Definition
AMPs	Antimicrobial peptides
AMR	Antimicrobial Resistance
CDS	Coding sequence
DMSO	Dimethyl sulfoxide
DNA	Deoxyribonucleic acid
dsDNA	double-stranded DNA
EEG	<i>Erwinia mcb</i> operon extra gene
EMB17	<i>Erwinia</i> microcin B17
EMO	<i>Erwinia mcb</i> operon
HPLC	High-performance liquid chromatography
IPTG	Isopropyl β - d-1-thiogalactopyranoside
LCMS	Liquid chromatography–mass spectrometry
ORF	Open reading frame
PCR	Polymerase chain reaction
qMSD	Quadrupole Mass Spectrometric Detector
RiPPs	Ribosomally synthesized and post-translationally modified peptides
SDS	Sodium dodecyl sulfate
SOC	Super Optimal broth with Catabolite repression
ssDNA	single-stranded DNA
TAE	Tris-acetate-EDTA

Chapter one

INTRODUCTION

1.0 Overview

Antibiotics are antimicrobial substances active against bacteria. The word *antibiotic* was first used in 1941 by Selman Waksman to mean a small molecule produced by a microorganism that prevents the growth of other microorganisms (Waksman, 1947). Penicillin (the first antibiotic, according to Waksman's definition) was discovered in 1928 by Alexander Fleming but was not available for clinical use until 1942 due to production and purification constraints (Fletcher, 1984). This has been followed by the development of different types and classes of antibiotics which have saved countless lives. Unfortunately, resistance to antibiotics emerged within a short period of their discovery, with an average of 13.3 years (calculation based on information provided by CDC) (CDC, 2019). From the information provided by CDC, some antibiotics showed resistance the same year they were introduced. The period between the 1950s and 1970s is referred to as the golden era of antibiotic discovery due to the number of antibiotic classes discovered (Aminov, 2010). Unfortunately, no new class has made it into the clinic since 1987; however, some promising discoveries have the potential of making it into the clinic, such as the discovery of teixobactin in 2015 (Gunjal *et al.*, 2020). Teixobactin is able to kill some Gram-positive bacteria; however, agents targeting Gram-negative organisms are hard to come by due to the nature of Gram-negative organisms' outer membranes (Silver, 2011).

Many bacteria are intrinsically resistant to antibiotics, and the rest are rapidly evolving resistance at an alarming rate. Based on this, it has been predicted that by 2050 annual death due to antibiotic resistant infections will be about 10 million (O'Neill, 2016). Researchers are considering numerous antimicrobial agents such as bacteriophage, probiotic bacteria and bacteriocins as alternatives to the limited antibiotic resources (Pag and Sahl, 2002). The first bacteriocin called 'colicine' because it killed *E. coli* was discovered in 1925 by André Gratia (Gratia, 1925). Bacteriocins are ribosomally synthesised antimicrobial peptides (AMPs), and they are known to be active and effective in the nanomolar range of concentration, and they have no known toxicity at these concentrations (De Vuyst and Vandamme, 1994; Parada *et al.*, 2007).

There are different categories of bacteriocin based on whether they are produced by Gram-positive or Gram-negative bacteria. Microcins are an antimicrobial peptide <10 kDa in weight synthesized by Gram-negative bacteria, especially from the family Enterobacteriaceae (Rebuffatt, 2013). Microcin B17 (MccB17), a bacteriocin synthesised by *Escherichia coli*, is a DNA gyrase inhibitor that acts in a similar way to quinolone antibiotics (such as ciprofloxacin and ofloxacin) while differing slightly in that MccB17 target gyrase subunit B in contrast, quinolones target gyrase subunit A (Heddle *et al.*, 2001).

Microcin B17 is encoded by seven genes on an operon carried on the 69 kb conjugative plasmid, pMccB17. pMccB17 previously known as pRYC17 (Baquero *et al.*, 1978) was originally found in *E. coli* LP12 (Baquero *et al.*, 1978) before it was transferred by conjugation to *E. coli* K-12 (Millan *et al.*, 1985). pMccB17 is an IncFII plasmid (Baquero *et al.*, 1978) in the same category as R100 and R1. Not much is known about this plasmid aside from the fact that it harbours genes that encode the antimicrobial microcin B17.

With the attention of researchers centred on finding new antibiotics or alternatives to antibiotics, this project focuses on homologues of microcin B17 that was discovered in *Erwinia persicina* using bioinformatic approach that started with BlastP search of *E. coli* McbA, as well the plasmid that carries the prototypical *mcb* operon.

1.1 Antibiotics

Antibiotics are a type of antimicrobial that is active against bacteria and used in the treatment and prevention of bacterial infections. As defined by Waksman (1947), antibiotics are small molecules produced by a microorganism which have the ability to prevent the growth of other microorganisms. While many antibiotics are currently made entirely by chemical synthesis, most of the antibiotics in clinical use have their origin from natural products, either (i) being a natural product from a microorganism, (ii) being semi synthetically produced from natural products or (iii) being chemically synthesised based on the structure of natural products (Demain, 2009).

1.1.1 Spectrum of activity of antibiotics

Antibiotics can be broad-spectrum or narrow-spectrum. Broad-spectrum antibiotics can act against a wide range of pathogenic bacteria; in some cases, they act against Gram-positive

and Gram-negative bacteria. They are administered when a bacterial infection is suspected and diagnosed based on symptoms, however, the type of bacterium responsible for the infection and its resistance profile is not known (this is called empirical therapy). Broad-spectrum antibiotics can also be utilised when there is a need to treat infection involving multiple groups of bacteria. The disadvantages involved in using broad-spectrum antibiotics include the development of antimicrobial resistance and the interference of normal microbiota (Rafii *et al.*, 2008). Broad-spectrum antibiotics (for example, ampicillin and tetracycline) do not only target the pathogenic bacteria causing a particular infection, but a wide range of bacteria, some of which can be beneficial; killing some of these beneficial bacteria will alter the normal microbiota which could lead to a disease state (Bull and Plummer, 2014).

On the other hand, narrow-spectrum antibiotics (for example, macrolides and vancomycin) can act against a limited range of species of bacteria, and these are administered when the causative pathogen of an infection has been identified. Since narrow-spectrum antibiotics target specific bacteria, they deal less damage to the microbiota by only killing or inhibiting the pathogenic bacteria and their relatives. Also, there is a lower tendency for resistance to develop because there is less selective pressure on commensal bacteria (Blaser, 2011; Melander *et al.*, 2017).

1.1.2 Antimicrobial resistance

Antimicrobial resistance (AMR) is the outcome when microbes evolve the mechanisms to protect themselves from the effect of drugs or agents designed to kill (CDC, 2022). AMR is a global threat that needs urgent attention; CDC reported that it killed 1.27 million people worldwide in 2019, and O'Neill in 2015 predicted that the number of deaths that will be attributed to AMR may rise to 10 million yearly by 2050 (O'Neill, 2016; CDC, 2022). Antibiotic resistance is a subset of AMR, and it is a specified resistance that is associated with bacteria. Antibiotic resistance can be defined microbiologically or clinically (i) microbiological resistance is defined as antibiotic resistance where the bacteria resist the mechanism associated with specific antibiotics, and (ii) clinical resistance is related to the failure of therapeutic techniques where bacteria become resistant after surviving a treatment that usually inhibits the bacteria (MacGowan and Macnaughton, 2017).

Antimicrobial resistance is a natural process; in bacteria, it can arise when there is a genetic mutation or when one species acquires resistance from another. However, this process can be

encouraged by the extended use of antimicrobials (Dabour *et al.*, 2016; Saha and Sarkar, 2021). In this context, narrow spectrum antibiotics are preferred over broad spectrum antibiotics because they are effective by targeting specific organisms, and as such, resistance is less likely to develop (Gerber *et al.*, 2017). In 2017, World Health Organization (WHO) published a list of antibiotic resistant priority pathogens for which new antibiotics are urgently needed (WHO, 2017). The majority of bacteria on the list are Gram-negative bacteria, especially Enterobacteriaceae, such as extended-spectrum beta-lactamases (ESBLs) producing bacteria and carbapenemase-producing Enterobacteriaceae (CRE). *Escherichia coli*, *Staphylococcus aureus*, *Klebsiella pneumoniae*, *Streptococcus pneumoniae*, *Acinetobacter baumannii*, and *Pseudomonas aeruginosa* have been highlighted as the six pathogens that cause the most antibiotic resistance related deaths (Murray *et al.*, 2022).

Bacteria can resist drugs or antibiotic agents through the following mechanisms.

- (i) Drug inactivation or modification - this could be enzymatic deactivation of the antibiotics, as seen in some penicillin resistance bacteria where β -lactamases are produced to deactivate penicillin or chemical modification of the drug by the addition of functional groups (Reygaert, 2018).
- (ii) Modification of drug target or binding site - bacterial cells are able to synthesise proteins that protect binding sites or that change the conformational shape of the target (Connell *et al.*, 2003).
- (iii) Modification of metabolic pathway - some antibiotic agent inhibits the synthesis of some crucial metabolic pathway precursors, and resistant bacteria can then utilize another precursor and alter the pathway.
- (iv) Reduced drug accumulation - this can be achieved by an increase in active efflux of the antibiotic agent out of the resistant bacterial cell or by an impermeable cell envelope. The antibiotic agents are pumped out of the cell before they do any damage (Aminov and Mackie, 2007).

1.2 Bacteriocins

One of the suggested alternatives to conventional antibiotics is bacteriocins. These are antimicrobial peptides (AMPs) which are ribosomally synthesized and possess an antimicrobial activity that inhibits the growth of closely related (or similar) species (De Vuyst

and Vandamme, 1994; Parada *et al.*, 2007; Cavera *et al.*, 2015). However, in a few cases, it has been reported that bacteriocins are active outside the circle of closely related species (Messi *et al.*, 2001; Lade *et al.*, 2006). Bacteriocins were probably discovered by André Gratia in 1925 when he found antagonisms between Enterobacteriaceae strains (Gratia, 1925) and in 1947, Fredericq and Levine identified the first bacteriocin, colicin (Fredericq and Levine, 1947). Bacteria produce bacteriocins for self-preservation and to have a competitive advantage over other bacteria in their ecological niches. Multiple bacterial species from Gram-negative and Gram-positive produce bacteriocins and have the associated genes forming clusters on plasmids, chromosomes, or transposable elements (Cavera *et al.*, 2015). Bacteriocins vary from bacterial species to species based on their different modes of action, structures, mechanisms of biosynthesis and self-immunity (Zimina *et al.*, 2020). Bacteriocins can interact with sensitive cells in two stages – physical interaction and biochemical interaction. In the first stage which is equivalent to physical adsorption, the bacteriocins are attached to cell-envelope receptors and there is no permanent physiological damage at this stage, removal of the bacteriocins will reverse the damage done. On the other hand, the second stage is brought about when there are biochemical damages over a measurable time and this results into irreversible pathological changes (Bindiya and Bhat, 2016).

1.2.1 Classification of bacteriocins

Bacteriocins are produced by a large number of bacteria that belong to different groups; as a result, there are many classification systems for bacteriocins. The commonest classification is based on cell type and is described below. The three main groups of bacteriocin producers are Archaea, Gram-positive and Gram-negative (Zimina *et al.*, 2020).

1.2.1.1 Bacteriocins from Archaea

Bacteriocins produced by Archaea are referred to as archaeocins which are divided into two classes:

Halocins are produced by members of the orders Halobacteriales. They vary in size from 3.6 to 35 kDa. Halocins less than 10 kDa are classified as peptide halocins, while those greater than 10 kDa are classed as protein halocins. Of the numerous types of halocins, 17 have been studied to date, examples include A4 and C8 halocins (Kumar *et al.*, 2021).

Sulfolobocins are produced by the members of the orders Sulfolobales. They are not secreted into a culture medium; rather, their activity is associated with cells (O'Connor and Shand, 2002).

1.2.1.2 Bacteriocins from Gram-positive bacteria

This group have three classes.

Class I – members are referred to as lantibiotics, with weight ranging between 2 – 5 kDa, and they contain lanthionine (thioether cross-links formed through the addition of a cysteine to a dehydroalanine) or methyllanthionine (thioether cross-links formed through addition of a cysteine to a dehydrobutyrine) which are produced post-transcriptionally to give the bacteriocins diverse structures (O'Connor and Shand, 2002; Zimina *et al.*, 2020).

Lantibiotics are further divided into class IA, which are lantibiotics that interact with cell membranes of the target cells, and class IB, which are immunologically active globular molecules that inhibits enzymes such as phospholipase A2 (O'Connor and Shand, 2002). Examples of class IA are nisin (which is widely used as a food preservative) and subtilin, while mersacidin is an example of class IB (Zimina *et al.*, 2020).

Class II –are unmodified (no post-translational modification), thermostable peptides with weight < 10 kDa. They are subdivided into four subclasses. Subclass IIA are pediocin-like or *Listeria*-active bacteriocin; they form pores in the cell membrane of the target cell and eventually result in the death of the target cell. Examples include Pediocin, enterocin and sakacin (O'Connor and Shand, 2002; Güllüce *et al.*, 2013). Subclass IIB, two-peptide bacteriocins, are comprised of two peptides (heterodimer) which form a strong bacteriocin. Individual peptides show weak activity, and the peptide pairs generally do not share ancestry. Examples of this pore forming bacteriocin include lactococcin G, plantaricin E/F and plantaricin J/K (O'Connor and Shand, 2002; Güllüce *et al.*, 2013). Subclass IIC, leaderless bacteriocins, they lack leader peptides and do not undergo post-translational modifications; a known example is lactococcin B (Meade *et al.*, 2020; Zimina *et al.*, 2020). Subclass IID, these are bacteriocins that are structurally different from subclasses IIA, B and C (not being post-translationally modified), and form a heterogeneous group of peptides (O'Connor and Shand, 2002).

Class III – they are composed of large (greater than 30 kDa) heat labile proteins. They have a different structure and a complex activity where the target cell wall is lysed by the N-

terminal part of the molecule while the C-terminal recognizes the target cell. Examples are helveticin J and acidophilus (Güllüce *et al.*, 2013; Meade *et al.*, 2020).

1.2.1.3 Bacteriocins from Gram-negative bacteria

Bacteriocins from Gram-negative are generally classified into four classes.

Colicins – they are heat sensitive and protease sensitive antimicrobial peptides synthesized by *E. coli* that carry the colicinogenic plasmid, pCol. Colicins are known to have high molecular weights (30 – 80 kDa) (Marković *et al.*, 2022; Zimina *et al.*, 2020). A typical colicin operon is made up of three genes: structural gene, the gene for immunity protein and the gene for lytic protein. Target cells are killed either by pore formation, nuclease degradation or inhibition of peptidoglycan synthesis. Examples include colicins A, B, E2 and E3 (Marković *et al.*, 2022; Zimina *et al.*, 2020).

Colicin-like bacteriocins – these are colicins synthesised by closely related bacteria that have a similar structure, size and function to *E. coli* colicins, for instance, klebicins produced by *Klebsiella* spp. or S-pyocins produced by *P. aeruginosa* (Marković *et al.*, 2022; Zimina *et al.*, 2020).

Phage-tail like bacteriocins – they are also known as tailocins and have large protein structures (20 – 100 kDa) that is made up of eight to fourteen polypeptide subunits which are similar to the tail modules of bacteriophage. A tailocin gene cluster (typically greater than 40 kbp) usually consists of genes encoding structural proteins, assembly enzymes, chaperones, regulatory genes, and lysis cassettes. The mechanism of action of these antimicrobial peptides is not fully understood, but it was suggested that they form pores in the membrane of a target cell. Examples include R- and F-pyocins (Marković *et al.*, 2022; Zimina *et al.*, 2020).

Microcins – Microcins are antimicrobial peptides having a molecular mass of less than 10 kDa and are produced by Enterobacteriaceae during a period of stress, mostly nutrient depletion (Braun *et al.*, 2002). Most microcins are heat, pH and protease stable, are required in nanomolar concentration or range to be active, and they are usually potent on a narrow spectrum (Duquesne *et al.*, 2007; Pavlova and Severinov, 2006; Rebuffat, 2013). Microcins can be encoded by gene clusters carried on plasmids or chromosomes and are ribosomally synthesized. Microcins can be distinguished from each other using their structures and modes of action, these differences are due to the diversity found in microcin gene clusters (genes

encoding microcin precursor, self-immunity factors, secretion/transport proteins and modification enzymes) (Arnison *et al.*, 2013).

According to Baquero *et al.* (2019), there are 15 known microcin peptides, of which 8 have been experimentally studied and structurally characterized. Microcins have been put into two groups based on three criteria: (i) the presence, nature and localization of the post-translational modifications, (ii) the gene cluster organization, (iii) the leader peptide sequences (Duquesne *et al.*, 2007).

Group I – this group consist of microcins that are less than 5 kDa that are subjected to post-translational modification. Examples include MccB17, MccC7/51, and MccJ25 (Duquesne *et al.*, 2007; Rebuffat, 2013).

Group II – they are microcins that have a molecular weight between 5 – 10 kDa; they are divided into two. Members of group IIa do not have post-translational modifications, and some of them contain disulfide bonds. Examples of members of this group are MccL, MccV and MccN. Group IIb microcins show evidence of C-terminal post-translational modification. Examples include MccE492, MccM, MccH47 and MccI47 (Duquesne *et al.*, 2007; Rebuffat, 2013).

1.2.2 Mode of actions of bacteriocins

Research has shown that bacteriocins can utilise some of the clinically validated antibacterial targets (Cavera *et al.*, 2015). These targets include (i) cell wall/peptidoglycan disruption; (ii) bacterial membrane disruption; (iii) inhibition of protein biosynthesis; (iv) inhibition of DNA replication and transcription; and (v) inhibition of folate biosynthesis (Cavera *et al.*, 2015; Penesyan *et al.*, 2015; Lewis, 2013). Four of these targets in addition to some novel pathways are utilized by bacteriocins, as shown in Fig. 1.1 (Cavera *et al.*, 2015).

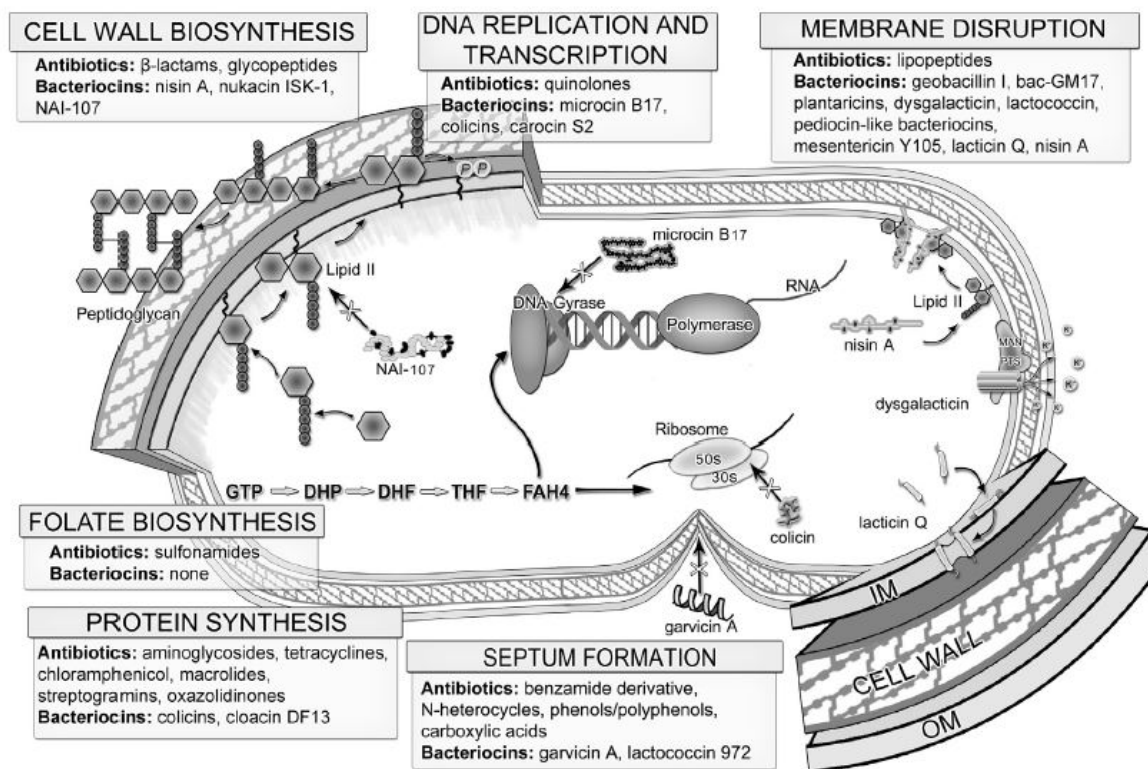


Figure 1.1 Targets of both antibiotics and bacteriocins, their general location, and examples of each capable of inhibition of these targets. Image taken from Cavera *et al.* (2015).

1.2.2.1 Cell wall/peptidoglycan disruption

The cell wall is an important target as it is crucial for bacterial survival in that it maintains cell integrity and morphology, especially during internal osmotic pressure fluctuations; this makes cell wall biosynthesis prevention a critical target. Cell wall synthesis can be targeted at four stages: (i) inhibition of the synthesis of lipid II; (ii) inhibition of the undecaprenyl carrier lipid; (iii) binding of lipid II; and (iv) binding and blocking of the active sites of penicillin-binding proteins (PBPs) (Cavera *et al.*, 2015; Lages *et al.*, 2013). Nisin A produced by *Lactococcus lactis*, nukacin ISK-1 (produced by *Staphylococcus warneri*), and lantibiotic NAI-107 (produced by *Microbispora* sp.) are examples of bacteriocins that inhibit cell wall synthesis by binding to a cell wall precursor, lipid II (Cavera *et al.*, 2015).

1.2.2.2 Bacterial membrane disruption

The bacterial membrane provides selective permeability for cellular homeostasis and metabolic energy-transduction. Destruction of membrane structures and functional impairment can be caused when antibiotic agents interact with some targets in the cell

membrane. For example, Bac-GM17 (produced by *Bacillus clausii* strain GM17) docks to lipid II, and dysgalactin (produced by *Streptococcus dysgalactiae* subsp. *equisimilis* strain W2580) binds to mannose phosphotransferase system (a transmembrane transporter protein complex) and their interaction with the membrane results into pore formation (Cavera *et al.*, 2015).

1.2.2.3 Inhibition of protein biosynthesis

Protein synthesis is an important cellular process that can also be targeted by antibiotic agents. These antibiotic agents bind to the bacterial 30S or 50S ribosomal subunit and inhibit the translocation of peptidyl-tRNA. Colicins E3, E4, and E6 and cloacin DF13 are examples of bacteriocins that utilize this mode of action (Cavera *et al.*, 2015; Kapoor *et al.*, 2017).

1.2.2.4 Inhibition of DNA replication and transcription

Some antibacterial agents, like Fluoroquinolone, target DNA gyrase to inhibit DNA replication by introducing negative supercoils. Microcin B17 produced by *E. coli* is an example of a bacteriocin that targets DNA gyrase (Cavera *et al.*, 2015; Collin *et al.*, 2013; Heddle *et al.*, 2001). On the other hand, some antibacterial agent targets RNA polymerase (RNAP) and inhibits transcription like rifamycin. MccJ25, another type of bacteriocin produced by *E. coli*, can bind to the RNA polymerase (RNAP) secondary channel and inhibits transcription (Telhig *et al.*, 2020).

1.2.2.5 Inhibition of folate biosynthesis

Antibiotic agents target molecules like dihydrofolate reductase (DHFR) and dihydropteroate synthase (DHPS) in the folate synthesis pathway, such as sulphonamides and trimethoprim, respectively. However, there are no bacteriocins that utilize this mode of action (Cavera *et al.*, 2015).

1.2.3 Current and Future Applications of bacteriocins

The wider potential for applications of bacteriocins as antimicrobials is gaining recognition; however, the food industry has found more uses for bacteriocin compared to any other industry. This section will cover the current and proposed uses of bacteriocins.

Lantibiotics are often produced by food-grade bacteria that are known to be safe for human consumption. Nisin has been approved as a food preservative by FDA and EU (Cotter *et al.*, 2005; EFSA ANS Panel *et al.*, 2017). It was suggested that it has the possibility to be used as

a treatment for peptic ulcer disease caused by *Helicobacter pylori* (Delves-Broughton *et al.*, 1996). Other potential uses of nisin include inhibiting the growth of multi-drug resistant pathogens like *Staphylococcus* and *Streptococcus* spp, treatment of catheters and tracheotomy tubes to offer protection against infection caused by Gram-positive bacteria (Severina *et al.*, 1998; Bower *et al.*, 2002). Juvenile acne caused by *Propionibacterium acnei* can be treated with epidermin and gallidermin, produced by *Staphylococcus epidermidis* and *S. gallinarum* respectively (Allgaier *et al.*, 1986; Kellner *et al.*, 1988). Duramycin and cinnamycin (by *Streptoverticillium* and *Streptomyces* spp.) have the potential to be used as an anti-inflammatory drug, and lacticin 3147 (by *Lactococcus lactis*) can be used to combat mastitis-causing bacteria *Streptococci* and *Staphylococci* (Märki *et al.*, 1991; Ryan *et al.*, 1998).

Microcin J25, microcin 24 and colicins E1, E4, E7, E8, K and S4 can be further explored pharmaceutically as they have the capacity to inhibit *E. coli* O157:H7 (the leading cause of haemorrhagic colitis and haemolytic uremic syndrome in humans) (Wooley *et al.*, 1999; Sable *et al.*, 2000; Jordi *et al.*, 2001). Microcin J25 and L can be used to inhibit *Salmonella enterica* serovars typhimurium, one of the causative agents of diarrhoea in humans (Pons *et al.*, 2004; Portrait *et al.*, 1999). Some bacteriocins, such as microcin E294 and colicin E9 have been reported to have anti-cancer properties and thus have the potential to be used as anti-cancer drugs as they can interact with target cell surface and have cytotoxic effect on certain cancerous cells (Hetz *et al.*, 2002; Walker *et al.*, 2002).

Bacteriocins can be used as a biological control to prevent diseases caused by bacterial plant pathogens. For instance, bacterial blight in rice's leaf was minimised by using the bacteriocin producing strain of *Xanthomonas campestris* to treat the leaves, fire blight caused by *Erwinia amylovora* can also be minimised by the colicin-like bacteriocin produced by *Serratia plymithicum* (Jabrane *et al.*, 2002; Sakthivel and Mew, 1991).

1.3 Microcin B17

Microcin B17 (MccB17), is a small (3093 Da), thiazole/oxazole-containing, ribosomally synthesized and post-translationally modified peptide (RiPP) that is produced by Enterobacteriaceae, majorly *E. coli* (Arnison *et al.*, 2013; Duquesne *et al.*, 2007). Their production is enhanced by environmental stresses, usually nutrient depletion or specifically nitrogen starvation, as Connell *et al.* established in their work in 1987. Microcin B17

primary function is presumably to help the producing strain to compete and survive in a stressed environment by attacking the DNA gyrase of competing bacteria. They are known to be stable at high temperatures and extreme pH (Liu, 1994).

Microcin B17 inhibits class II topoisomerases similar to quinolone antibiotics but targets gyrase subunit B, whereas quinolones attack gyrase subunit A (Heddle *et al.*, 2001; Pierrat and Maxwell, 2003). Microcin B17 inhibits DNA supercoiling by the gyrase, which initiates DNA cleavage. With ongoing research into how MccB17 inhibits DNA gyrase, it was proposed that MccB17 binds to the gyrase (A₂B₂)-DNA complex, and this interrupts the strand-passage reaction between the DNA and the gyrase (Collin and Maxwell, 2019).

1.3.1 *Microcin B17 Biosynthesis*

Biosynthesis of MccB17 is carried out by the products of 7 genes (*mcbA – G*) arranged in an operon found on the plasmid pMccB17 and encoding a precursor peptide, synthase, immunity and export functions (Collin and Maxwell, 2019; Garrido *et al.*, 1988; Genilloud *et al.*, 1989; Yorgey *et al.*, 1994). The promicrocin McbA is a 69 amino acid precursor peptide that is post-translationally modified by the microcin B17 synthetase complex (McbB, McbC and McbD), resulting in the formation of heterocycles (thiazole and oxazole) in the sequence (Parks *et al.*, 2007; Rebuffat, 2013). Thiazoles are heterocyclic compounds that contain both sulphur and nitrogen, while oxazoles are compounds with oxygen and nitrogen separated by one carbon. This modified peptide sequence is exported from the host cytoplasm by the product encoded by *mcbE* and *mcbF*. The last gene, *mcbG* (whose product binds to the host DNA gyrase), acts together with *mcbE* and *mcbF* function to offer immunity to the host cell (Collin *et al.*, 2013). MccB17 is one of the best studied microcins and is gaining recognition as a promising template for developing new antibacterial agents (Collin and Maxwell, 2019; Ghilarov *et al.*, 2019).

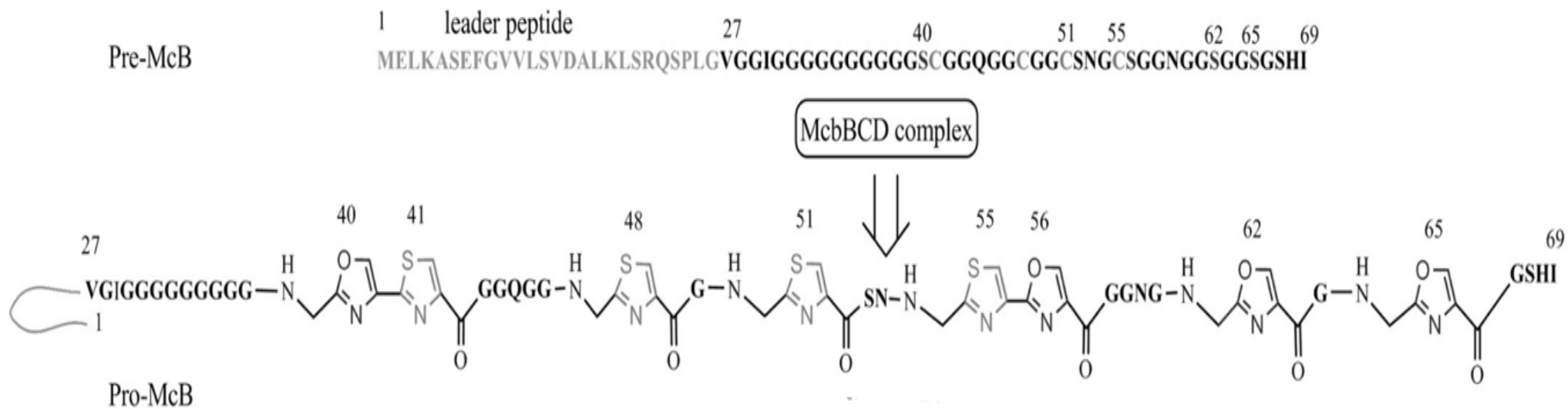


Figure 1.2 Amino acid sequence of MccB17 showing the location of the post-translational modifications. Image taken from (Ghilarov *et al.*, 2011).

The 69 amino acid McbA, a glycine-rich peptide, is made up of two parts: the leader peptide which is a 26 amino acid peptide (N- terminal) and functions as a recognition site for post-translational enzymes (Madison *et al.*, 1997), while the second part forms the mature and active microcin (Fig. 1.2).

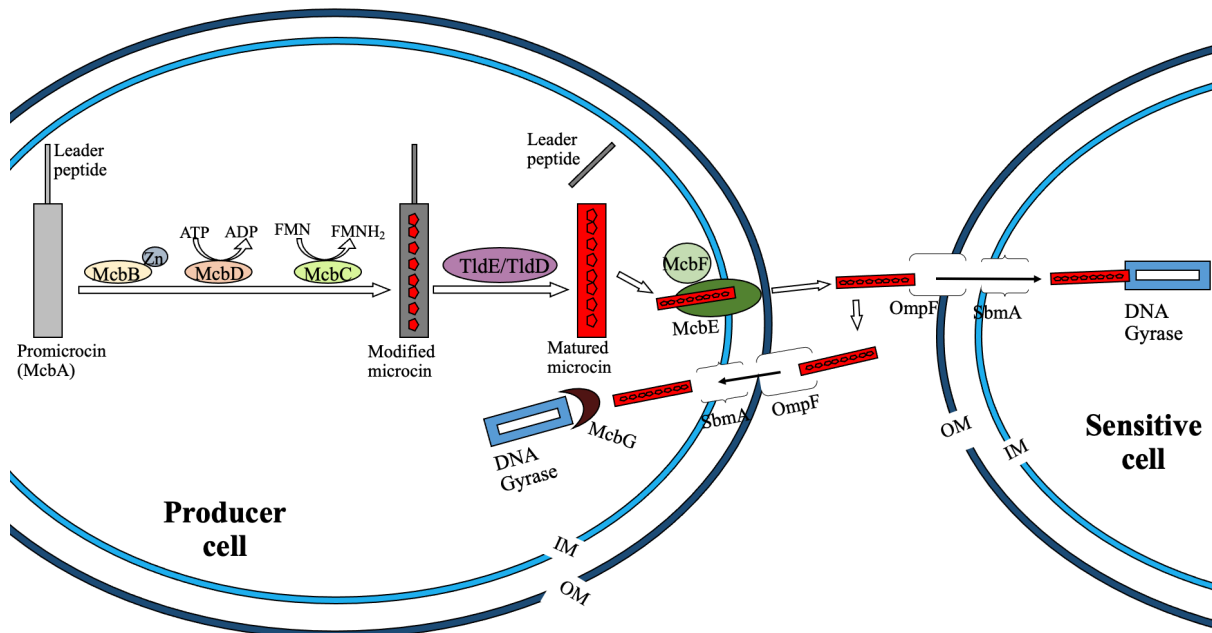


Figure 1.3 Biosynthesis of MccB17 and mechanism of action in sensitive cells. In modified and matured microcin pentagons indicate heterocycles; IM – Inner membrane; OM – Outer membrane. FMN is flavin mononucleotide and FMNH₂ is dianion of reduced flavin mononucleotide. Information to produce this image were from Duquesne *et al.* (2007) and Ghilarov *et al.* (2019).

The product of the synthase complex McbBCD acts in three steps during the post-translational modification, starting with McbB (a zinc-dependent subunit), followed by McbD (an ATP-dependent subunit) and McbC (a flavine-dependent subunit) (Fig 1.3). McbB first binds to the leader peptide and the polyglycine linker (Gly30 to Gly39). The polyglycine linker is proposed to function as a spacer region to allow correct heterocyclisation positioning (Sinha Roy *et al.*, 1998). The ATP-dependent subunit cyclodehydratase McbD subunit is responsible for the actual cyclization reaction. The cysteine and serine residues having a glycine preceding them are cyclized to form thiazole and oxazole respectively. This heterocyclisation is brought about by posttranslational modification (an important feature in

group 1 of Gram-negative bacteriocins). In total, there are eight rings in microcin B17, 2 are thiazole, 2 are oxazole, in addition, there is the formation of a tripeptide between Gly-Cys-Ser and Gly-Ser-Cys to form thiazole-oxazole and oxazole-thiazole rings respectively. McdD cyclizes cysteine and serine to form thiazoline and oxazoline respectively; flavin-dependent McdC acts to desaturate these compounds to thiazole and oxazole. The resultant is a post-translationally modified microcin peptide (Collin and Maxwell, 2019).

A mature microcin is obtained as a result of the cleavage of the leader peptide from the modified microcin by chromosomally encoded TldE and TldD peptidases (Rodríguez-Sáinz *et al.*, 1990). The mature microcin is exported outside the host cell via an ABC-like transporter system by the action of McdE and McdF (Garrido *et al.*, 1988). McdE spans the inner membrane, and McdF contains a nucleotide binding domain typical of ABC transporter proteins. These two proteins work together to pump out MccB17 to the periplasmic space.

MccB17 is usually taken up by susceptible cells that have a microcin receptor in its membrane protein, as MccB17 use porins or high affinity receptors for recognition, as reported by Duquesne *et al.* (2007). On the outer membrane, OmpF is the receptor, OmpF is a passive diffusion pore across the outer membrane, and it is crucial for the uptake of nutrients in nutritionally poor media (Liu and Ferenci, 1998). Nutritionally poor media is likely to enhance susceptibility to MccB17. The protein responsible for the uptake of MccB17 is SbmA. SbmA is an inner membrane found in Gram-negative bacteria and was established to be the most significant factor of sensitivity to several unrelated RiPPs from Gram-negative bacteria as they rely on it for transport (Ghilarov *et al.*, 2021).

1.4 Ribosomally synthesized and post-translationally modified peptide (RiPP)

Microcin B17 belongs to an important group of natural products, the ribosomally synthesized and post-translationally modified peptides (RiPPs), also called ribosomal peptide natural products (RPNPs). RiPPs are linear peptides (less than 10 KDa) that are synthesised by the ribosome and then undergo enzymatic post-translational modifications. The propeptide is made up of a core peptide, an N-terminal leader peptide (that serves as a recognition site by many of the post-translational modification enzymes), and in some cases, a C-terminal follower peptide. There are different post-translational modifications that are involved in the

biosynthesis of RiPPs and are important for the functionality of biological activity of the various RiPPs. These modifications include dehydration, cyclodehydration, cyclization, glycosylation, and phosphorylation (Arnison *et al.*, 2013; Li and Rebuffat, 2020).

1.4.1 Classes of RiPPs

Lanthipeptides – these are modified peptides produced by Gram-positive bacteria and have antimicrobial activity against Gram-positive bacteria but not against Enterobacteriaceae and other Gram-negative bacteria. Lanthipeptides have lanthionine and 3-methylanthionine residues, which characterizes this class. The post-translational modification of the lanthipeptides is in two steps: dehydration of serine and threonine in the core peptide to form dehydroalanine (Dha) and dehydrobutyrine (Dhb) residues, respectively, and cyclisation with cysteines by the formation of thioether linkages. Examples include nisin, actagardine, duramycin, and labyrinthopeptin (Arnison *et al.*, 2013; Li and Rebuffat, 2020).

Linaridins – these modified peptides are similar to lanthipeptides in that they also possess the thioether linkages in their structure, but they are generated by a different pathway, and the cyclization in their structures is formed by two cysteines. Also, they do not possess lanthionine and 3-methylanthionine residues, which is a characteristic feature of lanthipeptides. Cypemycin and epidermin are examples of linaridins (Arnison *et al.*, 2013).

Proteusins – they are also known as polytheonamides, and they are the most extensively modified RiPPs with 48 postribosomal modification steps; in addition, they have a large leader peptide. They are synthesised by an-yet uncultivated bacterium that is in a symbiotic relationship with the sponge *Theonella swinhoei* (Arnison *et al.*, 2013).

Cyanobactins – these are natural products with diverse metabolites isolated from cyanobacteria with N – C macrocyclic peptides. The precursor peptide of cyanobactin is designated as “E” in contrast to other precursor peptides in most RiPPs gene clusters. Some cyanobactins carry bear thiazolines/oxazolines, thiazoles/oxazoles, and methylations depending on the modification enzymes involved in their modifications (Leikoski *et al.*, 2009; Arnison *et al.*, 2013).

Bottromycins – they are mainly produced by *Streptomyces* spp., (for examples *Streptomyces bottropensis* and *Streptomyces scabies*) and they contain a C-terminal decarboxylated thiazole and a macrocyclic amidine (Arnison *et al.*, 2013; Shimamura *et al.*, 2009).

Bottromycins differ from other RiPPs in that they lack an N-terminal leader peptide; instead, there is an extension of 35 – 37 amino acids at the C-terminal of the precursor peptide that acts as a recognition sequence for the post-translational modification enzymes. Examples include bottromycin A₂, B₂, C₂, D, D₂ and E₂ (Gomez-Escribano *et al.*, 2012).

Lasso peptides – these modified peptides have a knotted arrangement in their structure called the lasso fold. Lasso peptides are short peptides that consist of an N-terminal macrolactam macrocycle, and a C-terminal tail is trapped within the macrocycle. This structure makes lasso peptides compact and stable, and they are known to be resistant to proteases and denaturing agents. Microcin J25 is the most extensively studied lasso peptide, other examples include capistrain, propetin, anantin and lariatins A and B (Arnison *et al.*, 2013; Maksimov and Link, 2014).

Linear azol(in)e-containing peptides (LAPs) – these are post-translationally modified peptides that have various combinations of thiazoles and oxazoles, or their reduced form, thiazoline and oxazoline. They are also called thiazole/oxazole-modified microcins (TOMMs). Thiazol(in)es are derived from cysteine residues, while (methyl)oxazol(in)es are derived from serine and threonine. The precursor peptide and the synthetase complex are the important components of this class of RiPPs. Biosynthesis of LAPs commences with substrate recognition, and this takes place at the N-terminal leader peptide of the precursor peptide, after this, the synthetase act on the substrate modifying with the outcome being the formation of the azol(in)e rings (Arnison *et al.*, 2013; Li and Rebuffat, 2020).

Microcins – this is the name given to all RiPPs synthesised by Enterobacteriaceae and having molecular weight <10 kDa. Microcins are generally known to be resistant to proteases, extreme pH and temperature. Their biosynthetic gene cluster has a conserved structure (i) a structural gene that encodes the precursor peptides, (ii) genes encoding modification enzymes or auxiliary proteins, (iii) genes encoding the export system, and (iv) a self-immunity gene. Microcin B17 can be classified as a member of both LAP and microcin classes (Arnison *et al.*, 2013).

1.5 Taxonomy of the Enterobacterales with a focus on *E. coli* and *E. persicina*

Microcins have been extensively studied and found to have a mode of action and structures that can be utilised to help in the fight against antibiotic resistance, as well as in other areas such as influencing human immune responses and as a pro-apoptotic peptide used in anticancer therapy (Baquero *et al.*, 2019). Some literature has stated that microcins are synthesised mainly by Enterobacteriaceae; however, with a recent adjustment to the family Enterobacteriaceae, we will expand this group of microcin producer to include the order Enterobacterales. This is because *E. persicina*, a microcin producer, does not belong to the family Enterobacteriaceae anymore.

1.5.1 Gammaproteobacteria

This is one of the classes of the phylum Pseudomonadota (also known as Proteobacteria). This is an important group of bacteria because a lot of medically, ecologically, and scientifically significant bacteria belong to this group, and within the prokaryotes, it has the highest number of genera, 250, according to Garrity *et al.* (2005). They are found in various terrestrial or marine environments, and some live in extreme environments. Their shapes vary from bacilli to cocci to filaments, and they can exist as free-living bacteria, biofilm formers, commensals, and symbionts. Due to the diverse nature of this group, they have different metabolism capabilities, including aerobic and anaerobic (obligate or facultative) species, photoautotrophs, heterotrophs, chemolithoautotrophs and chemoorganotrophs (Williams *et al.*, 2010).

This group of bacteria plays a major role in maintaining the Earth's ecosystems, as some of its members are involved in nutrient cycling and bioremediation (Evans *et al.*, 2008). They have members that play roles in keeping our environment safe; for instance, some members of Gammaproteobacteria can synthesise Poly- β -hydroxyalkanoate (PHA), which can be used in the production of biodegradable plastics, the phototrophic purple sulphur bacteria can be used to treat wastewater (de la Haba *et al.*, 2010; Kobayashi and Tchan, 1973). A lot of Gammaproteobacteria species can synthesise metabolites with antibacterial properties. A large number of the member of Gammaproteobacteria are pathogens of humans, animals or plants and are of clinical and scientific importance, such as *Escherichia coli*, *Pseudomonas aeruginosa*, *Vibrio cholerae*, *Yersinia pestis*, *Salmonella* spp., *Xanthomonas axonopodis* pv. *citri* and *Xylella fastidiosa*.

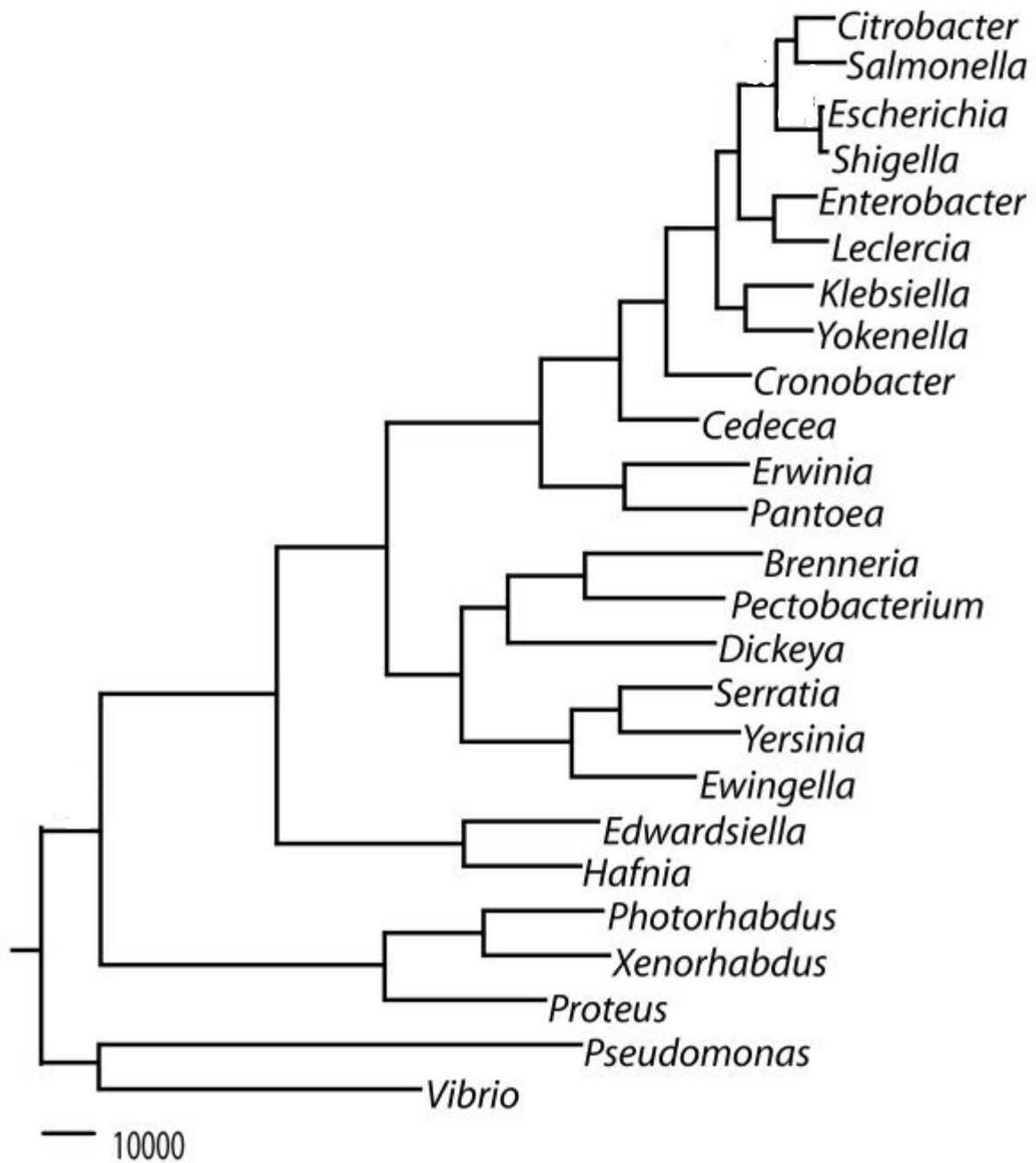


Figure 1.4 Phylogenetic tree of the old Enterobacteriaceae family. The old Enterobacteriaceae family included genera like *Erwinia* before the taxonomy was emended in 2016. Image adapted from Baumlér *et al.* (2013).

1.5.1.1 Enterobacteriales

Enterobacteriales is an order of the class Gammaproteobacteria. Bacteria in this order are Gram-negative, rod-shaped, non-spore forming, facultative anaerobes containing a large, diverse group of species and having an array of biochemical characteristics. Many members of this order are either human pathogens (such as *Escherichia coli* and *Salmonella enterica*) or plant pathogens (such as *Erwinia* and *Pantoea*) (Fig. 1.4).

Before 2016, *Escherichia* and *Erwinia* were classified as belonging to the same family, Enterobacteriaceae, which was the sole family under the Order "Enterobacteriales". However, the taxonomy was emended, "Enterobacteriales" was renamed Enterobacteriales and divided into seven families placing both *Escherichia* and *Erwinia* into different families (Adeolu *et al.*, 2016; Janda and Abbott, 2021). *Escherichia* was assigned to the Family Enterobacteriaceae, while *Erwinia* was assigned to the Family Erwiniaceae. Other families that make up the Order Enterobacteriales are (Budviciaceae, Hafniaceae, Morganellaceae, Pectobacteriaceae and Yersiniaceae (Adeolu *et al.*, 2016).

1.5.1.1.1 Enterobacteriaceae

Enterobacteriaceae are commonly called "enteric bacteria" because a lot of their members live in the intestines of animals; however, some are found in soil and water (Garrity *et al.*, 2005; Ferreira da Silva *et al.*, 2007). Members of this family are non-spore forming, rod-shaped Gram-negative bacteria and range in length from 1 – 5 µm. They have some members that are motile, and they move about using flagella; in addition, they are facultative anaerobes (Edwards and Ewing, 1972). Enterobacteriaceae is one of the families under the order Enterobacteriales, and it includes some harmless members as well as many common pathogens such as *Salmonella*, *Escherichia coli*, *Klebsiella*, *Shigella*, *Enterobacter* and *Citrobacter*.

Prior to being amended the Enterobacteriaceae contained an extensive array of members that have different biochemically distinct species, and as such, the biochemical description of the family was complex. The phenol red, tryptone broth, phenylalanine agar, methyl red, Voges-Proskauer, catalase, or oxidase tests have classically been used to identify genera of this family (Madigan *et al.*, 2006). However, the emendation ensured that the family Enterobacteriaceae must include genera that are directly related to the type genus, which are the enteric species (Adeolu *et al.*, 2016).

Escherichia is a genus of the family Enterobacteriaceae and order Enterobacterales containing Gram-negative, non-spore-forming, facultative anaerobic bacteria. Some species are commensal inhabitants of the gastrointestinal tracts of warm-blooded animals, while some are opportunistic or true pathogens. Their motility is through flagella, and they produce gas from fermentable carbohydrates. Species in the *Escherichia* genus include *E. albertii*, *E. fergusonii*, *E. hermannii*, *E. marmotae* and the model organism and clinically significant *E. coli* (Madigan *et al.*, 2006). Pathogenic strains of *Escherichia* cause urinary tract infections, gastrointestinal diseases (diarrhoea and dysentery), colonic escherichiosis, central nervous system infections (meningitis), bloodstream infections and inflammation in humans (Chaudhury *et al.*, 1999; Ronald, 2003).

Members of this genera are known to be catalase-positive and oxidase-negative. Members of the family Enterobacteriaceae are notable from other members of the order Enterobacterales as they are capable of forming a different monophyletic cluster in genome- and multi-gene-based phylogenetic trees.

1.5.1.1.2 *Erwiniaceae*

Erwiniaceae is one of the new families that was crafted when the old Enterobacteriaceae family was divided into seven novel families based on comparative genomic analyses. Members of this family are in the order Enterobacterales in the class Gammaproteobacteria of the phylum Pseudomonadota, and they mainly include plant pathogens and insect endosymbionts. *Erwinia* is the type genus of this family (Adeolu *et al.*, 2016). Bacteria found in this family are oxidase-negative, catalase-positive, and positive for the Voges-Proskauer test (except for some *Erwinia* species), and they do not produce indole or hydrogen disulphide. The family Erwiniaceae contains eight validly published genus under the ICNP: *Erwinia*, *Buchnera*, *Mixta*, *Pantoea*, *Phaseolibacter*, *Tatumella*, *Wigglesworthia* and *Kalamiella* (Adeolu *et al.*, 2016).

Erwinia is a genus of the family Erwiniaceae and order Enterobacterales containing Gram-negative rod-shaped bacteria that are distantly related to *Escherichia coli*, *Shigella* and *Salmonella*. *Erwinia* are mostly plant pathogens that infect plants and cause four types of infection: (i) rapid necrosis, (ii) progressive tissue maceration (soft rot), (iii) occlusion of vessel elements (vascular wilt), and (iv) gall or tumour formation. For example, *E. amylovora* causes fire blight on apples and pears, and *E. tracheiphila* causes bacterial wilt on cucurbit species (Kado, 2006; Toth *et al.*, 2003). All members of the *Erwinia* genus can ferment

substrates anaerobically, produce acid from sugar and can ferment and are motile with the aid of peritrichous flagella (Kado, 2006).

1.6 Plasmids

The biosynthetic gene cluster for the prototypical microcin MccB17 is known to be carried on a conjugative plasmid rather than the chromosome of its *E. coli* host, and hence it is potentially mobile between species and even more widely between Enterobacteriaceae family. Plasmids are extrachromosomal genetic elements that play a crucial role in horizontal gene transfer in bacteria and archaea. Plasmids can exist independently of the chromosomes and are usually circular but can also be in a linear form, such as in *Streptomyces* (Ahsan and Kabir, 2013; Willey *et al.*, 2008). They contribute to genetic diversity, evolution, and adaptation in bacteria by facilitating the exchange of genetic materials which confer functions like heavy metal resistance, bacteriocin production, antibiotic resistance and virulence on a recipient bacterium (Brooks *et al.*, 2019; Smillie *et al.*, 2010). There are naturally occurring plasmids, as well as those that have been designed to serve as vectors for gene cloning. When used as a vector, plasmids can introduce, modify, or delete genes, making plasmids an essential tool for molecular biology and genetic engineering (Shintani *et al.*, 2015). Naturally occurring plasmids are known to have “backbone” genes which are also referred to as essential features; these include the replication genes, genes that encode stable inheritance functions, and genes that encode addiction systems. The genes that encode mobilization and mating pair formation functions form the secondary core genes, while other genes that play a role in the phenotype (antibiotic resistance, metabolic properties (Thomas and Summers, 2020)) of the plasmid are termed accessory or “cargo” genes (Thomas and Frost, 2014).

1.6.1 Plasmid replication systems

There are basically three types of replication systems for circular plasmids: theta type, rolling circle (RC), and strand displacement. Linear plasmids, on the other hand, can replicate through concatemeric intermediate mechanism or by protein-priming mechanism (del Solar *et al.*, 1998).

1.6.1.1 *Theta Replication*

Theta replication involves the parental strands melting, synthesis of primer RNA (pRNA), and initiation of DNA synthesis by the pRNA. One strand of DNA is the leading strand as DNA synthesis is continuous while it is discontinuous on the other strand, the lagging strand (del Solar *et al.*, 1998; Zavitz and Marians, 1991). Synthesis of both strands is believed to be coupled, being carried out by two core PolIII enzymes, with one enzyme continuously replicating the leading strand and the second enzyme replicating the lagging strand discontinuously (in segments) (Zavitz and Marians, 1991). Theta replication can be unidirectional or bidirectional and often requires a plasmid-encoded Rep initiator protein (del Solar *et al.*, 1998; Kim *et al.*, 2020). During conjugation, the donor plasmid replicates via a rolling circle mechanism (Clark and Pazdernik, 2013). Examples of plasmids that uses the theta replication are R1, RK2, R6K, pSC101, pPS10, F, and P (Lilly and Camps, 2015)

1.6.1.2 *Strand displacement replication*

DNA replication in this system is initiated by three plasmid-encoded proteins (RepA, RepB, and RepC). Replication can proceed in either direction by a strand displacement mechanism (Scherzinger *et al.*, 1991). Only one strand is replicated at a time which in turn is copied into a double-stranded DNA. This is common in the IncQ family, whose example is RSF1010 (del Solar *et al.*, 1998).

1.6.1.3 *Rolling circle replication*

This is a unidirectional replication system with the synthesis of the leading strand and lagging strand being uncoupled, making the process an asymmetric one (del Solar *et al.*, 1998). RC replication is initiated by a plasmid-encoded Rep protein which nicks one strand of the double-stranded, circular DNA molecule at a region termed double-stranded origin (DSO). The 3' OH end is free which serves as a primer for the leading-strand synthesis by DNA polymerase III, while the 5' phosphate end is bound to the initiator protein. The unnicked DNA is used as a template, and replication continues around the circular DNA, displacing the nicked strand as single stranded DNA in the process. The linear lagging strand can be made double stranded in some steps with help from RNA polymerase and DNA polymerase III (del Solar *et al.*, 1998; Ruiz-Masó *et al.*, 2015). Examples of plasmid that uses RC replication are pRQ7, pMC24, pRKU1, pGT5, pT181, pC194, pMV158 and pCG4 (Ruiz-Masó *et al.*, 2015).

1.6.2 Classification of Plasmids

Plasmids can be grouped into different classes based on parameters such as incompatibility, conjugative ability, host range and function of the plasmid.

1.6.2.1 Incompatibility group

Incompatibility (Inc) features of plasmids mean that two or more plasmids from the same Inc group cannot stably co-exist in a cell (Novick, 1987; Velappan *et al.*, 2007). This is brought about by the plasmid replication system or partitioning system of the plasmids involved (Novick, 1987). Incompatibility was first described in the 1970s by Datta and Hedges (1973). There are at least 27 different incompatibility (*Inc*) groups recognized in *Enterobacteriaceae* (Carattoli *et al.*, 2018; Mutai *et al.*, 2019). Some of these include IncF, IncA, IncII, IncB/O, IncK, IncHI1, IncHI2, IncL/M, IncN, IncP, IncQ, IncT, IncU, IncW, IncX, IncY (Couturier *et al.*, 1988).

1.6.2.1.1 IncF plasmids

IncF plasmids are usually low copy number plasmids found in *Enterobacteriaceae*. The following are essential for their replication; DNA gyrase, DnaB, DnaC, DnaG, single strand binding and DNA polymerase III proteins (Villa *et al.*, 2010). On most occasions, IncF carries more than one replicon, and this has been reported to help narrow host range to accomplish broad host range replication, which in the long run has helped them achieve great intracellular adaptation, which has contributed to why *Enterobacteriaceae* are clinically relevant (Johnson *et al.*, 2007; Osborn *et al.*, 2000).

1.6.2.2 Conjugative functions

Conjugation is the exchange of genetic materials between bacterial cells, which takes place via a pilus. It is a type of horizontal gene transfer. One bacterium is termed the donor of the genetic material, while the other is the recipient. The donor has the fertility factor or F-factor, which allows the donor to produce a tubelike structure, a pilus, through which the genetic material moves to the recipient. Conjugative features of plasmids deal with the transmissible (mobility) of the plasmid to another cell. A plasmid can be classified as a conjugative plasmid or a non-conjugative plasmid. Conjugative plasmids possess the *Tra* operons, which are divided into two: *Tra1* region (*traABCDEFGHIJKLMNPQRSTUVWXYZ*) which is involved in DNA processing and transfer; and *Tra2* (*trbBCDEFGHIL*) which is involved in mating pair formation (Virolle *et al.*, 2020; Zatyka and Thomas, 1998). Functions of these genes include

assembly of conjugative pili, which helps to initiate contacts between mating pairs and assist the transfer of genetic information; inner membrane, periplasmic and surface exclusion proteins production; DNA transfer and mating pair stabilization. Examples of conjugative plasmids are RK2, pMcB17 and F plasmid (Mayo *et al.*, 1988; Loh *et al.*, 1989; Norman *et al.*, 2009). Non-conjugative plasmids can use the transfer system of a co-resident conjugal plasmid to transfer genetic material from the host cell to the recipient cell or through the process of conduction where the plasmid enters the recipient cell along with a conjugal plasmid (Yin and Stotzky, 1997). Examples of non-conjugative plasmids are RSF1010 and ColE1 (Zhou and Li, 2015).

1.6.2.3 Host range

Based on the host type that supports the survival and maintenance of bacterial plasmids, Datta and Hedges (1972) classified bacterial plasmids into two types. Plasmids where the replicon or replicating system can only be maintained in a single or closely related bacterial host are referred to as “narrow-host-range”, while plasmids that are able to replicate in different ranges of bacterial genera that are unrelated are referred to as “broad-host-range” (Zhong *et al.*, 2005). IncF plasmids are examples of narrow-host-range plasmids as they are usually limited to host species from the family Enterobacteriaceae (Villa *et al.*, 2010). The pMccB17 plasmid from this project is also an example of a narrow-host-range plasmid. IncP plasmids are mainly broad-host-range plasmids as they can be transferred and replicated in a wider range of hosts belonging to Alphaproteobacteria, Betaproteobacteria and Gammaproteobacteria (Popowska and Krawczyk-Balska, 2013; Yano *et al.*, 2013).

1.6.3 Plasmid copy number

In naturally occurring plasmids, copy numbers can be low or high, ranging from 1-15 per cell for low copy number plasmids to 200 copies for high copy number plasmids (Ilhan *et al.*, 2018), while for synthetic plasmids (vectors) such as pUC19, it can be as high as 1000 (Boros *et al.*, 1984). Copy number control is very important to ensure the survival of the plasmid in the host progeny because all the critical genes essential for regulation of plasmid replication are on the plasmid, so it is important that at least a copy is present in the daughter cell (Watve *et al.*, 2010). High copy number plasmids end up in both daughter cells randomly with certainty that there will be at least one plasmid in each daughter cell; however, a greater metabolic burden is exerted on the host than a low copy number plasmid. Low copy number plasmids have a high probability of one daughter cell being plasmid-less as opposed to a high

copy number plasmid. In order to avoid or minimize plasmid-less daughter cells, bacterial plasmids have some maintenance systems in place, such as post segregation killing system and partitioning system.

1.6.4 Plasmid encoded maintenance systems

These are features possessed by plasmids to ensure that daughter cells each receive at least one plasmid. There are three types of stable maintenance systems commonly found in bacterial plasmids: multimer resolution system, active partitioning system, and addiction system. Low copy number plasmids may require all three systems to guarantee that plasmids are transferred to daughter cells (Guynet and de la Cruz, 2011; Sengupta and Austin, 2011). These are summarised below.

1.6.4.1 Multimer resolution system

Plasmid dimers or multimers can arise during or after cell replication, and this has a negative effect on plasmid maintenance by decreasing the number of plasmid copies available to daughter cells. This is dealt with by changing the multimers to monomers utilizing a recombinase system which comprises specific recombinase genes and a recombination site for their action (Sengupta and Austin, 2011; Thomas and Summers, 2020).

1.6.4.2 Active partitioning systems

The partitioning (*par*) system is a mechanism employed by plasmids to ensure that daughter cells inherit copies of plasmids in a better way than random distribution. The copies of the plasmid are moved to the two ends of the plane of cell division before the commencement of cell division; in this way, daughter cells are guaranteed a copy of the plasmid (Thomas and Frost, 2014). A related mechanism also exists in most bacteria to ensure correct partitioning of the bacterial chromosome, but it is referred to as chromosome segregation (Jalal and Le, 2020). This system is present in many low copy number plasmids (Fedorec *et al.*, 2019), while high copy number plasmids do not encode partition systems (Baxter and Funnell, 2014). The system is composed of three components: a nucleoside triphosphate (NTPase), a centromere binding protein (CBP) and a centromere-like site on the DNA (Vecchiarelli *et al.*, 2013). Partitioning systems generally work in three steps. In the first step, the centromere binding proteins (CBPs) bind to the centromere repeats (or the partition binding site or

centromeric sites) to form the partition complex (segrosome), the partition complex initiates the nucleotide triphosphate (NTPase) in the second step, and the NTPase works to bring about the separation of the plasmid to opposite poles of the bacterial cell (Pióro and Jakimowicz, 2020; Schumacher, 2012). Schumacher (2012) in their work described three types of *par* system- their classification was based on the type of NTPase (Gerdes *et al.*, 2000):

1.6.4.2.1 *Type I partition system*

The NTPase encoded in this system is a walker-type ATPase. An example is the ParABS system which is the most common type of partition system, and it is usually found in low copy number plasmids (Gerdes *et al.*, 2000). Based on the size and sequence of CBPs and NTPase, type I is further divided into two, type Ia and type Ib (Schumacher, 2012). An example of type Ia is found in *E. coli* plasmid P1 (ParABS) and plasmid F (SopABC), while type Ib is found in *Salmonella enterica* plasmid TP228 (ParFGH) and *E. coli* plasmid pB171 (ParABC) (Baxter and Funnell, 2014).

1.6.4.2.2 *Type II partition system*

The NTPase present in the type II system is described as an actin-like ATPase because of its filamentous nature (Salje *et al.*, 2010). ParM filaments help transport the attached plasmids in reverse directions, thereby driving the DNA apart. Examples of this system are found in the *E. coli* R1 plasmid (ParMRC), *S. aureus* plasmid pSK41 and *B. subtilis* pLS20 plasmid (Baxter and Funnell, 2014).

1.6.4.2.3 *Type III partition system*

The NTPases in type III are tubulin-like GTPase, with examples found in *B. anthracis* plasmid pXO1 (TubZRC) and *B. thuringiensis* plasmid pBtoxis (Baxter and Funnell, 2014).

The CBPs in type Ib, II and III autoregulate the partition operons at the transcriptional level, while in type Ia, this is performed by the NTPase (Bingle and Thomas, 2001; Schumacher, 2012). Guynet and de la Cruz (2011) suggested a type IV partition system with an example in plasmid R388 (StbABC). StbA is a binding protein similar to ParM (typeII); StbA is a walker type ATPase (type I), while the function of StbC is unknown.

1.6.4.3 Post segregation killing systems

There are times when one of the daughter cells after binary fission, does not take up the plasmid even with a mechanism such as the partition system in place. When this happens, the plasmid has systems that kill this type of plasmid-less cells. This mechanism is called a post segregation killing system, and it works by the mechanism of toxin-antitoxin (TA). The plasmid produces a toxin whose effect (preventing proliferation or viability) is cancelled by the antitoxin also encoded by the plasmid (Hernandez *et al.*, 2015). The toxin is a stable protein component, whereas the antitoxin is less stable and can be either a protein or regulatory RNA. The difference in stability brings about the activation of the TA system (Fedorec *et al.*, 2019; Hernandez *et al.*, 2015). In cells that have the plasmid, there is a balance between the effect of toxin and antitoxin produced. However, in daughter cells that have no plasmid, this balance is not maintained, and the effect of the toxin will be expressed, which can be bactericidal or bacteriostatic (Fedorec *et al.*, 2019).

There are five main types of toxin-antitoxin systems based on the nature and activity of the antitoxins.

- Type I antitoxins are RNA based and inhibit the translation of the toxin. An example of this is the *hok-sok* system in *E. coli* (Hernandez *et al.*, 2015).
- Type II – in this case, the antitoxin is protein based and directly neutralizes the toxins. This is present in plasmid R1 (*kis-kid* system), plasmid F (*ccd* system) and plasmid RK2 (*parDE*) (Hernandez *et al.*, 2015).
- Type III antitoxins are also RNA based, like type I, but they act by neutralizing the toxin. This is found in *Pectobacterium atrosepticum* (*toxIN*) (Hernandez *et al.*, 2015).
- Type IV antitoxins are also proteins but does not associate directly with the toxin; instead, it interacts with the target of the toxin, thereby protecting it from the effect of the toxin (Masuda *et al.*, 2012).
- Type V – the antitoxin is a protein and functions by cleaving and disabling the mRNA of the toxin (Wang *et al.*, 2012).

1.7 Aims and objectives of this work

A major aim of this research was to characterize, purify, and genetically modify bacterial microcin peptides produced by *E. persicina*. More specifically, we will attempt to: (i) Clone the microcin operon of *E. persicina* into an *E. coli* host (mostly completed by M. Boutflower during a previous UG research project) and investigate the antimicrobial activity of the cloned genes. (ii) Purify the *E. persicina* microcin from the *E. coli* host and assess its spectrum of activity. (iii) Carry out “mix and match” analyses between *E. persicina* and *E. coli* microcin operon genes, e.g. swapping *E. coli mcbA* for *Erwinia mcbA*. The corresponding research questions are: (i) Does the *Erwinia* operon encode a functional microcin, and if so, what are its functional properties? (ii) What is the specificity of the synthase systems from *E. coli* and *Erwinia*, e.g. can the *Erwinia* synthase complex act on the *E. coli* McbA propeptide and vice-versa?

An additional aim was to characterise the conjugative plasmid that carries the prototypical *E. coli* McB17 operon by sequencing the plasmid genome, completing bioinformatic analysis and plasmid genome annotation. The corresponding research question was: how does the genetic context of the two operons compare?

Chapter two

Materials and methods

2.1 Materials

2.1.1 Buffers

Tris-acetate-EDTA (TAE) buffer: 40 mM Tris-acetate, 1 mM EDTA, pH 8.5. TAE was used as a solvent to dissolve agarose gel and as the running buffer during gel electrophoresis.

Tris: 10 mM Tris-HCl Buffer pH 7.4. This was used to elute DNA during plasmid extraction.

2.1.2 Media

Super Optimal broth with Catabolite repression (SOC): 20 g/l Tryptone, 5 g/l Yeast extract, 5 g/l NaCl. This is a rich medium used in the final step of *E. coli* transformation to obtain maximum transformation efficiency (Hanahan, 1983).

Lysogeny Broth (LB) Miller's formulation: 10 g/l tryptone, 5 g/l yeast extract, 10 g/l NaCl, pH = 7.5. This rich medium was used to grow bacteria cells, especially for molecular biology procedures (Bertani, 1951).

M63: 13.6 g/l KH_2PO_4 , 2 g/l $(\text{NH}_4)\text{SO}_4$, 0.5 mg $\text{FeSO}_4 \cdot 7\text{H}_2\text{O}$, 0.2 mM $\text{MgSO}_4 \cdot 7\text{H}_2\text{O}$, 0.2% glucose, 0.2% yeast extract. This minimal medium was used for bioassay and purification procedures (Atlas, 2010).

Soft agar: 0.6% w/v Bacteriological agar. This was used as an overlay during the bioassay experiments.

Media were made up in distilled water and autoclaved for 15 min at 121° C, 15 psi. Some components (such as glucose) that are heat sensitive were filter sterilised before adding them to the autoclaved media.

2.1.3 Enzymes

Restriction enzymes: These enzymes cleave DNA near or at specific recognition sites and produce fragments. As a result, they can be used to modify DNA as well as mapping DNA. The following Type II restriction enzymes (these cleave DNA at fixed positions in relation to the recognition sequence) from NEB were used in this project: HindIII, XmnI, AleI, EcoRI, Sall, BamHI and KpnI.

DNA Polymerase: This enzyme plays a role in synthesising new DNA molecules by adding nucleotides to the 3' end of a DNA strand. This is used in PCR to make several copies of a gene or DNA. The polymerase used in this project were Q5 High-Fidelity DNA Polymerase (NEB) which was generally used when PCR products were to be cloned or used in plasmid construction and IMMOLASE DNA Polymerase (Bioline).

T4 DNA Ligase: this is an enzyme that catalyses the formation of phosphodiester bonds between juxtaposed 5' phosphate and 3' hydroxyl termini in duplex DNA or RNA. It functions to join DNA with blunt and cohesive end termini. Blunt/TA Ligase Master Mix (NEB) was used for blunt end termini, and for cohesive end termini, Instant Sticky-end Ligase Master Mix (NEB) was used.

T4 polynucleotide kinase (T4 PNK): is an enzyme that catalyses the transfer of gamma phosphate from ATP to the 5' -hydroxyl terminus of polynucleotides or mononucleotides. This enzyme is used to phosphorylate RNA, DNA, and synthetic oligonucleotides. T4 polynucleotide kinase from Thermo Scientific was used in this project.

Alkaline phosphatase: this is an enzyme that catalyses the removal of phosphate groups from DNA, RNA, nucleotides, and proteins. It is used to dephosphorylate cloning vectors to prevent recircularization during ligation. Fast AP thermosensitive alkaline phosphatase (Thermo Scientific) was used in this project.

2.1.4 Oligonucleotides

All PCR primers were manufactured by Integrated DNA Technologies (IDT), Belgium. The stock solutions were 100µM in IDTE buffer, working stocks of 10µM were prepared by diluting the stock solution with nuclease free water. Both of these solutions were stored at -20°C.

Table 2.1 List of PCR primers

pMCCB17 gap-closing primers	
pMccB17-Nd1.27-F	CTGATACAGCCACAACCTGC
pMccB17-Nd1.27-R	CGTTTGAGAACCTTCGCCG
pMccB17-Nd1.52-F	CGATACAGCCACAACCTGC
pMccB17-Nd1.52-R	GTTTGAGAACCTTCGCCG

pMccB17-Nd2.1-F	AGAAATACCGGATTCAGTCCG
pMccB17-Nd2.1-R	GAAATCAGGACTGTCCTGGC
<i>mcbA</i> amplification primers	
EcoR-Ep-mcbA-F	GCGAATTCTCTAGAGAAAGAGGAGAAATACTAGATGGAA TTAAACACGACTGAATTTGGTG
Sal-Ep-mcbA-R	TACGTTAGTCGACTCAGATATGCGAACCACTCCCACCGCT
EcoR-Ec-mcbA-F	GCGAATTCTCTAGAGAAAGAGGAGAAATACTAGATGGAA TTAAAAGCGAGTGAATTTGGTG
Sal-Ec-mcbA-R	TACGTTAGTCGACTCAGATATGTGAACCACTCCGCCGCT G
<i>mcbA</i> deletion primers	
del-mcbA-Ep-F	TGACACGCTGAACTGACCGATC
del-mcbA-Ep-R	AAATTCAGTCGTGTTTAATTCCATAAG
del-mcbA-Ec-F	TGATACGTTGAATTAACCGTTCAG
del-mcbA-Ec-R	AAATTCAGTCGCTTTTAATTCCATAAG
<i>mcbB</i> amplification primers	
EcoR-Ep-mcbB-F	GCGAATTCTCTAGAGAAAGAGGAGAAATACTAGATGCTG AATATTCTGCCATTTGAG
Sal-Ep-mcbB-R	TACGTTAGTCGACTTACCTCTCCACGCAACCACAAG
EcoR-Ec-mcbB-F	GCGAATTCTCTAGAGAAAGAGGAGAAATACTAGATGGTG CTCCCTGATATTA AAAAAGG
Sal-Ec-mcbB-R	TACGTTAGTCGACTTATCTCTCCAGACAGCTGCAAGCCTG
<i>mcbB</i> deletion primers	
del-mcbB-Ep-F	GGTTGCGTGGAGAGGTAAAATG
del-mcbB-Ep-R	CTCAAATGGCAGAATATTCAGCATATT
del-mcbB-Ec-F	AGCTGTCTGGAGAGATAACATG
del-mcbB-Ec-R	CC AAA CGG CAG AAT ATT GAT CAT ATC
Sanger sequencing primers for pSEVA234 clones	

ptcr-for	TCATCCGGCTCGTATAATG
M13-F24	CGCCAGGGTTTTCCCAGTCACGAC
Sanger sequencing primers for pMBB-3Δ<i>mcbA</i>	
B-Ep-mcb-US-F	ATCGGGATCCATTTTCATAGCCCTCCGTCCCTGC
Primers for checking mutation in <i>mcbB</i>	
mcbB-for	GATATGTGAACCACTTCCGCCGCTG
mcbB-rev	CACAGATCCTGAAGTTCTGGCC
Primers for checking mutation in <i>mcbD</i>	
mcbD-for	GGCGCTACAGTCCTCTGATCCAG
mcbD-rev	GGAACAAATCAGAACCCCTTTCAGC
Primers for checking mutation in <i>mcbF</i>	
mcbF-for	CGTCATCAGAGACAACCTCCTGG
mcbF-rev	CACAAAACGACACCTCAGTGAAAGTG

2.1.5 Bacterial strains and plasmids

Table 2.2 List of organisms

Strain	Description	Source
<i>E. coli</i> K-12 strain RYC1000 [<i>F-araDI39AlacUI69Arib-7 rpsL relA thiA recA56</i>]	Host of the plasmid pMccB17	Provided by Professor Roberto Kolter (Harvard Medical School, USA)
<i>E. coli</i> K-12 BW25113 [<i>rrnB3 ΔlacZ4787 hsdR514 Δ(araBAD)567 Δ(rhaBAD)568 rph-I</i>]	Keio knockout parent strain	Purchased from Horizon Discovery
<i>E. coli</i> K-12 BW25113Δ <i>tldD</i> [<i>rrnB3 ΔlacZ4787 hsdR514</i>]	Keio knockout strain	Purchased from Horizon Discovery

<i>Δ(araBAD)567 Δ(rhaBAD)568 rph-1 Km^R</i>		
<i>E. coli</i> K-12 BW25113Δ <i>ompF</i> [<i>rrnB3 ΔlacZ4787 hsdR514 Δ(araBAD)567 Δ(rhaBAD)568 rph-1 Km^R</i>]	Keio knockout strain	Purchased from Horizon Discovery
<i>E. coli</i> K-12 BW25113Δ <i>sbmA</i> [<i>rrnB3 ΔlacZ4787 hsdR514 Δ(araBAD)567 Δ(rhaBAD)568 rph-1 Km^R</i>]	Keio knockout strain	Purchased from Horizon Discovery
<i>Pseudomonas aeruginosa</i> (NCIMB 8626)	Gram negative organism to study the activity range of the microcins	Provided by University of Sunderland Microbiology lab.
<i>Klebsiella pneumonia</i> (NCTC 10314)	Gram negative organism to study the activity range of the microcins	Provided by University of Sunderland Microbiology lab.
<i>Salmonella enterica</i> Tinda (NCTC 8147)	Gram negative organism to study the activity range of the microcins	Provided by University of Sunderland Microbiology lab.
<i>Staphylococcus epidermidis</i> (NCTC 11047)	Gram positive organism to study the activity range of the microcins	Provided by University of Sunderland Microbiology lab.
<i>Staphylococcus aureus</i> (NCTC 10788)	Gram positive organism to study the activity range of the microcins	Provided by University of Sunderland Microbiology lab.
<i>Bacillus subtilis</i> 168 (ATCC 23857)	Gram positive organism to study the activity range of the microcins	Provided by University of Sunderland Microbiology lab.
<i>E. coli</i> BL21 (DE3) [F ⁻ <i>ompT gal dcm lon hsdS_B(r_B⁻m_B⁻) λ (DE3 [lacI lacUV5-T7p07 ind1 sam7 nin5]) [malB⁺]_{K-12}(λ^S)]</i>	Bioassay indicator <i>E. coli</i> strain	Purchased from NEB

<i>E. coli</i> DH5 α [F ⁻ ϕ 80 <i>lacZ</i> Δ M15 Δ (<i>lacZYA-argF</i>) U169 <i>recA1 endA1 hsdR17</i> (r _K ⁻ , m _K ⁺) <i>phoA supE44 λ-thi-1 gyrA96 relA1</i>]	<i>E. coli</i> strain used for transformation experiments	Purchased from NEB
<i>E. coli</i> CC118 [F ⁻ Δ (<i>ara-leu</i>), <i>araD</i> , Δ <i>lacX74</i> , <i>galE</i> , <i>galK</i> , <i>phoA</i> , <i>thi1</i> , <i>rpsE</i> , <i>rpoB</i> , <i>argE</i> (Δ m), <i>recA1</i>]	<i>E. coli</i> strain that carried the SEVA234 plasmid.	Provided by SEVA Collection Team

Table 2.3 List of plasmids

Plasmids	Description	Source
pMccB17	Original plasmid carrying the <i>mcb</i> operon	Provided by Professor Roberto Kolter (Harvard Medical School, USA)
pMBB-1	Plasmid carrying <i>emo</i>	Bingle and Boutflower
pMBB-2	Plasmid carrying <i>emo</i>	Bingle and Boutflower
pMBB-3	Plasmid carrying <i>emo</i>	Bingle and Boutflower
pMBB-4	Plasmid carrying <i>emo</i>	Bingle and Boutflower
pPY113	Expression plasmid of <i>mcb</i> operon	Provided by Professor Roberto Kolter (Harvard Medical School, USA)
pMBB-3 Δ <i>mcbA</i>	A version of pMBB-3 lacking <i>mcbA</i>	This project
pMBB-3 Δ <i>mcbB</i>	A version of pMBB-3 lacking <i>mcbB</i>	This project
pPY113 Δ <i>mcbA</i>	A version of pPY113 lacking <i>mcbA</i>	This project
pPY113 Δ <i>mcbB</i>	A version of pPY113 lacking <i>mcbB</i>	This project
pSEVA234	Cloning vector	Provided by SEVA Collection Team
pSEVA234 <i>mcbA</i> EP	Plasmid carrying <i>mcbA</i>	This project
pSEVA234 <i>mcbA</i> EC	Plasmid carrying <i>mcbA</i>	This project
pSEVA234 <i>mcbB</i> EP	Plasmid carrying <i>mcbB</i>	This project
pSEVA234 <i>mcbB</i> EC	Plasmid carrying <i>mcbB</i>	This project
pUC19	Cloning vector	Purchased from NEB

pUC18	Cloning vector	Purchased from NEB
-------	----------------	--------------------

2.2 Methods

2.2.1 Isolation of plasmid DNA

This project used the QIAGEN Plasmid midi Kit and Monarch Plasmid Miniprep Kit (NEB) to purify plasmids, following the manufacturer's recommended protocol. The kits protocol is based on the alkaline lysis methods of Birnboim and Doly (1979). The process starts by growing and harvesting cells when sufficient growth has been achieved. The harvested cells are resuspended in a solution that contains EDTA, Tris, glucose, and RNase A. The lysis buffer has two key components: NaOH and SDS, which break down the cell wall and denature DNA and proteins. A neutralizer containing potassium acetate decreases the mixture's alkalinity and favours re-nature of small ssDNA to dsDNA. The single-stranded gDNA and denatured proteins form a precipitate, while the dsDNA easily dissolves in solution. These are separated using centrifugation and solid phase extraction (spin columns) to cleanse the DNA further, following manufacturer's recommendations. The quality (and quantity) of the plasmid DNA were assessed by agarose gel electrophoresis by comparing the sample DNA intensity to that of a sized DNA quantitation standard. In addition, absorbance method were used, where the DNA concentration was assessed using UV light at 260 nm and the purity is assessed using the ratio of absorbance at 260 nm and 280 nm, by utilising a microvolume spectrophotometer (Nanodrop Lite, Thermo Fisher).

2.2.2 Illumina sequencing of pMccB17 DNA

The plasmid, pMccB17, was sequenced twice by MicrobesNG (Birmingham, UK; <https://www.microbesng.com/>) using a standard Illumina sequencing platform with 2x250 bp paired-end reads, firstly together with the host bacterial chromosome, then plasmid DNA was purified as described in section 2.2.1 and sequenced separately.

2.2.3 Annotation and Visualization of pMccB17

The draft genome sequence was annotated automatically by MicrobesNG (Birmingham, UK; <https://www.microbesng.com/>) using Prokka 1.11 (Seemann, 2014). Prokka works by homology searching of translated protein sequences using BLAST, HMMER and some other databases (AMR gene and transposon databases). The annotations generated by Prokka were

manually checked using BLASTn (Zhang *et al.*, 2000) and BLASTp (Altschul, 1997), retaining the default parameters. CDSs without proper annotation (y-genes and hypothetical proteins) were manually assigned one where possible using the BLAST result as a guide and Artemis version v18.0.0 to edit the annotation (Berriman, 2003; Rutherford *et al.*, 2000).

Genetic sequences of the three pMccB17 contigs were compared using Artemis Comparison Tool (ACT) version v18.0.0 to match the contigs together in order to assemble a complete plasmid genome sequence (Carver *et al.*, 2005). The ACT comparison files used in ACT were generated by using the NCBI tBLASTx page, the Fasta files of sequence to be compared were uploaded and a “hit table” format containing the comparison was generated and downloaded. The circular diagram of the complete plasmid genome (Fig. 3.2) was generated by DNAPlotter version v18.0.0 (Carver *et al.*, 2009).

For uploading the complete plasmid genome sequence to GenBank, GB2sequin was used to create a five-column, tab-delimited feature table for the pMccB17 sequence (<https://chlorobox.mpimp-golm.mpg.de/GenBank2Sequin.html#>; Lehwark and Greiner, 2019). In addition to the pMccB17 sequence in Fasta format, the feature table was submitted to BankIt, and the pMccB17 was assigned an accession number (<https://www.ncbi.nlm.nih.gov/WebSub/>).

2.2.4 *Plasmid replicon and conjugation system typing*

Replicon typing was carried out using the IncF RST (replicon sequence typing) scheme, as implemented by pubMLST version 1 (<https://pubmlst.org/plasmid/>; Jolley *et al.*, 2018) and pMLST version 2.0 (<https://cge.cbs.dtu.dk/services/pMLST/>; Carattoli *et al.*, 2014). Initial classification of the relaxase was carried out using MOBscan (<https://castillo.dicom.unican.es/mobscan/>; Garcillán-Barcia *et al.*, 2020). This was manually refined by pairwise comparisons with selected relaxase proteins using EMBOSS Needle (https://www.ebi.ac.uk/Tools/psa/emboss_needle/; Needleman and Wunsch, 1970), which uses the Needleman-Wunsch global alignment algorithm to locate the optimum alignment of two sequences along their entire length (gaps inclusive).

2.2.5 *Phylogenetic tree building*

The phylogenetic tree of Pbf protein homologs was built using a maximum-likelihood method (Felsenstein, 1981) in MEGA X version 10.1.8 (Kumar *et al.*, 2018). The tree used the LG + G model (Le and Gascuel, 2008) for amino acid substitution. Initial tree(s) for the

heuristic search were obtained automatically by applying Neighbor-Join and BioNJ algorithms to a matrix of pairwise distances estimated using a JTT model, and then selecting the topology with superior log likelihood value. A discrete Gamma distribution was used to model evolutionary rate differences among sites (5 categories (+G, parameter = 3.2901)). The tree is drawn to scale, with branch lengths measured in the number of substitutions per site. This analysis involved 24 amino acid sequences. All positions with less than 95% site coverage were eliminated, i.e., fewer than 5% alignment gaps, missing data, and ambiguous bases were allowed at any position (partial deletion option). There were a total of 169 positions in the final dataset. Evolutionary analyses were conducted in MEGA X. Bootstrapping is a technique for assessing the accuracy of the phylogenetic tree; generally, above 75% is considered reliable. Bootstrap for the phylogenetic tree in this project was set to 1000 (the tree was redrawn 1000 times based on random subsampling of the phylogenetic data set, with replacement, and the proportion of these trees with each clade displayed was calculated).

2.2.6 Identification of Resistance Genes, Insertion Sequences and Virulence factors

Identification of resistance genes was performed by submitting the complete plasmid nucleotide sequence to ResFinder web server version 4.0 (<https://cge.cbs.dtu.dk/services/ResFinder/>; Zankari *et al.*, 2012). ResFinder uses BLAST to search in whole-genome sequence data for acquired antimicrobial resistance genes, as collected in the ResFinder database. The nucleotide sequence was also submitted to ISSaga web server version 2.0 (<https://www.issaga.biotoul.fr/>; Varani *et al.*, 2011) and VFDB web server (<http://www.mgc.ac.cn/VFs/>; Liu *et al.*, 2019) to identify insertion sequences and virulence factors respectively. TOPCONS web server (www.topcons.net; Tsirigos *et al.*, 2015) was used to predict transmembrane regions. TOPCONS works by predicting the topology of proteins. This is done using five topology prediction algorithms: OCTOPUS, Philius, PolyPhobius, SCAMPI and SPOCTOPUS, which give a consensus prediction for the proteins, along with a reliability score that is built on the agreement of the included methods. ISSaga, VFDB and webFlaGs work by using BLAST search similar to the ResFinder.

The webFlaGs web server (<http://130.239.193.227/html/webFlaGs.html>; Saha *et al.*, 2021) was used to predict proteins encoded by genomic neighbours of Pbf homologues. HHpred server was used to detect protein homology domains (<https://toolkit.tuebingen.mpg.de/tools/hhpred>; Söding *et al.*, 2005), and EMBOSS Pepstats

was used to calculate protein molecular weight (https://www.ebi.ac.uk/Tools/seqstats/emboss_pepstats/; Madeira *et al.*, 2022).

2.2.7 PCR amplification of DNA.

For standard PCR (using IMMOLASE DNA polymerase) and proofreading PCR (using Q5 polymerase), the following reaction mixtures were used: 25 μ l of Mastermix (Q5 or IMMOLASE), 0.2 μ M (for IMMOLASE) or 0.5 μ M (for Q5) of forward and reverse primer, 50 ng/ μ l of template DNA and the reaction made up to 50 μ l by adding autoclaved, deionised water. The thermocycling parameters used were based on manufacturer's recommendation which can vary based on sample's annealing temperature and the expected product size. All PCR reactions were carried out in a T100 thermal cycler (BioRad).

2.2.8 Agarose gel electrophoresis

To prepare the gel, the required w/v % of agarose was added to TAE buffer, and the mixture was boiled until the agarose had completely dissolved. GelRed fluorescent DNA stain (Biotium) was added to 1X concentration, and the gel was allowed to cool down before casting. The solidified gel was placed in the gel tank and flooded with TAE buffer. Ladder(s) and samples (20 μ l) were loaded into the well, and the gel was run at 70 V for 150 minutes using wide Mini-Sub cell GT or Mini-Sub cell GT electrophoresis tanks (Bio-Rad). The gel was visualized using UV transillumination in a ChemiDoc MP, Universal Hood III Gel Documentation System (Bio-Rad).

2.2.9 Purification of PCR products

PCR products showing a single amplicon (one PCR product) on agarose gel were purified from solution using Monarch PCR and DNA Cleanup Kit (NEB). While the expected band was excised under blue light illumination from agarose gel showing more than one amplified product, the excised DNA was purified from the agarose slice using Monarch DNA Gel Extraction Kits (NEB). The kits employ a silica-based spin column (the silica will bind DNA in the presence of chaotropic salts, such as guanidinium hydrochloride, present in the binding buffer), in addition to wash buffers to remove short primers, dNTPs, enzymes, short-failed PCR products, and salts.

2.2.10 Restriction enzyme digestion

The restriction enzyme digestion reaction setup varied depending on whether it was a single or double digest. The reaction was generally made up of 2 μl of 10X Restriction enzyme buffer, up to 1 μg of DNA, 10 U of each restriction enzyme and the final volume is made up to 20 μl by adding nuclease free water. Following the manufacturer's recommendations, the reactions were incubated, and heat inactivated (if appropriate) at the recommended temperature for the required time.

2.2.11 Alkaline dephosphorylation

Thermosensitive alkaline phosphatase was used to dephosphorylate the pSEVA234 plasmids during the cloning process to prevent recirculation during ligation. 1 U of thermosensitive alkaline phosphatase (FastAP, Thermo Scientific) was added to the restriction enzyme digest setup of pSEVA234 plasmid, incubated at 37°C for 10 minutes and heat inactivated at 65°C for 15 minutes.

2.2.12 Phosphorylation

Thermo Scientific T4 Polynucleotide Kinase was used to phosphorylate PCR products where amplicon was self-ligated. The reaction mixture consisted of 2 μl of 10X Buffer A for T4 Polynucleotide Kinase, 10 U of T4 Polynucleotide Kinase, 1-20 pmol of 5'-termini 10-50 pmol of DNA, 2 μl of ATP 10mM and volume made up to 20 μl by adding nuclease-free Water. The reaction setup was mixed thoroughly and incubated at 37°C for 20 minutes, after which it was heated at 75°C for 10 minutes.

2.2.13 Ligation

An Instant Sticky-end Ligase Master Mix (NEB) or Blunt/TA Ligase Master Mix (NEB) was used for joining DNA samples, depending on whether the sample had cohesive or blunt DNA ends. They both use the same protocol. A vector DNA (20–100 ng) was combined with a 3-fold molar excess of the DNA insert, and the volume was adjusted to 5 μl with distilled water. 5 μl of ligase master mix was added and mixed thoroughly.

2.2.14 Transformation

NEB 5-alpha chemically competent *E. coli* (High Efficiency) was used for all transformation-related experiments where a high transformation efficiency was required. The competent cells were thawed on ice, 1 μl of plasmid DNA was added to 25 μl of the competent cells and

mixed gently, placed on ice for 30 minutes, heat shocked at 42 °C for 45 seconds and placed back on ice for 2 minutes. 300 µl of SOC was added, and the mixture was incubated at 37 °C for 60 minutes with shaking at 200 rpm. Samples of the mixture were spread onto a plate (with the right antibiotic for selection of the plasmid) and incubated at 37 °C overnight.

E. coli BW25113 and *E. coli* BW25113 Δ tldD were treated chemically to make them competent in order to transform microcin plasmids into them to study the effect of TldD on microcin production. New colonies of cells to be made competent were grown in LB broth (with antibiotics if needed) overnight at 37 °C with shaking at 200 rpm. The overnight culture was diluted 1:100 with fresh LB broth. These were incubated at 37 °C with shaking at 200 rpm until OD_{600 nm} =0.5. The suspension was centrifuged at 4500 x g for 7 minutes at 4 °C, and the cell pellet was resuspended in prechilled 100 mM CaCl₂ (2 ml per 5 ml culture) and incubated on ice for 20 minutes. The suspension was centrifuged at 4500 x g for 7 minutes at 4 °C, and the cell pellet was gently resuspended in prechilled 100 mM CaCl₂ (0.5 ml per 5ml culture). The newly made competent cells were stored at 4 °C.

2.2.16 *In vivo* determination of microcin activity (bioassay)

Following the method of Shkundina *et al.* (2014), 3 µl of an overnight culture (grown in LB at 37 °C with shaking at 200 rpm) of *E. coli* DH5 α transformed with microcin, or control plasmids were spotted onto M63 agar. After overnight incubation, the M63 plates were overlaid with 5 ml of 0.6% soft agar containing 100µl overnight microcin sensitive *E. coli* BL21 (DE3) culture (grown in LB at 37 °C with shaking at 200 rpm). Diameters of growth inhibition zones were determined by measuring from the front of the plate using a ruler after overnight growth at 25 °C.

2.2.17 Microcin purification steps

The procedure described by Yorgey *et al.* (1994) was followed with some adaptations. A colony from freshly streaked plates of microcin producer cells was grown in 5 ml LB (with ampicillin, 100 µg/ml) and incubated overnight. A 500 ml M63 broth (with ampicillin) in a 2 L conical flask was inoculated with the starter culture and incubated for 24 hours at 37°C with shaking at 200 rpm. Cells were collected by centrifugation at 4250xg for 15 mins, and the cell pellet was boiled in 20 ml of 100mM acetic acid, 1mM EDTA for 10 minutes with stirring and centrifuged again. The supernatant was decanted into a clean centrifuge tube and centrifuged again at 4250xg for 30 minutes to get a clarified suspension. A solid phase

extraction (SPE) was set up using the Varian Bond Elut LRC C18 cartridge, which was equilibrated by applying 5 ml methanol and 5 ml water onto the cartridge. 2 ml of the clarified supernatant was loaded onto the C18 cartridge, and the cartridge was washed with 5 ml water / 0.1 % formic acid and 12 % methanol / 0.1 % formic acid. Fractions containing the microcin B were eluted in 50% methanol/0.1% formic acid; these were vacuum dried and redissolved in 100 μ l of DMSO.

2.2.18 Liquid chromatography–mass spectrometry analysis

Agilent 1290 Infinity LC – System coupled with Agilent 6120 Single Quadrupole LC/MS system was used for the LCMS analysis in addition to ACE 5 C18, 250 X 4.6 MM column. A quality control pooled analysis was set up because individual runs of the SPE samples on LCMS gave little to no peaks. With an injection size of 10 μ l, 3 runs of mobile phase blank (mobile phase – 50% methanol in water), 3 runs of SPE EMB17 and SPE MccB17 pooled together, 3 runs of SPE biological blank (*E. coli* DH5 α [pUC19]), followed by 1 run of mobile phase blank, 3 runs of SPE MccB17, 1 run of mobile phase blank, 3 runs of SPE EMB17 and 1 run of mobile phase blank.

The reason for this experiment was to establish the peaks as well as the masses that were likely to represent microcin products. Mobile phase blank was run after each sample run to wash the column, and the 3 runs for the pooled SPE MccB17 and SPE EMB17, SPE MccB17, and SPE EMB17 were to allow for quality control.

Each LCMS run was for 70 minutes; compounds were eluted from the column in a linear gradient of methanol in water from 10% to 90%, with 0.1% formic acid, a flow rate of 4 ml/min and were detected by measuring absorbance at 254 nm.

2.2.19 Preparative High pressure liquid chromatography analysis

Agilent 1260 Infinity Preparative HPLC System in addition to ACE 5 C18, 250 X 10.0 MM (ACE-121-2510) column, was used to purify the microcin peptide. A sample of SPE EMB17 was run on the HPLC system in order to purify the masses of interest that the LCMS picked up. A linear gradient elution method of methanol in water from 10% to 90% with formic acid for 25 minutes, a flow rate of 5 ml/min and an absorbance of 254 nm was used. The mass spectrometer was set to detect masses between 900 – 1200. There was no collection for the first run, as the aim was to establish when the masses of interest were retained.

The second run followed the same method as the first in addition to collecting the peaks of interest. A time-based collection program was set up using the first run as a template. The collected fractions were vacuum-dried, resuspended in DMSO and stored at -20°C.

Chapter Three

pMccB17

3.0 Overview

This chapter describes the sequencing and genomic analysis of the *E. coli* plasmid pMccB17 that carries the microcin B17 gene cluster. Some plasmid typing, as well as bioinformatic analysis of the plasmid backbone sequences was carried out to characterise the plasmid. The microcin B17 operon was compared to some published microcin B17 operon sequences. An additional spotlight was on *yffB*, a gene of interest whose function is not known.

3.1 Introduction

The conjugative plasmid pMccB17, previously known as pRYC17, was discovered in *E. coli* strain LP17 isolated from the intestinal tract of a healthy neonate in Spain and transferred by conjugation to a K-12 strain of *E. coli* (Baquero *et al.*, 1978). This is a low copy number plasmid (approximately 2 copies per chromosome) belonging to the IncFII group assigned using a phenotypic incompatibility study that includes the archetypes R100 and R1 (Baquero *et al.*, 1978). Plasmid pMccB17 is not known to possess any conventional antibiotic resistance markers, and its size has been previously estimated as 70 kb (Millan *et al.*, 1985).

Here we report the complete sequence of pMccB17, as well as a comparative genomic analysis. We highlight some noteworthy features, in particular, a ParB-fusion protein of unknown function (encoded by the *yffB* gene), homologous of which are widespread in plasmids of IncF type.

3.2 Results and discussion

3.2.1 Purification and sequencing of plasmid pMccB17

The *E. coli* K-12 strain ZK0005 (RYC1000 [pMccB17]) carrying the plasmid pMccB17 was isolated by alkaline lysis using a Plasmid Purification Midi kit (Qiagen) according to the manufacturer's recommendations for "very low-copy" plasmids.

The plasmid, pMccB17, was sequenced twice by MicrobesNG (Birmingham, UK; <https://www.microbesng.com/>) using an Illumina short-read next-generation method. Firstly, the plasmid, pMccB17, and the host bacterial chromosome were sequenced; later, the plasmid DNA was sequenced separately after purification.

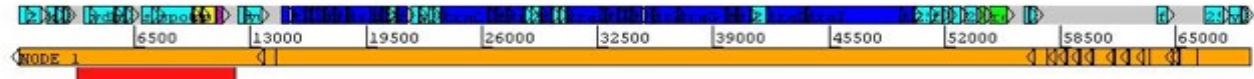
3.2.2 Plasmid sequence assembly

Plasmid contigs from the draft sequence were scaffolded as follows. Two contigs, 1.27 and 1.52, from the first draft sequence (of plasmid and chromosomal DNA) were discovered to have standard plasmid-related features such as plasmid replicon, maintenance genes and conjugation genes by examining the sequence annotations. Contig 1.27 (60265 bp) had 71 genes, including the microcin B17 operon, conjugative transfer system, replicon system and some plasmid maintenance genes (*stbA*, *stbB*, *parE*), while contig 1.52 (8894 bp) had 12 genes, half of which have plasmid maintenance related functions, i.e. *ssb*, *pbf*, *psiA*, *psiB* and *hok/sok*. Contig 2.1, the first contig in the second draft sequence (from purified plasmid DNA), contained all the genes present in contigs 1.27 and 1.52, and when the three contigs were aligned (Fig. 3.1) using the Artemis Comparison Tool (Carver *et al.*, 2008; Carver *et al.*, 2005), it was evident that contigs 1.27 and 1.52 make up contig 2.1. The ends of contig 2.1 were found to be within a CDS (coding sequence), gene_81, found in contig 1.27. Finally, a few nucleotides missing from contig 2.1 but present in contig 1.27 (gene_81) were added manually using Artemis, resulting in a complete genome sequence for pMccB17. Contig 2.1 was edited to be a complete linearized plasmid sequence of pMccB17. The nucleotide depth of coverage for the plasmid sequences in contig 1.27, contig 1.52 and contig 2.1 were 165X, 170X and 680X respectively. In order to have a complete circular plasmid, the ends (1 gap) of contig 2.1 were closed and using contigs 1.27 and 1.52 as a template where needed, the two gaps in their ends were closed as well in the section below.

(i) Contig 1.27 {



Contig 2.1 {



Contig 1.52 {



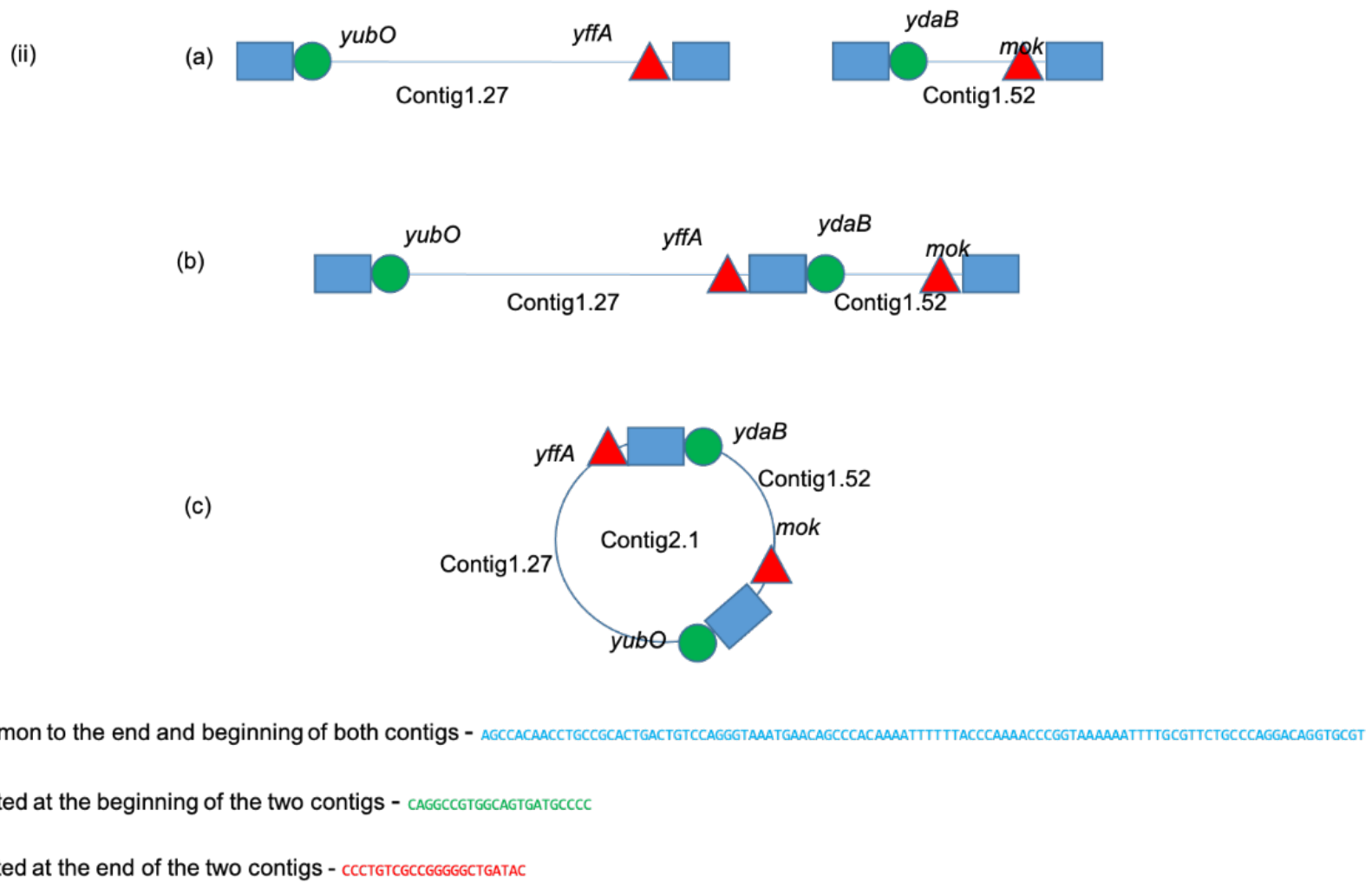


Figure 3.1 Relationship between the first and second draft sequence of pMcC17. (i) shows the result of comparing the first and second drafts using Artemis Comparison Tool (ACT). The red trapezoidal regions between the contigs show identical sequences and genes. (iia) shows the two contigs (1.27 and 1.52) in the first draft (iib) shows the assembly of contigs 1.27 and 1.52 into one linear sequence (iic) shows the assembly of contigs 1.27 and 1.52 into a circular form which is similar to what was seen in the second draft, contig 2.1.

3.2.3 PCR and plasmid gap closing

PCR primers were designed to flank gaps in the plasmid sequence, and the corresponding oligonucleotides synthesized by Integrated DNA Technologies (Table 2.1) were employed to close the plasmid gaps by PCR amplification followed by Sanger sequence of amplicons. ImmoMix (Bioline) or Q5 High Fidelity (New England Biolabs) PCR master mixes were used following the manufacturer's recommendations for the reaction setup and the thermocycler parameters. Products of the expected sizes were excised from the gel and purified using Monarch DNA Gel Extraction Kit (NEB) according to the manufacturer's instructions. Sanger DNA sequencing of PCR amplicons was carried out by DBS Genomics (Durham, UK). Manual inspection of the sequences allowed plasmid genome gap resolution, resulting in a complete finished circular genome for pMccB17.

3.2.4 Annotation of pMccB17 sequence and submission to GenBank

The sequence was annotated automatically by MicrobesNG (Birmingham, UK; <https://www.microbesng.com/>) while CDS without proper annotation were assigned one using BLAST as a guide and Artemis to edit the annotation (Berriman, 2003; Rutherford *et al.*, 2000) (Table 3.1). BLAST best fit were decided based on the percentage similarities and coverage which were close to 100% in most cases. Examples are *psiA* and *psiB* which had no annotation from the first sequence but using BLASTp, names were assigned and their functions predicted.

Table 3.1 CDS identified in pMccB17 and origin of annotation

Coding Sequence (CDS)	Position (bp)	Protein function	Origin of annotation
<i>ycfA</i>	496..1266	hypothetical protein	manual
<i>klcA</i>	1682..2107	antirestriction protein	Prokka
<i>gene_78</i>	2154..2576	hypothetical protein	Prokka
<i>yffA</i>	2573..2764	hypothetical protein	manual
<i>ydaB</i>	3802..4032	hypothetical protein	manual
<i>ydbA</i>	4084..5445	hypothetical protein	manual
<i>rsmA</i>	5492..6055	ribosomal RNA small subunit methyltransferase A	Prokka
<i>gene_73</i>	6055..6321	hypothetical protein	Prokka

<i>ssb</i>	6964..7503	plasmid-derived single-stranded DNA-binding protein	Prokka
<i>yubL</i>	7565..7798	hypothetical protein	manual
<i>yfjB</i>	7863..9821	partitioning protein	manual
<i>psiB</i>	9876..10310	plasmid SOS inhibition protein B	manual
<i>psiA</i>	10307..11026	plasmid SOS inhibition protein A	manual
<i>mok</i>	11248..11460	modulation of hok protein	manual
<i>yubO</i>	12391..12678	hypothetical protein	manual
<i>yubP</i>	12797..13618	hypothetical protein	manual
<i>gene_64</i>	complement (13913..14422)	invasion protein	Prokka
<i>traM</i>	14838..15221	relaxosome protein	Prokka
<i>traJ</i>	15415..16086	conjugal transfer transcriptional regulator	manual
<i>traY</i>	16223..16450	relaxosome protein tray	Prokka
<i>traA</i>	16483..16842	pilin TraA	Prokka
<i>traL</i>	16857..17168	peripheral membrane protein TraL	manual
<i>traE</i>	17190..17756	bacterial sex pilus assembly and synthesis proteins TraE	manual
<i>traK</i>	17743..18471	pilus assembly TraK	manual
<i>traB</i>	18471..19922	involved in F pilus assembly	manual
<i>traP</i>	19912..20478	conjugation protein	manual
<i>trbD</i>	20417..20788	conjugal transfer protein TrbD	manual
<i>gene_53</i>	20781..21290	hypothetical protein	Prokka
<i>traV</i>	21283..21795	conjugative transfer system lipoprotein	manual
<i>yafA</i>	22144..22557	hypothetical protein	manual
<i>yfhA</i>	22550..23023	hypothetical protein	manual
<i>orfH</i>	23104..23322	hypothetical protein	manual
<i>orfI</i>	23350..23697	hypothetical protein	manual
<i>traC</i>	23823..26453	conjugative transfer ATPase	manual
<i>trbI</i>	26450..26836	type-F conjugative transfer system protein	manual
<i>traW</i>	26833..27465	type-F conjugative transfer system protein	manual

<i>traU</i>	27462..28454	conjugal DNA transfer protein	manual
<i>ygeA</i>	28481..28789	hypothetical protein	manual
<i>gene_42</i>	28852..29370	hypothetical protein	Prokka
<i>trbC</i>	29397..30035	type-F conjugative transfer system pilin assembly protein	manual
<i>gene_40</i>	30032..30403	hypothetical protein	Prokka
<i>gene_39</i>	30428..30850	hnh nucleases	Prokka
<i>traN</i>	30847..32697	type-F conjugative transfer system mating-pair stabilisation protein	manual
<i>trbE</i>	32724..32981	type IV conjugative transfer system protein	manual
<i>traF</i>	32974..33717	type-F conjugative transfer system pilin assembly protein	manual
<i>trbA</i>	33731..34075	conjugal transfer repressor	manual
<i>traQ</i>	34194..34478	type-F conjugative transfer system pilin chaperone	manual
<i>trbB</i>	34465..35010	type-F conjugative transfer system pilin assembly thiol-disulphide isomerase	manual
<i>trbF</i>	35268..35660	plasmid mobilisation	manual
<i>traH</i>	35647..37020	type IV conjugative transfer system protein	manual
<i>traG</i>	37017..39842	TraG	manual
<i>traS</i>	39839..40348	conjugal transfer entry exclusion protein	manual
<i>traT</i>	40362..41093	lipoprotein TraT	manual
<i>gene_27</i>	41296..42033	hypothetical protein	Prokka
<i>traD</i>	42084..44282	type IV conjugative transfer system, coupling protein trad	Prokka
<i>traI</i>	44282..49552	conjugal transfer relaxase protein	Prokka
<i>traX</i>	49572..50318	type-F conjugative transfer system pilin acetylase	manual
<i>gene_23</i>	50377..51237	2,6-dihydropseudooxynicotine hydrolase	Prokka
<i>finO</i>	51340..51900	fertility inhibition protein	Prokka
<i>gene_21</i>	52045..52287	hypothetical protein	Prokka

<i>gene_20</i>	52486..52947	thermonuclease nuc	Prokka
<i>gene_19</i>	53241..53831	hypothetical protein	Prokka
<i>repA2</i>	54071..54328	replication regulatory protein	manual
<i>repA1</i>	54630..55487	plasmid replication initiator	manual
<i>parD</i>	56557..56820	toxin-antitoxin system	manual
<i>parE</i>	56810..57109	toxin-antitoxin system	manual
<i>gene_14</i>	complement (57145..57810)	hypothetical protein	Prokka
<i>mcbG</i>	complement (58380..58943)	microcin B17 immunity protein, McbG	manual
<i>mcbF</i>	complement (58947..59678)	microcin B17 export protein, McbF	manual
<i>mcbE</i>	complement (59675..60400)	microcin B17 export protein, McbE	manual
<i>mcbD</i>	complement (60410..61600)	microcin B17 synthase complex protein, McbD	manual
<i>mcbC</i>	complement (61581..62399)	microcin B17 synthase complex protein, McbC	manual
<i>mcbB</i>	complement (62401..63288)	microcin B17 synthase complex protein, McbB	manual
<i>mcbA</i>	complement (63315..63524)	microcin B17 peptide	Prokka
<i>fdtC</i>	64054..64506	dtdp-3-amino-3,6-dideoxy-alpha-d-galactopyranose 3-n-acetyltransferase	Prokka
<i>stbB</i>	complement (64909..65262)	type IV partitioning system	manual
<i>stbA</i>	complement (65262..66224)	type IV partitioning system	manual
<i>gene_3</i>	66736..67662	hypothetical protein	Prokka
<i>yhdJ</i>	68047..68730	DNA adenine methyltransferase YhdJ	Prokka
<i>gene_1</i>	68731..68952	hypothetical protein	Prokka
<i>Gene_81</i>	17..448	hypothetical protein	Prokka

The complete annotated genome sequence of pMccB17 was uploaded to the GenBank database at National Center for Biotechnology Information (NCBI), and the sequence was assigned an accession number ON989342.

3.2.5 Visualization

The circular diagram of the plasmid genome (Fig. 3.2) was generated by DNAPlotter version v18.0.0 (Carver *et al.*, 2009; <https://www.sanger.ac.uk/tool/dnaplotter/>).

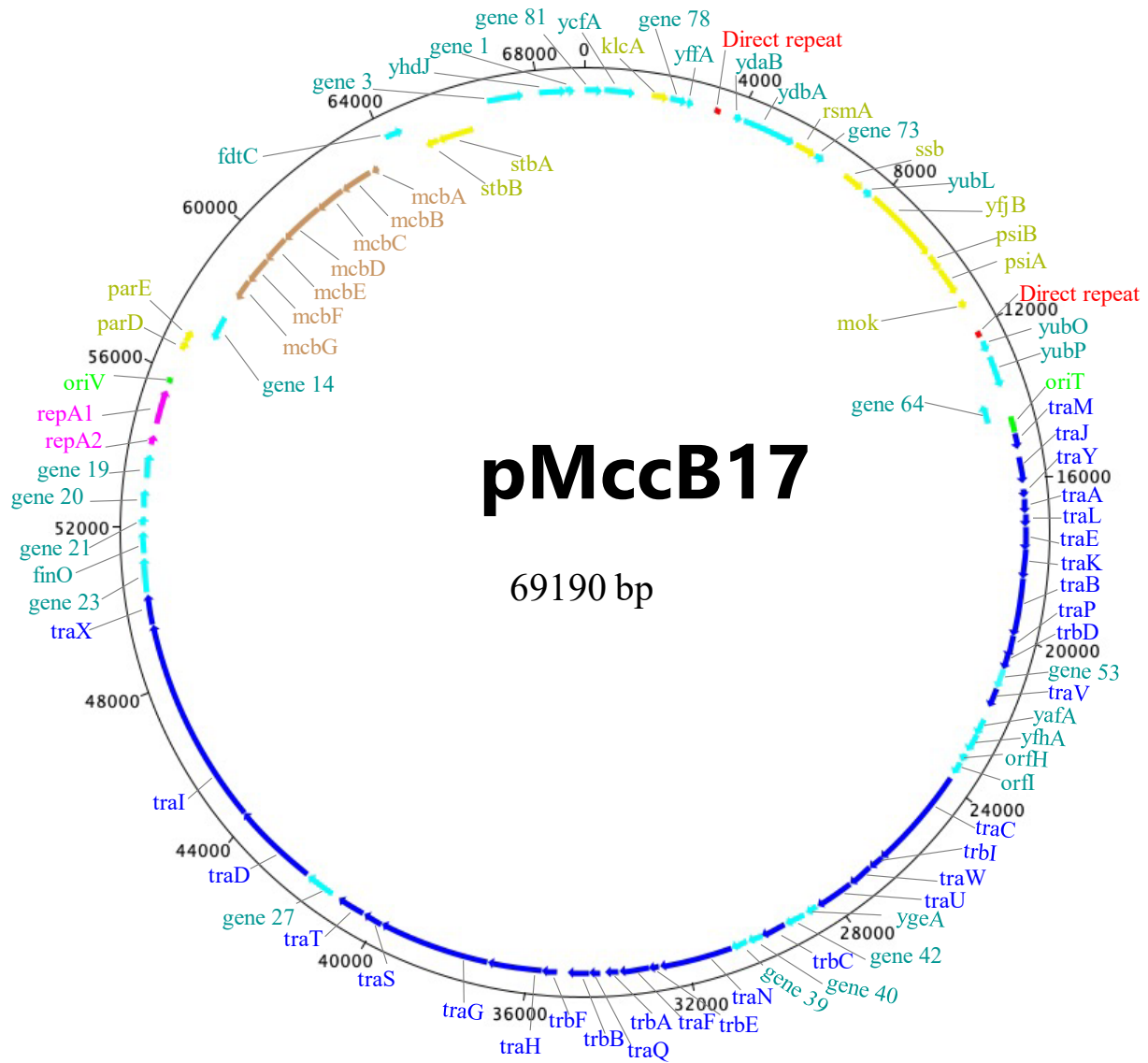


Figure 3.2 Circular Map of pMccB17. The outer ring shows the size of the plasmid, each tick representing 4kb. The microcin operon is shown in brown, plasmid maintenance systems are shown in yellow, the replicon is shown in pink, the conjugative transfer system is represented in blue, hypothetical proteins are shown in turquoise, origins of replication are in green, and the 149 bp direct repeats are shown in red.

3.2.6 Direct repeats

In the process of aligning contigs 1.27 and 1.52 with contig 2.1, to ascertain the complete plasmid sequence of pMccB17, a 149 bp direct repeat in contig 2.1 was discovered, with the sequence as below;

```
CCCTGTCGCCGGGGGCTGATACAGCCACAACCTGCCGCACTGACTGTCCAGGGT
AAATGAACAGCCCACAAAATTTTTTACCCAAAACCCGGTAAAAAATTTTGCGTTC
TGCCCAGGACAGGTGCGTCAGGCCGTGGCAGTGATGCCCC. These repeats are
comprised of the sequences represented by a triangle, rectangle, and circle in Fig. 3.1(ii). This
149 bp sequence is repeated twice (8639 bp apart) in the sequence of pMccB17 (Fig. 3.1(ii)
and Fig. 3.2). These repeats are identical, and each has an intragenic location: one is located
between yffA and ydaB, while the other is located between mok and yubO (Fig. 3.1(iic)). The
repeats have 12 genes in between them, including ssb, yffB, psiAB and the hok/sok TA system
(Fig 3.3). This approximately corresponds to the plasmid leading region originally defined for
F plasmid (Loh et al., 1989).
```

Searches using BLASTn against the nucleotide collection (nr/nt) database at NCBI revealed that the repeated sequence is present in other plasmids of Gammaproteobacteria, mainly in the Enterobacterales, with a copy number ranging from 1 to 3 in 20 plasmids examined (having query coverage and identity of 100%). The exact function of these direct repeats is unknown, but several genes are commonly flanked by these repeats in other plasmids, including *ssb*, *yffB* and *psiAB* (Fig. 3.3).

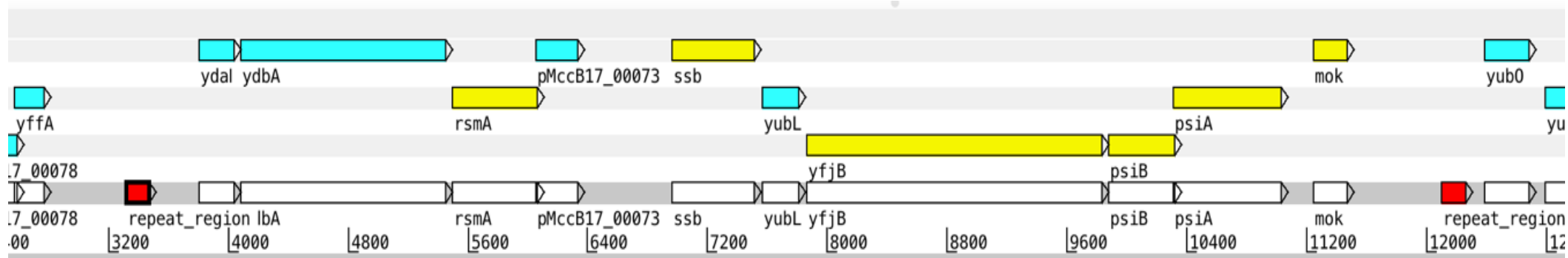


Figure 3.3 Genomic view of the regions between the direct repeats in pMccB17 showing the location of the 149 bp repeats as well as the genes present between them. The red boxes represent the 149 bp direct repeats, while the turquoise and yellow boxes represent genes encoding hypothetical proteins and plasmid maintenance proteins respectively.

3.2.7 Plasmid replicon and conjugation system typing

Plasmid replicons and conjugation system typing are important concepts in the study of plasmids and their transfer mechanisms, providing insights into the dissemination of genetic material among bacteria and contributing to our understanding of antibiotic resistance and other traits of clinical significance. Plasmid replicon system typing involves identifying and categorising different types of plasmids based on the specific replication mechanisms they employ. Plasmids can be put into incompatibility groups and typed more accurately using in-silico replicon typing and subtyping tools, but there are limitations to this as plasmids are often multireplicon or mosaic molecules (Orlek *et al.*, 2017). The replicon sequence type was assigned using the IncF RST scheme as implemented by pubMLST and pMLST, indicating a FAB type of F2:A-B-, so pMccB17 has an FII replicon (allele 2) without additional FIA or FIB replicons which is the same replicon type present in the plasmids R100 (*S. flexner*), R136 (*S. enterica* Typhimurium), R1-16 (*S. enterica* Paratyphi B), p48165T (*E. coli*) (Villa *et al.*, 2010). In addition to replicon typing, plasmids can also be typed and subtyped according to their conjugation systems. Initial classification of the conjugative relaxase using MOBscan assigned it to the MOB_F family (Garcillán-Barcia *et al.*, 2020). The first 300 N-terminal amino acids of the TraI relaxase / helicase protein (i.e. the relaxase domain) were then compared in a pairwise manner with archetypal IncF plasmid relaxases - the pMccB17 relaxase domain was 99.3 % identical to that of the IncFII plasmid R100 and 91.3 % identical to that of F plasmid, placing it in the MOB_{F12} type (Garcillán-Barcia *et al.*, 2011), or group A according to a recent phylogeny of IncF relaxases (Fernandez-Lopez *et al.*, 2016).

3.2.8 pMccB17 overview

pMccB17 is a circular typical IncFII plasmid (69190 bp) with a complete IncF conjugation system, an average GC content of 51% and 81 coding sequences (CDS) (Fig. 3.2 and Table 3.1). The assigned GenBank accession number for the nucleotide sequence is ON989342. This plasmid encodes the biosynthetic and immunity functions of MccB17, having the *mcB* operon. The *tra* region (*tra*MJYALEKBPVCWUNFQHGSTDIX and *trb*DICEABF) encodes conjugative transfer functions and takes up about half of the plasmid backbone. pMccB17 encodes two types of toxin-antitoxin (TA) / plasmid addiction system for eliminating plasmid-free cells: a type I system which comprises the post-segregation killing protein Hok (together with its modulator Mok), which is encoded downstream of *psiA* and a type II system ParDE which is encoded downstream of the replication gene (*repA*). The plasmid has

no known antibiotic resistance genes, virulence factors or transposable elements, according to ResFinder, IS-Finder and VFDB web-servers respectively.

3.3 The microcin B17 operon of pMccB17

The Microcin B17 operon, MccB17, consists of seven genes *mcbA-G* (see chapter 1 for full details) and is located between *gene_14* and *fdtC*. The nucleotides, as well as the amino acids sequences of genes and proteins present in the *mcb* operon found in pMccB17 (accession number ON989342), had some variations when compared to some historical published sequences (Table 3.2 and Table 3.3). The M24253 sequence was derived by Sanger sequencing of a cloned DNA fragment that has four MccB17 genes and is the first available sequence of *mcbA*, *mcbB*, *mcbC* and *mcbD* (Genilloud *et al.*, 1989). The X07875 sequence, also obtained via Sanger sequencing of a cloned fragment, is a continuation of the M24253 sequence having the remaining three genes *mcbE*, *mcbF* and *mcbG* (Garrido *et al.*, 1988). The ON98934 sequence, in theory, is meant to be identical to the M24253 and X07875 sequence as the plasmid, pMccB17, containing the operon was provided by one of the publishing authors (Roberto Kolter) of the M24253 and X07875 sequence (Garrido *et al.*, 1988; Genilloud *et al.*, 1989). The FM877811 sequence (located in the chromosome rather than on plasmid) provides a complete *mcb* operon and was selected because of this, although it is not the original *mcb* operon but rather is from the genome sequence (derived by pyrosequencing) of *E. coli* strain L1000, isolated from human faeces (Zihler *et al.*, 2009). There are lots of whole genome sequences of the *mcb* operon available in the databases, but the M24253 and X07875 sequences are of high importance to this work.

Table 3.2 Comparison of the genes in *mcb* operon in pMccB17 (ON989342) with some published sequences of *mcb* operon.

Gene	Predicted Function	Nucleotides Length				Identity (%) to ON989342		
		ON989342	FM877811	M24253	X07875	FM877811	M24253	X07875
McbA	Microcin precursor	210	210	210	– ^a	100	100	–
McbB	Microcin B synthase enzyme complex	888	888	888	–	99.9	99.9	–
McbC		819	819	819	–	99.9	100	–
McbD		1191	1191	1191	–	99.9	99.9	–
McbE	Self-immunity and export of MccB17	726	726	–	726	99.7	–	100
McbF	Self-immunity and export of MccB17	732	732	–	744	99.9	–	97.4
McbG	Self-immunity.	564	564	–	564	99.8	–	100
		Average				99.9		

^a The – indicates that no data was available.

Table 3.3 Comparison of the amino acid in MccB17 operon in pMccB17 (ON989342) with some published sequences of MccB17 operon.

Protein	Predicted Function	Amino Acid Length				Identity (%) to ON989342			Similarity (%) to ON989342		
		ON989342	FM877811	M24253	X07875	FM877811	M24253	X07875	FM877811	M24253	X07875
McbA	Microcin precursor	69	69	69	– ^a	100	100	–	100	100	–
McbB	Microcin B synthase enzyme complex	295	295	295	–	99.7	99.7	–	100	99.7	–
McbC		272	272	272	–	100	100	–	100	100	–
McbD		396	396	396	–	99.7	99.7	–	99.7	99.7	–
McbE	Self-immunity and export of MccB17	241	241	–	241	100	–	100	100	–	100
McbF	Self-immunity and export of MccB17	243	243	–	247	100	–	91.5	100	–	94.7
McbG	Self-immunity.	187	187	–	187	99.5	–	100	99.5	–	100

^a The – indicates that no data was available.

3.3.1 Sequence variation

Table 3.4 shows the sequence variations that exist when the *mcb* operon in the pMccB17 sequence (ON989342) is compared to selected published sequences. The variations that exist in the M24253 and X07875 sequence in comparison to the ON989342 sequence are noticed in *mcbB*, *mcbD* and *mcbF* and could be attributed to single nucleotide substitution error (GC swap) (Fig. 3.4), while the major variation was noticed towards the tail end of *mcbF* where there is a 13 base indel in between the ON989342 and X07875 sequence, thus making the X07875 sequence longer. After comparing the sequences using EMBOSS Needle, a deletion of a base between positions 687 and 688 in the X07875 sequence was observed, which might have resulted in a frameshift and thus the difference between the two *mcbF* genes (Fig. 3.5).

3.3.2 Confirming the variations

To confirm that the sequence variation observed in our ON989342 sequence was not due to an illumina sequencing error, PCR primers were designed to amplify *mcbB*, *mcbD* and *mcbF* of the ON989342 sequence and the amplicons were sent for Sanger sequencing. The Sanger sequencing data obtained unambiguously confirmed that our ON989342 sequence was correct (Fig. 3.4). Also, *mcbF* of the ON989342 and X07875 sequences were compared to sequences in the nr/nt database using BLASTn, and the ON989342 sequence has 100% identity match with other sequences (Fig. 3.6). In comparison, the only 100% identity match of X07875 was the hit from the X07875 sequence present in the database (Fig. 3.7). We suspect artefactual substitutions or deletions in the M24253 and X07875 sequences (i.e., Sanger sequencing errors or mutations in the cloned fragments that were sequenced). One of the limitations of Sanger sequencing compared to next-generation sequencing is low sensitivity; also, the enormous depth of next-generation sequencing coverage makes the result more reliable (Arsenic *et al.*, 2015).

Table 3.4 Details of the sequence variation between the ON989342 sequence and the published sequences.

Gene	Published sequence accession number	DNA sequence change^a	Amino acid change^b
<i>mcbB</i>	M24253	359G>C	C117S
	FM877811	594T>G	D198E
<i>mcbC</i>	FM877811	621C>T	
<i>mcbD</i>	M24253	512C>G	T171R
	FM877811	337A>G	T113A
<i>mcbE</i>	FM877811	474C>A	
		621T>C	
<i>mcbF</i>	FM877811	558T>G	
	X07875	127C>G	P43G
		128C>G	
		327C>G	D109E
		328G>C	V110L
		439G>C	V147L
		687_688delC	Q230P
		731_744ins CACATGGATATAA	K231E
			R232K
			H233A
			L234F
			L235I
			K237E
			L238V
			W239M
	E240G		
	Q241T		
	E242G		
	T243D		
<i>mcbG</i>	FM877811	334A>T	S112C

^a The first nucleotide belongs to the published sequences while the second base is the ON989342 sequence (this project)

^b The first amino acid belongs to the previously published sequences while second amino acid is the ON989342 sequence.

(A)

Artemis Entry Edit: Plasmid pMcC17 microcin (mcbA, mcbB, mcbC, and mvbD) genes, complete cds, and imm gene, 5' end.gb

entry: Plasmid pMcC17 microcin (mcbA, mcbB, mcbC, and mvbD) genes, complete cds, and imm gene, 5' end.gb ✓

4 selected bases on forward strand: 978..991

>>

mcbB

mcbA

mcbC

mcbD

300 600 900 1200 1500 1800 2100 2400

<<

S P L W V N S V I I S D V P H L F S N A R E Q W K C D G V F V S H I I D I K D N N

H H L S G # T Q S # # A M F H I F S V M P G N N G N V M V F L F L I S L I # K I I I

I T S L G K L S H N K R C S T S F Q # C P G T M E M * W C F C F S Y H * Y K R # # Y

ATCACCTCTCTGGGTAACCTCAGTCATAATAAGCGATGTTCCACATCTTTTCAGTAATGCCCGGAACAATGGAAATCTGATGGTGTITTTGTTCTCATATCATTGATATAAAAGATAATAATA

920 940 960 980 1000 1020

TAGTGGAGAGACCCATTGAGTCAGTATTATTCGCTACAAGGTGTAGAAAAGTCATTACGGGCCCTTGTACCTTTACTACCACAAAAAACAAGAGTATAGTAACATATTTTCTATTATTAT

M V E R P L S L * L L R H E V D K * Y H G P V I S I H H H K Q K E Y * Q Y L L Y Y Y

D G R Q T F E T M I L L S T G C R K L L A R S C H F H S P T K T E * I M S I F S L L I

: * R E P Y V * D Y Y A I N W M K E T I G P F L P F T I T N K N R M D N I Y F I I I

<<

source	1	3810
gene	408	617
CDS	408	617
gene	643	1530
CDS	643	1530
gene	1532	2350
CDS	1532	2350
gene	2331	3521
CDS	2331	3521
gene	3531	3810
CDS	3531	>3810

G366:C380 - 1 B01_190422b_LB_McbB_McbB-for.ab1 Search

T A A T G C C G G G A A C A A T G G A A A T C T G A T G G T G T T T T T G T T T C T C A T

360 370 380 390 400

1.0x

(B)



Figure 3.4 A chromatogram (from Sanger sequencing of pMccB17 PCR amplicons) and a sequence view in Artemis of the original *mcb* operon sequence (M24253) showing the single base substitution errors (GC swaps) in the original *mcb* operon. The nucleotide base in red circles indicates where the substitution took place. (a) the CG swap in *mcbB* (b) the GC swap in *mcbD*

```

mcbF_ON989342 501 TCAGTACAGAATGATGTTATGGGAACGATAAAATAAGATCACCGCTGACG 550
      |||
mcbF_X07875   501 TCAGTACAGAATGATGTTATGGGAACGATAAAATAAGATCACCGCTGACG 550

mcbF_ON989342 551 GGAAAACGGTATTTTTCTCAACTCATATTTTGTGAATTAACAAGAGAT 600
      |||
mcbF_X07875   551 GGAAAACGGTATTTTTCTCAACTCATATTTTGTGAATTAACAAGAGAT 600

mcbF_ON989342 601 AAAATCCCTTCATATGCTTTCAAAAAATAGTATAAATCGTTATTCTGA 650
      |||
mcbF_X07875   601 AAAATCCCTTCATATGCTTTCAAAAAATAGTATAAATCGTTATTCTGA 650

mcbF_ON989342 651 TATGAGTGATTTTATTCAATCCAATAATGAAACAAC687TCCAGAAAAGGCAT 700
      |||
mcbF_X07875   651 TATGAGTGATTTTATTCAATCCAATAATGAAACAAC688-CAGAAAAGGCAT 699

mcbF_ON989342 701 TTATTAAGAAGTTATGGGAACAGGAGACTGA----- 732
      |||
mcbF_X07875   700 TTATTAAGAAGTTATGGGAACAGGAGACTGACACATGGATATAA 744

```

Figure 3.5 Deletion in *mcbF* of the X07875 sequence, which resulted into frameshift. The red rectangle shows the location of the deletion and where the frameshift started from while the blue shows the incorrect extension of the predicted gene.

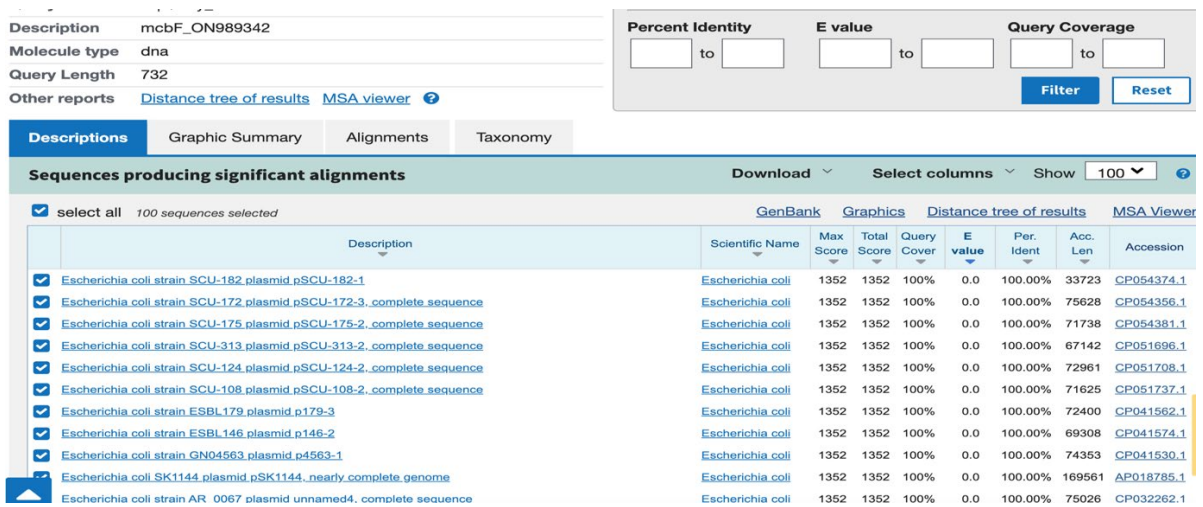


Figure 3.6 A BLASTn result of *mcbF* ON989342 sequence showing other homologues with 100% identity in the nr/nt database.

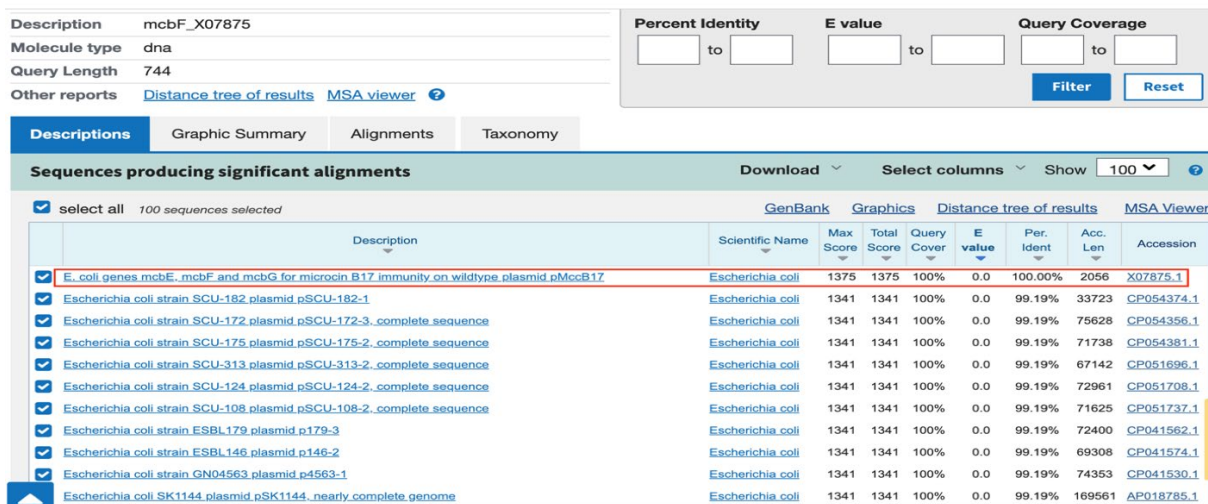


Figure 3.7 A BLASTn result of *mcbF* X07875 sequence showing homologous sequences in the nr/nt database. In red rectangle is a 100% identity match of the same sequence.

3.4 ParB-fusion protein (Pbf).

Upstream of *psiB* is the gene *yjfB* encoding a 652 amino acid protein that we refer to as “*pbf*” (ParB fusion protein) as it features a ParB-like N-terminal domain region (CBP of type I partition system) joined to a C-terminal region that does not include any known conserved domain. The result of the BLASTp search of its amino acids against the non-redundant protein sequences (nr) database detected a conserved domain annotated as “ParB/RepB/Spo0J family partition protein” in the N-terminal half (the first 250 amino acids) of the sequence (Fig. 3.8). So Pbf is clearly evolutionarily, if not functionally, related to proteins that have interactions with a NTPase partner (ParA) and a centromere-like partition site on the DNA (*parS*) that are involved in the active partitioning of bacterial chromosomes and low copy-number plasmids.

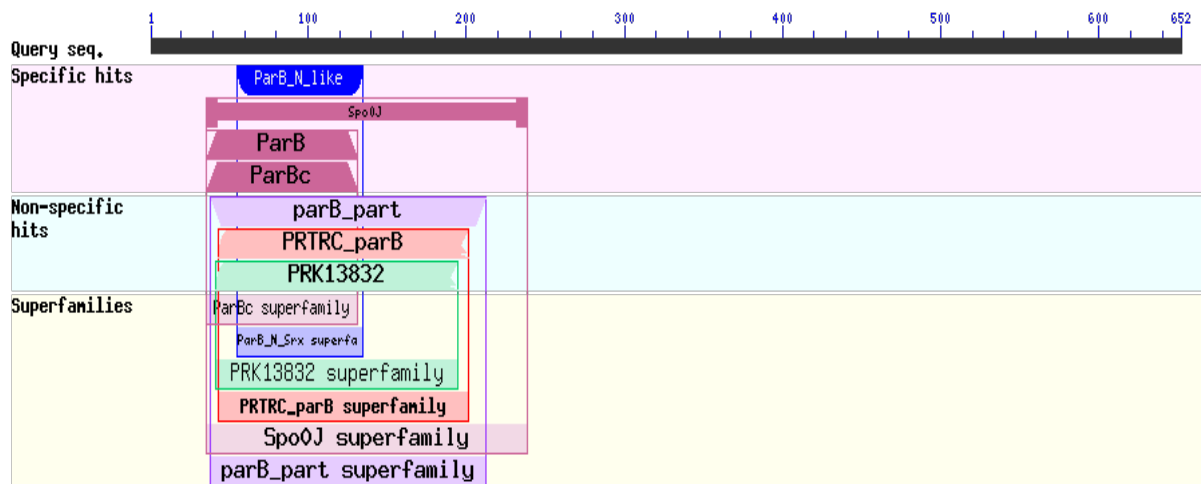
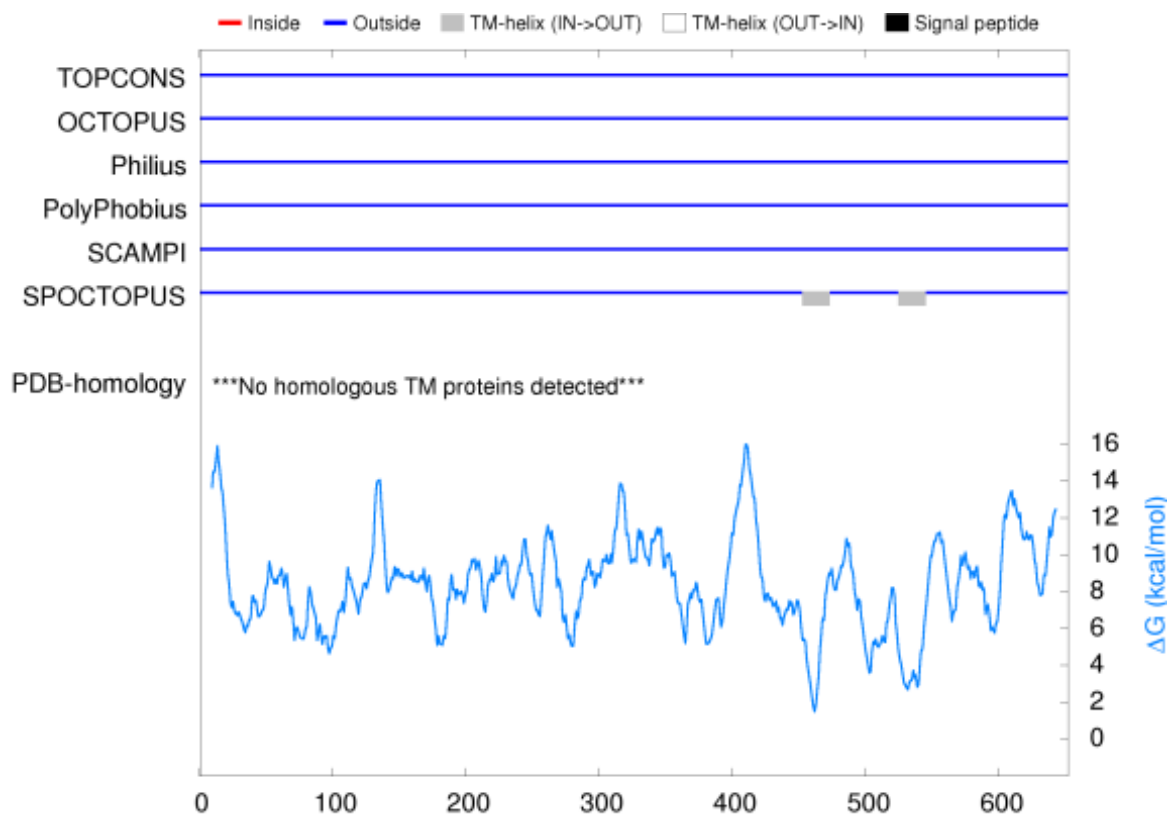


Figure 3.8 Location of known conserved domains in the ParB-fusion protein. Numbers refer to amino acid sequence of the 652 amino acid Pbf. The image was taken from Conserved Domain Database (Marchler-Bauer *et al.*, 2017).

The classical plasmid F carries a homologous gene, *orf652*, which is 94% identical to *pbf*, and the similarity of the encoded protein to ParB was first described by Manwaring *et al.* (1999). The C-terminal region (amino acids 250-652) does not include any known conserved domain, currently there is no hint as to its function.

Using TopPred II software, Manwaring *et al.* (1999) predicted the presence of two transmembrane segments in the C-terminal region of 3 (F plasmid, pO157, pMT1) out of the 4 Pbf homologous they identified. This analysis was repeated using the TOPCONS server, which predicts transmembrane protein by compiling the predictions of 5 other transmembrane prediction packages: Philius, PolyPhobius, SCAMPI and SPOCTOPUS. We analysed 13 ParB family proteins, included were 9 Pbf homologous, Spo0J from *Bacillus subtilis*, Spo0C from *Thermus thermophilus*, KorB from plasmid RK2 and ParB from P1 phage and found that a transmembrane segment was only predicted in those that were used by Manwaring *et al.* (1999), F plasmid, pMT1 and pNL1, with F plasmid having the 2 predicted transmembrane while pMT1 and pNL1 have only one predicted transmembrane region, as opposed to the two initially predicted for them. Overall, the evidence for transmembrane domains in the C-terminal of these proteins is weak: only one of the methods compiled by TOPCONS (SPOCTOPUS) indicated the presence of transmembrane segments as noted above, and this was only with low reliability (Fig. 3.9).

(A)



(B)

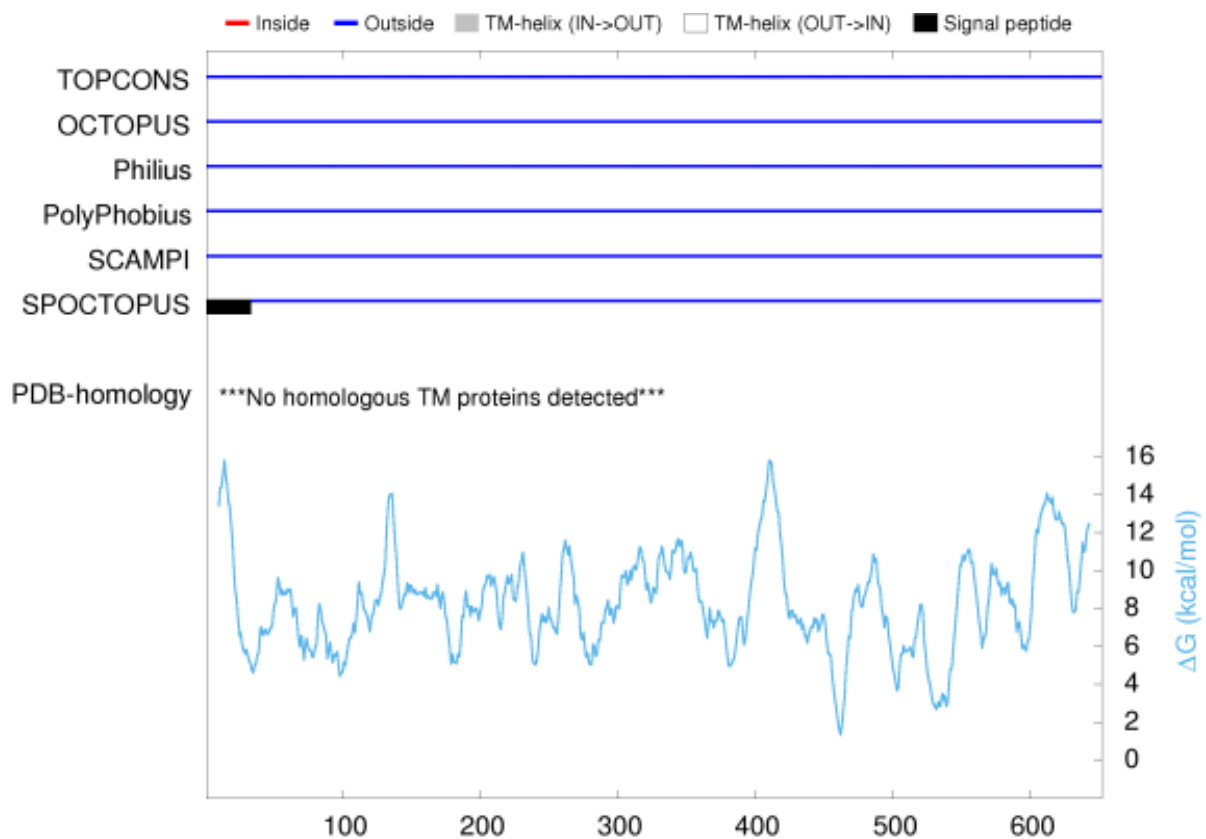


Figure 3.9 Prediction of transmembrane segment in selected Pbf proteins. The location of predicted transmembrane region(s) if any are represented by the colour code on the chart. (a) F plasmid (b) pMccB17

The N-terminal region of Pbf has two features that are functionally important in ParB family proteins: a partly conserved arginine-rich ParB Box II (consensus GXXRXXA; Fig. 3.10 and Fig. 3.11) essential for bridging of the ParB-*parS* complexes; and the helix-turn-helix (HTH) motif required for specific binding of the centromere-like DNA site (Manwaring *et al.*, 1999; Graham *et al.*, 2014; Pióro and Jakimowicz, 2020). This strongly suggests that Pbf is functionally related to ParB family proteins, and a role in partitioning was suggested by Manwaring *et al.* (1999). This might seem somewhat redundant given that some plasmids have both *pbf* and another partitioning system (in most cases, the type Ia partitioning system: *parAB* or *sopAB*). For instance, F plasmid and pEK499 both have the type Ia *sopAB*

partitioning system; pSLT, and pST1007-1D have the type Ib *parAB* system, while pMccB17 has *stbAB*, a type II system. It is worth noting that functional dual partitioning systems of different types have been reported in some plasmids pINV, type Ia and II in *Shigella flexneri* (McVicker *et al.*, 2019) and pB171, type Ib and II in *E. coli* (Baxter and Funnell, 2014). However, a question remains whether partitioning systems of the same type can exist and function together in a plasmid. Pbf protein is most similar to the type I partitioning system owing to its key features (ParB BoxII motif and HTH). With Pbf being present in some plasmids that already have a type I partition system, we are uncertain if the ParB-fusion protein might be playing a role outside plasmid partitioning or contributing to plasmid partitioning.

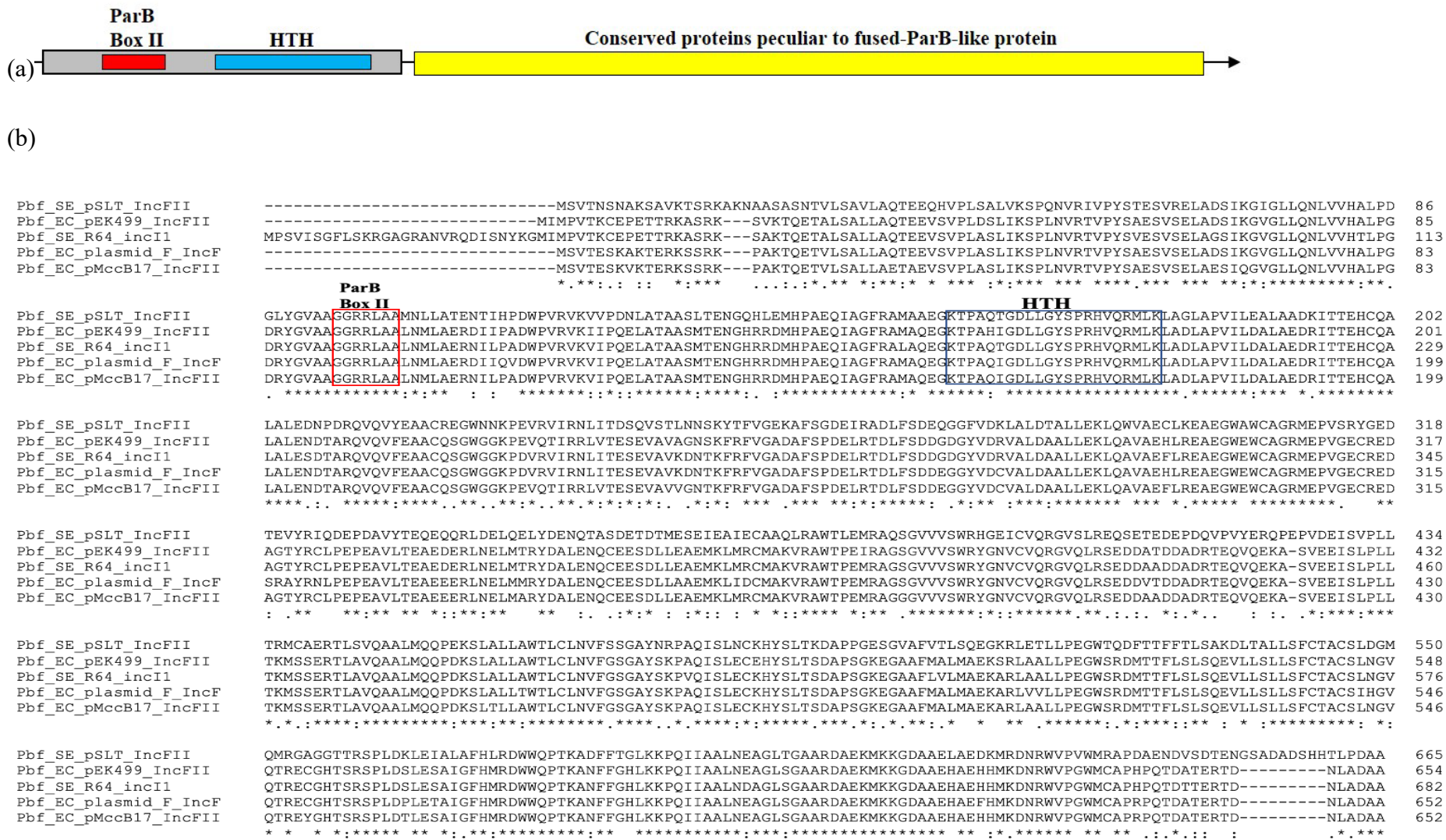


Figure 3.11 (a) An overview of the fused-ParB-like protein. The red box represents ParB Box II, while the blue represents helix-turn-helix domain (HTH), the grey box shows areas that are similar to a typical ParB protein, and the yellow box shows the area that is peculiar only to the fused-ParB-like protein. (b) MSA of selected fused-ParB-like protein showing location and sequences of ParB Box II in red and HTH in blue. In the amino acid sequence name, EC = *Escherichia coli* and SE = *Salmonella enterica*.

Table 3.5 Neighbouring genes common to Pbf (*yfbB*)

Plasmid	Genes upstream		Genes downstream	
	<i>yubL</i>	<i>ssb</i>	<i>psiAB</i>	<i>hok/sok</i>
pMccB17	✓	✓	✓	✓
F plasmid	✓	✓	✓	✓
pSLT	✓	✓	✓	✓
pEK499	✓	✓	✓	✓
pEFC36a	✓	✓	✓	✓
pJIE512b	✓	✓	✓	✓
R64	✓	✓	✓	✓
pP6qnrS1	✓	✓	✓	✓

Pbf (652 amino acids) is significantly longer than the typical ParB protein (about 300 amino acids) due to the large C-terminal domain of unknown function (Fig. 3.11a). The Pbf protein was mostly found in IncF plasmids but not restricted to this group as they were also found in IncI1 and rarely in IncR plasmids (one representative present in available sequence databases), and predominantly in the order Enterobacterales with few homologous found in Gram-positive bacteria (Fig. 3.12). In the phylogenetic tree (calculated using Maximum Likelihood method and Le and Gascuel (L+G) model) in Fig. 3.12, all the represented Pbf homologous were closely related except for the one found in plasmid pSLT, an out-group, which is 64% identical to Pbf. The Pbf protein encoded by pMccB17 has the following genetic context: *yubl*, *ssb*, *gene_73* and *rsmA* are immediately upstream, while *psiAB* and *mok* are downstream (Fig. 3.3). In other plasmids that have the ParB-fusion protein, this arrangement is absolutely conserved upstream and downstream of the Pbf protein (Table 3.5). This conserved genomic arrangement was also confirmed using the webFlaG tool for graphical visualization of the gene neighbourhood and its conservation (Fig. 3.13) (Saha *et al.*, 2021). Notably, this conserved genetic neighbourhood does not include an obvious ATPase gene (i.e. *parA* which encodes an essential protein partner and hence is an expected genomic neighbour of *parB*).

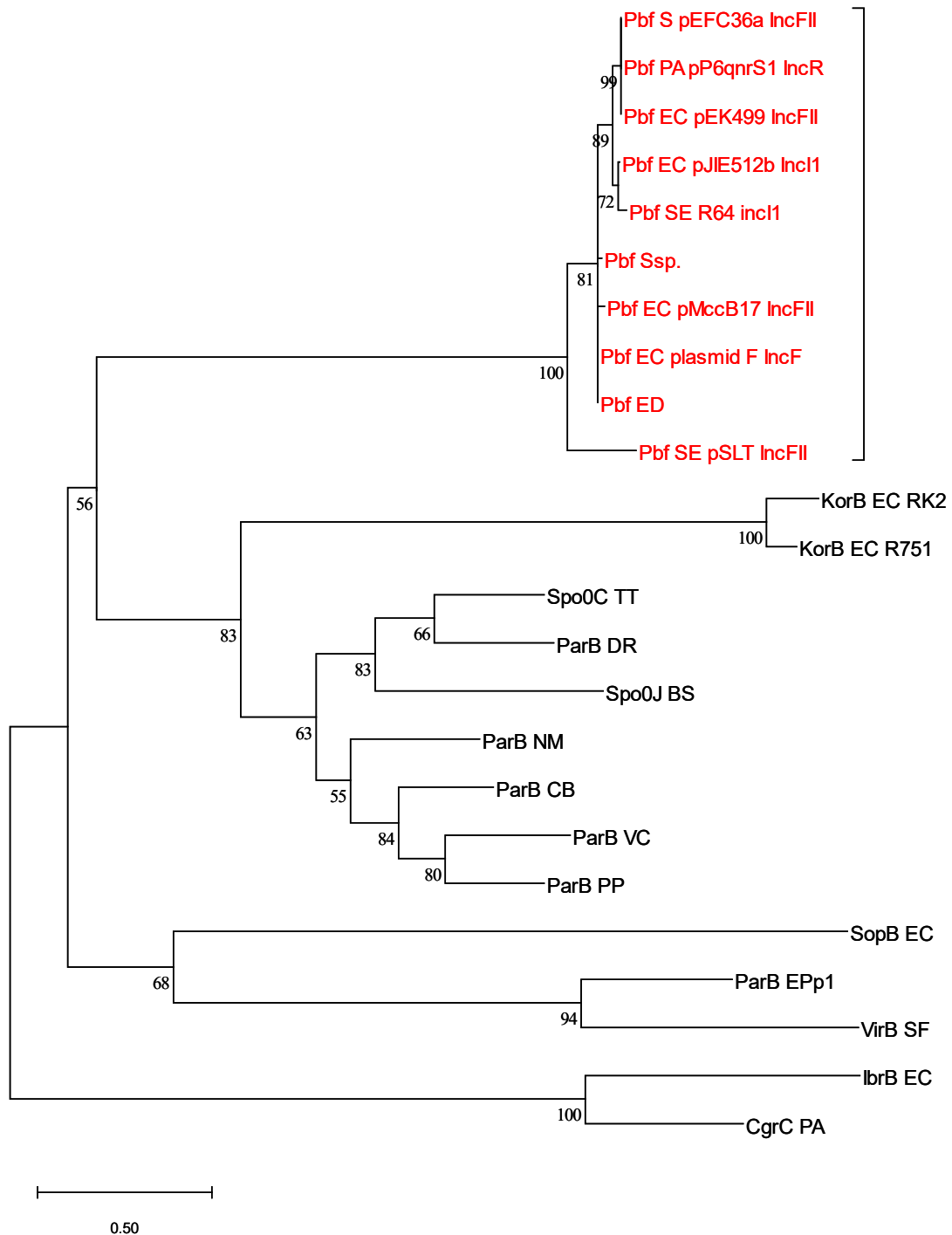


Figure 3.12 Phylogenetic analysis of ParB family proteins. Pbf proteins are highlighted in red. Protein sequences were obtained from EBI database. The clades are named in this order:

name of the ParB family protein, two letters representing the organism it is found, plasmid it is found in, and plasmid incompatibility group (only in Pbf). EC = *Escherichia coli*, SE = *Salmonella enterica*, S = *Salmonella*, PA = *Pseudomonas aeruginosa*, Ssp. = *Streptomyces sp.*, ED = *Enterococcus durans*, TT = *Thermus thermophilus*, BS = *Bacillus subtilis*, EPp1 = *Escherichia phage P1*, SF = *Shigella flexneri*, CB = *Coxiella burnetii*, DR = *Deinococcus radiodurans*, VC = *Vibrio cholerae*, NM = *Neisseria meningitidis*, PP = *Pseudomonas putida*.

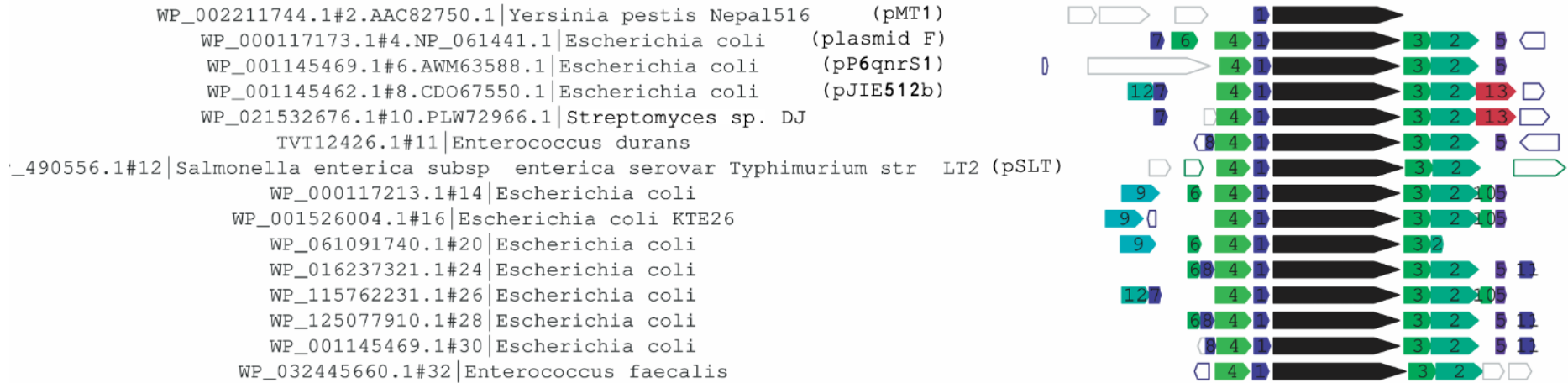


Figure 3.13 Conservation of neighbouring genes of *pbf*. *pbf* is shown in black while neighbouring genes are coloured and numbered: 1-*yubL*, 2-*psiA*, 3-*psiB* and 4-*ssb*. Image was generated using webFlaGs.

A BLASTp homology search, excluding Enterobacterales, revealed that Pbf is also encoded in other bacterial and eukaryotic genomes, although rarely. The BLASTp search (excluding uncultured models, coverage above 50% and E value ≤ 0.1) produced a total of 4826 hits, this was reduced to 8 hits by filtering the result to include only proteins that have a percentage identity greater than 40%. Two of these homologous proteins belong to Gammaproteobacteria (*Pseudomonas aeruginosa*, *Vibrio parahaemolyticus*), i.e. the same class as Enterobacterales, while 3 belonged to Gram-positive bacteria (*Streptomyces sp. DJ*, *Enterococcus durans*, *Listeria monocytogenes*), 2 to Eukaryotes (*Clonorchis sinensis*; liver fluke and *Callosobruchus maculatus*; cowpea weevil) and the last homologue is only defined as bacterial. We suspect that the presence of *pbf* homologous in eukaryotic genomes may represent “contamination” by sequences from gut bacteria, as this seems the most parsimonious explanation for their presence. Amidst these eight extra-Enterobacterales *Pbf* homologous, three are partial (shotgun) sequences, and the genomic neighbours they have depends on the parts (N or C terminal) they have missing. For instance, in *Listeria monocytogenes* and *Clonorchis sinensis*, the N-terminal region is missing as well as the neighbouring genes upstream (*yubL* and *ssb*), but the neighbouring downstream genes (*psiAB*) are present; while in *Vibrio parahaemolyticus*, the C-terminal is missing as well as neighbouring genes downstream while the neighbouring genes upstream are present. The only *pbf* homologous observed to lack the above-mentioned neighbouring genes is from the beetle *Callosobruchus maculatus* (again presumably a bacterial contaminant or symbiont or possibly an instance of horizontal genetic transfer from bacterium to beetle), where the homologous (50% identity) appears to be alone and not in a group of genes, with the adjacent genes upstream and downstream located over 6.5kb away. We currently have no indication as to the role of the gene *pbf*; however, from this work, it seems likely that it is involved in plasmid conjugation or maintenance.

Chapter four

Cloning and analysis of *Erwinia persicina* microcin operon.

4.0 Overview

This chapter focuses on a gene cluster found in *Erwinia persicina*, NBRC 102418 (the *Erwinia* microcin operon, *emo*), which is homologous to the *mcb* operon encoding biosynthetic pathway genes for microcin MccB17. These biosynthetic genes were hypothesized to direct the biosynthesis of a metabolite analogous to microcin B17 (here called EMB17). The metabolite, EMB17, was subjected to *in vivo* biological analyses, which provided further information about the metabolite, such as its optimal production conditions and activity spectrum. It was also subjected to chemical analysis to purify and determine its mass.

4.1 Discovery of *Erwinia microcin* operon (*emo* operon)

As there are lots of class-specific protein motifs, bioinformatics can be used to predict novel bacteriocins. One such bioinformatic tool recommended by de Jong *et al.* (2011) for bacteriocin prediction is BLAST, Basic Local Alignment Search Tool, which is used to search for similarities between protein or nucleic acid sequences (Altschul *et al.*, 1990).

BLASTp was used to search the non-redundant (nr) NCBI protein database for homologues of *E. coli* McbA. The search aimed to find a new bacteriocin related to the *E. coli* microcin B17. From the initial search results, *E. persicina* McbA was of interest for two reasons, (i) its low percentage identity, 88%, in relation to other results from the search and (ii) *E. persicina* was one of the least closely related organisms to *E. coli* in comparison to the list of organisms in the BLAST result. For these reasons, Lewis Bingle worked with Matthew Bouflower (a BSc student) and decided to look further at the McbA of *E. persicina*. In their study, they discovered that *Erwinia persicina* (NBRC 102418) had a gene cluster similar to the *mcb* operon in *E. coli*. Henceforth, we refer to them as *Erwinia mcb* operon (*emo* operon), and the synthesized microcin is referred to as *Erwinia* microcin B17 (EMB17) in this project (Fig. 4.1). These two operons differ in several ways aside from the obvious difference in the sequence of their *mcbA* gene (and its propeptide, McbA). *mcb* operon for a long time has been associated with *E. coli*; however, Metelev *et al.* (2013) reported homologous *mcb* operons in some species of *Pseudomonas* (*P. syringae* pv. *Glycinea* and *P. syringae* pv.

aesculi), and this was the model for our search. In this project, we report a homologous operon in *E. persicina*. Another major difference is that the *mcb* operon is found on a plasmid, whereas the *Erwinia* operon is located on the chromosome, similar to what is found in *Pseudomonas*.

4.1.1 Previous background work

Before this project commenced, some preliminary works were carried out by Bingle and Boutflower, which are summarized in this section. Establishing the similarities between the *E. coli* *mcb* operon and the *emo* operon using bioinformatic comparison tools to compare the sequences, a bioassay to investigate microcin production in *E. persicina*, in which the result showed that *E. persicina* is capable of producing an antimicrobial. The *emo* operon was amplified using PCR and ligated with pUC vectors (see more details in section 4.4). Bioassay experiments were set up to confirm that the cloned *emo* operons maintained their antimicrobial production ability, this bioassay was repeated in this project, and the results were in agreement.

4.2 An extra ORF included in the *E. persicina* microcin operon

An open reading frame (ORF) close to the *emo* operon, located downstream of *mcbG* was noticed during the examination of the *Erwinia mcb* operon (Fig. 4.1 and Fig. 4.6). The effect (if any) this predicted gene has on the microcin production was unknown. Further studies using bioassay experiments (Fig. 4.7) reveal that the presence of this extra gene in the *Erwinia mcb* operon seems not to affect microcin production significantly. Comparing the sequence of the extra ORF with the sequences in the non-redundant databases using BLASTp maintaining default parameters, it was found that this ORF is present mainly in the *Erwiniaceae* family (Table 4.1). From the result, there was a total of 106 hits; 1 was for *Enterobacter cloacae*, 15 were organisms from the genus *Kosakonia*, and the remaining belonged to the *Erwiniaceae* family. The hit results outside the *Erwiniaceae* family have a low percentage similarity of <55%. The ORF is without a known or predicted function, and the accession number for its encoded protein is WP_062748500. For this project, the gene will be referred to as *Erwinia mcb* operon extra gene (*eeg*).

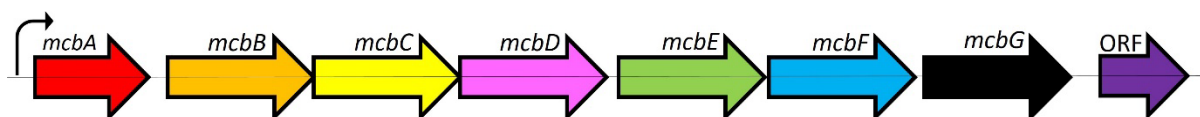


Figure 4.1 Overview of *Erwinia* microcin operon (*emo* operon) showing the location of the *eeg*. The black curve arrow represents the presumed location of *mcbA* promoter. Red, orange, yellow, pink, green, blue, and black arrows represent *mcbABCDEFG* respectively and the purple arrow represents the *Erwinia mcb* operon extra gene (*eeg*). Image adapted from Severinov *et al.* (2007).

Table 4.1 BLASTp search of *eeg* located next to the *E. persicina* microcin operon

Organisms*	Function	E value	Percent identity
<i>Erwinia persicina</i>	Hypothetical protein	6e-92	100.00%
<i>Erwinia billingiae</i>	Hypothetical protein	2e-88	98%
<i>Pantoea agglomerans</i>	Hypothetical protein	4e-86	94%
<i>Pantoea agglomerans</i>	Hypothetical protein	1e-85	94%
<i>Pantoea sp. Paga</i>	Hypothetical protein	1e-85	93%
<i>Pantoea agglomerans</i>	Hypothetical protein	2e-85	94%
<i>Pantoea agglomerans</i>	Hypothetical protein	2e-85	94%
<i>Pantoea vagans</i>	Hypothetical protein	2e-85	93%
<i>Pantoea agglomerans</i>	Hypothetical protein	2e-85	92%
<i>Pantoea agglomerans</i>	Hypothetical protein	2e-85	93%

* Results shown are the first 10 hits from the BLASTp search.

4.3 Comparison of *E. coli* and *E. persicina* microcin operon

The *E. coli* *mcb* operon has seven genes *mcb*ABCDEFG, with the promoter of the operon located upstream of *mcbA*; this configuration is also found in the *emo* operon (Fig. 4.1). Each gene in the *emo* operon encodes a protein highly similar to the corresponding *mcb* gene. However, some differences were noticed in their amino acid sequence as described below.

EMBOSS NEEDLE (Needleman and Wunsch, 1970) was used to compare the protein products from the two microcin operons using the default parameters. Overall, there is an average of 87% similarity between the amino acid sequences of the protein products from the two microcin operons from *E. persicina* and *E. coli* across the seven genes in the operon. The degree of similarity suggested that they perform similar functions and encode for the biosynthesis of similar RiPPs (Table 4.2).

Table 4.2 Functional description and comparison of proteins encoded by microcin operons of *E. coli* and *E. persicina*

Protein	Predicted Function	Length		Similarity (%)	Identity (%)
		<i>Erwinia</i>	<i>E. coli</i>		
McbA	Microcin precursor	72	69	91.7	88.9
McbB	Microcin B synthase enzyme complex	284	295	83.1	69.2
McbC		268	272	88.2	81.6
McbD		395	396	89.4	81.1
McbE	Self-immunity and export of MccB17	241	241	84.6	75.9
McbF	Self-immunity and export of MccB17	242	247	84.2	73.7
McbG	Self-immunity.	189	187	85.7	74.1
Average				86.7	77.8

4.3.1 Comparison of McbA peptides

McbA is an important product from the microcin operon because it is the promicrocin that is modified to form the matured microcin. There is a difference of 3 amino acids between the length of the two McbA peptides: *E. coli* McbA has 69 amino acids while the McbA from *E. persicina* has 72, as shown in Fig. 4.2. The additional 3 amino acids in *E. persicina* McbA are glycines (Gly) which are inserted within a polyglycine linker. This polyglycine linker in *E. coli* MccB17 (Gly30 to Gly39) is believed to act as a spacer that influences the correct positioning of the heterocycles (Sinha Roy *et al.*, 1998). This spacer is elongated in *E. persicina* McbA, ranging from Gly30 to Gly42. The effect this will have on *E. persicina* microcin is yet to be discovered; however, Sinha Roy *et al.* (1998) found that there is a 2-fold increase in the rate of product formation where the polyglycine linker has an extra glycine compared to wild type (from G₁₁SC to G₁₀SC) and shortening the polyglycine linker (for instance, from G₁₀ to G₉) reduces product formation. Since *Erwinia* promicrocin, McbA, has 3 extra glycines more than *E. coli* promicrocin, we expect that the product formation rate will be correspondingly higher in *Erwinia*.

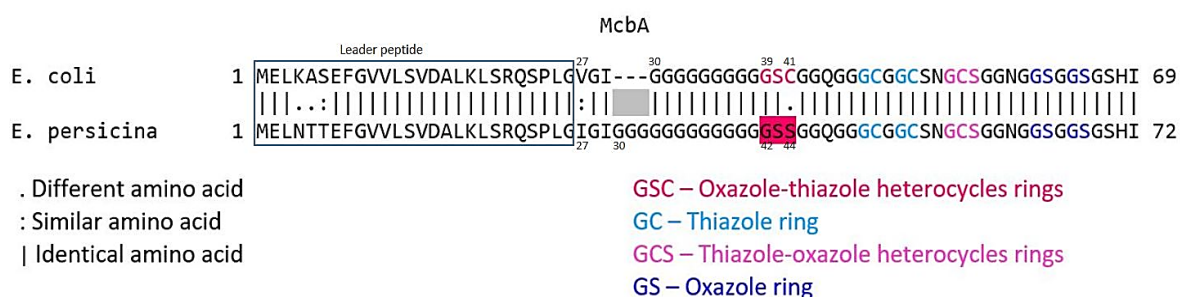


Figure 4.2 Alignment of McbA from *E. coli* and *E. persicina* amino acid sequences. Insertion of 3 glycines in the *Erwinia* McbA homologues relative to *E. coli* is highlighted in grey. The dipeptides involved in ring formation (GC and GS) are in blue while the tripeptides involved in ring formation (GSC and GCS) are in red (single letter amino acid code).

There are 5 amino acid differences between the two McbA peptides. Two of these are similar in chemical properties, and as such, these substitutions may not cause any tangible difference in how the peptides operate. Two of the remaining three are located in the leader peptide of the McbA peptides and would not be expected to cause any differences between the processed microcins peptides because the leader peptide is detached later in the biosynthesis

of MccB17. The *E. coli* McbA propeptide has a motif FxxxL, between F⁸ and L¹², responsible for leader peptide binding to the synthetase complex (Ghilarov *et al.*, 2019). This motif is conserved in *E. persicina*, which is another indication that the difference noticed in the leader peptides of the two McbA peptides might not significantly affect the microcin produced or the process involved. This was also shown in chapter 4, where non-cognate microcin operons (containing microcin biosynthetic genes from *E. persicina* and *E. coli*) were compared with cognate operons (containing microcin biosynthetic genes from a single source) to observe microcin production from both types of operons. We found that *E. persicina* synthase (McbBCD) can interact with *E. coli* McbA and vice versa to synthesise microcin. This confirmed that the difference between the two McbA peptides would not affect the synthase's specific recognition of the promicrocin. The remaining amino acid stands out based on its position also because it is not considered to be similar to the amino acid it is substituting in *E. coli*. Serine in position 44 in *E. persicina* McbA and Cysteine in position 41 in *E. coli* McbA both forms part of the heterocycle; the effect these differences will have on the microcins is not fully known, but some useful predictions can be made.

The predicted heterocycle formed at this position in *E. persicina* is bisoxazole, which results from G⁴² S⁴³ S⁴⁴ (Sinha Roy *et al.*, 1999). Sinha Roy *et al.* (1999) mentioned that there is a lower yield when microcin contains monooxazole or bisoxazole at a site that originally had either oxazole-thiazole bisheterocycle or thiazole-oxazole bisheterocycle (Fig. 4.3). We can see from the amino acid sequence alignment that the oxazole-thiazole bisheterocycle G³⁹ S⁴⁰ C in *E. coli* McbA peptide was replaced by a bisoxazole G⁴² S⁴³ S⁴⁴ in *E. persicina*, this could therefore mean that there will be a low yield in the EMB17 when compared to MccB17 unless there are compensatory mutations in the synthase. In line with this, it was reported that the cyclization of serine to oxazole is slower than the cyclization of cysteine to thiazole by 100-fold (Belshaw *et al.*, 1998). Calculating the mass of the matured and active EMB17 using EMBOSS Pepstats (Madeira *et al.*, 2022) and considering that there is a loss of 20 Da for every heterocycle that is formed (Metelev *et al.*, 2013), the mass of matured and active EMB17 will be about 3094 Da which is close to the mass of MccB17 3093 Da. It is important to point out that the mass of matured EMB17 given here is a calculated and not an experimentally-derived value to compare the masses of the two peptides.

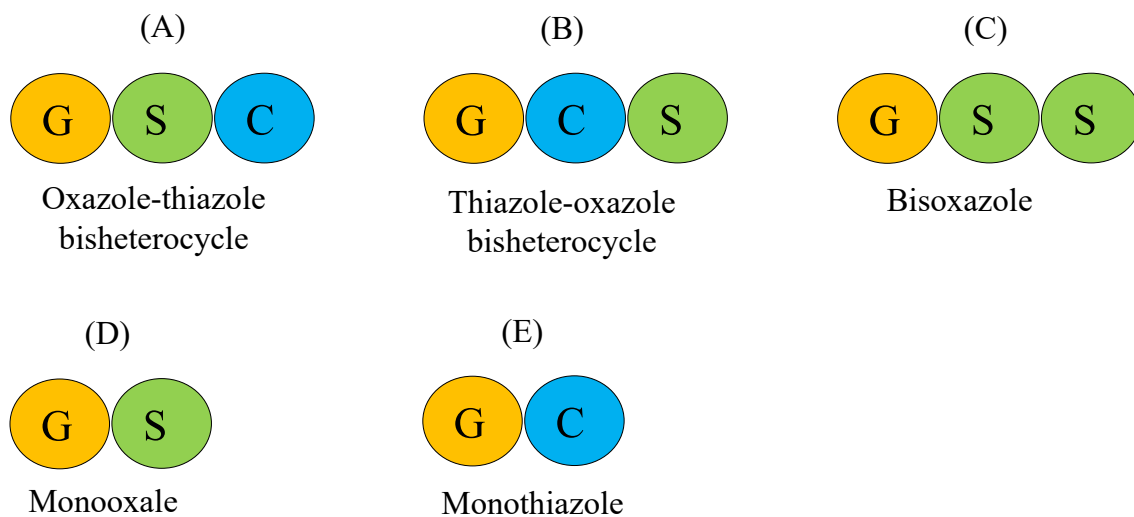


Figure 4.3 Amino acid composition of the types of rings (heterocycles) formed during cyclization of *Erwinia* and *E. coli* McbA. A, B and C are tripeptides that are modified to form heterocycles, while D and E are dipeptides that result into the named heterocycles after modification.

4.3.2 Comparison of McbBs proteins

From the comparison result in Table 4.2, McbB from both *E. persicina* and *E. coli* have the lowest percentage identity, and as such, we took an interest in examining the amino acid sequences of the two McbBs. The main difference between the 2 McbBs is the gap of 11 amino acids absent at the beginning of *E. persicina* McbB, which explains the difference in length of the two McbB proteins (295 in *E. coli*, 284 in *E. persicina*) (Fig. 4.4). This is in addition to the 39 different amino acids substituted between the two that do not have similar chemistry.

E. coli McbB is reported to be a homologue of E1/ubiquitin ligase-like proteins, which contains RiPP recognition element (RRE) that functions to mediate the recognition of leader peptide in McbA (Ghilarov *et al.*, 2019). The *E. persicina* McbB was checked for the RRE domain using the HHpred Web tool (<https://toolkit.tuebingen.mpg.de/tools/hhpred>; Söding *et al.*, 2005), and the result showed that this domain is present in *E. persicina* McbB as well. This is an indication that the *E. Persicina* McbB retains its function despite the difference with *E. coli* McbB.

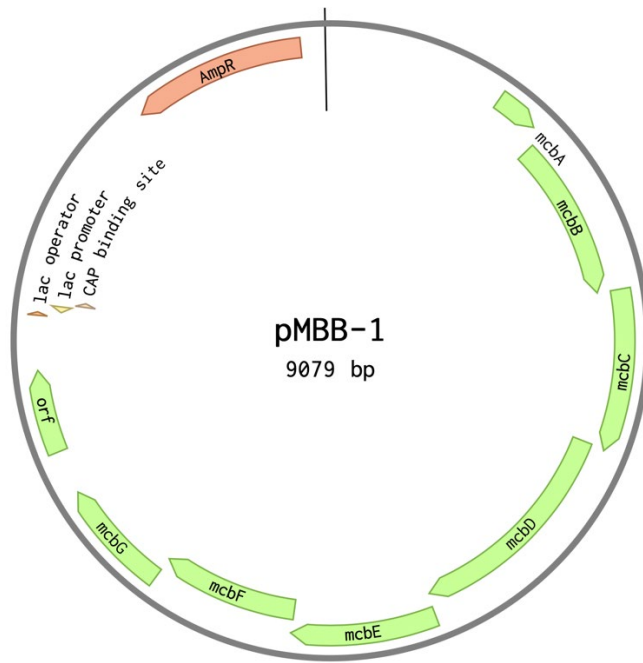
4.4 Cloning of *Erwinia persicina* microcin operon

The *Erwinia mcb* operon along with flanking intragenic regions, was amplified by PCR. BamH and KpnI restriction enzyme recognition sequences were introduced at either end of the amplicon. As it was unclear from the initial sequence analysis whether the *eeg* was part of the *emo* operon, the *eeg* was included in some of the clones that were constructed. The amplicons were ligated with pUC18 and pUC19 cloning vectors. Thus, there were 4 different resulting constructs (Table 4.3 and Fig. 4.5). To confirm that the cloning was successful, the constructs were digested with BamHI and KpnI and ran on agarose gel, they were also sequenced using Sanger methodology and next-generation short read technology (illumina) to check the correct insertion.

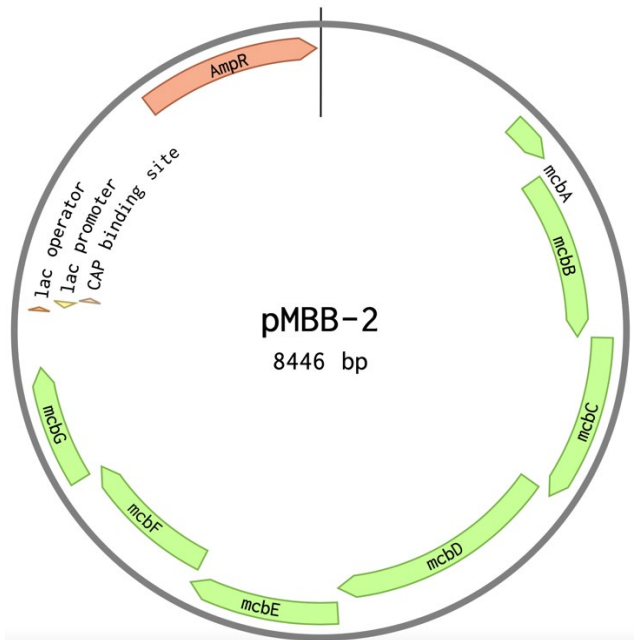
Table 4.3 Assigned names of cloned vectors with and without the *eeg*

Clones name	Vector	<i>Erwinia mcb</i> operon
pMBB-1	pUC18	<i>emo</i> operon + <i>eeg</i>
pMBB-2	pUC18	<i>emo</i> operon
pMBB-3	pUC19	<i>emo</i> operon + <i>eeg</i>
pMBB-4	pUC19	<i>emo</i> operon

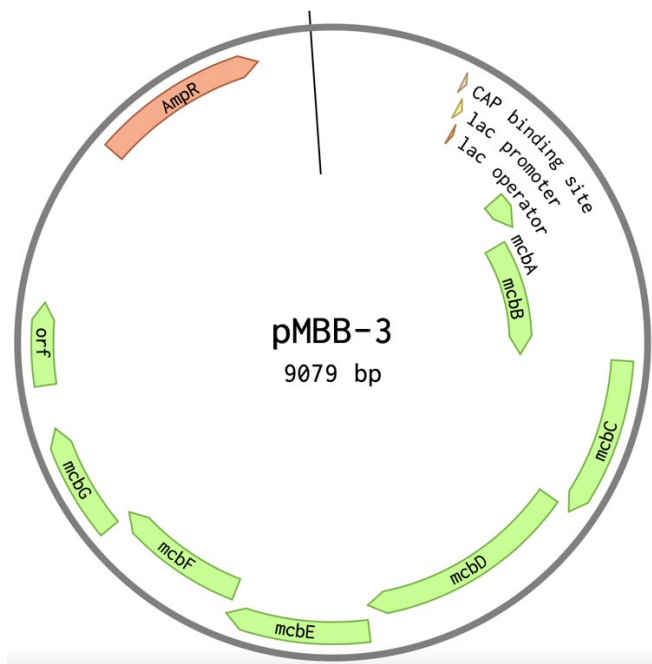
(A)



(B)



(C)



(D)

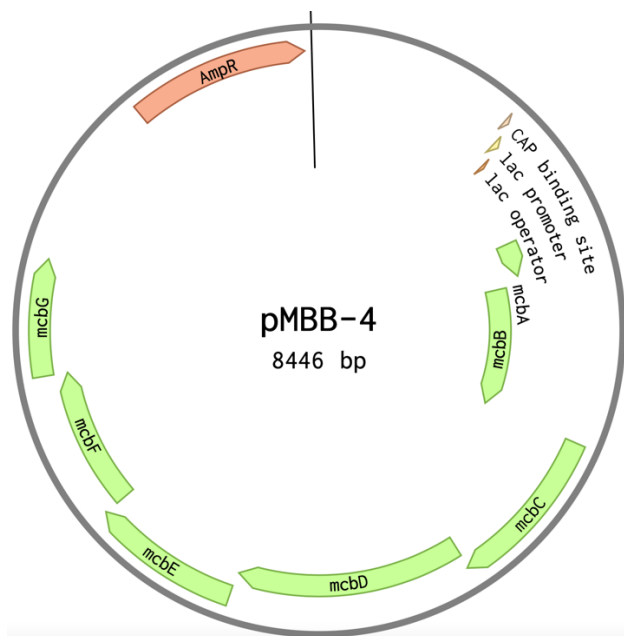


Figure 4.5 Maps of the cloned *emo* operon showing complete operon plasmid constructs. (A) pMBB-1 (B) pMBB-2 (C) pMBB-3 (D) pMBB-4. In red is the ampicillin gene (AmpR)

which is part of the pUC vector, in green is the microcin gene cluster from *E. persicina*. A and B are derived from pUC18 vector which differs in orientation from C and D (pUC19) in relation to the *lac* promoter. The *lac* promoter is in opposite direction to the operon in the pUC18 clones as opposed to being in the same direction as the pUC19 clones. Images were generated using Benchling.

4.5 Next generation sequencing of pMBB-3

The plasmid pMBB-3 (which consists of the *emo* operon with *eeg* cloned into the pUC19 vector) was extracted using a plasmid miniprep kit and sequenced by MicrobesNG, UK using a next-generation short read technology (illumina). The genes from the cloned BamHI-KpnI fragment derived from *Erwinia* in the resulting sequence were compared to sequences in the nr/nt database using BLASTn maintaining default parameters, and there was a 100% identity match with the published sequence of *E. persicina* NBRC 102418. This confirmed that all the genes and sequences present in the cloned operon 100% match those in *E. persicina* NBRC 102418. This was more detailed as it would have flagged up mutations if any were present compared to section 4.3, where Sanger sequencing was used only to check if the operon was cloned successfully. With the sequence from the NGS, Artemis was used to generate an image of the arrangement of the *emo* operon to include the *eeg* (Fig. 4.6).

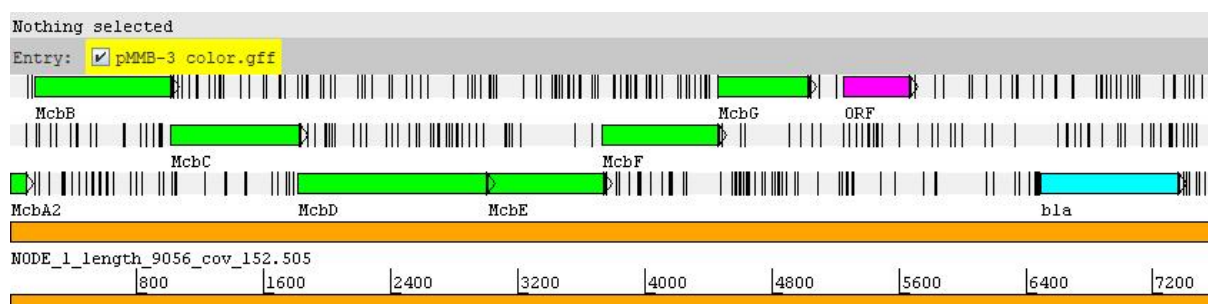


Figure 4.6 pMBB-3 showing the microcin genes (*mcbA – mcbG*) and the *eeg* cloned into pUC19 vector. bla represent the ampicillin resistance gene marker in pUC19 plasmid vector.

4.6 Bioassay of antimicrobial activity of the EMB17

The bioassay experiment was designed to determine the potency of EMB17 against lawns of sensitive cells by spotting 3 μ l of aliquot on the lawns and observing for halos of growth inhibition on the lawns. Preliminary data were obtained in 2017 by Bingle and Boutflower, and this was verified as described below. The *emo* operon clones (Table 4.3) were transformed into a heterologous host, *E. coli* DH5 α cells, for expression of EMB17. Although the bioinformatic analysis above in section 4.4 suggests that *emo* operon can produce MccB17-like products, however, the biological activity of these metabolites was unknown – hence the need for bioassays (see session 2.2.16 for more details on bioassay experiments).

A pUC19 plasmid vector without an insert was transformed into *E. coli* BL21 and was employed as the indicator strain. The reason for this is that all the bacterial test samples (pMBB-1, pMBB-2, pMBB-3, pMBB-4 and pPY113) except for the pMccB17 strain are ampicillin resistant. The presence of pUC19 in the indicator strain thus eliminates from the result interference of the antibiotic used to select for producer plasmid carriage. Fig. 4.7 shows the result of the bioassay experiment setup to confirm microcin production in the 4 clones. The pUC19 clones (pMBB-3 and pMBB-4) show a slightly higher level of microcin production when compared to the pUC18 clones (pMBB-1 and pMBB-2). Using a 2 tailed t-test, with $p \leq 0.05$, the means of the zone of inhibition between the pUC18 and pUC19 clones (pMBB-1 and pMBB-3; pMBB-1 and pMBB-4; pMBB-2 and pMBB-3; and pMBB-2 and pMBB-4) are significantly different; however, there is no significant difference between the clones of the same vector types (pMBB-1 / pMBB-2, and pMBB-3 / pMBB-4). This variation in the microcin level could be due to the different arrangement of the lac promoter in the 2 vectors – in pUC19, the lac promoter is in the same orientation as the cloned operon and so many boost transcription (Fig. 4.5). The *eeg* in the cloned pMBB-1 and pMBB-3 seems not to affect microcin production or activity. From the four clones, pMBB-3 was chosen for further studies because the pUC19-derived clones produce slightly higher microcin activity than the pUC18 clones.

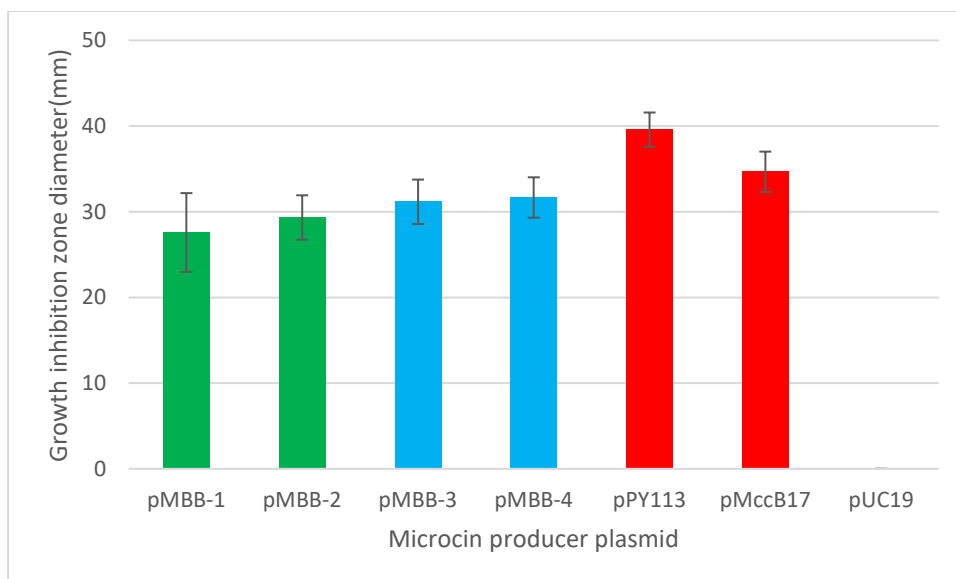


Figure 4.7 *In vivo* activity of microcin production using *E. coli* BL21[pUC19] as the indicator strain. In green are the pUC18 clones, in blue are the pUC19 clones, in red are the positive controls and to the right of the chart is pUC19 with no insert to serve as a negative control. Mean values and standard deviations (error bars) were obtained from two experiments, each performed in triplicate.

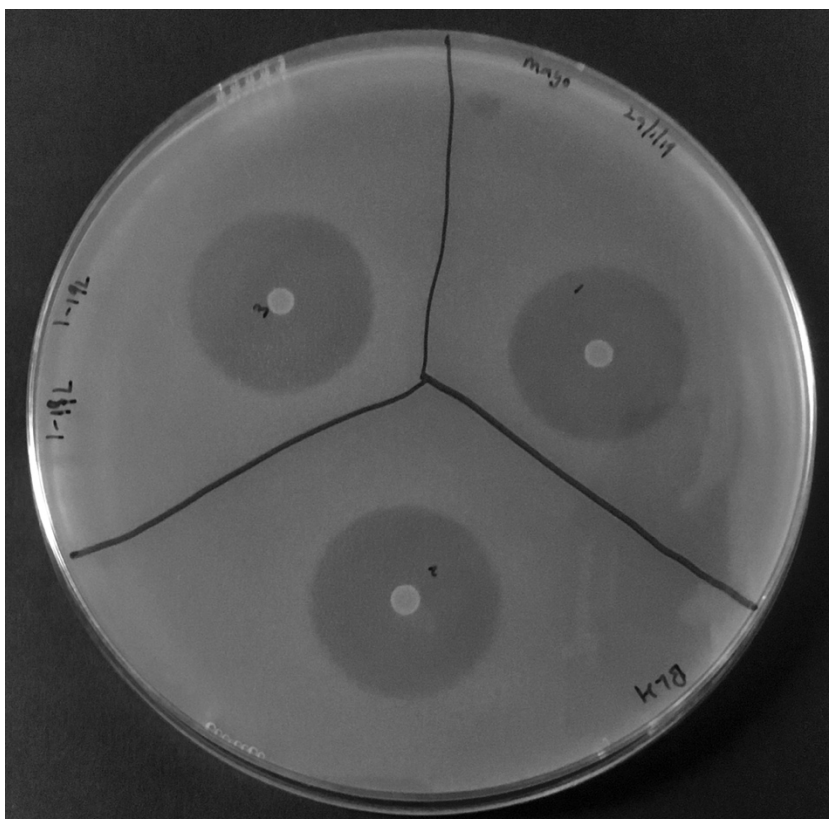


Figure 4.8 Bioassay of microcin production by cloned *emo* operon. 3 μ l drops of an overnight culture of *E. coli* DH5 α [pMBB-3] were spotted onto M63 agar. After overnight growth, the M63 plates were overlaid with 100 μ l overnight *E. coli* BL21 culture in 5ml soft agar. Diameters of growth inhibition zones were determined after overnight growth at room temperature.

Other bioassay experiments conducted on the four clones involve replacing the indicator strain with some microcin producer cells *E. coli* BL21[pPY113], *E. coli* BL21[pMBB-3] and *E. coli* _RYC1000[pMccB17] to observe if a known microcin producer will be inhibited by another microcin producer. When these three microcin producers (*E. coli* BL21[pPY113], *E. coli* BL21[pMBB-3] and *E. coli*-RYC1000[pMccB17]) were used as an indicator strain in the experimental bioassay setup, no zone of inhibition was noticed on the lawns. This is presumably because of the protective functions of McbE, McbF and McbG, which work collectively to offer self-immunity to the producer cell and confirms that these proteins also confer immunity to the related non-cognate microcin.

4.7 Characterization of *Erwinia microcin B17* (EMB17)

In order to better understand the production and activity of EMB17, additional bioassay experiments were conducted to characterize the growth conditions of the host cell carrying the *emo* operon and optimal production conditions for EMB17, as well as its spectrum of activity.

4.7.1 Effects of IPTG on microcin production

The *emo* operons were cloned into pUC18 and pUC19 vectors, which both have *lac* operon. This section will examine the effect of chemical induction with isopropyl β -D-1-thiogalactopyranoside (IPTG) on the expression of EMB17 by the *emo* operon clones. Three concentrations of IPTG 0mM, 0.1mM and 1mM were used for this study (Fig. 4.9). Using pairwise comparisons, 2 tailed t-test, with $p \leq 0.05$, indicated that there were no significant differences in inhibition zone diameter due to IPTG concentration IPTG for each clone.

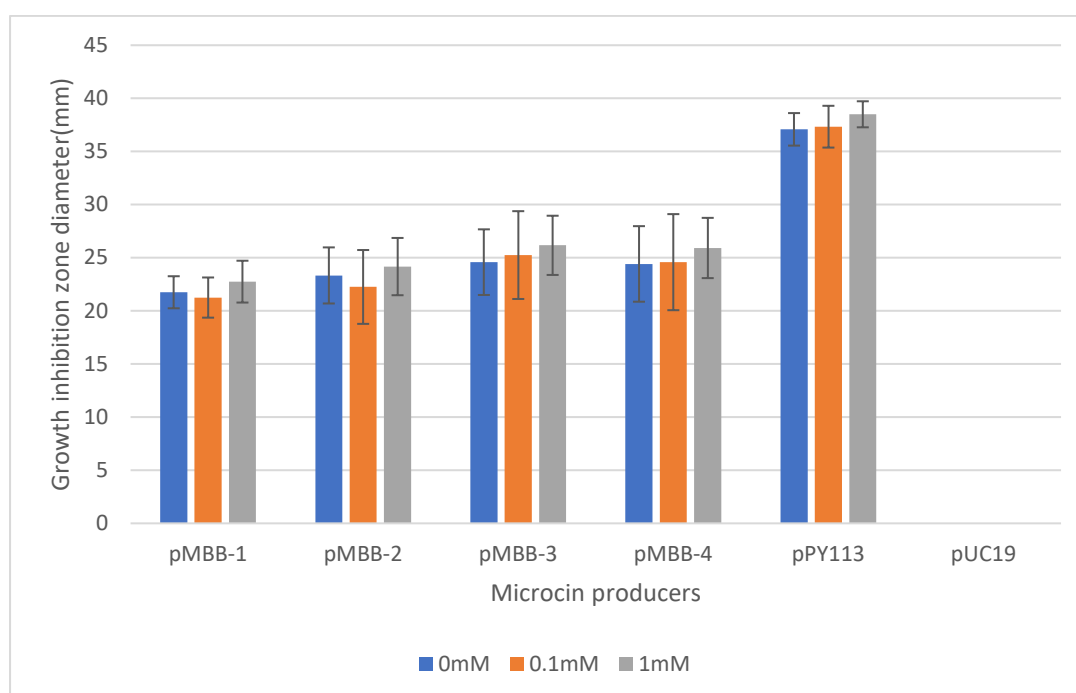


Figure 4.9 Bioassay experiment to study the effect of IPTG on microcin production for the *emo* operon clones. Each clone was exposed to IPTG concentrations. Mean values and standard deviations (error bars) were obtained from three experiments, each performed in duplicate.

4.7.2 Time course optimization experiment

To identify the optimum growth conditions for the cloned *emo* operon in a heterologous host (*E. coli* DH5 α [pMBB-3]) as well as the optimum conditions for microcin production, a time course microcin production experiment was set up. Two media, LB (rich medium) and M63 (minimal medium), were employed for this experiment, as well as the addition of 0.1mM IPTG to the media to examine the effect of induction of the *lac* promoter on the *emo* operon. Growth was determined by taking the optical density (OD₆₀₀) of the growth culture at the stipulated time.

The growth curve shows that the cloned *emo* operon in a heterologous host (*E. coli* DH5 α [pMBB-3]) have a longer doubling time in LB medium and M63 minimal medium, as well as lower cell density in stationary phase when compared to *E. coli* DH5 α [pPY113] and *E. coli* DH5 α [pUC19] (Fig. 4.10). This could be because the *emo* operon is not in its original host and thus the plasmid containing the operon exerts some burden on the full functioning of the heterologous *E. coli* host. Another reason could be because the *E. persicina* McbG does not work well with the *E. coli* gyrase and does not provide full immunity. The addition of IPTG to the growth medium had no noticeable effect on the growth of the heterologous host, as seen in Fig 4.10. Producer cells grown in M63 minimal medium have a lower cell density than those grown in LB medium (Fig. 4.10 and Fig. 4.11). LB media supports optimal growth of the cloned operon as it is a rich medium compared to M63, a minimal medium. Cell density was highest at the 24th hour in each case and plateaued after that.

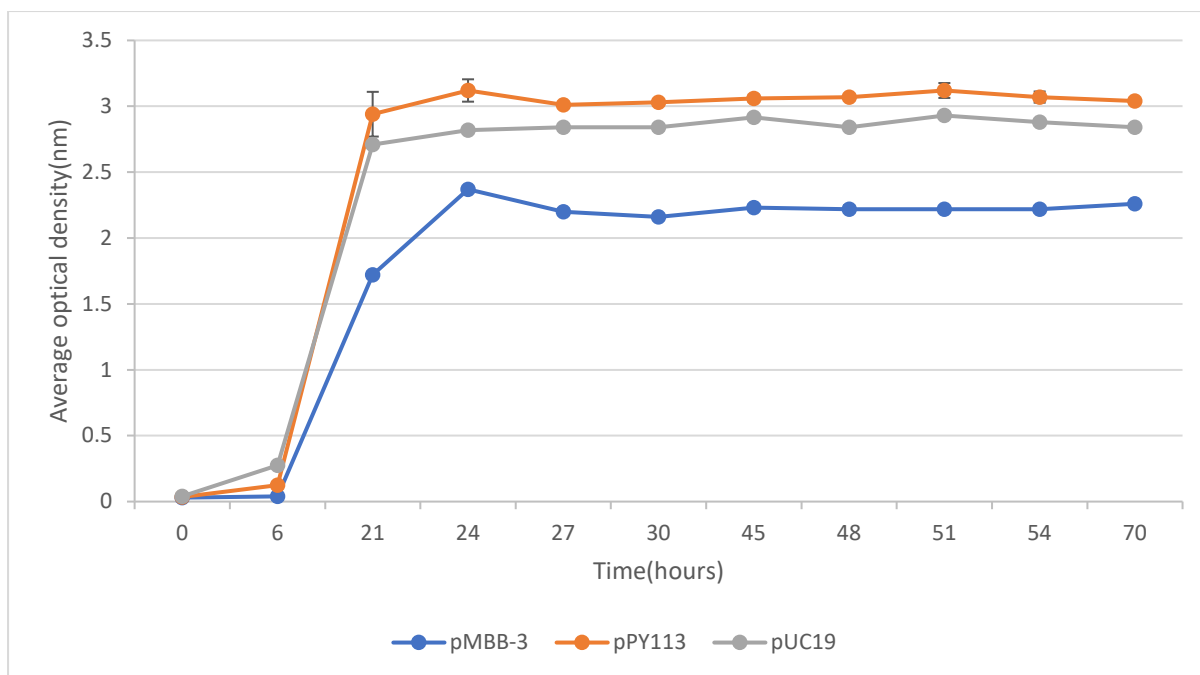


Figure 4.10 Bacterial growth curve for *E. coli* carrying pMBB-3, pPY113 and pUC19 in LB medium. pMBB-3 is a cloned *E. persicina* microcin operon, pPY113 is a known microcin producer (positive control) and pUC19 does not encode microcin (negative control). Mean values and standard deviations (error bars) were obtained from two experiments, each performed in triplicate.

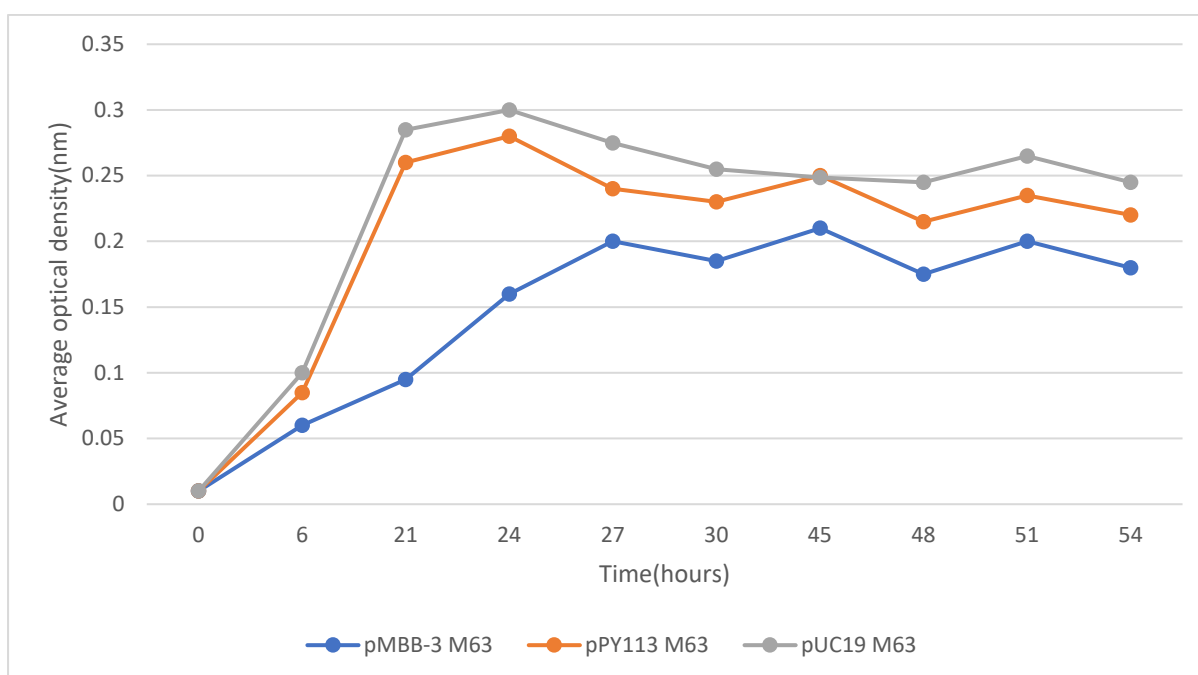


Figure 4.11 Bacterial growth curve in M63 medium. (A)pMBB-3 (B)pPY113 (C)pUC19. Mean values and standard deviations (error bars) were obtained from two experiments, each performed in triplicate.

Samples were collected at the same times as the optical density was read for the bacterial culture growth during the determination of cell growth density above. These samples were used to set up bioassay experiments to determine the zone of growth inhibition on sensitive cells. MccB17 (positive control) as with previous bioassay results, showed higher activity than EMB17 (Fig. 4.12). For the first 20th hour, there was no microcin production from the *E. coli* or *E. persicina* microcin producers, and this is the same for the two media used, LB and M63 minimal medium (Fig. 4.12). It appears that the microcin production dropped at the 24th hour for EMB17 and the 27th hour for MccB17 in LB medium, while the decline was noticed at the 30th hour for both microcins in M63 minimal medium. In both microcins across the two media, production plateaued from the 48th hour onward. From the result, we were able to draw an inference that the optimum microcin production time is between the 21st to 24th hour, which appears to be the start of the stationary phase (Fig. 4.12). Previous studies have shown

that MccB17 biosynthesis is associated with the stationary growth phase (Pierrat and Maxwell, 2003; Sinha Roy *et al.*, 1999).

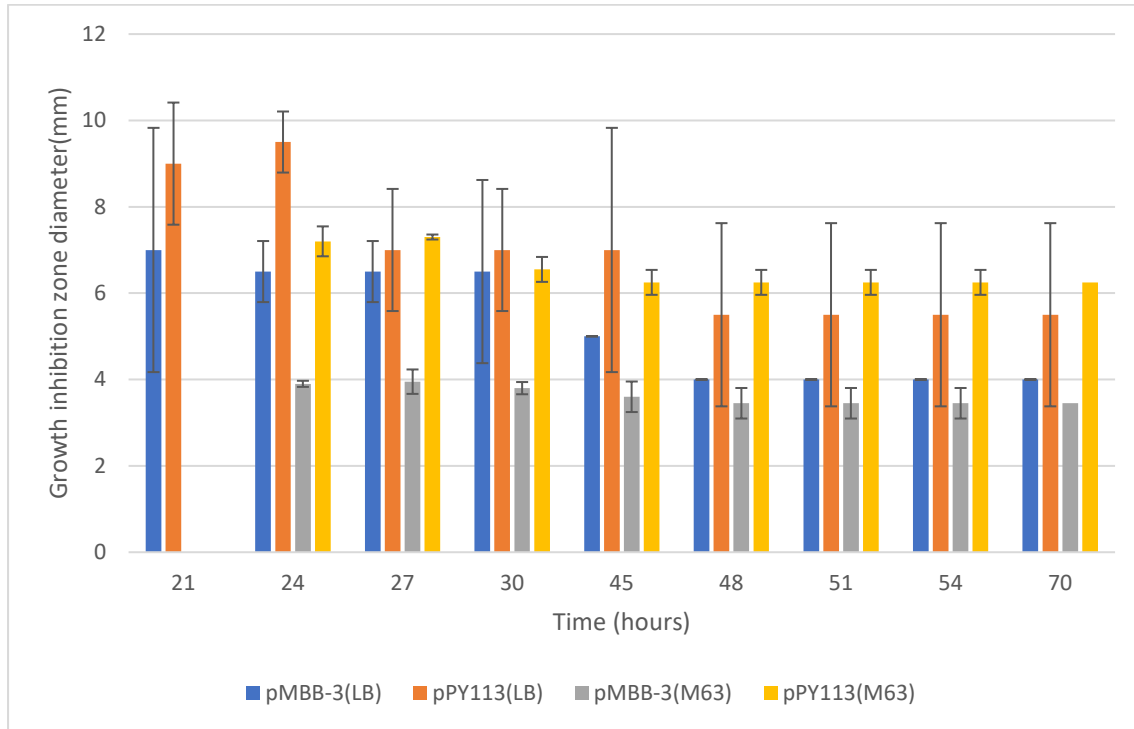


Figure 4.12 Time course study of microcin production assessed by bioassay in LB and M63 media. pMBB-3 is a cloned *E. persicina* microcin operon and a producer of EMB17, pPY113 is a known MccB17 producer (positive control) and pUC19 does not encode any microcin (negative control). Mean values and standard deviations (error bars) were obtained from 2 replicates.

LB medium supported a better yield of microcin production at the stationary phase, as shown in Fig. 4.13, where the microcin productions for pMBB-3 in both LB and M63 media were compared. The stationary phase seems to be reached at a similar time, and cell density is much higher (~10x) in LB medium. However, for ease microcin purification, minimal media are preferred to rich media because with the composition of rich media, there is high tendency of the column becoming clogged.

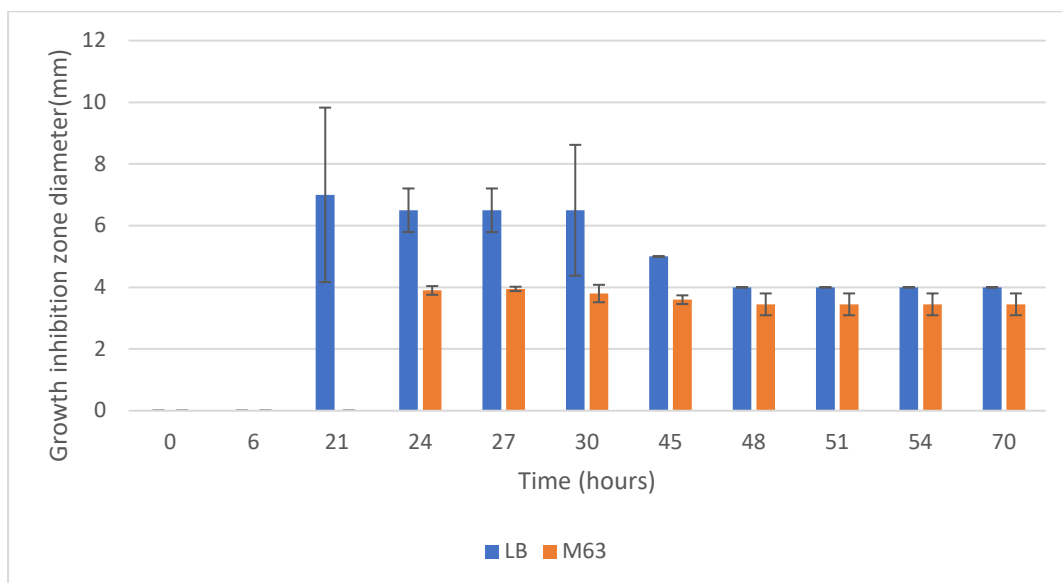


Figure 4.13 Time-course study comparing *in vivo* activity of pMBB-3 microcin production with time in LB and M63 medium. Mean values and standard deviations (error bars) were obtained from 2 replicates.

4.7.3 Examining mechanism of action via bioassay of deletion mutant strains

E. coli MccB17 has been extensively studied, and as such, there is a better understanding of how it works. In this section, we want to compare how EMB17 works in relation to what we know about MccB17 by utilizing some mutant strains (of genes that are involved in microcin metabolism or mechanism of action) in bioassay experiments. The *E. coli* Keio Knockout mutant collections were used for the bioassay in this section.

4.7.3.1 Study with $\Delta tldD$ producer cells

During MccB17 biosynthesis, the leader peptide is detached from the microcin by the action of TldD/TldE. The effect of the presence of TldD was studied on EMB17 production and compared to MccB17. The plasmids carrying the microcin operon in *E. persicina* and *E. coli* (pMBB-3 and pPY113 respectively), as well as pUC19 plasmid (control), were transformed into *E. coli* BW25113 cells (wild type strain from which the mutant strains were made) and *E. coli* BW25113 $\Delta tldD$ cells (Keio knockout with gene exchanged for KmR gene). The bioassay experiment was carried out using *E. coli* BL21[pUC19] as the indicator strain. Producer strains that lack TldD have lower production of active microcin. This could be

because TldD is partly redundant due to the absence of TldE. The result indicates that the presence of *tldD* increases microcin production (Fig. 4.14)

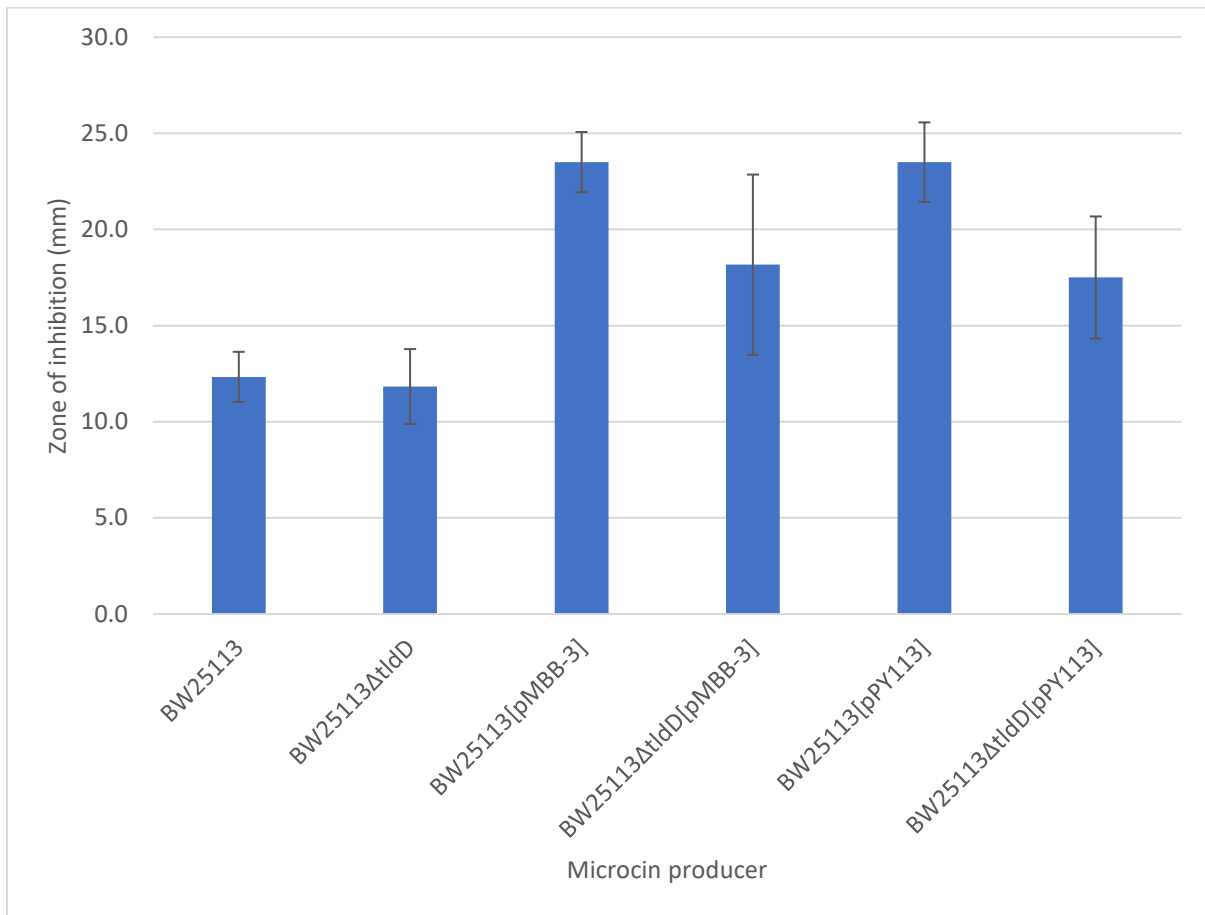


Figure 4.14 Effects of *tldD* on microcin production. Mean values and standard deviations (error bars) were obtained from 2 replicates.

4.7.3.2 Study with $\Delta ompF$ and $\Delta sbmA$ strains

Matured and post-translationally modified MccB17 gains entrance into the sensitive strain by passing through its outer membrane porin protein *OmpF* and inner membrane protein *SbmA*. Bioassay experiments were set up using the *E. coli* Keio Knockout mutant collection strains: *E. coli* BW25113 $\Delta ompF$ and *E. coli* BW25113 $\Delta sbmA$ as indicator strains to confirm that these membrane proteins function similarly in the *E. persicina* EMB17 as they do in the *E. coli* MccB17 to allow movement of MccB17 into their cytoplasm. *E. coli* BW25113 cell (Wild type) was used as a control. With *E. coli* BW25113 $\Delta sbmA$ being used as an indicator strain, there was no zone of inhibition, while the wild type and *E. coli* BW25113 $\Delta ompF$ both

have similar zone sizes (Fig. 4.15). During our study, we focus on mutation in the *ompF* gene; however, two genes, *ompF* (structural gene for OmpF) and *ompR* (a positive regulator gene), are responsible for the production of the outer membrane protein OmpF. This might explain why there was no reduced sensitivity when *E. coli* BW25113 $\Delta ompF$ was used as the indicator strain. A similar study revealed that some strains with the *ompF* or *ompR* mutation were partially resistant to MccB17 and that some strains with the *sbmA* mutation were completely resistant (Lavina *et al.*, 1986). It was proposed that SbmA recognises the heterocycles in the microcin structure, and therefore, its absence will make a sensitive strain resistant (Duquesne *et al.*, 2007; Yorgey *et al.*, 1994).

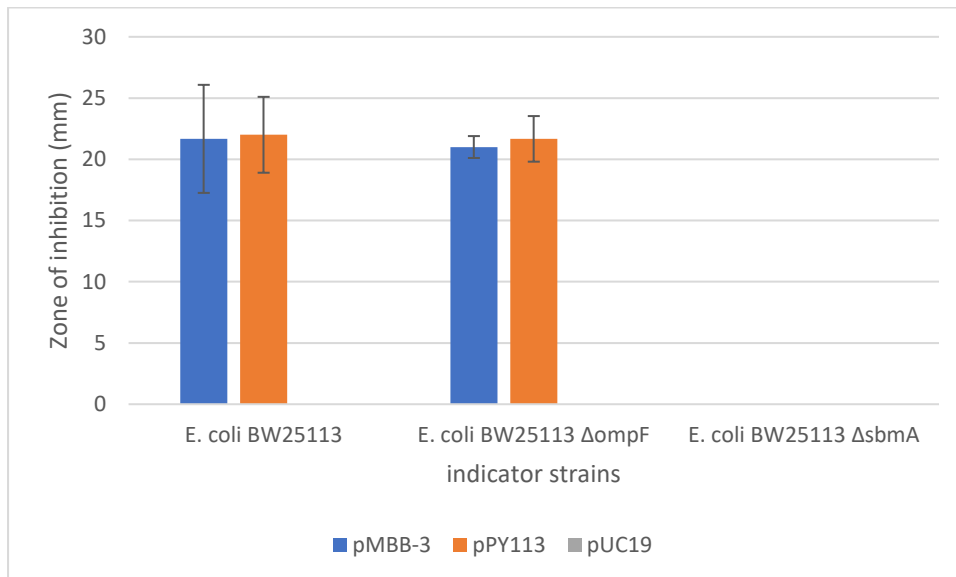


Figure 4.15 Effects of *ompF* and *sbmA* on the transport of EMB17 into sensitive strains. pMMB-3 is the cloned *emo* operon that synthesises EMB17, pPY113 synthesises MccB17 and serves as a positive control, and pUC19 is an empty plasmid vector without an insert and serves as a negative control. Mean values and standard deviations (error bars) were obtained from two experiments, each performed in triplicate.

4.7.4 Spectrum of action of EMB17

To have a clearer picture of the activity range of the EMB17, bioassay experiments were set up to include Gram-positive and Gram-negative bacteria as indicator organisms. From previous studies, it was reported that MccB17 actively inhibited the growth of the following genera *Escherichia*, *Citrobacter*, *Klebsiella*, *Salmonella*, *Shigella* and *Pseudomonas*

(Baquero and Moreno, 1984; Duquesne *et al.*, 2007). Similarly, EMB17 inhibited the growth of *Escherichia* (from results of a previous bioassay where *E. coli* BL21 was used as an indicator strain), *Klebsiella* and *Pseudomonas* but not *Salmonella* (Table 4.4). A phylogenetic tree of Enterobacteriaceae shows that *Salmonella* is closer to *Escherichia* than it is to *Erwinia* (Baumler *et al.*, 2013); this may account for the reason why EMB17 had no inhibition activity on *Salmonella* as microcins are mainly narrow spectrum and in most cases kill or inhibit closely related organisms. With *Pseudomonas aeruginosa* and *Klebsiella pneumoniae*, MccB17 showed higher activity in comparison to EMB17. The higher activity of MccB17 against *Klebsiella pneumoniae* is expected as *Klebsiella* is closer to *Escherichia* than it is to *Erwinia*. However, from taxonomy classification and the phylogenetic tree cited above, *Pseudomonas* is not close to *Escherichia* or *Erwinia*, the high activity (especially of MccB17) noticed against it was therefore unexpected. As expected, both EMB17 and MccB17 had no inhibitory activity on Gram-positive bacteria.

Table 4.4 Activity spectrum of EMB17

Test organisms	Gram variable	Microcin activity ^c		
		pMBB-3 (EMB17)	pPY113 (MccB17)	pUC19
<i>Pseudomonas aeruginosa</i> (NCIMB 8626)	Negative	++ ^a	+++	- ^b
<i>Klebsiella pneumonia</i> (NCTC 10314)	Negative	+	++	-
<i>Salmonella enterica</i> Tinda (NCTC 8147)	Negative	-	+	-
<i>Staphylococcus epidermidis</i> (NCTC 11047)	Positive	-	-	-
<i>Staphylococcus aureus</i> (NCTC 10788)	Positive	-	-	-
<i>Bacillus subtilis</i> 168 (ATCC 23857)	Positive	-	-	-

^a + gives the range of the diameter of the halo formed during the bioassay, +(5-15 mm), ++(15.5-29), +++(30>)

^b -no noticeable halo .

^c - Mean values were obtained from triplicate experiments.

4.8 Analysis and Purification of EMB17

From previous sections, we are aware that the *emo* operon is active when expressed in its natural producer (*E. persicina*) as well as in a heterologous host (*E. coli*). In this section, using a method based on (Parks *et al.*, 2007; Sinha Roy *et al.*, 1999), we attempted the

purification of crude EMB17 so that we could examine the activity of purified EMB17, as well as determining the size of EMB17 in relation to MccB17 using analytical techniques including liquid chromatography with mass spectrometry, and the information obtained from its amino acid sequence. The purification process involved a hot acid extraction step in harvesting the producer cells, solid phase extraction and preparative HPLC based on previously published approaches (Ghilarov *et al.*, 2011; Parks *et al.*, 2007; Shkundina *et al.*, 2014).

4.8.1 Antimicrobial activity of DMSO on *E. coli* BL21[pUC19]

DMSO is often utilized as a solvent to resuspend biological compounds. There have been reports that at high concentrations, DMSO has antimicrobial effects on microorganisms (Gadebusch and Basch, 1968; Kirkwood *et al.*, 2018; Wadhvani *et al.*, 2009). As a result, we were concerned that at high concentrations, it might interfere with the antimicrobial activity of some biological compounds on indicator organisms and therefore lead to unreliable results. Some reports had suggested concentrations or percentages of DMSO that are safe to resuspend biological compounds during antimicrobial testing against microorganisms without the DMSO having antimicrobial effects. An example is a study carried out by Ansel *et al.* (1969), where they examined the antimicrobial effect of DMSO on 3 organisms *E. coli*, *Pseudomonas aeruginosa*, and *Bacillus megaterium*; they concluded that an increase in concentrations to about 15% of DMSO decreases the growth rate of these organisms while concentrations above this level eliminate the organisms. In this project, DMSO was used to resuspend the dried extract from the solid phase extraction sample as well as the fractions collected after purifying EMB17 from prep HPLC. To rule out the effect of DMSO on the bioassay result in the sections below, a bioassay experiment was set up to study the effect DMSO will have on the *E. coli* BL21[pUC19] that was used as an indicator strain.

Four concentrations of DMSO (100%, 50%, 10% and 1%) and three sample volumes (1 μ l, 5 μ l and 10 μ l) were used for this study. No zone of inhibition was observed for the four concentrations with the change in sample volumes. This result was different from what Ansel *et al.* (1969) reported, where antimicrobial effects were observed on organisms with an increase in concentrations of DMSO. However, a similar study of 5 organisms *Staphylococcus aureus*, *Enterococcus faecalis*, *Escherichia coli*, *Pseudomonas aeruginosa* and *Proteus mirabilis* conducted by Brito *et al.* (2017) showed that concentrations of DMSO had no antimicrobial effects on these strains. As a result, the dried samples were resuspended in 100% DMSO.

4.8.2 Mass spectrometry analysis of EMB17

Mass spectrometry was used to elucidate the chemical properties of EMB17. After the microcin producer cells were harvested and solid phase extraction steps were completed, 2 μL of the partially purified EMB17 was injected on a liquid chromatography-mass spectrometer (LCMS) Agilent 1290 Infinity UPLC, 6120 Single qMSD (see sessions 2.2.17 and 2.2.18 for more details on solid phase extraction and LCMS). Tiny peaks were noticed in the chromatograms (data not shown); as a result, the subsequent injections were increased to 10 μL . In addition to the increase in the volume of injection, a pooled quality control sample containing EMB17 and MccB17 was made. The order in which the pooled QC samples along with other samples were run is shown in Table 4.5, and this is a similar method used by Broadhurst *et al.* (2018) in their studies. Following cell harvesting and solid phase extraction, the partially purified microcin from *E. coli* DH5 α [pMBB-3] and *E. coli* DH5 α [pPY113] cells were assessed for activity using a bioassay experiment to check that the extraction process had not damaged the microcin (Table 4.6).

The partially purified microcins were run on LCMS using an injection volume of 10 μL following the sampling order in Table 4.5. The pooled QC sample, biological blank, sample 1 (MccB17) and sample 2 (EMB17) were injected 3 times, and the column was washed with mobile phase blank to minimise compounds from different samples mixing and thus interfering with the result. The data generated from the third injection of pooled QC sample, biological blank, sample 1 (MccB17) and sample 2 (EMB17) were used to plot a single chromatogram (Fig. 4.16). The resulting chromatogram in Fig. 4.16 shows two peaks that were unique to the EMB17 sample at the 29th minute and 34th minute. When further observed, it was discovered that these peaks have the same mass of 1088.9 m/z in positive mode (Fig. 4.17). Further study of the fragmentation pattern (locations and types of bonds between molecules) of the peaks using the LCMS would have been a step further in determining the structure of EMB17; however, this was precluded due to instrument availability.

Table 4.5 Analytical run order of the pooled QC EMB17 and MccB17 sample.

Injection number	Sample type
1	Mobile phase blank ^a
2	Mobile phase blank
3	Mobile phase blank
4	Pooled QC sample ^b
5	Pooled QC sample
6	Pooled QC sample
7	Biological blank ^c
8	Biological blank
9	Biological blank
10	Mobile phase blank
11	Sample 1 – MccB17
12	Sample 1 – MccB17
13	Sample 1 – MccB17
14	Mobile phase blank
15	Sample 2 – EMB17
16	Sample 2 – EMB17
17	Sample 2 – EMB17
18	Mobile phase blank

^a Mobile phase blank (50:50 MeOH-Water)

^b Pooled QC sample (partially purified EMB17 + partially purified MccB17 in the ratio 1:1)

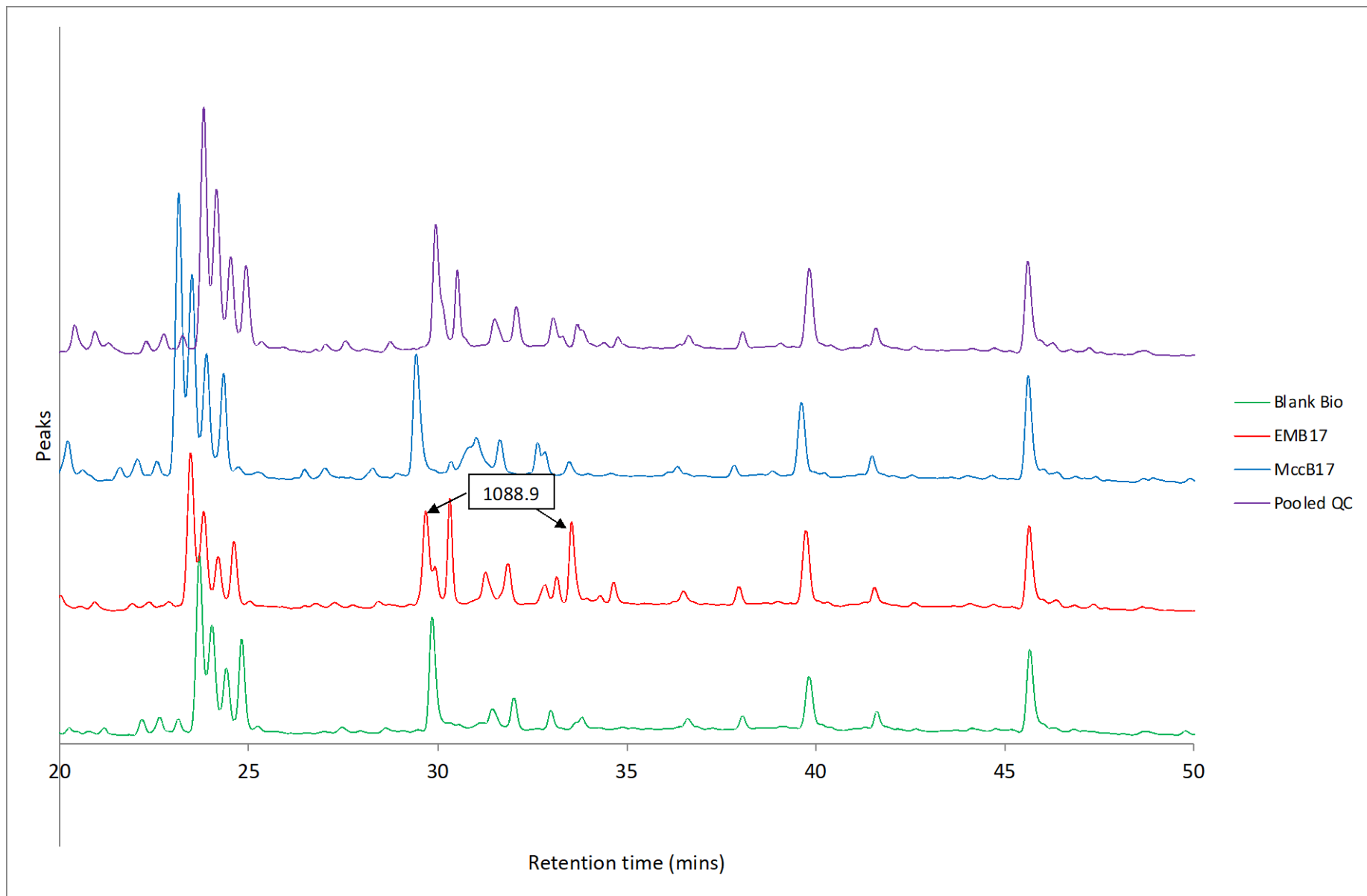
^c Biological blank - a bacteria cell (*E. coli* DH5 α [pUC19]) extraction that does not synthesis microcin.

Table 4.6 Bioassay result of partial purified microcin.

Microcin producer strain	Zone of inhibition (mm)
<i>E. coli</i> DH5 α [pMBB-3]	17
<i>E. coli</i> DH5 α [pPY113]	15
<i>E. coli</i> DH5 α [pUC19]	0

3 μ l drops of the partially purified microcins were spotted onto M63 agar and overlaid with 100 μ l overnight *E. coli* BL21 culture in 5ml soft agar. Diameters of growth inhibition zones were determined after overnight growth at room temperature. Mean values were obtained from triplicate experiments.

(A)



(B)

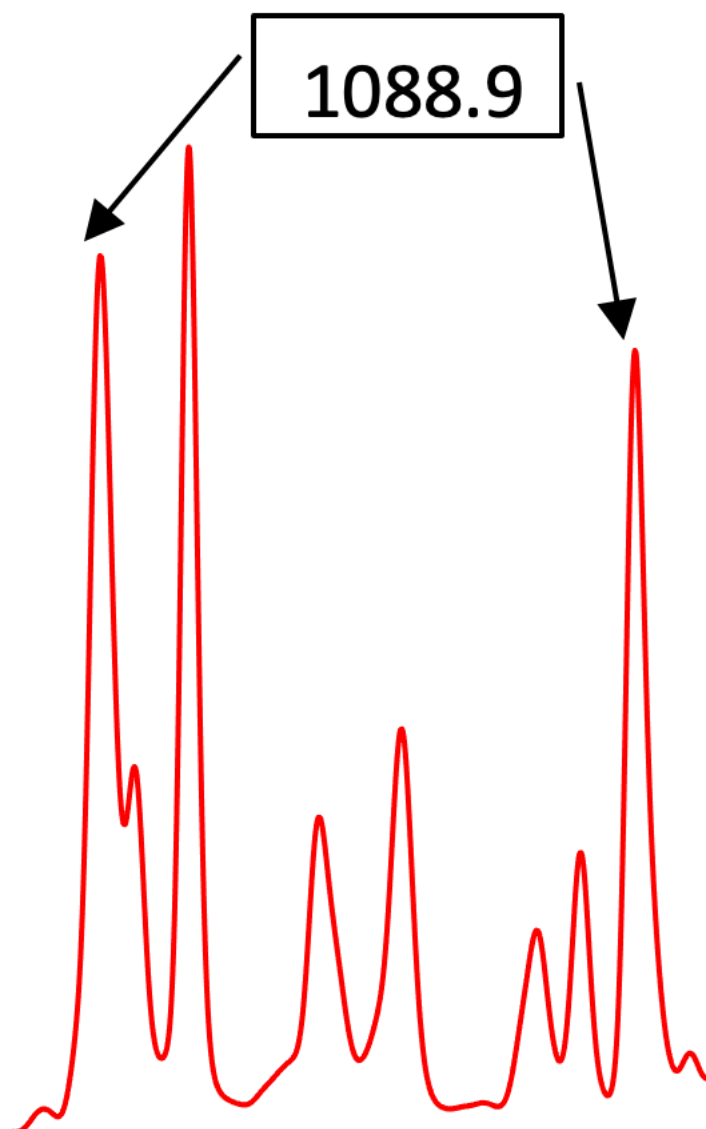


Figure 4.16 Chromatogram showing the peaks of EMB17. (A) Chromatogram comparing the peaks of EMB17 in relation to other closely related samples. The black arrows are pointing to the peaks unique to EMB17 at the 29th and 34th minutes. (B) Expanded view of EMB17 chromatogram from 28th to 35th minutes showing details of the peaks. Injection volumes of 10 μ L were made. The column was an ACE 5 C18 Column, 250 X 4.6 MM and the mobile phases used were 10% CH₃OH, 90% H₂O with 0.1% CH₂O₂ and 90% CH₃OH, 10% H₂O with 0.1% CH₂O₂. The flow rate was 0.4 mL/min

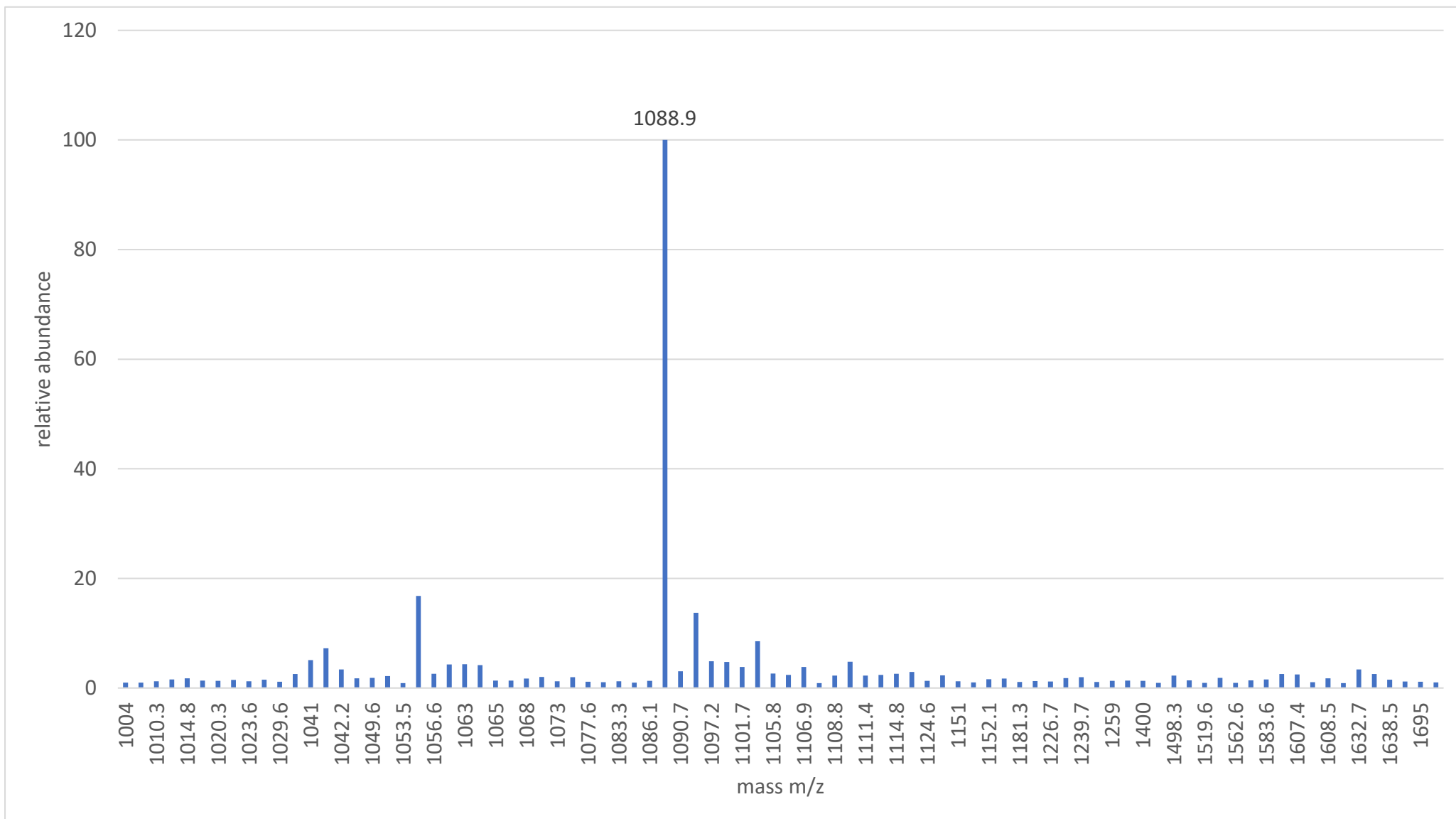


Figure 4.17 The mass spectrum of the molecular ion of EMB17 from the peak at 29th minute showing the m/z 1088.9.

4.8.3 Preparative HPLC-MS analysis

The activity of the purified EMB17 was investigated against standard indicator bacteria, *E. coli* BL21[pUC19]. The partially purified EMB17 fractions from the solid phase extraction were collected from the preparative HPLC (Agilent 1260 Infinity) and examined for activity. The sensitivity of the m/z 1088.9 picked up on LCMS was enhanced by narrowing the mass range on the preparative HPLC to 900 – 1200 m/z. This was achieved by splitting the flow from the preparative HPLC to the MSD so that the target mass could be detected in the fraction collector.

Five μL of the partially purified EMB17 sample was injected on preparative HPLC twice. The first run was to determine the retention time of peaks of interest since the sample is being run on a different system (Fig. 4.18) (see section 2.2.19). The second run was set up to mimic the first run with an additional feature of a time-based collection for the peaks of interest. The new peaks detected on the preparative HPLC were checked to examine the masses of particles that were present. Masses in the range of 1088.9 m/z were present in all the peaks (see Fig. 4.19). This could be an active component in EMB17 because it was peculiar to EMB17 as it was not found in MccB17. The collected compounds were tested for activity using bioassay experiment (Table 4.7 and Fig 4.21). All the collected fractions showed some level of activity against *E. coli* BL21[pUC19], but fraction D which is part of the 4th peak, showed the highest zone of inhibition. The size of the zone of inhibition could be related to the activity of the compound from fraction D, or it could be a concentration effect (that is, all the extracts have the same type of compound, but it is more concentrated in fraction D).

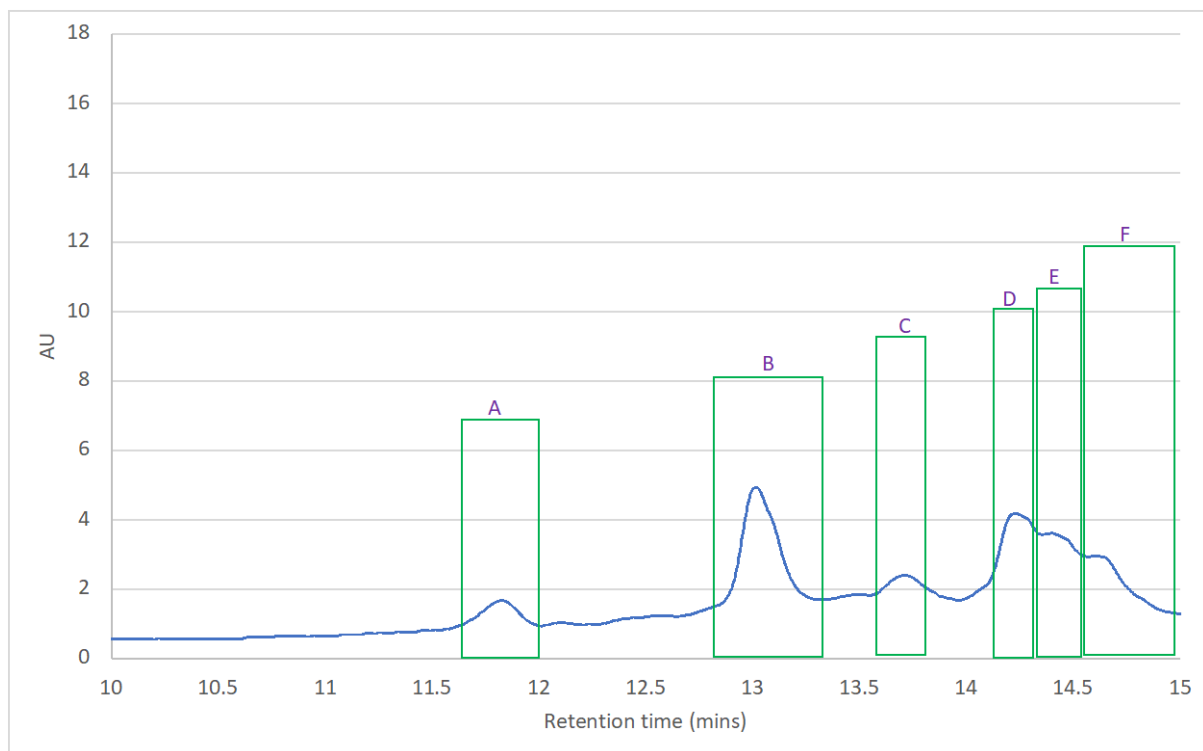
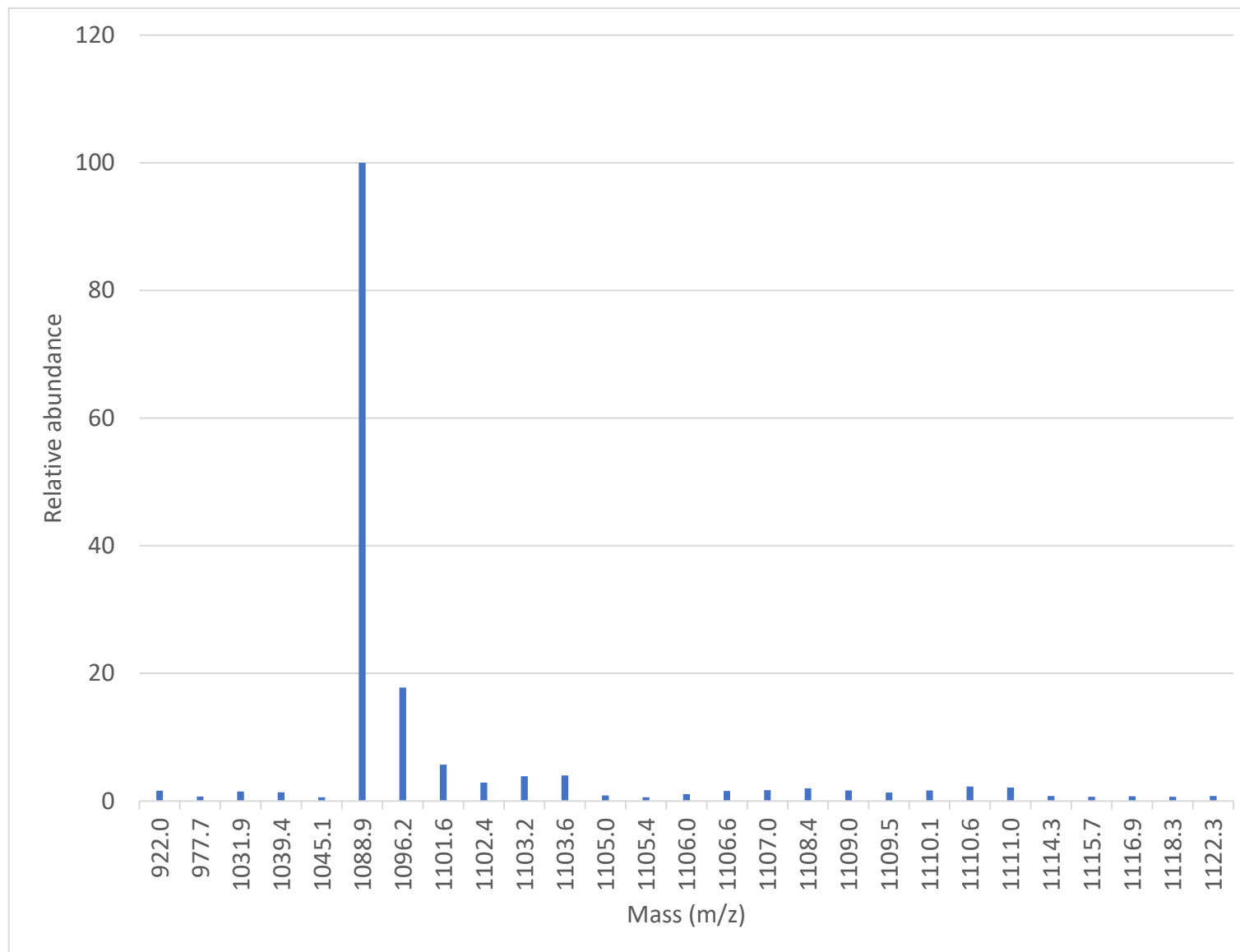
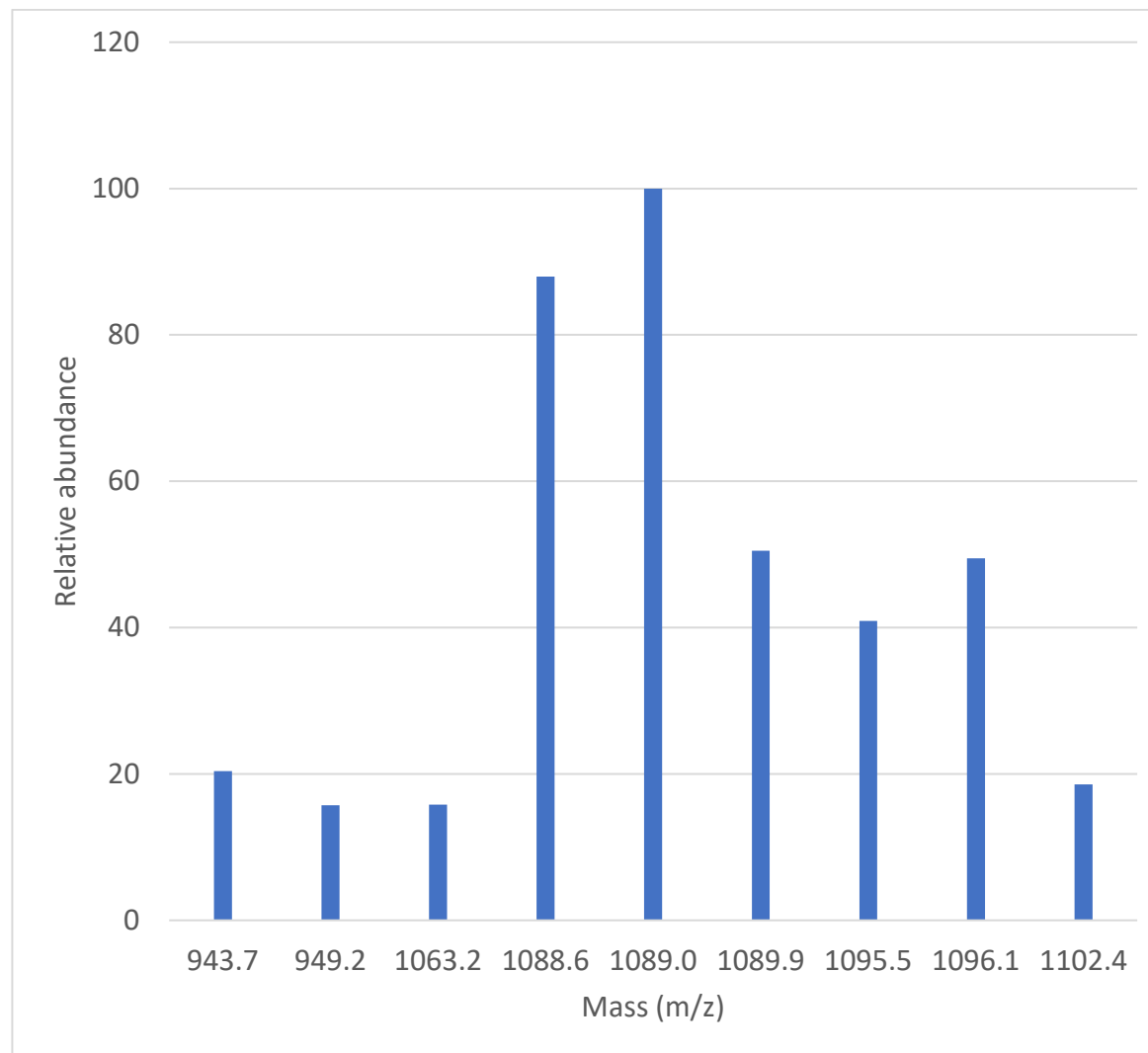


Figure 4.18 Chromatogram showing time-based collection of EMB17 sample. The collection time was split between 4 peaks. (A) The collection time is between 11.70 – 12.00 minutes. (B) Collection time was between 12.85 – 13.25 minutes, because of the width of the peak, the extract was collected into 2 collection tubes B1 and B2. (C) The collection time is between 13.60 – 13.85 minutes. (D, E and F) the 4th peak was collected between 14.09 – 15.00 minutes and split into 3 sections.

(A)



(B)



(C)

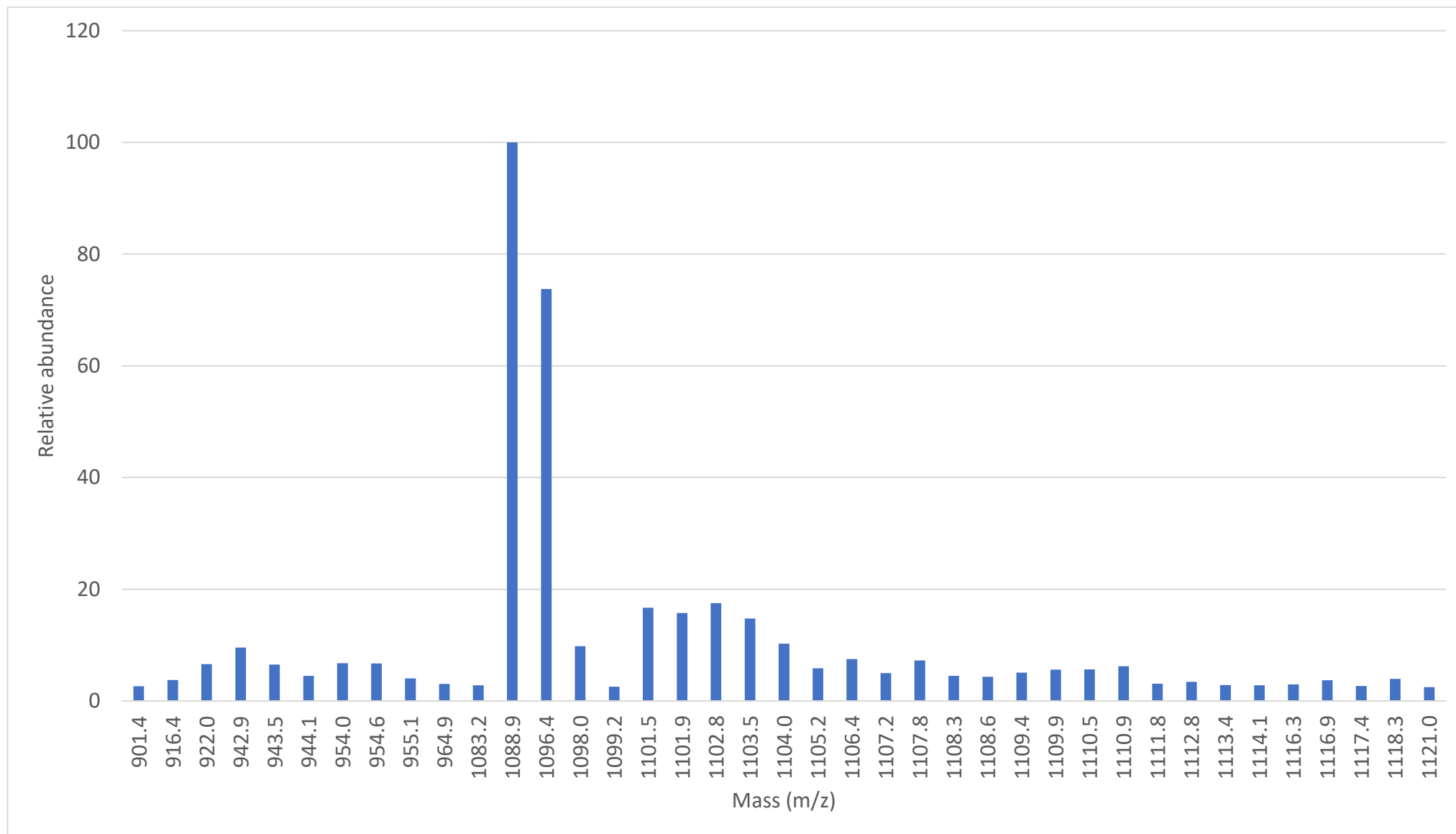


Figure 4.19 The mass spectrum of the selected peaks from the preparative HPLC analysis of EMB17. (A) Mass spectrum at 12.89 minute from the 2nd peak (B) Mass spectrum at 13.67 minute from the 3rd peak (C) Mass spectrum at 14.17 minute from the 4th peak

Table 4.7 Bioassay for HPLC time-based collection

Fraction	Zone of inhibition
A	5
B1*	5
B2*	5
C	7
D	12
E	7
F	5

3µl drops of the purified microcins were spotted onto M63 agar and overlaid with 100µl overnight *E. coli* BL21 culture in 5ml soft agar. Diameters of growth inhibition zones were determined after overnight growth at room temperature. Mean values were obtained from triplicate experiments.

* B was split into 2 (B1 and B2) because the volume was more than the collection tube because of the width of the 2nd peak. They were treated independently during the bioassay, but they have the same result.

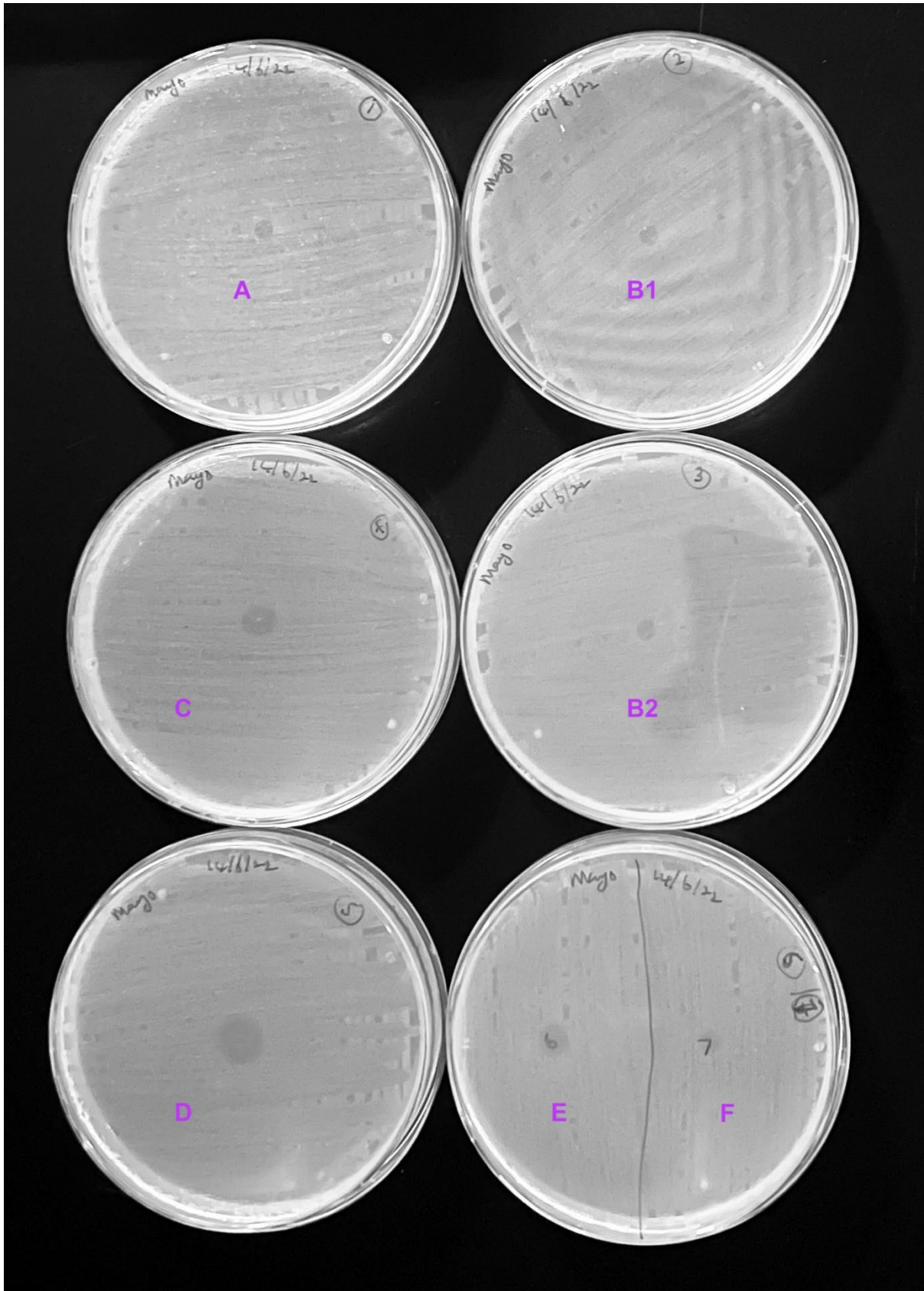


Figure 4.20 Image of Time-based collection bioassay. (A) Collection time between 11.70 – 12.00 minutes. (B) Collection time between 12.85 – 13.25 minutes, because of the width of

the peak, the extracts were collected into 2 collection tubes B1 and B2. (C) Collection time between 13.60 – 13.85 minutes. (D, E and F) the 4th peak was collected between 14.09 – 15.00 minutes and split into 3 sections.

Chapter five

Genetic analysis of promicrocin – synthase interactions of *E. persicina* and *E. coli* microcin gene clusters

5.0 Overview

This chapter focuses on studying how components of the microcin biosynthetic pathway, the promicrocin substrate (*McbA*) and synthase complex (*McbBCD*), can affect microcin production. Recombinant DNA technology was used as a tool to separate or delete these components from the *emo* and *mcb* operons. These components were cloned into a vector under the control of an inducible promoter to create plasmid constructs for expression. Using a mix and match selection approach, the plasmid constructs were paired in a *cis* and *trans* complementary way to form a complete microcin operon comprising plasmid constructs from *Erwinia* and *E. coli* microcin operons. This inspiration was from the experiments of Shkundina *et al.* (2014), where they created a double-plasmid *MccB17* production system that comprised of a pUC-based plasmid with *mcbA*, and another compatible plasmid pACYC-based with *mcbABCDEFG*. Bioassay experiments were carried out to examine the effects of this complementary pairing on microcin production.

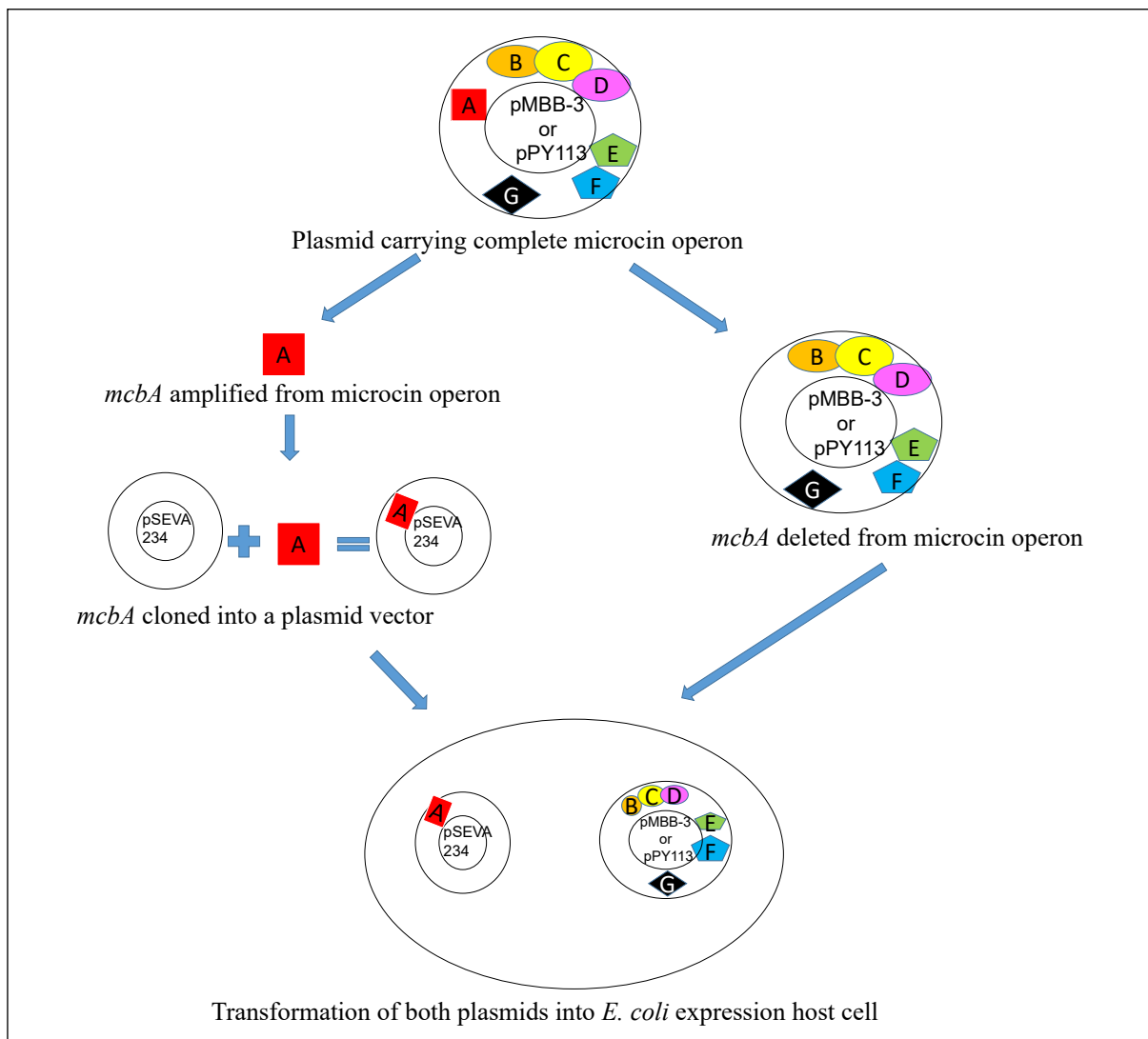
5.1 Strategy for genetic analysis of modified microcin gene clusters

Plasmids containing altered microcin gene clusters were cloned to create a collection of microcin biosynthesis genes from both *E. persicina* and *E. coli* microcin operon, which can be paired to another compatible plasmid containing the complementary gene(s) to form a complete microcin operon (Fig. 5.1). The idea was to delete a gene from the microcin operon as well as having the same gene cloned into a plasmid vector, thus having two types of plasmid constructs arising from one microcin operon. For instance, *mcbA* (*E. coli*) and *mcbBCDEFG* (*E. coli*) will exist in different plasmids.

Plans were to carry out this study on the promicrocin gene, *mcbA*, and the microcin synthetase genes, *mcbB*, *mcbC* and *mcbD*. However, due to restrictions on laboratory access during the global coronavirus pandemic, two genes were examined, *mcbA* and *mcbB*. These two genes were individually deleted from the microcin operons of *E. persicina* and *E. coli*,

and in a different procedure, these genes were cloned into a plasmid vector (Fig. 5.1(A)). The resulting plasmid constructs were mixed and paired complementarily to make up a complete microcin operon (cognate and non-cognate). The paired plasmid constructs were transformed into *E. coli* DH5 α cells for expression (Fig. 5.1(B)). Subsequent sections in this chapter will focus on how these genes were deleted or cloned into a plasmid vector, as well as their expression and bioassay to confirm their activity.

(A)



(B)

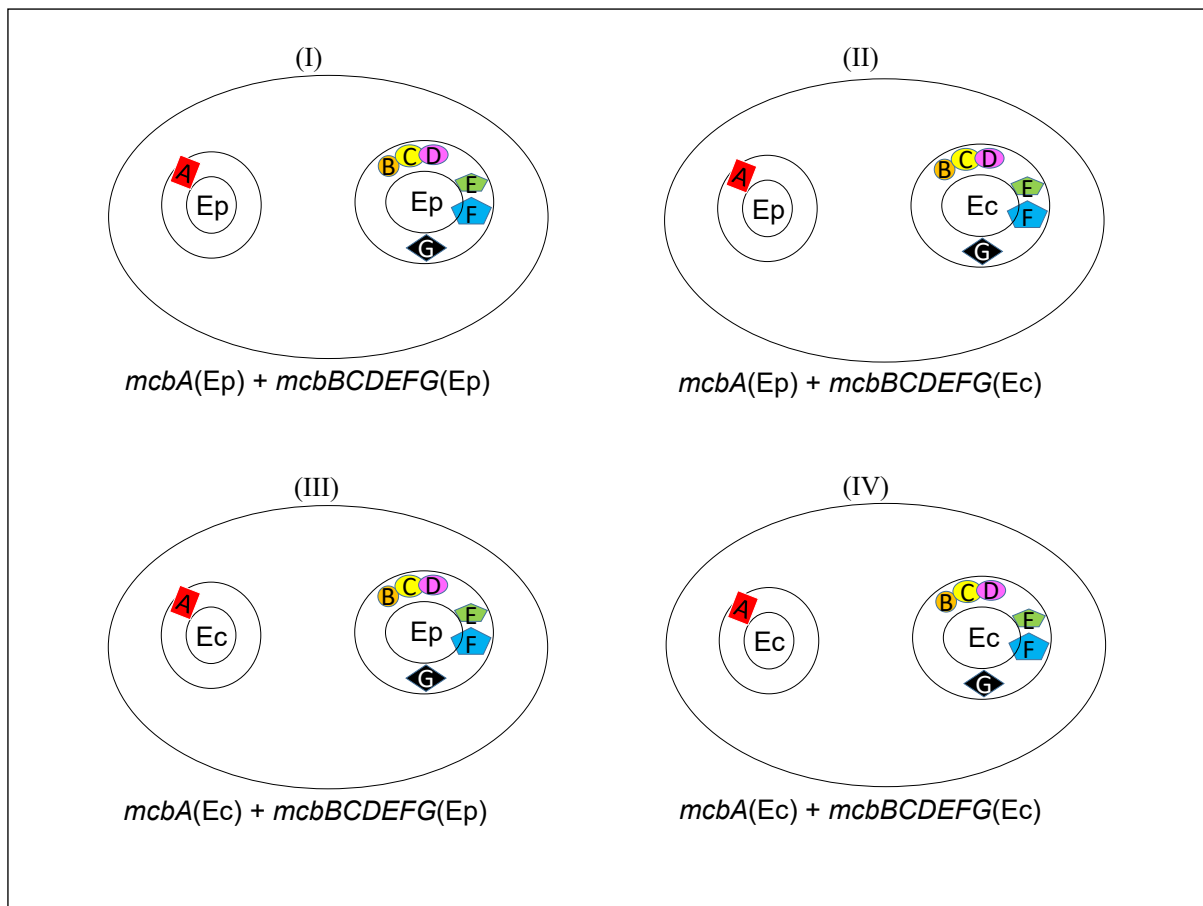


Figure 5.1 Schematic illustration of how plasmid constructs were made and used for the mix and match pairing using *mcbA* as an example. (A) In two different experiments, *mcbA* was cloned into a vector and in the other experiment, it was deleted from the operon. This generated 2 types of constructs each from *Erwinia* and *E. coli*. (B) Choice of plasmid constructs pairing to make up complete microcin operon. Ep stands for *Erwinia persicina* and Ec stands for *E. coli*.

5.1.1 Deletion of *mcbA* from *E. persicina* and *E. coli* microcin operon

A method based on inverse PCR mutagenesis was used, where primers were designed to delete unwanted genes while the remaining genes in the *emo* and *mcb* operons were amplified (Dorrell *et al.*, 1996). The forward primers, del-*mcbA*-Ep-F and del-*mcbA*-Ec-F for *Erwinia* and *E. coli*, respectively, were designed to include the stop codon of *mcbA* and sequences downstream of *mcbA* (Table 2.1). The reverse primers, del-*mcbA*-Ep-R and del-*mcbA*-Ec-R for *Erwinia* and *E. coli*, respectively, were designed by reverse complement of sequences to include the beginning of *mcbA* as well as the region upstream of *mcbA* (Fig. 5.2; Table 2.1). This in-frame deletion left a very truncated residue of *mcbA* sequence in the operon, but it was non-functional and did not result into a polar effect; this was later confirmed by a bioassay experiment where there was no inhibition zone formed by producer cells that carried the operon with deleted *mcbA*.


A 0.6 w/v % agarose gel was set up to check the size of the PCR products (Fig. 5.3). It confirmed the expected band size of the amplicons, about 8.9 kb for pMBB-3Δ*mcbA* and about 10.4 kb for pPY113Δ*mcbA* (Fig. 5.6). The PCR products (pMBB-3Δ*mcbA* and pPY113Δ*mcbA*) were purified from gel slices using DNA gel extraction kit (section 2.2.9). They were later phosphorylated because fragments created by PCR amplification for cloning purposes can lack the necessary 5' phosphate terminal. The phosphorylated PCR products were self-ligated with the aid of T4 DNA ligase. The resultant plasmids were transformed into *E. coli* DH5α. These cells were grown in LB media containing AMP to select for the plasmid constructs which were purified afterwards using a plasmid mini prep kit.

(A)

```
cttATGGAATTAAACACGACTGAATTTGGTGTAGTTCTGTCCGTAGATGCCCTCAAGCTTT  
CCCGACAGTCCCCGTTAGGAATTGGTATTGGTGGGGGTGGCGCGGTGGAGGCGGTGGCGG  
CGGCGGAAGCAGTGGCGGCCAGGGCGGCGGCTGTGGCGGTTGCAGCAACGGTTGCAGTGGA  
GAAATGGTGGCAGCGGTGGGAGTGGTTCGCATATCTGAcacgctgaactgaccgatc
```

(B)

Point of deletion



```
cttATGGAATTAAACACGACTGAATTT | TGAcacgctgaactgaccgatc
```

Figure 5.2 The deletion pattern of *mcbA* and the residue that remained after deletion. (A) The full length of *Erwinia mcbA* showing the primers used for the deletion. Highlighted in yellow are the primers used for the deletion of *mcbA*, while the deleted sequences from *mcbA* are highlighted in grey. In lower cases are sequences that flank *mcbA*, while sequences in upper cases are those that make up the CDS, *mcbA*. (B) The remnant of *mcbA* after deletion is shown in the green box with the black arrow showing the point where the deletion took place.

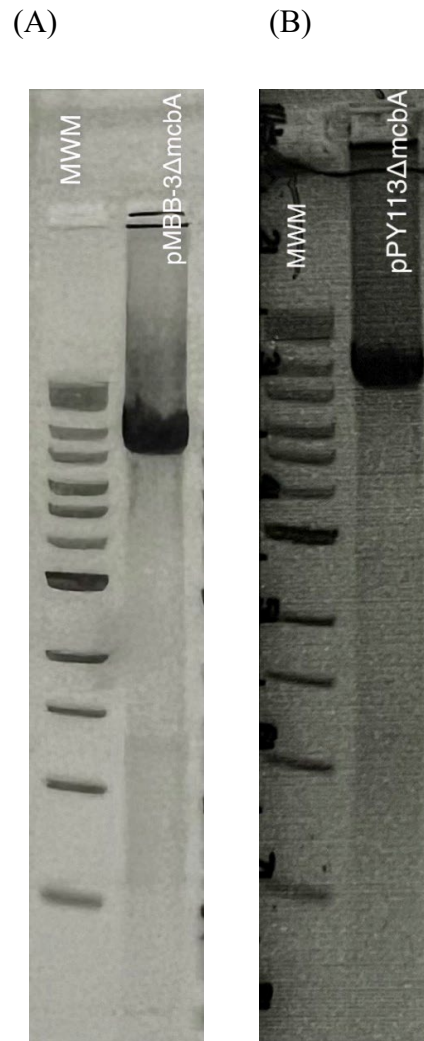


Figure 5.3 Agarose gel electrophoresis of PCR amplicons after the deletion of *mcbA* from the microcin operons. (A) Agarose gel of pMBB-3Δ*mcbA* showing the expected band at 8.9 kb (B) Agarose gel of pPY113Δ*mcbA* showing the expected band at 10.4 kb. Gel concentration is 0.6% agarose gel in TAE buffer. MWM, molecular weight marker (1 kb Extend DNA Ladder, NEB).

A virtual digest was run on Benchling to help with the choice of restriction enzymes and to give an idea of the expected bands (Fig. 5.4B and Fig. 5.5B). From the virtual digest, it was clear that HindIII will digest pMBB-3 twice while it digests pMBB-3Δ*mcbA* once. Also, XmnI will digest pPY113 four times while it digests pPY113Δ*mcbA* three times. With the band size from the respective restriction enzyme digest, we had a way to confirm the deletion of *mcbA* from both microcin operons. pMBB-3Δ*mcbA* and pMBB-3 were digested with HindIII, and as expected, pMBB-3Δ*mcbA* was cut once into a linear DNA of 8888bp while

pMBB-3 was cut twice into two DNA fragments (a 481 bp and an 8598 bp fragments). The 481 bp fragment was not visible on the agarose gel; this could be because the relative mass of DNA in this fragment is low, and hence fluorescence is very low (Fig. 5.4A). The 8888 bp pMBB-3 Δ *mcbA* was slightly higher than the 8598 bp pMBB-3 DNA on both the agarose gel image and the virtual digest image. Also, pPY113 Δ *mcbA* and pPY113 were digested with XmnI, pPY113 Δ *mcbA* was cut into 3 fragments (1932 bp, 2581 bp and 5896 bp) while pPY113 was cut into 4 fragments (509 bp, 1932 bp, 2255 bp and 5896 bp). The restriction digest and the agarose gel image confirmed the deletion of *mcbA* from both plasmids (pMBB-3 and pPY113).

For further confirmation that the gene *mcbA* has been deleted from the pMBB-3 plasmid, the purified plasmid construct was sent to DBS Genomics (Durham) for Sanger sequencing using the primer B-Ep-mcb-US-F (Table 2.1). The result of the sequencing showed the deletion of *mcbA* sequences outside of the expected amplicon (Fig. 5.2). The purified plasmid constructs were also sent to MicrobesNG (Birmingham, UK) for next-generation short read technology sequencing (Illumina) to confirm the deletion as well as to check the plasmid constructs for mutations. The variant calling file data showed that no mutations were detected in pMBB-3 Δ *mcbA* or pPY113 Δ *mcbA*, in addition to verifying the deletion of *mcbA*. In addition, the plasmid constructs were sent to plasmidsaurus (U.S.A.) for long Oxford Nanopore reads; the sequencing results also confirmed the deletion of *mcbA* from both pMBB-3 and pPY113 (Fig. 5.7).

(A)



(B)

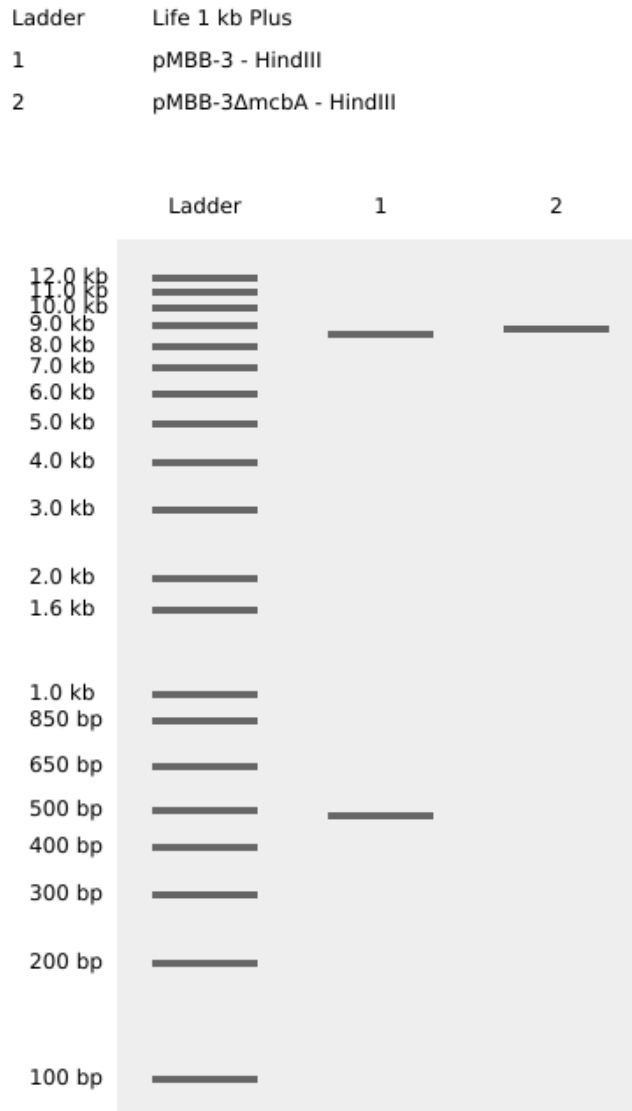


Figure 5.4 Agarose gel electrophoresis to confirm deletion of *mcbA* by restriction enzyme digest of pMBB-3Δ*mcbA* and pMBB-3. (A) Digestion of pMBB-3Δ*mcbA* and pMBB-3 with HindIII (B) Virtual digest of pMBB-3Δ*mcbA* and pMBB-3. Gel concentration is 0.8% agarose gel in TAE buffer. MWM, molecular weight marker (1 kb Extend DNA Ladder, NEB).

(A)

(B)

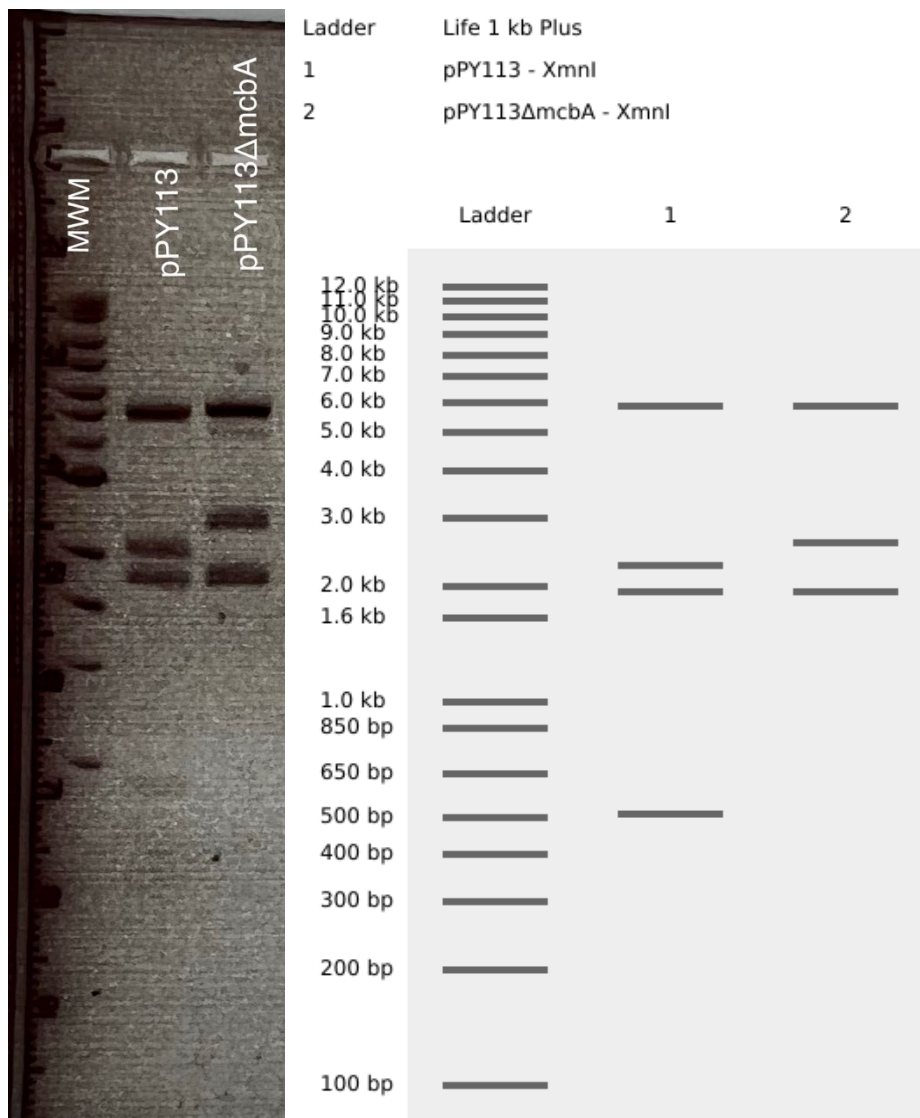


Figure 5.5 Agarose gel electrophoresis to confirm deletion of *mcbA* by restriction enzyme digest of pPY113Δ*mcbA* and pPY113. (A) Digestion of pPY113Δ*mcbA* and pPY113 with XmnI (B) Virtual digest of pPY113Δ*mcbA* and pPY113. The 509 bp fragment was very faint on the agarose gel (A) compared to the virtual digest (B), which is due to its low mass. Gel concentration is 0.6% agarose gel in TAE buffer. MWM, molecular weight marker (1 kb Extend DNA Ladder, NEB).

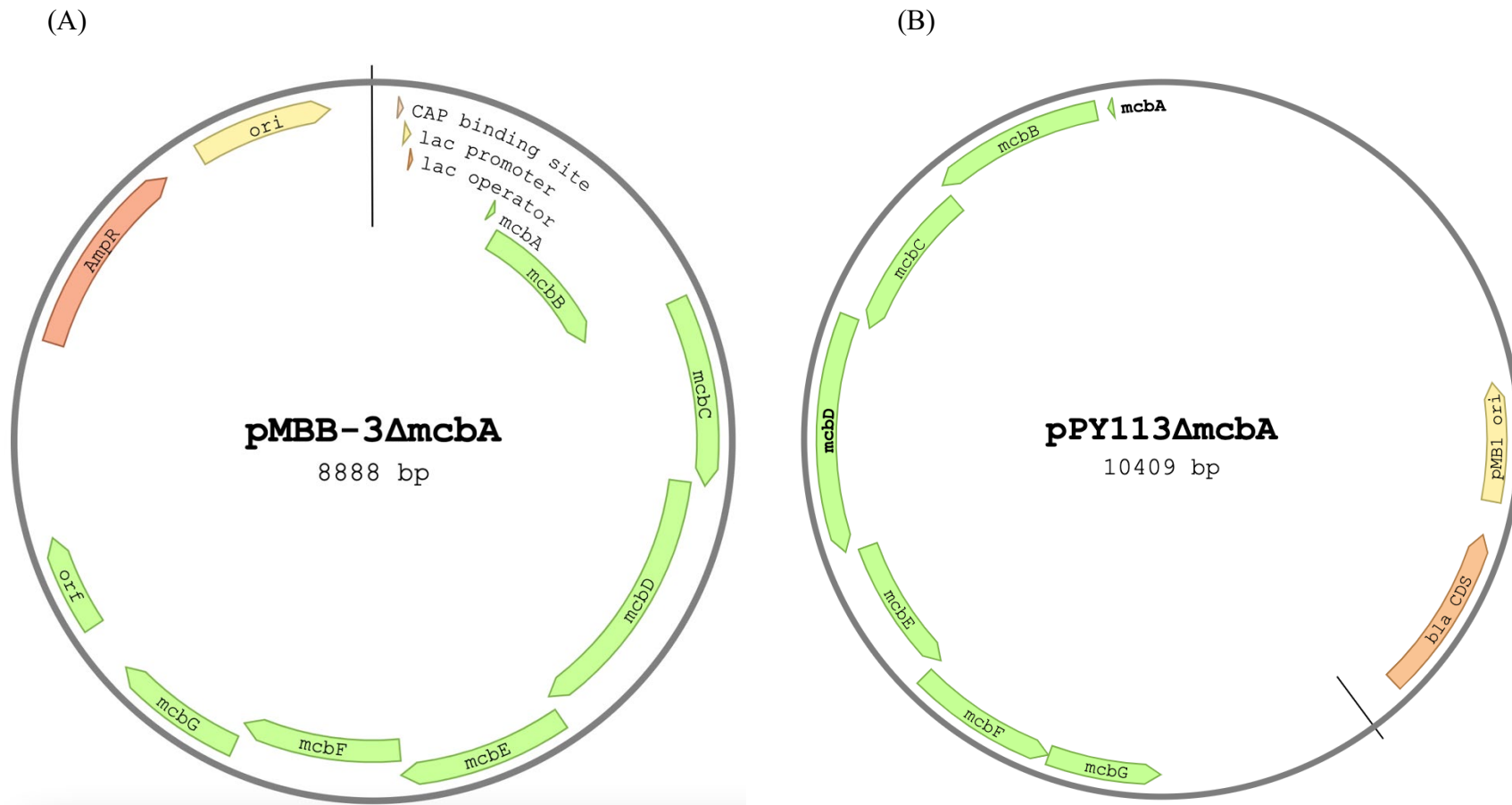
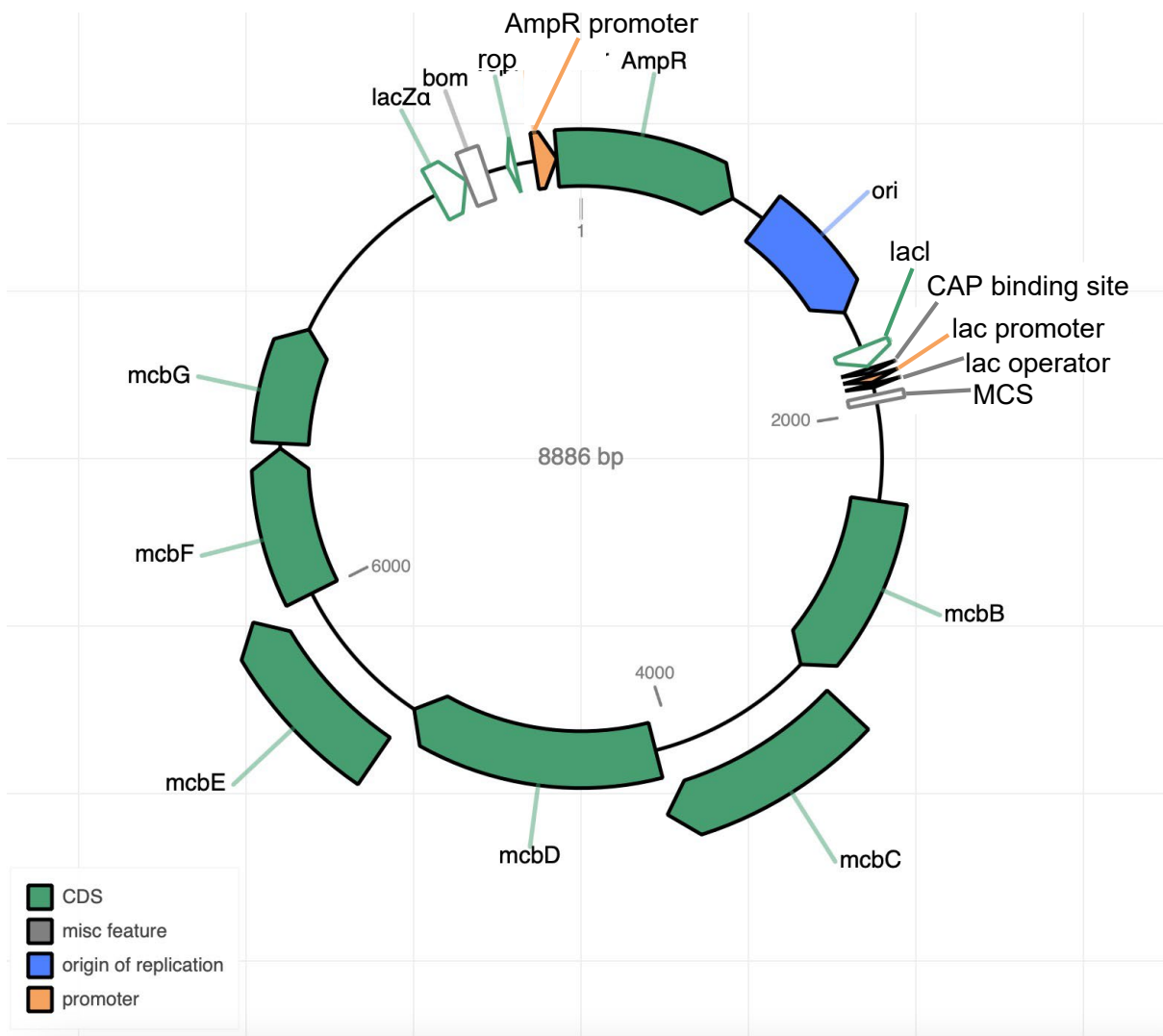


Figure 5.6 Plasmid maps of the cloned microcin operons showing deletion of *mcbA* and incomplete microcin operon. (A) pMBB-3Δ*mcbA*, *Erwinia* microcin operon without *mcbA* (B) pPY113Δ*mcbA*, *E. coli* microcin operon without *mcbA*.

(A)



(B)

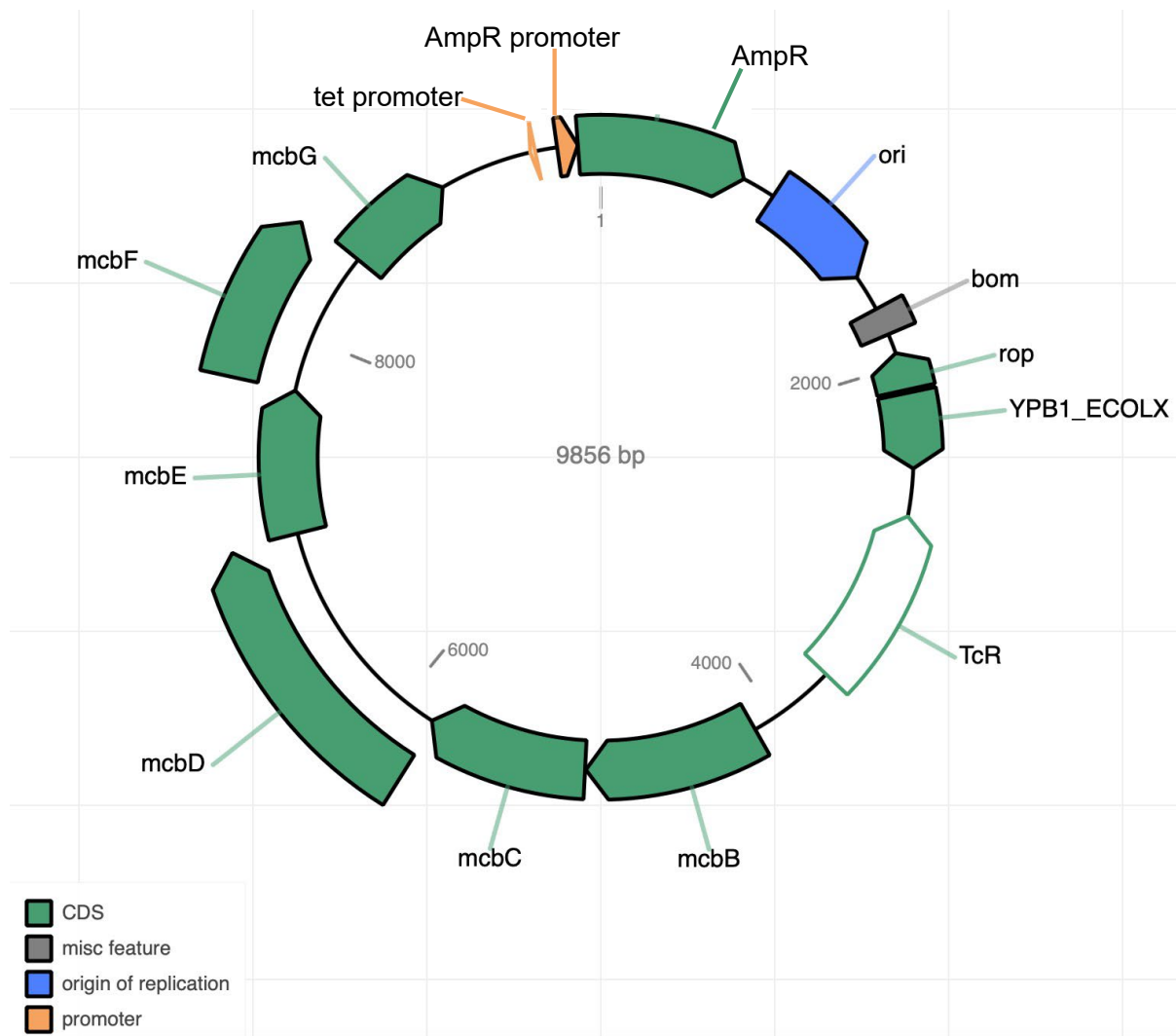


Figure 5.7 Plasmid maps of cloned microcin operon deletion mutants of *mcbA*. (A) pMBB-3Δ*mcbA*, *Erwinia* microcin operon without *mcbA* (B) pPY113Δ*mcbA*, *E. coli* microcin operon without *mcbA*. These were derived from whole-plasmid sequencing data by Plasmidsaurus using the pLannotate tool (McGuffie and Barrick, 2021).

5.1.2 Cloning of *mcbA* into SEVA234 plasmid vector

This section focuses on amplifying a *mcbA* gene cassette from the microcin operon and cloning it into an inducible expression vector. PCR primers were designed to amplify *mcbA* from the *emo* and *mcb* operons to achieve this. The forward primers, EcoR-Ep-mcbA-F and EcoR-Ec-mcbA-F for *Erwinia* and *E. coli*, respectively, were designed to have an EcoRI recognition site close to the beginning of the 5' end (Table 2.1). In addition, the forward primers were fitted with a Ribosome Binding Site (RBS) after the EcoRI recognition site. The reverse primers, Sal-Ep-mcbA-R and Sal-Ec-mcbA-R for *Erwinia* and *E. coli*, respectively, were designed to have a SalI recognition site included close to the beginning of the 5' end (Table 2.1). After the PCR amplification of *mcbA*, the amplified product was run on agarose gel to confirm the size of the amplicons (Fig. 5.8). The agarose gel image showed the expected band size of about 200 bp for both the *Erwinia* and *E. coli mcbA*. The PCR products were purified from gel using a DNA gel extraction kit, and the purified *mcbA* gene cassettes were digested with EcoRI and SalI in preparation for ligation with a plasmid vector.

Plasmid pSEVA234 was chosen as the plasmid vector to clone *mcbA* into because it has an IPTG inducible promoter, *P_{trc}*, that can be used to drive the expression of the cloned genes. It was also chosen because it has a different antibiotic gene marker, *neo*, which encodes neomycin phosphotransferase (which allows kanamycin to be used as the selective agent) as against *bla*, which encodes beta-lactamase in pMBB-3Δ*mcbA* and pPY113Δ*mcbA*. In addition, the origin of replication of pSEVA234, pBBR1, is different from (and hence compatible with) the one found in both pMBB-3Δ*mcbA* and pPY113Δ*mcbA*, pMB1. As a result of the differences in the origin of replication, pSEVA234 is compatible with either pMBB-3Δ*mcbA* or pPY113Δ*mcbA* and can coexist with them in an expression host cell. These considerations are essential in selecting the cells that will carry the two types of plasmid constructs later in this project. The pSEVA234 was extracted from *E. coli* CC118 by plasmid DNA miniprep and was digested with EcoRI and SalI in preparation for *mcbA* insertion. The digested pSEVA234 DNA was dephosphorylated by adding alkaline phosphatase to prevent recirculation of pSEVA234 during ligation. The digested *mcbA* was ligated with the digested dephosphorylated pSEVA234 using Instant sticky-end ligase. The result of the ligation was pSEVA234*mcbA*EP and pSEVA234*mcbA*EC for the *Erwinia* and *E. coli* constructs respectively (Fig. 5.9). These were transformed into *E. coli* DH5α cells, after which the cloned plasmids were extracted for further analysis.

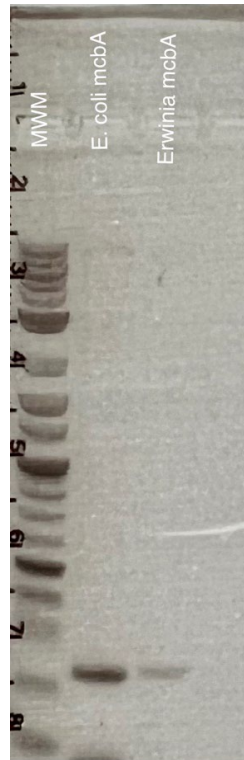
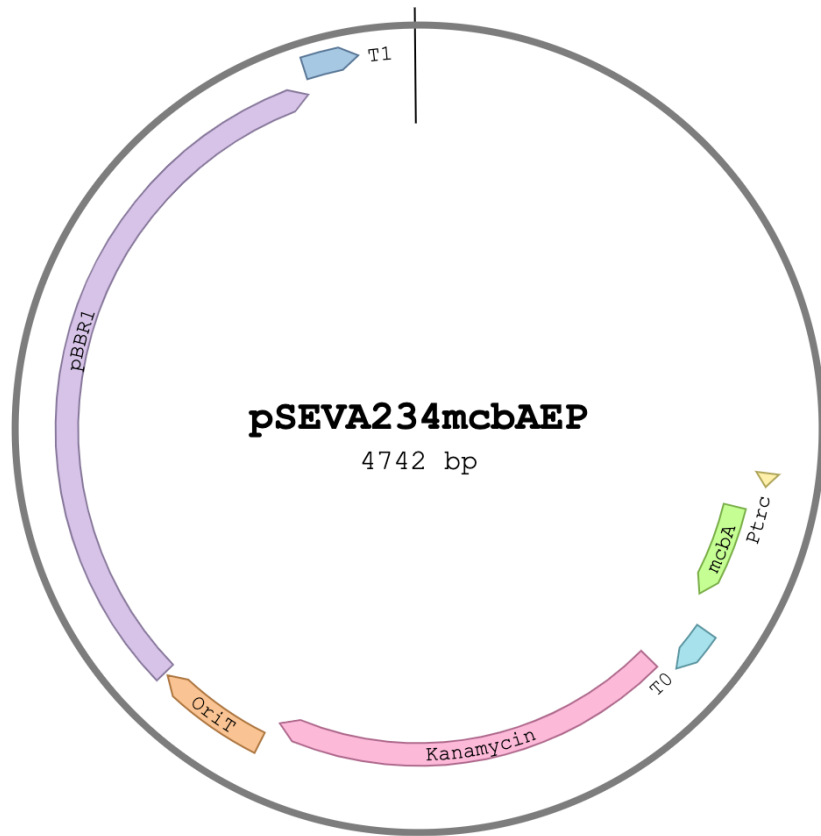


Figure 5.8 Agarose gel electrophoresis of PCR amplicons, EcoRI – SalI *mcbA* cassette (~200bp) from *E. coli* and *Erwinia* microcin operons. Gel concentration is 1.2 % agarose gel in TAE buffer. MWM, molecular weight marker (2 log DNA ladder, NEB).

(A)



(B)

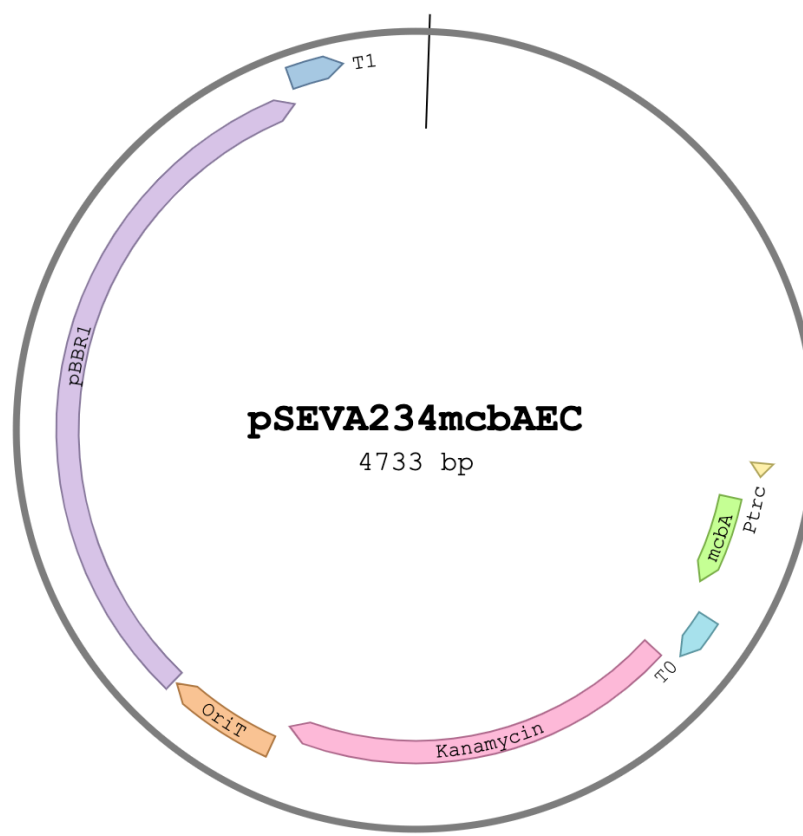


Figure 5.9 Plasmid maps of pSEVA234 derivatives with *mcbA* insert (A) pSEVA234*mcbA*EP (B) pSEVA234*mcbA*EC. EP and EC represent *Erwinia persicina* and *E. coli* respectively.

To confirm that *mcbA* was correctly cloned into the plasmid vector, pSEVA234, the cloned plasmids (pSEVA234*mcbA*EP and pSEVA234*mcbA*EC) were digested with EcoRI and Sall and run on a 1.2% agarose gel. This was compared with pSEVA234, which was also digested with EcoRI and Sall. No noticeable difference was seen in the bands shown on the agarose gel image; this could be due mainly to the small size and low concentration of the cloned gene, *mcbA*, after digestion (Fig. 5.10). When other restriction enzymes were used, the result was the same. Another method of confirmation that *mcbA* from the *emo* and *mcb* operon were successfully cloned into the plasmid vector, pSEVA234, was to sequence the cloned plasmids. Using both the Sanger method (with primers ptrc-for and M13-F24) and a full-length plasmid sequencing service, we were able to confirm that *mcbA* from both microcin operons were successfully cloned into pSEVA234. Fig. 5.11 shows a map of pSEVA234 with *mcbA* insertion.

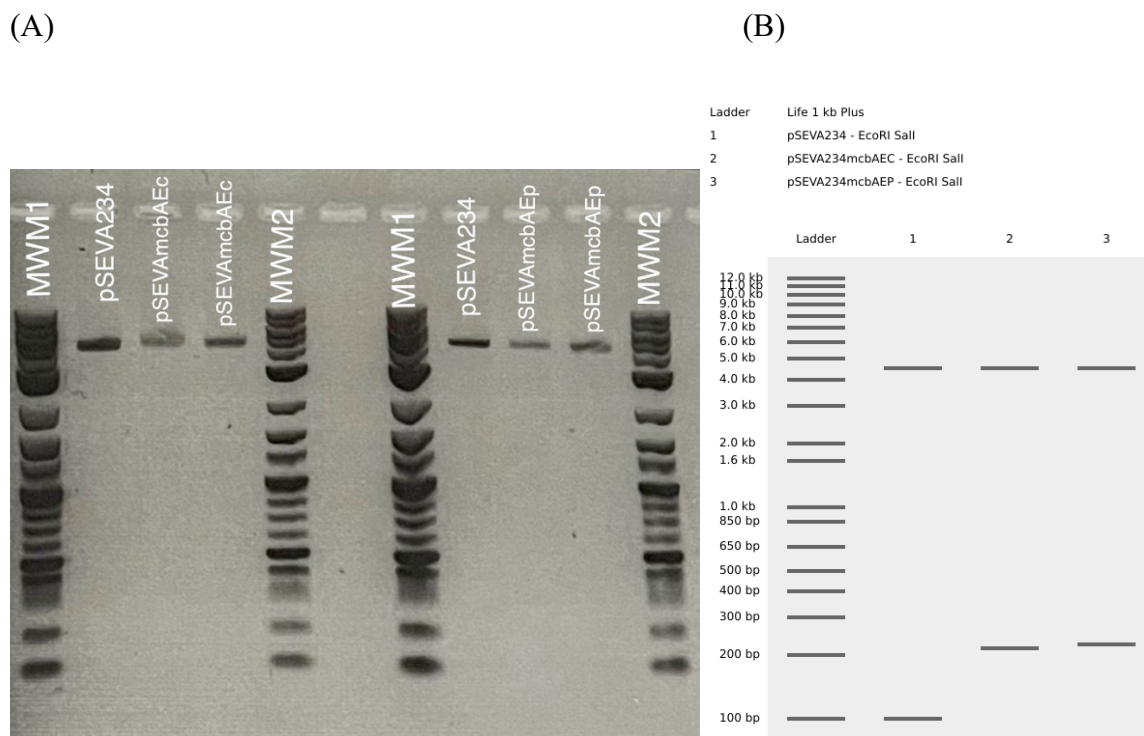
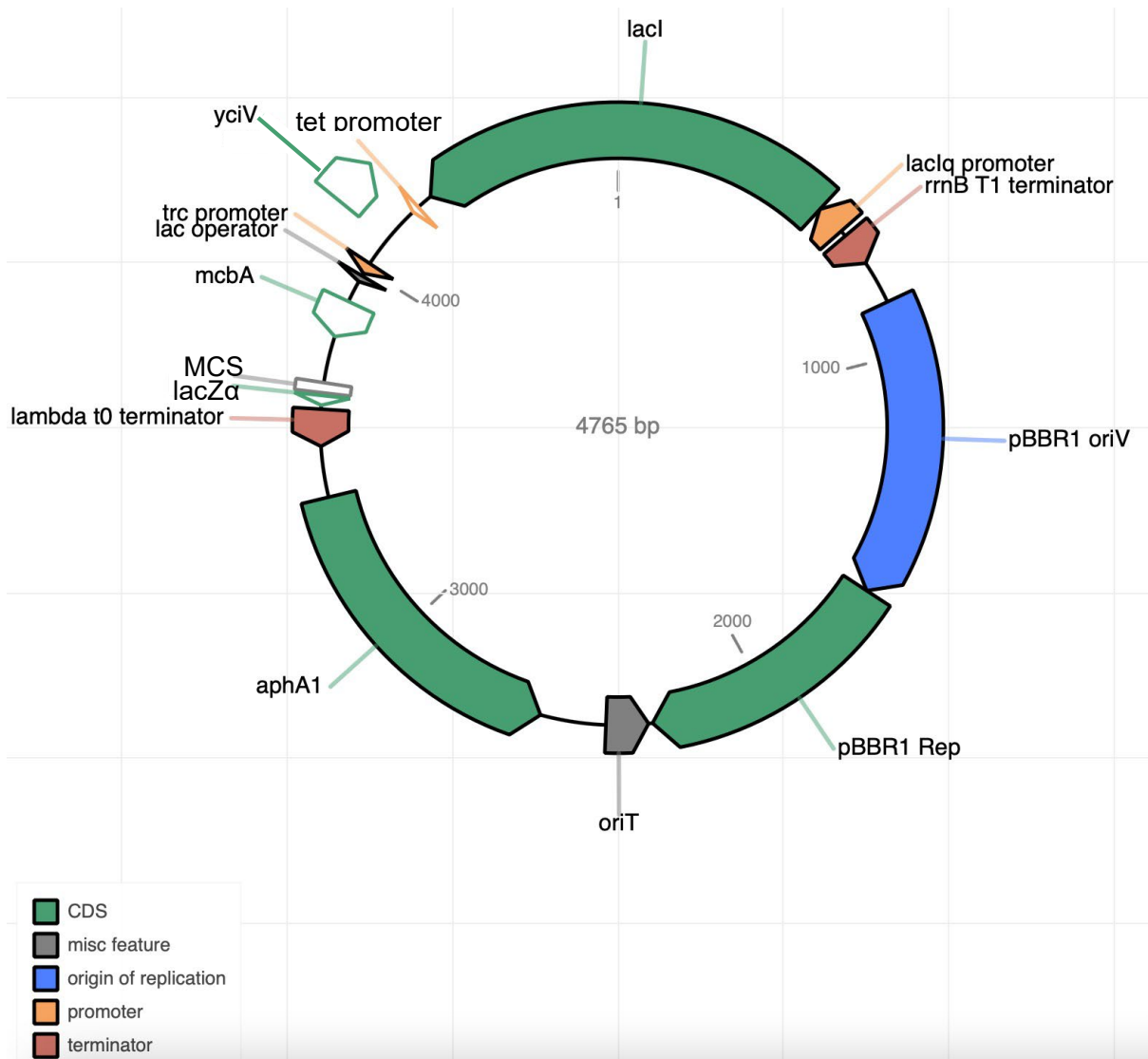


Figure 5.10 Confirmation of cloning of *mcbA* into pSEVA234. (A) Agarose gel image of the digestion of pSEVA234, pSEVA234*mcbA*EC and pSEVA234*mcbA*EP with EcoRI and Sall. There were no noticeable differences in the bands present. (B) Virtual digest of pSEVA234,

pSEVA234*mcbA*EC and pSEVA234*mcbA*EP with EcoRI and SalI. Gel concentration is 1.2 % agarose gel in TAE buffer. MWM, molecular weight marker (MWM1 - 2 log DNA ladder, NEB and MWM2 - 1 kb plus DNA ladder, NEB).

(A)



(B)

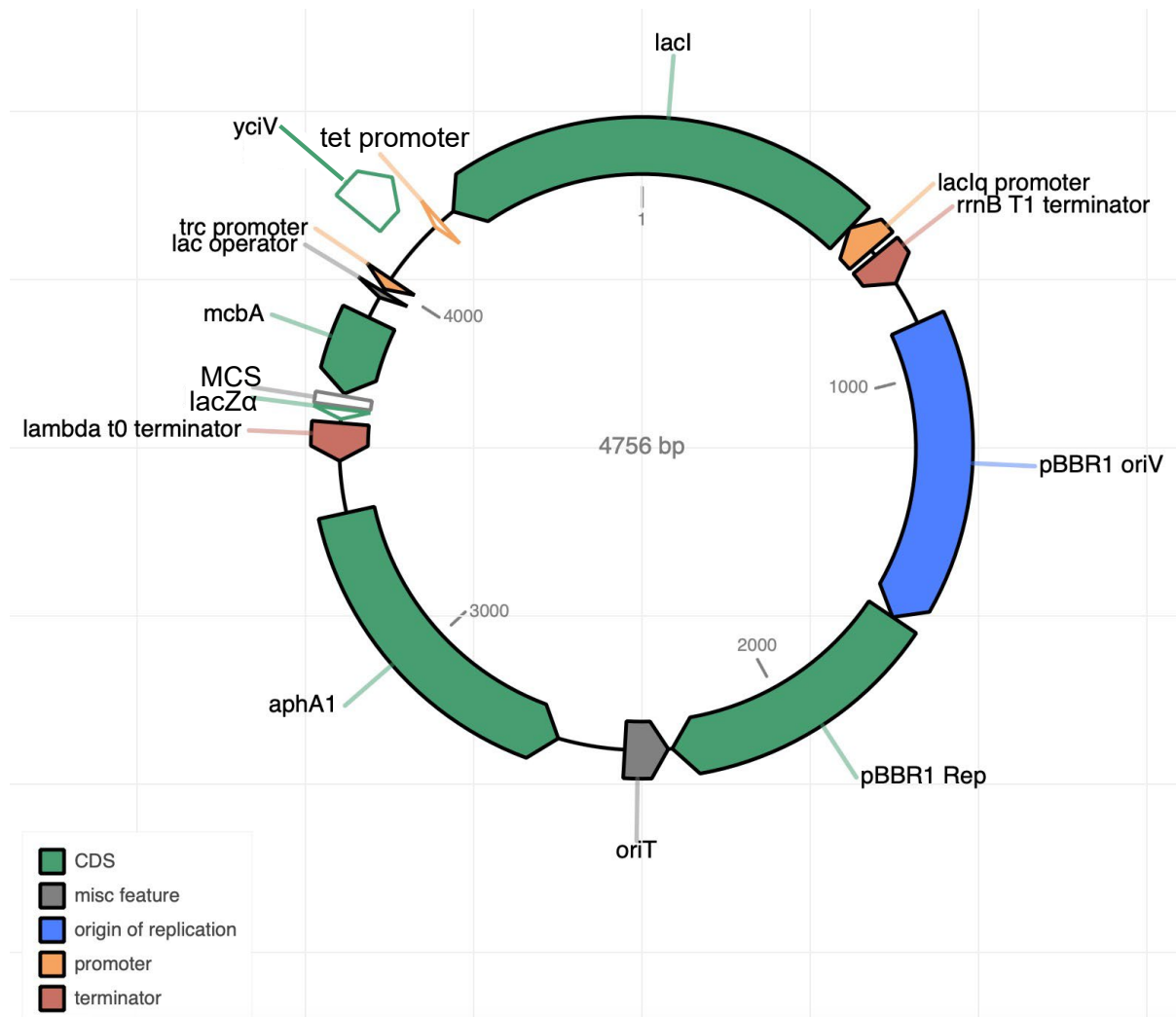


Figure 5.11 Plasmid maps of cloned *mcbA* into pSEVA234. (A) pSEVA234*mcbA*EP (B) pSEVA234*mcbA*EC. These were derived from whole-plasmid sequencing data by Plasmidsaurus using the pLannotate tool (McGuffie and Barrick, 2021).

5.1.3 Deletion of *mcbB* from *E. persicina* and *E. coli* microcin operon

Similarly, to deleting *mcbA*, the primers del-mcbB-Ep-F, del-mcbB-Ep-R, del-mcbB-Ec-F, and del-mcbB-Ec-R for deleting *mcbB* from both *emo* and *mcb* operons were designed following the same method with the only variation being that these primers were synthesised with 5' phosphate (Table 2.1). As such, the phosphorylation step in the deletion of *mcbA* was skipped. After the PCR mediated plasmid DNA deletion of *mcbB* from both microcin operons, the PCR products (pMBB-3 Δ *mcbB* and pPY113 Δ *mcbB*) were run on a 0.6% agarose gel to confirm the size of the amplicons before further analyses were carried out (Fig. 5.12). From Benchling virtual gel tool, we deduced the expected band size for pMBB-3 Δ *mcbB* and pPY113 Δ *mcbB* to be about 8.3 kb and 9.8 kb respectively (Fig. 5.13). Fig. 5.12 gives a rough estimate of the band sizes of the amplicons where the pPY113 Δ *mcbB* bands are just below the 10 kb ladder mark, and the pMBB-3 Δ *mcbB* bands were further down.

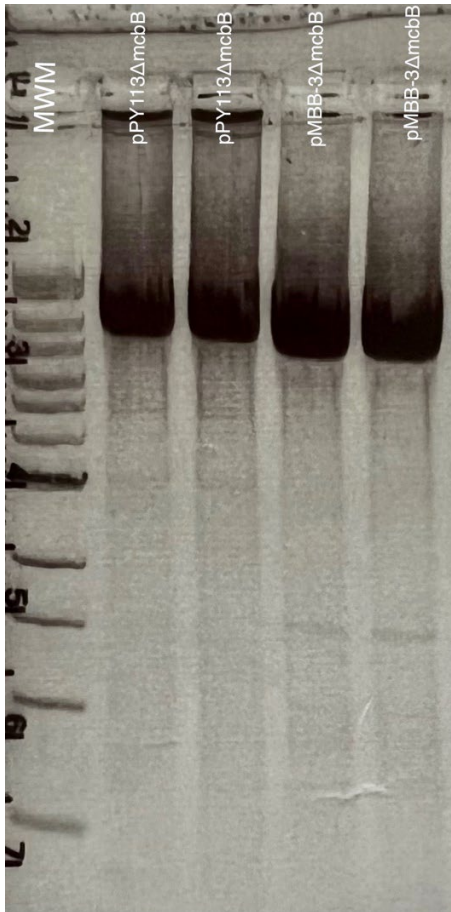


Figure 5.12 Agarose gel electrophoresis of PCR amplicons after the deletion of *mcbB* from the microcin operons. The red line represents 10 kb band on the ladder. Bands of pPY113 Δ *mcbB* were just a little below the 10 kb mark while pMBB-3 Δ *mcbA* is further down, these sizes correlate with the expected size of the amplicons. Gel concentration is 0.6 % agarose gel in TAE buffer. MWM, molecular weight marker (1 kb Extend DNA Ladder, NEB).

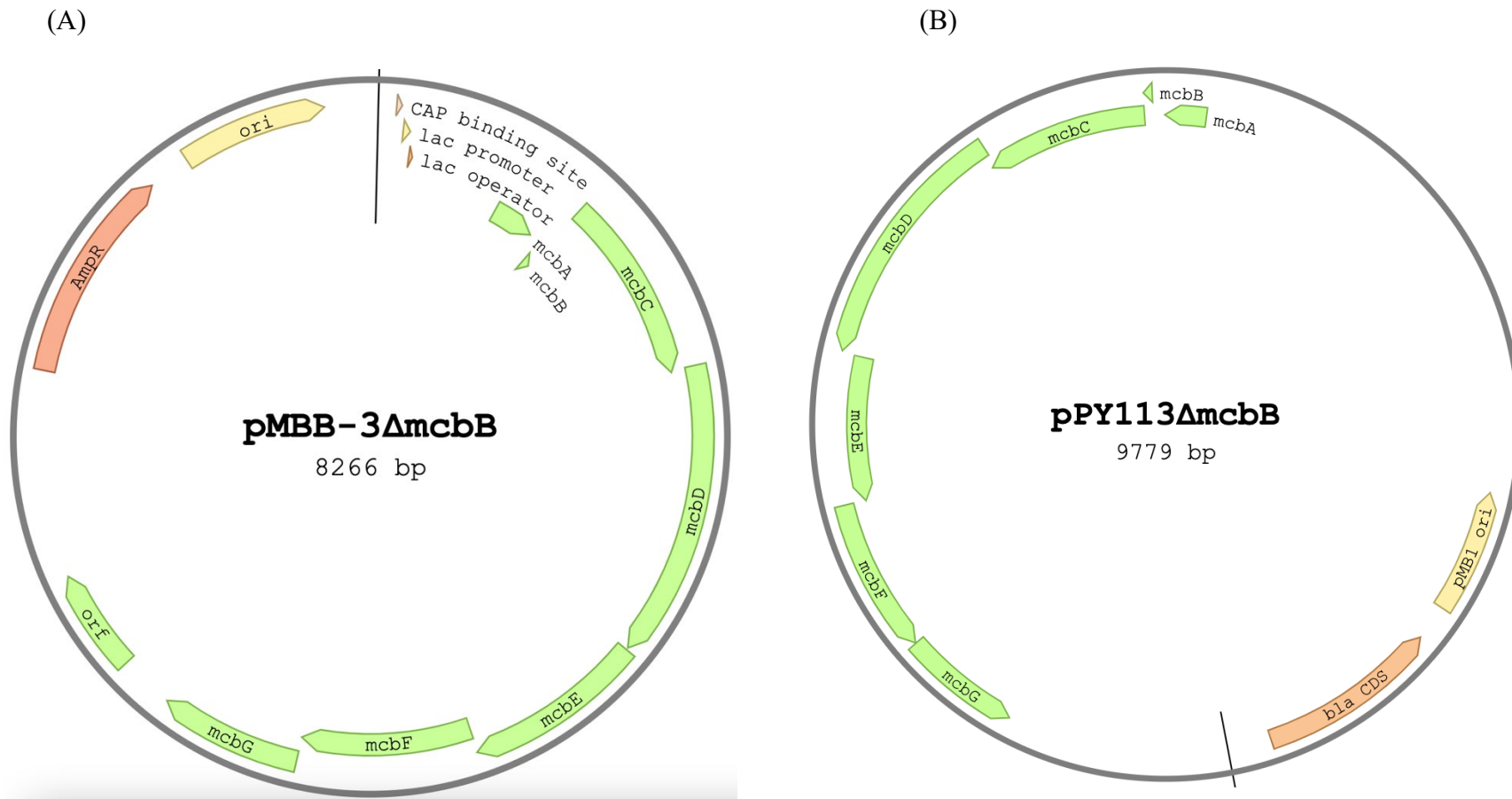
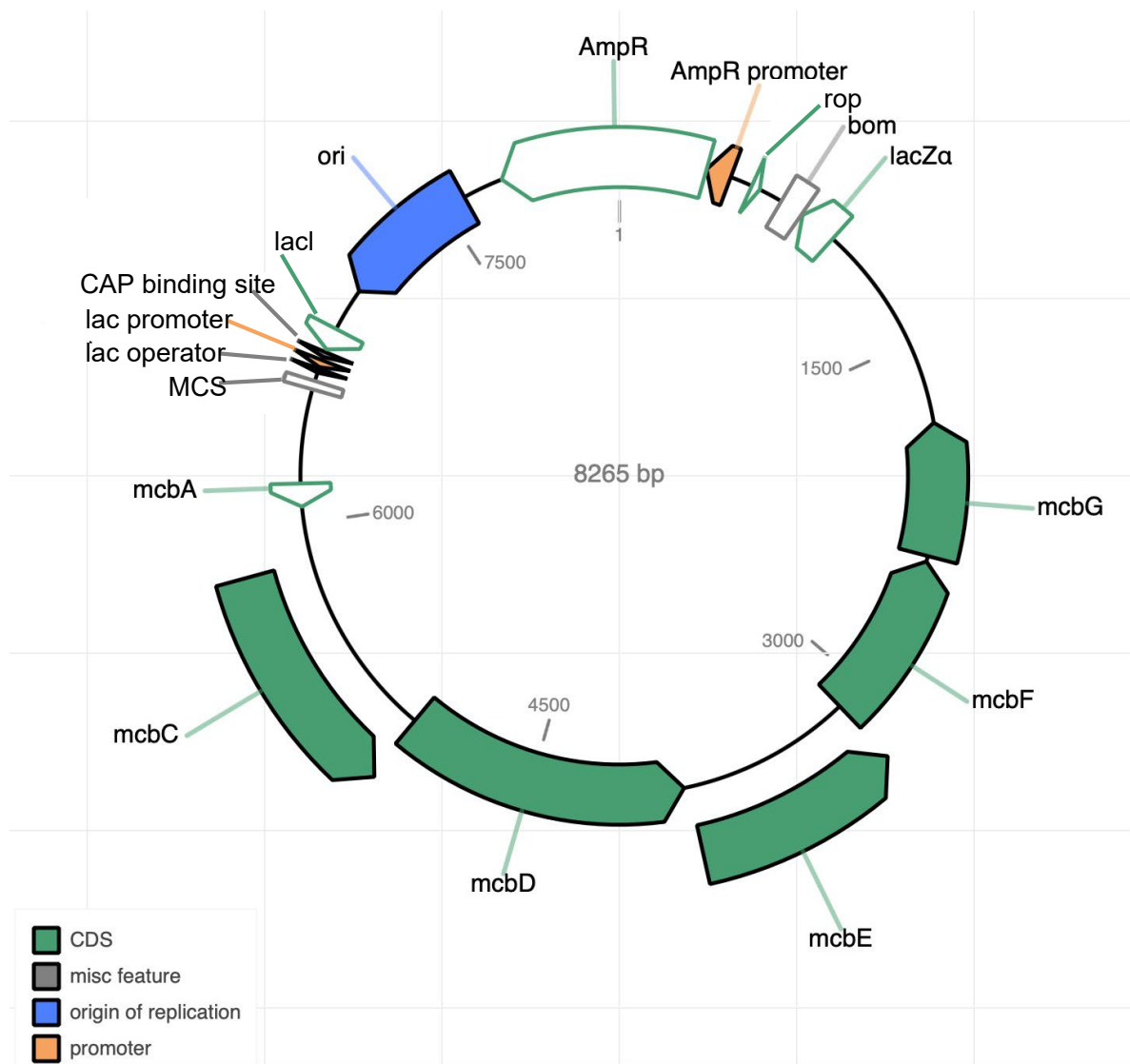


Figure 5.13 Plasmid maps of the cloned microcin operons showing deletion of *mcbB* and incomplete microcin operon. (A) pMBB-3Δ*mcbB*, *Erwinia* microcin operon without *mcbB* (B) pPY113Δ*mcbB*, *E. coli* microcin operon without *mcbB*

The PCR products were purified from gel and self-ligated using TA DNA ligase. The results are circular plasmid constructs, pMBB-3 Δ *mcbB* and pPY113 Δ *mcbB*, shown in Fig. 5.13. The plasmid constructs were transformed into *E. coli* DH5 α cells, where they were expressed, extracted, and sent out to be sequenced. Long Oxford Nanopore reads for full length plasmid sequencing (by plasmidsaurus, USA) was used to confirm the deletion of *mcbB* from the *emo* and *mcb* operon (Fig. 5.14). Additional confirmation was carried out by digesting the plasmids carrying the complete microcin operons, pMBB-3 and pPY113, as well as the plasmid constructs, pMBB-3 Δ *mcbB* and pPY113 Δ *mcbB*, with restriction enzyme *AleI*. The agarose gel image from the digestion was compared with a virtual digestion carried out on Benchling. *AleI* cuts both pMBB-3 Δ *mcbB* and pMBB-3 once. However, the band of pMBB-3 Δ *mcbB* is slightly lower than pMBB-3 because of the deletion of *mcbB*, 855 bp (Fig. 5.15). On the other hand, *AleI* cuts pPY113 Δ *mcbB* once but cuts pPY113 twice, as seen in both the agarose gel image and the virtual digest (Fig. 5.16).

(A)



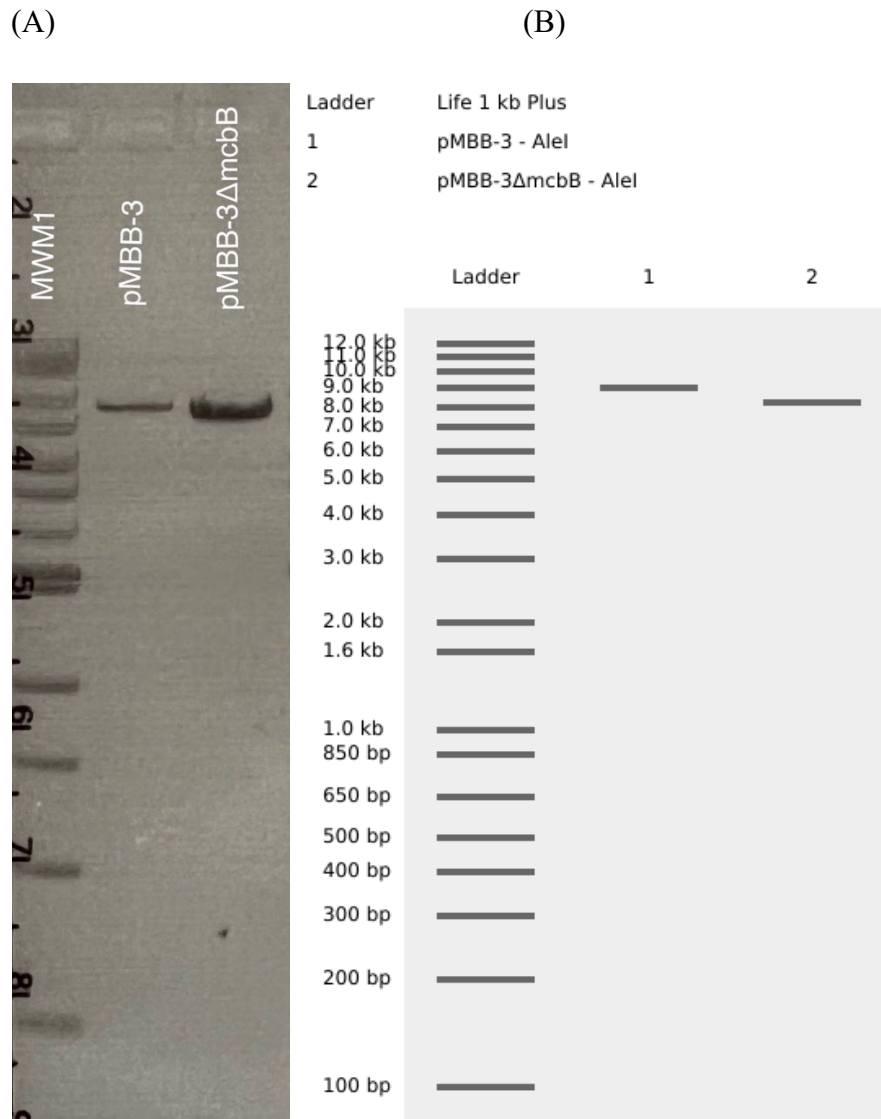


Figure 5.15 Agarose gel electrophoresis to confirm deletion of *mcbB* by restriction enzyme digest of pMBB-3Δ*mcbB* and pMBB-3. (A) Digestion of pMBB-3Δ*mcbB* and pMBB-3 with AleI (B) Virtual digest of pMBB-3Δ*mcbB* and pMBB-3. Gel concentration is 0.7 % agarose gel in TAE buffer. MWM, molecular weight marker (1 kb Extend DNA Ladder, NEB).

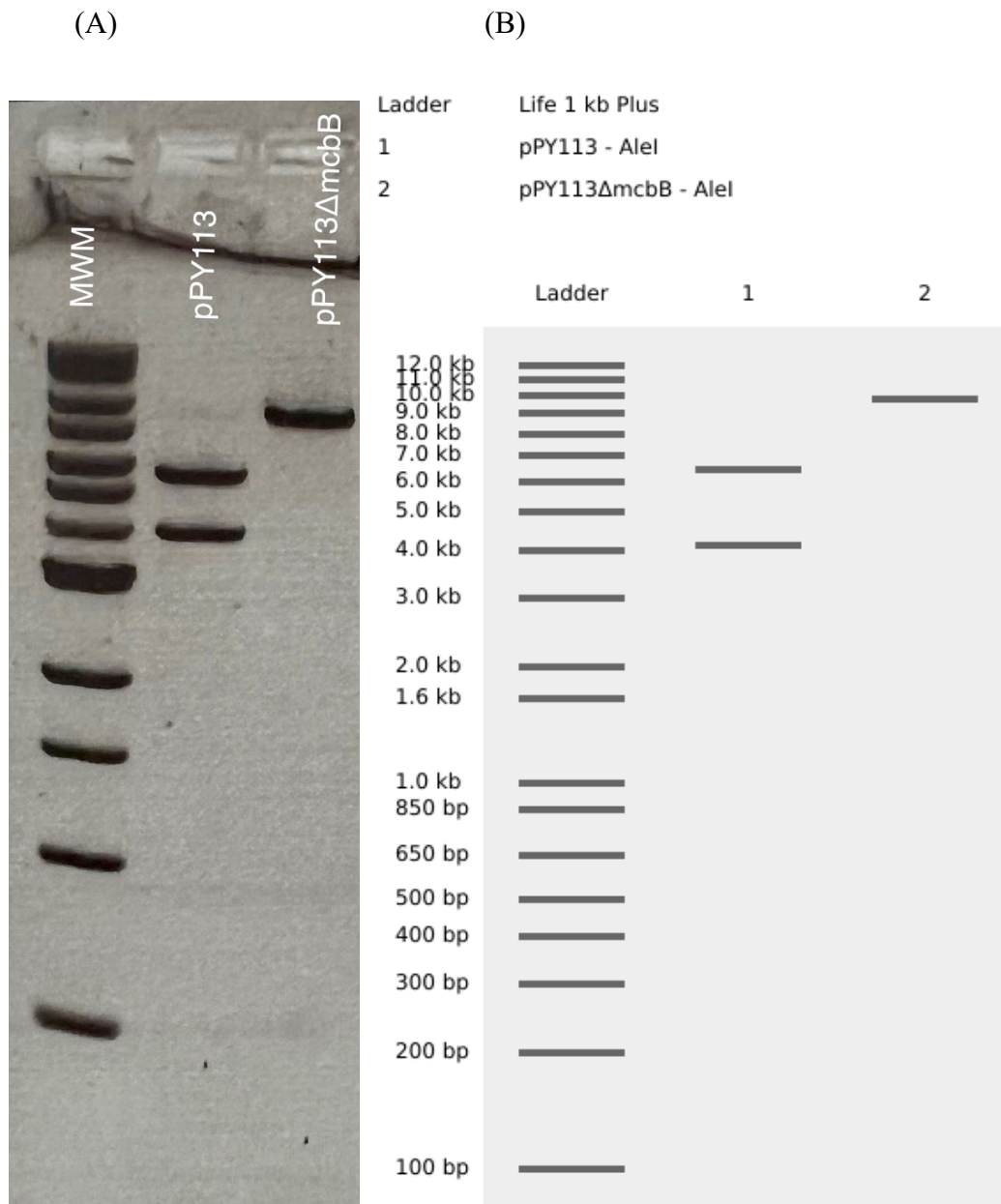


Figure 5.16 Agarose gel electrophoresis to confirm deletion of *mcbB* by restriction enzyme digest of pPY113Δ*mcbB* and pPY113. (A) Digestion of pPY113Δ*mcbB* and pPY113 with AleI (B) Virtual digest of pPY113Δ*mcbB* and pPY113. Gel concentration is 0.7 % agarose gel in TAE buffer. MWM, molecular weight marker (1 kb Extend DNA Ladder, NEB).

5.1.4 Cloning of *mcbB* into SEVA234 plasmid vector

The same approach used above for cloning *mcbA* into plasmid pSEVA234 was repeated for cloning *mcbB* into pSEVA234. PCR primers EcoR-Ep-*mcbB*-F, Sal-Ep-*mcbB*-R, EcoR-Ec-*mcbB*-F, and Sal-Ec-*mcbB*-R were designed to amplify *mcbB* from the *emo* and *mcb* operon (Table 2.1). A 1% agarose gel was run to check the amplicons which were thereafter purified from solution using a PCR and DNA cleanup kit. The purified PCR products were digested with EcoRI and SalI in preparation for ligation with pSEVA234. The pSEVA234 was also digested with EcoRI and SalI, dephosphorylated and ligated with the inserts, *mcbB*. The ligation process produced cloned plasmids pSEVA234*mcbB*EP and pSEVA234*mcbB*EC (Fig. 5.17).

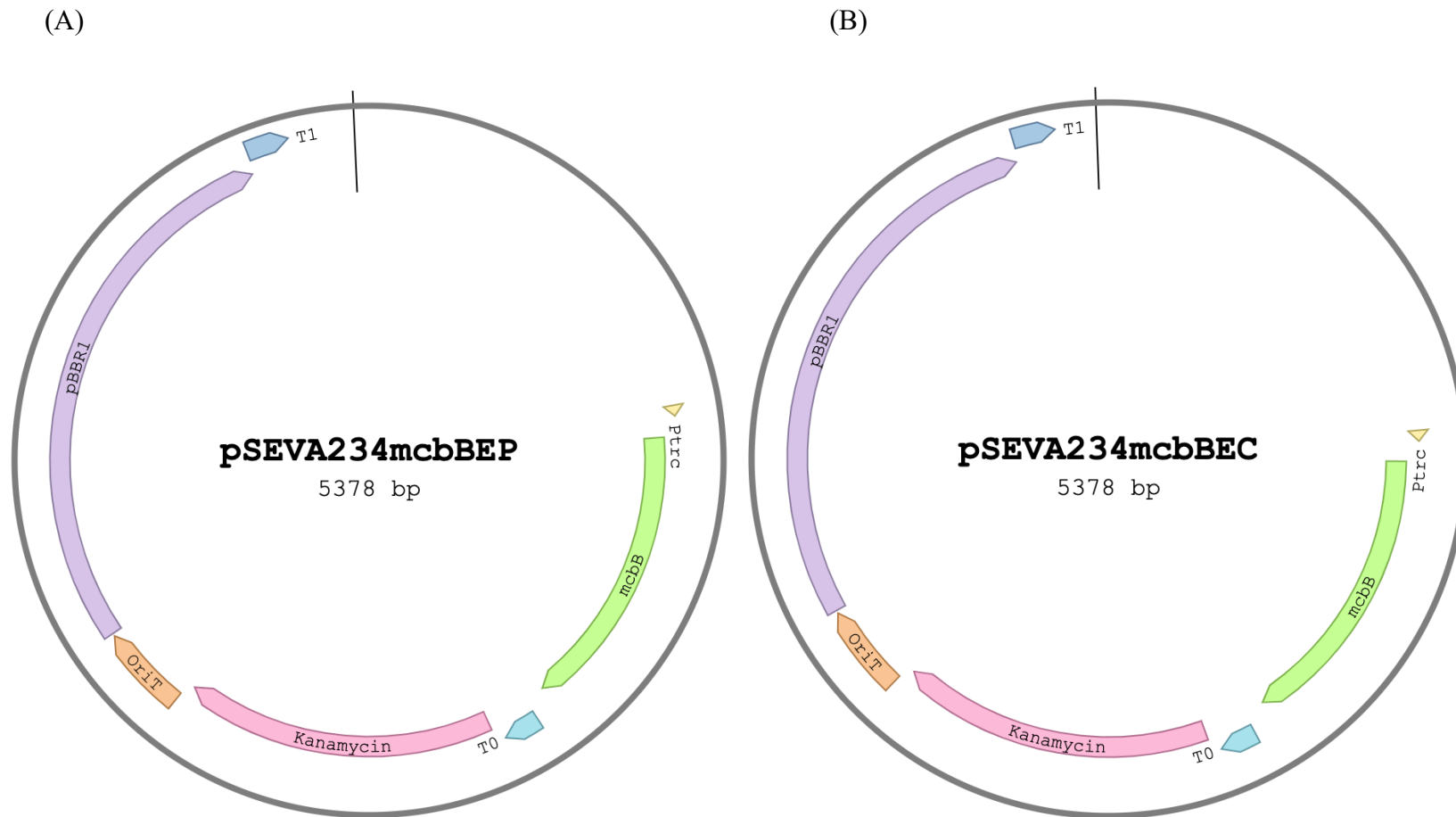
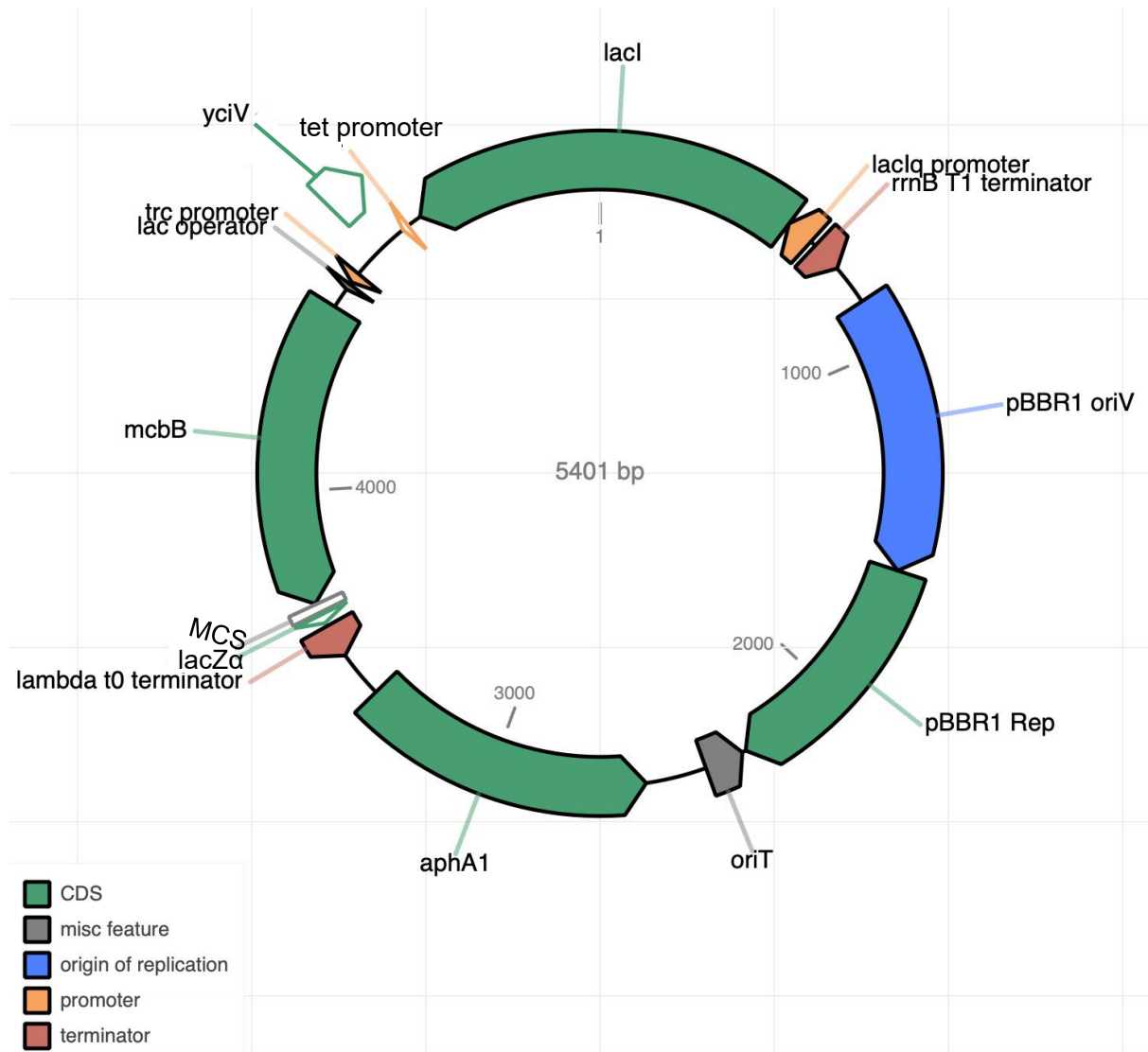


Figure 5.17 Plasmid maps of pSEVA234 derivatives with *mcbB* insert (A) pSEVA234*mcbBEP* (B) pSEVA234*mcbBEC*. EP and EC represent *Erwinia persicina* and *E. coli* respectively.

The cloned plasmids were transformed into *E. coli* DH5a cells, the transformant cells were selected, and their plasmids were extracted. These plasmids were sent to plasmidsaurus for Long Oxford Nanopore reads for full length plasmid sequencing to confirm that *mcbB* from both microcin operons were correctly cloned into pSEVA234. Fig. 5.18 shows the result of full length plasmid sequencing as a map of pSEVA234 with *mcbB* inserts to confirm the successful cloning of *mcbB* into pSEVA234.

(A)



(B)

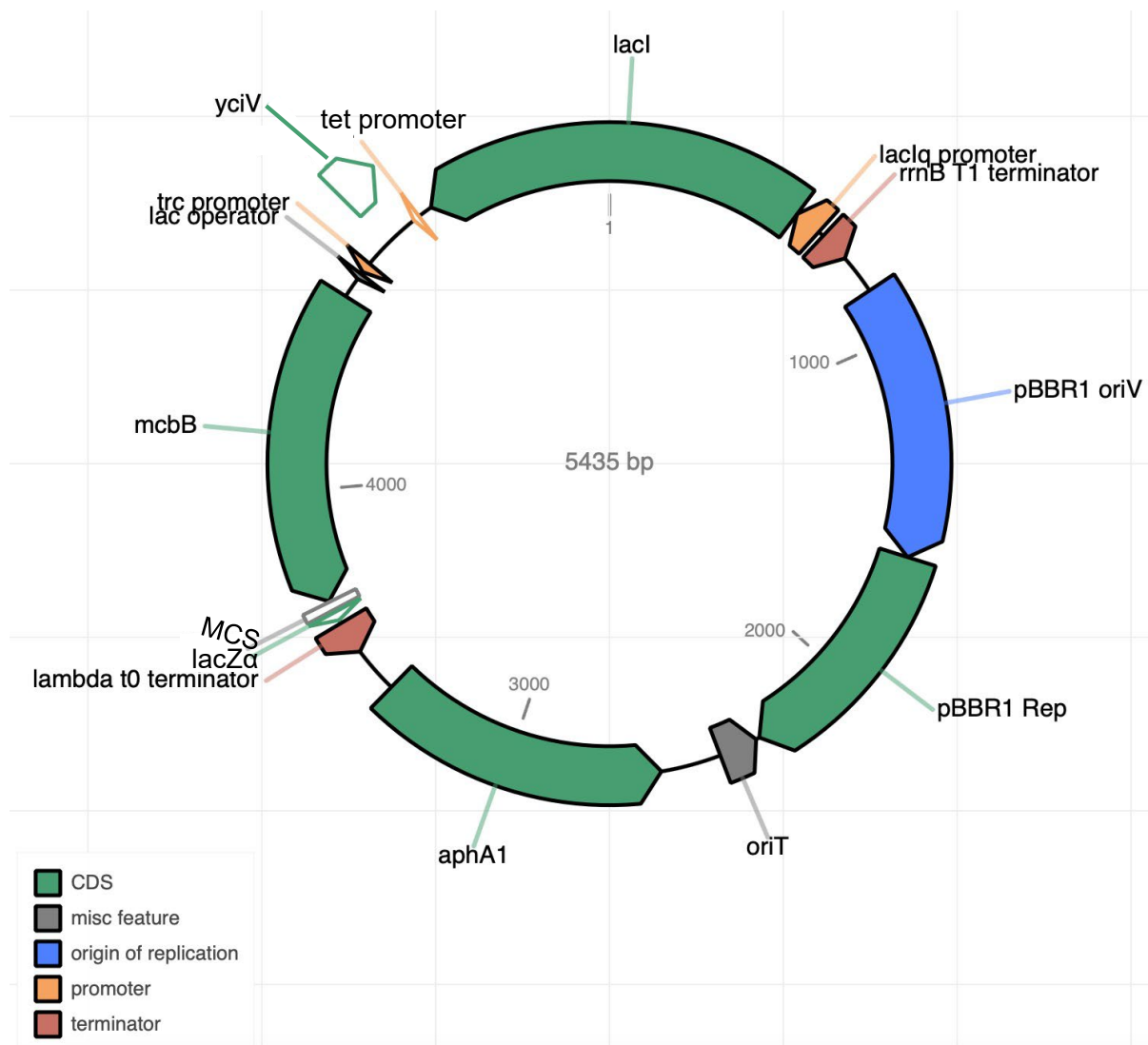


Figure 5.18 Plasmid maps of cloned *mcbB* into pSEVA234. (A) pSEVA234*mcbB*E_P (B) pSEVA234*mcbB*E_C. These were derived from whole-plasmid sequencing data by Plasmidsaurus using the pLannotate tool (McGuffie and Barrick, 2021).

5.2 Pairing of complementary gene(s) in the microcin operon (mix and match)

This section seeks to understand how promicrocin and synthase will affect microcin production, as well as to find out if the synthases are promicrocin specific (do the synthase interacts specifically with their cognate *mcbA*), and if the synthase complex will function while having component(s) from another synthase complex (for instance, having *E. coli*

mcbB to replace the *mcbB* in *emo* operon). The cloned microcin genes / operons were paired in various combinations, cognate (where all the genes are from the same source) and non-cognate (where the genes are a mixture of *E. persicina* and *E. coli* genes), and transformed into *E. coli* DH5 α cells for expression (Fig. 5.1B). This section will focus on the pairing of the plasmid constructs made in the sections above and the result of their expression. For ease of pairing, plasmids of interest were classified into two categories, A and B (Table 5.1). Category A is the SEVA234 plasmid with or without genes of interest, *mcbA* or *mcbB*. Category B are plasmids that either carry the complete operon, an incomplete operon, or an empty plasmid vector without the microcin operon.

Empty plasmid vectors without microcin operon, pUC19 for *Erwinia* and pBR322 for *E. coli*, are needed as a control in the bioassay experiment to study the expression of the plasmids that will be used during the mix and match pairing of the genes. pPY113 (pMccB17 expression cell) is a derivative of pBR322 that lacks Tc^R gene, and we had to engineer pPY113 by deleting the *mcb* operon previously inserted into it to get a copy of pBR322 that also lacks Tc^R gene (pBR322 Δ T) like pPY113 so that it would be a matched control.

Table 5.1 Pairing selection of microcin biosynthesis gene(s) used in the mix and match experiment.

Category A	Category B	Choice of Antibiotics	Given names
	pPY113	Amp	pPY113
pSEVA234	pPY113 Δ <i>mcbA</i>	Km + Amp	ECM01
pSEVA234 <i>mcbA</i> ECC	pPY113 Δ <i>mcbA</i>	Km + Amp	ECM02
pSEVA234 <i>mcbA</i> ECC	pBR322 Δ T	Km + Amp	ECM03
	pMBB-3	Amp	pMBB-3
pSEVA234	pMBB-3 Δ <i>mcbA</i>	Km + Amp	ECM04
pSEVA234 <i>mcbA</i> ECC	pMBB-3 Δ <i>mcbA</i>	Km + Amp	ECM05
pSEVA234 <i>mcbA</i> ECC	pUC19	Km + Amp	ECM06
pSEVA234 <i>mcbA</i> ECC	pMBB-3 Δ <i>mcbA</i>	Km + Amp	ECM07
pSEVA234 <i>mcbA</i> ECC	pPY113 Δ <i>mcbA</i>	Km + Amp	ECM08
pSEVA234	pPY113 Δ <i>mcbB</i>	Km + Amp	ECM09
pSEVA234 <i>mcbB</i> ECC	pPY113 Δ <i>mcbB</i>	Km + Amp	ECM10
pSEVA234 <i>mcbB</i> ECC	pBR322 Δ T	Km + Amp	ECM11
pSEVA234	pMBB-3 Δ <i>mcbB</i>	Km + Amp	ECM12
pSEVA234 <i>mcbB</i> ECC	pMBB-3 Δ <i>mcbB</i>	Km + Amp	ECM13
pSEVA234 <i>mcbB</i> ECC	pUC19	Km + Amp	ECM14
pSEVA234 <i>mcbB</i> ECC	pMBB-3 Δ <i>mcbB</i>	Km + Amp	ECM15
pSEVA234 <i>mcbB</i> ECC	pPY113 Δ <i>mcbB</i>	Km + Amp	ECM16

Highlighted in yellow are selections that do not contain a complete set of microcin biosynthetic genes.

5.2.1 Deriving pBR322 from pPY113

The plasmid pPY113 is a derivative of pBR322 made by cloning the *mcb* operon into the pBR322 Tc^R gene; as a result, the Tc^R gene was non-functional in pPY113 (Yorgey *et al.*, 1993). To get pBR322 from pPY113, pPY113 was digested with BamHI and EcoRI; this resulted in 3 fragments of DNA, a fragment of about 5.5 kb and a 0.8 kb, which sum up the size of the microcin operon that was cloned into the third fragment with a size of about 4.3 kb

(Fig. 5.19). The 4.3 kb fragment (pBR322) was excised, gel purified, ligated and transformed into *E. coli* DH5 α . This fragment was called pBR322 Δ T because it lacks Tc^R gene like pPY113. The transformant cells were selected, and the plasmid pBR322 Δ T was extracted for further usage.

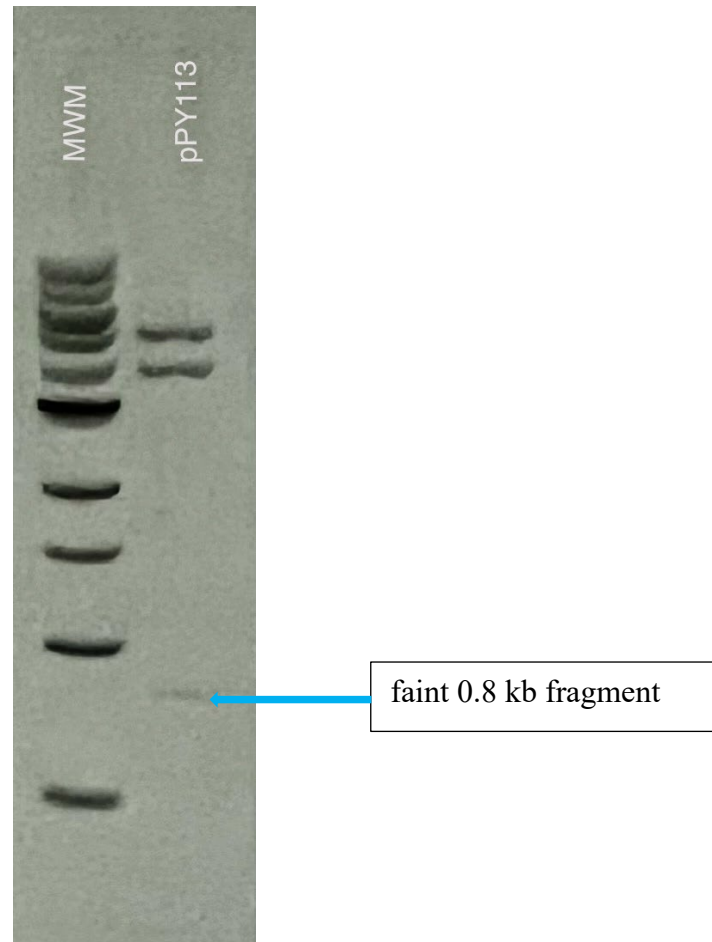


Figure 5.19 Digestion of pPY113 with BamHI and EcoRI to get pBR322. Highlighted by the blue arrow is the faint 0.8 kb fragment. Gel concentration is 0.8 % agarose gel in TAE buffer. MWM, molecular weight marker (1 kb plus DNA Ladder, NEB).

5.2.2 Bioassay analysis of the paired microcin genes

Plasmids in category A were co-transformed with plasmids in category B into *E. coli* DH5 α cell following the pairing arrangement in Table 5.1. These cells were tested for microcin activity by using bioassay experiments to assess them. pSEVA234 and pUC19 were transformed into *E. coli* BL21 cells and named SUC19. *E. coli* BL21(SUC19) was used as the indicator strain for this round of bioassay experiments conducted to rule out kanamycin and ampicillin interference on the indicator strain as pSEVA234 contains Km^R gene and pUC19 contains Amp^R gene. With pSEVA234 having a *lac* promoter, the bioassay experiment was set up also to study the effect IPTG will have on microcin production of the altered microcin operons (Fig. 5.20). Before the bioassay, some pairing selections were perceived to lack the ability to synthesise microcin because they either have no microcin operon present in their plasmid(s) or they do not have the complete genes responsible for microcin production, and they were used as controls. These selections are highlighted in yellow in Table 5.1, and results from their bioassay experiments confirmed that no zone of inhibition was produced.

An overview of the bioassay observing the effect of 0.1mM IPTG on microcin production of the paired selection revealed that there is a significant difference between the selections that contained IPTG in their agar plate and those selections that lacked IPTG. This effect which varies with different pairing selections is due to the inducible *P_{trc}* promoter in the vector pSEVA234. For instance, the ECM02, ECM05, ECM07 and ECM13 selections showed increased microcin production in the presence of IPTG, while the ECM10, ECM15 and ECM16 selections exhibited increased microcin production in the absence of IPTG. The presence of IPTG did not affect microcin production in pPY113, pMBB-3 and ECM08. Microcin production in the *mcb* operon appears to be generally greater when the operon is in a single plasmid (pPY113) compared to when the genes are split between two plasmids ECM02 (*mcbA* on a different plasmid in this selection) and ECM10 (*mcbB* on a separate plasmid in this selection). The only exception to this observation is pECM10 without IPTG, where the microcin production is more than that in pPY113. On the other hand, microcin production in the *Erwinia* operon appears better when the microcin operon is split between two plasmids, as seen in ECM05 and ECM13; the exception is ECM13 in the absence of IPTG.

Four selections resulted in a non-cognate microcin operon, ECM07, ECM08, ECM15 and ECM16. It was observed between these four pairs that there is better microcin production

when category A of the selection is from *E. coli* and category B of the selection is from *Erwinia*, ECM07 (pSEVA234*mcbA*E_{EC} and pMBB-3Δ*mcbA*) and ECM15 (pSEVA234*mcbB*E_{EC} and pMBB-3Δ*mcbB*). The reason for this is unknown, but it could be due to *E. coli mcbA*, which is common to both selections. In chapter 3, it was mentioned that the presence of heterocycles formed by GSS tripeptide, which is found in *Erwinia McbA*, has a lower yield than the heterocycles formed by either GSC or GCS tripeptide which are present in *E. coli McbA* based on the experiments of Sinha Roy *et al.* (1999). This difference in amino acid sequence could be one of the reasons for the production difference between the group with the non-cognate microcin operon.

EC05 and ECM07 both in the presence of IPTG, show the highest level of microcin production; the link between them is the category B plasmid (pMBB-3Δ*mcbA*) in their mix and match selection, as well as the presence of IPTG in their expression. As pMBB-3 was derived from pUC19, and pUC19 has a *lac* promoter, we can infer that the presence of IPTG in the medium increase the expression of *McbA*. There appears to be no synergy between the pairing of pSEVA234*mcbA*E_{EP} and pPY113Δ*mcbA* to form ECM08, as their combination produced the lowest microcin.

All non-cognate pairings (ECM07, ECM08, ECM15 and ECM16) tried in this study worked, indicating that the protein-protein interaction between the promicrocin and the synthase is not species specific. This could be largely due to the two operons having conserved motifs in the promicrocin and conserved active sites across their synthase. In chapter 4, it was shown that the motif present in *McbA* that is responsible for the leader peptide to bind to the synthase, and the active site of *McbB* that mediates the recognition of leader peptide in *McbA* are found in both *mcb* and *emo*. Further bioinformatic examination has revealed that the active sites present in *McbC* and *McbD* of *mcb*, as reported by Ghilarov *et al.* (2019), are also present in *emo*. In *McbC*, there are Arg82, Arg117 and Arg181 of the *McbC* subunit, which contributes to the binding of FMN cofactor molecules, as well as Lys201 and Tyr202, which function as catalytic residues (Fig. 5.21). In *McbD*, the conserved elements were reported to be residues 147–151 and the C-terminal tail (393–396) in *mcb* (Fig. 5.22). All these sites are conserved in *emo*; we have confirmed that swapping the promicrocin or the synthases between *emo* and *mcb* will not halt microcin production.

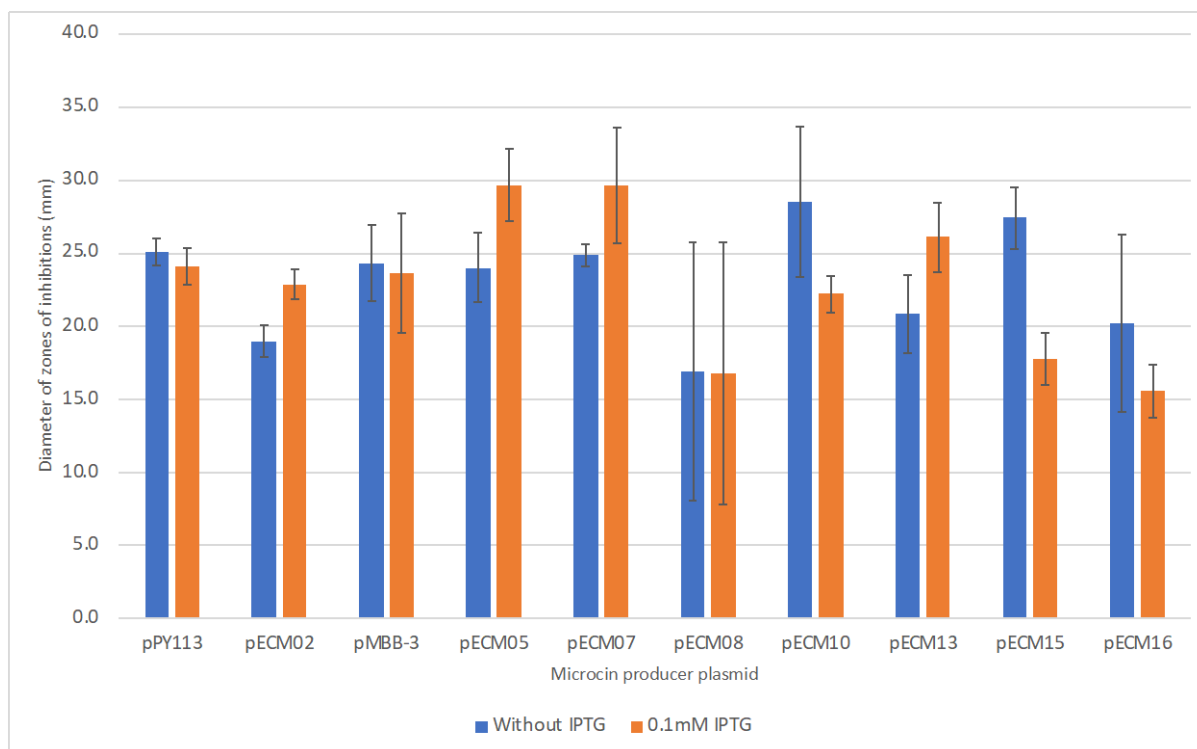


Figure 5.20 Activity of microcin production by the cognate and non-cognate microcin operons. *E. coli* BL21(SUC19) was used as the indicator strain. Mean values and standard deviations (error bars) were obtained from three experiments, each performed in triplicate.

E. coli	1	MINVYSNLSAWPATMAMSPKLNRMPTFSQIWDYERITPASAAGETLKSIQGAIGEYFERRHFFNEIVTGGQKTLYEMPPSAAKAFTEAFFQISSLTRDEIITHKFKTVRAFNLFSLEQQEIPAVIIALDNITAADDLKFYPDR ¹⁴⁷ DTCG	150
E. persicina	1	MINVYSNLSAWPATMVMSPKLNRMPTFSQVWDYERITPTSAAGETMKSQGAIGEYFERRHFFNEIVANGKTLYDMMSPHAADAFTRAFVQTSSLAEDVVRTHQFKTTRAFNLFSLEKQEIPAVIIALDNITAASDLTIYPDR DTCG	150
E. coli	151	CSFHGSLNDAIEGSLCEFMETQSLLLYWLQKGKANTEISSEIVTGINHIDEILLALRSEGDIFDITLPGAPGHAVLTLYGTKNKISRIKYSTGLSYANSLKKALCKSVELWQSYICLHNFLIGGYTDDDIIDSYQRHFMSCKYESFT	300
E. persicina	151	CSFHGNLNDAMDGALCEFMERQSLLLYWLQGQANTEISGEIITGINYIDEIMSALRSEGEIRIFDITLPGAPGHAVLTLYGTTNEDSQIKYSTGLSYANSMKALCKSVTELWQSYICMHNFLIGGYTDEDIIDYQRHFLSCNKYESFT	300
E. coli	301	DLCENTVLLSDDVKLTLEENITSDTNLLNYLQQISDNIFVYYARERVSNSLVWYTKIVSPDFFLHMNNSGAININNKIYHTGDGKIVRESKM ³⁹³ VPFP	350
E. persicina	301	DLCENTLFRQSDVVLSSL-ENPPSDNDLLSYLEKISDNIFVYYARERASENLFWYTKVSPDFFLHMNGSGAININNKLRYAGNGIKHREFKM VPFP	349

. Different amino acid
: Similar amino acid
| Identical amino acid

Figure 5.22 Alignment of McdB from *E. coli* and *E. persicina* amino acid sequences. Highlighted in red are the conserved elements in McdB.

Chapter six

Discussion and conclusions

6.1 Plasmid pMccB17

While earlier studies involving pMccB17 focused on the *mcb* operon, this study focused on the whole plasmid, including the *mcb* operon. Plasmid pMccB17 is a typical IncFII conjugative plasmid. Baquero *et al.* (1978) studies supported this when they successfully transferred the plasmid from its original host *E. coli* strain LP17 to a new host *E. coli* strain, BM21, through mating. The transconjugant microcin positive cells displayed the same microcin properties as the donor cell used for the conjugation process. Sequencing pMccB17 and bioinformatically analysing it, revealed all the expected conjugation genes present in the plasmid, confirmed the previous assignment (based on transconjugant loss of microcin attributes after mating with strains that contained IncFII plasmid) to the incompatibility group FII (Baquero *et al.*, 1978). In addition, a replicon typing system utilizing the FAB formula was used to classify pMccB17 more precisely into the F2:A-:B- type. The F2:A-:B- replicon type is very common in *E. coli* and is strongly associated with the carriage of antimicrobial resistance genes, e.g. ESBL gene (Kondratyeva *et al.*, 2020; Puangseree *et al.*, 2022). Based on analysis of the relaxase (the protein that starts and terminates conjugative DNA processing), we were also able to place the plasmid into the MOB_{F12} group. This group is a subclade of MOB_{F1} (an important clade to the Gammaproteobacteria which includes plasmids of the classical IncF, IncN, IncP9 and IncW incompatibility group) (Garcillán-Barcia *et al.*, 2009). It has previously been reported that pMccB17 has no antibiotic resistance gene marker; bioinformatic analysis of the pMccB17 sequence during this study also supported this (Mayo *et al.*, 1988; Millan *et al.*, 1985).

One of the interesting qualities of this plasmid was the presence of two direct repeat regions that bordered some important plasmid maintenance genes. The region between these two repeats in pMccB17 corresponds to the plasmid leading region originally defined for F plasmid (Loh *et al.*, 1989). Whether these repeats' roles are related to functional sequence tasks for the plasmid, or serve as a form of mobile genetic sequence elements (like insertion sequence or transposons) remains unknown. However, checking the plasmid sequence for insertion sequences revealed that no known type was present. Using a targeted mutagenesis methodology (such as a lambda Red recombineering approach; Lee *et al.*, 2009), these repeats can be deleted in future experiments in order to ascertain their function (if any) in

relation to things like gene transcription (via RNAseq analysis) and plasmid replicative, conjugative or maintenance functions (by assaying of plasmid stability, copy number and conjugative transfer frequency and range).

The Pbf protein was another feature of interest noticed in the plasmid pMccB17, whose precise function was unknown. This protein has an N-terminal region similar to ParB proteins, and the C-terminal region lacks any conserved domain. Based on the similarity the N-terminal region has with ParB proteins and bioinformatic analysis, we predicted that this protein could be involved in plasmid conjugation or maintenance. Further recommended experiments will be to knock out the gene that encodes this protein and observe the outcome on the plasmid if the plasmid can survive, replicate, conjugate and remain stable using the approaches outlined above. In a case where the gene is an essential gene, deletion mutation will make the plasmid harbouring it non-viable, and as such, the gene can be overexpressed from an inducible expression vector *in trans* to pMccB17 to observe the effect that will have on plasmid attributes such as copy number, conjugation.

Possibly the most important thing about the plasmid pMccB17 is that it carries the *mcb* operon. It was mentioned that cells harbouring only the four production genes (*mcbABCD*) without the export and immunity genes (*mcbEFG*) would become inviable because they lack protection from the synthesised MccB17 (Garrido *et al.*, 1988). One question we would like to answer is what will be the effect on plasmid stability, replication (and conjugation) if the entire *mcb* operon was to be lacking from the pMccB17 plasmid? A knockout experiment on the *mcb* operon is suggested as further work to find out if the operon with its immunity function will act as a plasmid addiction system (similar to a toxin-antitoxin module) and the effect a lack of *mcb* operon will have on pMccB17 functions. An alternative approach will be to clone the operon into an unstable plasmid along with an antibiotic resistance marker (to help with selection) and observe if the operon will make it more stable when the antibiotic selection is removed. The minor sequence variations that were noticed between some genes (*mcbB*, *mcbD* and *mcbF*) from the first published sequence of *mcb* operon and the sequence from this work could be due to mutations in cloned copies of these genes or simple errors in Sanger sequencing of the cloned fragments (containing the production, export and immunity genes of MccB17) from the pMccB17 plasmid (Garrido *et al.*, 1988; Genilloud *et al.*, 1989b).

6.2 The *Erwinia* microcin, EMB17

Initial observation of the regions of both microcin operons (*emo* and *mcb*) would have made it easy to predict that the presence of the *eeg*, a gene of unknown function, found downstream of the *emo* operon in comparison to the *mcb* operon would be one of the main differences between the two operons. However, bioinformatic analysis of the *eeg* reveals that it has no known function and using bioassay to find out if its function is to increase microcin production showed that this is not the case as it has no obvious effect on microcin production, at least in the *E. coli* heterologous host used here. We conclude that it is a gene that has nothing to do with microcin production, even though it is close to the microcin gene cluster in *E. persicina*.

The two microcin operons (*emo* and *mcb*) and their products (EMB17 and MccB17) were similar in a lot of ways: from the number of genes that made up their operon (seven in both) down to the nucleotides of the genes in the operon (87% similarity), to the fact that microcin production for both starts at stationary phase – the expression of the *emo* operon in heterologous *E. coli* host might have an effect on this (Pierrat and Maxwell, 2003; Sinha Roy *et al.*, 1999), to their spectrum of action where they both do not inhibit any Gram-positive bacteria but inhibited *E. coli*, *Klebsiella* and *Pseudomonas* confirming both have a narrow-spectrum of action (Baquero and Moreno, 1984; Duquesne *et al.*, 2007). However, with all these similarities, we observed some differences between corresponding genes from each operon, but this did not affect the conserved domains and functions of these genes. The main difference between the two operons was in the product of their first gene, McbA, where there is one amino acid difference in the substrate peptide converted by the synthase into a heterocycle in MccB17. The heterocycle in EMB17 at this location did appear to have a lower activity when compared with the heterocycle at the same position in MccB17. Previous studies where a mutant variant of MccB17 (having a similar structure to EMB17) was synthesised and compared to the wild type MccB17 showed that this mutant variant of MccB17 has a lower activity in comparison to the wild type MccB17 (Collin *et al.*, 2013; Shkundina *et al.*, 2014; Sinha Roy *et al.*, 1999). The bioassay result in this work also supported this, as MccB17 had a larger zone of inhibition when compared to EMB17. However, we do not know that production levels of microcin are the same in both cases, so we would need to confirm this finding using purified (and quantified) microcin. The polyglycine linker from both McbA peptides also differs by three additional glycines present in EMB17. Studies suggested that the longer the polyglycine linker, the more product

formation increases (Sinha Roy *et al.*, 1998). This could be a way to at least partly compensate for the deficit caused by the “weaker heterocycle” in EMB17.

MccB17 gains entry into target cells through the OmpF porin situated in the outer membrane and the SbmA protein located across the inner membrane of target cells (Duquesne *et al.*, 2007). Through a bioassay experiment where target cells with mutations in *ompF* and *sbmA* were used, we were able to confirm that EMB17 gains entry into target cells using the OmpF and SbmA proteins. We were also able to confirm that the inner membrane protein, SbmA, is crucial for EMB17 (as well as MccB17) uptake into the cell. The absence of this protein in the sensitive cell will make the cell fully resistant to EMB17 and MccB17, as seen in Ghilarov’s and their team studies and in this work. (Ghilarov *et al.*, 2021). In contrast, the absence of OmpF porin will offer partial resistance (Lavina *et al.*, 1986).

Initial attempts to purify EMB17 were unsuccessful, and a QC sample was made by mixing EMB17 and MccB17 extracts (from solid phase extraction) at a 1:1 ratio; this was run on LCMS to pick up peaks of important masses. The mass of 1088.9 m/z in positive mode was picked up, and this was used in preparative HPLC-MS to purify EMB17. The purified EMB17 was assessed for microcin activity against sensitive *E. coli* and showed a zone of inhibition. Although the zones of inhibition formed before the microcins were purified was bigger than those observed after the purification of the microcins. One of the reasons for this could be the poor physico-chemical properties of MccB17 reported by (Collin *et al.*, 2013). However, we were able to confirm the presence of active microcin after the purification steps. Future work on the purified EMB17 should be aimed at standardizing the concentration and measure of specific activity (i.e. activity per µg microcin peptide).

This work showed that some of the genes responsible, specifically the *mcbA* encoding the promicrocin and *mcbB* encoding one of the synthase components, for microcin production can function in a non-cognate setting and are not species specific. This lack of specificity or stringency of molecular interactions will potentially be useful for the future development of genetically engineered microcins with enhanced or altered properties, e.g. feeding novel peptides to the synthase by altering the sequence of *mcbA*.

MccB17 (and the EMB17 in this work) holds lots of potential, one of which is that they can be used as a source of biological control to fight and control plant pathogens (Gillor *et al.*, 2005). EMB17 being a microcin produced by *E. persicina*, might help inhibit some plant

pathogens from the Erwiniaceae family to curb some plant diseases, such as fire blight caused by *Erwinia amylovora* and bacterial wilt commonly caused by *Erwinia tracheiphila*.

However, limitation such as solubility is hampering the uses of these peptides. One way to address the limitations and improve the efficacy of bacteriocins is through manipulation (Cavera *et al.*, 2015). This work has provided a framework that could be built upon to help with the manipulation of microcin, which might eventually help address some limitations, such as stability and solubility. The genetic manipulations performed in this work have confirmed that it is possible to have the microcin operon split across two separate plasmids, which can be utilised in target or random mutations (in McbA, for instance) with the aim of improving the microcin's activity or physico-chemical properties.

References

- Adeolu, M., Alnajar, S., Naushad, S., and Gupta, R. (2016). Genome-based phylogeny and taxonomy of the 'Enterobacteriales': Proposal for Enterobacterales ord. nov. divided into the families Enterobacteriaceae, Erwiniaceae fam. nov., Pectobacteriaceae fam. nov., Yersiniaceae fam. nov., Hafniaceae fam. nov., Morganellaceae fam. nov., and Budviciaceae fam. nov. *International Journal of Systematic and Evolutionary Microbiology*, 66(12), 5575–5599.
- Ahsan, S., and Kabir, M. S. (2013). Linear plasmids and their replication. *Stamford Journal of Microbiology*, 2(1), 1–5.
- Allgaier, H., Jung, G., Werner, R. G., Schneider, U., and Zähler, H. (1986). Epidermin: Sequencing of a heterodet tetracyclic 21-peptide amide antibiotic. *European Journal of Biochemistry*, 160(1), 9–22.
- Altschul, S. (1997). Gapped BLAST and PSI-BLAST: A new generation of protein database search programs. *Nucleic Acids Research*, 25(17), 3389–3402.
- Altschul, S. F., Gish, W., Miller, W., Myers, E. W., and Lipman, D. J. (1990). Basic local alignment search tool. *Journal of Molecular Biology*, 215(3), 403–410.
- Aminov, R. I. (2010). A Brief History of the Antibiotic Era: Lessons Learned and Challenges for the Future. *Frontiers in Microbiology*, 1.
- Aminov, R. I. and Mackie, R. I. (2007). Evolution and ecology of antibiotic resistance genes. *FEMS Microbiology Letters*, 271(2), 147–161.
- Ansel, H. C., Norred, W. P., and Roth, I. L. (1969). Antimicrobial activity of dimethyl sulfoxide against *Escherichia coli*, *Pseudomonas aeruginosa*, and *Bacillus megaterium*. *Journal of Pharmaceutical Sciences*, 58(7), 836–839.
- Arnison, P. G., Bibb, M. J., Bierbaum, G., Bowers, A. A., Bugni, T. S., Bulaj, G., Camarero, J. A., Campopiano, D. J., Challis, G. L., Clardy, J., Cotter, P. D., Craik, D. J.,

- Dawson, M., Dittmann, E., Donadio, S., Dorrestein, P. C., Entian, K.-D., Fischbach, M. A., Garavelli, J. S., ... van der Donk, W. A. (2013). Ribosomally synthesized and post-translationally modified peptide natural products: Overview and recommendations for a universal nomenclature. *Natural Product Reports*, 30(1), 108–160.
- Arsenic, R., Treue, D., Lehmann, A., Hummel, M., Dietel, M., Denkert, C., and Budczies, J. (2015). Comparison of targeted next-generation sequencing and Sanger sequencing for the detection of PIK3CA mutations in breast cancer. *BMC Clinical Pathology*, 15(1), 20.
- Atlas, R. M. (2010). *Handbook of microbiological media* (4th ed). ASM Press; CRC Press/Taylor and Francis.
- Baquero, F., Bouanchaud, D., Martinez-Perez, M. C., and Fernandez, C. (1978). Microcin Plasmids: A Group of Extrachromosomal Elements Coding for Low-Molecular-Weight Antibiotics in *Escherichia coli*. *Journal of bacteriology*, 135(2), 342–347.
- Baquero, F., Lanza, V. F., Baquero, M.-R., del Campo, R., and Bravo-Vázquez, D. A. (2019). Microcins in Enterobacteriaceae: Peptide Antimicrobials in the Eco-Active Intestinal Chemosphere. *Frontiers in Microbiology*, 10, 2261.
- Baquero, F., and Moreno, F. (1984). The microcins. *FEMS Microbiology Letters*, 23, 117–124.
- Baumler, D. J., Ma, B., Reed, J. L., and Perna, N. T. (2013). Inferring ancient metabolism using ancestral core metabolic models of enterobacteria. *BMC Systems Biology*, 7(1), 46.
- Baxter, J. C., and Funnell, B. E. (2014). Plasmid Partition Mechanisms. *Microbiology Spectrum*, 2(6) 1-20.

- Belshaw, P. J., Roy, R. S., Kelleher, N. L., and Walsh, C. T. (1998). Kinetics and regioselectivity of peptide-to-heterocycle conversions by microcin B17 synthetase. *Chemistry and Biology*, 5(7), 373–384.
- Berriman, M. (2003). Viewing and annotating sequence data with Artemis. *Briefings in Bioinformatics*, 4(2), 124–132.
- Bertani, G. (1951). STUDIES ON LYSOGENESIS I: The Mode of Phage Liberation by Lysogenic *Escherichia coli*. *Journal of Bacteriology*, 62(3), 293–300.
- Bindiya, E. S., and Bhat, S. G. (2016). Marine Bacteriocins: A Review. *Journal of Bacteriology and Mycology: Open Access*, 2(5):140–147.
- Bingle, L. E. H., and Thomas, C. M. (2001). Regulatory circuits for plasmid survival. *Current Opinion in Microbiology*, 4(2), 194–200.
- Birnboim, H. C., and Doly, J. (1979). A rapid alkaline extraction procedure for screening recombinant plasmid DNA. *Nucleic Acids Research*, 7(6), 1513–1523.
- Blaser, M. (2011). Antibiotic overuse: Stop the killing of beneficial bacteria. *Nature*, 476(7361), 393–394.
- Boros, I., Pósfai, G., and Venetianer, P. (1984). High-copy-number derivatives of the plasmid cloning vector pBR322. *Gene*, 30(1), 257–260.
- Bower, C. K., Parker, J. E., Higgins, A. Z., Oest, M. E., Wilson, J. T., Valentine, B. A., Bothwell, M. K., and McGuire, J. (2002). Protein antimicrobial barriers to bacterial adhesion: In vitro and in vivo evaluation of nisin-treated implantable materials. *Colloids and Surfaces B: Biointerfaces*, 25(1), 81–90.
- Braun, V., Patzer, S. I., Hantke, K. (2002). Ton-dependent colicins and microcins: modular design and evolution. *Biochimie*. 84(5-6):365-80.
- Brito, R., Silva, G., Farias, T., Ferreira, P., and Ferreira, S. (2017). Standardization of the Safety Level of the Use of DMSO in Viability Assays in Bacterial Cells. *Proceedings*

of MOL2NET 2017, International Conference on Multidisciplinary Sciences, 3rd Edition, 4980.

Broadhurst, D., Goodacre, R., Reinke, S. N., Kuligowski, J., Wilson, I. D., Lewis, M. R., and Dunn, W. B. (2018). Guidelines and considerations for the use of system suitability and quality control samples in mass spectrometry assays applied in untargeted clinical metabolomic studies. *Metabolomics*, 14(6), 72.

Brooks, L., Kaze, M., and Sistro, M. (2019). A Curated, Comprehensive Database of Plasmid Sequences. *Microbiology Resource Announcements*, 8(1), e01325-18.

Bull, M. J., and Plummer, N. T. (2014). Part 1: The Human Gut Microbiome in Health and Disease. *Integrative Medicine*, 13(6).

Carattoli, A., Villa, L., Fortini, D., and García-Fernández, A. (2021). Contemporary IncII plasmids involved in the transmission and spread of antimicrobial resistance in Enterobacteriaceae. *Plasmid*, 118, 1-10.

Carattoli, A., Zankari, E., García-Fernández, A., Larsen, M. V., Lund, O., Villa, L., Aarestrup, F. M., and Hasman, H. (2014). In Silico Detection and Typing of Plasmids using PlasmidFinder and Plasmid Multilocus Sequence Typing. *Antimicrobial Agents and Chemotherapy*, 58(7), 3895–3903.

Carver, T., Berriman, M., Tivey, A., Patel, C., Böhme, U., Barrell, B. G., Parkhill, J., and Rajandream, M.-A. (2008). Artemis and ACT: Viewing, annotating, and comparing sequences stored in a relational database. *Bioinformatics*, 24(23), 2672–2676.

Carver, T. J., Rutherford, K. M., Berriman, M., Rajandream, M.-A., Barrell, B. G., and Parkhill, J. (2005). ACT: The Artemis comparison tool. *Bioinformatics*, 21(16), 3422–3423.

Carver, T., Thomson, N., Bleasby, A., Berriman, M., and Parkhill, J. (2009). DNAPlotter: Circular and linear interactive genome visualization. *Bioinformatics*, 25(1), 119–120.

- 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(5), 494–501.
- Centers for Disease Control and Prevention (CDC). *What Exactly is Antibiotic Resistance?* Centers for Disease Control and Prevention. Accessed 05/10/2022.
<https://www.cdc.gov/drugresistance/about.html>
- Centers for Disease Control and Prevention, Office of Infectious Disease (CDC). (2019). *Antibiotic resistance threats in the United States, 2019*.
- Chaudhury, A., Nath, G., Tikoo, A., and Sanyal, S. C. (1999). Enteropathogenicity and Antimicrobial Susceptibility of New Escherichia Spp. *Journal of Diarrhoeal Diseases Research*, 17(2), 85–87.
- Clark, D. P., and Pazdernik, N. J. (2013). Chapter e25—Bacterial Genetics. In D. P. Clark and N. J. Pazdernik (Eds.), *Molecular Biology (Second Edition)* (pp. e641–e646). Academic Press.
- Collin, F., and Maxwell, A. (2019). The Microbial Toxin Microcin B17: Prospects for the Development of New Antibacterial Agents. *Journal of Molecular Biology*, 431(18), 3400–3426.
- Collin, F., Thompson, R. E., Jolliffe, K. A., Payne, R. J., and Maxwell, A. (2013). Fragments of the Bacterial Toxin Microcin B17 as Gyrase Poisons. *PLoS ONE*, 8(4), e61459.
- Connell, N., Han, Z., Moreno, F., & Kolter, R. (1987). An E. coli promoter induced by the cessation of growth. *Molecular microbiology*, 1(2), 195–201.
- Connell, S. R., Tracz, D. M., Nierhaus, K. H., and Taylor, D. E. (2003). Ribosomal Protection Proteins and Their Mechanism of Tetracycline Resistance. *Antimicrobial Agents and Chemotherapy*, 47(12), 3675–3681.

- Cotter, P. D., Hill, C., and Ross, R. P. (2005). Bacteriocins: developing innate immunity for food. *Nature reviews. Microbiology*, 3(10), 777–788
- Couturier, M., Bex, F., Bergquist, P. L., and Maas, W. K. (1988). Identification and Classification of Bacterial Plasmids. *Microbiological reviews*, 52, 21.
- Dabour, R., Meirson, T., and Samson, A. O. (2016). Global antibiotic resistance is mostly periodic. *Journal of Global Antimicrobial Resistance*, 7, 132–134.
- Datta, N., and Hedges, R. W. (1972). Host Ranges of R Factors. *Journal of General Microbiology*, 70(3), 453–460.
- Datta, N., and Hedges, R. W. (1973). R Factors of Compatibility Group A. *Journal of General Microbiology*, 74(2), 335–336.
- de Jong, A., van Heel, A. J., and Kuipers, O. P. (2011). Genome Exploitation and Bioinformatics Tools. In D. Drider and S. Rebuffat (Eds.), *Prokaryotic Antimicrobial Peptides* (pp. 75–80). Springer New York.
- de la Haba, R. R., Sánchez-Porro, C., Márquez, M. C., and Ventosa, A. (2010). Taxonomic study of the genus *Salinicola*: Transfer of *Halomonas salaria* and *Chromohalobacter salarius* to the genus *Salinicola* as *Salinicola salarius* comb. nov. and *Salinicola halophilus* nom. nov., respectively. *International Journal of Systematic and Evolutionary Microbiology*, 60(4), 963–971.
- De Vuyst, L., and Vandamme, E. J. (1994). Lactic acid bacteria and bacteriocins: their practical importance. In De Vuyst, L. and Vandamme, E.J. (Eds.). *Bacteriocins of Lactic Acid Bacteria: Microbiology, Genetics and Applications*. Blackie Academic & Professional, London, pp. 1-11.
- del Solar, G., Giraldo, R., Ruiz-Echevarría, M. J., Espinosa, M., and Díaz-Orejas, R. (1998). Replication and Control of Circular Bacterial Plasmids. *Microbiology and Molecular Biology Reviews*, 62(2), 434–464.

- Delves-Broughton, J., Blackburn, P., Evans, R. J., and Hugenholtz, J. (1996). Applications of the bacteriocin, nisin. *Antonie van Leeuwenhoek*, *69*(2), 193–202.
- Demain, A. L. (2009). Antibiotics: Natural products essential to human health. *Medicinal Research Reviews*, *29*(6), 821–842.
- Dorrell, N., Gyselman, V. G., Foyne, S., Li, S.-R., and Wren, B. W. (1996). Improved Efficiency of Inverse PCR Mutagenesis. *BioTechniques*, *21*(4), 604–608.
- Duquesne, S., Destoumieux-Garzon, D., Peduzzi, J., and Rebuffat, S. (2007). Microcins, gene-encoded antibacterial peptides from enterobacteria. *Natural Product Reports*, *24*(4), 708.
- Edwards, P. R., and Ewing, W. H. (1972). *Identification of enterobacteriaceae* (3rd ed.). pp. 109-110. Burgess Publishing Company.
- EFSA Panel on Food Additives and Nutrient Sources added to Food (ANS) F., Younes, M., Aggett, P., Aguilar, F., Crebelli, R., Dusemund, B., Filipič, M., Frutos, M. J., Galtier, P., Gundert-Remy, U., Kuhnle, G. G., Lambré, C., Leblanc, J.-C., Lillegaard, I. T., Moldeus, P., Mortensen, A., Oskarsson, A., Stankovic, I., Waalkens-Berendsen, I., Woutersen, R. A., Wright, M., Herman, L., Tobback, P., Pizzo, F., Smeraldi, C., Tard, A., Papaioannou, A., Gott, D. (2017). Safety of nisin (E 234) as a food additive in the light of new toxicological data and the proposed extension of use. [EFSA://doi.org/10.2903/j.efsa.2017.5063](https://doi.org/10.2903/j.efsa.2017.5063)
- Evans, F. F., Egan, S., and Kjelleberg, S. (2008). Ecology of type II secretion in marine gammaproteobacteria. *Environmental Microbiology*, *10*(5), 1101–1107.
- Fedorec, A. J. H., Ozdemir, T., Doshi, A., Ho, Y. K., Rosa, L., Rutter, J., Velazquez, O., Pinheiro, V. B., Danino, T., & Barnes, C. P. (2019). Two New Plasmid Post-segregational Killing Mechanisms for the Implementation of Synthetic Gene Networks in Escherichia coli. *iScience*, *14*, 323–334.

- Felsenstein, J. (1981). Evolutionary trees from DNA sequences: A maximum likelihood approach. *Journal of Molecular Evolution*, 17(6), 368–376.
- Fernández-García, L., Blasco, L., Lopez, M., Bou, G., García-Contreras, R., Wood, T., & Tomas, M. (2016). Toxin-Antitoxin Systems in Clinical Pathogens. *Toxins*, 8(7), 227
- Fernandez-Lopez, R., de Toro, M., Moncalian, G., Garcillan-Barcia, M. P., and de la Cruz, F. (2016). Comparative genomics of the conjugation region of F-like plasmids: Five shades of F. *Frontiers in Molecular Biosciences*, 3, 1-14.
- Ferreira da Silva, M., Vaz-Moreira, I., Gonzalez-Pajuelo, M., Nunes, O. C., and Manaia, C. M. (2007). Antimicrobial resistance patterns in Enterobacteriaceae isolated from an urban wastewater treatment plant. *FEMS Microbiology Ecology*, 60(1), 166–176.
- Fletcher, C. (1984). First clinical use of penicillin. *British medical journal*, 289(6460), 1721–1723.
- Fredericq, P., and Levine, M. (1947). Antibiotic Interrelationships among the Enteric Group of Bacteria. *Journal of Bacteriology*, 54(6), 785–792.
- Gadebusch, H. H., and Basch, H. (1968). In vitro antimicrobial activity of dimethylsulfoxide. *Applied microbiology*, 6(12), 3–4.
- Garcillán-Barcia, M. P., Alvarado, A., and Cruz, F. de la. (2011). Identification of bacterial plasmids based on mobility and plasmid population biology. *FEMS Microbiology Reviews*, 35(5), 936–956.
- Garcillán-Barcia, M. P., Francia, M. V., and de La Cruz, F. (2009). The diversity of conjugative relaxases and its application in plasmid classification. *FEMS Microbiology Reviews*, 33(3), 657–687.
- Garcillán-Barcia, M. P., Redondo-Salvo, S., Vielva, L., and de la Cruz, F. (2020). MOBscan: Automated Annotation of MOB Relaxases. In F. de la Cruz (Ed.), *Horizontal Gene Transfer: Methods and Protocols* (pp. 295–308). Springer US.

- Garrido, M. C., Herrero, M., Kolter, R., and Moreno, F. (1988). The export of the DNA replication inhibitor Microcin B17 provides immunity for the host cell. *The EMBO Journal*, 7(6), 1853–1862.
- Garrity, G. M., Bell, J. A., and Lilburn, T. (2005). Gammaproteobacteria class. Nov. In D. J. Brenner, N. R. Krieg, J. T. Staley, and G. M. Garrity (Eds.), *Bergey's Manual of Systematic Bacteriology* (2nd ed., Vol. 2). Springer.
- Genilloud, O., Moreno, F., and Kolter, R. (1989). DNA sequence, products, and transcriptional pattern of the genes involved in production of the DNA replication inhibitor microcin B17. *Journal of Bacteriology*, 171(2), 1126–1135.
- Gerber, J. S., Ross, R. K., Bryan, M., Localio, A. R., Szymczak, J. E., Wasserman, R., Barkman, D., Odeniyi, F., Conaboy, K., Bell, L., Zaoutis, T. E., and Fiks, A. G. (2017). Association of Broad- vs Narrow-Spectrum Antibiotics with Treatment Failure, Adverse Events, and Quality of Life in Children with Acute Respiratory Tract Infections. *Journal of the American Medical Association* 318(23), 2325–2336.
- Gerdes, K., Møller-Jensen, J., and Jensen, R. B. (2000). Plasmid and chromosome partitioning: Surprises from phylogeny: Phylogeny of partitioning ATPases. *Molecular Microbiology*, 37(3), 455–466.
- Ghilarov, D., Inaba-Inoue, S., Stepien, P., Qu, F., Michalczyk, E., Pakosz, Z., Nomura, N., Ogasawara, S., Walker, G. C., Rebuffat, S., Iwata, S., Heddle, J. G., and Beis, K. (2021). Molecular mechanism of SbmA, a promiscuous transporter exploited by antimicrobial peptides. *Science Advances*, 7(37), 1-10.
- Ghilarov, D., Serebryakova, M., Shkundina, I., and Severinov, K. (2011). A Major Portion of DNA Gyrase Inhibitor Microcin B17 Undergoes an *N, O* -Peptidyl Shift during Synthesis. *Journal of Biological Chemistry*, 286(30), 26308–26318.

- Ghilarov, D., Stevenson, C. E. M., Travin, D. Y., Piskunova, J., Serebryakova, M., Maxwell, A., Lawson, D. M., and Severinov, K. (2019). Architecture of Microcin B17 Synthetase: An Octameric Protein Complex Converting a Ribosomally Synthesized Peptide into a DNA Gyrase Poison. *Molecular Cell*, 73(4), 749-762.e5.
- Gillor, O., Nigro, L., and Riley, M. (2005). Genetically Engineered Bacteriocins and their Potential as the Next Generation of Antimicrobials. *Current Pharmaceutical Design*, 11(8), 1067–1075.
- Gomez-Escribano, J. P., Song, L., Bibb, M. J., and Challis, G. L. (2012). Posttranslational β -methylation and macrolactamidation in the biosynthesis of the bottromycin complex of ribosomal peptide antibiotics. *Chemical Science*, 3(12), 3522.
- Graham, T. G. W., Wang, X., Song, D., Etson, C. M., van Oijen, A. M., Rudner, D. Z., and Loparo, J. J. (2014). ParB spreading requires DNA bridging. *Genes and Development*, 28(11), 1228–1238.
- Gratia, A. (1925). Sur un remarquable exemple d'antagonisme entre deux souches de colibacille. *Comptes Rendus des Seances de la Societe de Biologie et de Ses Filiales.*, 93, 1040–1043.
- Güllüce, M., Karadayı, M., and Barış, Ö. (2013). *Bacteriocins: Promising Natural Antimicrobials*. In: Méndez-Vilas, A. (Ed.) *Microbial pathogens and strategies for combating them: science, technology and education* (pp. 1016-1027). Formatex
- Gunjal, V. B., Thakare, R., Chopra, S., and Reddy, D. S. (2020). Teixobactin: A Paving Stone toward a New Class of Antibiotics? *Journal of Medicinal Chemistry*, 63(21), 12171–12195.
- Guyenet, C., and de la Cruz, F. (2011). Plasmid segregation without partition. *Mobile Genetic Elements*, 1(3), 236–241.

- Hanahan, D. (1983). Studies on transformation of *Escherichia coli* with plasmids. *Journal of Molecular Biology*, 166(4), 557–580.
- Heddle, J. G., Blance, S. J., Zamble, D. B., Hollfelder, F., Miller, D. A., Wentzell, L. M., Walsh, C. T., and Maxwell, A. (2001). The antibiotic microcin B17 is a DNA gyrase poison: Characterisation of the mode of inhibition. *Journal of Molecular Biology*, 307(5), 1223–1234.
- 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*, 99(5), 2696–2701.
- Ilhan, J., Kupczok, A., Woehle, C., Wein, T., Hülter, N. F., Rosenstiel, P., Landan, G., Mizrahi, I., and Dagan, T. (2018). Segregational Drift and the Interplay between Plasmid Copy Number and Evolvability. *Molecular Biology and Evolution*, 36(3), 472–486.
- Jabrane, A., Sabri, A., Compère, P., Jacques, P., Vandenberghe, I., Van Beeumen, J., and Thonart, P. (2002). Characterization of Serracin P, a Phage-Tail-Like Bacteriocin, and Its Activity against *Erwinia amylovora*, the Fire Blight Pathogen. *Applied and Environmental Microbiology*, 68(11), 5704–5710.
- Jalal, A. S. B., and Le, T. B. K. (2020). Bacterial chromosome segregation by the ParABS system. *Open Biology*, 10(6), 200097.
- Janda, J. M., and Abbott, S. L. (2021). The Changing Face of the Family Enterobacteriaceae (Order: “Enterobacterales”): New Members, Taxonomic Issues, Geographic Expansion, and New Diseases and Disease Syndromes. *Clinical Microbiology Reviews*, 34(2), e00174-20.
- Johnson, T. J., Wannemuehler, Y. M., Johnson, S. J., Logue, C. M., White, D. G., Doetkott, C., and Nolan, L. K. (2007). Plasmid Replicon Typing of Commensal and Pathogenic

- Escherichia coli* Isolates. *Applied and Environmental Microbiology*, 73(6), 1976–1983.
- Jolley, K. A., Bray, J. E., and Maiden, M. C. J. (2018). Open-access bacterial population genomics: BIGSdb software, the PubMLST.org website and their applications. *Wellcome Open Research*, 3, 124.
- Jordi, B. J. A. M., Boutaga, K., van Heeswijk, C. M. E., van Knapen, F., and Lipman, L. J. A. (2001). Sensitivity of Shiga toxin-producing *Escherichia coli* (STEC) strains for colicins under different experimental conditions. *FEMS Microbiology Letters*, 204(2), 329–334.
- Kado, C. I. (2006). *Erwinia* and Related Genera. In M. Dworkin, S. Falkow, E. Rosenberg, K.-H. Schleifer, and E. Stackebrandt (Eds.), *The Prokaryotes* (pp. 443–450). Springer New York.
- Kapoor, G., Saigal, S., and Elongavan, A. (2017). Action and resistance mechanisms of antibiotics: A guide for clinicians. *Journal of Anaesthesiology, Clinical Pharmacology*, 33(3), 300–305.
- Kellner, R., Jung, G., Hörner, T., Zähler, H., Schnell, N., Entian, K.-D., and Götz, F. (1988). Gallidermin: A new lanthionine-containing polypeptide antibiotic. *European Journal of Biochemistry*, 177(1), 53–59.
- Kim, J. W., Bugata, V., Cortés-Cortés, G., Quevedo-Martínez, G., and Camps, M. (2020). Mechanisms of Theta Plasmid Replication in Enterobacteria and Implications for Adaptation to Its Host. *EcoSal Plus*, 9(1),
- Kirkwood, Z., Millar, B., Downey, D., and Moore, J. (2018). Antimicrobial effect of dimethyl sulfoxide and N, N-Dimethylformamide on *Mycobacterium abscessus*: Implications for antimicrobial susceptibility testing. *International Journal of Mycobacteriology*, 7(2), 134.

- Kobayashi, M., and Tchan, Y. T. (1973). Treatment of industrial waste solutions and production of useful by-products using a photosynthetic bacterial method. *Water Research*, 7(8), 1219–1224.
- Kondratyeva, K., Salmon-Divon, M., and Navon-Venezia, S. (2020). Meta-analysis of Pandemic *Escherichia coli* ST131 Plasmidome Proves Restricted Plasmid-clade Associations. *Scientific Reports*, 10, 36.
- Kumar, S., Stecher, G., Li, M., Knyaz, C., and Tamura, K. (2018). MEGA X: Molecular Evolutionary Genetics Analysis across Computing Platforms. *Molecular Biology and Evolution*, 35(6), 1547–1549.
- Kumar, V., Singh, B., van Belkum, M. J., Diep, D. B., Chikindas, M. L., Ermakov, A. M., and Tiwari, S. K. (2021). Halocins, natural antimicrobials of Archaea: Exotic or special or both? *Biotechnology Advances*, 53, 107834.
- Lade, H. S., Chitanand, M. P., Gyanaanath, G. and Kadam, T. A. (2006). Studies on some properties of bacteriocins produced by *Lactobacillus* species isolated from agro-based waste. *The Internet J. Microbiol.* 2(1), 1-4
- Lages, M. C. A., Beilharz, K., Morales Angeles, D., Veening, J.-W., and Scheffers, D.-J. (2013). The localization of key *Bacillus subtilis* penicillin binding proteins during cell growth is determined by substrate availability. *Environmental Microbiology*, 15(12), 3272–3281.
- Lavina, M., Pugsley, A., and Moreno, F. (1986). Identification, Mapping, Cloning and Characterization of a Gene (*sbmA*) Required for Microcin B17 Action on *Escherichia coli* K12. *Journal of General Microbiology*, 132, 1685–1693.
- Le, S. Q., and Gascuel, O. (2008). An Improved General Amino Acid Replacement Matrix. *Molecular Biology and Evolution*, 25(7), 1307–1320.

- Lee, D. J., Bingle, L. E., Heurlier, K., Pallen, M. J., Penn, C. W., Busby, S. J., and Hobman, J. L. (2009). Gene doctoring: A method for recombineering in laboratory and pathogenic *Escherichia coli* strains. *BMC Microbiology*, 9(1), 252.
- Lehwark, P., and Greiner, S. (2019). GB2sequin—A file converter preparing custom GenBank files for database submission. *Genomics*, 111(4), 759–761.
- Leikoski, N., Fewer, D. P., and Sivonen, K. (2009). Widespread Occurrence and Lateral Transfer of the Cyanobactin Biosynthesis Gene Cluster in Cyanobacteria. *Applied and Environmental Microbiology*, 75(3), 853–857.
- Lewis, K. (2013). Platforms for antibiotic discovery. *Nature Reviews Drug Discovery* 12, 371–387.
- Li, Y., and Rebuffat, S. (2020). The manifold roles of microbial ribosomal peptide-based natural products in physiology and ecology. *Journal of Biological Chemistry*, 295(1), 34–54.
- Lilly, J., and Camps, M. (2015). Mechanisms of Theta Plasmid Replication. *Microbiology Spectrum*, 3(1), PLAS-0029-2014.
- Liu, J. (1994). Microcin B17: Posttranslational modifications and their biological implications. *Proceedings of the National Academy of Sciences*, 91(11), 4618–4620.
- Liu, X., and Ferenci, T. (1998). Regulation of porin-mediated outer membrane permeability by nutrient limitation in *Escherichia coli*. *Journal of bacteriology*, 180(15), 3917–3922.
- Liu, B., Zheng, D., Jin, Q., Chen, L., and Yang, J. (2019). VFDB 2019: a comparative pathogenomic platform with an interactive web interface. *Nucleic acids research*, 47(D1), D687–D692.

- Loh, S., Cram, D., and Skurray, R. (1989). Nucleotide sequence of the leading region adjacent to the origin of transfer on plasmid F and its conservation among conjugative plasmids. *Molecular and General Genetics MGG*, 219(1–2), 177–186.
- MacGowan, A., and Macnaughton, E. (2017). Antibiotic resistance. *Medicine*, 45(10), 622–628.
- Madeira, F., Pearce, M., Tivey, A. R. N., Basutkar, P., Lee, J., Edbali, O., Madhusoodanan, N., Kolesnikov, A., and Lopez, R. (2022). Search and sequence analysis tools services from EMBL-EBI in 2022. *Nucleic Acids Research*, 50(W1), W276–W279.
- Madigan, M. T., Martinko, J. M., and Brock, T. D. (2006). *Brock biology of microorganisms* (11th ed). Pp. 494-496. Pearson Prentice Hall.
- Madison, L. L., Vivas, E. I., Li, Y.-M., Walsh, C. T., and Kolter, R. (1997). The leader peptide is essential for the post-translational modification of the DNA-gyrase inhibitor microcin B17. *Molecular Microbiology*, 23(1), 161–168.
- Maksimov, M. O., and Link, A. J. (2014). Prospecting genomes for lasso peptides. *Journal of Industrial Microbiology and Biotechnology*, 41(2), 333–344.
- Manwaring, N. P., Skurray, R. A., and Firth, N. (1999). Nucleotide Sequence of the F Plasmid Leading Region. *Plasmid*, 41(3), 219–225.
- Marchler-Bauer, A., Bo, Y., Han, L., He, J., Lanczycki, C. J., Lu, S., Chitsaz, F., Derbyshire, M. K., Geer, R. C., Gonzales, N. R., Gwadz, M., Hurwitz, D. I., Lu, F., Marchler, G. H., Song, J. S., Thanki, N., Wang, Z., Yamashita, R. A., Zhang, D., ... Bryant, S. H. (2017). CDD/SPARCLE: Functional classification of proteins via subfamily domain architectures. *Nucleic Acids Research*, 45(D1), D200–D203.
- Märki, F., Hänni, E., Fredenhagen, A., and van Oostrum, J. (1991). Mode of action of the lanthionine-containing peptide antibiotics duramycin, duramycin B and C, and

- cinnamycin as indirect inhibitors of phospholipase A2. *Biochemical Pharmacology*, 42(10), 2027–2035.
- Marković, K. G., Grujović, M. Ž., Koraćević, M. G., Nikodijević, D. D., Milutinović, M. G., Semedo-Lemsaddek, T., and Djilas, M. D. (2022). Colicins and Microcins Produced by Enterobacteriaceae: Characterization, Mode of Action, and Putative Applications. *International Journal of Environmental Research and Public Health*, 19(18), 11825.
- Masuda, H., Tan, Q., Awano, N., Wu, K.-P., and Inouye, M. (2012). YeeU enhances the bundling of cytoskeletal polymers of MreB and FtsZ, antagonizing the CbtA (YeeV) toxicity in *Escherichia coli*: YeeU enhances bundling of MreB and FtsZ filaments. *Molecular Microbiology*, 84(5), 979–989.
- Mayo, O., Hernández-Chico, C., and Moreno, F. (1988). Microcin B17, a novel tool for preparation of maxicells: Identification of polypeptides encoded by the IncFII minireplicon pMccB17. *Journal of Bacteriology*, 170(5), 2414–2417.
- McGuffie, M. J., and Barrick, J. E. (2021). pLannotate: Engineered plasmid annotation. *Nucleic Acids Research*, 49(W1), W516–W522.
- McVicker, G., Hollingshead, S., Pilla, G., and Tang, C. M. (2019). Maintenance of the virulence plasmid in *Shigella flexneri* is influenced by Lon and two functional partitioning systems. *Molecular Microbiology*, 111(5), 1355–1366.
- Meade, Slattery, and Garvey. (2020). Bacteriocins, Potent Antimicrobial Peptides and the Fight against Multi Drug Resistant Species: Resistance Is Futile? *Antibiotics*, 9(1), 32.
- Melander, R. J., Zurawski, D. V., and Melander, C. (2017). Narrow-spectrum antibacterial agents. *Medicinal Chemistry Communications*, 9(1), 12–21.
- Messi, P., Bondi, M., Sabia, C., Buttini, R. and Manicardi, G. (2001). Detection and preliminary characterisation of a bacteriocin (plantaricin 35d) produced by a *Lactobacillus plantarum* strain. *Int. J. Food. Microbiol.* 64:193-198.

- Meteliev, M., Serebryakova, M., Ghilarov, D., Zhao, Y., and Severinov, K. (2013). Structure of Microcin B-Like Compounds Produced by *Pseudomonas syringae* and Species Specificity of Their Antibacterial Action. *Journal of Bacteriology*, 195(18), 4129–4137.
- Millan, J. L. S., Hernandez-Chico, C., Pereda, P., and Moreno, F. (1985). Cloning and Mapping of the Genetic Determinants for Microcin B17 Production and Immunity. *Journal of Bacteriology* ., 163, 7.
- Murray, C. J., Ikuta, K. S., Sharara, F., Swetschinski, L., Robles Aguilar, G., Gray, A., Han, C., Bisignano, C., Rao, P., Wool, E., Johnson, S. C., Browne, A. J., Chipeta, M. G., Fell, F., Hackett, S., Haines-Woodhouse, G., Kashef Hamadani, B. H., Kumaran, E. A. P., McManigal, B., Agarwal, R., Akech, S., Albertson, S., Amuasi, J., Andrews, J., Aravkin, A., Ashley, E., Bailey, F., Baker, S., Basnyat, B., Bekker, A., Bender, R., Bethou, A., Bielicki, J., Boonkasidecha, S., Bukosia, J., Carvalheiro, C., Castañeda-Orjuela, C., Chansamouth, V., Chaurasia, S., Chiurchiù, S., Chowdhury, F., Cook, A. J., Cooper, B., Cressey, T. R., Criollo-Mora, E., Cunningham, M., Darboe, S., Day, N. P.J., De Luca, M., Dokova, K., Dramowski, A., Dunachie, S. J., Eckmanns, T., Eibach, D., Emami, A., Feasey, N., Fisher-Pearson, N., Forrest, K., Garrett, D., Gastmeier, P., Giref, A. Z., Greer, R. C., Gupta, V., Haller, S., Haselbeck, A., Hay, S. I., Holm, M., Hopkins, S., Iregbu, K. C., Jacobs, J., Jarovsky, D., Javanmardi, F., Khorana, M., Kisson, N., Kobeissi, E., Kostyanev, T., Krapp, F., Krumkamp, R., Kumar, A., Kyu, H. H., Lim, C., Limmathurotsakul, D., Loftus, M. J., Lunn, M., Ma, J., Mturi, N., Munera-Huertas, T., Musicha, P., Mussi-Pinhata, M. M., Nakamura, T., Nanavati, R., Nangia, S., Newton, P., Ngoun, C., Novotney, A., Nwakanma, D., Obiero, C. W., Olivas-Martinez, A., Olliaro, P., Ooko, E., Ortiz-Brizuela, E., Peleg, A. Y., Perrone, C., Plakkal, N., Ponce-de-Leon, A., Raad, M., Ramdin, T., Riddell,

- A., Roberts, T., Robotham, J. V., Roca, A., Rudd, K. E., Russell, N., Schnall, J., Scott, J. A. G., Shivamallappa, M., Sifuentes-Osornio, J., Steenkeste, N., Stewardson, A. J., Stoeva, T., Tasak, N., Thaiprakong, A., Thwaites, G., Turner, C., Turner, P., van Doorn, H. R., Velaphi, S., Vongpradith, A., Vu, H., Walsh, T., Waner, S., Wangrangsimakul, T., Wozniak, T., Zheng, P., Sartorius, B., Lopez, A. D., Stergachis, A., Moore, C., Dolecek, C., Naghavi, M. (2022). Global burden of bacterial antimicrobial resistance in 2019: A systematic analysis. *The Lancet*, 399(10325), 629–655.
- Mutai, W. C., Waiyaki, P. G., Kariuki, S., and Muigai, A. W. T. (2019). Plasmid profiling and incompatibility grouping of multidrug resistant *Salmonella enterica* serovar Typhi isolates in Nairobi, Kenya. *BMC Research Notes*, 12(1), 422.
- Needleman, S. B., and Wunsch, C. D. (1970). A general method applicable to the search for similarities in the amino acid sequence of two proteins. *Journal of Molecular Biology*, 48(3), 443–453.
- Norman, A., Hansen, L. H., and Sørensen, S. J. (2009). Conjugative plasmids: Vessels of the communal gene pool. *Philosophical Transactions of the Royal Society B: Biological Sciences*, 364(1527), 2275–2289.
- Novick, R. P. (1987). Plasmid Incompatibility. *Microbiological Reviews*, 51(4), 381–395.
- O'Connor, E., and Shand, R. (2002). Halocins and sulfolobicins: The emerging story of archaeal protein and peptide antibiotics. *Journal of Industrial Microbiology and Biotechnology*, 28(1), 23–31.
- O'Neill J. (2016). Tackling drug-resistant infections globally: Final report and recommendations. Review on Antimicrobial Resistance, chaired by Jim O'Neill. HM Government and Wellcome Trust, London.

- Orlek, A., Stoesser, N., Anjum, M. F., Doumith, M., Ellington, M. J., Peto, T., Crook, D., Woodford, N., Walker, A. S., Phan, H., and Sheppard, A. E. (2017). Plasmid Classification in an Era of Whole-Genome Sequencing: Application in Studies of Antibiotic Resistance Epidemiology. *Frontiers in Microbiology*, 8, 1-10.
- Osborn, A. M., Pickup, R. W., and Saunders, J. R. (2000). Mosaic plasmids and mosaic replicons: Evolutionary lessons from the analysis of genetic diversity in IncFII-related replicons. *Microbiology (Reading)*, 146 (Pt 9):2267-2275
- Pag, U., and Sahl, H. G. (2002). Multiple activities in lantibiotics--models for the design of novel antibiotics? *Curr Pharm Des.*, 8(9):815-33.
- Parada, J. L., Caron, C. R., Medeiros, A. B. P., and Soccol, C. R. (2007). Bacteriocins from Lactic Acid Bacteria: Purification, properties and use as Biopreservatives. *Brazilian Archives of Biology and Technology* 50 (3):521-542
- Parks, W. M., Bottrill, A. R., Pierrat, O. A., Durrant, M. C., and Maxwell, A. (2007). The action of the bacterial toxin, microcin B17, on DNA gyrase. *Biochimie*, 89(4), 500–507.
- Pavlova, O. A., and Severinov, K. V. (2006). Posttranslationally modified microcins. *Russian Journal of Genetics*, 42(12), 1380–1389.
- Penesyan, A., Gillings, M., and Paulsen, I. T. (2015). Antibiotic discovery: combatting bacterial resistance in cells and in biofilm communities. *Molecules (Basel, Switzerland)*, 20(4), 5286–5298.
- Pierrat, O. A., and Maxwell, A. (2003). The Action of the Bacterial Toxin Microcin B17: Insight into the cleavage-religation reaction of DNA gyrase. *Journal of Biological Chemistry*, 278(37), 35016–35023.

- Pióro, M., and Jakimowicz, D. (2020). Chromosome Segregation Proteins as Coordinators of Cell Cycle in Response to Environmental Conditions. *Frontiers in Microbiology*, *11*, 588.
- Pons, A.-M., Delalande, F., Duarte, M., Benoit, S., Lanneluc, I., Sablé, S., Van Dorselaer, A., and Cottenceau, G. (2004). Genetic Analysis and Complete Primary Structure of Microcin L. *Antimicrobial Agents and Chemotherapy*, *48*(2), 505–513.
- Popowska, M., and Krawczyk-Balska, A. (2013). Broad-host-range IncP-1 plasmids and their resistance potential. *Frontiers in Microbiology*, *4*, 44.
- Portrait, V., Gendron-Gaillard, S., Cottenceau, G., and Pons, A. M. (1999). Inhibition of pathogenic *Salmonella enteritidis* growth mediated by *Escherichia coli* microcin J25 producing strains. *Canadian Journal of Microbiology*, *45*(12), 988–994.
- Puangsee, J., Prathan, R., Srisanga, S., Angkittitrakul, S., and Chuanchuen, R. (2022). Plasmid profile analysis of *Escherichia coli* and *Salmonella enterica* isolated from pigs, pork, and humans. *Epidemiology and Infection*, *150*, e110.
- Rafii, F., Sutherland, J. B., and Cerniglia, C. E. (2008). Effects of treatment with antimicrobial agents on the human colonic microflora. *Therapeutics and Clinical Risk Management*, *4*(6), 1343–1358.
- Rebuffat, S. (2013). Microcins. In *Handbook of Biologically Active Peptides* (pp. 129–137). Elsevier.
- Reygaert, W. C. (2018). An overview of the antimicrobial resistance mechanisms of bacteria. *AIMS Microbiology*, *4*(3), 482–501.
- Ronald, A. (2003). The etiology of urinary tract infection: Traditional and emerging pathogens. *Disease-a-Month*, *49*(2), 71–82.
- Rodríguez-Sáinz, M. C., Hernández-Chico, C., & Moreno, F. (1990). Molecular characterization of pmbA, an *Escherichia coli* chromosomal gene required for the

- production of the antibiotic peptide MccB17. *Molecular microbiology*, 4(11), 1921–1932.
- Ruiz-Masó, J. A., Machón, C., Bordanaba-Ruiseco, L., Espinosa, M., Coll, M., and Del Solar, G. (2015). Plasmid Rolling-Circle Replication. *Microbiology Spectrum*, 3(1), 3.1.16.
- Rutherford, K., Parkhill, J., Crook, J., Horsnell, T., Rice, P., Rajandream, M.-A., and Barrell, B. (2000). Artemis: Sequence visualization and annotation. *Bioinformatics*, 16(10), 944–945.
- Ryan, M. P., Meaney, W. J., Ross, R. P., and Hill, C. (1998). Evaluation of Lacticin 3147 and a Teat Seal Containing This Bacteriocin for Inhibition of Mastitis Pathogens. *Applied and Environmental Microbiology*, 64(6), 2287–2290.
- Sable, S., Pons, A.-M., Gendron-Gaillard, S., and Cottenceau, G. (2000). Antibacterial Activity Evaluation of Microcin J25 against Diarrheagenic *Escherichia coli*. *Applied and Environmental Microbiology*, 66(10), 4595.
- Saha, C. K., Sanches Pires, R., Brolin, H., Delannoy, M., and Atkinson, G. C. (2021). FlaGs and webFlaGs: Discovering novel biology through the analysis of gene neighbourhood conservation. *Bioinformatics*, 37(9), 1312–1314.
- Saha, M., and Sarkar, A. (2021). Review on Multiple Facets of Drug Resistance: A Rising Challenge in the 21st Century. *Journal of Xenobiotics*, 11(4), 197–214.
- Sakthivel, N., and Mew, T. W. (1991). Efficacy of bacteriocinogenic strains of *Xanthomonas oryzae* pv. *Oryzae* on the incidence of bacterial blight disease of rice (*Oryza sativa* L.). *Canadian Journal of Microbiology*, 37(10), 764–768.
- Salje, J., Gayathri, P., and Löwe, J. (2010). The ParMRC system: Molecular mechanisms of plasmid segregation by actin-like filaments. *Nature Reviews Microbiology*, 8(10), 683–692.

- Scherzinger, E., Haring, V., Lurz, R., and Otto, S. (1991). Plasmid RSF1010 DNA replication in vitro promoted by purified RSF1010 RepA, RepB and RepC proteins. *Nucleic Acids Research*, 19(6), 1203–1211.
- Schumacher, M. A. (2012). Bacterial plasmid partition machinery: A minimalist approach to survival. *Current Opinion in Structural Biology*, 22(1), 72–79.
- Seemann T. (2014). Prokka: rapid prokaryotic genome annotation. *Bioinformatics (Oxford, England)*, 30(14), 2068–2069.
- Sengupta, M., and Austin, S. (2011). Prevalence and Significance of Plasmid Maintenance Functions in the Virulence Plasmids of Pathogenic Bacteria. *Infection and Immunity*, 79(7), 2502–2509.
- Severina, E., Severin, A., and Tomasz, A. (1998). Antibacterial efficacy of nisin against multidrug-resistant Gram- positive pathogens. *Journal of Antimicrobial Chemotherapy*, 41(3), 341–347.
- Severinov, K., Semenova, E., Kazakov, A., Kazakov, T., and Gelfand, M. S. (2007). Low-molecular-weight post-translationally modified microcins. *Molecular Microbiology*, 65(6), 1380–1394.
- Shimamura, H., Gouda, H., Nagai, K., Hirose, T., Ichioka, M., Furuya, Y., Kobayashi, Y., Hirono, S., Sunazuka, T., and Ōmura, S. (2009). Structure Determination and Total Synthesis of Bottromycin A2: A Potent Antibiotic against MRSA and VRE. *Angewandte Chemie International Edition*, 48(5), 914–917.
- Shintani, M., Sanchez, Z. K., and Kimbara, K. (2015). Genomics of microbial plasmids: Classification and identification based on replication and transfer systems and host taxonomy. *Frontiers in Microbiology*, 6, 1-16.

- Shkundina, I., Serebryakova, M., and Severinov, K. (2014). The C-Terminal Part of Microcin B Is Crucial for DNA Gyrase Inhibition and Antibiotic Uptake by Sensitive Cells. *Journal of Bacteriology*, *196*(9), 1759–1767.
- Silver, L. L. (2011). Challenges of Antibacterial Discovery. *Clinical Microbiology Reviews*, *24*(1), 71–109.
- Sinha Roy, R., Belshaw, P. J., and Walsh, C. T. (1998). Mutational Analysis of Posttranslational Heterocycle Biosynthesis in the Gyrase Inhibitor Microcin B17: Distance Dependence from Propeptide and Tolerance for Substitution in a GSCG Cyclizable Sequence. *Biochemistry*, *37*(12), 4125–4136.
- Sinha Roy, R., Kelleher, N. L., Milne, J. C., and Walsh, T. (1999). *In vivo processing and antibiotic activity of microcin B17 analogs with varying ring content and altered bisheterocyclic sites. Chem Biol*, *6*, 305–318.
- Smillie, C., Garcillan-Barcia, M. P., Francia, M. V., Rocha, E. P. C., and de la Cruz, F. (2010). Mobility of Plasmids. *Microbiology and Molecular Biology Reviews*, *74*(3), 434–452.
- Söding, J., Biegert, A., and Lupas, A. N. (2005). The HHpred interactive server for protein homology detection and structure prediction. *Nucleic Acids Research*, *33*, W244–W248.
- Telhig, S., Ben Said, L., Zirah, S., Fliss, I., and Rebuffat, S. (2020). Bacteriocins to Thwart Bacterial Resistance in Gram Negative Bacteria. *Frontiers in Microbiology*, *11*(586433), 1-25.
- Thomas, C. M., and Frost, L. S. (2014). Plasmid Genomes, Introduction to. In E. Bell (Ed.), *Molecular Life Sciences* (pp. 1–20). Springer New York.
- Thomas, C. M., and Summers, D. (2020). Bacterial Plasmids. *Encyclopedia of Life Sciences*, *1*(2), 240–250.

- Toth, I. K., Bell, K. S., Holeva, M. C., and Birch, P. R. J. (2003). Soft rot erwiniae: From genes to genomes. *Molecular Plant Pathology*, 4(1), 17–30.
- Tsirigos, K. D., Peters, C., Shu, N., Käll, L., and Elofsson, A. (2015). The TOPCONS web server for consensus prediction of membrane protein topology and signal peptides. *Nucleic Acids Research*, 43(W1), W401–W407.
- Varani, A. M., Siguier, P., Gourbeyre, E., Charneau, V., and Chandler, M. (2011). ISsaga is an ensemble of web-based methods for high throughput identification and semi-automatic annotation of insertion sequences in prokaryotic genomes. *Genome biology*, 12(3), R30.
- Vecchiarelli, A. G., Hwang, L. C., and Mizuuchi, K. (2013). Cell-free study of F plasmid partition provides evidence for cargo transport by a diffusion-ratchet mechanism. *Proceedings of the National Academy of Sciences*, 110(15), E1390–E1397.
- Velappan, N., Sblattero, D., Chasteen, L., Pavlik, P., and Bradbury, A. R. M. (2007). Plasmid incompatibility: More compatible than previously thought? *Protein Engineering Design and Selection*, 20(7), 309–313.
- Villa, L., García-Fernández, A., Fortini, D., and Carattoli, A. (2010). Replicon sequence typing of IncF plasmids carrying virulence and resistance determinants. *Journal of Antimicrobial Chemotherapy*, 65(12), 2518–2529.
- Virolle, C., Goldlust, K., Djermoun, S., Bigot, S., and Lesterlin, C. (2020). Plasmid Transfer by Conjugation in Gram-Negative Bacteria: From the Cellular to the Community Level. *Genes*, 11(11), 1239.
- Wadhvani, T., Desai, K., Patel, D., Lawani, D., Bahaley, P., Joshi, P., and Kothari, V. (2009). Effect of various solvents on bacterial growth in context of determining MIC of various antimicrobials. *The Internet Journal of Microbiology*, 7(1).

- Waksman, S. A. (1947). What is an Antibiotic or an Antibiotic Substance? *Mycologia*, 39(5), 565–569.
- Walker, D. C., Georgiou, T., Pommer, A. J., Walker, D., Moore, G. R., Kleanthous, C., and James, R. (2002). Mutagenic scan of the H-N-H motif of colicin E9: Implications for the mechanistic enzymology of colicins, homing enzymes and apoptotic endonucleases. *Nucleic Acids Research*, 30(14), 3225–3234.
- Wang, X., Lord, D. M., Cheng, H.-Y., Osbourne, D. O., Hong, S. H., Sanchez-Torres, V., Quiroga, C., Zheng, K., Herrmann, T., Peti, W., Benedik, M. J., Page, R., and Wood, T. K. (2012). A new type V toxin-antitoxin system where mRNA for toxin GhoT is cleaved by antitoxin GhoS. *Nature Chemical Biology*, 8(10), 855–861.
- Watve, M. M., Dahanukar, N., and Watve, M. G. (2010). Sociobiological control of plasmid copy number in bacteria. *PloS one*, 5(2), e9328
- Willey, J. M., Sherwood, L., Woolverton, C. J., and Prescott, L. M. (2008). *Prescott, Harley, and Klein's microbiology* (7th ed). Pp. 53. McGraw-Hill Higher Education.
- Williams, K. P., Gillespie, J. J., Sobral, B. W. S., Nordberg, E. K., Snyder, E. E., Shallom, J. M., and Dickerman, A. W. (2010). Phylogeny of Gammaproteobacteria. *Journal of Bacteriology*, 192(9), 2305–2314.
- Wooley, R. E., Gibbs, P. S., and Shotts, E. B. (1999). Inhibition of *Salmonella typhimurium* in the chicken intestinal tract by a transformed avirulent avian *Escherichia coli*. *Avian Diseases*, 43(2), 245–250.
- World Health Organization (WHO). (2017). WHO publishes list of bacteria for which new antibiotics are urgently needed. <https://www.who.int/news/item/27-02-2017-who-publishes-list-of-bacteria-for-which-new-antibiotics-are-urgently-needed> (accessed 14 December 2022)

- Yano, H., Rogers, L. M., Knox, M. G., Heuer, H., Smalla, K., Brown, C. J., and Top, E. M. (2013). Host range diversification within the IncP-1 plasmid group. *Microbiology*, *159*(Pt 11), 2303–2315.
- Yin, X., and Stotzky, G. (1997). Gene Transfer Among Bacteria in Natural Environments. In S. L. Neidleman and A. I. Laskin (Eds.), *Advances in Applied Microbiology* (Vol. 45, pp. 153–212). Academic Press.
- Yorgey, P., Davagnino, J., and Kolter, R. (1993). The maturation pathway of microcin B17, a peptide inhibitor of DNA gyrase. *Molecular Microbiology*, *9*(4), 897–905.
- Yorgey, P., Lee, J., Kordel, J., Vivas, E., Warner, P., Jebaratnam, D., and Kolter, R. (1994). Posttranslational modifications in microcin B17 define an additional class of DNA gyrase inhibitor. *Proceedings of the National Academy of Sciences*, *91*(10), 4519–4523.
- Zatyka, M., and Thomas, C. M. (1998). Control of genes for conjugative transfer of plasmids and other mobile elements. *FEMS Microbiology Reviews*, *21*(4), 291–319.
- Zavitz, K. H., and Marians, K. J. (1991). Dissecting the functional role of PriA protein-catalysed primosome assembly in *Escherichia coli* DNA replication. *Molecular Microbiology*, *5*(12), 2869–2873.
- Zhang, Z., Schwartz, S., Wagner, L., and Miller, W. (2000). A Greedy Algorithm for Aligning DNA Sequences. *Journal of Computational Biology*, *7*(1–2), 203–214.
- Zhong, Z., Helinski, D., and Toukdarian, A. (2005). Plasmid host-range: Restrictions to F replication in *Pseudomonas*. *Plasmid*, *54*(1), 48–56.
- Zhou, X., and Li, Y. (Eds.). (2015). Basic Biology of Oral Microbes. In *Atlas of Oral Microbiology* (pp. 1–14). Academic Press.
- Zihler, A., Le Blay, G., de Wouters, T., Lacroix, C., Braegger, C. P., Lehner, A., Tischler, P., Rattei, T., Hächler, H., and Stephan, R. (2009). *In vitro* inhibition activity of different

bacteriocin producing *Escherichia coli* against *Salmonella* strains isolated from clinical cases. *Letters in Applied Microbiology*, 49(1), 31–38.

Zimina, M., Babich, O., Prosekov, A., Sukhikh, S., Ivanova, S., Shevchenko, M., and Noskova, S. (2020). Overview of Global Trends in Classification, Methods of Preparation and Application of Bacteriocins. *Antibiotics*, 9(9), 553.