Investigating the Bacteriocin Library of *Lactobacillus plantarum* A-1

Patrick James Main

2014

Investigating the Bacteriocin Library *Lactobacillus plantarum* A-1

Presented in partial fulfilment of the requirements for the degree of:

Master of Science

in

Microbiology

at Massey University, Manuwatū Campus

Palmerston North

New Zealand

Patrick James Main

2014

Abstract

Bacteriocins are a highly diverse group of ribosomally synthesised, antimicrobial polypeptides produced by nearly all bacterial and archaeal species. Individual bacteriocins typically exhibit a narrow phylogenetic range of activity, but collectively inhibit a wide range of species through a variety of mechanisms. Glycocins are uncommon bacteriocins with rare, *S*-linked glycosidic bond;, currently there are only four characterised glycocins.

Preliminary characterisation of the bacteriocin ASM1 from *Lactobacillus plantarum* A-1 was reported by Hata *et al.* in 2010. ASM1 is structurally similar to GccF, a glycocin produced by *L. plantarum* KW30. Like GccF, ASM1 has two covalently linked N-acetylglucosamine moieties, one of which is attached through a rare *S*-glycosidic bond.

Due to its structural similarity, it was hypothesised that ASM1 would have similar inhibitory activity to GccF. Experiments showed that the two bacteriocins have almost identical inhibitory activity and both glycocins rely on their GlcNAc moieties to inhibit target cells.

The range of species inhibited by ASM1 was shown to be wider than previously thought. The inhibitory activity, however, varied considerably even between strains in a species.

The ASM1 gene cluster was established by sequencing and Southern hybridisation to be located on a 11,905 bp plasmid pA1_ASM1. An *in silico* analysis of the ASM1 gene cluster showed it to have the same operonic organisation similar as the GccF cluster, and overall DNA sequence identity of 75%.

A second active bacteriocin gene cluster was detected in the *L. plantarum* A-1 genome. A synthetic peptide, named ASM2, corresponding to this bacteriocin was partially characterised. ASM2 has 82% amino acid sequence identity to the recently identified bactofencin A produced by *L. salivarius* DPC6502.

A brief *in silico* analysis of proteins from the A-1 bacteriocin library and their orthologues provided some evolutionary context for the glycocins of *Lactobacillus* and gave hints as to the evolutionary history of the four characterised glycocins.

ASM1 is one of four characterised glycocins and this work has increased the overall knowledge of glycocins. Identification of a novel secondary bacteriocin in *L. plantarum* A-1 shows the complexity of bacteriocin systems and provides many avenues for future research regarding bacteria that produce multiple bacteriocins.

Acknowledgements

Special thanks to the following people

Dr. Mark Patchett
&
Dr. Gillian Norris

For their guidance, support and encouragement

Trevor Loo
&
Soyean Ahn

For their assistance in experiments and novel techniques

Ami Main

For her patience, love and tolerance

Andrew Kerr
&
Rebecca Smith

For reducing the monotony of long incubation times

The rest of X-Lab

For advice, camaraderie and further technical assistance

The Institute of Fundamental Sciences at Massey University

For providing the facilities and tools that made it all possible

Contents

List of Figures

List of Tables

List of Abbreviations

aa	Amino Acid
ABC	ATP Binding Cassette
ATCC	American Type Culture Collection
ATP	Adenosine Triphosphate
bac^-	Bacteriocin minus (mutants lacking expected bacteriocin production)
BLAST	Basic Local Alignment Search Tool
bp	Base Pairs
cd	Conserved Domain
Da	Dalton
dH_2O	Distilled Water
DNA	Deoxyribonucleic Acid
dNTP	Deoxyribose nucleotide triphosphate
DSM	Deutsche Sammlung von Mikroorganismen (German Collection of Microorganisms)
DUF	Domain of Unknown Function
et al.	*et alteri* (and others)
EtOH	Ethanol
g	standard gravity (9.81 ms^{-2})
GccF	Glycocin F
gDNA	Genomic DNA
GlcNAc	N-Acetyl Glucosamine
GRAS	Generally Regarded As Safe
GTase	Glycosyltransferase
HPK	Histidine Protein Kinase
HTH	Helix-Turn-Helix
IC_{50}	Concentration required to inhibit 50% of the population
LAB	Lactic Acid Bacteria
LB	Luria Broth
M	Molar (moles per litre)
min	minute
MUSCLE	Multiple Sequence Comparison by Log-Expectation
MRS	de Man, Rogosa and Sharpe medium
MRSA	Methicillin Resistant *Staphylococcus aureus*
NCBI	National Centre of Biotechnology Information
NZRM	New Zealand Reference Culture Collection, Medical Section
OD_{600}	Optical Density at 600 nm wavelength
ORF	Open Reading Frame

PCR	Polymerase Chain Reaction
pH	Negative $\log_{10}$ of proton concentration
RP-HPLC	Reverse-Phase High Pressure Chromatography
rpm	revolutions per minute
RT	Room Temperature
SDS-PAGE	Sodium Dodecyl Disulfate Polyacrylamide Gel Electrophoresis
T-COFFEE	Tree-based Consistency Objective Function For alignment Evaluation
TE	Tris-EDTA
TRX	Thioredoxin
TSB	Tryptone soya broth
UV	Ultraviolet
VRE	Vancomycin resistant *Enterococci*
w/v	Weight per volume

1. Introduction

1.1 Lactic acid bacteria

Lactic Acid Bacteria (LAB) are a group of gram-positive, non-sporulating bacteria defined less by phylogeny than by their metabolic capability to produce lactic acid as a by-product of fermentation (Pfeiler & Klaenhammer, 2007). These bacteria populate a wide variety of ecological niches, from milk and vertebrate GI tracts to plants, soils and shellfish (Stiles & Hastings, 1991; Lee *et al.*, 2010; Salvetti *et al.*, 2013). The presence of LAB in foodstuffs is one of the primary reasons that they play a critical role in food production, both as starter cultures in various fermentations and as additives to improve both flavour and texture of foods (Cintas *et al.*, 2001). In addition to these roles, there has been a greater recognition of the use of LAB as a "natural" preservative in food, that maintains flavour and quality without the use of harsh treatments or the use of synthetic preservatives (Nes *et al.*, 1996).

Another characteristic of LAB is their ability to produce one or more, small, diverse, anti-microbial peptides (Cintas *et al.*, 2001) collectively known as bacteriocins.

1.2 An introduction to bacteriocins

Bacteriocins are small, ribosomally-synthesised peptides (Cotter *et al.*, 2005) that exhibit antimicrobial activity against certain strains of microbes. Whilst many bacteriocins have only a narrow spectrum of activity several bacteriocins have been characterised that inhibit the growth of a larger range of targets (Cleveland *et al.*, 2001; Cotter *et al.*, 2005). It is interesting to note that those shown to possess a wider

phylogenetic range of activity are more often sourced from gram positive bacteria but that the reasons for this trend are uncertain (Nes *et al.*, 2007). Due to the fact that bacteriocins are produced by almost all bacterial and archaeal species (Klaenhammer, 1993), it is highly probable that there are many variations within this ensemble of peptides that are yet to be discovered.

To produce a fully functional bacteriocin, the host cell must produce numerous proteins in a coordinated fashion. The expression of these genes is normally triggered in response to quorum sensing molecules which result in a signal transduction cascade that leads to the activation of genes involved in bacteriocin production, maturation and secretion by specific response regulators (Trmčić *et al.*, 2011). This repertoire of proteins is required to produce the peptide, modify it, and then export it out of the cell. In addition to this the cell must protect itself from the deleterious effects of its own bacteriocin; usually through the use of an immunity protein specific to the bacteriocin it produces (Kjos *et al.*, 2011).

The transporter proteins produced in bacteriocin gene clusters are usually ATP-binding cassette (ABC) transporters (Todorov, 2009) that have two functions. They initially cleave the leader peptide from the bacteriocin precursor peptide, which activates the peptide, then export the mature peptide into the environment. Bacteriocin ABC transporters consist of three domains: a proteolytic N-terminal domain, a hydrophobic domain embedded in the cell membrane, and an ATP-binding domain at the C-terminus (Havarstein *et al.*, 1995). The N-terminal proteolytic domain is the site that recognises and cleaves the leader sequence of the bacteriocin (often identified by a double-glycine motif) from the mature peptide. The ATP binding cassette is a region

conserved between species, and provides energy through ATP binding and hydrolysis to transport the mature bacteriocin out of the cell and into the extracellular environment (Fath & Kolter, 1993).

Groups of bacteriocins are named for the species that produces them, such as the plantaricins produced by *Lactobacillus plantarum* strains and the lactococcins produced by *Lactococcus lactis.* The bacteriocins of *Escherichia coli* (colicins) were the first to be characterised and studied (Lazdunski, 1988). Since then, the study of bacteriocins has grown into a large and rapidly growing field of research, on the basis of their potential use as natural food preservatives (Ben Omar *et al.*, 2008; Mills *et al.*, 2011) and as an alternative treatment for antibiotic resistant bacteria (Giacometti *et al.*, 1999; Brumfitt *et al.*, 2002).

In a natural environment where a mixture of bacterial strains exists, a network of interaction between three types of species occurs. This network can be simplified by comparing it to a game of rock, paper, scissors where strains A, B and C represent rock, paper and scissors respectively (Figure 1.2.1). Strain A, the rock, is a bacteriocin producer; this gives it a clear advantage over strain C, the scissors, as it inhibits the growth of strain C, eliminating the competition for nutrients. However; strain B, the paper, is immune (or at least resistant) to strain A's inhibitory effects. As it does not have the metabolic burden of bacteriocin production it can grow faster with less energy expenditure. However, in the absence of strain A, strain C, which does not produce an immunity protein, can outcompete strain B (Kerr *et al.*, 2002). This variability in the environment results in an ever-changing mix of strains in the

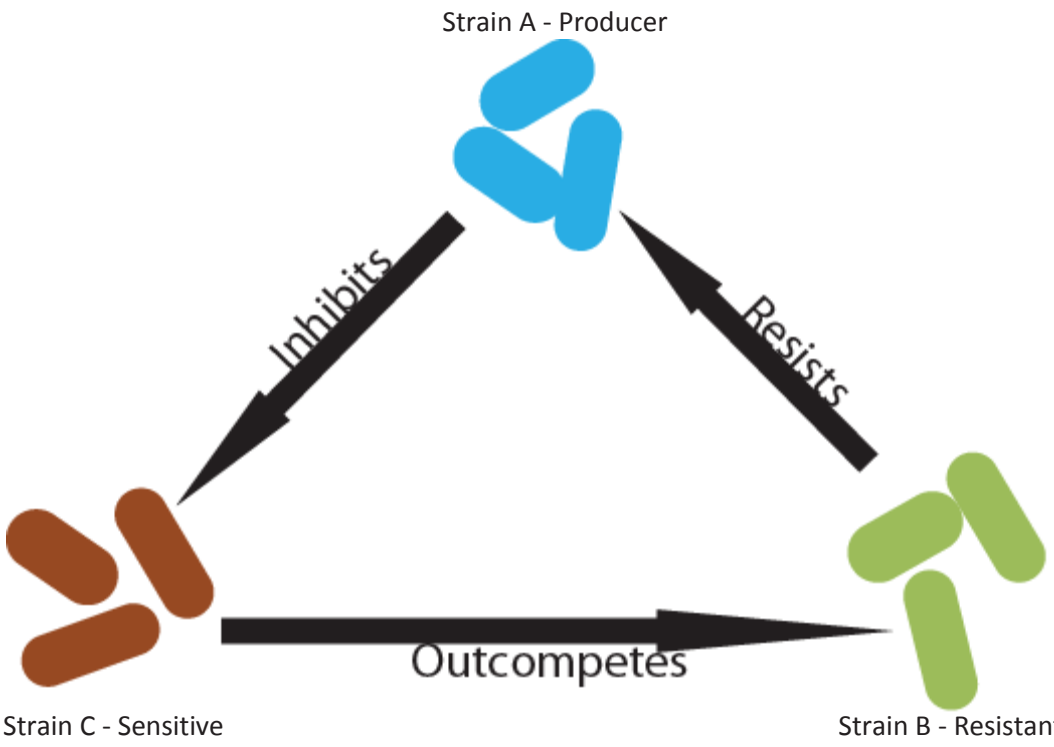


Figure 1.2.1 Schematic of the environmental dynamics of bacteriocin producing bacteria
A mixed culture of bacterial strains that either produce a bacteriocin, are resistant to a bacteriocin or are sensitive to a bacteriocin. Figure adapted from Hibbing *et al.* (2010). Adapted with permission from Nature Publishing Group.

population which means that production of bacteriocins may not always be the direct path to ecological dominance.

Amongst bacteriocins, those produced by LAB have been of particular interest due to the natural presence of LAB in many foodstuffs (Cleveland *et al.*, 2001) including dairy produce (Sip *et al.*, 2012), many low processed grain products (Hata *et al.*, 2010) and meats (Cleveland *et al.*, 2001). They are also generally regarded as safe (GRAS) for human consumption allowing them to be considered as potential candidates for food additives. This longstanding consumption of food products that contain LAB has provided an important base for the use of bacteriocins in modern food processing to limit the frequency of spoilage and reduce the microbial load of pathogenic bacteria (Cleveland *et al.*, 2001).

1.3 Classification of bacteriocins

Bacteriocins are, as mentioned, a diverse group of peptides that have been subjected to many classification schemes over the years. This is not only due to the abundance and diversity of currently known bacteriocins, but also because the field of research is not particularly old. Thus novel bacteriocins are continuously being found that sometimes have unpredicted modes of activity or structural features and do not adhere to any previously described classification system.

One of the first systems generated by Klaenhammer in 1993 has been modified, debated and reconstructed over the last twenty years with an expanding set of classes and subclasses to encompass the wide variety of bacteriocins being isolated and characterised. One of the main reasons for this level of disagreement is that many research groups propose classification systems for small subsets of bacteriocins. Klaenhammer's original classification system was intended only for the bacteriocins of *Lactobacillus* and did not include other bacterial species (Klaenhammer, 1993). Very recently a simple classification system was proposed that categorises the bacteriocins into two main classes: the modified bacteriocins and the unmodified bacteriocins, as shown in Table 1.3.1 (Cotter *et al.*, 2013).

While it does not suit all bacteriocins, it accurately encompasses the most numerous classes. In the earlier three and four class systems the first class was given over to lantibiotics, bacteriocins with post-translational modifications that result in the inclusion of lanthionine and methyl-lanthionine (Todorov, 2009; Lee & Kim, 2011).

Group	Distinctive feature	Examples
Class I (modified)		
MccC7-C51-type bacteriocins	Is covalently attached to a carboxy-terminal aspartic acid	MccC7-C51
Lasso peptides	Have a lasso structure	MccJ25
Linear azole- or azoline-containing peptides	Possess heterocycles but no other modifications	MccB17
Lantibiotics	Possess lanthionine bridges	nisin, planosporicin, mersacidin, actagardine
Linaridins	Have a linear structure and contain dehydrated amino acids	Cypemycin
Proteusins	Contain multiple hydroxylations, epimerizations and methylations	Polytheonamide A
Sactibiotics	Contain sulphur–α-carbon linkages	Subtilosin A, thuricin CD
Patellamide-like cyanobactins	Possess heterocycles and undergo macrocyclization	Patellamide A
Anacyclamide-like cyanobactins	Cyclic peptides consisting of proteinogenic amino acids with prenyl attachments	Anacyclamide A10
Thiopeptides	Contain a central pyridine, dihydropyridine or piperidine ring as well as heterocycles	Thiostrepton, nocathiacin I, GE2270 A, philipimycin
Bottromycins	Contain macrocyclic amidine, a decarboxylated carboxy-terminal thiazole and carbon-methylated amino acids	Bottromycin A2
Glycocins	Contain *S*-linked glycopeptides	Sublancin 168
Class II (unmodified or cyclic)		
IIa peptides (pediocin PA-1-like bacteriocins)	Possess a conserved YGNGV motif (in which N represents any amino acid)	Pediocin PA-1, enterocin CRL35, carnobacteriocin BM1
IIb peptides	Two unmodified peptides are required for activity	ABP118, lactacin F
IIc peptides	Cyclic peptides	Enterocin AS-48
IId peptides	Unmodified, linear, non-pediocin-like, single-peptide bacteriocins	MccV, MccS, epidermicin NI01, lactococcin A
IIe peptides	Contain a serine-rich carboxy-terminal region with a non-ribosomal siderophore-type modification	MccE492, MccM

Table 1.3.1 Class I and II bacteriocins

Recent, comprehensive classification system for bacteriocins from Cotter *et al.* (2013). Reproduced with permission from Nature Publishing Group.

The second class contained unmodified, heat-stable peptides which, in fact, often included post-translational modifications such as disulfide bonds (Cleveland *et al.*, 2001; Lee & Kim, 2011). The most recent classification system proposed by Cotter *et al.* (2013) removes a lot of the complications present in previous classification systems. The simple division between unmodified and modified bacteriocins is defined by unusual post-translational modifications such as the presence of lanthionine and methyl-lanthionine; sugar moieties and sulphur-α-carbon linkages (Cotter *et al.*, 2013). This system also allows for the naming of the subgroups of class I by their defining characteristics. An example is the glycocin subclass which incorporates bacteriocins that contain S-linked glycopeptides, such as sublancin 168 and glycocin F (GccF) (Oman *et al.*, 2011; Stepper *et al.*, 2011). The second class in this system incorporates the unmodified or cyclic bacteriocins, the former being divided into four other subclasses by specific characteristics such as those with conserved motifs and those that require two-peptides for activity (Cotter *et al.*, 2013).

Another group of peptides, previously included as type III bacteriocins are large (>30 kDa), heat-labile proteins (Cleveland *et al.*, 2001). It has been suggested that these don't meet the standard criteria characteristic of bacteriocins. Firstly, they cause cell lysis by cell-wall hydrolysis of the sensitive cell, rather than by forming pores. Secondly, that they do not always express an immunity gene specific to themselves but may instead rely on modifications to their own cell wall to provide resistance (Cotter *et al.*, 2005).

Another proposed method of classification and organisation of bacteriocins is classification by amino acid sequence, grouping bacteriocins based on their conserved

regions, rather than their activity (Zouhir *et al.*, 2010). This classification system was proposed to be less ambiguous than those based on structure and function. However, it is challenged by the appearance of novel bacteriocins that do not share sequences and could incorrectly group bacteriocins that rely on post-translational modifications for activity (such as the glycocins). Groups of these could have little sequence similarity beyond the modified residues or they could have similar protein scaffolds that are modified in different ways to exhibit significantly different activities.

1.4 Range of activity

Although bacteriocins are predominantly active against species that are closely related to the producer strain (Klaenhammer, 1993) it has been observed that some affect a very wide spectrum of species. The plantaricin MG affects many gram positive bacteria (including, but not limited to; *Micrococcus luteus, Listeria monocytogenes, Staphylococcus aureus*, and *Bacillus subtilis*) as well as several gram negative bacteria Including *Pseudomonas putida,* several different disease causing *Escherichia coli* strains (Enteroinvasive, enterotoxic and enteropathogenic) as well as *Salmonella typhimurium (Gong et al., 2009b)*. The activity of plantaricin MG is such that it causes a more than 95% reduction in viable cell count of *S. typhimurium* after 9 hours of incubation (Gong *et al.*, 2009b). Other plantaricins have also been shown to inhibit the growth of gram negative pathogens (Gong *et al.*, 2009a; Sip *et al.*, 2012). Because of this potentially wide range of activity exhibited by LAB derived bacteriocins, the idea of using them as starter cultures in cheese and other fermented dairy products has been explored with several groups investigating changes in the microbial load of products such as Gouda

cheese when they are manufactured using *Lactobacillus* spp. as starter cultures (Mills *et al.*, 2011; Sip *et al.*, 2012)

1.5 Modes of Activity

While many of the well-known bacteriocins, such as nisin, are bactericidal bacteriocins (Matsusaki *et al.*, 1998), it has been observed that others, such as Bac 51, are bacteriostatic (Yamashita *et al.*, 2011). Instead of lysing or otherwise killing the target cell, these bacteriocins inhibit the growth of the cells, removing competition for resources. The advantage of this strategy over killing competing organisms is not clear.

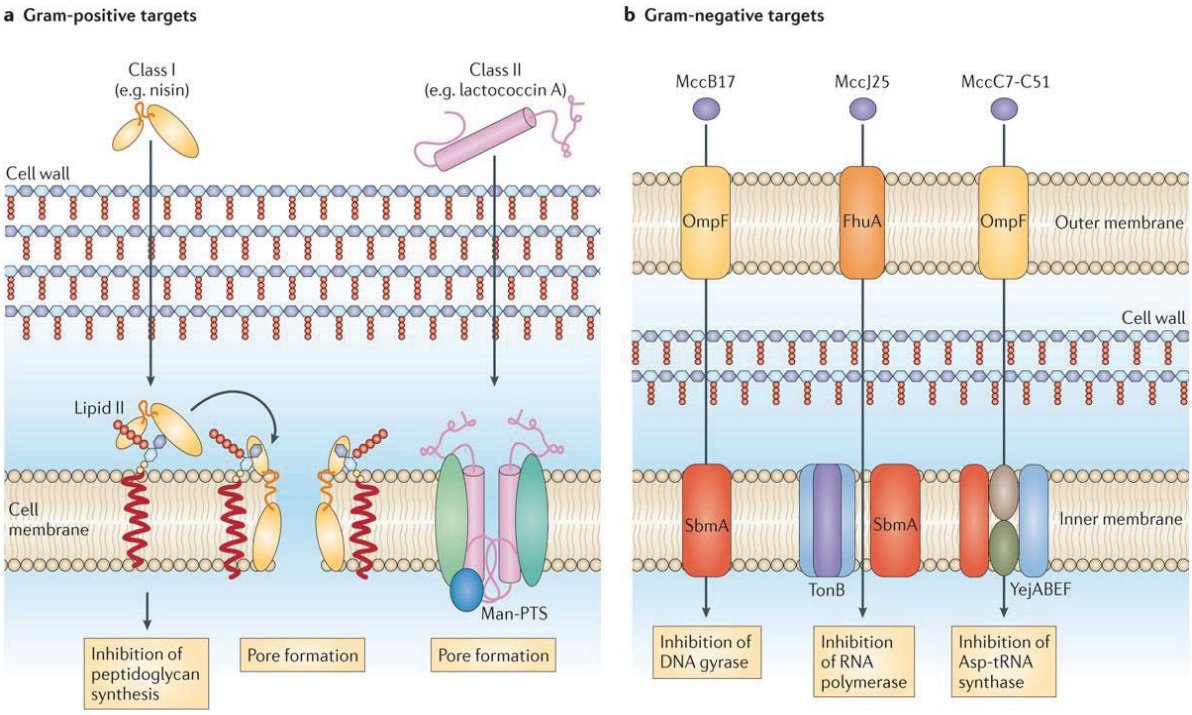


Figure 1.5.1 Diagrams showing bacteriocin activity mechanisms
(A) Binding mechanisms of gram-positive-targeting bactericidal bacteriocins (B) Binding mechanisms of gram negative targeting bacteriocins. Figure from Cotter *et al.* (2013). Reproduced with permission from Nature Publishing Group

Lytic bacteriocins were originally thought to cause lysis by interacting exclusively with the target cell membrane to form pores causing the leakage of intracellular fluid (Figure 1.5.1A) (Zendo *et al.*, 2010). More recently it has been shown that many

bactericidal bacteriocins also act to directly degrade the cell membrane/cell wall to bind to intracellular proteins such as DNA gyrase, required for DNA repair and cell replication (Figure 1.5.1B) (Destoumieux-Garzon *et al.*, 2002). For pore-forming, lytic, bacteriocins, the actual cause of cell death of target cells depends on the size of the pore formed. Cell death can be due to ATP/ion loss (Breukink *et al.*, 1999), leakage of large molecules (Eijsink *et al.*, 1998; Yoneyama *et al.*, 2009b), or other changes in the cell's internal environment, such as changes in internal pH or changes in membrane potential (Gong *et al.*, 2009b; Pingitore *et al.*, 2009).

Bacteriostatic bacteriocins are less well characterised, it has been proposed that they cause the cessation of cell growth by binding to surface receptors and triggering a signal cascade that halts (or heavily reduces) the biosynthesis of structural polymers such as peptidoglycan (Kamensek & Zgur-Bertok, 2013). It has been hypothesised that GccF disrupts essential intracellular processes by hijacking the regulatory phosphorelay cascades of GlcNAc-specific phosphotransferase transport systems (Norris & Patchett, 2014)

The bactericidal actions caused by bacteriocins often require a specific target molecule. In the case of nisin, a cell wall precursor known as lipid II anchored in the cell-membrane is central to the inhibition of the target cell (Breukink *et al.*, 1999). It is thought that the cell poration leading to cell death is initiated by the N-terminus of nisin binding to the pyrophosphate moiety of lipid II (Figure 1.5.2). The peptide then bends around the flexible central hinge region of lipid II allowing its C-terminal region to penetrate the membrane of the cell (Hsu *et al.*, 2004). The binding of nisin to the

cell-wall precursor also inhibits peptidoglycan biosynthesis, providing a dual mode of activity of the bacteriocin (Wiedemann *et al.*, 2001).

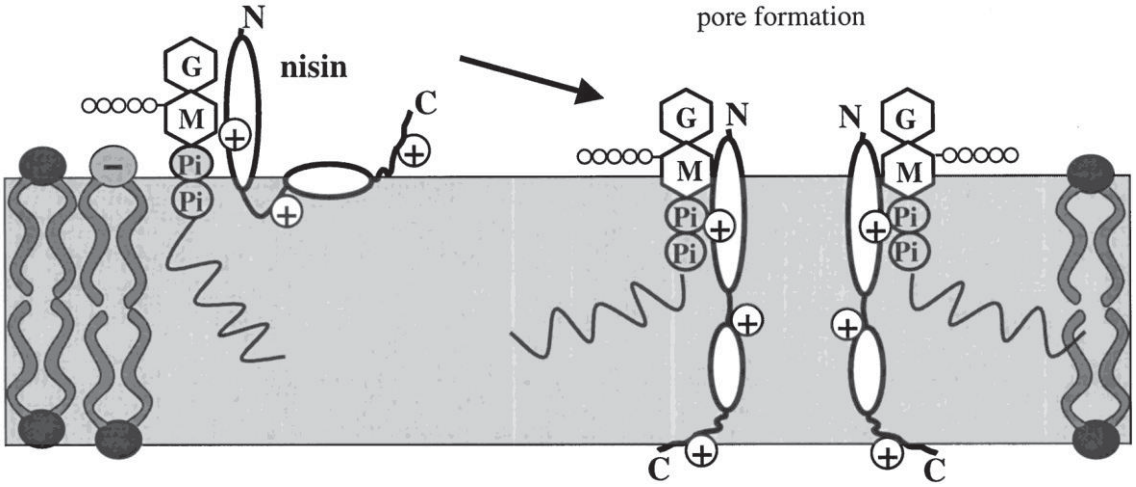


Figure 1.5.2 Schematic of the proposed activity of nisin binding to Lipid II.
Several complexes of nisin and lipid II are proposed to form the pores, resulting in relatively large, stable pores with low selectivity. Figure from Wiedemann, *et al.* (2001*)*. Reproduced with permission from The American Society for Biochemistry and Molecular Biology

It was shown by Diep *et al.* (2007), that several subclasses of bacteriocins can be identified that rely on interaction with a mannose PTS transporter for anti-microbial activity. These include the class IId bacteriocins which are unmodified linear bacteriocins such as lactococcin A and B and the class IIb bacteriocins, plantaricin EF and lactococcin G. However in light of the sheer variety of bacteriocins that have so far been characterised, it is not surprising that they exhibit a wide variety of binding mechanisms. Indeed some, such as the bacteriocin lacticin Q, function without a specific receptor protein (Yoneyama *et al.*, 2009a).

1.6 Applications of Bacteriocins

Bacteriocins are potentially very useful in a wide variety of fields for the selective elimination of specific species of bacteria. There are wide applications in the food industry where an increasing demand for naturally preserved food has arisen over the last decade, this has coincided with an increase in the amount of prepared or processed food requiring long-term storage and transportation. Bacteriocins have been found to enhance marketability, especially the use of nisin and pediocin as these two bacteriocins are active against multiple *Bacillus* species that can cause gastroenteritis if ingested (Cabo *et al.*, 2009). Although nisin and pediocin inhibit different species, combined they have been shown to reduce the number of cells in produce to such an extent that after 90 days of storage no vegetative cells could be detected. In contrast, high levels of vegetative cells were detected in untreated samples (Cabo *et al.*, 2009).

It has been shown that bacteriocins produced by bacterial strains such as *Lactobacillus salivarius* UCC118 can reduce the chance of host infection by *Listeria monocytogenes* in animal models (Corr *et al.*, 2007). It is therefore possible that the consumption of selective mixtures of bacteria that are able to co-exist in the gastro-intestinal tract (GIT) and produce bacteriocins without harming the pre-existing microflora might act as a prophylactic treatment or, even a potential cure for infections of the GIT caused by food-borne pathogens. Due to the diversity and range of targets of bacteriocin producing strains, this area is poised for exploitation. Currently, isolation of the producer strain can be relatively simple and allows for the analysis of the bacteriocin(s) produced by the strain (characterisation of activity, specificity, function and structure).

The purification and mass production of pure bacteriocins is, however, currently considered a challenge, necessitating the investigation of methods to improve both quality and yield.

Modern industrial farming techniques have resulted in swathes of land being given over to large monocultures of plants. However, high density plantings provide a breeding ground for plant diseases such as blights that will, if successful in adapting to the defences of a specific species, have a high chance of infecting the entire local population and potentially spreading further. Current methods used to combat these potential blights usually involve the use of antibacterials and pesticides (the latter to prevent the transmission of insect-borne bacterial diseases). At the same time consumer demand for natural produce not treated with pesticides or fungicides is increasing. The use of bacteriocins is thus a potential solution for reducing plant pathogens in the soil.

The application of bacteriocins to control plant pathogens is an idea that has been around for several decades with early studies providing promising results. For example, the reduction of crown galls in tomato plants caused by *Agrobacterium radioacter* var. *tumefaciens* was achieved by utilising a closely related strain of *Agrobacterium* which produces a bacteriocin that inhibits growth of the pathogenic strain (Kerr & Htay, 1974). More recently, it has been shown that a purified bacteriocin produced by *Pseudomonas syringae* pv. *ciccaronei* can reduce the mass of knots characteristic of Olive Knot Disease by up to 81% (Lavermicocca *et al.*, 2002). Utilisation of bacteriocins to combat such diseases could have a massive commercial impact as the difference in fruit yield between “light” infections and “moderate”

infections has been shown to be approximately 28% (Schroth *et al.*, 1973). This translates to a potential difference in yield of more than half a ton per acre in a mature olive orchard (Vossen *et al.*, 2011). Though research into bacteriocin production by plant pathogens is limited, studies with soy beans, rice and potatoes have shown that they are a veritable treasure trove of biological control systems for the control of diseases of valuable commercial food crops (Heu *et al.*, 2001; Holtsmark *et al.*, 2007).

Due to the rising frequency of pathogenic bacterial strains that are immune to conventional (and, increasingly, less conventional) antibacterials (Deurenberg *et al.*, 2007; David & Daum, 2010) it has been suggested that bacteriocins may provide a substantial advantage in the treatment of drug-resistant infections. It has been shown that the use of bacteriocins to counter the growth of *Pseudomonas aeruginosa*, a species often resistant to common antibiotics, can result in a six log-cycle reduction in viable cell counts within 20 minutes (Giacometti *et al.*, 1999). Similar results have been documented from experiments when the use of nisin in combination with more traditional antibiotics reduced the viable cell counts of Methicillin Resistant *Staphylococcus aureus* (MRSA) and Vancomycin Resistant *Enterococci* (VRE) strains by three log-cycles within an hour (Brumfitt *et al.*, 2002). Other bacteriocins, such as those produced by *L. paracasei* subsp. *paracasei* BMK2005 have been shown to have bactericidal activity against multidrug-resistant pathogens such as *Klebsiella oxytoca*, MRSA, and enteropathogenic *E. coli* (Bendjeddou *et al.*, 2012).

Finally, bacteriocins can also be used as a method of food preservation. As many of the bacteriocin strains are isolated from food sources such as dairy (Gong *et al.*, 2009a), meats (Cleveland *et al.*, 2001), and fermented grain products (Ben Omar *et al.*, 2008)

their ability to reduce the microbial load in their source food product is often investigated thoroughly. The bacteriocin-producing strains have been shown to be highly efficient at reducing the levels of pathogens present in their normal environment. An example of this is *L. plantarum* LMG P-26358, when used as a starter culture in Gouda cheese the post-maturation cell count of *Listeria inoccua* in the cheese was 3 log cycles lower than that with only the normal *L. lactis* DPC4268 starter culture (Mills *et al.*, 2011).

1.7 Genetics of Bacteriocins

Bacteriocins are generally encoded in all-inclusive operons containing several genes. Individual genes or groups of genes within the operon may have their own promoter regions, allowing varying levels of expression at different stages of cell growth (Klaenhammer, 1993). Bacteriocin operons express the structural gene for the bacteriocin, genes for any proteins required for modifications specific to the bacteriocin, a secretion protein for the bacteriocin, a regulatory protein that controls the expression of the bacteriocin, and, normally, an immunity protein (Klaenhammer, 1993).

In most bacteriocins the structural gene encodes the basic structure of the bacteriocin with the signal peptide attached (if any). This is generally referred to as the precursor peptide or prebacteriocin and is often inactive without subsequent modifications (Todorov, 2009).

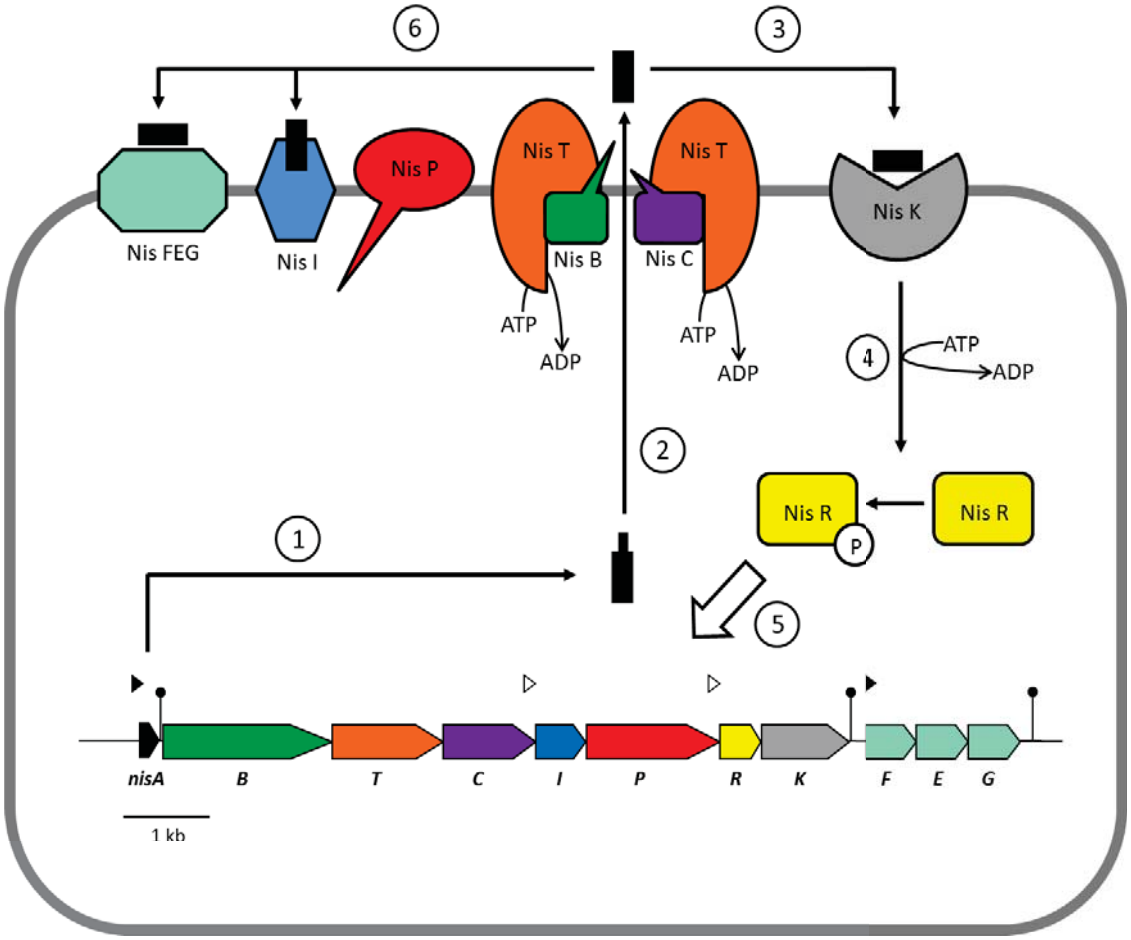


Figure 1.7.1 Schematic of nisin operon, component localisation and bacteriocin synthesis
Genes are matched by colour and name with the protein they encode. (1) Formation of the prebacteriocin (2) Modification of the prebacteriocin and transport to extracellular environment (3) Detection of bacteriocin by producer cell (4) Activation of response regulator (5) Activation of regulated gene transcription (6) Protection of producer cell by immunity-related ABC – transporter NisFEG and secondary immunity peptide NisI. Image adapted from Trmčić *et al*. (2011) (with permission from Elsevier) and Chen & Hoover (2003) (With permission from Wiley)

In nisin-producing *Lactococcus lactis* strains, the fully characterised bacteriocin operon is known to contain 11 genes (*nisABTCIPRKFEG*) (Figure 1.7.1) that are under the control of a mixture of promoter regions, allowing for expression of different components of the bacteriocin system at different stages during cell growth. Trmčić *et al.* (2011) showed that in the early stages of bacteriocin production, only the gene encoding the nisin A precursor peptide, *nisA* is expressed. Although *nisA* shares a promoter with the group *nisBTCIP*, the two transcripts are expressed at different

concentrations. This, combined with the presence of putative transcription terminator loops, suggests that the *nisA* transcripts are individually expressed and that expression of the *nisBTCIP* cluster occurs afterwards. This delay is because *nisB* (a dehydratase that prepares the prepeptide for *nisC*), *nisC* (A cyclase responsible for formation of the lanthionine rings (Koponen *et al.*, 2002)), *nisT* (An ABC transporter responsible for secretion of nisin), and *nisP* (A membrane anchored protease responsible for removing the signal sequence and activating the peptide) are not required until the nisin pre-peptide is present in the cell. *nisR* and *nisK* are genes that encode regulatory proteins that detect the presence of extracellular nisin and activate promoters upstream from *nisA* and *nisF*. It has been proposed that the *nisRK* expression is constitutive as they are regulated by a nisin-independent promoter. Such constitutive expression permits the strain to respond to changes in the extracellular levels of nisin as they change. The nisin operon has three immunity genes named *nisFEG*; the expression of this group is controlled by a nisin-dependent promoter. The expression of these immunity genes produces peptides that protect the nisin-producing cells from self-destruction (Trmčić *et al.*, 2011).

The immunity genes associated with bacteriocins (including *nisI* and *nisFEG* for nisin-producing species) were originally thought to be a requirement for classification as a bacteriocin (Nes *et al.*, 1996). As more bacteriocins from gram-negative strains are becoming characterised, it has become apparent that many bacteriocin producing bacteria do not utilise specific immunity proteins but instead rely on specialised ABC transporters that allow the cells to counter the inhibitory effect of their own bacteriocin(s) (Diaz *et al.*, 2003). The realisation that not all bacteriocin producing

strains have a distinct immunity protein has improved the potential for the use of bacteriocins in food preservation as it had previously been observed that some closely related bacterial strains (often the target of the bacteriocin in question) could acquire the immunity gene without acquiring the ability to produce the bacteriocin (Ben Omar *et al.*, 2008). If the source of host immunity is more complex than a single gene in a cluster, it can safely be assumed that the risk of "transference" of immunity is significantly reduced, increasing the utility of the bacteriocin in commercial environments.

Bacteriocin gene clusters are usually encoded on either plasmid or chromosomal DNA although divisions in gene clusters have also been observed. In these cases elements of the bacteriocin operon can be partially encoded on a plasmid and partially on chromosomal DNA. For example, *Carnobacterium piscicola* LV17 expresses a bacteriocin from both plasmid and chromosomal DNA; the bacteriocin structural and immunity genes are encoded chromosomally but full bacteriocin production relies on specific genetic elements encoded on a plasmid (Saucier *et al.*, 1997).

Plasmids encoding bacteriocins have been observed in *Lactobacillus acidophilus* (Vandervossen *et al.*, 1994); *Lactococcus lactis* (Kojic *et al.*, 2006); and *L. plantarum* (Van Reenen *et al.*, 2006; Karthikeyan & Santosh, 2009). These strains usually produce multiple bacteriocins with one or more situated on a plasmid. In some strains that produce multiple bacteriocins (such as *L. lactis* subsp. *lactis* biovar *diacetylactis* DPC938) it has been noted that all of the bacteriocin genes are located on the same plasmid, as are the immunity, transportation and modification genes relevant to those bacteriocins (Morgan *et al.*, 1995). Plasmid encoded bacteriocins are theoretically

easier to manipulate and characterise *via* the engineering of genetic knockouts or mutants to probe the mechanisms of action, pathways to maturation and the functionalities of the bacteriocins. Nisin, mentioned earlier as the prime candidate for modern food preservation, is a chromosomally encoded bacteriocin for which the genetic elements are situated in a large conjugative transposon (Steen *et al.*, 1991; Trmčić *et al.*, 2011).

It is worth noting that a single genome can be host to multiple bacteriocin clusters and that often, more than one of these clusters is active. This poses problems as far as analysis of activity is concerned as many studies base preliminary results of findings on the activity of crude supernatant which may contain more than one active bacteriocin and thus may produce an inaccurate assessment of the activity of what is assumed to be a single bacteriocin (Sawa *et al.*, 2013).

1.8 Glycosylated bacteriocins

In eukaryotic organisms, glycosylation is one of the most common post-translational modifications. In prokaryotes, however, it is less common and as a result is less well characterised. It was previously thought that glycosylation occurred only through *O*-linked or *N*- linked glycosidic bonds but several discoveries within the bacteriocin field have challenged this (Oman *et al.*, 2011; Stepper *et al.*, 2011).

Initially, the isolation of a bacteriocin produced by *B. subtilus* characterised was characterised as a lantibiotic and named sublancin 168 (Paik *et al.*, 1998). However, further characterisation using FT-ICR-MS showed a difference in mass when compared to the sequence prediction consistent with two disulfide bonds and a hexose molecule.

Further investigation showed that there is a hexose moiety *S*-linked to the Cys22 residue. A discovery that lead to its reclassification as a glycocin (Oman *et al.*, 2011; Stepper *et al.*, 2011).

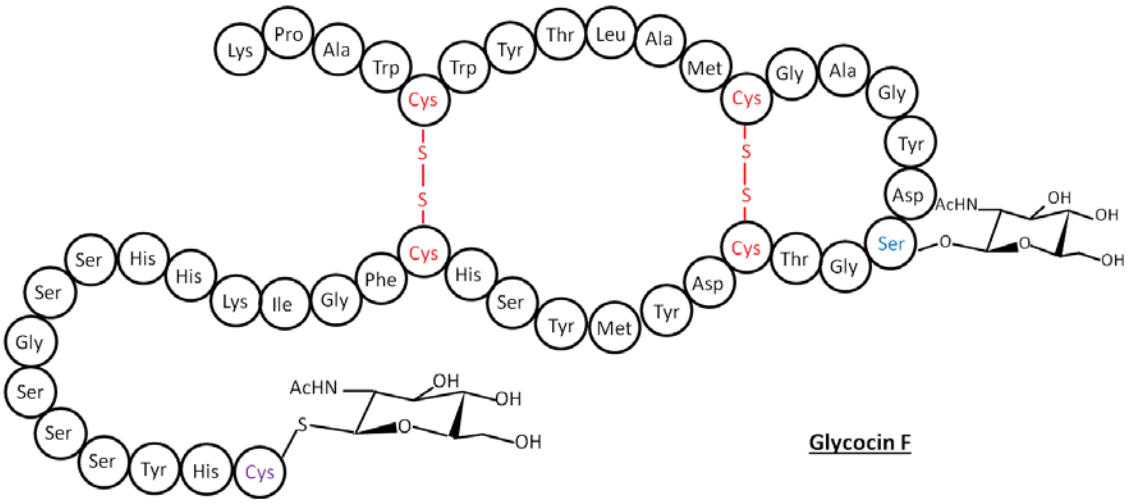


Figure 1.8.1 Schematic of the primary structure of the glycosylated bacteriocin Glycocin F
Disulfide bonds forming nested cysteines (Red) and N-acetyl-D-glucosamine moiety *O*-linked to serine (Blue) and *S*-linked to the C-terminal cysteine (Purple). Figure from Arnison *et al.* (2013). Adapted with permission from the Nature Publishing Group

A similar modification was found on GccF, a bacteriocin produced by *L. plantarum* KW30 that was initially isolated from Kaanga Wai or fermented corn (Kelly *et al.*, 1996), and is a post-translationally modified bacteriocin with two glycosylation sites. The first is an *O*-linked N-Acetylglucosamine (GlcNAc) at Ser18 and the second is an unusual *S*-linked GlcNAc at Cys43 (Figure 1.8.1) (Venugopal *et al.*, 2011). The thioester linkage used to attach one of the GlcNAc moieties is proposed to be highly advantageous to the bacteriocin as it offers greater stability than the *N*- and *O*-linkages at both high and low pH and is less susceptible to hydrolysis (Zhu *et al.*, 2004; Liang *et al.*, 2009). It has been shown that deglycosylation does not affect sublancin's activity but severely reduces the activity of GccF (Stepper *et al.*, 2011; Wang & van der Donk, 2011).

These bacteriocins are currently the main members of the class I subclass, the glycocins. The glycocin subclass is currently defined as bacteriocins containing *S*-linked glycopeptides but may be re-worked to include other glycopeptide linkages as they are characterised. In addition to the unusual mode of glycosylation, the first four cysteines in glycocin F are situated in an α/α fold, more specifically a $C–X_6–C–X_n–C–X_6–C$ motif with two nested sulphide bonds between cysteine residues 12 and 21 and 5 and 28 constraining residues 1-32 in a hairpin structure (Figure 1.8.1). This conserved structure is another notable feature of glycocins (Stepper *et al.*, 2011).

It has been previously noted that several bacteriocins utilise mannose-PTS transporters as a target for their initial interaction (Diep *et al.*, 2007). However, it has been shown than previously susceptible bacterial strains can become resistant, or even immune, through the down-regulation of the phosphotransferase system that renders them susceptible, thus reducing the inhibitory effect by decreasing the prevalence of binding targets available (Kjos *et al.*, 2011). This modification of sugar transport systems could imply that specific bacteriocins, particularly those with glycopeptide modifications, are produced in an effort by their producer cell to reduce competition for a specific carbon source, rather than to explicitly kill off competitors. By using sugar moieties as bait molecules the producer cell can specifically knock-down direct competitors for carbon sources, increasing the efficiency of bacteriocin production.

A notable factor about glycocins is that unlike lantibiotics and plantaricins their operons tend to be quite simple, the GccF operon in *L. plantarum* KW30, for example, consists of 8 ORFs (Stepper *et al.*, 2011) compared to nisin, which has 11 ORFs (Trmčić

et al., 2011) and the highly variable plantaricin cluster with up to 25 genes, as in *L. plantarum* C11 (Diep *et al.*, 2009).

1.9 ASM1

ASM1 is a glycocin produced by *L. plantarum* A-1 isolated from tortilla, a Mexican bread derived from corn (Hata *et al.*, 2010). It was established by Hata *et al.* (2010) approximately 88% sequence identity with GccF and work done by Man *et al.* (Unpublished) showed it to be similarly modified. The presence of both GlcNAc moieties on ASM1 is required for full bacteriostatic activity and interestingly, as is the case for GccF, the addition of GlcNAc into growth media protects cells normally sensitive to the growth inhibition caused by GccF (Stepper *et al.*, 2011). A third glycocin with a similar basic structure is the previously mentioned sublancin 168.

The ASM1 peptide is highly heat stable, capable of being heated to 90°C for fifteen minutes without a significant loss in inhibitory activity (Hata *et al.*, 2010). In addition to this it is also highly resistant to increased pH, unlike many bacteriocins produced by *Lactobacillus* species (including nisin), retaining approximately 90% inhibitory activity after being treated at pH 8.5 (Hata *et al.*, 2010). This activity in more alkaline environments could imply that ASM1 will be useful for food preservation in a separate range of food products to nisin.

ASM1 was also tested against several LAB species in an effort to analyse its range of activity. It was found to have activity against only a few of the strains tested and was therefore deemed to have a narrow range of activity, affecting only closely related *Lactobacillus* species (Hata *et al.*, 2010).

It is currently uncertain whether or not the ASM1 bacteriocin is encoded in a similar fashion to the chromosomally-based GccF operon as only a partial sequencing of the *asm1* cluster has been completed (Hata *et al.*, 2010). From the open reading frames identified so far it appears to have a similar genetic layout to that seen in GccF (Figure 1.9.1). It is uncertain however, whether or not the putative operon is plasmid-based or whether it is situated on the chromosomal DNA of *L. plantarum* A-1.

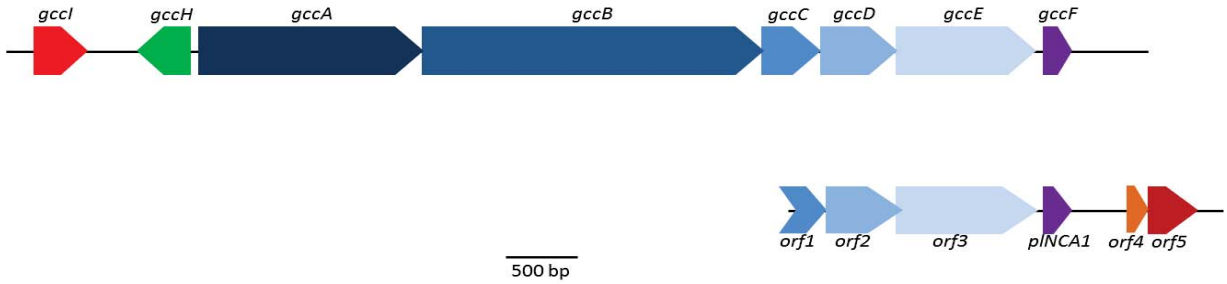


Figure 1.9.1 Proposed alignment of the sequenced glycocin F cluster and the ASM1 cluster
Structural genes of both clusters are in purple, homologous sequences are matched by colour. *orf1* in ASM1 is a partial ORF with a partial conserved thioredoxin (TRX) domain, like *gccC*, *orf2* shows homology to both TRX and bacteriocin transport accessory proteins, *orf3* shows homology with LytTR response regulators. *orf4* and *orf5* did not show significant homology to known proteins. Figure adapted from Stepper *et al.* (2011) (With permission from Elsevier) and Hata *et al.* (2010) (With permission from Elsevier)

1.10 Aims of Research

Genetics of the ASM1 cluster:

Due to the similarities in primary sequence between ASM1 and glycocin F it can be hypothesised that the ASM1 gene cluster is similar to that present in *L. plantarum* KW30 encoding GccF. In addition to this, artefacts have arisen in preliminary sequencing that suggest the gene cluster is present in multiple copies, indicating that the gene cluster is present on a plasmid in the *L. plantarum* A-1 strain.

Specific Aims:

- To extend the core sequence of the ASM1 gene cluster provided by Hata *et al.* (2010).
- To determine whether ASM1 is located on a plasmid or a chromosome in *L. plantarum* A-1.
- To clone or otherwise insert the ASM1 gene cluster into a heterologous host to provide a basis for experiments to study the roles of individual genes through knockout and mutation experiments.
- To genetically identify secondary bacteriocin gene clusters in the A-1 genome.

Activity of the ASM1 bacteriocins

Because of the sequence similarities between the two glycocins it is probable that there are high levels of similarity in their inhibitory effects against sensitive bacteria; range of species affected, and range of environmental conditions in which the glycocins are active. Additionally, the characterisation of bac-minus mutants generated by treatment with novobiocin (a known mutagen and plasmid curing agent) will aid in the characterisation of the immunity – activity relationship within the strain.

Specific Aims

- To characterise the activity of the ASM1 protein and compare its activity to that of the homologous glycosylated bacteriocin Glycocin F.
- To investigate novobiocin-induced ASM1-minus mutants with regards to their potential production of additional bacteriocins and the reason for their susceptibility to ASM1.
- To characterise the activity of any secondary bacteriocins in *L. plantarum* A-1.

2. Methods and Materials

2.1 Bacterial Methods

2.1.0 *Lactobacillus* overnight cultures

Unless otherwise stated, overnight cultures were made by inoculating 1 mL sterile MRS media (Accumedia)(Lansing, MI, United States) from either frozen glycerol stock of the intended strain or a single colony from a plate. Samples were taken from their source with a flame sterilised loop.

2.1.1 *Enterococcus faecalis* overnight cultures

Unless otherwise stated, overnight cultures of *E. faecalis* were made by inoculating 1 mL sterile TSB (Oxoid)(Basingstoke, United Kingdom) from a glycerol stock using a flame sterilised loop to scrape a small volume from the stock.

2.1.2 Glycerol stocks

Glycerol stocks were created by growing the bacteria in 15 mL of media to an OD_{600} of 0.6 – 0.8, pelleting the cells at 3,500 x g for 10 minutes and then re-suspending them in 800 µL of media and 200 µL glycerol (Carl Roth®)(Karlsruhe, Germany) before snap freezing. The cells were then stored at -80°C for up to 12 months.

2.1.3 Novobiocin mutants

Novobiocin generated mutants were formed by using overnight cultures of *Lactobacillus plantarum* A-1 to inoculate 1 mL volumes of MRS containing 0.2 µg/mL novobiocin with or without 2% GlcNAc. These cultures were then incubated at 40°C for

48 hours, during which time they were mixed by brief vortexing twice daily. They were then diluted by dosing 5 mL of MRS media with 5 µL of culture and then using between 2.5 - 10 µL of each solution to inoculate a further 10 mL of MRS media (resulting in a final dilution of 5 $x10^5$ - $2x10^6$). 50 µL of each final dilution was dispersed in 10 mL of MRS +1% agar at 40°C, poured into plates and incubated at 30°C for 40 hours. Each plate was then overlaid with 14 mL of MRS +1% agar containing 70 µL of an overnight culture of *L. plantarum* ATCC 14917 (a known ASM1-sensitive strain of *L. plantarum*) and incubated overnight at 30°C.

2.1.4 *L. plantarum* growth and bacteriocin production

50 µL of an overnight culture of the required strain of *L. plantarum* was used to inoculate two 15 mL falcon tubes containing 10 mL of sterile MRS media to an $OD_{600\ nm}$ of ~0.1. These cultures were then incubated at 30°C and 0.6 mL aliquots taken every two hours for eight hours with the final sample taken at 24 hours. The $OD_{600\ nm}$ of each sample was immediately determined after which the cells were removed from the supernatant by centrifugation at 13,000 x g. The pH of the supernatant was recorded and the sample was stored at 4°C until all samples had been acquired.

Bacteriocin production was measured by treating a plate of MRS media + 1% agar inoculated with 50 µL of an overnight culture of a sensitive strain of *L. plantarum* with 4 µL samples of the supernatant from different time points and comparing the size of the “halos” of bacterial inhibition caused by the supernatant.

2.1.5 Bacteriocin activity assays (broth and agar)

The effects of both GccF and ASM1 on bacterial growth were tested in both liquid assays and by soft agar assays. The former was used to determine the level of activity provided by a specific dose of the bacteriocin, as well as the effects of variables like temperature and pH upon activity. Soft agar assays were mostly used to determine if the bacteriocin inhibited the growth of specific strains of bacteria.

For liquid-based activity assays, 10 mL of sterile MRS media was inoculated with 500 µL of an overnight culture of *L. plantarum* ATCC 8014, then incubated at 30°C until the $OD_{600\ nm}$ reached 0.6 – 1. This was then used diluted to an $OD_{600\ nm}$ of ~0.2 and 1 mL aliquots were added to a number of 1 mL plastic cuvettes. These cuvettes were then inserted into the multi-sample holder of a Varian (Palo Alto, CA, United States) Cary™ 300 Bio at a set temperature (regulated by the Varian Cary Temperature Regulator accessory) and the $OD_{600\ nm}$ was measured every two minutes. After approximately 30 minutes, bacteriocin was added to each tube, mixed by gentle pipetting to avoid excessive oxygenation and then further incubated for another 120-180 minutes under the same conditions.

Soft agar assays utilised media with 0.8% agar that contained 50 µL of an overnight culture of an indicator strain dispersed in 15-20 mL of agar while it was still liquid at 40°C. These plates, when set were then spotted with small volumes (2 - 5 µL) of bacteriocin at different concentrations or cell-free supernatant from a bacteriocin producer strain and incubated at 30°C (or less, if required) overnight. Areas of bacteriocin activity were indicated by clearance “halos” in the agar in where cell

growth had not occurred or had been significantly attenuated. These could be measured both in terms of completeness and diameter of clearance.

2.1.6 GlcNAc protection

It had been observed that the presence of N-Acetylglucosamine (GlcNAc) severely reduced the activity of GccF against most *Lactobacillus* species (Stepper *et al.*, 2011). To confirm that GlcNAc had the same effect on ASM1 the following experiments were carried out.

Firstly, 50 µL of *L. plantarum* ATCC 8014 was dispersed in 20 mL of MRS + 0.8% agar at 40°C, poured into a plate and allowed to set. 3 µL aliquots of mixtures of a monosaccharide and purified ASM1 at concentrations of 0.5 M of a sugar (from a selection of 15 different sugars) and 25 µM ASM1 were then spotted onto the plate. When all the excess liquid had evaporated the plate was then incubated overnight at 30°C.

Following this, liquid assay experiments were carried out following the basic method described in section 2.1.5. However, 30 minutes after the addition of the bacteriocin GlcNAc was added to the culture to a range of final concentrations to determine whether or not it could induce recovery from the effects of ASM1.

2.1.7 Competent *Escherichia coli* cells

To make *E. coli* XL1 cells competent, the following method was used modified from Sambrook and Russel (2001).

A 1 mL aliquot of LB media was inoculated with *E. coli* XL1 from a glycerol deep using a sterile wire loop and incubated overnight at 37°C. This overnight culture was then used to inoculate two x 50 mL volumes of LB media which were then incubated at 37°C for 2.5 - 3 hrs with shaking at 200 rpm. The cultures were then cooled on ice for 5 minutes before being subjected to centrifugation at 3,500 x g for 10 minutes at 4°C to create a loose pellet at the bottom of each tube. The supernatant from each tube was gently removed before each pellet was washed with 5 mL of ice-cold 0.1M $CaCl_2$ by gentle swirling. Following this step, the cells were re-centrifuged as before and each pellet gently re-suspended in 1.5 mL of ice-cold 0.1M $CaCl_2$ then stored, on ice, overnight at 4°C. Glycerol was added to each suspension to a final concentration of 10% and the cells were stored in 100 µL aliquots at -80°C.

2.1.8 Competent *Lactobacillus* cells

To make *Lactobacillus plantarum* A-1 or NC8 cells competent, a modified version of the method from Aukrust & Blom (1992) was utilised. 1 mL of MRS media containing 1.2% glycine was inoculated with the desired bacterial strain from a stored glycerol stock and incubated overnight at 30°C without shaking. The 1 mL culture was used to inoculate 2x 50 mL volumes of MRS containing 1.2% glycine and incubated at 30°C for 3-4 hours. Both cultures were then placed on ice for 10 minutes before being centrifuged at 3,000 x g for 10 minutes (less if possible, though NC8 does not pellet easily). The supernatant was removed and each pellet gently washed in 50 mL of $MgCl_2$ before being subjected to centrifugation, as before. The supernatant was then carefully removed and the cells gently re-suspended in 50 mL of 30% PEG 1500 before being pelleted a final time. After removal of the supernatant the cells were re-

suspended in 1 mL of 30% PEG 1500, and stored in 50 µL aliquots at -80°C for no more than 2 months.

2.1.9 Transformation of *E. coli* XL1

Competent cells were transformed by thawing a 100 µL aliquot on ice before adding 20 to 40 ng of DNA and incubating the mixture on ice for 30 minutes. After this time the cells were transferred to a 42°C water bath for exactly 2 minutes then put on ice for a further 5 minutes. The chilled cells were used to inoculate 1 mL of SOC media, which was then incubated for an hour at 37°C, with shaking at 200 rpm to allow for the synthesis of the recombinant protein that allows for antibiotic selection. The incubated cells were then plated onto LB agar plates that contained the appropriate antibiotic using a glass spreader. For each sample at least two plates were inoculated, a control plate with no antibiotic selection and a plate containing antibiotics for the selection of successful transformants.

2.1.10 Transformation of *L. plantarum* strains

Competent *L. plantarum* cells suspended in 30% PEG were transformed using electroporation. This was done by adding 2.5 µL or 5 µL of DNA between 50 ng/µL and 150 ng/µL to a pre-chilled electroporation cuvette (0.2 cm gap) (Bio-Rad)(Hercules, CA, United States), then adding a 50 µL aliquot of chilled competent cells. Cells and DNA were mixed by vigorous tapping then subjected to electroporation at 2.5 kV, 200 Ω of parallel resistance and a capacitance of 25 µF. Time intervals were recorded for each sample as these varied, depending on DNA concentrations. After each electroporation, each sample was quickly re-suspended in chilled MRS media, then incubated at 30°C

for 2.5 hours before plating on selective MRS plates. The plates were then incubated for 36-48 hours at 30°C before colony counting.

2.2 DNA Methods

2.2.0 Agarose gels

Agarose gels for DNA analysis were made by dissolving the required weight of Agarose (Apex) into 1 x TBE and melting it in a microwave. The solution was allowed to cool to ~40°C before being poured into a gel mould. When set, this solidified gel was transferred to a gel rig (Bio-Rad Sub-Cell MINI GT) and submerged in 1 x TBE Buffer so that the gel was covered to a depth of approximately 5 mm with the solution.

Before loading, 10 µL aliquots of DNA solutions were diluted to between 40 – 100 ng/µL with dH_2O. 6x DNA loading dye was made with 2.5 mg Bromophenol Blue, 2.5 mg Xylitol Green, 0.4 g sucrose and made to 1 mL with dH_2O.

Gel lanes were loaded in 6 µL aliquots comprised of 5 parts diluted DNA solution and 1 part loading dye. This gel rig was connected to a PowerPac™ Basic volt box (Bio-Rad) to control voltage and run-time. 0.6% (w/v) agarose gels were used for pDNA and gDNA analysis and were run at 70 V for 2 hours at 4°C. 1.0 - 1.5% (w/v) gels were used for PCR product analysis and were run at 80 V for 45 minutes either at 4°C or at room temperature. Once run, gels were removed from the apparatus and placed in a solution of 0.5 mg/mL ethidium bromide for 15 minutes, then de-stained in distilled H_2O for five minutes with gentle agitation before being visualised under UV light in a UV-Trans Illuminator Gel Doc™ XR^+ System (Bio-Rad). To obtain an estimate of the size

of the DNA bands an Invitrogen (Carlsbad, CA, United States) 1 kb+ ladder was used as a molecular size standard.

2.2.1 PCR

PCR mixtures were comprised of 1x PCR reaction Buffer (Invitrogen), 1 mM $MgCl_2$, 0.5 mM dNTPs (New England Biolabs)(Ipswich, MA, United States), 1 unit of *Taq* Polymerase (Invitrogen) and 40 pM of each primer (forward/reverse) from each pair involved in the reaction (Table 2.2.1). This was then made up to a volume of 25 µL with dH_2O. Temperatures used varied depending on the primers used. Protocol I was used for *asmA* amplifications, protocol II for pTRK669 amplifications and protocol III for *asm* operon products to be sent for sequencing (Appendix 1.1). A Biometra (Goettingen, Germany) TGradient thermocycler was used for all reactions. Primers used are shown in Table 2.2.1 and were a kind gift from M.L Patchett and S. Ahn

Primer Name	Primer Sequence	Source
AsmA_Nde_f	GGTCTTCATATGAAGAATAAAACAACTCATACCG	(Ahn, 2012a)
A1_gDNA_Gtase_Direct	TTCTTCGTGAACCATTTCTTGTA	(Patchett, 2012a)
ASM1_R1	GTAGTTTTTTTGCTTTTACTGT	(Patchett, 2012b)
ASM1_F1	AGATGGTGAAATTATWWCAAG	(Patchett, 2012b)
ASM1_R2	AAAAAGGGTTATACTCATCGT	(Patchett, 2012b)
ASM1_F2	CTTTATCATGATGGATATGATA	(Patchett, 2012b)
A1_DirectRLastplus1	CACTAAGCTTGAATGAATAGT	(Patchett, 2012b)
A1_DirectFLastPlus1	CCAATATGTTAAACAAGTTGT	(Patchett, 2012b)
A1_DirectRlastplus2	AAATTAACGATTGAAGACTGT	(Patchett, 2012b)
A1_DirectFLastplus2	CGATACTTAATGGTTGGAGC	(Patchett, 2012b)
gcc_asm_Rev	TACCAACACCATGCAGGTTTTCC	(Patchett, 2012a)
gcc_asm_Fwd	TTGGGAACATTGGGAATATCATCAAC	(Patchett, 2012a)
A1_FinalPCR Rev	GCGATTGTAAGACTGTCAACACTTT	(Patchett, 2012b)
A1_FinalPCR Fwd	CAAGCTCCACAACCACATTATC	(Patchett, 2012b)
TRKCmSeqF	TCAGATAGGCCTAATGACT	(Patchett, 2013a)
TRKCmSeqR2	TTCATTCAGTCATCGGCTTT	(Patchett, 2013a)

Table 2.2.1 PCR primers used and their source
ASM1_F1 contains ambiguous residues noted as “W” that is a mixture of A or T bases

2.2.2 Restriction digests

Restriction enzymes used included *Afe*I, *Msc*I, *Pst*I, *Sal*I and *Sap*I from New England Biolabs, these were chosen because for each enzyme there was only one probable cut site in the ASM1 plasmid. The reactions were carried out using the buffers recommended by the manufacturer with 400 ng of DNA, 1X Buffer, 1U Restriction Enzyme and made up to 20 µL with dH_2O. The digest was incubated for 1 hour at 37°C and then stored at 4°C until used.

2.2.3 Southern blotting

Southern blotting hybridisation was used in an effort to identify the plasmid containing the *asm1* gene cluster in the *L. plantarum* A-1 plasmid profile. pDNA extracted from *L. plantarum* A-1 (Section 2.2.6) was digested with *Msc*I, then separated by gel electrophoresis in 0.7% agarose at 30 V for 26 hours at 4°C. A positive control (the *asmA* PCR product) was added at 24 hours to ensure the band of interest did not over-run the gel. The resulting agarose gel was then stained with ethidium bromide and visualised using a UV-Trans-Illuminator Gel Doc™ XR^+ System (Bio-Rad). Once the appropriate band separation had been confirmed, the bands on the gel were transferred to a GE Healthcare (Little Chalfont, United Kingdom) Amersham Hybond-N+ kit (Nylon Membrane) using passive diffusion with a high-salt transfer buffer, according to the manufacturer's instructions.

An *asmA* PCR product to which a thermostable alkaline phosphatase (GE Healthcare) had been covalently linked as *per* manufacturer's instructions was used as the probe. The nylon membrane was hybridised to the probe overnight, following the

manufacturer's instructions, then subjected to stringency washes detailed in the manufacturer's instructions**.** Detection was achieved using a Fujifilm Typhoon FLA-9000 by applying 700 µL of the ECF substrate to the membrane, incubating at RT for 60 seconds then visualising using excitation at 532 nm as this provided the clearest image.

2.2.4 Sequencing

PCR products were submitted to Massey Genome Services (Massey University, Palmerston North) with the associated primers for sequencing.

L. plantarum A-1 genomic DNA extract was submitted to Massey Genome Services for Next Generation Sequencing (NGS) using the Illumina (San Diego, CA, United States) MiSeq instrument. Contigs from the sequencing were assembled *de novo* using CLC Genomic Workbench (Version 6).

2.2.5 Bioinformatics

The NCBI ORF finder was used to detect open reading frames in the sequence data. These ORFs were then analysed using BLASTX 2.2.29+ or BLASTP 2.2.29+ (Altschul *et al.*, 1990; Altschul *et al.*, 1997). Conserved domains identified in blast were referenced by number from the NCBI Conserved Domain Database (CDD) (Marchler-Bauer *et al.*, 2013).

Sequence similarity dot-plots were made with YASS Dot-plot generator (Noé & Kucherov, 2005).

Plasmid maps were drawn with the aid of GENtle (Manske, 2007).

Peptidase_C39 domains for large-scale alignment were acquired from the Sanger pfam database (Punta *et al.*, 2012).

Protein alignments were done using T-COFFEE (Notredame *et al.*, 2000).

Phylogenetic trees were predicted using MUSCLE (Edgar, 2004) and were then visualised and edited for printing with FigTree 1.4.0 (Rambaut, 2006).

2.2.6 Plasmid DNA extraction from *Lactobacillus* species.

Extraction of plasmid DNA (pDNA) from *Lactobacillus* species was achieved using the method of O'Sullivan & Klaenhammer (1993) in which a 100 µL of an overnight culture of *L. plantarum* was used to inoculate 15 mL of sterile MRS media. This was then grown at 30°C to $O.D._{600\ nm}$ of ~2.0, and the cells pelleted by centrifugation (4,000 x g at 4°C for 15 minutes) then re-suspended in 200 µL of 25% sucrose containing 40 mg/mL hen egg white (HEW) lysozyme (Serva Electrophoresis)(Heidelberg, Germany). This mixture was transferred to a 1.5 mL microcentrifuge tube and incubated at 37°C for 45 minutes. 400 µL of alkaline SDS solution (3.0% SDS + 0.2 N NaOH) was then added to the tube and mixed by inversion before the tube was incubated at room temperature for 7 minutes. 300 µL of ice-cold 3M potassium acetate (pH 4.7) was then added to the tube, mixed by inversion then subjected to centrifugation at 17,000 x g for 15 minutes at 4°C. The supernatant was decanted and 650 µL of propan-2-ol added and mixed by inversion before centrifugation at 17,000 x g for 15 minutes at 4°C. Following this, the supernatant was removed and the pellet was re-suspended in 320 µL distilled H_2O + 200 µL 7.5 M ammonium acetate containing 0.5 mg/mL ethidium bromide. Once the pellet was completely resuspended, 350 µL of phenol : chloroform :

Isoamyl Alcohol (IAA) (25:24:1) was added to the tube and the contents thoroughly mixed before being subjected to centrifugation at 17, 000 x g for 5 minutes at room temperature. The upper phase of the solution was transferred to a new tube, checking to make sure no precipitate was present. If some of the interface layer was accidently transferred, the previous step was repeated before 1 mL of -20°C absolute ethanol was added. The contents were mixed by inversion, centrifuged at 17,000 x g at 4°C for 10 minutes to pellet the DNA and the supernatant removed by pipette to avoid pellet disturbance. The pellet was then washed twice with 70% ethanol at room temperature, and the DNA was re-isolated by centrifugation, as before. The pellet was then air-dried and re-suspended in 49 µL of TE buffer (10 mM Tris/HCl, 1 mM EDTA, pH 8.5) to which 1 µL of 10 mg/mL RNAse was added. DNA was stored at 4°C for up to 6 months.

2.2.7 Plasmid DNA extraction from *E. coli*

3 mL overnight cultures of plasmid bearing *E. coli* XL1 strains in LB media with the relevant selective antibiotic agent were harvested by centrifugation at 17,000 x g for 10 minutes. Plasmid extraction was carried out using a High Pure Plasmid Isolation Kit (Roche)(Basel, Switzerland) according to the manufacturer's instructions. The resulting plasmid DNA was stored in TE buffer for up to 6 months at -80°C.

2.2.8 Genomic DNA extraction from *Lactobacillus* species.

Extraction of genomic DNA (gDNA) from *Lactobacillus* species was carried out using a method that utilised the Wizard Genomic DNA purification Kit (Promega)(Madison, WI, United States) modified from Stepper (2009). Briefly, a 1mL overnight culture of the

target *Lactobacillus* strain was used to inoculate 35 mL of sterile MRS media, which was then incubated at 30°C until the $OD_{600\ nm}$ reached 2.0. Cells were harvested by centrifugation at 4,000 x g for 10 minutes at 4°C, and the pellet resuspended in 2 mL of ice-cold 50 mM EDTA (pH 7.8) containing 10 mM Tris-HCl (pH 8). The 2 mL cell suspension was equally divided between two 1.5 mL microcentrifuge tubes and 400 µL of 40 mg/mL lysozyme added to each tube, resulting in a final lysozyme concentration of 11.4 mg/mL. Both tubes were then incubated at 37°C for 1 hour, mixing by inversion every 10 minutes. The suspension was then centrifuged at 17,000 x g for 1 minute and the supernatants quickly removed by pipette. The pellets were re-suspended in 800 µL of Nuclei Lysis Solution (Wizard Kit) at RT, and then subjected to 3 cycles of freeze/thaw by storing in the -80°C freezer for 5 minutes before placing in an 80°C heat block for 3 minutes. The tubes were cooled to ~35°C before the addition of 5 µL of 4 mg/mL RNAse solution and incubation at 37°C for 45 minutes, mixing by inversion every 15 minutes. The tubes were then chilled on ice before the addition of 270 µL of protein precipitation solution (Wizard Kit) at RT and mixed thoroughly by inversion. The tubes were then incubated on ice for 5 minutes before insoluble material was pelleted by centrifugation (17,000 x g, 5 minutes at RT). The supernatant was collected and re-centrifuged at 17,000 x g for 3 minutes. This step was repeated until the resulting supernatant of ~800 µL was completely transparent. At this stage 620 µL of propan-2-ol was added and the tubes were gently mixed by inversion until DNA threads were visible. These were separated from the solution by centrifugation at 17,000 x g for 5 minutes after which the pellets were washed twice with 70% ethanol. The resulting DNA pellet was re-suspended in TE buffer and stored at 4°C for up to 6 months.

2.2.9 DNA quantification

Final products from DNA extractions and purifications were quantified using a Nanodrop ND1000 (Thermo Scientific).

2.2.10 Plasmid isolation from agarose gels

Digested pDNA from A-1 (sections 2.2.3 and 2.2.7) was separated on an agarose gel (using the method described in section 2.2.1) except a reduced voltage and elongated run-time were used (30V, 7 hours) to increase resolution and the separation of individual bands. Before visualisation under UV light the gel was marked in two locations by nicking the agarose with a sterile scalpel, these physical marks were used to determine the location of the band of interest by matching them with the location of the band in the picture of UV fluorescence. The band of interest was then excised using a sterile scalpel and placed into a sterile microcentrifuge tube. This method of band excision was used as the alternative (excision whilst visualising under UV) can cause degeneration of the DNA in question and can make it harder to purify from the agarose.

After gel excision the DNA was purified from the agarose using a Zymoclean™ Gel DNA Recovery Kit (Zymo Research)(Irvine, CA, United States) with a slight modification to the manufacturer's instructions in that, during the elution step, the TE Buffer was added to the membrane and incubated at RT for 30 seconds before being subjected to centrifugation to increase the yield. The eluted product was then quantified (As described in section 2.2.10).

2.3 Protein Methods

2.3.0 Protein purification

50 mL of MRS media was inoculated with *L. plantarum* A-1 from a glycerol deep and left to grow overnight at 25°C, without shaking. The overnight culture was used to inoculate 6 L of MRS which was incubated for four days at 25°C under static conditions. The cells were subjected to centrifugation at 6,000 x g for 30 minutes and the supernatant was retained. At this stage the pH of the supernatant was measured and reduced to approximately pH 4.5 using 0.1 M formic acid if necessary.

50 mL of SP-Sephadex (GE Healthcare) (charged with Cl^- and equilibrated in 50 mM Sodium formate at pH 3.8) were mixed with the 6 L of supernatant and stirred overnight at room temperature. The mixture was then packed into a Bio-Rad Econo-Column® (5 x 20 cm) by gravity and, once almost all of the supernatant had drained through, was eluted stepwise using 1 L of 20 mM sodium formate (pH 3.8), then 1 L of 20 mM MOPS (pH 7.2), followed by 0.5 L 20 mM Ammonium Bicarbonate (AmBic) and finally 0.5 L 20 mM AmBic in 60% Acetonitrile (MeCN). 80 mL fractions were collected during the entire process and from each elution a 10 µL sample was taken, half of which was spotted onto an indicator plate which was incubated overnight to check for activity, the second half of the aliquot was retained in case the initial indicator plates did not grow.

Once fractions with ASM1 were identified (in AmBic + 60% MeCN) the fraction was frozen at -80°C and then lyophilised using a Dura-Dry MP Freeze-dryer. The lyophilised powder was dissolved in the minimum amount of distilled water and adjusted to pH

4.0 using 0.1 M acetic acid. The solution was then filtered using a Minisart 0.45 µm syringe filter (Sartorius)(Goettingen, Germany) before being subjected to Reversed Phase High Pressure Chromatography (RPHPLC)**.** The column used was a Jupiter C18 10 x 250mm, particle size 5 µM, pore size 300 Å. Solvents used were A: 0.1% Trifluoroacetic Acid (TFA) in dH_2O B: 0.08% TFA in MeCN. The flow rate used was 4 mL / minute. Protein was eluted from the column using the gradient shown in Table 2.3.1

Stage	Time (Start)	Time (Finish)	Start %A	Start %B	Finish %A	Finish %B
Normalise	0	5	80	20	75	25
De-salt	5	15	75	25	75	25
Protein Elution	15	35	75	25	50	50
Wash	35	42	0	100	0	100
Normalise	42	60	80	20	80	20

Table 2.3.1 HPLC Purification protocol buffers
% of each buffer (A and B) used in each step of the RPHPLC purification of ASM1; concentrations were altered over the course of the run-time of each step.

2 mL samples were injected manually using an external Rheodyne valve and peak fractions were collected manually. Fractions were tested for activity using indicator plates as previously described and the active fractions were pooled, lyophilised then stored at -20°C.

Liquid fractions were pooled then lyophilised and the powder weighed to determine the final yield. Solution concentrations were verified using absorbance at 205 nm and 280 nm, as measured using the Varian Cary™ 300 Bio.

2.3.1 Bradford assays

Determination of supernatant protein concentration was done using the method established by Bradford (Bradford, 1976). 1 mL of Bradford solution was added to 100

µL aliquots of a range of concentrations of a Bovine Serum Albumin (BSA) and two 100 µL aliquots of 2-fold diluted supernatant from the culture used for the protein purification (See 2.3.1). These samples were incubated at room temperature for 10 minutes and then the absorbance at 595 nm was measured using a BioRad SmartSpec™ Spectrophotometer.

BSA standards were used to create a standard curve and the protein concentration of the supernatant was calculated from this standard curve.

2.3.2 SDS-PAGE protein separation

To determine protein size and purity the purified proteins were run on a tricine gel based on the method described by Schägger and Von Jagow (1987). Separating gels were 16.5% acrylamide and bisacrylamide (T) and 3% crosslinker (C) relative to the total concentration of T (Derived from a premixed Merck (White House Station, NJ, United States) solution). Stacking gel was 4% T and 3% C, no spacer gel was used.

Once cast, gels were placed in a Bio-Rad Mini PROTEAN® 3 CELL gel holder and Anode and Cathode Buffers, as described in Schägger and Von Jagow (1987), were added to the holder. The gel was loaded with approximately 130 µg of protein in 25 µL of a solution comprised of equal parts dH_2O and 2x Tricine loading dye, as described in Schägger and Von Jagow (1987). The protein standard used for size comparisons was the Invitrogen SeeBlue® Plus2 Pre-stained Protein Standard.

Electrophoresis was carried out using controlled currents at room temperature, as shown in Table 2.3.2.

Stage	Current (mA)	Time (min)
Stacking	25	30
Separating	45	90

Table 2.3.2 Protein gel electrophoresis protocols

Following electrophoresis, the gel was stained overnight with Coomassie Blue dye (50% MeOH, 10% HoAC, 40% dH_2O with 0.25% Coomassie Blue R-250) with gentle agitation. Destaining was carried out using 5% MeOH, 7.5% HoAC and 87.5% H_2O in three half-hour washes, decanting the destaining solution every half-hour and re-filling.

Chapter 3 - ASM1

3.1 ASM1 Purification

Purified ASM1 peptide was purified from the cell-free supernatant of a *L. plantarum* A-1 culture with a total protein content of 1.7 mg/mL (Determined by Bradford Assay (section 2.3.0) through Ion-exchange column purification and RP-HPLC analysis and sampling, as described in section 2.3.1.

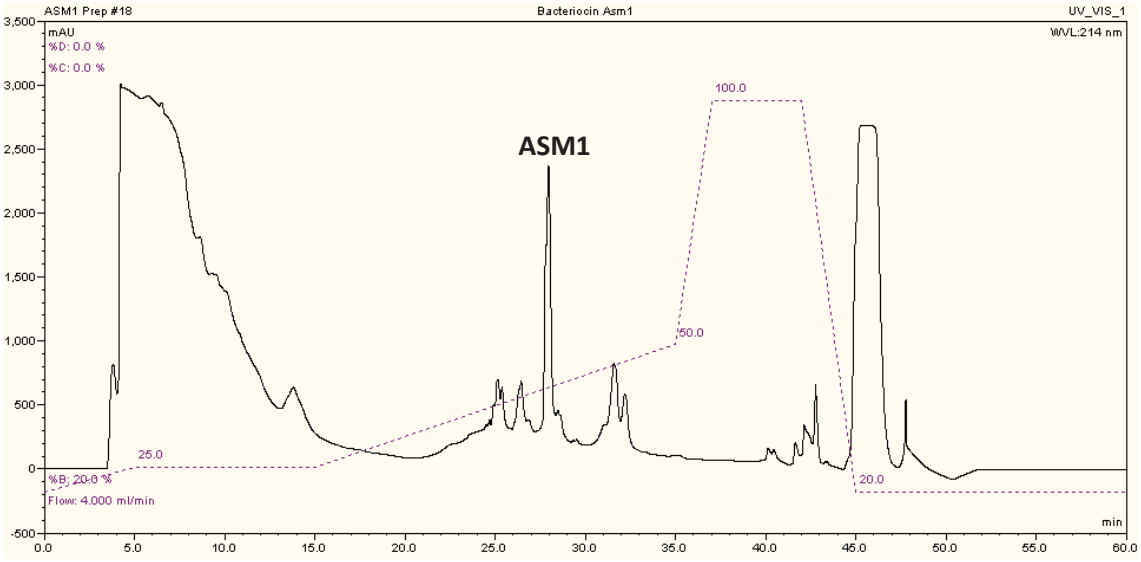


Figure 3.1.1 RP-HPLC profile with absorbance peaks
Absorbance of UV light at 214 nm recorded. ASM1 peak at 27 minutes (as labelled) was collected manually. Image from Chromeleon chromatography data system.

This process resulted in the extraction of 9.1 mg of ASM1 protein powder. Purity was determined based on the sharpness of the peak in the HPLC profile (Figure 3.1.1).

The purified product was then run on a tricine gel and compared to the homologous bacteriocin GccF (as described in 2.3.2). The gel showed that the ASM1 peptide was the only protein present in the solution, supporting the single peak from the HPLC purification (Figure 3.1.1). The gel shows a large, darkly stained band at the correct molecular mass for each bacteriocin that would initially suggest an excess of protein

had been loaded. However it has been noted that the purification method used will often leave detergent in the solution, which can cause the band in the tricine gel to diffuse slightly and bind the Coomassie blue dye more strongly.

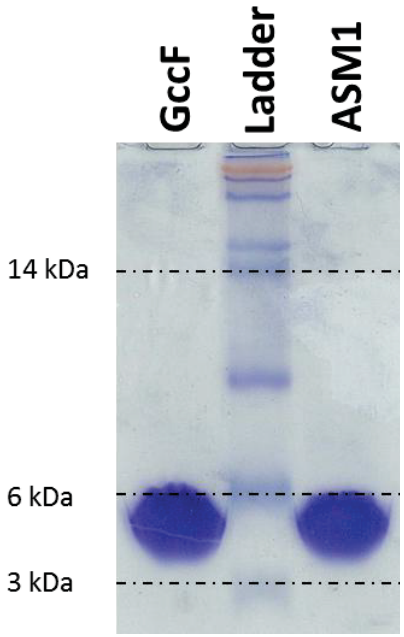


Figure 3.1.2 SDS-PAGE gel showing relative sizes of ASM1 and GccF
Tricine gel stained with Coomassie Brilliant Blue, protein standard (L). ASM1 migrates slightly further than GccF, as predicted by its lower mass.

Although the proteins have slightly different masses (GccF is 5199 Da (Stepper *et al.*, 2011) and ASM1 is 5042 Da (Hata *et al.*, 2010)), the GccF has several serine residues where ASM1 has glycine residues, and the increased density of polar residues in GccF could enhance the slight differences between the electrophoretic mobility of the two peptides.

3.2 ASM1 Characterisation

3.2.0 Data Processing and Presentation of Growth Curves

The following data sets concerning the growth of *Lactobacillus* in liquid media have been simplified for ease of interpretation in three different ways. The first processing step involved the elimination of extra time points. The CARY 300 Bio used to measure

the growth of the cultures was set to read data every 2 minutes for the duration of the run. This read frequency was used to determine any immediate changes caused by the addition of the bacteriocin but the data were smoothed around every 10th data point to allow for easier interpretation of graphs.

The second processing step was to normalise the starting OD for each run to a common value. This was done by averaging the starting OD across all the data in each group (all of the 30°C runs at pH 6.5, for example) and then using that value to adjust the curve (each data point was adjusted so that it started from the averaged optical density value). This allowed easier visualisation of the effects on growth rate, reduced overcrowding of the graph and also prevented small differences in OD from being obscured by different starting ODs.

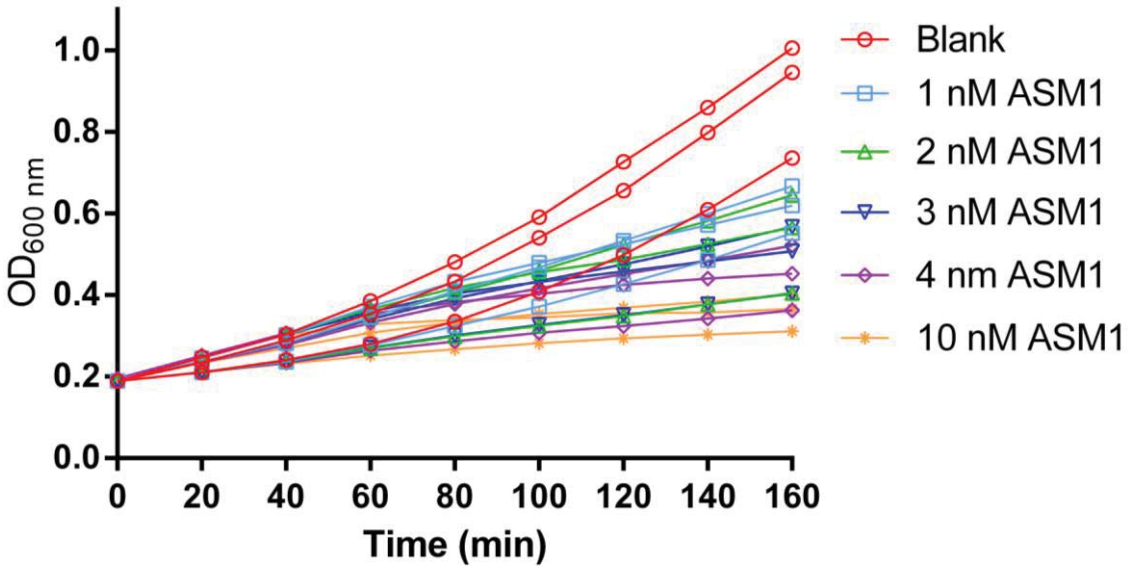


Figure 3.2.1 ASM1 IC_{50} at 30°C and pH 6.5 without data averaging or smoothing
Growth curves of *L. plantarum* ATCC 8014 treated with various concentrations of purified ASM1 added at 30 minutes. No data averaging,

Finally, the replicates were collated after processing and averaged to show the overall trend for each treatment but without overcrowding the graph, making it easier to interpret the overall trend. Figure 3.2.1 shows the IC_{50} of ASM1 at 30°C and pH 6.5

with all three processed and refined data sets, showing the extent of the variation amongst samples even under identical conditions. Such variation is consequence of the method used and is caused by a number of uncontrollable factors including the age of the sensitive strain (the duration of the overnight culture before subculturing), the duration of the subculture and minor variations in the starting OD, before normalisation.

3.2.1 IC_{50} of ASM1

Using *L. plantarum* ATCC 8014 as the indicator strain, live growth assays were done to determine the concentration required to inhibit the growth of half of the population (the IC_{50}) determined by reductions in growth rates (as described in 2.1.5). A variety of environmental conditions were used to determine the effects of heat and pH on the ASM1 peptide. The "baseline" for these was the determination of the IC_{50} at pH ~6.5 at 30°C. Results obtained (Figure 3.2.2) allowed for the calculation of the IC_{50} in these conditions to be slightly less than 2 nM.

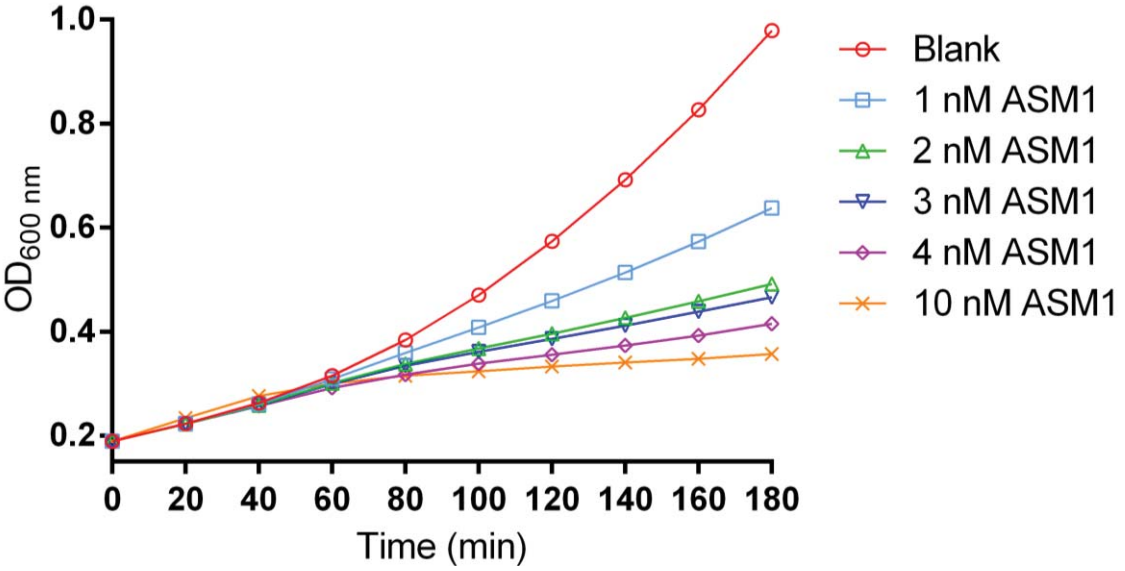


Figure 3.2.2 ASM1 IC_{50} at 30°C and pH 6.5
Growth curves of *L. plantarum* ATCC 8014 treated with various concentrations of purified ASM1 added at 30 minutes. Data averaged from three runs.

Concentrations of ASM1 higher than 10 nM almost completely stopped the growth of the target strain within the time frame of the assay and doses as low as 1 nM caused a significant reduction in the growth rate (Figure 3.2.2).

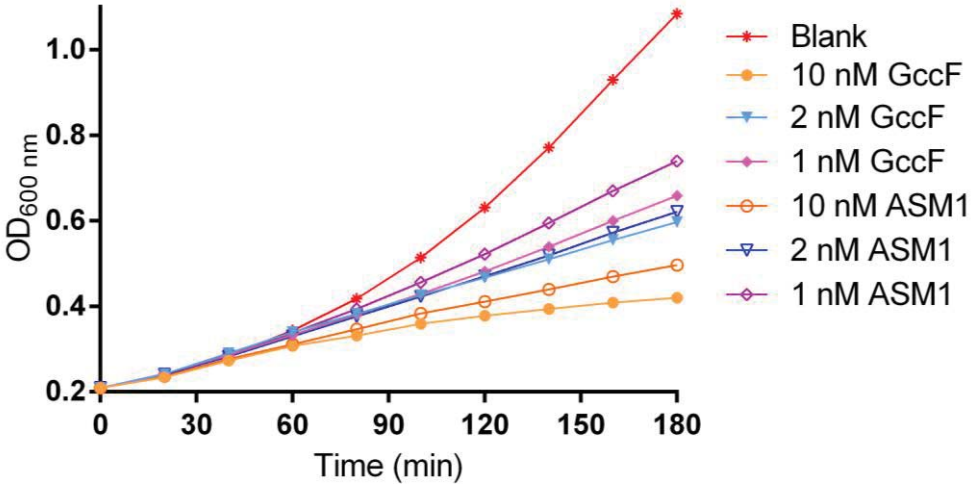


Figure 3.2.3 Comparison of ASM1 and GccF at 30°C and pH 6.5
Growth curves of *L. plantarum* ATCC 8014 treated with various concentrations of purified ASM1 and GccF. Both added at 30 minutes. Data averaged over three to four runs.

The results of a direct comparison agree with previous results in establishing the IC_{50} of GccF (Stepper *et al.*, 2011) and show that both bacteriocins have an IC_{50} of approximately 2 nM. However, it appears that ASM1 peptide is less active at higher and lower concentrations (Figure 3.2.3). This activity difference is most likely caused by sequence differences in the flexible C-terminal tail of ASM1.

3.2.2 Environmental effects

Though the stability effects of pH and temperature on the ASM1 protein had been examined by Hata *et al.* (2010), the effects of different environmental conditions on the activity of the bacteriocin had not been determined. To do this, the incubation temperature and media pH were adjusted and the effects of ASM1 on an indicator strain grown in these conditions were determined. However, small changes in the pH

often caused a change in the growth rate of the strain being tested. The reasons for this were unclear, although it is possible that alterations to the environmental salt level could have been the cause.

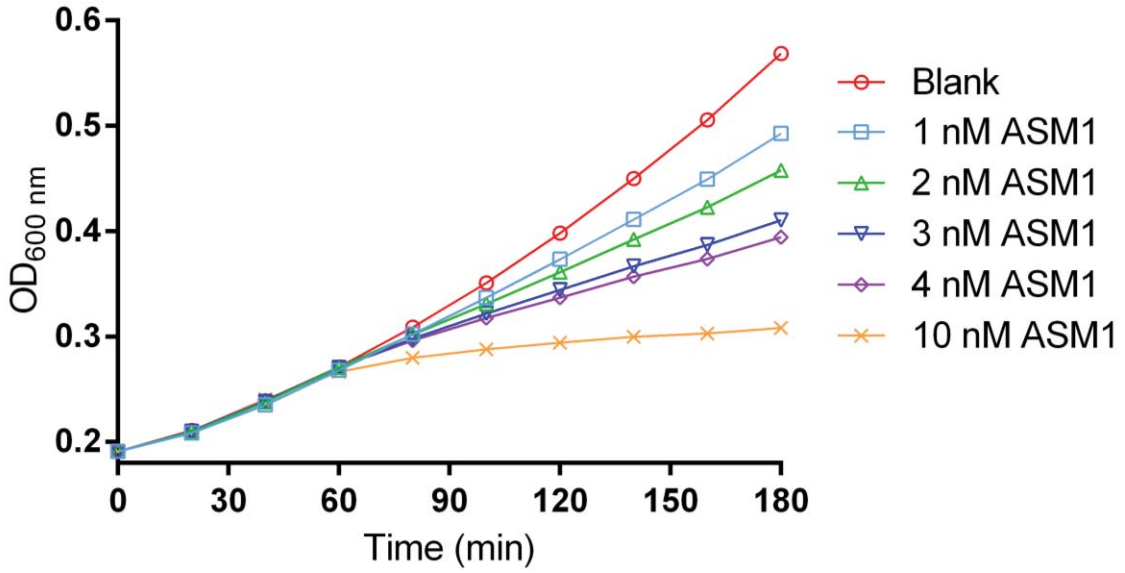


Figure 3.2.4 ASM1 IC_{50} at 30°C and pH 5.0
Growth curves of *L. plantarum* ATCC 8014 treated with various concentrations of purified ASM1 added at 30 minutes. Sensitive strain grown in pH adjusted media overnight to avoid exposing it to a rapid change in pH. Data averaged from three runs.

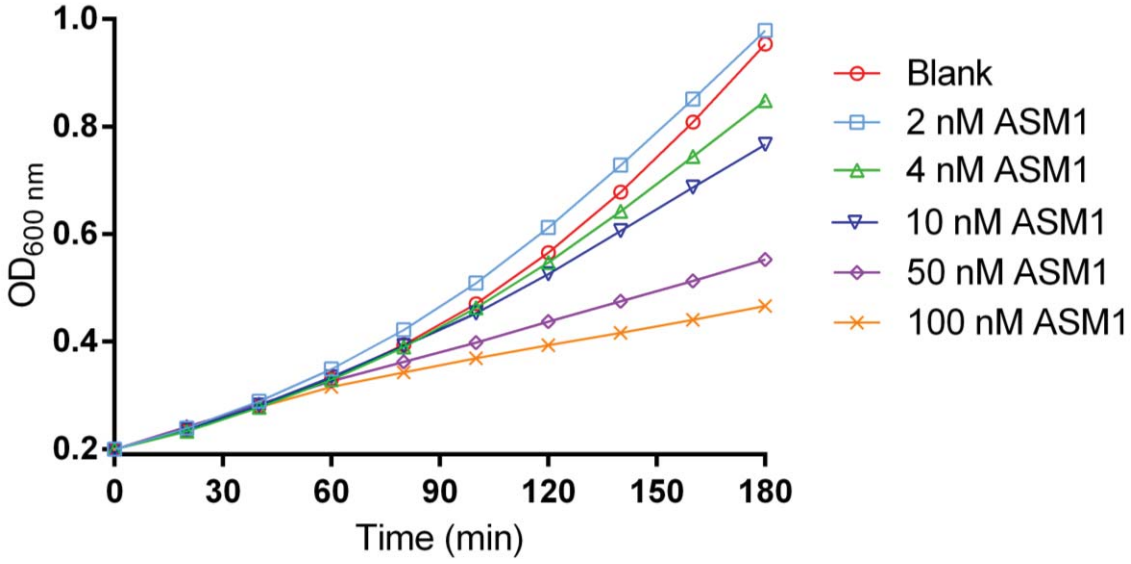


Figure 3.2.5 ASM1 IC_{50} at 30°C and pH 7.5
Growth curves of *L. plantarum* ATCC 8014 treated with various concentrations of purified ASM1 added at 30 minutes. Sensitive strain grown in pH adjusted media overnight to avoid exposing it to a rapid change in pH. Data averaged from three runs.

Despite the effect on growth rate, it was apparent that the alkalinity affects the activity of ASM1. Although the IC_{50} at pH 5.0 is largely unchanged from that at pH 6.5 (Figure 3.2.4), it is increased to approximately 25 nM at pH 7.5 (Figure 3.2.5). This is unusual as the *L. plantarum* A-1 strain is found in a pH neutral corn fermentation (Hata *et al.*, 2010), which appears to render the ASM1 less effective than it is at more acidic pH levels. The reason for this is unclear but could reflect the fact that, when compared to the natural secretion level, the concentrations tested here are quite low. In soft agar assays, the normal activity of the *L. plantarum* A-1 supernatant appears to be approximately equivalent to an equal volume of 1 µM purified ASM1, 10 fold higher than the highest concentration tested at pH 7.5 (Figure 3.2.5). This would suggest that the activity of ASM1 in the pH neutral environment is adequate in reducing competition but if the population is already occupied or comes under more intense competition (as a higher pH would indicate, based on higher acidic waste products) then the bacteriocin becomes more effective at eliminating competition.

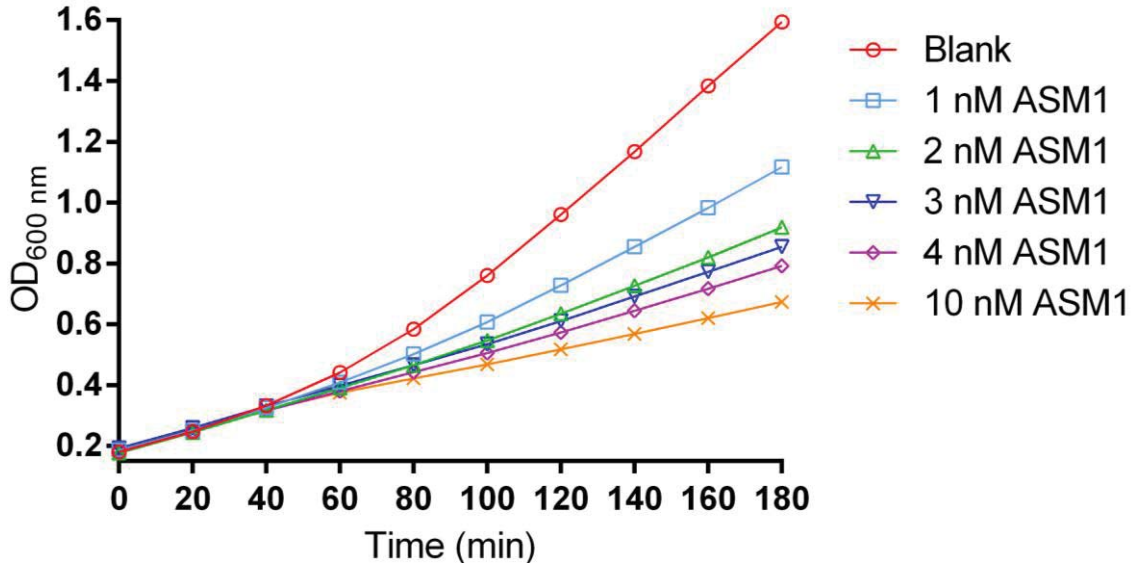


Figure 3.2.6 ASM1 IC_{50} at 37°C and pH 6.5
Growth curves of *L. plantarum* ATCC 8014 incubated at 37°C and treated with various concentrations of purified ASM1 added at 30 minutes. Data averaged from five runs.

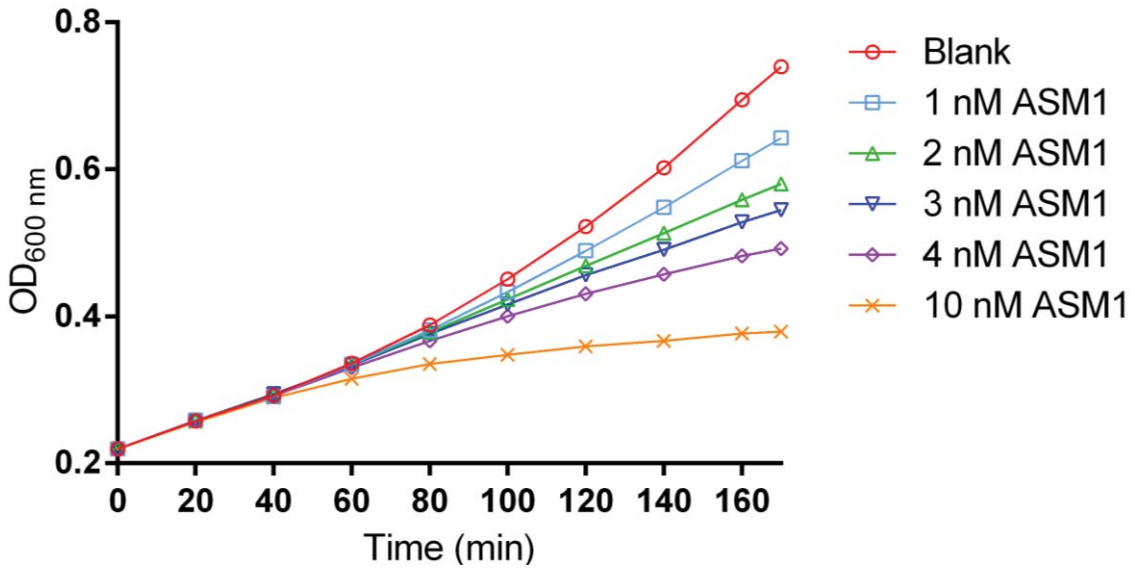


Figure 3.2.7 ASM1 IC_{50} at 25°C and pH 6.5
Growth curves of *L. plantarum* ATCC 8014 incubated at 25°C and treated with various concentrations of purified ASM1 added at 30 minutes. Data averaged from three runs.

Changes in the incubation temperature not only affected the base growth rate of the sensitive strain, but also provided relatively clear changes to the activity of the bacteriocin. Increases in the temperature up to 37°C lowered the IC_{50} to ~1 nM (Figure 3.2.6). Despite this, the effects of higher concentrations of ASM1 appeared to have a lesser effect, as shown by the higher growth rate of the sample treated with 10 nM ASM1 at 37°C when compared with the same treatment at 30°C (Figure 3.2.2). On the other hand, lowering the incubation temperature to 25°C caused an increase in the IC_{50} to ~4 nM, reducing the effects of small doses of the ASM1 protein but still causing dramatic reductions in growth rates at the higher doses of 10 nM (Figure 3.2.7).

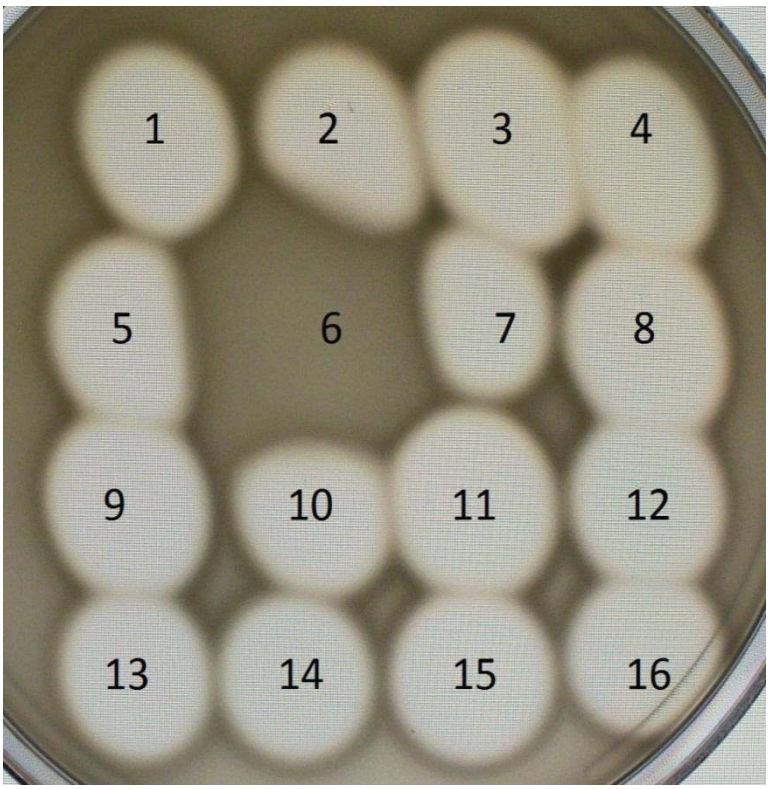
This change in activity with temperature is possibly merely a consequence of an increase in the rate of growth. Because the cells are growing faster at higher temperatures they will take up more GlcNAc and other nutrients from the media upregulating the expression GlcNAc-PTS transporters, the putative binding targets for ASM1. This would have two possible outcomes; firstly it would provide more binding

sites for the ASM1 GlcNAc moieties, so that small concentrations bind more rapidly causing an increase in the reduction of growth rate. Secondly, the large number of binding sites available means that it is less likely that a low concentration of ASM1 could occupy all of the available sites. Figure 3.2.4 shows that a low dose of 1 nM causes a reduction in growth of the target strain by nearly 50% whilst the highest concentration tested, 10 nM caused a significant reduction in growth but did not halt it, indicating that many cells are still growing, albeit at a reduced rate. This difference at higher growth rates is apparent when compared to the effects of the equivalent dose at 30°C (Figure 3.2.2).

3.2.3 GlcNAc Protection

It was shown that the activity of GccF can be severely attenuated by the presence of N-Acetylglucosamine (GlcNAc), to the point where previously sensitive cells are no longer affected by the bacteriocin (Stepper *et al.*, 2011). While the mechanism for this inhibition is not known, it could be due to the free GlcNAc competing with the GlcNAc moieties on the bacteriocin for potential binding sites on the target of GccF. As ASM1 is a homologue of GccF it was probable that GlcNAc would have a similar effect on ASM1 activity.

Plate assays (as described in section 2.1.6) treated with a mixture of ASM1 and several different monosaccharides showed that the inhibitory activity of ASM1 is attenuated by the addition of GlcNAc to the media. They also showed that the protection is GlcNAc specific as none of the other fourteen monosaccharides provided any protection against ASM1 (Figure 3.2.8). These are identical results to those reported from earlier studies with GccF and are a good indicator that the GlcNAc moieties on ASM1 are essential for its function.



Key: ASM1 mixed with -

1: Arabinose
2: Fructose
3: Fucose
4: Galactose
5: N-Acetylgalactose
6: N-Acetylglucosamine
7: Glucosamine
8: Glucose
9: Mannose
10: Melezitose
11: N-Acetylmannose
12: Rhamnose
13: Sorbitol
14: Trehalose
15: Xylose
16: H_2O

Figure 3.2.8 Soft agar assay showing specific protection by N-Acetylglucosamine
MRS Media with *L .plantarum* ATCC 8014 suspension treated superficially with various ASM1 + Sugar mixtures, as denoted by Key to the right. Pale areas are those where cell growth has been inhibited, darker areas indicate cell growth. Plate representive of repeat experiments.

Next, the ability of GlcNAc to rescue cells from the inhibitory effect of ASM1 was investigated (section 2.1.6). The addition of a range of concentrations of GlcNAc to cultures already treated with 15 nM ASM1 showed that after sufficient GlcNAc is

added, the cells start to grow again within 40 minutes. For a complete recovery, the final concentration of GlcNAc had to be more than 70,000 times that of ASM1. When present in concentrations only 700 times greater than that of ASM1, GlcNAc leads to only a partial recovery of the cells (Figure 3.2.9).

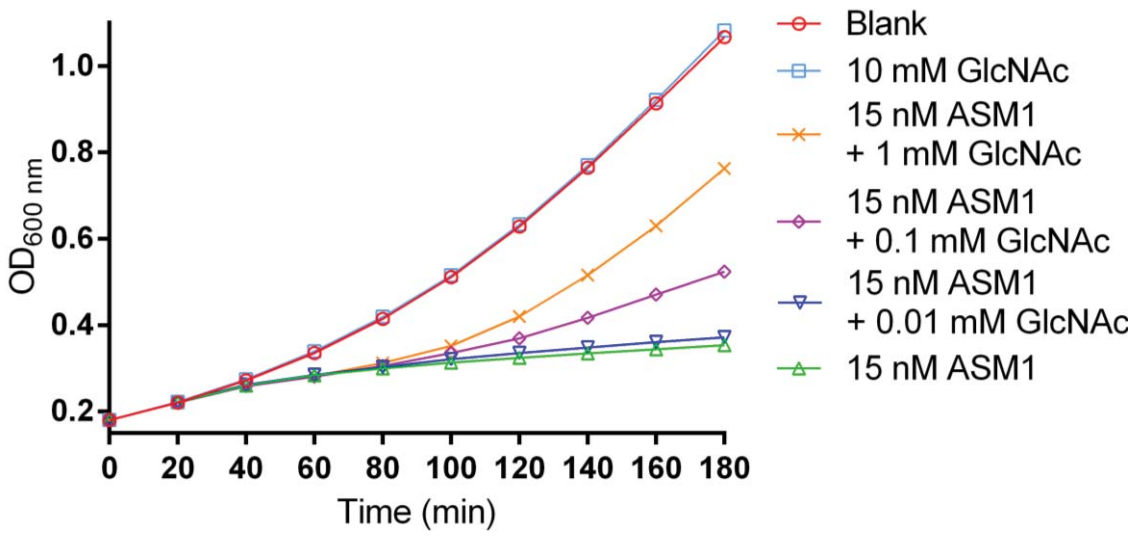


Figure 3.2.9 GlcNAc – induced recovery of *L. plantarum* ATCC 8014
Growth curves of *L. plantarum* ATCC 8014 incubated at 30°C and treated with 15 nM ASM1 and then a range of concentrations of GlcNAc, as shown. ASM1 was added at 30 minutes, GlcNAc was added at 60 minutes. Data averaged from four repeats.

The experiment also showed that ASM1 is bacteriostatic rather than bactericidal; this is because the cells are rescuable. If they were lysed or otherwise killed by ASM1 the optical density would increase slowly as few, if any cells would survive, but this is clearly not the case. It can be seen in figure 3.2.9 that the cells begin to recover within 30 minutes of GlcNAc addition, suggesting that the inhibition and recovery do not take place at a genetic level but rather at the physical level, possibly by interfering with cell-wall processes, possibly by preventing peptidoglycan synthesis.

3.2.4 ASM1 Fragment Analysis

Several different peptides were created by enzymatically dissecting the ASM1 peptide; these were the glycosylated N-terminal domain of ASM1 comprising of residues 1-32, the glycosylated C-terminal domain comprised of residues 33-43 and the *O*-deglycosylated peptide, residues 1-43, with only the *O*-linked GlcNAc moiety removed. These were tested in liquid assays to ascertain which regions of the peptide are required for function and to confirm that the GlcNAc moieties are essential for activity, as they are in GccF.

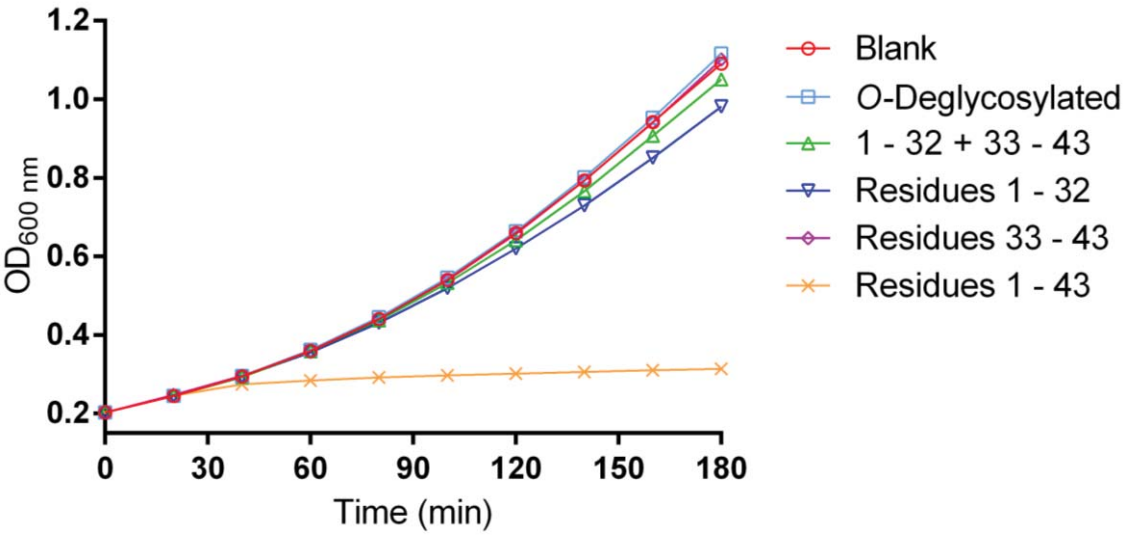


Figure 3.2.10 Residual activity of ASM1 peptide fragments
Growth curves of *L. plantarum* ATCC 8014 incubated at 30°C and treated with solutions of 50 nM of peptide, added at 30 minutes. Fragments are all naturally glycosylated apart from the *O*-Deglycosylated fragment which has one glycosylated residue. Data averaged over three runs.

Liquid assays showed that the *O*-deglyosylated peptide has no activity, supporting the hypothesis that the *O*-GlcNAc moiety is essential for activity as it is in GccF. The glycosylated N-terminus (1-32) has a mild inhibitory effect on the tested strain, but nowhere near the effect that the equal concentration of whole peptide (1-43) exhibits (Figure 3.2.10). The C-terminus (33-43) did not cause any apparent inhibition, despite also having a GlcNAc residue. These results are similar to those reported when

enzymatically-dissected fragments of GccF were tested against the same strain (Stepper *et al.*, 2011). The combination of the N- and C-terminals did not cause a greater reduction in growth than the N-terminal alone, suggesting that the structure of the bacteriocin as a whole is important in its interaction with the target. Though there is some inhibition caused by the mixture, it is most likely that this is due to the N-terminal fragment, rather than to any synergystic effect.

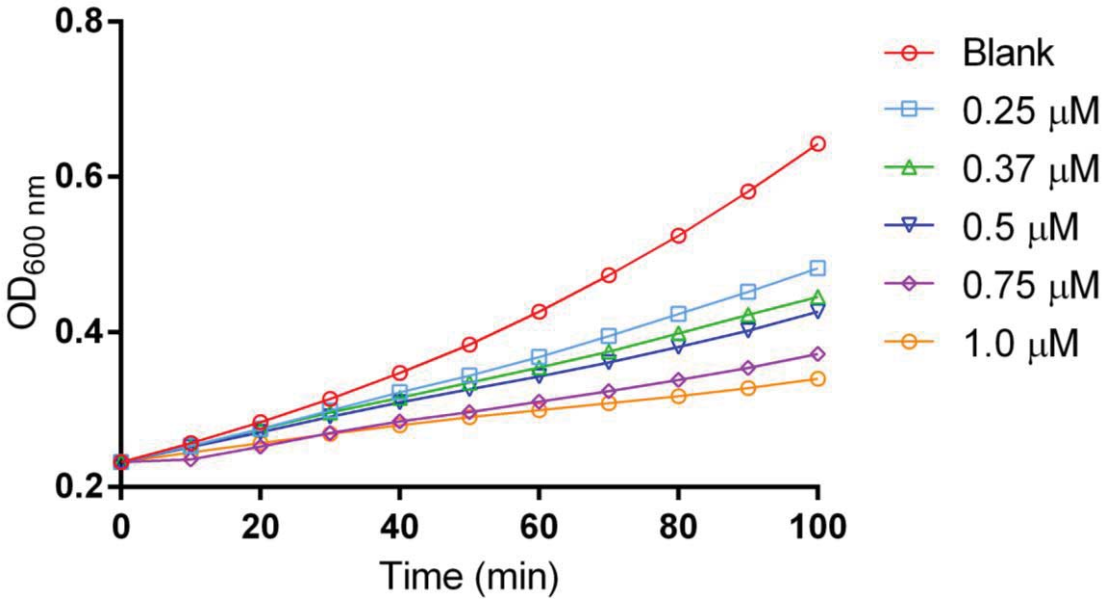


Figure 3.2.11 Residual activity of the N-terminal fragment (1-32) of the ASM1 peptide
Growth curves of *L. plantarum* ATCC 8014 incubated at 30°C and treated with a range of concentrations of the ASM1 N-terminal fragment, added at 30 minutes. Data averaged over three runs.

The levels of inhibition caused by the N-terminal fragment are low in comparison to the full peptide but some residual activity is clearly present, further testing across a range of concentrations allowed for the calculation of an IC_{50} of 370 nM, approximately 180-fold greater than the IC_{50} of the whole peptide (Figure 3.2.11). This is also comparable to the levels of activity shown by the GccF N-terminal fragment which showed an IC_{50} of about 350 nM.

3.3 ASM1 and *Lactobacillus* species

Purified ASM1 protein was tested against a wide array of *Lactobacillus* species using soft agar assays (see 2.1.5). The results of these plate assays show that the ASM1 bacteriocin affects a moderate range of *Lactobacillus* species (Table 3.3.1). Several species that have previously exhibited resistance to ASM1 (Hata *et al.*, 2010) were observed to be sensitive to the bacteriocin in these assays, suggesting that resistance to the bacteriocin is strain-based, rather than species-based. This observation was supported by the variation in resistance between *L. plantarum* species tested.

L. plantarum				***L. lactis***	
KW30	-	Lb790	+++	LF_6a	-
KW30 bac$^-$	++	DSM 2601	+	Corn_7	-
A-1	-	ATCC 8014	++	Corn_28	-
WCFS1	+	DSM 20205	++	***Lactococcus gar***	+
DSM 13273	++	DSM 2648	+	***L. brevis* 7**	+
DSM 20174	++	NC8	++	***L. rhamnosus***	-
DSM 20246	+	Lb965	-	***L. acidophilus* NCFM**	++

Table 3.3.1 Range of activity of the ASM1 peptide against *Lactobacillus* species
Results from soft agar assays testing the activity of purified ASM1 against an array of *Lactobacillus* species. Species names are in bold and refer to the strains listed below them (or just themselves if only one strain from that species was tested).
- indicates no clearing
+ indicates weak clearing, less than 12 mm in diameter
++ indicates mid-strength clearing (but not transparent) up to 15 mm in diameter
+++ indicates transparent clearing of any size.
Assays were done in triplicate.

When compared with similar experiments using GccF, some differences were observed in the inhibition caused by ASM1. This was most apparent against *L. plantarum* WCFS1 and *L. brevis* 7. GccF caused severe inhibition of WCFS1 but did not inhibit the *L. brevis* strain at all, ASM1 on the other hand caused mild inhibition of the growth of both of these strains (Kerr, 2013). This difference is most likely a result of the sequence

differences in the C-terminal, flexible tail of the bacteriocin as the proteins are identical from residues 1-30.

3.4 ASM1 and *E. faecalis*

Previous experiments had shown that GccF was active against more than just *L. plantarum* strains and in fact inhibited the growth of quite a wide array of *E. faecalis* strains and other species at varying levels of effectiveness. In addition to this, it had been found that the effects of GlcNAc were also highly variable (Kerr, 2013). For some strains, GlcNAc provided protection against the bacteriocin activity, as observed for *L. plantarum*, whereas for other strains it induced increased sensitivity to the bacteriocin.

	Treatment		
Strain (NZRM)	GlcNAc	ASM1	ASM1 + GlcNAc
89	-	-	-
1106	-	+	++
1240	-	++	++
2262	-	-	++
3178	-	+	+
3488	-	+	++
3601	-	+	-
4061	+	+	++
V583*	-	+	++

Table 3.4.1 *E. faecalis* response to ASM1 and GlcNAc

Tabulated results of plate assays testing *E. faecalis* strain responses to GlcNAc, ASM1 and a combination of the two.

+	minor clearing, barely translucent
++	semi-translucent clearing
+++	completely transparent, no cell growth
-	no effect visible

Data is averaged over three plates for each sample tested, this was due to variability within the responses.

**E. faecalis* V583 does not have a NZRM number

Experiments were therefore carried out to test the effects of ASM1, GlcNAc and a combination of the two reagents on various *E. faecalis* strains (see 2.1.4). The results from these in-plate growth assays showed a wide array of sensitivities to both ASM1, GlcNAc alone and their combined effect (Table 3.4.1).

Many of the strains were sensitised by the addition of GlcNAc to ASM1 with only two being unaffected by the addition of GlcNAc (NZRM 89 and NZRM 3178). When compared with results from experiments with GccF it appears that there are some notable differences between the activities of the two bacteriocins with ASM1 not affecting some strains that the GccF inhibited quite strongly (notably NZRM 2262). However, the synergistic or inhibitory effect provided by GlcNAc is consistent between the two bacteriocins (Kerr, 2013). This suggests that the effects of GlcNAc protection are a result of the mechanisms within the target cell, as would be expected.

In only one of the nine strains tested did the addition of GlcNAc protect the cells from ASM1 activity. It is reasonable to assume that the difference in effects between strains (and between *E. faecalis and L. plantarum*) is due to differences in the GlcNAc metabolic pathway. Strains such as *E. faecalis* NZRM 89 either do not rely on GlcNAc uptake in normal situations or have a drastically different uptake system to that of *E. faecalis* V583, one of the strains that becomes more sensitive to ASM1 when it is administered alongside GlcNAc. V583 and other strains that are sensitised to ASM1 by the presence of GlcNAc, most likely respond to the presence of GlcNAc in the media by increasing the expression of GlcNAc-PTS transporters responsible for GlcNAc uptake. As a result they present more binding sites for the GlcNAc moieties of ASM1.

It is worth noting that, although most *E. faecalis* strains were sensitive to ASM1, those tested were not as sensitive to the bacteriocin as some *L. plantarum* strains tested (Table 3.3.1). This is not entirely surprising as bacteriocins are often most effective against closely-related strains.

3.5 pA1_ASM1 Sequencing

3.5.0 Extension of core ASM1 sequence

When ASM1 was initially identified and characterised by Hata *et al.* in 2010, preliminary sequencing revealed a cluster of 5 ORFs (including the structural gene) and one partial ORF. When compared to the GccF operon, high sequence identity was found for three of the complete ORFs and the partial ORF, the ASM 1 structural gene showed high identity to *gccF.* The first upstream ORF showed high homology to *GccE* (putative transcription regulator), the second showed partial homology to *GccD* (putative thioredoxin) and the partial ORF showed high homology to *GccC* (putative thioredoxin) (See Figure 1.9.1).

Using the core sequence from Hata *et al.* (2010) primers were designed to permit the elongation of the known sequence in both directions by primer walking (see 2.2.4). This sequence elongation was necessary to establish the boundaries of the ASM1 cluster and to examine its similarity to the GccF cluster. It was also necessary to determine whether or not the genes (and putative operon) were plasmid-based.

3.5.1 The ASM1 operon

The seven open reading frames sequenced formed a fully functional operon that includes all of the genes necessary for bacteriocin synthesis, maturation and transport, as well as the associated immunity protein (Figure 3.5.1). All of these genes have orthologues in the *gcc* cluster that encodes GccF. The primary difference is that in the *gcc* operon, a gene is present near *gccH* that was thought to be the immunity gene and was termed *gccI*. The absence of this in the *asm* cluster suggests that it is probable that *gccH* and *asmH* are the primary immunity genes in their respective clusters.

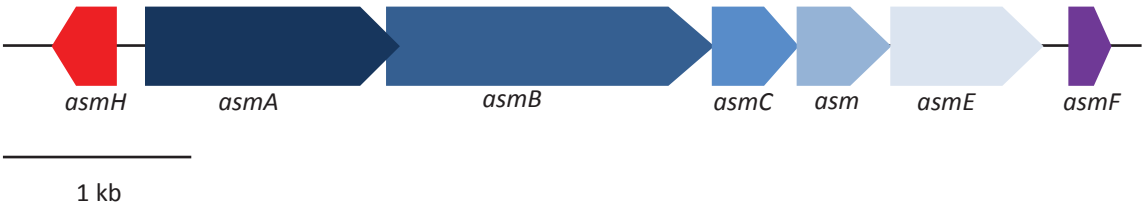


Figure 3.5.1 Schematic of sequenced ASM1 gene cluster in *L. plantarum* A -1
asmF (Structural gene) in purple, *asmH* (Putative immunity gene) in red.

A dot plot analysis (see 2.2.5) of the sequence similarity between pA1_ASM1 and the *gcc* gene cluster is very high within the cluster (*asmH* - *asmF*) but drops significantly to the normal background similarity present in *L. plantarum* DNA either side of the cluster (excluding *gccI*), suggesting that the gene cluster has undergone a translocation either from the plasmid to the chromosome or vice-versa (Figure 3.5.9).

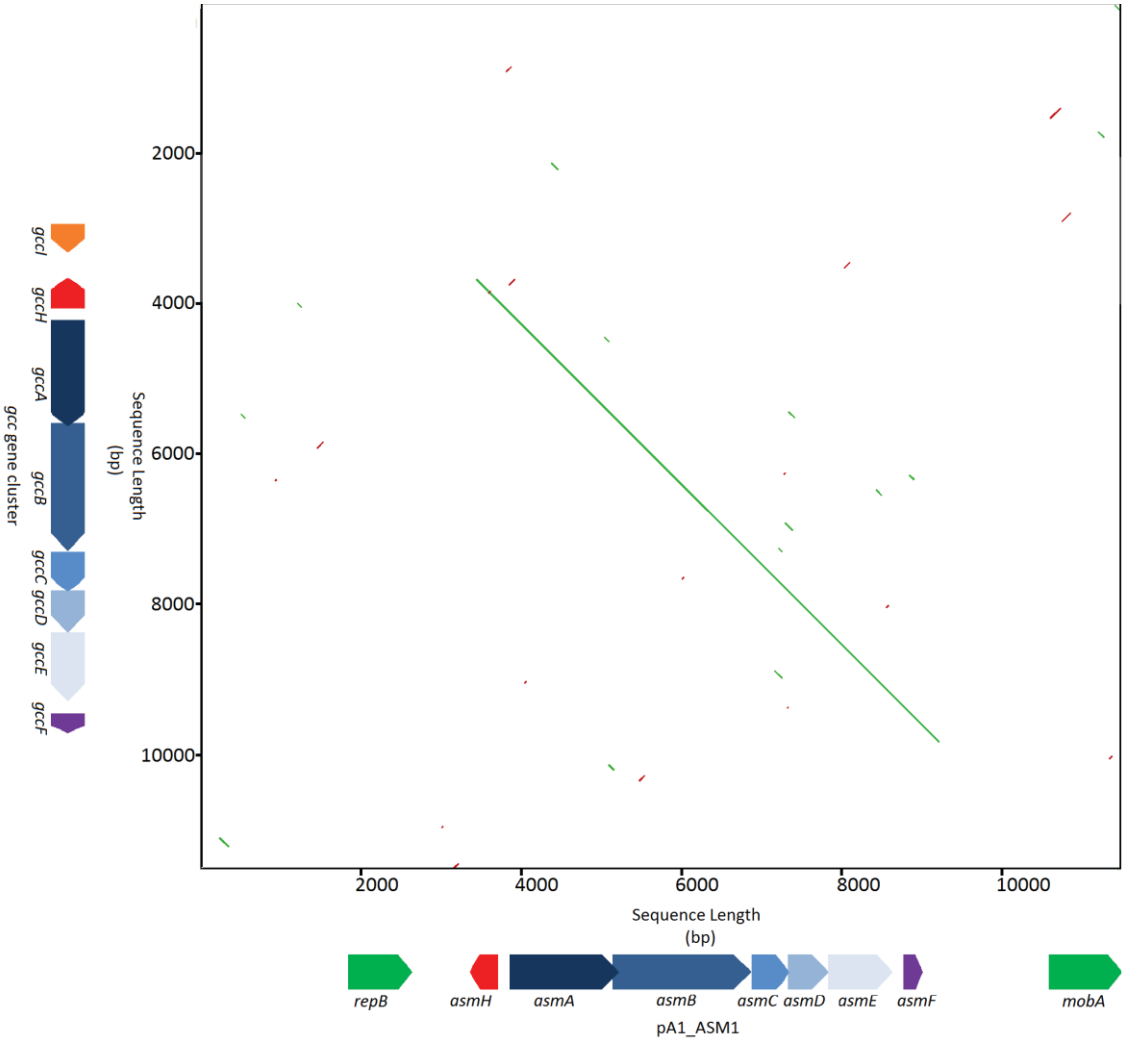


Figure 3.5.2 Dot-plot showing sequence similarity between pA1_ASM1 and the *gcc* cluster
The *gcc* gene cluster plus 2885 bp of the flanking DNA was compared to the pA1_ASM1 sequence. The sequence similarity reduces to background similarity levels at the stop codon of *asmH* but retains high similarity for approximately 150 bp after the stop codon of *asmF*.

Using BLAST analysis it was shown that the sequence similarity between the gcc and asm gene clusters ends at the *asmH* stop codon but carries on 150 bp after the *asmF* stop codon (in the opposite direction due to *asmH* being in the reverse frame). This would suggest that there was originally a *gccI* orthologue in the ASM1 cluster that has since been lost in a previous ancestor, as opposed to being a novel gene in the GccF cluster. The sequence of each ORF was analysed using BLASTP searches to detect any conserved domains and identify similar sequences (see 2.2.5).

3.5.2 *asmH* – The putative immunity gene

The open reading frame identified at the far end of the cluster from the structural gene was termed *asmH* due to its homology to the *gccH* gene in the GccF operon.

```
ASMH     MITKICVIPAVVTGIQESSISGNVIIDSAITAHFTHRLKKALHEISDTLKIFSLSEFE  58
         M+T IC+IP+ V GIQESSISG++IIDSAI+AHFTHRLKK+L  + +  KIF+LSEFE
GccH     MVTNICIIPSAVVGIQESSISGSMIIDSAISAHFTHRLKKSLQGVPNISKIFTLSEFE  58

ASMH     PDLTPDSNTLYVFSENVYHNLPASFLKNAKYVKFRNSQIFSSSLEDLAHAINIKLLQY  116
          +  P+ NT+Y+F +NVYH LPASFL  AKYVKF NS+IFSS L++L+HAI IKL QY
GccH     SNWEPNPNTIYIFPDNVYHVLPASFLNKAKYVKFPNSKIFSSDLKELSHAIKIKLFQY  116
```

Figure 3.5.3 Alignment of proteins encoded by *asmH* and *gccH*, putative immunity genes
Identical matches denoted by conserved residue symbol, conserved properties indicated by "+", dissimilar residues indicated by empty space.

No conserved domains were found in the *asmH* ORF but it shared a 66% identity with *gccH*, a gene of unknown function in the GccF cluster (Figure 3.5.3). Like *gccH* it is also in the reverse frame compared to the rest of the operon. In the GccF cluster, a gene further upstream of *gccH* was sequenced and termed *gccI*. This gene was initially proposed as the immunity gene as it was close to the GccF operon and also in the same orientation as the majority of the genes. However no orthologue of *gccI* was found in the ASM1 cluster.

gccI has only been identified in a small subset of *L. plantarum* genomes, including Lb965, MP-10 and *L. plantarum* 16 with very high conservation (>80% identity). The idea that it could be an immunity gene was initially strengthened by its presence in Lb965, a strain that is resistant to the glycocins of KW30 and A-1. Though a *gccI* orthologue is absent from the pA1_ASM1 plasmid, a copy is found in on a contig from the full genome sequence of A-1 with a sequencing coverage of 677-fold, indicating that it is possibly plasmid-based, though there is nothing else on the contig that seems relevant. The presence of a *gccI* orthologue in the A-1 genome would support the

hypothesis that it is involved in glycocin immunity and could be explained by the sudden decrease in sequence similarity between the *asm* and *gcc* operons though it is also possible that it has a secondary immunity function, such as *nisI* in the nisin system (Chen & Hoover, 2003), with *asm/gccH* being the primary immunity proteins.

3.5.3 *asmA* –The glycosyltransferase

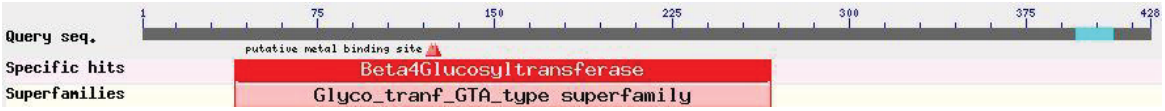


Figure 3.5.4 Conserved domains of *asmA* as predicted by BLASTP 2.2.29 (NCBI)
Highlighted conserved GTase domain cd02511, part of the glycosyltransferase superfamily.

The second open reading frame was termed *asmA* and has a conserved glycosyltransferase domain (cd02511) belonging to glycosyltransferase family 2 (GT2) from residues 40 to 267 (Figure 3.5.4) with an E-value of $7.02e^{-37}$ (predicted by NCBI conserved database). Furthermore the gene product shares 64% identity with the KW30 orthologue GccA. An important difference between the *gcc* and *asm* sequences is the overlap between the *asm*/*gccA* and *asm*/*gccB*. *asmA* is 18 base pairs longer than *gccA*, meaning that the overlap between *asmA* and *asmB* 18 base pairs longer than the 12 bp overlap between *gccA* and *gccB.* Despite this increased sequence overlap the ASMB amino acid sequence is identical to the GccB amino acid sequence (the the genes have some slight differences, as seen in Section 3.5.9).

3.5.4 *asmB* – The ABC Transporter

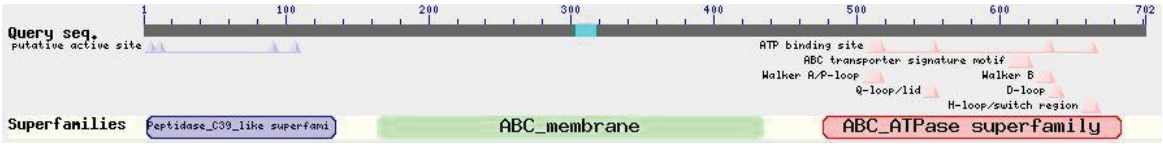


Figure 3.5.5 Conserved domains of *asmB* as predicted by BLASTP 2.2.29 (NCBI)
Highlighted conserved domains of a full ABC Transporter gene, Peptidase_C39 (pfam03412) in blue, Transmembrane domain (pfam00644) in green and ATP Binding Cassette (cd03254) in red

As mentioned, the *asmA* gene overlaps with the beginning of the *asmB* gene which encodes a full ABC transporter gene. There are three conserved domains within the protein sequence (Figure 3.5.5). The first is a Peptidase_C39 domain associated with the conserved group pfam03412 (E value of $3.59e^{-16}$) from residues 3 to 135. This is a common domain in the ABC transporters of bacteriocin producers as it is required to cleave the leader sequence from the immature bacteriocin peptide before it is secreted from the cell. The second conserved domain is the ABC transmembrane domain, a subunit of the ABC transporter that contains six transmembrane helices, associated with the conserved sequence pfam00644 (E value of $2.64e^{-07}$) which sits between residues 165-434. The third conserved domain is the ABC-ATPase subunit that gives the transporter its name, and is associated with the conserved domain cd03254 (E value $2.49e^{-40}$). This conserved domain occupies residues 476 to 685 of the protein and shares 77% identity with its GccB orthologue. Not surprisingly most of the variation occurs outside of the conserved domains.

The multi-domain protein belongs to the SunT superfamily which is a conserved protein associated with the transport of bacteriocins with a double-glycine leader sequence. This superfamily is named for the sublancin ABC transporter that transports sublancin 168, a glycocin of *B. subtilis*.

3.5.5 *asmC* – A putative thioredoxin

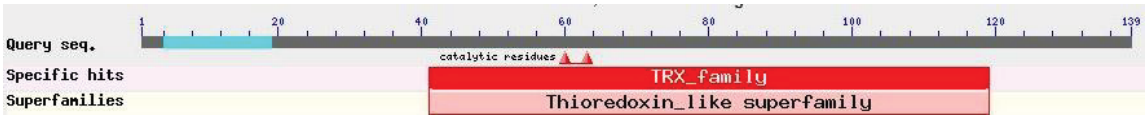


Figure 3.5.6 Conserved domains of *asmC* as predicted by BLASTP 2.2.29 (NCBI)
Thioredoxin-like domain (cd02947) underlined in red, conserved catalytic residues indicated by arrows

The *asmC* gene was partially sequenced by Hata *et al.*(2010) and presented as a partial ORF in the initial ASM1 gene cluster. This is the first of two ORFs which contain thioredoxin-like domains. In *asmC* the thioredoxin-like domain is between residues 41 and 119 and shows high similarity to conserved domain cd02947 (E value $2.16e^{-10}$) and contains the CXXC motif that is necessary for activity and is always present in functional thioredoxins (Figure 3.5.6). ASMC shares a 74% identity with its GccC orthologue and the genes are the same length in both strains.

3.5.6 *asmD* – A second thioredoxin protein

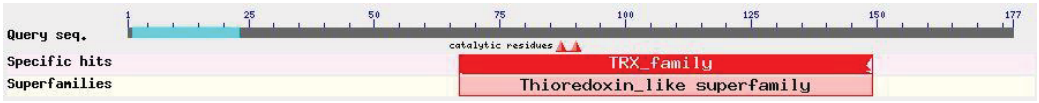


Figure 3.5.7 Conserved domains of *asmD* as predicted by BLASTP 2.2.29 (NCBI)
Thioredoxin-like domain (cd02947) underlined in red and catalytic residues denoted by red arrows

The second gene containing a thioredoxin-like domain is *asmD*. In this gene the TRX conserved domain occupies residues 67 to 149 (Figure 3.5.7). The conserved domain is cd02947 (E-value $1.91e^{-09}$) the same as for *asmC*, and also contains the functional CXXC motif. It is uncertain whether one of these genes is more important than the other in the process of disulfide bond formation in the ASM1 peptide or if they are both equally important or, indeed, if they even have a role in disulfide bond formation. The protein shares 66% identity with *GccD* with most of the conserved sequence being in the TRX-like domain of GccD with the N- and C- terminals being less similar.

When compared to the putative ORF from Hata *et al.* the sequences are very different, this is because two sequence errors introduce a frame shift that alters the sequence from residue 148 onwards results in the putative product being 158 residues in length as opposed to 177 (Hata *et al.*, 2010).

3.5.7 *asmE* – A putative response regulator

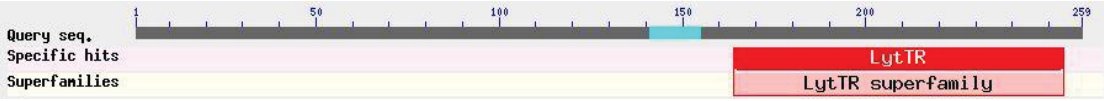


Figure 3.5.8 Conserved domains of *asmE* as predicted by BLASTP 2.2.29 (NCBI)
LytTR conserved region (pfam04397) underlined in red.

The *asmE* gene encodes a protein 259 aa in length that contains a LytTR domain, at the C-terminal end of the protein (Figure 3.5.6); these domains are common in bacteriocin clusters like the plantaricin cluster (Diep *et al.*, 2009) and are involved in the regulation of bacteriocin gene transcription. They receive input from membrane-bound proteins that detect extracellular pheromones (sometimes the bacteriocin) and increase the expression of the bacteriocin from the cell in response to increased pheromone levels.

The LytTR domain in *asmE* is probably involved in regulation of bacteriocin production, much like PlnC and PlnD in the plantaricin system (Diep *et al.*, 2009). The LytTR domain is similar to the conserved domain pfam04397 (E-value $6.06e^{-13}$) and occupies residues 164 - 254. The *asmE* ORF was sequenced by Hata *et al.*, but a single sequencing error meant that the previous predicted sequence gave alanine at residue 23 instead of proline (A23P).

The N-terminal region of the protein does not contain any recognised conserved domains, this is similar to what is seen in the *gcc* cluster. Normally, the N-terminus of a

protein containing the LytTR domain would contain a region that recognises signals from a histidine protein kinase (phosphorylation), thus acting as an input domain (Chen & Hoover, 2003).

3.5.8 *asmF* – The ASM1 bacteriocin

The structural gene in the ASM1 cluster, termed *asmF* does not contain any recognised conserved domains, but the gene product shares 86% identity to the GccF immature peptide with the variation occurring only in the leader sequence (cleaved during transport to make the mature peptide) and the C-terminal 13 amino acids (Figure 3.5.9). Despite the differences in the tail sequence the 4 C-terminal residues are identical for both peptides, indicating that the C-terminal region may be highly involved in function. The sequence is identical to that provided Hata *et al.* in 2010.

```
ASMF     MSKLVKTLTVDEISKIQTNGGKPAWCWYTLAMCGAGYDSGTCDYMYSHCFG  51
         |||||||||  ||||  |  ||||||||||||||||||||||||||||||||
GccF     MSKLVKTLTISEISKAQNNGGKPAWCWYTLAMCGAGYDSGTCDYMYSHCFG  51

ASMF     VKHSSGGGGSYHC  64
          ||  |  |  ||||
GccF     IKHHSSGSSSYHC  64
```

Figure 3.5.9 Alignment of the ASM1 and GccF prepeptides
Residues 1-21 (blue) make up the leader sequence and residues 22-64 constitute the mature peptide sequence.

3.5.9 Overlapping genes in the *asm* cluster

Though the overall sequence similarity between the operons is quite high, a few important differences have been noted. In the *asm* cluster, the *asmA* gene overlaps the beginning of the *asmB* gene by 32 base-pairs, including the stop codon of *asmA* whereas in the *gcc* cluster the *gccA* overlaps the *gccB* gene by only 14 base-pairs although again, this includes the stop codon of *gccA*. This difference is due to a point

mutation where T is changed to A, making what is a stop codon (TAA) in the *gcc* cluster a lysine (AAA) in the *asm* cluster (Figure 3.5.10)

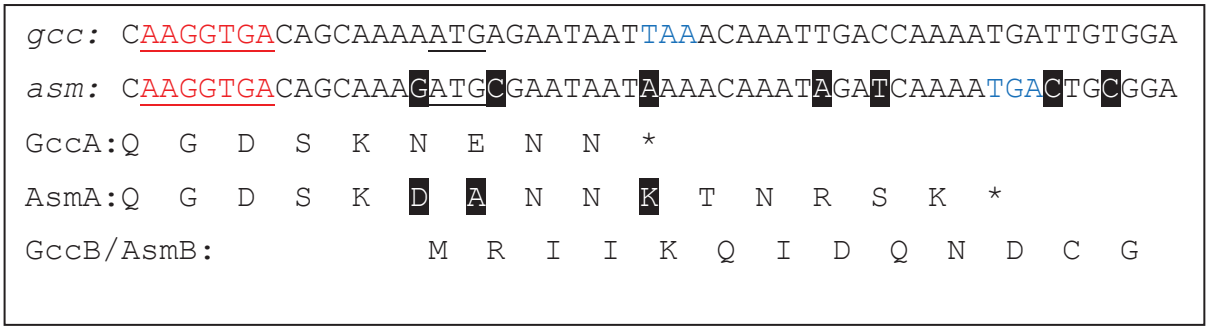


Figure 3.5.10 Sequence differences between *gcc* and *asm* clusters with protein translations.
Ribosome binding site in red, amino acid and nucleic acid differences in white. Start codons underlined. Stop codons in blue.

Overlapping genes are relatively common, especially in prokaryotes where genes are often transcribed as groups within the operon. They can be useful in determining the phylogeny of species due to their comparatively high conservation. A mutation has to make sense in two genes, not just one, resulting in greater selective pressure. It has been noted that although short phase +2 overlaps are more common overall, phase +1 overlaps (such as the one between *asmA* and *asmB*) are more common in larger overlaps (>7 bp) (Cock & Whitworth, 2010).

Whilst the presence of overlapping genes is a very useful tool in determining which strain is the origin of a gene or gene cluster, there is not enough data to determine whether *asm* or *gcc* is the older of the two clusters. There are two main options to describe the origins of the sequence. The first is that it could be assumed that the original gene cluster lacked the TAA stop codon present in *gccA* and the resulting A -> T mutation caused a C-terminal truncation, lessening the overlap between the two genes in the *gcc* cluster. The other option is that the ancestral GTase protein ended in –NN* with the T -> A mutation leading to a lengthening of the ASM GTase protein,

leading to a gain of three positive charges in the C-terminal region of the protein and an increased overlap between *asmA* and *asmB*. However, in the latter idea two T -> A mutations would be required to add the three positively-charged residues to the C-terminus of the AsmA protein.

3.5.10 Circularising the sequence

Sequencing beyond the *asm* gene cluster was done in several stages and resulted in the discovery that the gene cluster is situated on a circular piece of DNA, a plasmid,

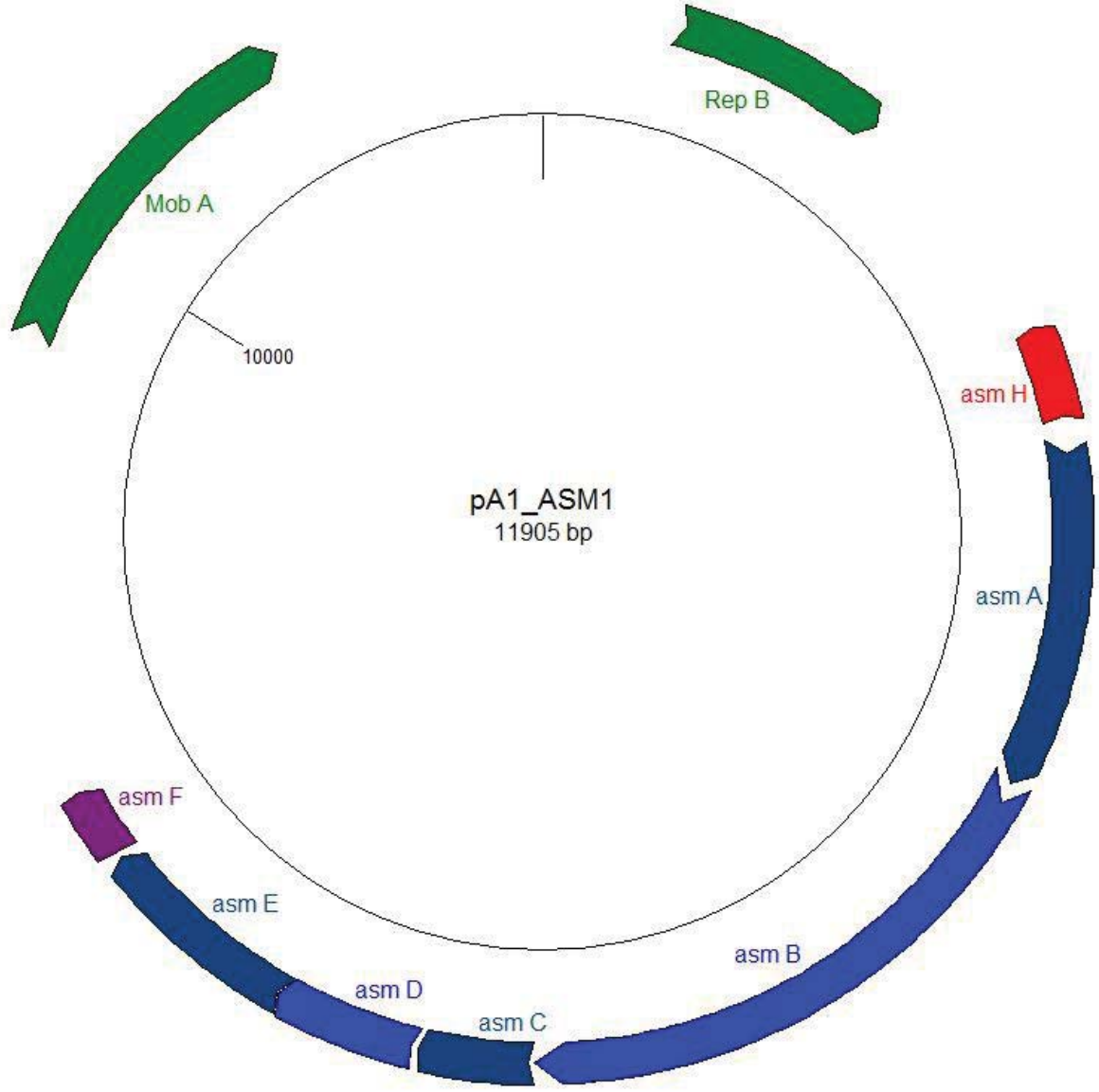


Figure 3.5.11 Plasmid map of recognised ORFs in pA1_ASM1
The ASM1 structural gene in purple, putative immunity gene in red and other ASM1 components in blue. Other plasmid ORFs are shown in green. Figure made in GENtle (see 2.2.5)

11,905 bp in length (Figure 3.5.13). Sequence analysis showed that the *asm* gene cluster is the only functional operon present, occupying slightly less than half of the sequence. Two other *ORFs* on the plasmid were identified. The first was a mobilisation protein (*mobA*) and the second as a replication protein (*repB*). Both of these ORFs are commonly found in *Lactobacillus* plasmids where they are genes responsible for the correct replication of the plasmid. Unfortunately an origin of replication could not be identified by homology.

3.6 Isolation of pA1_ASM1

3.6.0 Determining the plasmid location

Before sequencing of the plasmid was completed, Southern blotting was used to verify the location of the ASM1 gene cluster using the *asmA* PCR product as a probe to highlight the location of the plasmid (see section 2.2.3) (Figure 3.6.1).

As part of the sequence was already known and extension beyond the core cluster had begun, it was determined that restriction enzyme *Msc*I would be a good choice to linearise the plasmid as a single restriction site had already been found in the sequence and the *MscI* restriction enzyme has relatively few restriction sites in *Lactobacillus* genomes because of their relatively low GC content.

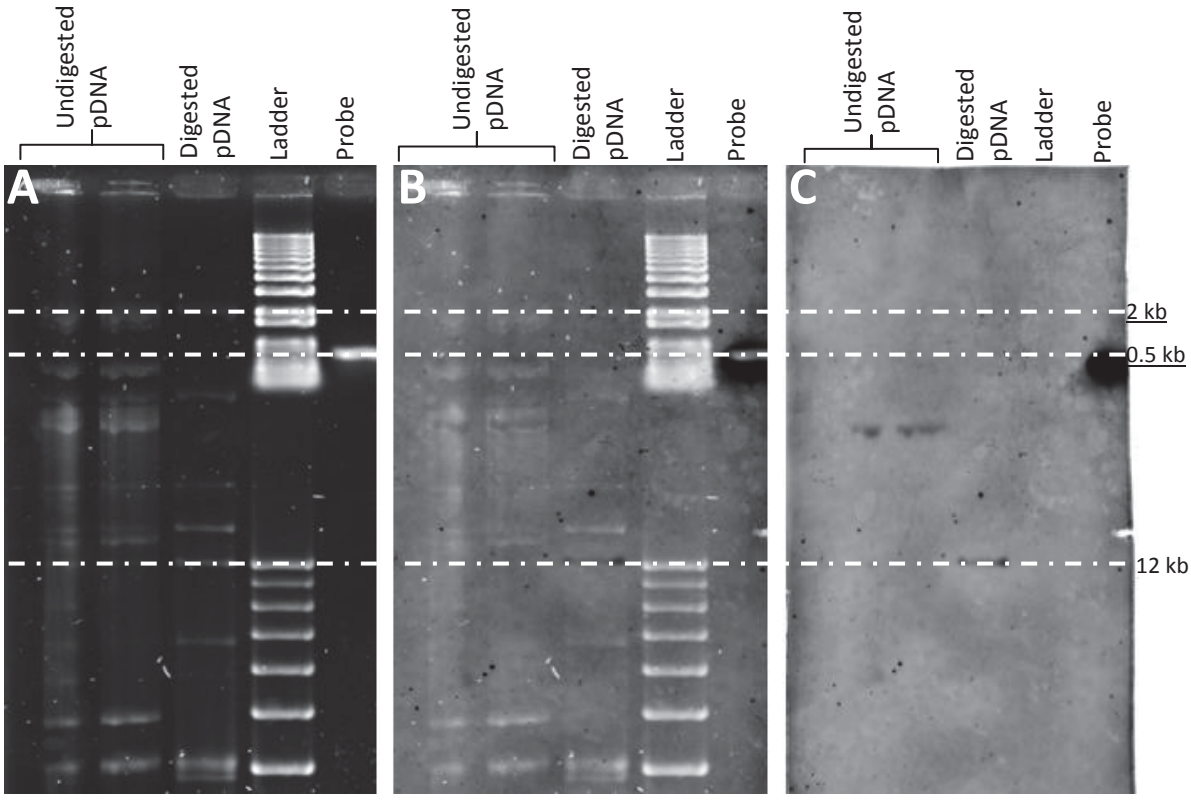


Figure 3.6.1 Southern Blots using *asmA* PCR product as fluorescent probe.
(A) 0.6% Agarose gel run at 30V for 26 hours at 4°C, then stained with Ethidium Bromide and visualised under UV light. B) Southern Blot fluorescence scan overlaid onto of the Agarose gel image, aligned using probe, physical nicking and positive result. C) Southern Blot fluorescence scan showing *asmA* probe hybridisation locations. Underlined size indicators refer to ladder loaded after 24 hours to ensure the probe is the correct size.

The Southern blots showed that the *asmA* PCR product hybridised exclusively to a 12 kb band in the digested A-1 pDNA sample and to a blurred band in the undigested sample that resolved around 20 kb (though this is unreliable due to the variable mobility of undigested DNA based on its conformation). The size of the linearised band that the *asmA* probe bound to was confirmed by the sequencing, completed shortly after the Southern Blot was finished.

3.6.1 Isolating the plasmid

Several attempts were made to isolate the pA1_ASM1 plasmid in order to transform it into a heterologous host (see 2.2.10). Gel purification kits were used following the

manufacturer's instructions but apparent yields were always < 10 ng/µL, substantially less than is required for transformation via electroporation and was regularly of poor quality. It was assumed this was because the plasmid size (~12 kb) was at the limit of the kit's capabilities. To try and improve transformations the plasmid was digested with both *Msc*I and *Stu*I, restriction enzymes that should cut either side of the ASM1 operon (*asmA* – asmF) to produce a fragment of DNA approximately 7 kb in length. This too was unsuccessful as the DNA of the correct molecular weight could not be extracted at any viable concentration. The process was therefore abandoned.

3.7 Transformation of *Lactobacillus* Strains

One of the primary goals of this research project was to transform the ASM1 gene cluster into a heterologous host so the transformation of plasmids into *Lactobacillus* strains was a major focus. The strain chosen for transformations was *L. plantarum* NC8, primarily because it is a naturally plasmid-free strain of *Lactobacillus* and also because it had been used previously in our lab and had been shown to be a relatively efficient recipient of DNA, compared to other *Lactobacillus* strains.

Plasmid	CFU per µg DNA ± SD
pTRK669	$5.47 \times 10^3 \pm 3.34 \times 10^3$
pSIP409	$6.34 \times 10^3 \pm 4.96 \times 10^3$

Table 3.7.1 Efficiency of transformation into *L. plantarum* NC8
Overall transformation efficiencies of two plasmids into *L. plantarum* NC8 using method described in 2.1.10. Results show colony forming units (CFU) per µg of DNA in original reaction ± Standard Deviation across all trials

Two plasmids were the main focus of NC8 transformation attempts, these were pTRK669 (Russell & Klaenhammer, 2001) and pSIP409 (Sørvig *et al.*, 2005). These two plasmids can replicate both in *E. coli* and *L. plantarum* and would potentially allow for

the insertion of individual genes from the ASM1 cluster for expression experiments, for example to test the hypothesis that the *asmH* gene encodes the immunity protein in the ASM1 gene cluster.

Multiple methods were tried before that described in 2.1.10 was used; this resulted in reliable transformation with adequate transformation efficiencies across two plasmids with different sizes, pSIP409 at ~5 kb and pTRK669 at ~2.5 kb (Table 3.7.1). These transformations also dispelled the assumption that the larger plasmid would cause a reduction in transformation efficiency when transformed by electroporation.

As the attempted transformation of a heterologous host with the ASM1 operon (either as pA1_ASM1 or ligated into a vector) was not successful, a number of attempts were made at transforming NC8 cells with an entire A-1 pDNA library (from an unprocessed pDNA extraction, see section 2.2.6). Though one of these attempts appeared to be successful, PCR results could not confirm the presence of pA1_ASM1 and selection with ASM1 protein proved unsuitable. This was because the background rate of resistance of NC8 WT to ASM1 seemed higher than the actual transformation efficiency indicating that a selected colony was more likely to have developed spontaneous resistance than to have acquired the pA1_ASM1 plasmid.

4. The Bacteriocins of L. plantarum A-1

4.1 Novobiocin-Induced Mutants

4.1.0 Mutant Selection by lack of inhibition

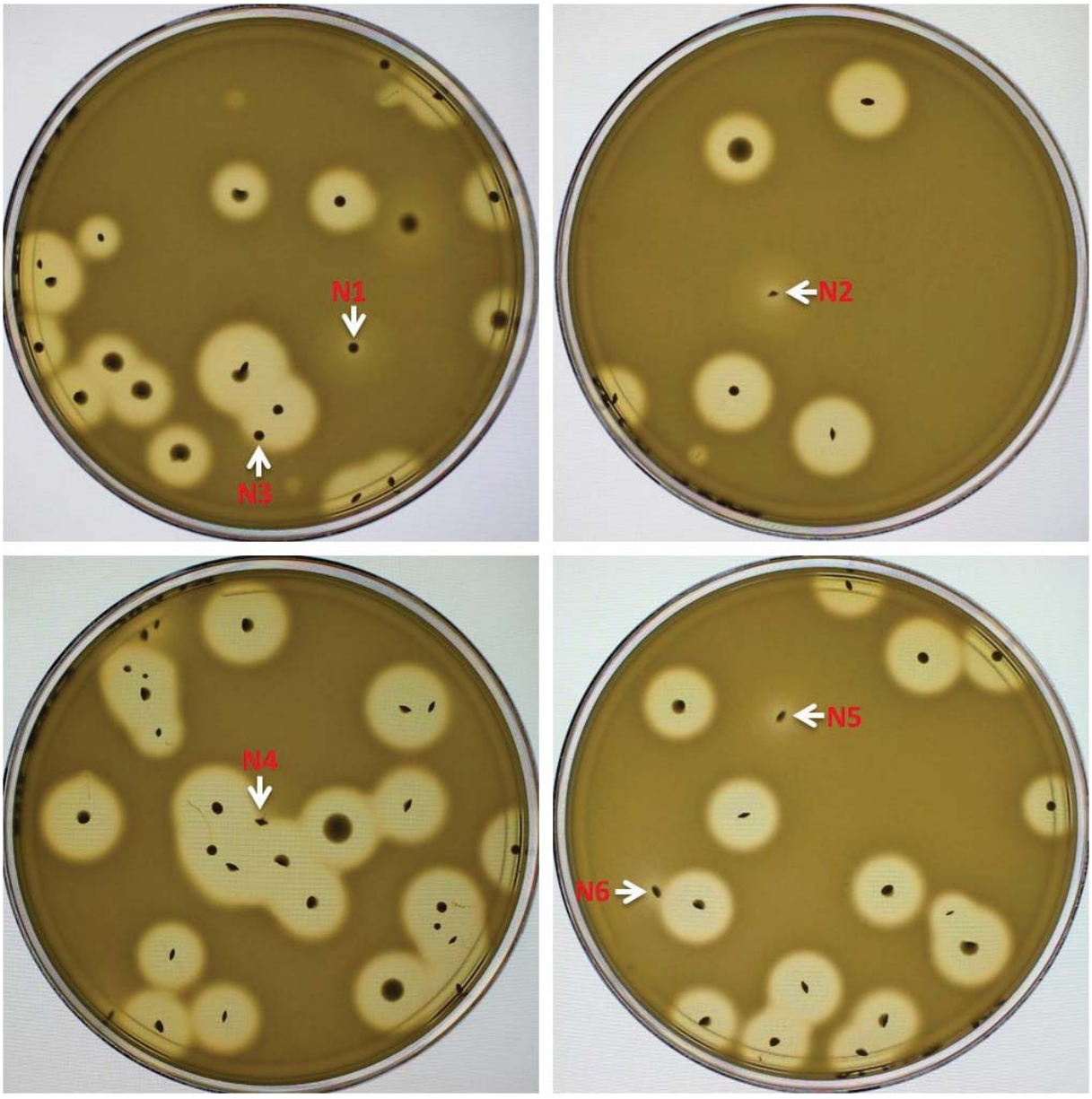


Figure 4.1.1 Novobiocin - treated *L. plantarum* A-1 cultures overlaid with a sensitive strain suspension
MRS agar plates of novobiocin - treated *L. plantarum* A-1 colonies overlaid with agar containing the ASM1 indicator strain *L. plantarum* ATCC 14917. Large inhibition halos of complete inhibition are present around most colonies but for some of the colonies formed by novobiocin treated cells the inhibition is visibly attenuated, six of these were selected and isolated for further use (Colonies labelled N1 - N6).

As described in section 2.1.3, *L. plantarum* A-1 was cultured for 48 hours in MRS medium, supplemented with 0.2 μg/mL novobiocin, with or without the addition of 2% (w/v) GlcNAc and then diluted and grown on MRS agar. These strains were then overlaid with agar containing a suspension of *L. plantarum* ATCC 14917, a sensitive strain and the plates were incubated further.

From these plates, six isolates with reduced or absent bacteriocin production were isolated (Figure 4.1.1). These six strains were termed N1 – N6 and were selected from two different treatment methods: N1, N2 and N3 were isolated from novobiocin treatment without 2% GlcNAc and N4, N5 and N6 the novobiocin treatment with 2% GlcNAc.

Novobiocin was used as the mutagenic product because, like other coumarins, it binds to the GyrB subunit of DNA Gyrase, inhibiting its function (Lewis *et al.*, 1996). The result of this activity means that, at non-lethal doses, novobiocin treatment results in a high (>4%) rate of plasmid loss (curing) in the treated cells (Karthikeyan & Santosh, 2010). This high rate of plasmid curing was evident when the rate of bac^- or bac attenuated mutants was calculated to be approximately 10%, far higher than the mutation-based novobiocin-induced bacteriocin attenuation in *L. plantarum* KW30 (with a chromosomal bacteriocin) which resulted in only 1 in 600 colonies losing bacteriocin function (Kelly *et al.*, 1996).

One of the proposed uses for the novobiocin-induced mutants was to determine whether the production of the ASM1 bacteriocin could be re-established in a mutated host, complementing the lost genetic data. To establish whether or not the selected mutants would be useable in these studies, bacteriocin production by each of the six

novobiocin mutants was tested using conventional plate assays in order to find out whether ASM1 production had been completely removed or merely attenuated.

4.1.1 Identification of secondary activity

To determine the activity of the A-1 derived novobiocin mutants, soft agar assays were carried out (as described in section 2.1.5) using the novobiocin strain supernatants to treat MRS agar containing a suspension of *L. plantarum* ATCC 14917.

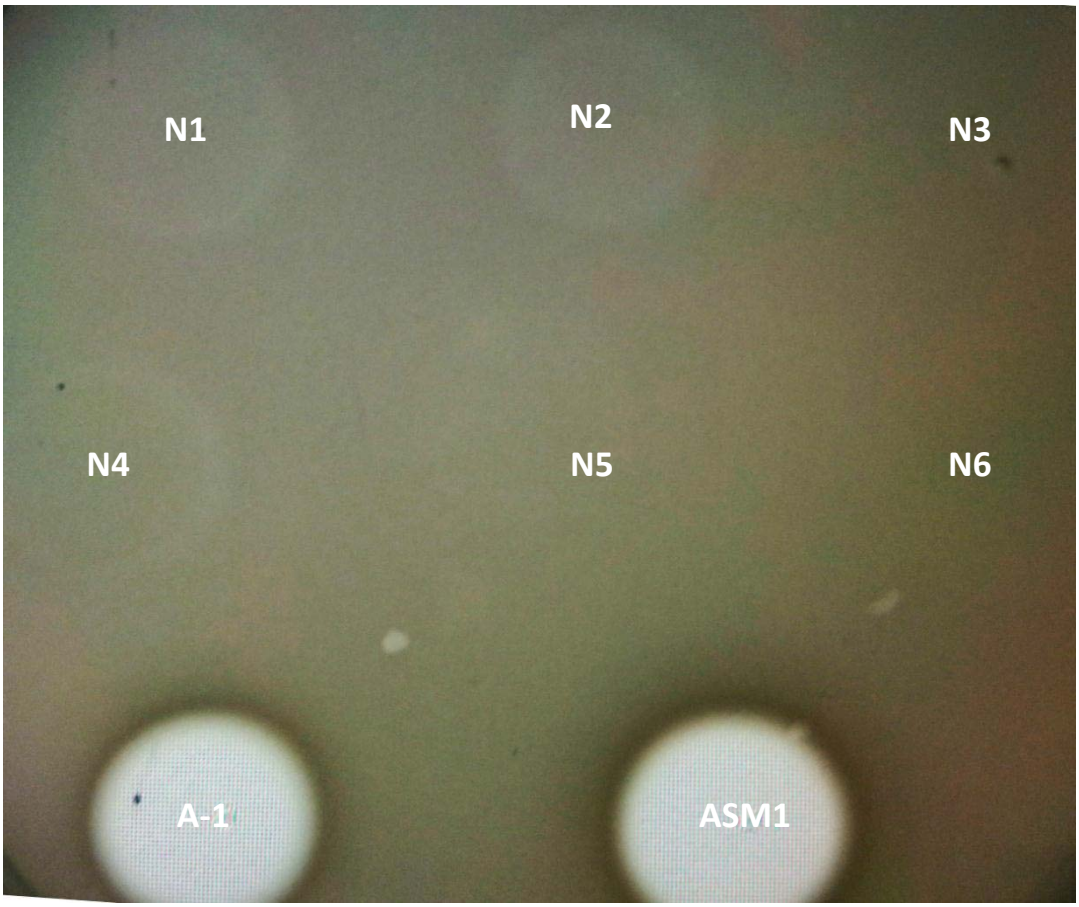


Figure 4.1.2 Residual bacteriocin production by novobiocin mutants N1 - N6
MRS plate with suspended *L. plantarum* ATCC 14917 culture dosed with 3 µL aliquots of supernatants of N1-N6 and A-1 as well as 3 µL of 1 µM purified ASM1 (as labelled). Faint inhibition halos present where N1, N2 and N4 supernatants were applied. Full inhibition present where A-1 supernatant and purified ASM1 were applied, indicative of an inhibitory activity in these three mutants not present in the other three.

These assays showed that the six isolated novobiocin mutants did not exhibit inhibition of bacterial growth to the same degree as the A-1 wild-type, but supernatant from the

N1, N2 and N4 strains produced faint inhibition halos when tested against *L. plantarum* ATCC 14917 suspended in MRS + 1% Agar (Figure 4.1.2).

When treated with purified ASM1, it was shown that the three novobiocin mutants that had residual bacteriocin activity (N1, N2 and N4) were as susceptible to ASM1 as were the three that did not exhibit residual activity (N3, N5 and N6). As the residual activity present in the N1, N2 and N4 strains was not likely to be attenuated ASM1 function (loss of the pA1_ASM1 plasmid would include loss of all immunity genes) it is most likely to be due toa second bacteriocin, already present in the A-1 genome. This sensitivity to ASM1 also strongly suggests that the immunity gene is no longer present in any of the six novobiocin mutants and, as immunity genes associated with bacteriocins are usually within the same cluster as the structural peptide (Ben Omar *et al.*, 2008), it supports the hypothesis that the ASM1 gene cluster is on a plasmid that has been cured from the novobiocin mutants.

4.1.2 Verification of plasmid absence

As a means to verify that the asm containing plasmid had been cured from the novobiocin mutants, PCR with primer pairs designed to amplify various regions of the pA1_ASM1 plasmid (both within the ASM1 operon and on the opposite side of the plasmid) was carried out (see 2.2.1). The absence of product indicates that at least part of the target sequence was no longer present. Results of this experiment should provide more evidence to support the hypothesis that A-1 produces a secondary inhibitory substance and that the ASM1 operon is plasmid-based. As expected the PCRs did not generate any products from the novobiocin mutants' pDNA but did from

the wild-type A-1 pDNA providing strong evidence that the plasmid has been cured from the novobiocin-treated strains (Figure 4.1.3).

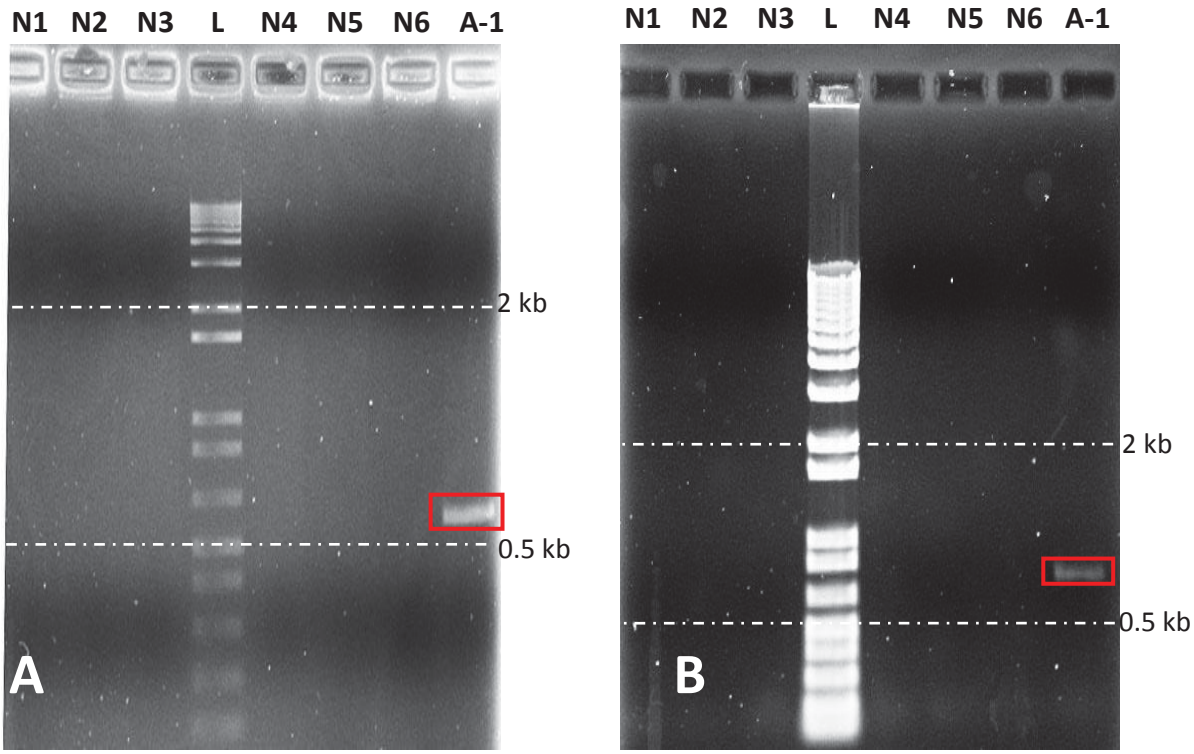


Figure 4.1.3 Agarose gels showing primer products from A-1 and N1 - N6 pDNA extracts
(A) 1.5 % Agarose gel of A1FinalPCR (Forward and Reverse) primer product from novobiocin mutants and A-1 wild-type plasmid DNA. Expected product at 590bp (red box) (B) 1.0% Agarose gel of PCR using AsmA_Nde_F + A1_gDNA_Gtase_direct primers. product from novobiocin mutants and A-1 wild-type plasmid DNA. Expected product at ~720 bp (red box).

4.2 Investigation of Other A-1 Bacteriocins

4.2.0 Secondary Substance Cluster Identification

Having identified that there was a secondary inhibitor, the genomic sequence of the A-1 strain (in randomly assembled contigs) was analysed, searching for ABC transporter genes with the ASM1 Peptidase_C39 domain of the *asmB* gene product. This is a well established marker for bacteriocins as the peptidase region is required for the removal

of the leader sequence. However, it is not fool-proof as not all bacteriocins have ABC transporters with these conserved peptidase regions (Havarstein *et al.*, 1995).

Using these methods three operons were found: the first, already known is the ASM1 operon; the second is on a contig labelled Contig_4 (17,296 bp in length) with an expect score of $3e^{-07}$; and the third is on Contig_57 (23,579 bp in length) with an expect score of $6e^{-06}$. The Peptidase_C39 of the ABC transporter on Contig_57 shares a 25% identity with that of the ASM1 Peptidase_C39 and matches to a well-studied bacteriocin ABC-transporter called PlnG which is part of a plantaricin operon found in many *L. plantarum* strains (Diep *et al.*, 2009). The bacteriocin associated with this operon is, however not likely to be responsible for the secondary activity observed as the strain used to identify the activity (*L. plantarum* ATCC 14917) is resistant to the plnEF/plnJK bacteriocins as shown in Figure 4.2.9. The other candidate, on Contig_4 is a Peptidase_C39 domain that shares a 57% identity (and 75% positive match) with an ABC transporter known as ComA, part of a proteolytic bacteriocin system in *L. brevis* KB290 (Fukao *et al.*, 2013).

4.2.1 ORF identification in ASM2

Having identified a potential site for an additional ASM operon (ASM2), further analysis of Contig_4 was carried out. Sequencing data showed that the sequence had a very high coverage, suggesting that it was present in multiple copies and, as such, probably a plasmid. The assembled operon is relatively simple for a bacteriocin cluster, containing only four genes (Figure 4.2.1). However, it does contain all of the elements required for bacteriocin function; a structural gene, genes encoding proteins required

for transport and maturation of the peptide and a domain of unknown function (DUF) which could be an immunity protein, a gene responsible for preventing self-inhibition.

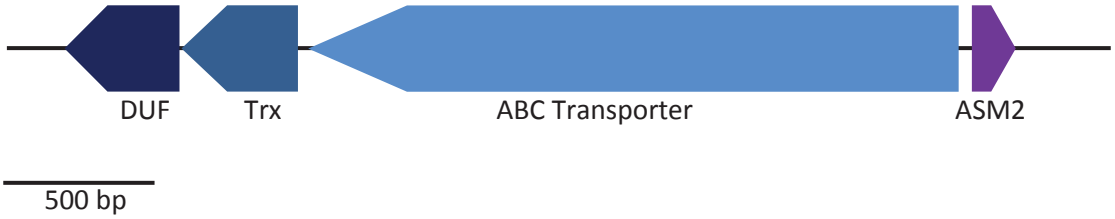


Figure 4.2.1 Putative ASM2 operon
Thioredoxin (Trx) gene, ABC Transporter with C39_Peptidase domain and ASM2 structural gene. The ORF encoding the domain of unknown function (DUF) is a putative immunity gene.

A putative immunity gene was found in the contig but is separated from the operon by a large amount of non-coding DNA and a putative mobilisation protein, indicating that it is probably not connected to the ASM2 operon. Sequence analysis of each ORF in the putative ASM2 cluster was carried out using BLASTP, as described in section 2.2.5.

4.2.2 *orf1* - Domain of unknown function

orf1 is a conserved domain of unknown function (DUF) (Figure 4.2.2) encoded in the opposite direction to the ASM2 structural gene.

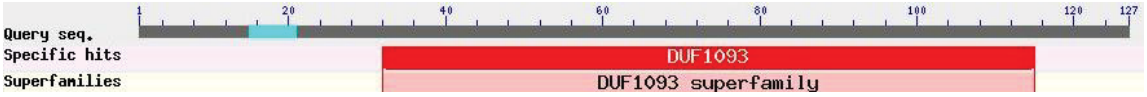


Figure 4.2.2 Conserved domain present in *orf1* of ASM2 predicted by BLASTP 2.2.29 (NCBI)
Domain of unknown function (pfam06486) underlined in red

The gene product contains a domain characteristic of the DUF superfamily (pfam06486 E-Value $3.75e^{-07}$), between residues 32-115, and shares 65% identity to a DUF in *L. plantarum* WJL. It is possible that this ORF encodes an immunity protein as these are often conserved between similar bacteriocins but do not have a recognisable function.

4.2.3 *orf2* - Putative thioredoxin

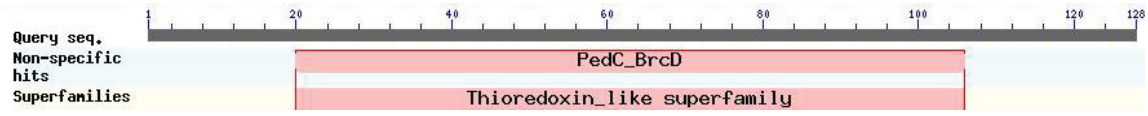


Figure 4.2.3 Conserved domain present in *orf2* of ASM2 predicted by BLASTP 2.2.29 (NCBI)
Thioredoxin-like conserved domain (TIGR01295) underlined in red

The second ORF was in the the same direction as *orf1* and contains a conserved domain from the trx superfamily (Figure 4.2.3).

The PedC_BrcD conserved domain TIGR01295 (E-value $5.21e^{-04}$) occupies the predicted gene product from residues 20 - 108 and is a conserved domain associated with bacteriocin transport and is generally found associated with pediocins. This is unusual as trx-domains are usually associated with keeping peptides reduced in an oxidising environment or forming disulfide bonds. The conserved CXXC active motif required for thioredoxin function is present between residues 31 - 34.

4.2.4 *orf3* - ABC Transporter

The third ORF is the largest by far and encodes three conserved domains that combine to make an ABC transporter (Figure 4.2.4).

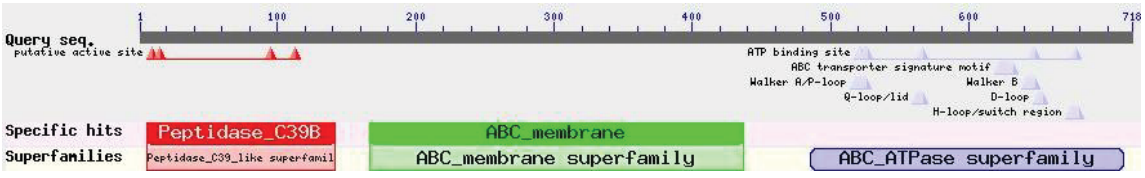


Figure 4.2.4 Conserved domains present in *orf3* of ASM2 predicted by BLASTP 2.2.29 (NCBI)
Conserved peptidase_C39 domain (cd02418) in red, conserved transmembrane domain (pfam00664) in green and the ATP binding cassette domain (cd03253) in blue.

This ORF which is oriented in the opposite direction to the ASM2 structural gene, translates to a gene product that contains the three conserved domains common in bacteriocin ABC transporters. The first conserved domain that constitutes residues 6 -

142 in the gene product is a Peptidase_C39B domain cd02418 (E-value 1.69e^{-50}). The second recognised domain (residues 173 - 443) is a transmembrane domain from pfam00664 (E-value 4.61e^{-41}). The third and final conserved domain is the ATP binding cassette domain of the ABC transporter (cd03253, E-value 4.85 e^{-69}). Despite the fact that the Peptidase_C39 domain in the *asm2* operon belongs to a differently conserved domain group to those in the glycocin transporters, the three domains combine to make an ATP transporter that has a similar overall structure to the SunT transporter from *B. subtilis* as they did in the *asm* operon.

4.2.5 *orf4* - Bacteriocin ASM2

The fourth and final ORF is the structural peptide, ASM2. This ORF doesn't contain any conserved domains but the putative gene product has extremely high sequence identitiy to a recently-characterised bacteriocin termed Bactofencin A (O'Shea *et al.*, 2013).

```
ASM2                MLKHLKKIVNYRKLSNQDLSKINGGKRKKHKCQIYNNGMPTGQYRWC    47
                          ||    |    |   || | ||||||||||  |||||||| ||||
Bacto* A      MFFNFMKKVDVKKNFGYKEVSRKDLAKVNGGKRKKHRCRVYNNGMPTGMYRWC    53
```

Figure 4.2.5 Sequence similarity between the ASM2 and Bactofencin A peptides
Leader sequence in blue. Mature 22 amino acid peptide in black. Amino acid differences are highlighted in black.Bacto* A is Bactofencin A peptide.

Though the leader sequence homology is very low, there are only four residues that differ between the two mature peptides and of those, three are similar to those in the orthologous peptide (Figure 4.2.5). Both peptides have the X-C-X_6-C-X motif (That is two cysteines separated by six non-specific residues) that suggests a disulfide bond holding a tight loop.

4.2.6 ASM2 activity spectrum

As the inhibitory effect of ASM2 was hard to visualise on the *L. plantarum* ATCC 14917 plates due to its weak inhibitory effect, the novobiocin mutants were tested against a wide array of *Lactobacillus* strains*, Lactococcus* strains and other species (Table 4.2.1)**.** Results from these experiments showed that the novel bacteriocin exhibits a very small range of activity against the strains tested and only inhibits one of these (*W. viridescens*) with any notable activity. The increased inhibition of *W. viridescens* was only observed after the plate had been incubated at 30°C for another 24 hours, suggesting that the strain only uses GlcNAc as a food source once its primary carbon source is exhausted.

It was also confirmed that only N1, N2 and N4 produce the secondary inhibitory substance. This is most likely due to the novobiovin plasmid curing knocking out the ASM2 plasmid as well as the ASM1 plasmid in the N3, N5 and N6 strains.

In addition to identifying a better indicator strain, this experiment showed that the inhibitory substance was not an attenuated form of ASM1 as there are multiple strains (such as *L. plantarum* ATCC 8014) that are strongly inhibited by ASM1 but are uninhibited by the this secondary bacteriocin. Another key observation was that all of the novobiocin mutants were inhibited by ASM1, indicating that the immunity gene associated with ASM1 was almost certainly removed with the plasmid. Interestingly there was one observed strain of the novobiocin mutants, N5, that was susceptible to the secondary substance. N5 was inhibited by the N1 supernatant, supporting the hypothesis that the plasmid containing ASM2 was cured from the cells through the same process as pA1_ASM1.

Plate Strain	PlnA 22	A-1 SN	N1 SN	N2 SN	N3 SN	N4 SN	N5 SN	N6 SN	Nisin	Baci-tracin	ASM1
KW30	-	-	-	-	-	-	-	-	+	++	-
KW30 N	-	+	-	-	-	-	-	-	+	+	++
A-1	-	-	-	-	-	-	-	-	++	+	-
WCFS1	-	-	-	-	-	-	-	-	++	+	+
Lb790	-	+	+	+	-	+	-	-	++	+	+++
DSM 20174	-	+	-	-	-	-	-	-	++	+	++
L. brevis	+	+	-	-	-	-	-	-	++	++	+
L. rhamnosus	-	-	-	-	-	-	-	-	+++	++	-
DSM 13273	-	+	+	+	-	+	-	-	++	+	++
DSM 2648	-	+	-	-	-	-	-	-	++	+	+
DSM 20246	-	+	+	+	-	+	-	-	++	+	+
ATCC 8014	+	++	-	-	-	-	-	-	++	+	++
DSM 2601	+	++	-	-	-	-	-	-	++	++	+
DSM 20205	-	+	-	-	-	-	-	-	++	+	++
NC8	-	++	+	+	-	+	-	-	++	+	++
Lb965	++	-	-	-	-	-	-	-	++	++	-
NCFM	-	++	-	-	-	-	-	-	+++	+	++
Weisella viridescens	+	++	++	++	-	++	-	-	++	+++	++
Lc_gar	-	++	+	+	-	+	-	-	++	++	+
LF_6a	+	-	-	-	-	-	-	-	+	++	-
Corn_7	-	-	-	-	-	-	-	-	+	++	-
Corn_28	-	-	-	-	-	-	-	-	+	++	-
E. faecalis V583	-	-	-	-	-	-	-	-	++	++	-
N1	-	++	-	-	-	-	-	-	+++	+	++
N2	-	+	-	-	-	-	-	-	+++	+	++
N3	-	++	-	-	-	-	-	-	+++	+	++
N4	-	+	-	-	-	-	-	-	+++	+	++
N5	-	++	+	-	-	-	-	-	+++	+	++
N6	-	++	-	-	-	-	-	-	+++	+	++

Table 4.2.1 Inhibition caused by bacteriocins and supernatants against a range of strains

Strains low-lighted in grey are particularly good indicator strains for ASM2.

- \- indicates no clearing caused.
- \+ indicates weak clearing, less than 12 mm in diameter.
- ++ indicates mid-strength clearing (but not transparent) up to 15 mm in diameter.
- +++ indicates transparent clearing of any size.

If a clearing area was weak or mid-strength but larger than the specified diameter it was shifted up a category

Plate images in Appendix 6.3 - Supernatant and bacteriocin testing plates for range of activity

4.2.7 Synthetic ASM2

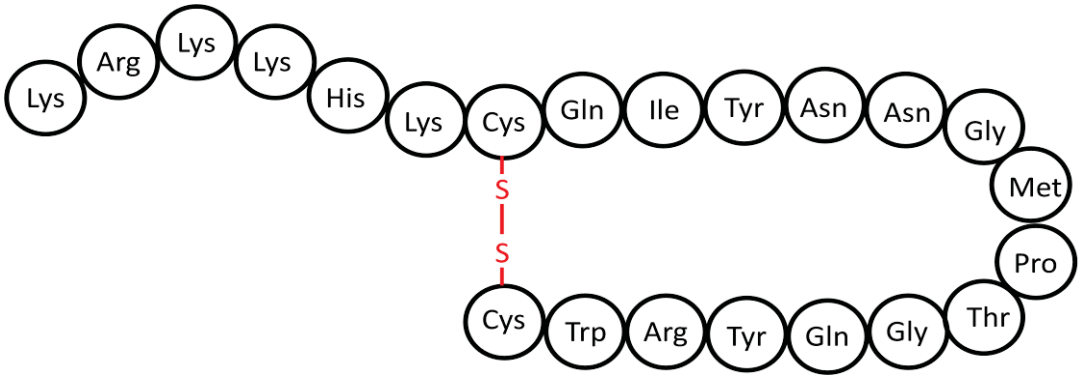


Figure 4.2.6 Putative structure of the ASM2 peptide with cysteine-cysteine disulfide bond.

Synthetic ASM2 using the predicted peptide sequence from the sequencing data was ordered from GL Biochem Ltd, Shanghai (See 4.2.2). A disulfide bond links the cysteine residues as shown. The presence of a gene with a thioredoxin-like domain in the putative operon suggests that the cysteine residues are linked (Figure 4.2.6).

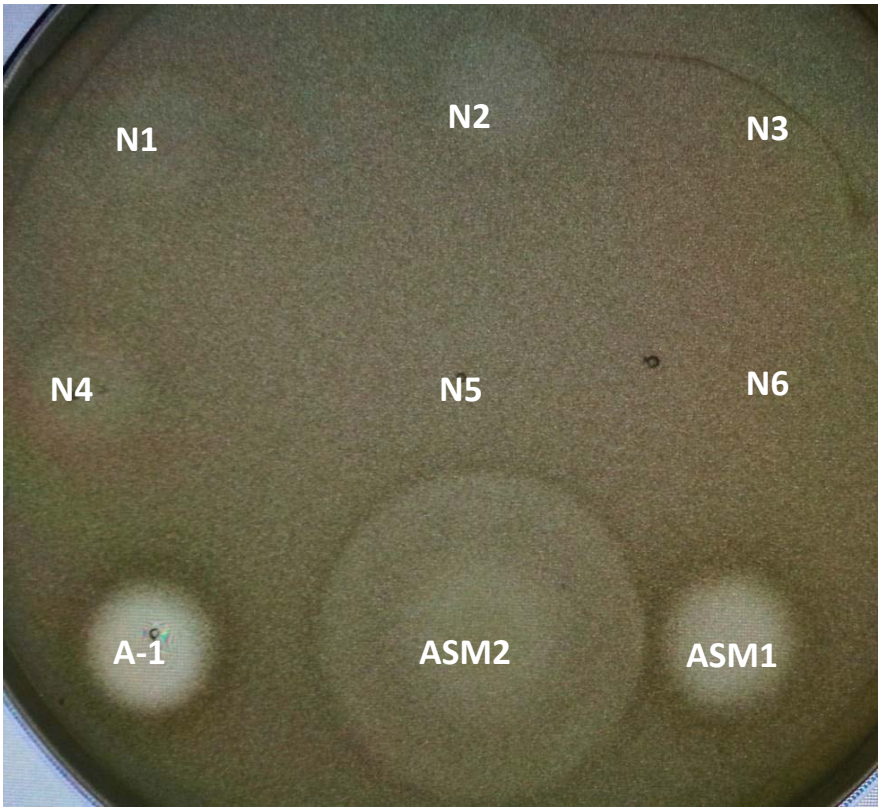


Figure 4.2.7 Agar plate treated with supernatants and pure bacteriocins
Soft agar assay testing the 3 µL aliquots of N1-N6 mutant supernatants, A-1 wild-type supernatant, 1 µM ASM1 and 10 mM ASM2 (as labelled) against *L. plantarum* DSM 13273.

Antimicrobial activity of the synthetic peptide was tested alongside that of the novobiocin mutants using one of the most sensitive strains using soft agar assays (see 2.1.5). This showed that ASM2 is active and has approximately the same inhibitory activity as the N1, N2 and N4 supernatants. These cause partial clearing, in contrast to the full clearing brought about by ASM1 and the A-1 supernatant (Figure 4.2.7). The large area of partial clearing was most likely caused by the high concentration of peptide used (10 mM) which will be much higher than its secreted concentration.

4.2.8 Comparison of A-1 and N1 growth and bacteriocin production

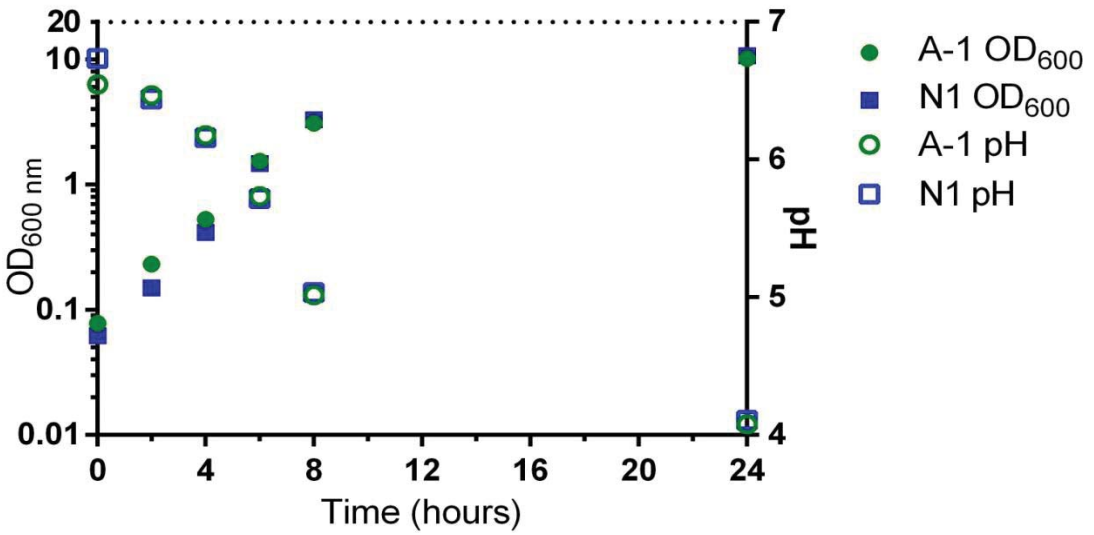


Figure 4.2.8 Graph of OD_{600} and pH of A-1 and N1 cultures over 24 hours
OD_{600} is on a logarithmic scale on the left Y-axis, pH is on a linear scale on the right Y-axis

Duplicates of N1 and A-1 cultures were grown and had both their OD_{600} and pH recorded over 24 hours. The results of this data collection showed that there is no significant difference between the two strains in growth rate, final OD_{600} in the conditions provided (final cell density) and production of acidic metabolic byproducts (Figure 4.2.8). A third measurement was taken at every time point by taking a 4 µL aliquot of cell-free supernatant and testing it for activity using soft-agar assays (See

section 2.1.5) testing the A-1 supernatant against *L. plantarum* ATCC 8014 and the N1 supernatant against *L. plantarum* DSM 13273.

The soft agar assays showed that the bacteriocin production in both strains starts to reach inhibitory concentrations at an OD_{600} of about 3 in A-1 but higher in N1, this could be because the ASM2 peptide is produced later in the growth of the culture or it could be an artefact of the lower activity of the ASM2 peptide as, even at putative full concentration after a 24 hour incubation, the inhibition "halo" caused by the N1 strain is very hard to capture in a photograph or scan but appears more clearly when looking through the backlit plates (Figure 4.2.9).

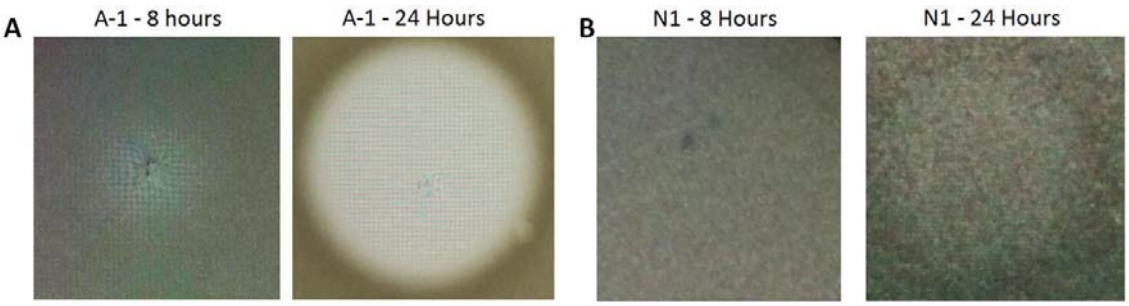


Figure 4.2.9 Segments of agar plates treated with supernatant from either A-1 or N1
(A) Soft agar assay with *L. plantarum* ATCC 8014 culture treated 3 µL of A-1 supernatant after 24 hours incubation and 8 hours incubation, note mild inhibition caused after 8 hours of incubation. (B) Soft agar assay with *L. plantarum* DSM13273 culture treated with 3 µL of N1 supernatant after 24 hours incubation and 8 hours incubation. Mild inhibitory activity of N1 visible after 24 hours but not at 8 hours.

4.2.9 Activating the plantaricin cluster in *L. plantarum* A-1

During the search for the secondary activity, a plantaricin cluster was found in the A-1 genome (Section 4.2.0). As the indicator strain used (*L. plantarum* ATCC 14917) is resistant to plantaricins, it is unlikely that the plantaricin is responsible for the secondary inhibition. However, it is known that the plantaricin cluster in the *L. plantarum* KW30 genome can be activated by the addition of plnA22 to the growth

media during incubation. PlnA22 is a signalling pheromone that can instigate plantaricin production in strains with an active plantaricin cluster.

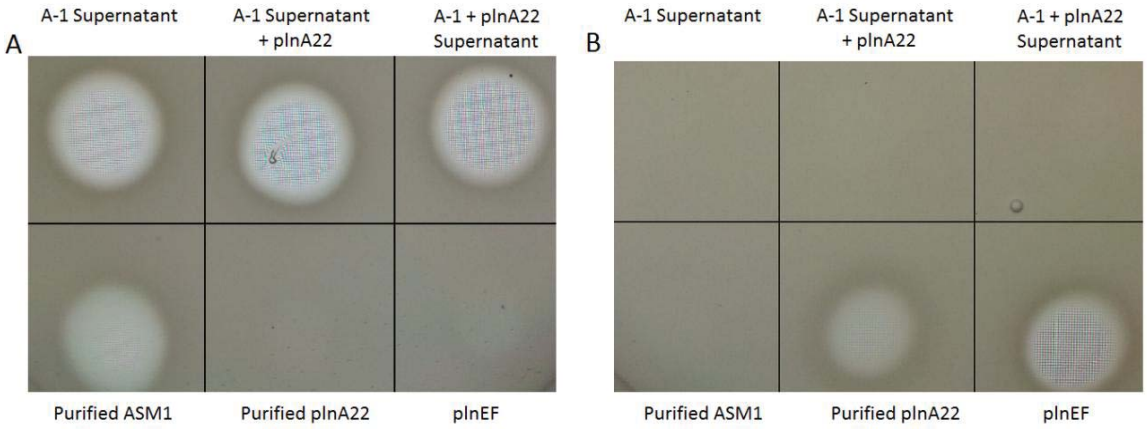


Figure 4.2.10 Agar plates showing inhibition caused by ASM1 and PlnEF
Soft agar assays showing lack of activation of the putative plantaricin cluster in *L. plantarum* A-1. (A) *L. plantarum* 14917 suspension treated with 4µL aliquots of bacteriocin and/or supernatant, as labelled, showing activity of ASM1. (B) *L. plantarum* 965 suspension treated with bacteriocin and/or supernatant, as labelled, showing A-1 does not produce a plantaricin peptide.

Soft agar assays using either a strain sensitive to glycocins and resistant to plantaricins (*L. plantarum* ATCC 14917) or a strain sensitive to plantaricins and resistant to glycocins (*L. plantarum* 965) were treated with either glycocin producer supernatant, glycocin producer supernatant and PlnA22, purified glycocin, purified PlnA22, PlnEF (the active, two-peptide plantaricin) or supernatant from a glycocin producer incubated in media containing PlnA22.

It was clearly shown that enough of the plantaricin cluster remains active to provide immunity to the plA22 solution but that the A-1 strain does not produce plnEF as the 965 strain is sensitive to plantaricin (Figure 4.2.10). Analysis of the flanking sequence in the contig that the plantaricin cluster was found in showed that only a few components of the plantaricin cluster remained. This included one of the three immunity proteins and the ABC transporter but no response regulator nor any

structural genes. A wider search of the genome for these sequences yielded no results, suggesting that most of the plantaricin cluster in *L. plantarum* A-1 is no longer present. In contrast, *L. plantarum* KW30 retains the cluster, and it was shown to be active using the same assay (Figure 4.2.11).

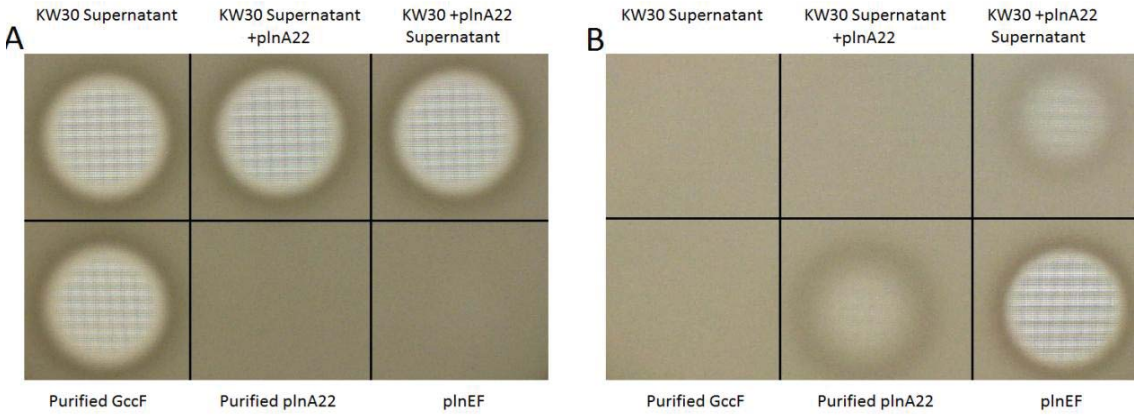


Figure 4.2.11 Agar plates showing inhibition caused by GccF and PlnEF
Soft agar assays showing the presence of anactive plantaricin cluster in *L. plantarum* KW30. (A) *L. plantarum* 14917 suspension treated with bacteriocin and/or supernatant, as labelled, showing activity of GccF bacteriocin. (B) *L. plantarum* 965 suspension treated with bacteriocin and/or supernatant, as labelled, showing KW30 production of plantaricin induced by plnA22. Figures from M. Patchett (2013b) (Personal Comm.)

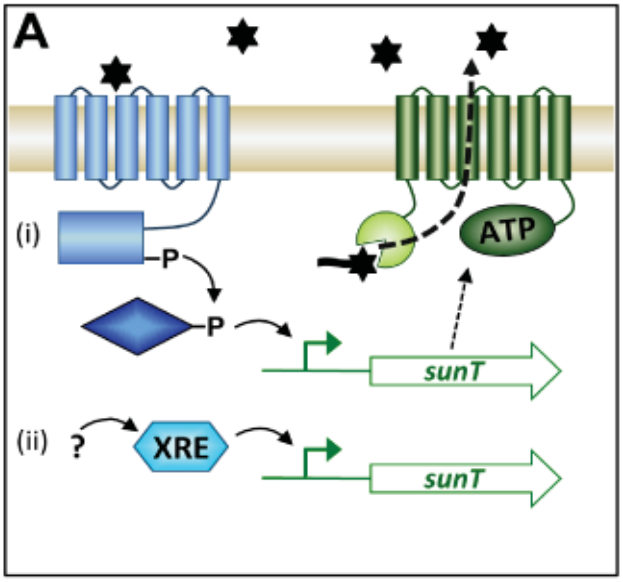
The loss of the plantaricin cluster from *L. plantarum* A-1 could suggest that the PlnEF was less effective in the environment that in which the strain evolved. It is possible that the ASM2 bacteriocin was acquired through horizontal gene transfer and was more useful as a cellular defense mechanism.

4.3 Bioinformatic analysis of selected bacteriocin-associated proteins in *L. plantarum* A-1

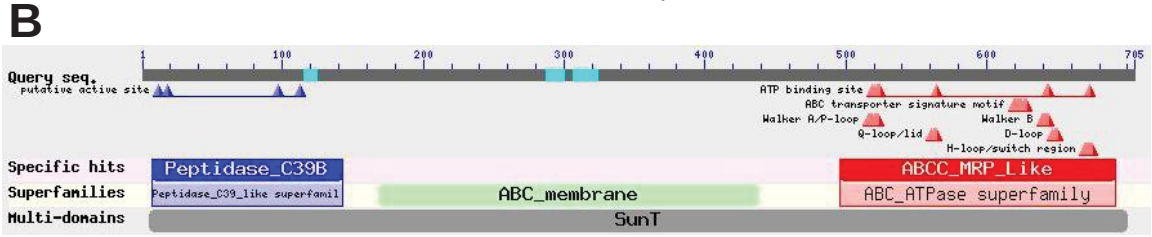
4.3.0 Bioinformatic analysis of bacteriocin peptidases

The ABC transporter encoded in many bacteriocin operons is a three-domain protein that contains a transmembrane region, the ATP-binding cassette that gives the transporter its name and, in bacteriocin operons, an N-terminal peptidase domain (Peptidase_C39) that is required for the removal of the leader sequence of the precursor peptide before it is exported from the cell as a mature bacteriocin (Figure 4.3.1). The Peptidase_C39 domain is useful in genome mining for bacteriocin clusters as it is a well conserved protein domain that is very strongly associated with bacteriocins (Dirix *et al.*, 2004; Diep *et al.*, 2009; Gebhard, 2012)

Figure 4.3.1 Relationship between the ABC transporter gene and its role in the cell



(A) Schematic of precursor peptide (**black**) cleavage and secretion by sunT transporter system (green) followed by activation of histidine kinase and response regulator (dark blue) resulting in a positive feedback loop. Note that a histidine kinase gene is not present in the *gcc* and *asm* operons, only the C-terminus of the response regulator has been identified (LytTR, see section 3.5.7). Figure from Gebhard (2012). Reproduced with permission from Wiley Online. (B) Schematic of the Sublancin SunT gene product, an ABC transporter. Domains predicted by BLASTP 2.2.29+ from peptide sequence.



As mentioned in section 4.2.1 the amino acid sequence of the ASM1 ABC Peptidase_C39 domain was used as the query in a protein BLAST (BLASTP) search to identify alternative putative bacteriocin clusters in the *L. plantarum* A-1 genome (see 2.2.5). This method resulted in the identification of two additional Peptidase_C39 domains. One of these (E-value $8e^{-09}$), shared 25% identity with the N-terminus of PlnG, the ABC transporter associated with the plantaricin cluster common to most *L. plantarum* strains (Diep *et al.*, 2009). The other (E-value $7e^{-11}$) shared 29% identity with part of ComA, an ABC transporter in *L. brevis* which is associated with a novel bacteriocin cluster, termed ASM2.

A T-COFFEE alignment of a small selection of Peptidase_C39 domains encoded by four plantaricin gene clusters was carried out (see 2.2.5). The alignment also included three glycocin clusters (ASM1, GccF and sublancin 168), a colicin cluster and one *E. faecalis* V583 cluster. Two highly conserved regions in the peptidase domain were apparent; one at the N-terminus and the other at the C-terminus (Figure 4.3.2).

In the alignment (Figure 4.3.2) the N-terminal conserved region holds two proposed catalytic residues; the first is a totally conserved cysteine residue and the second is a conserved glutamine, 6 residues N-terminal to the cysteine (residues 4 - 10 in ASM1). In addition to the catalytic site there is a fully conserved alanine residue (A12 in ASM1). The C-terminal conserved region contains a completely conserved histidine residue (H91 in ASM1) and a conserved aspartic acid and proline (D108 and P109, respectively, in ASM1). C-terminal of the second conserved region is a conserved aromatic amino acid (W/F) that may associate with the transmembrane domain of the ABC transporter (W130 in ASM1).

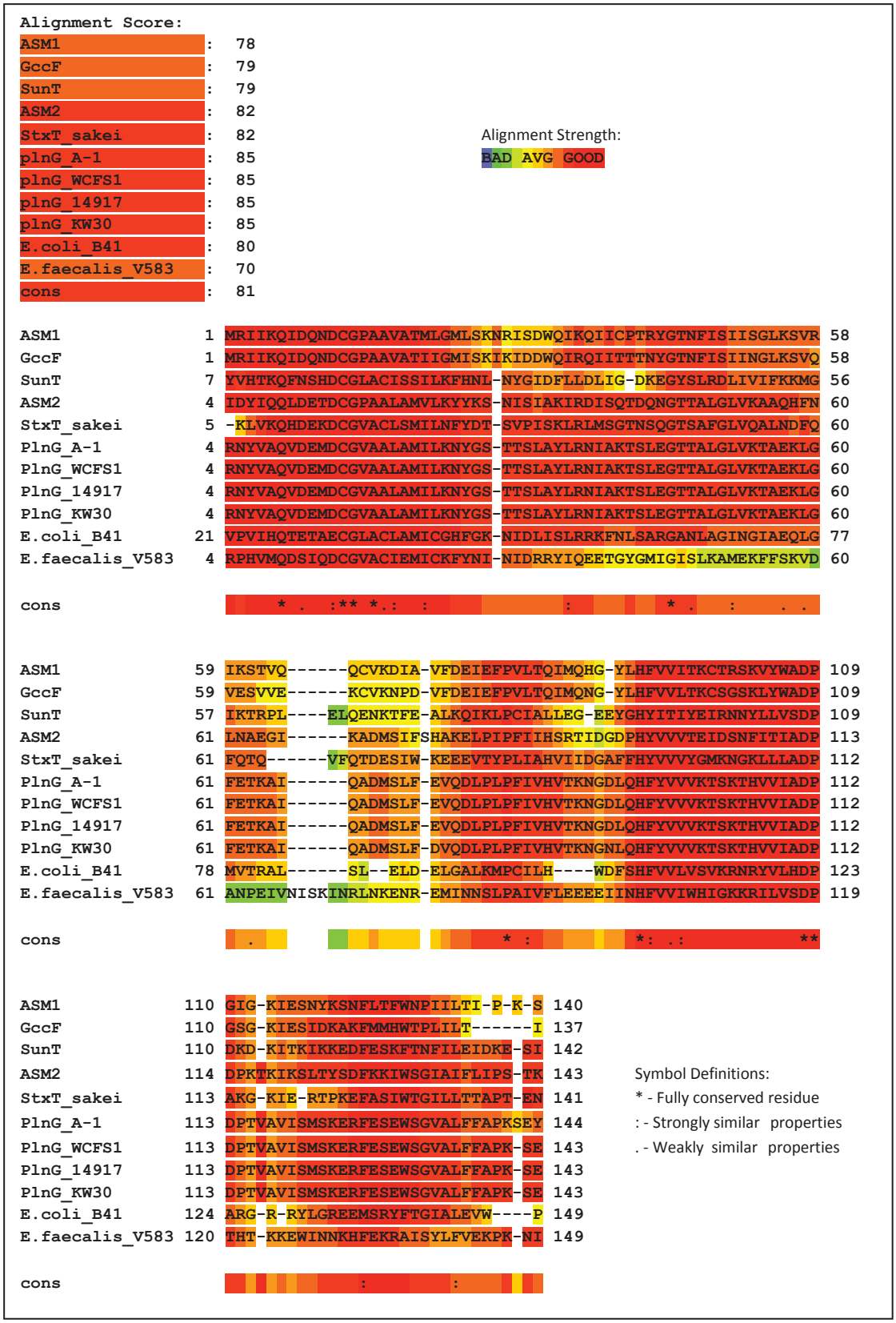


Figure 4.3.2 T-COFFEE multiple sequence alignment of bacteriocin Peptidase_C39 domains

Bacteriocin ABC Peptidase_C39 domains from *L. plantarum* A-1 (ASM1, ASM2 and PlnG_A-1), *L. plantarum* KW30 (GccF and plnG_KW30), *L. plantarum* WCFS1 (PlnG_WCFS1), *L. plantarum* ATCC 14917 (PlnG_14917), *B. subtilis* 168 (Sublancin_168), *E. coli* B41 and *E. faecalis* V583 with the conserved residues at the bottom of the alignment. Two dark red segments indicate highly conserved regions (1-24 and 92-110 in ASM1). Alignment score shows how well each sequence fits the set.

The sequence alignment also revealed that the glycocin (ASM1, GccF and sublancin 168) Peptidase_C39 sequences are less similar to the consensus sequence than the *E. coli* B41 colicin Peptidase_C39 sequence, as determined by the overall alignment scores (Figure 4.3.2). If the plantaricin cluster is the only bacteriocin cluster that developed within *L. plantarum* the glycocin Peptidase_C39 originates from a different species as, theoretically the *L. plantarum* derived sequence should have a high degree of similarity to the plantaricin sequences.

The peptidase domain of bacteriocin ABC transporters cleaves the leader sequence of the prepeptide before secretion. This would suggest that bacteriocins with similar Peptidase_C39 domains would have similar leader sequences on their precursor peptide because the Peptidase_C39 domain has to recognise and then cleave the leader sequence. To try and ascertain a reason for the differences between the glycocin peptidases and the ASM2 peptidase the leader sequences from the precursor peptides were compared and investigated using a T-COFFEE alignment (section 2.2.5).

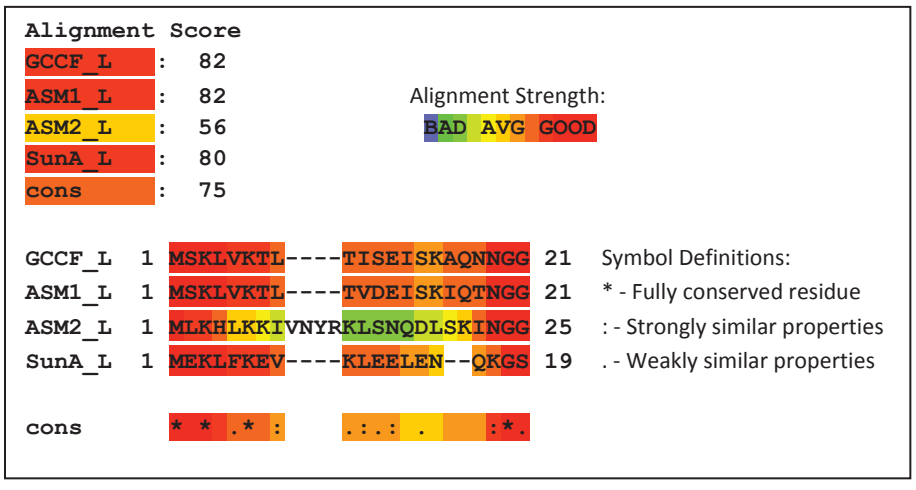


Figure 4.3.3 T-COFFEE Alignment of bacteriocin leader sequences
Leader sequences from ASM1, ASM2, GccF and Sublancin (SunA). Conserved double Glycine (GS in SunA) where the mature peptide sequence begins. Conserved sequence below aligned leader sequences.

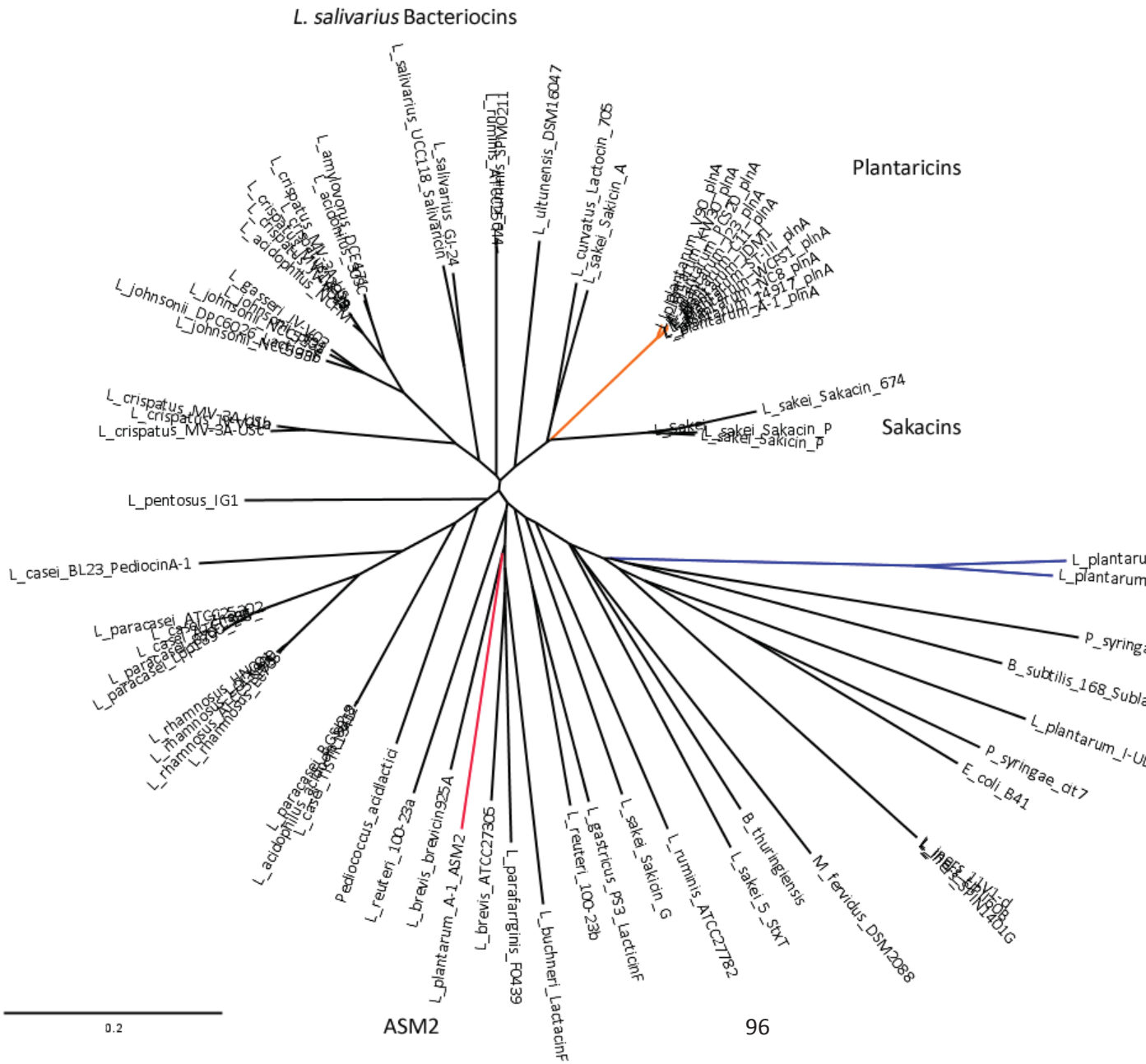
The alignment of leader sequences showed that, aside from the conserved double-glycine motif and N-terminus, the ASM2 leader sequence is notably different from the leader sequences for the glycocin precursor peptides, both in length and amino acid composition (Figure 4.3.3). This supports the differences found within the Peptidase_C39 domain as the two sequences are functionally linked in the bacteriocin secretion system.

The relationship between the peptidases of each bacteriocin was further investigated and a better evolutionary relationship between the sequences was established by acquiring a wider range of peptidase_C39 sequences using the Sanger Pfam database (Punta *et al.*, 2012). All of the Lactobacillae peptidase_C39 sequences in the database were collected and then all duplicated database entires were removed to leave a non-redundant set of sequences in which each unique C39 domain sequence from each organism occurs only once. This was done to avoid unbalanced rooting of the tree as the duplicated sequences can cause an incorrect consensus sequence to be generated, providing a misrepresentation of the average protein. These sequences were then combined with two Peptidase_C39 domains from ABC Transporters from characterised glycocins (SunT from *B. subtilis* 168 and the peptidase (locus tag bthur0009_56260) from *B. thuringiensis* BGSC 4AW1), and several other unrelated bacteriocin peptidases including two gram-negative Peptidase_C39 domains (from *Pseudomonas syringeae* and *E. coli* B41) and an outlier from an archaeal species (*Methanothermus fervidus* DSM 2088). The sequence data was combined with the ASM1, ASM2, and A-1 plnG Peptidase_C39 sequence data (GccF Peptidase_C39 sequence data is already part of the Sanger pfam database (see 2.2.5)) and used to

generate a phylogenetic tree based on the multiple sequence alignment created with MUSCLE (See 2.2.5). The resulting tree (generated using FigTree 1.4.0 (see 2.2.5)) allowed for clearer visualisation and easier interpretation of the relationship between the peptidase regions than provided by the T-COFFEE alignment provided as well as visualisation of the variety in the sequence (Figure 4.3.4).

In this larger set of data, it is shown that the ASM2 peptidase is no longer grouped in with the *L. plantarum* peptidases. These are clustered tightly into a well defined group due to most of the plantaricin ABC transporters sharing a 99% amino acid identity, regardless of their strain. It is apparent that many peptidase domains are highly conserved within their species (the transporters of *L. paracasei, L. rhamnosus* and the plantaricins of *L. plantarum*, for example) but between these groups there is a large diversity in the sequence no doubt reflecting of the massive variation in bacteriocin leader sequences recognised and cleaved by the Peptidase_C39 domain.

The ASM2 ABC peptidase region is on the fringes of the *Lactobacillus* groups, in a relatively diverse group with transporters from *L. brevis, L. reuteri* and others. The fact that the closest relative of the ASM2 peptidase in this tree differs by slightly more than 0.2 residues per site (a greater than 20% difference) suggests that the ASM2 bacteriocin did not originate from the *L. plantarum* species, and belongs to a group of bacteriocins that has not been fully characterised yet. The one known orthologue to ASM2, Bactofencin A, was isolated from *L. salivarius* DPC6502 (O'Shea *et al.*, 2013) but the Peptidase_C39 domain present in the *asm2* cluster of A-1 does not appear to be related to the Peptidase_C39 domains of other *L. salivarius* bacteriocins suggesting that it may not be native to the *L. salivarius* species either.


L. salivarius Bacteriocins
Plantaricins
Sakacins
L_salivarius_UCC118_Salivaricin
L_salivarius_G1-24
L_ultunensis_DSM16047
L_cuvatus_Lactocin_705
L_sakei_Sakicin_A
L_sakei_Sakacin_674
L_sakei_Sakacin_P
L_sakei_Sakicin_P
L_crispatus_MV-3A-US1a
L_pentosus_IG1
L_casei_BL23_PediocinA-1
L_paracasei_ATCC25302
L_rhamnosus_LRHMDP2
L_rhamnosus_ATCC53103
L_rhamnosus_LRHMDP3
L_acidophilus_NCFM
L_paracasei_BG66013
L_casei_ATCC334
Pediococcus_acidlactici
L_reuteri_100-23a
L_brevis_brevicin925A
L_plantarum_A-1_ASM2
L_brevis_ATCC27305
L_parafarraginis_F0439
L_buchneri_LactacinF
L_reuteri_100-23b
L_gastricus_PS3_LacticinF
L_sakei_Sakicin_G
L_ruminis_ATCC27782
L_sakei_5_StxT
B_thuringiensis
M_fervidus_DSM2088
E_coli_B41
P_syringae_cit7
B_subtilis_168_Subla
L_plantarum_I-UL
L_plantaru
L_plantarum
0.2
ASM2
96

The two C39-peptidases of the *Lactobacillus* glycocins, ASM1 and GccF, show a greater separation from the plantaricin peptidases than the ASM2 peptidase. The peptidases are not particularly similar (about 70% identity and 90% positive matches) and they are well separated from the nearest peptidase sequence, that from *P. syringae* B728a. The other glycocin peptidases (from *B. thuringiensis* and *B. subtilis*) are similarly separated from both one another and the *Lactobacillus* glycocin peptidases. This suggests independent evolution of the peptidases of these glycocins and potentially independent evolution of the *Bacillus* and *Lactobacillus* glycocins. Despite this, it is apparent that the A-1 and KW30 Peptidase_C39 came to the *Lactobacillus* genus as a result of horizontal gene transfer and subsequently partitioned into different strains, then diverged from there as they are closely related compared to other peptidases around them.

Despite these conclusions, it is acknowledged that the layout of these relationship trees can be changed by the addition of a few sequences. It is therefore possible that the discovery and genetic characterisation of more glycocins might alter the structure of the tree. It is also possible that their grouping could be an artefact of their relative rarity. The recently characterised homologue of ASM2, bactofencin A from *L. salivarius* is probably processed by a peptidase domain that shares homology with the ASM2 peptidase. A comparison of these could provide more clues as to the origin of the ASM2 bacteriocin but the amino acid sequence of the bactofencin ABC transporter is not yet available (O'Shea *et al.*, 2013).

4.3.1 Bioinformatic analysis of bacteriocin thioredoxin-like peptides

Thioredoxins (trx) are a diverse group of small, disulfide-containing redox proteins involved primarily in the reduction of proteins by cysteine thiol-disulfide exchange. This allows them to keep protein residues in the reduced state, preventing the formation of disulfide bonds early, as well as preventing unwanted oxidising of proteins. The formation of the correct disulfide bonds is essential for GccF, sublancin 168 and ASM1 and is potentially an important feature in ASM2. All four of these bacteriocins have putative thioredoxin-encoding genes in their operons, which would tend to suggest that the trx-like domain is related to the bacteriocin function. The thurandacin thioredoxin, the AsmC/GccC and the AsmD/GccD thioredoxins and the sublancin thioredoxin share the same conserved domain (cd02947) although the AsmD/GccD peptides are missing the last four residues of the conserved domain. It has been shown that in the GccF gene cluster, *gccC* and *gccD* are expressed at the same time as the structural gene *gccF* (Ahn, 2012b)(Personal Comm.). The ASM2 thioredoxin is not the same as the others (TIGR01295) which could suggest that the putative thioredoxin has a different role in the ASM2 peptide development than it does in the glycocins.

The glycocin thioredoxins and the ASM2 thioredoxin were used as input for a T-COFFEE multiple alignment to determine sequence similarity and conserved regions and to see if there were any major differences present in the sequences that may influence disulfide bond formation in the peptide. BsuBdbB (the second trx gene in sublancin) was not included in this alignment as the conserved domain is not a thioredoxin-like

domain but a disulfide oxoreductase and does not share much similarity with the trx-like domains.

The multiple sequence alignment shows that the two pairs of thioredoxin-like proteins (GccC/AsmC and GccD/AsmD) in ASM1 and GccF are highly conserved between the two species. AsmD and GccD are significantly longer than the other trx-domain containing peptides analysed. The extra amino acids contribute to a very long, very hydrophobic N-terminal region (L3 – I24) while the second, larger insert adds a long chain of predominantly charged or polar residues (T25 – T45) (Figure 4.3.5) which may possibly be involved with interactions with other proteins in the cell. Both of these inserts occur N-terminal of the first conserved region, and may affect the function of the trx-domain by altering the protein's interaction with substrates and neighbouring proteins.

The first of the two highly conserved regions contains the CXXC motif characteristic of thioredoxins. This region is less well conserved in the ASM2 trx. Although it still contains the CXXC active site motif, the amino acid sequence before this region has reduced sequence similarity. If these sequence changes result in the ASM2 trx being less active then that may explain the observation of O'Shea *et al.* (2013) that the disulfide bond was not required for bactofencin A activity, assuming that the bactofencin A gene cluster contains a trx protein.

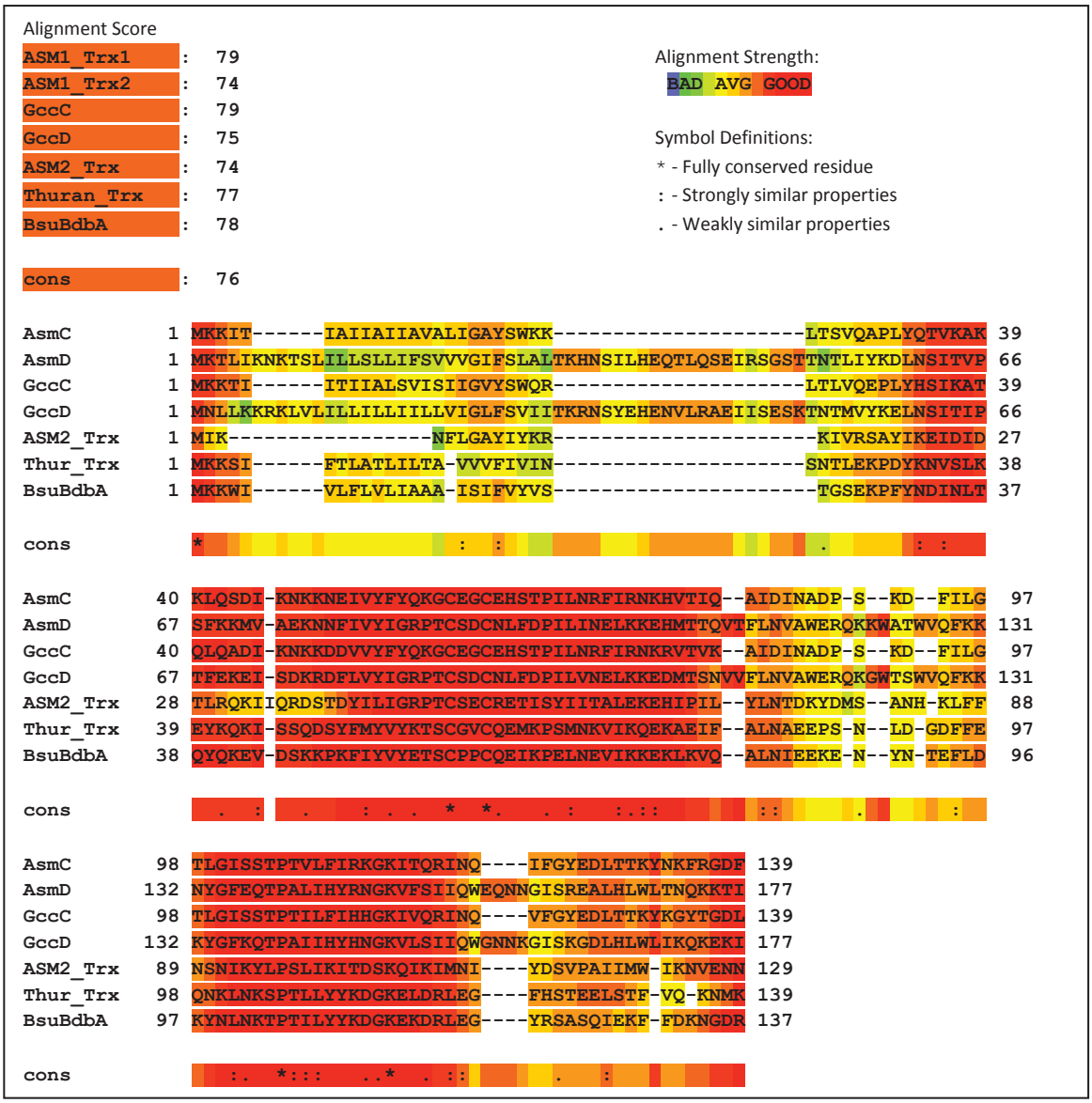


Figure 4.3.5 T-COFFEE multiple sequence alignment of thioredoxin-like domains

Sequences from the ASM1, ASM2, GccF, thurandacin and sublancin gene clusters with the conserved residues denoted below the alignment. Two dark red regions in the centre of each sequence denote highly conserved regions.

The second conserved region contains two fully conserved residues, a proline (P105 In AsmC) and a lysine (K114 in AsmC). These do not have roles in the formation of the disulfide bonds but are possibly involved in binding to the substrate. Though not as likely, the second conserved region could associate with a transport protein. The ASM2 trx-domain is recognised as TIGR01295 in the NCBI CDD (See 2.2.5) and this conserved domain has been associated with the transport of bacteriocins in *Streptococcus*

pneumoniae WU2 (Hardy *et al.*, 2000). As the ASM2 thioredoxin shares this sequence conservation with the other bacteriocin thioredoxins, it could be that this proposed association with transporters may be the role of this conserved region.

Although there are two redox proteins in the sublancin operon, only BdbA shares homology with the thioredoxin-like proteins. BdbB has a very different amino acid sequence and its predicted conserved domain is not a thioredoxin-like domain but rather a putative disulfide oxidoreductase (PRK03113). This protein has the same role in the cell, i.e. the reduction of amino acid residues, and even has the highly conserved CXXC motif but beyond that has neglible amino acid sequence similarity with the trx-like domains encoded by the two *Lactobacillus* glycocin gene clusters. A study into the thioredoxin and disulfide oxidoreductase in *B. subtilis* showed that BdbB is required for sublancin activity whereas BdbA can be knocked out without a reduction in sublancin activity (Dorenbos *et al.*, 2002). This would suggest that the trx-domain in BdbA is inactive or at least does not interact with the sublancin peptide. Taking these two pieces of information together with the high sequence conservation between BdbA and the glycocin thioredoxins suggests that the roles of the trx-like domains in the ASM1 and GccF in the maturation of their respective peptides is questionable. It may be that the extra amino acids in the AsmD/GccD proteins are involved in recognition of the substrate and that the AsmC and GccC proteins play a different role in the protein maturation and secretion, or even immunity

4.3.2 Bioinformatic analysis of bacteriocin structural peptides

In section 3.2.1 it was shown that ASM1 has a similar IC_{50} to GccF but slightly different activity at high and low concentrations. A T-COFFEE alignment of the two mature structural peptides without their leader peptide sequences (briefly analysed in 4.3.1) showed that the first 30 amino acids are identical (Figure 4.3.6).

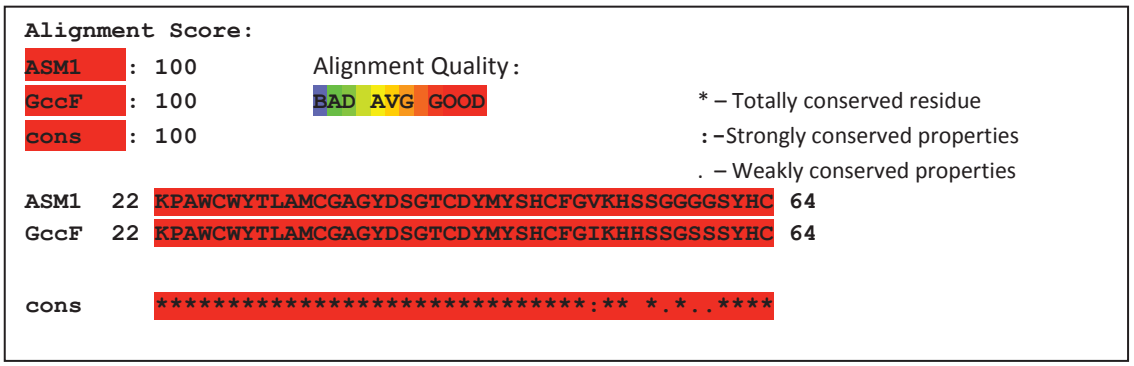


Figure 4.3.6 T-COFFEE alignment of ASM1 and GccF structural peptides
Identical residues from 22-53 (1-30 of mature peptide) apparent as * in conserved sequence (below the alignment)

All five amino acid sequence differences are in the flexible C-terminal tails of the peptides and could possibly affect mobility. While the function of this region is not known any loss of the two C-terminal residues and the *S*-linked GlcNAc moiety significantly reduce activity (Stepper *et al.*, 2011). It is possible that, changes in the sequence may affect mobility of the C-terminal tail and that the increased flexibility provided by the smaller residues in the ASM1 peptide may well reduce the conservation of structure in the peptide and cause the slightly reduced activity seen in section 3.2.1.

A manual alignment of the ASM1 and GccF structural peptides with two other known glycocins (sublancin 168 and disglycosylated thurandacin) was made in an effort to establish the presence of regions that may have been conserved between distantly

related species. A manual alignment was done as T-COFFEE alignments did not recognise the nested cysteine disulfides (C–X_6–C) as conserved regions in the peptides and attempted to match the peptides across their entire length, leading to misalignment.

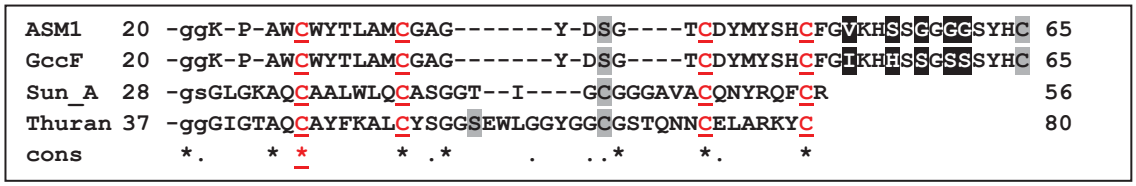


Figure 4.3.7 Manual sequence alignment of four glycocins
ASM1, GccF, sublancin 168 (SunA) and thurandacin (Thuran) using the C–X_6–C motifs as fixed references, cysteine residues shown in red. Glycosylated residues highlighted in grey and sequence differences between ASM1 and GccF in white

The manual sequence alignment (Figure 4.3.7) showed that, beyond the conserved C–X_6–C motifs, little sequence conservation is present in the four known glycocins. The totally conserved alanine residue two residues away from the first cysteine could suggest an important role in the formation of the disulfide bonds. Additionally, the conserved glycine next to the glycosylated residue on every glycocin could indicate that a small residue is important in allowing glycosylation of the cysteine/serine.

It is apparent from the manual alignment that the *Lactobacillus* glycocins are only related to the *Bacillus* glycocins distantly, if at all. The *Lactobacillus* glycocins have much shorter loops and have long C-terminal tails after the second C-X_6-C motif, suggesting different mechanisms of action and/or perhaps an additional inhibitory mechanism as they have a secondary sugar moiety attached to the C-terminal cysteine at the end of the flexible tail (Venugopal *et al.*, 2011).

The results of an *in silico* T-COFFEE alignment emphasise how different the sequences of these glycocins are, as it fails to align the second conserved C–X_6–C motif and rather

attempts to optimise the alignment of peptides in a region where there is negligible sequence conservation (Figure 4.3.8).

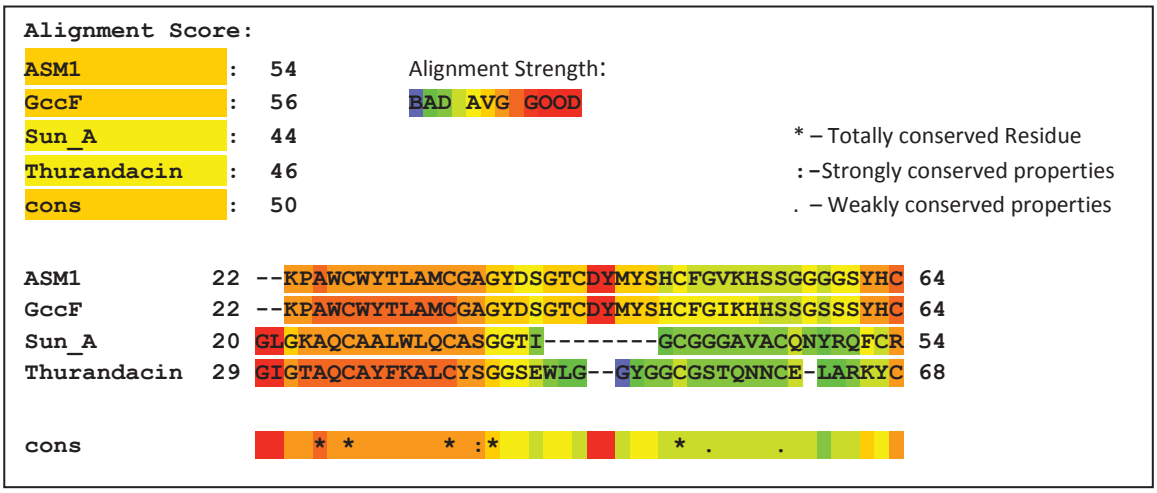


Figure 4.3.8 T-COFFEE alignment of the four glycocin structural peptides
Leader sequences not included as those are substantially different in length. Sequence conservation is below the four aligned sequences.

Although the first dozen residues are relatively well conserved in the T-COFFEE alignment, the sequence conservation begins to deteriorate as the loop structure of the *Bacillus* glycocins overlaps with the second C–X_6–C motif of the *Lactobacillus* glycocins. The conservation of the C-terminal cysteine would initially suggest that it has an important role in the bacteriocin function but is an artefact of the unanchored alignment as the C-terminal cysteines in the *Bacillus* glycocins are involved in the disulfide bonds and those in the *Lactobacillus* glycocins are glycosylated and at the end of flexible peptide tails.

It was shown that the N-terminal fragment of ASM1 (residues 22-53 in figure 4.3.8) is predominantly responsible for the inhibitory activity of the bacteriocin but has its activity greatly reduced if the C-terminal region (residues 54-64 in figure 4.3.8) is removed (see section 3.2.4). This would suggest that the sublancin and thurandacin (glycocins without flexible C-terminal tails) rely on a slightly different mechanism for

target inhibition. It is possible that their higher pI (8.3 for thurandacin and 11.8 for sublancin, compared to ~7 for the *Lactobacillus* glycocins) could be involved in the initial attraction to the target cell, whereas this role is fulfilled by the flexible tail in the *Lactobacillus* glycocins. Alternatively, it is possible full activity relies upon a flexible region with a glycosylated residue and that the longer loops of thurandacin and sublancin are flexible enough to allow for this activity.

4.3.3 Bioinformatic analysis of bacteriocin glycosyltransferases

The four known glycocins that have rare *S*-glycosidic bonds are glycosylated by a glycosyltransferases (GTases) encoded in the bacteriocin gene cluster. It has been shown that the thurandacin GTase (ThuS) can glycosylate both sublancin and thurandacin but that the sublancin GTase (SunS) cannot glycosylate thurandacin. This suggests that the GTase proteins are quite similar but that the SunS GTase has higher substrate specificity than the ThuS GTase (Wang *et al.*, 2014).

In ASM1 and GccF, the GTase is encoded upstream of the bacteriocin gene and is separated from the structural gene by several genes (Stepper *et al.*, 2011). In the sublancin and thurandacin operons the GTase is encoded immediately downstream of the bacteriocin gene (Garcia De Gonzalo *et al.*, 2014; Wang *et al.*, 2014). In addition to this, the thurandacin, ASM1 and GccF GTases can create both *S*- and *O*- glycosidic bonds (Stepper *et al.*, 2011; Wang *et al.*, 2014). A multiple sequence alignment using T-COFFEE was carried out to investigate the similarity of the of the GTase proteins.

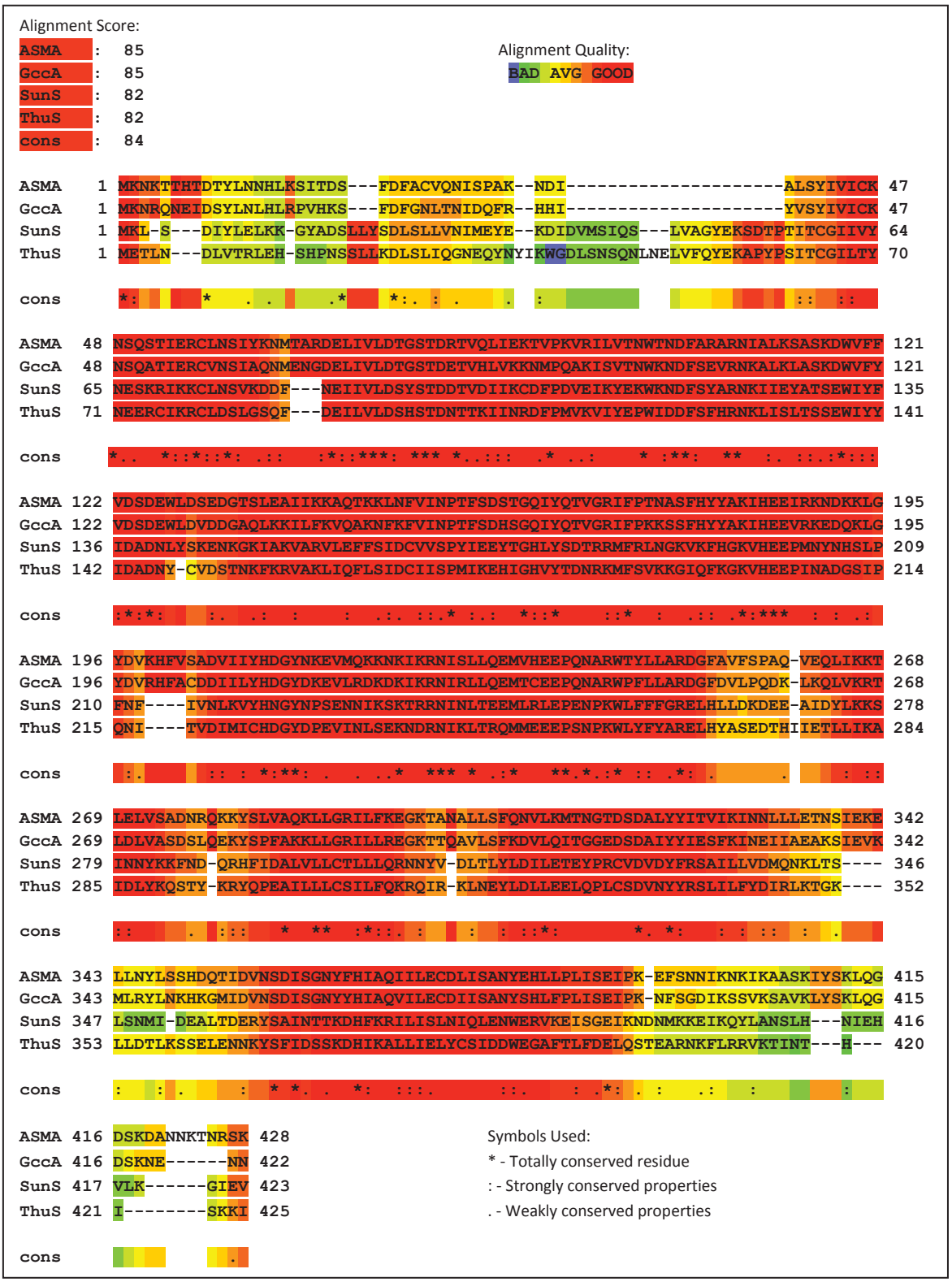


Figure 4.3.9 T-COFFEE multiple sequence alignment of glycocin Glycosyltransferases

GTase sequences from ASM1, GccF, sublancin and thurandacin glycosyltransferases. Large conserved region visible with multiple totally conserved residues and several small less conserved gaps (residues 44-331 in ASMA). A second, much smaller conserved region is also present (residues 347 - 392 in ASMA) Sequence conservation below aligned sequences (cons).

This *in silico* analysis showed a very high degree of sequence conservation, both overall (the alignment score indicating how well the sequence in question fits the set as a whole) and in the conserved GTase domains within the protein (Figure 4.3.9).

The alignment highlights the relationship between the GTase proteins, showing notable intergenus differences visible in the poorly conserved regions (Residues 9-22 in ASMA and GccA and residues 42-57 in SunS are a good example of this). This supports the previous evidence from the alignments of the peptidase regions and the bacteriocin peptides that the two species' bacteriocins evolved independently. Despite this, there is a large, highly conserved region present in the protein (residues 44-331 in ASMA) that contains a large number of totally conserved amino acids. This conservation extends beyond the Beta4Glucosyltransferase conserved domain by approximately 80 residues, suggesting that those 80 residues have an important role in the function of the GTase in the glycocin post-translational modification. A second, much smaller conserved region is present at residues 347-392 in ASMA. This region is less well conserved and does not appear to have any obvious characteristics that may suggest its role in the protein.

As there are only a few small regions in the conserved region of the protein that differ between these proteins (Residues 147-150 and 268-279 in ThuS, for example) these could potentially be the subjects of future studies to determine the reasons for the differences in substrate specificity between SunS and ThuS established by Wang and van der Donk (2011) and Wang *et al.* (2014).

5.0 Conclusions and looking forwards

The results presented in this thesis expand the knowledge of the glycocin ASM1, produced by *Lactobacillus plantarum* A-1, and also open the door to new bacteriocin research on a second bacteriocin (ASM2), the activity of which was masked by the greater activity of ASM1 against the standard *Lactobacillus* indicator strains.

Characterisation of the activity of the ASM1 bacteriocin showed that the glycocin has a very similar specific inhibitory activity to GccF (as measured by IC_{50}), an orthologous glycocin produced by *Lactobacillus plantarum* KW30. This similarity is perhaps not surprising given that the two glycocins have identical amino acid sequences for the first 30 of 43 residues. It was also shown that addition of free GlcNAc to MRS media reduces the inhibitory activity of ASM1, as it does with GccF. Finally, modifications to the bacteriocin (de-glycosylation and proteolytic cleavage) showed that the fragments of ASM1 have the same activity as their GccF counterparts, despite some differences in C-terminal sequence.

Investigations into the effects of physiochemical conditions on ASM1 activity demonstrated increased activity at lower pH (more acidic values) and temperatures closer to room temperature; these conditions probably mimic the environment the A-1 strain was originally isolated from (Hata *et al.*, 2010).

Investigations into the variety of species that ASM1 inhibits showed that the range of activity exhibited by ASM1 was as broad as that of GccF but the two glycocins have slightly different levels of activity against strains within this range (Kerr, 2013). It was also shown that the bacteriocin's range of activity was greater than previously thought

(Hata *et al.*, 2010), affecting a variety of *Lactobacillus* species and *E. faecalis* strains. However, this activity was observed to be highly strain specific, possibly reflecting differences in particular GlcNAc processing and signalling pathways within the cell as it is probable that the *Lactobacillus* glycocins target the GlcNAc-PTS transporters on cells, using the GlcNAc moieties as a sort of bait.

DNA sequence analysis revealed that the gene for the ASM1 bacteriocin is part of complete bacteriocin gene cluster in the 11,905 bp pA1_ASM1 plasmid in contrast to the orthologous *gcc* gene cluster which is chromosomally encoded. As is the case with the *gcc* cluster, the *asm* cluster does not contain a histidine protein kinase gene though it does contain the other half of the two-component regulatory system, the LytTR transcription regulation domain. The *asm* and *gcc* gene clusters share a high degree of similarity both in gene organisation and the amino acid sequence of the encoded proteins, although there are sequence differences that suggest the clusters have been evolving separately for quite some time.

Transferring other elements of the gene cluster to a shuttle vector that can replicate in both *E. coli* and *Lactobacillus* would potentially enable the production of ASM1 in a heterologous host and allow for more comprehensive analysis of the regulation, expression, and maturation of the protein. The ability to express individual genes from the glycocin cluster would allow for the analysis of activity of the proteins *in vivo*. This would allow for the determination of the role of the thioredoxin-like proteins, as their role in the cell has been called into question during this work. In addition to this the single-gene expression could potentially answer questions about the immunity protein. The lack of a *gccI* orthologue in the *asm* cluster suggested that it may not be an

immunity protein involved in the glycocin systems of KW30 or A-1 and that the *gcc/asmH* could instead be the immunity protein. Further studies in expressing the *H* gene in heterologous strains to determine whether it induces immunity would resolve this question. The presence of a *gccI* orthologue in the A-1 genome could indicate that it plays a secondary role in bacteriocin immunity.

The novobiocin-induced generation of ASM1-minus mutants of *L. plantarum* A-1 led to the identification of a second bacteriocin, ASM2, that is orthologous to a recently characterised *L. salivarius* bacteriocin called bactofencin A (O'Shea *et al.*, 2013).

It was determined that ASM2 is encoded in a minimalistic bacteriocin gene cluster that contains only four ORFs and is probably encoded on another plasmid in the A-1 strain. Though the *asm2* gene cluster contains the bare essentials for bacteriocin production it does not encode a histidine protein kinase or a transcription regulator, suggesting that it is either regulated elsewhere or that the bacteriocin is constitutively expressed. ASM2 bears similarity in structure to the ASM1 peptide as it, too, is a short peptide bound into a tight loop with a disulfide bond, though only one disulfide bond is present in ASM2. Synthetic disulfide-bonded ASM2 was shown to inhibit fewer of the tested indicator species than ASM1 and the inhibition seen in plate assays was less complete in almost all cases.

Bioinformatic analyses of a selection of bacteriocin-related proteins encoded in genome suggest that the *L. plantarum* glycocins (GccF and ASM1) are most likely a result of horizontal gene transfer from other species, due to their dissimilarity to other *Lactobacillus* proteins. In addition to this, it was proposed that the plasmid-based ASM2 cluster also originated outside the *L. plantarum* species as it is apparently as

unrelated to the glycocin clusters as it is to the archetypal *L. plantarum* plantaricin cluster.

Analysis of the four proven glycocins (GccF, ASM1, sublancin & thurandacin) (Table 5.0.1) suggested that, despite their nested disulfide bond architecture, $(C–X_6–C)_2$, the peptides from *L. plantarum* were markedly different to those in *Bacillus*, and are potentially not evolutionarily related indicating that the glycocin scaffold could have arisen independently on multiple occasions, and that there is a strong selective pressure towards this conformation, similar to the selective pressure for $(C–X_3–C)_2$ architecture seen in scorpion toxins (Chagot *et al.*, 2005).

Organism	Bacteriocin	Glycosyl-ation	Disulfide Bonds	Location of gene cluster	Active	Reference
***L. plantarum* A-1**	ASM1	*O*- and *S*-	2	Plasmid	Yes	(Hata *et al.*, 2010)
	ASM2	None	1	Plasmid	Yes	This work
	PlnEF/JK	N/A	N/A	Chromosome	No	This work
***L. plantarum* KW30**	GccF	*O*- and *S*-	2	Chromosome	Yes	(Stepper *et al.*, 2011)
	PlnEF/JK	None	0	Chromosome	Yes	(Diep *et al.*, 2009)
***L. salivarius* DPC6502**	Bactofencin A	None	1	Chromosome	Yes	(O'Shea *et al.*, 2013)
***B. subtilis* 168**	Sublancin 168	*S*-	2	Chromosome	Yes	(Paik *et al.*, 1998)
***B. thuringiensis* BGSC 4AW1**	Thurandacin	*O*- and *S*-	2	Unknown	Yes	(Wang *et al.*, 2014)

Table 5.0.1 Summary of the bacteriocins important to this work
Bacteriocins listed next to the strain of origin. Characteristics of *L. plantarum* A-1 PlnEF/JK could not be ascertained as the structural genes are not present in the genome.

indicates that the inhibitory action of the supernatant of A-1 and potentially any other bacteriocin-producing strain may not be due solely to the first bacteriocin identified in

the strain, suggesting that supernatant testing may give an inaccurate picture of the phylogenetic range and overall strength of the inhibition caused by a novel bacteriocin. Thus it is preferable, where possible, to test purified bacteriocins. In addition to this, further characterisation of the ASM2 bacteriocin with regards to its range of activity and inhibitory activity could be carried out, as well as determining the role of the disulfide bond in the peptide (O'Shea, *et al.,* 2013).

The brief analysis of bacteriocin-related proteins has opened up an avenue for future research with regards to the origins and genetic context of bacteriocin genes. The variation present between established groups of bacteriocins combined with the proposed diversity of bacteriocins within strains, suggests that the "bacteriocin library" of a single strain may be broader than previously thought as well as being varied between strains of the same species (Table 5.0.1).

6.0 Appendices

6.1 Appendix - PCR Protocols

Protocol I - PCR Probing

Step	Purpose	Temperature (°C)	Duration (seconds)
1	Initial Melt	94	60
2	Melt	94	30
3	Annealing	56	40
4	Elongation	72	35
5	Final Elongation	72	300
6	Storage	10	Indefinite

Table 6.1.1 PCR probing protocols
Protocols for analysing content of the A-1 genome. Repeat steps 2-5 for 28 cycles

Protocol II - pTRK669 confirmation

Step	Purpose	Temperature (°C)	Duration (seconds)
1	Initial Melt	94	60
2	Melt	94	30
3	Annealing	62	40
4	Elongation	72	60
5	Final Elongation	72	300
6	Storage	10	Indefinite

Table 6.1.2 PCR pTRK669 probing protocols
Protocols for confirming the presence of pTRK669 in transformants. Repeat steps 2-5 for 30 cycles

Protocol III- PCR Products for sequencing

Step	Purpose	Temperature (°C)	Duration (seconds)
1	Initial Melt	94	60
2	Melt	94	30
3	Annealing	64	40
4	Elongation	72	90
5	Final Elongation	72	300
6	Storage	10	Indefinite

Table 6.1.3 PCR protocols for sequencing products
Protocols to generate lengths of pA1_ASM1 DNA for sequencing. Repeat steps 2-5 for 35 cycles

6.2 Appendix – Raw CARY Data

See CD with Excel files included in printed copy of thesis

6.3 Images of ASM2 spectrum of activity plates

See CD for all appropriately labelled figures in .pdf format

7.0 References

Ahn, S. (2012a). [PCR Primers for analysis of *L. plantarum* A-1 genome].

Ahn, S. (2012b). [Premilinary findings from rtPCR of genes in the *gcc* cluster of *Lactobacillus plantarum* KW30].

Altschul, S. F., Gish, W., Miller, W., Myers, E. W., & Lipman, D. J. (1990). Basic local alignment search tool. *Journal of Molecular Biology, 215*(3), 403-410.

Altschul, S. F., Madden, T. L., Schäffer, A. A., Zhang, J., Zhang, Z., Miller, W., & Lipman, D. J. (1997). Gapped BLAST and PSI-BLAST: a new generation of protein database search programs. *Nucleic Acids Research, 25*, 3389-3402.

Arnison, P. G., Bibb, M. J., Bierbaum, G., Bowers, A. A., Bugni, T. S., Bulaj, G., . . . 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. doi: 10.1039/c2np20085f

Aukrust, T., & Blom, H. (1992). Transformation of *Lactobacillus* strains used in meat and vegetable fermentations. *Food Research International, 25*(4), 253-261. doi: 10.1016/0963-9969(92)90121-k

Ben Omar, N., Abriouel, H., Keleke, S., Valenzuela, A. S., Martinez-Canamero, M., Lopez, R. L., . . . Galvez, A. (2008). Bacteriocin-producing *Lactobacillus* strains isolated from poto poto, a Congolese fermented maize product, and genetic fingerprinting of their plantaricin operons. *International Journal of Food Microbiology, 127*(1-2), 18-25. doi: 10.1016/j.ijfoodmicro.2008.05.037

Bendjeddou, K., Fons, M., Strocker, P., & Sadoun, D. (2012). Characterization and purification of a bacteriocin from *Lactobacillus paracasei* subsp. *paracasei* BMK2005, an intestinal isolate active against multidrug-resistant pathogens. *World Journal of Microbiology & Biotechnology, 28*(4), 1543-1552. doi: 10.1007/s11274-011-0958-1

Bradford, M. M. (1976). Rapid and sensitive method for quantitation of microgram quantities of protein utilizing principle of protein-dye binding. *Analytical Biochemistry, 72*(1-2), 248-254. doi: 10.1006/abio.1976.9999

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

Brumfitt, W., Salton, M. R. J., & Hamilton-Miller, J. M. T. (2002). Nisin, alone and combined with peptidoglycan-modulating antibiotics: activity against methicillin-resistant *Staphylococcus aureus* and vancomycin-resistant enterococci. *Journal of Antimicrobial Chemotherapy, 50*(5), 731-734. doi: 10.1093/jac/dkf190

Cabo, M. L., Torres, B., Herrera, J. J. R., Bernardez, M., & Pastoriza, L. (2009). Application of nisin and pediocin against resistance and germination of Bacillus spores in sous vide products. *Journal of Food Protection, 72*(3), 515-523.

Chagot, B., Dai, L., Pil, J., Tytgat, J., Nakajima, T., Corzo, G., . . . Ferrat, G. (2005). An unusual fold for potassium channel blockers: NMR structure of three toxins from the scorpion *Opisthacanthus madagascariensis*. *Biochemical Journal, 388*, 263-271. doi: 10.1042/bj20041705

Chen, H., & Hoover, D. G. (2003). Bacteriocins and their food applications. *Comprehensive Reviews in Food Science and Food Safety, 2*(3), 82-100. doi: 10.1111/j.1541-4337.2003.tb00016.x

Cintas, L. M., Casaus, M. P., Herranz, C., Nes, I. F., & Hernandez, P. E. (2001). Review: Bacteriocins of lactic acid bacteria. *Food Science and Technology International, 7*(4), 281-305. doi: 10.1006/r8de-p6hu-clxp-5ryt

Cleveland, J., Montville, T. J., Nes, I. F., & Chikindas, M. L. (2001). Bacteriocins: safe, natural antimicrobials for food preservation. *International Journal of Food Microbiology, 71*(1), 1-20. doi: 10.1016/s0168-1605(01)00560-8

Cock, P. J. A., & Whitworth, D. E. (2010). Evolution of relative reading frame bias in unidirectional prokaryotic gene overlaps. *Molecular Biology and Evolution, 27*(4), 753-756. doi: 10.1093/molbev/msp302

Corr, S. C., Li, Y., Riedel, C. U., O'Toole, P. W., Hill, C., & Gahan, C. G. M. (2007). Bacteriocin production as a mechanism for the antfinfective activity of *Lactobacillus salivarius* UCC118. *Proceedings of the National Academy of Sciences of the United States of America, 104*(18), 7617-7621. doi: 10.1073/pnas.0700440104

Cotter, P. D., Hill, C., & Ross, R. P. (2005). Bacteriocins: Developing innate immunity for food. *Nature Reviews Microbiology, 3*(10), 777-788. doi: 10.1038/nrmicro1240

Cotter, P. D., Ross, R. P., & Hill, C. (2013). Bacteriocins - a viable alternative to antibiotics? *Nature Reviews Microbiology, 11*(2), 95-105. doi: 10.1038/nrmicro2937

David, M. Z., & Daum, R. S. (2010). Community-associated methicillin-resistant *Staphylococcus aureus*: epidemiology and clinical consequences of an emerging epidemic. *Clinical Microbiology Reviews, 23*(3), 616. doi: 10.1128/cmr.00081-09

Destoumieux-Garzon, D., Peduzzi, J., & Rebuffat, S. (2002). Focus on modified microcins: structural features and mechanisms of action. *Biochimie, 84*(5-6), 511-519. doi: 10.1016/s0300-9084(02)01411-6

Deurenberg, R. H., Vink, C., Kalenic, S., Friedrich, A. W., Bruggeman, C. A., & Stobberingh, E. E. (2007). The molecular evolution of methicillin-resistant *Staphylococcus aureus*. *Clinical Microbiology and Infection, 13*(3), 222-235. doi: 10.1111/j.1469-0691.2006.01573.x

Diaz, M., Valdivia, E., Martinez-Bueno, M., Fernandez, M., Soler-Gonzalez, A. S., Ramirez-Rodrigo, H., & Maqueda, M. (2003). Characterization of a new operon, as-48EFGH, from the as-48 gene cluster involved in immunity to enterocin AS-48. *Applied and Environmental Microbiology, 69*(2), 1229-1236. doi: 10.1128/aem.69.2.1229-1236.2003

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

Diep, D. B., Straume, D., Kjos, M., Torres, C., & Nes, I. F. (2009). An overview of the mosaic bacteriocin pln loci from *Lactobacillus plantarum*. *Peptides, 30*(8), 1562-1574. doi: 10.1016/j.peptides.2009.05.014

Dirix, G., Monsieurs, P., Marchal, K., Vanderleyden, J., & Michiels, J. (2004). Screening genomes of gram-positive bacteria for double-glycine-motif-containing peptides. *Microbiology-Sgm, 150*, 1121-1126. doi: 10.1099/mic.0.27040-0

Dorenbos, R., Stein, T., Kabel, J., Bruand, C., Bolhuis, A., Bron, S., . . . van Dijl, J. M. (2002). Thiol-disulfide oxidoreductases are essential for the production of the lantibiotic sublancin 168. *Journal of Biological Chemistry, 277*(19), 16682-16688. doi: 10.1074/jbc.M201158200

Edgar, R. C. (2004). MUSCLE: multiple sequence alignment with high accuracy and high throughput. *Nucleic Acids Research, 32*(5), 1792-1797.

Eijsink, V. G. H., Skeie, M., Middelhoven, P. H., Brurberg, M. B., & Nes, I. F. (1998). Comparative studies of class IIa bacteriocins of lactic acid bacteria. *Applied and Environmental Microbiology, 64*(9), 3275-3281.

Fath, M. J., & Kolter, R. (1993). ABC Transporters - Bacterial exporters. *Microbiological Reviews, 57*(4), 995-1017.

Fukao, M., Oshima, K., Morita, H., Toh, H., Suda, W., Kim, S. W., . . . Yajima, N. (2013). Genomic analysis by deep sequencing of the probiotic *Lactobacillus brevis* KB290 harboring nine

plasmids reveals genomic stability. *Plos One, 8*(3). doi: e6052110.1371/journal.pone.0060521

Garcia De Gonzalo, C. V., Zhu, L., Oman, T. J., & van der Donk, W. A. (2014). NMR structure of the S-Linked glycopeptide sublancin 168. *ACS Chemical Biology*. doi: 10.1021/cb4008106

Gebhard, S. (2012). ABC transporters of antimicrobial peptides in Firmicutes bacteria - phylogeny, function and regulation. *Molecular Microbiology, 86*(6), 1295-1317. doi: 10.1111/mmi.12078

Giacometti, A., Cirioni, O., Barchiesi, F., Fortuna, M., & Scalise, G. (1999). In-vitro activity of cationic peptides alone and in combination with clinically used antimicrobial agents against *Pseudomonas aeruginosa*. *Journal of Antimicrobial Chemotherapy, 44*(5), 641-645. doi: 10.1093/jac/44.5.641

Gong, H. S., Meng, X. C., & Wang, H. (2009a). Plantaricin MG active against Gram-negative bacteria produced by *Lactobacillus plantarum* KLDS1.0391 isolated from "Jiaoke", a traditional fermented cream from China. *Food Control, 21*(1), 89-96. doi: 10.1016/j.foodcont.2009.04.005

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

Hardy, G. G., Caimano, M. J., & Yother, J. (2000). Capsule biosynthesis and basic metabolism in *Streptococcus pneumoniae* are linked through the cellular phosphoglucomutase. *Journal of Bacteriology, 182*(7), 1854-1863. doi: 10.1128/jb.182.7.1854-1863.2000

Hata, T., Tanaka, R., & Ohmomo, S. (2010). Isolation and characterization of plantaricin ASM1: A new bacteriocin produced by *Lactobacillus plantarum* A-1. *International Journal of Food Microbiology, 137*(1), 94-99. doi: DOI 10.1016/j.ijfoodmicro.2009.10.021

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

Heu, S., Oh, J., Kang, Y., Ryu, S., Cho, S. K., Cho, Y., & Cho, M. (2001). gly gene cloning and expression and purification of glycinecin A, a bacteriocin produced by *Xanthomonas campestris* pv. *glycines* 8ra. *Applied and Environmental Microbiology, 67*(9), 4105-4110. doi: 10.1128/aem.67.9.4105-4110.2001

Hibbing, M. E., Fuqua, C., Parsek, M. R., & Peterson, S. B. (2010). Bacterial competition: surviving and thriving in the microbial jungle. *Nature Reviews Microbiology, 8*(1), 15-25. doi: 10.1038/nrmicro2259

Holtsmark, I., Mantzilas, D., Eijsink, V. G. H., & Brurberg, M. B. (2007). The tomato pathogen *Clavibacter michiganensis* ssp *michiganensis*: producer of several antimicrobial substances. *Journal of Applied Microbiology, 102*(2), 416-423. doi: 10.1111/j.1365-2672.2006.03091.x

Hsu, S. T. D., Breukink, E., Tischenko, E., Lutters, M. A. G., de Kruijff, B., Kaptein, R., . . . van Nuland, N. A. J. (2004). The nisin-lipid II complex reveals a pyrophosphate cage that provides a blueprint for novel antibiotics. *Nature Structural & Molecular Biology, 11*(10), 963-967. doi: 10.1038/nsmb830

Kamensek, S., & Zgur-Bertok, D. (2013). Global transcriptional responses to the bacteriocin colicin M in *Escherichia coli*. *Bmc Microbiology, 13*. doi: 4210.1186/1471-2180-13-42

Karthikeyan, V., & Santosh, S. W. (2009). Isolation and partial characterization of bacteriocin produced from *Lactobacillus plantarum*. *African Journal of Microbiology Research, 3*(5), 233-239.

Karthikeyan, V., & Santosh, S. W. (2010). Comparing the efficacy of plasmid curing agents in *Lactobacillus acidophilus*. *Beneficial Microbes, 1*(2), 155-158. doi: 10.3920/bm2009.0038

Kelly, W. J., Asmundson, R. V., & Huang, C. M. (1996). Characterization of plantaricin KW30, a bacteriocin produced by *Lactobacillus plantarum*. *Journal of Applied Bacteriology, 81*(6), 657-662. doi: 10.1111/j.1365-2672.1996.tb03561.x

Kerr, A., & Htay, K. (1974). Biological-control of crown gall through bacteriocin production. *Physiological Plant Pathology, 4*(1), 37-&. doi: 10.1016/0048-4059(74)90042-3

Kerr, A. P. (2013). *Expanding the microbiological characteristics of the bacteriocin, glycocin F, produced by Lactobacillus plantarum KW30.* Masters in Science, Massey University, Palmerston North.

Kerr, B., Riley, M. A., Feldman, M. W., & Bohannan, B. J. M. (2002). Local dispersal promotes biodiversity in a real-life game of rock-paper-scissors. *Nature, 418*(6894), 171-174. doi: 10.1038/nature00823

Kjos, M., Borrero, J., Opsata, M., Birri, D. J., Holo, H., Cintas, L. M., . . . Diep, D. B. (2011). Target recognition, resistance, immunity and genome mining of class II bacteriocins from Gram-positive bacteria. *Microbiology-Sgm, 157*, 3256-3267. doi: 10.1099/mic.0.052571-0

Klaenhammer, T. R. (1993). Genetics of bacteriocins produced by lactic-acid bacteria. *FEMS Microbiology Reviews, 12*(1-3), 39-86. doi: 10.1111/j.1574-6976.1993.tb00012.x

Kojic, M., Strahinic, I., Fira, D., Jovcic, B., & Topisirovic, L. (2006). Plasmid content and bacteriocin production by five strains of *Lactococcus lactis* isolated from semi-hard homemade cheese. *Canadian Journal of Microbiology, 52*(11), 1110-1120. doi: 10.1139/w06-072

Koponen, O., Tolonen, M., Qiao, M. Q., Wahlstrom, G., Helin, J., & Saris, P. E. J. (2002). NisB is required for the dehydration and NisC for the lanthionine formation in the post-translational modification of nisin. *Microbiology-Sgm, 148*, 3561-3568.

Lavermicocca, P., Lonigro, S. L., Valerio, F., Evidente, A., & Visconti, A. (2002). Reduction of olive knot disease by a bacteriocin from *Pseudomonas syringae* pv. *ciccaronei*. *Applied and Environmental Microbiology, 68*(3), 1403-1407. doi: 10.1128/aem.68.3.1403-1407.2002

Lazdunski, C. J. (1988). Pore-forming colicins - synthesis, extracellular release, mode of action, immunity. *Biochimie, 70*(9), 1291-1296. doi: 10.1016/0300-9084(88)90197-6

Lee, H., & Kim, H. Y. (2011). Lantibiotics, class I bacteriocins from the genus *Bacillus*. *Journal of Microbiology and Biotechnology, 21*(3), 229-235. doi: 10.4014/jmb.1010.10017

Lee, H. I., Kim, M. H., Kim, K. Y., & So, J. S. (2010). Screening and selection of stress resistant Lactobacillus spp. isolated from the marine oyster (Crassostrea gigas). *Anaerobe, 16*(5), 522-526. doi: 10.1016/j.anaerobe.2010.07.002

Lewis, R. J., Singh, O. M. P., Smith, C. V., Skarzynski, T., Maxwell, A., Wonacott, A. J., & Wigley, D. B. (1996). The nature of inhibition of DNA gyrase by the coumarins and the cyclothialidines revealed by X-ray crystallography. *Embo Journal, 15*(6), 1412-1420.

Liang, C. F., Yan, M. C., Chang, T. C., & Lin, C. C. (2009). Synthesis of S-linked alpha(2 -> 9) octasialic acid via exclusive alpha *S*-glycosidic bond formation. *Journal of the American Chemical Society, 131*(9), 3138-+. doi: 10.1021/ja808353m

Man, P., Novak, P., Hata, T., Main, P. L., Trevor S., Havlíček, V., Patchett, M. L., & Norris, G. E. (Unpublished). *Determining post-translational modifications on ASM1*.

Manske, M. (2007). GENtle (Version 1.9.4). Cologne.

Marchler-Bauer, A., Zheng, C., Chitsaz, F., Derbyshire, M. K., Geer, L. Y., Geer, R. C., . . . Bryant, S. H. (2013). CDD: conserved domains and protein three-dimensional structure. *Nucleic Acids Research, 41*(D1), D348-D352. doi: 10.1093/nar/gks1243

Matsusaki, H., Sonomoto, K., & Ishizaki, A. (1998). Characterization of lantibiotic nisin Z produced by *Lactococcus lactis* IO-1. *Journal of the Faculty of Agriculture Kyushu University, 43*(1-2), 181-193.

Mills, S., Serrano, L. M., Griffin, C. M., O'Connor, P. M., Schaad, G., Bruining, C., . . . Meijer, W. C. (2011). Inhibitory activity of *Lactobacillus plantarum* LMG P-26358 against *Listeria innocua* when used as an adjunct starter in the manufacture of cheese. *Microbial Cell Factories, 10*(Suppl. 1), 11pp.

Morgan, S., Ross, R. P., & Hill, C. (1995). Bacteriolytic activity caused by the presence of a novel lactococcal plasmoid encoding lactococcin-a, lactococcin-b, and lactococcin-m. *Applied and Environmental Microbiology, 61*(8), 2995-3001.

Nes, I. F., Diep, D. B., Havarstein, L. S., Brurberg, M. B., Eijsink, V., & Holo, H. (1996). Biosynthesis of bacteriocins in lactic acid bacteria. *Antonie Van Leeuwenhoek International Journal of General and Molecular Microbiology, 70*(2-4), 113-128. doi: 10.1007/bf00395929

Nes, I. F., Diep, D. B., & Holo, H. (2007). Bacteriocin diversity in *Streptococcus* and *Enterococcus*. *Journal of Bacteriology, 189*(4), 1189-1198. doi: 10.1128/jb.01254-06

Noé, L., & Kucherov, G. (2005). YASS: enhancing the sensitivity of DNA similarity search. *Nucleic Acids Research, 33*(suppl 2), W540-W543. doi: 10.1093/nar/gki478

Norris, G. E., & Patchett, M. (2014). *Glycosylated ribosomally-synthesised peptide toxins: discovery, characterisation and applications*. Natural products analysis: instrumentation, methods, and applications (Wiley).

Notredame, C., Higgins, D. G., & Heringa, J. (2000). T-Coffee: A novel method for fast and accurate multiple sequence alignment. *Journal of Molecular Biology, 302*(1), 205-217.

O'Shea, E. F., O'Connor, P. M., O'Sullivan, O., Cotter, P. D., Ross, R. P., & Hill, C. (2013). Bactofencin A, a new type of cationic bacteriocin with unusual immunity. *Mbio, 4*(6). doi: e00498-1310.1128/mBio.00498-13

O'Sullivan, D. J., & Klaenhammer, T. R. (1993). Rapid Mini-Prep isolation of high-quality plasmid DNA from *Lactococcus* and *Lactobacillus* Spp. *Applied and Environmental Microbiology, 59*(8), 2730-2733.

Oman, T. J., Boettcher, J. M., Wang, H. A., Okalibe, X. N., & van der Donk, W. A. (2011). Sublancin is not a lantibiotic but an S-linked glycopeptide. *Nature Chemical Biology, 7*(2), 78-80. doi: 10.1038/nchembio.509

Paik, S. H., Chakicherla, A., & Hansen, J. N. (1998). Identification and characterization of the structural and transporter genes for, and the chemical and biological properties of, sublancin 168, a novel lantibiotic produced by *Bacillus subtilis* 168. *Journal of Biological Chemistry, 273*(36), 23134-23142. doi: 10.1074/jbc.273.36.23134

Patchett, M. (2012a). [PCR Primers for the analysis of the *L. plantarum* A-1 genome].

Patchett, M. (2012b). [PCR Primers for the sequencing of *L. plantarum* A-1 sequence around ASM1].

Patchett, M. (2013a). [PCR Primers for the detection of pTRK669].

Patchett, M. L. (2013b). [Activation of the plantaricin cluster in *Lactobacillus plantarum* KW30].

Pfeiler, E. A., & Klaenhammer, T. R. (2007). The genomics of lactic acid bacteria. *Trends in Microbiology, 15*(12), 546-553. doi: http://dx.doi.org/10.1016/j.tim.2007.09.010

Pingitore, E. V., Hebert, E. M., Nader-Macias, M. E., & Sesma, F. (2009). Characterization of salivaricin CRL 1328, a two-peptide bacteriocin produced by *Lactobacillus salivarius* CRL 1328 isolated from the human vagina. *Research in Microbiology, 160*(6), 401-408. doi: 10.1016/j.resmic.2009.06.009

Punta, M., Coggill, P. C., Eberhardt, R. Y., Mistry, J., Tate, J., Boursnell, C., . . . Finn, R. D. (2012). The Pfam protein families database. *Nucleic Acids Research, 40*(D1), D290-D301. doi: 10.1093/nar/gkr1065

Rambaut, A. (2006). FigTree (Version 1.4.0). Edinburgh. Retrieved from http://tree.bio.ed.ac.uk

Russell, W. M., & Klaenhammer, T. R. (2001). Efficient system for directed integration into the *Lactobacillus acidophilus* and *Lactobacillus gasseri* chromosomes via homologous recombination. *Applied and Environmental Microbiology, 67*(9), 4361-4364. doi: 10.1128/aem.67.9.4361-4364.2001

Salvetti, E., Fondi, M., Fani, R., Torriani, S., & Felis, G. E. (2013). Evolution of lactic acid bacteria in the order Lactobacillales as depicted by analysis of glycolysis and pentose phosphate pathways. *Systematic and Applied Microbiology, 36*(5), 291-305. doi: 10.1016/j.syapm.2013.03.009

Sambrook, J., & Russell, D. (2001). *Molecular Cloning: A Laboratory Manual* (3rd ed. Vol. 1-3): Cold Spring Harbor Laboratory Print.

Saucier, L., Paradkar, A. S., Frost, L. S., Jensen, S. E., & Stiles, M. E. (1997). Transcriptional analysis and regulation of carnobacteriocin production in *Carnobacterium piscicola* LV17. *Gene, 188*(2), 271-277. doi: 10.1016/s0378-1119(96)00822-0

Sawa, N., Koga, S., Okamura, K., Ishibashi, N., Zendo, T., & Sonomoto, K. (2013). Identification and characterization of novel multiple bacteriocins produced by *Lactobacillus sakei* D98. *Journal of Applied Microbiology, 115*(1), 61-69. doi: 10.1111/jam.12226

Schagger, H., & Vonjagow, G. (1987). Tricine sodium dodecyl-sulfate polyacrylamide-gel electrophoresis for the separation of proteins in the range from 1-kDa to 100-kDa. *Analytical Biochemistry, 166*(2), 368-379. doi: 10.1016/0003-2697(87)90587-2

Schroth, M. N., Osgood, J. W., & Miller, T. D. (1973). Quantitative assessment of effect of olive knot disease on olive yield and quality. *Phytopathology, 63*(8), 1064-1065.

Sip, A., Wieckowicz, M., Olejnik-Schmidt, A., & Grajek, W. (2012). Anti-Listeria activity of lactic acid bacteria isolated from golka, a regional cheese produced in Poland. *Food Control, 26*(1), 117-124. doi: 10.1016/j.foodcont.2012.01.014

Sørvig, E., Mathiesen, G., Naterstad, K., Eijsink, V. G. H., & Axelsson, L. (2005). High-level, inducible gene expression in *Lactobacillus sakei* and *Lactobacillus plantarum* using versatile expression vectors. *Microbiology-Sgm, 151*, 2439-2449. doi: 10.1099/mic.0.28084-0

Steen, M. T., Chung, Y. J., & Hansen, J. N. (1991). Characterization of the nisin gene as part of a polycistronic operon in the chromosome of L*actococcus lactis* TCC-11454. *Applied and Environmental Microbiology, 57*(4), 1181-1188.

Stepper, J. (2009). *The molecular and cellular characterisation of the first glycocin: plantaricin KW30 : a thesis presented in partial fulfilment of the requirements for the degree of Doctor of Philosophy in biochemistry at Massey University, Palmerston North, New Zealand*: Massey University, Palmerston North.

Stepper, J., Shastri, S., Loo, T. S., Preston, J. C., Novak, P., Man, P., . . . Norris, G. E. (2011). Cysteine S-glycosylation, a new post-translational modification found in glycopeptide bacteriocins. *Febs Letters, 585*(4), 645-650. doi: 10.1016/j.febslet.2011.01.023

Stiles, M. E., & Hastings, J. W. (1991). Bacteriocin production by lactic acid bacteria: potential for use in meat preservation. *Trends in Food Science & Technology, 2*(10), 247-251. doi: 10.1016/0924-2244(91)90706-o

Todorov, S. D. (2009). Bacteriocins from *Lactobacillus plantarum* - production, genetic organization and mode of action. *Brazilian Journal of Microbiology, 40*(2), 209-221.

Trmčić, A., Monnet, C., Rogelj, I., & Matijasic, B. B. (2011). Expression of nisin genes in cheese-A quantitative real-time polymerase chain reaction approach. *Journal of Dairy Science, 94*(1), 77-85. doi: 10.3168/jds.2010-3677

Van Reenen, C. A., Van Zyl, W. H., & Dicks, L. M. T. (2006). Expression of the immunity protein of plantaricin 423, produced by *Lactobacillus plantarum* 423, and analysis of the plasmid encoding the bacteriocin. *Applied and Environmental Microbiology, 72*(12), 7644-7651. doi: 10.1128/aem.01428-06

Vandervossen, J., Vanherwijnen, M. H. M., Leer, R. J., Tenbrink, B., Pouwels, P. H., & Huisintveld, J. H. J. (1994). Production of Acidocin-B, a bacteriocin of *Lactobacillus acidophilus* m46 is a plasmid-encoded trait - plasmid curing, genetic marking by in-vivo plasmid integration, and gene-transfer. *Fems Microbiology Letters, 116*(3), 333-340. doi: 10.1016/0378-1097(94)90056-6

Venugopal, H., Edwards, P. J. B., Schwalbe, M., Claridge, J. K., Libich, D. S., Stepper, J., . . . Pascal, S. M. (2011). Structural, dynamic, and chemical characterization of a novel S-Glycosylated bacteriocin. *Biochemistry, 50*(14), 2748-2755. doi: 10.1021/bi200217u

Vossen, P. M., Elkins, R. B., Bianchi, M. L., Klonsky, K. M., Livingstone, P., & De Moura, R. L. (2011). *Sample Costs to Establish a Medium-Density Olive Orchard and Produce Bottled Olive Oil*. University of California.

Wang, H., Oman, T. J., Zhang, R., De Gonzalo, C. V. G., Zhang, Q., & van der Donk, W. A. (2014). The glycosyltransferase involved in thurandacin biosynthesis catalyzes both *O*- and *S*-glycosylation. *Journal of the American Chemical Society, 136*(1), 84-87. doi: 10.1021/ja411159k

Wang, H., & van der Donk, W. A. (2011). Substrate selectivity of the sublancin *S*-glycosyltransferase. *Journal of the American Chemical Society, 133*(41), 16394-16397. doi: 10.1021/ja2075168

Wiedemann, I., Breukink, E., van Kraaij, C., Kuipers, O. P., Bierbaum, G., de Kruijff, B., & Sahl, H. A. (2001). Specific binding of nisin to the peptidoglycan precursor lipid II combines pore formation and inhibition of cell wall biosynthesis for potent antibiotic activity. *Journal of Biological Chemistry, 276*(3), 1772-1779.

Yamashita, H., Tomita, H., Inoue, T., & Ike, Y. (2011). Genetic organization and mode of action of a novel bacteriocin, bacteriocin 51: Determinant of VanA-type vancomycin-resistant *Enterococcus faecium*. *Antimicrobial Agents and Chemotherapy, 55*(9), 4352-4360. doi: 10.1128/aac.01274-10

Yoneyama, F., Imura, Y., Ichimasa, S., Fujita, K., Zendo, T., Nakayama, J., . . . Sonomoto, K. (2009a). Lacticin Q, a lactococcal bacteriocin, causes high-level membrane permeability in the absence of specific receptors. *Applied and Environmental Microbiology, 75*(2), 538-541. doi: 10.1128/aem.01827-08

Yoneyama, F., Imura, Y., Ohno, K., Zendo, T., Nakayama, J., Matsuzaki, K., & Sonomoto, K. (2009b). Peptide-lipid huge toroidal pore, a new antimicrobial mechanism mediated by a Lactococcal bacteriocin, Lacticin Q. *Antimicrobial Agents and Chemotherapy, 53*(8), 3211-3217. doi: 10.1128/aac.00209-09

Zendo, T., Yoneyama, F., & Sonomoto, K. (2010). Lactococcal membrane-permeabilizing antimicrobial peptides. *Applied Microbiology and Biotechnology, 88*(1), 1-9. doi: 10.1007/s00253-010-2764-3

Zhu, X. M., Haag, T., & Schmidt, R. R. (2004). Synthesis of an *S*-linked glycopeptide analog derived from human Tamm-Horsfall glycoprotein. *Organic & Biomolecular Chemistry, 2*(1), 31-33. doi: 10.1039/b311583f

Zouhir, A., Hammami, R., Fliss, I., & Ben Hamida, J. (2010). A new structure-based classification of gram-positive bacteriocins. *Protein Journal, 29*(6), 432-439. doi: 10.1007/s10930-010-9270-4