Understanding the mechanism of action of the glycosylated bacteriocin glycocin F

A thesis presented in partial fulfilment of the requirements for the degree of

Doctor of Philosophy
in
Biochemistry

at Massey University, Manawatu, New Zealand.

Sean William Bisset

2019

Abstract

With the increasing threat posed by antibiotic-resistant bacteria, efforts must be made to find new antimicrobial agents. One growing area of promise is the bacteriocins, which are a diverse group of antimicrobial peptides produced by bacteria. This thesis focuses on determining the mechanism of action of one of these peptides, glycocin F (GccF). GccF is produced by the bacterium *Lactobacillus plantarum* and is modified with two N-acetyl glucosamine (GlcNAc) sugar moieties, one located on an interhelical loop region, and the other at the end of a flexible C-terminal 'tail'. It has also been shown to exhibit a unique effect on susceptible bacteria, putting them into a reversible state of hybernation as opposed to outright killing them. However, little is known about the roles of the structural features of GccF, how it triggers bacteriostasis in target cells, or even what part(s) of bacterial cells it targets. This work addresses these questions using three main approaches: studying the structure-function relationship of different parts of GccF with chemically synthesised analogues; looking at the transcriptional response of a pathogenic bacteria, *Enterococcus faecalis*, to GccF; and trying to identify binding partners of GccF and its respective immunity protein, GccH. The results presented here highlight different roles of the GlcNAcs attached to GccF, with both the interhelical loop and presence of GlcNAc on this loop being vital for activity, while the sugar at the C-terminal position is important, but not crucial for the peptide's activity. Additionally, a role of the GlcNAc phosphotransferase system on the mechanism of GccF is strongly indicated, with evidence from both the transcriptional studies and the protein interaction studies of GccF's immunity protein. Taken together, the results allow for two theoretical models of GccF's mechanism of action to be proposed. These models presented here should serve to increase the understanding of other glycocin-class bacteriocins and their mechanisms of action, and possibly contribute towards the creation of a blueprint for a new class of antimicrobial agents.

Acknowledgements

A PhD is a lot like visiting the dentist: it takes longer than you'd expect, is often not painless, and puts a dent in your finances. And you generally don't seek out either unless there's something wrong with you.

With that being said, I would first like to thank my supervisors, Gill Norris and Mark Patchett, for all of their encouragement, support, and ideas throughout the years of this work. The enthusiasm they bring to their research is infectious, and it has been a an exciting journey working with them. I would also like to thank Dave Wheeler for his ideas and support in an area that I felt I was stepping into completely blind, and Jasna Rakonjac, for encouraging me to take this opportunity. And many thanks to Trevor Loo, for all of the help, advice, expertise and patience with poorly-worded questions.

I would also like to thank the administration staff of the Institute/ School of Fundamental Sciences, especially Ann, Cynthia and Debbie, who keep everything running so smoothly and who are always happy to provide help with even the most inane requests.

Thank you to Professors Gregory Cook and Lyn Hancock for kindly providing bacterial strains for testing, and of course to the Maurice Wilkins Centre for supporting and financing myself and this work.

As with any endeavour, a good support network makes all the difference. A big shout out to the Lab Ladies and other members of X-Lab, past and present, who made the lab such a friendly, fun and supportive work environment, Travis and Chelsea for all your kindness and fun over the years, and all the other people I've met over this time. And thank you to mum, whose support, love, and encouragement is second to none.

Last, but certainly not least, Shannon: thank you for all your love and support over the years. Thank you for sharing this journey with me, you made the world of difference!

Abbreviations

General abbreviations

aa-tRNA	Aminoacyl-tRNA
ABC (*chemical*)	Ammonium bicarbonate
ABC (*transporter*)	ATP-binding cassette
APS	Ammonium persulfate
ATP	Adenosine tri-phosphate
BAGEL	Bacteriocin genome mining tool
BLAST	Basic Local Alignment Search Tool
CD	Circular dichroism
CDM	Chemically-defined media
CFU	Colony-forming units
COG	Cluster of gene ontologies
DEG	Differentially expressed gene
DEPC	Diethyl pyrocarbonate
DHPS	dihydropteroate synthase
DNA	Deoxyribonucleic acid
DTT	Dithiothreitol
EBP	Enhancer binding protein
EDTA	Ethylenediaminetetraacetic acid
fL	Flag-linker
GccF	Glycocin F
GlcNAc	*N*-acetyl-D-glucosamine
GlcNDAz	Diazirine-labelled GlcNAc
GO	Gene ontology
HPLC	High pressure liquid chromatography
HRP	Horse radish peroxidase
IC_{50}	Inhibitory concentration 50%
IPP	Isoprenyl pyrophosphate

Continued on next page

General abbreviations (*continued*)

IPTG	Isopropyl β-D-1-thiogalactopyranoside
KEGG	Kyoto Encyclopedia of Genes and Genomes
LAB	Lactic acid bacteria
MIC	Minimum inhibitory concentration
MurNAc	*N*-acetylmuramic acid
NCL	Native chemical ligation
NMR	Nuclear magnetic resonance
NZGL	New Zealand Genomics Limited
OD_{600}	Optical density at 600 nm
ORF	Open reading frame
PAGE	Polyacrylamide gel electrophoresis
PBP	Penicillin-binding protein
PCA	Principal component analysis
PCR	Polymerase chain reaction
PEP	Phosphoenol pyruvate
PRD	PTS regulation domain
PTS	Phosphotransferase system
PVDF	Polyvinylidine difluoride
RNA	Ribonucleic acid
RNA-seq	RNA sequencing
RODEO	Rapid ORF description and evaluation online
RT	Room temperature
SDS	Sodium dodecyl sulphate
SNP	Single-nucleotide polymorphism
TEMED	Tetramethylethylenediamine
TFA	Trifluoro acetic acid
UV	Ultraviolet

Amino acid abbreviations

Full name	Three-letter code	Single-letter code
Alanine	Ala	A
Arginine	Arg	R
Asparagine	Asn	N
Aspartic acid	Asp	D
Cysteine	Cys	C
Glutamine	Gln	Q
Glutamic acid	Glu	E
Glycine	Gly	G
Histidine	His	H
Isoleucine	Ile	I
Leucine	Leu	L
Lysine	Lys	K
Methionine	Met	M
Phenylalanine	Phe	F
Proline	Pro	P
Serine	Ser	S
Threonine	Thr	T
Tryptophan	Trp	W
Tyrosine	Tyr	Y
Valine	Val	V

Contents

List of Figures

List of Tables

1 | Introduction

Preface

Bacteria were surviving, multiplying, and killing each other long before multicellular life emerged. Over the millennia they have evolved a diverse array of combat strategies, including antimicrobial weaponry, and as these weapons have evolved, so too have their resistance mechanisms. By the time eukaryotes entered the scene, bacteria had already adapted their weapons to exploit and predate them too. Until the early 20th century, humankind simply accepted these microscopic pathogens as a force of nature, and expected infection and illness to be inevitable and often fatal[10,177]. But the discovery of antibiotics changed this balance, and humans were no longer at the mercy of microbes - injuries no longer meant the loss of a limb, mortality rates of surgeries steeply declined allowing more complex procedures to be carried out, and the average human lifespan increased. However, the 21st century has ushered in a new era in which the tables are once again turning, with the emergence of multi drug resistant strains of pathogens that threaten to return us to the pre-antibiotic era[78,114].

There is no single solution to overcoming antibiotic resistance. It is highly unlikely that a single drug will be created that will selectively kill all pathogens, while preventing the development of resistance. Rather, the most effective method for combating bacteria will be to continually discover new molecular targets in the microbial genome or proteome what can be exploited by antimicrobial agents, and to never rely on one single compound for extensive periods of time. In particular, a comprehensive investigation into mechanisms that bacteria use to inhibit each other has the potential to provide us with the broadest range of new tools. This thesis will focus on just one of these targets, which falls into an expanding category of presently under-utilised molecules: the bacteriocins[50]. It will focus on attempts made to characterise the mechanism of action of an unusual bacteriocin, glycocin F. To place the work in context, chapter 1 will begin with an overview of the history of the current state of antibiotics, followed by descriptions of their general mechanisms of action, as well as some associated risks with their use. This will be followed by an introduction to a class of antimicrobial compounds collectively known as the bacteriocins, including what is currently understood about their mechanisms of action, and their physiological roles. It will then focus on a particular class of bacteriocins (**The RiPPs**), specifically the glycocins. Lastly, what is known about one member of this sub class, glycocin F, will be described in terms of the aims of this study.

1.1 Antibiotics - a history and summary

1.1.1 The discovery of antibiotics

The early days, pre-1950's

The discovery of penicillin in 1929 by Alexander Fleming, where he witnessed that growth of mould on an agar plate was inhibiting the growth of bacteria, is usually considered the beginning of the 'Antibiotic Era'[68]. This mould was identified as *Penicillium notatum*, and the antibacterial compound it produced was later characterised[177]. However, Fleming was not the first person to notice the disease-curing effects of certain compounds. Work by Emmerich and Low in 1899 showed that extracts from *Pseudomonas aeruginosa* had inhibitory activity against a range of pathogenic bacteria, although not without detrimental effects to the treated patients[66,10]. Meanwhile, in 1910, Paul Ehrlich and others pioneered a screening program in an effort to find compounds active against syphilis (caused by the bacteria *Treponema pallidium*) that resulted in the discovery of the arsenic compound Arsphenamine (marketed as Salvarsan)[177]. These investigations were all part of the search for compounds with which to combat and control pathogenic bacteria, and led to what has since been referred to as the 'Golden Age' of antibiotic discovery.

The 'Golden Age' of antibiotic discovery, 1950-1970

The 'Golden Age' of antibiotic discovery is said to have occurred between 1950 and 1970, when almost all the classes of antibiotics currently in use were discovered[10,104]. A major factor in the discovery of these molecules was the screening method developed by Selman Waksman, whereby soil-derived microbes were screened against 'test' bacteria using agar overlay plates[104]. The first antibiotic discovered using this method was streptomycin in 1943, leading to the clinical introduction of the aminoglycoside class of antibiotics in 1946[138,104]. Waksman was subsequently awarded the Nobel prize in Physiology or Medicine in 1952. A number of other classes of antibiotics were discovered before the end of the 1960's (Table 1.1), with the final discovery of this period being that of fosfomycin in 1969[67,85]. The success of this method of screening soil bacteria was short-lived, however, and once it became apparent that pre-existing antibiotics were being rediscovered the search was abandoned[104].

Table 1.1: Discovery of the main classes of antibiotics during the 'Golden Age' of discovery

Class	**Antibiotic**	**Year discovered**	**Producing organism**	**Reference**
β-lactams[†]	Penicillin	1928	*Penicillium notatum*	Fleming 1928[68]
Polypeptides	Tyrocidine	1939	*Bacillus brevis*	Roskoski *et al.* 1970[134]
Aminoglycosides	Streptomycin	1943	*Streptomyces griseus*	Schatz *et al.* 1944[138]
Amphenicols	Chloramphenicol	1947	*Streptomyces venezuelae*	Ehrlich *et al.* 1947[66]
Lipopeptides	Colistin	1947	*Paenibacillus polymyxa*	Storm *et al.* 1977[147]
Tetracyclines	Tetracycline	1948	*Streptomyces aureofaciens*	Duggar 1948[65]
Macrolides	Erythromycin	1949	*Saccharopolyspora erythraea*	Garrod 1957[75]
Oxazolidinones	Cycloserine	1952	*Streptomyces* strain K-300	Aminov 2017[9]
Streptogramins	Streptogramin	1953	*Streptomyces* spp.	Charney *et al.* 1953[37]
Glycopeptides	Vancomycin	1953	*Streptomyces orientalis*	McGuire *et al.* 1955[112]
Ansamycins	Rifamycin	1957	*Amycolatopsis rifamycinica*	Margalith and Beretta 1960[108]
(Fluoro)quinolones	Nalidixic acid	1961	Synthesised *in vitro*	Deitz *et al.* 1963[55]
Lincosamides	Lincomycin	1963	*Streptomyces lincolnensis*	MacLeod *et al.* 1964[106]
Fosfomycins	Fosfomycin	1969	*Streptomyces fradiae*	Hendlin *et al.* 1969[85]

[†]Cephalosporins and carbapenems fall under this class.

The end of soil-screening did not spell the end of antibiotic development, however, as the focus shifted towards modifying the existing classes of antibiotics[44]. It began in 1970[177] and was primarily initiated by the observation of growing antibiotic resistance in hospitals. For example, the development of methicillin and oxacillin (narrow-spectrum antibiotics targeting *Staphylococcus aureus*) was spurred by the need for antibiotics with resistance to β-lactamase, an enzyme that can hydrolyse the β-lactam ring of penicillin and β-lactam antibiotics, rendering them ineffective[177]. Additionally, a second class of antibiotics that has been extensively developed is the cephalosporins, which have been modified continuously since being made available to the public in 1962, with the fourth generation of molecules being released in 1983[177]. However, aside from the modifications of pre-existing molecules, no new classes of antibiotics were discovered in the years between fosfomycin in 1969 and linezolid a member of the oxazolinedone class in 2000[44].

1.1.2 Antibiotic mechanisms of action

An overview of the current classes of antibiotics, along with their founding members, is shown in Table 1.1. These classes are structurally distinct and function by inhibiting one of five main cellular processes: cell wall synthesis, DNA-templated RNA synthesis (transcription), DNA replication, protein translation, and folate synthesis (Figure 1.1)[44,41]. The first four targets are the most commonly exploited by clinically available antibiotics[104].

1.1.2.1 Cell wall inhibition

Synthesis of peptidoglycan and cell wall assembly occurs in three main stages: (i) cytoplasmic synthesis of sugar units and attachment to pentapeptides, (ii) transfer of the peptidoglycan building blocks across the cell membrane, and (iii) polymerisation of the glycans and attachment to the pre-existing cell wall structure[41,27]. Key steps in the cytoplasm include the synthesis of UDP-*N*-acetylmuramic acid (UDP-MurNAc), and the subsequent sequential attachment of five amino acids onto the sugar unit. Following this, the UDP-MurNAc-L-Ala-D-Glu-*meso*-A_2pm-D-Ala-D-Ala building block is attached to undecaprenyl (C_{55} isoprenyl) phosphate, a lipid-bound carrier. *N*-acetyl-D-glucosamine (GlcNAc) is then attached to the MurNAc-pentapeptide building block, creating a GlcNAc-MurNAc-pentapeptide,

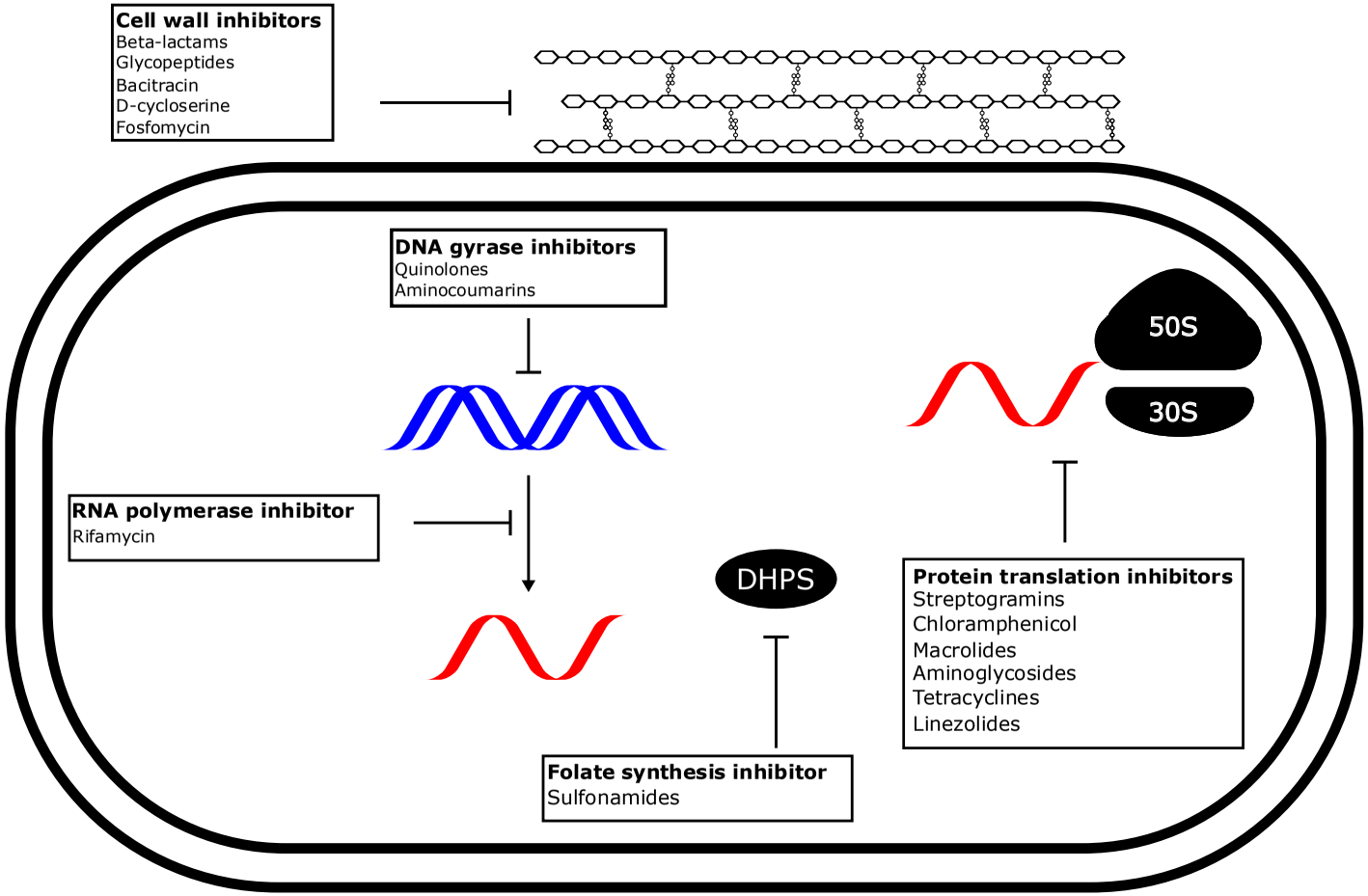


Figure 1.1: **Common antibiotic mechanisms of action**
The five main targets of antibiotics are highlighted: Cell wall synthesis, protein translation, DNA replication/ DNA gyrase, RNA transcription, and folate biosynthesis. Double-stranded DNA is indicated in blue, and single-stranded RNA is indicated in red. DHPS = dihydropteroate synthase.

and the structure is then flipped through the membrane to sit on the extracellular face[27]. The final steps involve the attachment of the peptidoglycan building block to the external pre-existing structure, firstly by transglycosylation of GlcNAc-MurNAc to a glycan chain. Finally a transpeptidase catalyses additional crosslinking of the pentapeptide group to other glycan chains (summarised in Figure 1.2)[41].

Five different classes of antibiotics target cell wall synthesis: two target specific steps in the cytoplasm (fosfomycin and oxazolinidones), one targets a step that occurs in the membrane (polypeptides), and two target the third stage of synthesis (β-lactams and glycopeptides) (Table 1.1; Figure 1.2). Fosfomycin is a phosphoenol pyruvate (PEP) analogue that inhibits the activity of MurA, the enzyme that catalyses the addition of PEP to UDP-GlcNAc in the cytoplasm, by binding to its active site cysteine[19]. D-cycloserine is a member of the oxazolinidone class of antibiotics and is a structural analogue of D-alanine. It disrupts the first stage of peptidoglycan synthesis by binding to and inhibiting D-Ala-D-Ala ligase[128]. Bacitracin is a member of the polypeptide class of antibiotics and binds to undecaprenyl pyrophosphate, preventing dephosphorylation by pyrophosphatase, thus decreasing the availability of free undecaprenyl phosphate necessary to carry sugars across the membrane[155]. Glycopeptides function by binding to the D-Ala-D-Ala dipeptide of the extracellular-facing pentapeptide, preventing crosslinking[172,61,18], while β-lactams function by binding to penicillin-binding proteins (PBPs) that facilitate transpeptidation[61].

Resistance to fosfomycin occurs through mutation of the active site of MurA (generally substitution of the active site Cys with Asp), mutations in the importers exploited by fosfomycin (such as the glycerol and carbohydrate transporters GlpT and UhpT), and inactivation of the antibiotic itself by glutathione-S-transferases[67]. Resistance to D-cycloserine in *Mycobacterium sp.* typically occurs through mutations in the *alr* gene (which encodes D-alanine racemase), although recent evidence also suggests that mutations in *ptsI* (which in encodes a furanose-specific transporter) also confers resistance, possibly through being the primary entry point for D-cycloserine[39]. Bacitracin resistance occurs through either over-expression of undecaprenyl phosphate, or over-expression of the ABC transporters believed to be involved in its export[155]. β-lactam resistance occurs most commonly through two primary mechanisms: expression of enzymes which hydrolyse the central β-lactam ring[79], and mutation of the PBPs to prevent antibiotic binding[35]. Developments to overcome the former mechanism resulted in the development of β-lactamase inhibitors, such as clavulanic acid and tazobactam, which are used in conjunction

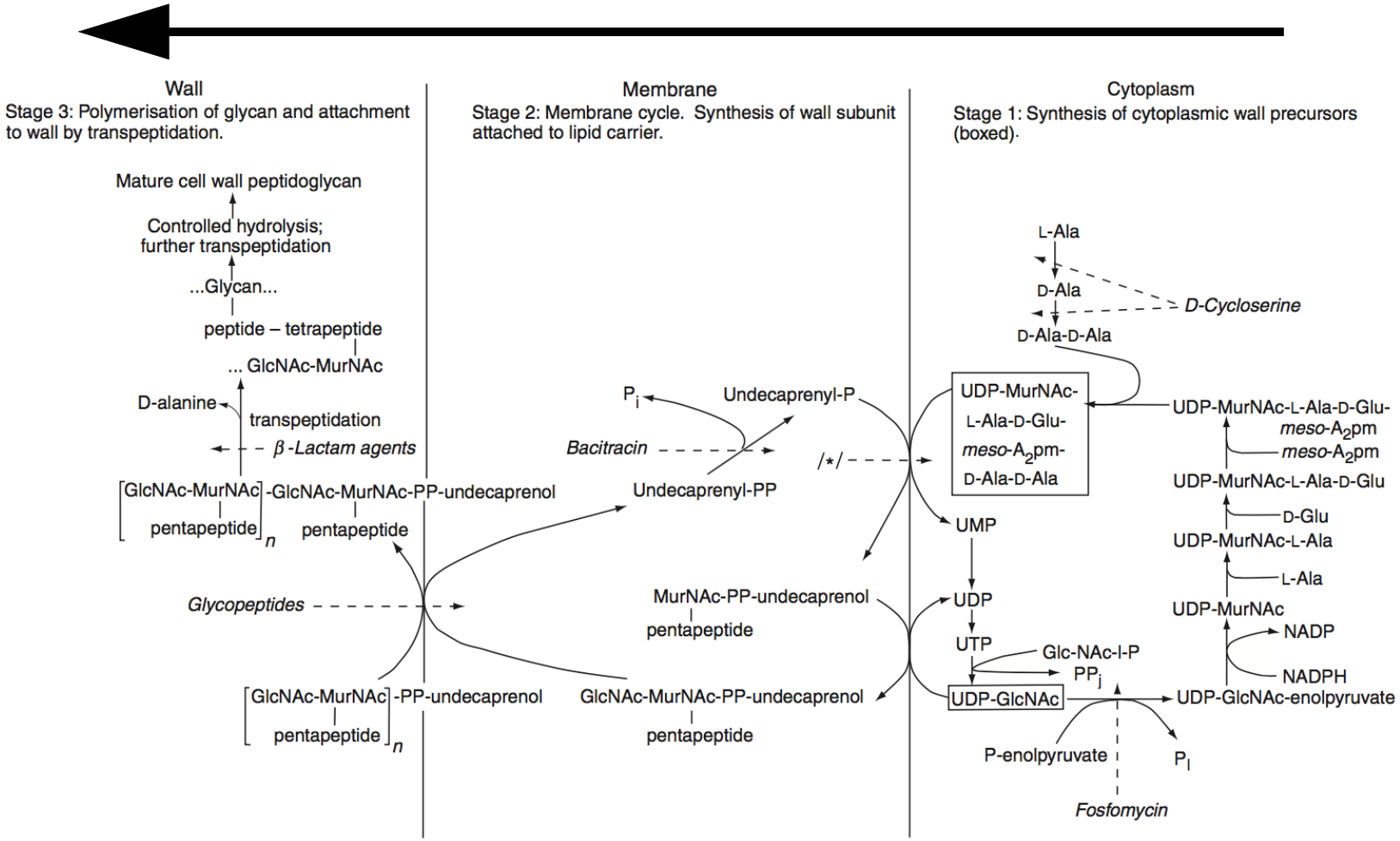


Figure 1.2: **Diagram of peptidoglycan synthesis stages**
Diagram of peptidoglycan synthesis, also showing the antibiotic classes that affect cell wall synthesis and the steps they act on. Reprinted by permission from Springer Nature, *European Journal of Clinical Microbiology and Infectious Diseases*, Structure, biochemistry and mechanism of action of glycopeptide antibiotics, P. E Reynolds, copyright© 1989.

with broad-spectrum penicillins[79,177]. Finally, resistance to glycopeptides such as vancomycin predominately occurs through either modification of the cross-linking dipeptide, replacing D-Ala-D-Ala with variants such as D-Ala-D-Lac or D-Ala-D-Ser, or *via* mutation of the ligase which facilitates this crosslinking[35].

1.1.2.2 Protein translation inhibition

The ribosome is a central component of all forms of life, and is highly conserved within each of the three Kingdoms (Bacteria, Archaea, and Eukarya), and is a target for specifically combating bacteria. The biological process controlled by ribosomes, translation of proteins from an mRNA template, is an essential and highly regulated activity that ultimately provides every structural, storage, and enzymatic protein to the cell. Unsurprisingly, the inhibition of protein translation is the most common process targeted by antibiotics, with numerous different classes inhibiting this process through a variety of different mechanisms. These have been extensively reviewed by Wilson (2009)[165], and will only be briefly covered here.

The bacterial ribosome is comprised of two subunits (30S and 50S) which are comprised of near-equal amounts of ribosomal RNA (rRNA) and protein. The process of translation also requires the use of other proteins which help assemble the ribosomal complex, deliver charged aminoacyl-tRNAs, facilitate peptidyl transfer and translocation, and release the mature peptide. These include proteins such as initiation factors (IFs), elongation factors (*e.g.* EF-Tu and EF-G), and release factors (RFs), to name but a few[165]. The process of protein translation is generally divided into four main phases: initiation, where the ribosome 50S and 30S subunits are assembled into a complex along with the mRNA and the initiating aminoacyl-tRNA (aa-tRNA), which is facilitated by initiation factors IF1, IF2, and IF3; elongation, where aa-tRNAs are transported to the ribosome by the elongation factor EF-Tu, the correct aa-tRNA is selected, and peptidyl transfer is carried out to add the amino acid to the growing chain; termination, where the peptide chain is released from the ribosome following recognition of a stop codon by the release factors RF1 or RF2; and recycling, where the ribosome·mRNA·deacylated tRNA complex is dissociated through the use of EF-G and ribosome recycling factor (RRF) to free the components for synthesis of a new peptide[103,130]. Almost each step of protein translation is inhibited by an antibiotic (see Figure 1.3)[165].

The initiation step of protein synthesis is inhibited by two main antibiotics, kasug-

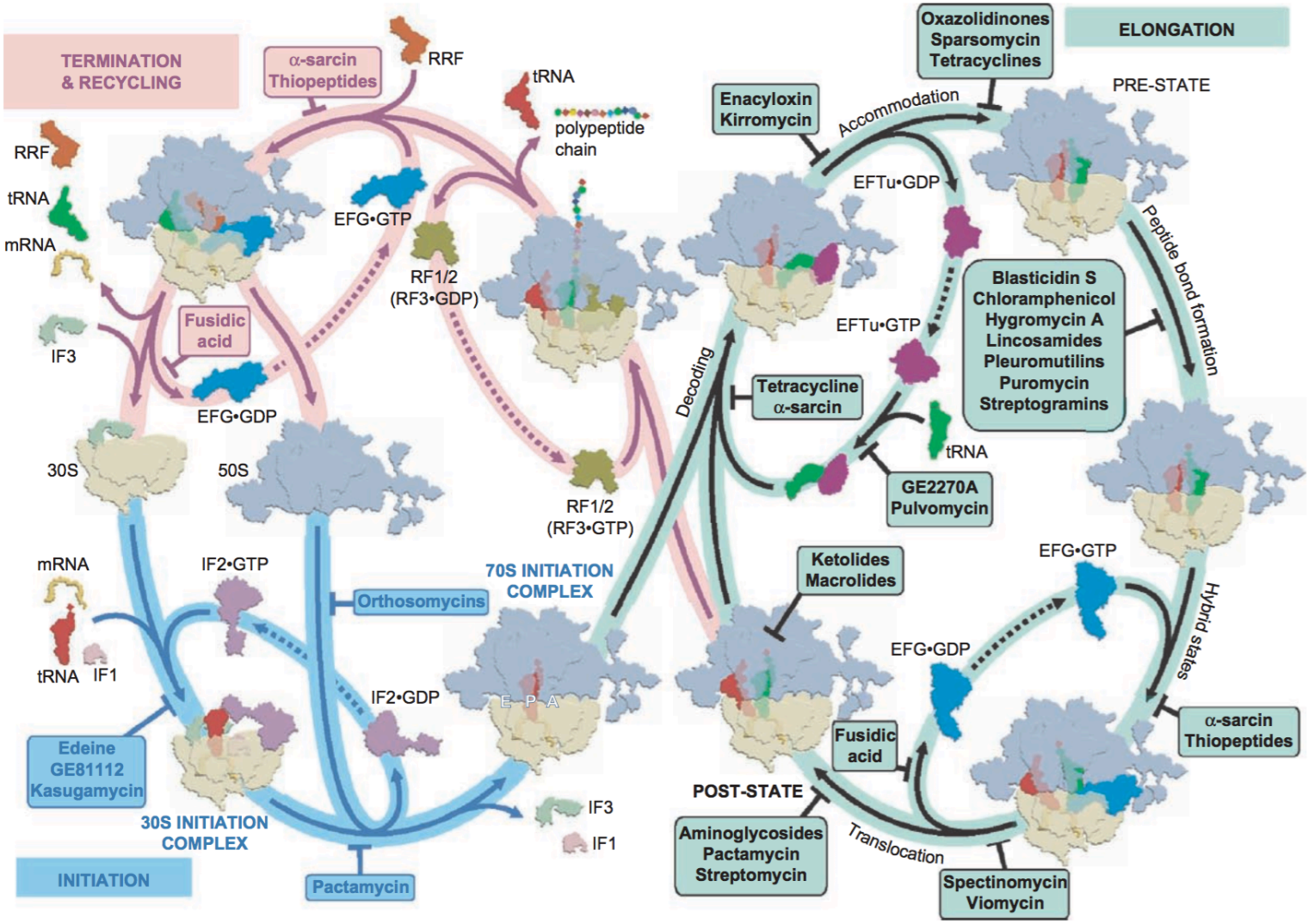


Figure 1.3: **Overview of stages of protein translation targeted by antibiotics**
The different stages of protein translation targeted by antibiotics are shown. The three main phases of translation are highlighted, with Initiation in blue, Elongation in green, and Termination and Recycling in pink. Figure reprinted from Wilson (2009) [165], with permission from Taylor & Francis.

amycin (largely used in combating fungal growth in plants) and edeine (a pentapeptide belonging to the amide class)[165]. Kasugamycin binds to two locations on the ribosome 30S subunit, the E-site, and a location that overlaps with the P-site. It is believed to function by preventing binding of the fMet-tRNA$_{fMet}$ to the nascent ribosome complex[139]. Edeine, on the other hand, binds to a single site on the 30S subunit, and functions by inhibiting the binding of aa-tRNAs to the 30S and 70S subunits[165].

Steps involved in the elongation stage of protein synthesis are most commonly targeted by antibiotics. Elongation is roughly divided into four steps: decoding, accommodation, peptide bond formation, and translocation[165]. Classes of antibiotics that inhibit elongation include the oxazolidines, tetracyclines, chloramphenicol, lincosamides, aminoglycosides, macrolides, and thiopeptides, although many other classes also target this process[165]. These function by a variety of processes, including blocking the sites of aminoacyl-tRNA binding (*e.g.* tetracyclines[31]), preventing peptidyl-transferase activity (*e.g.* chloramphenicol[140]), blocking extension of the growing peptide chain by blocking the ribosome exit pore (*e.g.* macrolides[77]), and increasing the rate of codon:anticodon mismatches interfering with the 'proofreading' nucleotides (*e.g.* aminoglycosides[120]). Mechanisms of resistance generally involve either mutation of the antibiotic binding sites (such as on the 30S and 50S ribosomal subunits), expression of enzymes which methylate or acetylate the antibiotic in question, or over-expression of efflux pumps to prevent the intracellular accumulation of the antibiotics[53,165].

1.1.2.3 DNA synthesis inhibition

A group of enzymes that are crucial to DNA replication and upkeep in the prokaryotic cells are the topoisomerases, which modify the supercoiling status of bacterial DNA. In particular, topoisomerase II (also called DNA gyrase) is a hetero-tetrameric enzyme composed of two GyrA subunits (responsible for DNA binding and cleavage) and two GyrB subunits (responsible for ATP hydrolysis), and functions to introduce negative supercoiling through double-stranded DNA breaks[28,84,36]. Topoisomerase II is the target of two classes of antibiotics: the quinolones (and modified variants) and the aminocoumarins[28,41]. Quinolones inhibit DNA gyrase activity by binding in a pocket formed between GyrA and the bound DNA, and stabilising the DNA-protein complex, stalling the enzyme and introducing double-stranded DNA breaks

that do not get repaired[41,84,97]. Aminocoumarins, on the other hand, function as competitive inhibitors of the ATPase activity of the GyrB subunit[154].

Resistance to quinolones most commonly arises through mutations in a conserved region of the GyrA subunit near the active site, known as the quinolone resistance-determining region (QRDR)[175]. Recently, however, mutations have also been found in *gyrB* which also facilitate resistance to quinolones. Analysis of the position of the GyrB mutations show that the affected residues sit close to the QRDR of GyrA, and presumably disrupt the binding of quinolone to the binding pocket[84,170,169]. Resistance to aminocoumarins is generally conserved, and has only been found as mutations on two residues of the GyrB subunit, as loss of function of DNA gyrase is generally lethal to cells[45].

1.1.2.4 RNA transcription inhibition

RNA transcription is carried out by RNA polymerase, a highly conserved enzyme which is composed of five different polypeptide chains: two α subunits, a β and β' subunit, and an ω subunit[41,116]. To date, only one class of antibiotics, the rifamycins, targets the bacterial RNA polymerase[41]. Rifamycin inhibits cells by binding to the DNA/RNA binding pocket of the β subunit of the RNA polymerase complex, which is involved in DNA binding and RNA synthesis. Crucially, rifamycin does not bind to the catalytic site of RNA polymerase, but rather binds in such a position that extension of the RNA chain is prevented from progressing past 2-3 nucleotides in length[33]. Whether this is all that the antibiotic is doing is a source of ongoing debate[97]. Unsurprisingly, the primary causes of resistance to the rifamycins are mutations to the *rpoB* gene which encodes the β subunit[41].

1.1.2.5 Folate synthesis inhibition

The synthesis of folate is a process that is largely restricted to bacteria, with mammalian cells being unable to synthesise folate, but rather have folate-uptake mechanisms[143]. As such, the inhibition of folate synthesis provides an exclusive target posing no danger to mammalian cells. However, only one class of antibiotics, the sulfonamides, targets this process[41,143]. They function by acting as a structural analogue of *para*-aminobenzoic acid, which is a key component of the folate synthesis pathway, and bind irreversibly to the dihydropteroate synthase (DHPS), which

catalyses the synthesis of dihydropteroic acid[41]. Resistance to sulfonamides occurs fairly rapidly, leading to the minimal use of this class of antibiotics in recent years. Resistance typically occurs through mutations in the *dhps*/*folP* (DHPS-encoding) gene[143].

1.1.3 Concerns with antibiotics

One rising issue that has been associated with the use of antibiotics, particularly broad-spectrum antibiotics, is the unintended damage of the commensal microbiome[101]. Overuse of antibiotics can cause a change in the balance and diversity of the taxa making up a person's microbiome, known as dysbiosis, and can increase the relative abundance of pathogenic bacteria[101,23]. This collateral damage can lead to health issues such as digestion problems, food allergies and increased prevalence of autoimmune conditions such as coeliac disease[109,101]. The gut microbiome is a particularly problematic target for this occurrence, where loss of bacterial diversity can lead to chronic *Clostridium difficile* infections, and in severe cases can be fatal if left untended[23]. While moving to completely different treatment regimes (such as phage therapy[101]) may be one way to minimise collateral damage of commensal microbiota, another option could be the discovery and development of a wider range of narrow-spectrum antimicrobials, to decrease the likelihood of off-target effects. Another option may be the development of bacteriostatic agents that slow the growth of pathogens enough to give the host immune system a chance to catch up.

1.2 Bacteriocins

Bacteriocins are ribosomally synthesised peptides produced by bacteria, which may be post-translationally modified, that convey an antimicrobial effect on closely-related strains or species of bacteria[34]. Those that are post-translationally modified belong to the group of molecules known as RiPPs (ribosomally synthesised and post-translationally modified peptides)[13]. Numerous bacteriocins have been characterised to date, and new members are continually being discovered and characterised. Due in part to this abundance, bacteriocins have become an increasingly popular source of compounds for antimicrobial research[34,50], and have been utilised in food production and preservation for many years[42]. Mechanisms of action of bacteriocins loosely

overlap with those of antibiotics (Figure 1.4), although bacteriocins do not share the same molecular targets as pre-existing antibiotics. In fact, a common mechanism of action of bacteriocins from Gram-positive bacteria (particularly those from lactic acid bacteria (LAB)) is to form lethal pores in the cell membrane (Table 1.2) [82].

Bacteriocin genes are typically harboured on plasmids, and are organised in operons containing the structural gene(s), specific transporters (typically ABC-type), immunity genes and genes encoding any other proteins necessary for their maturation [153]. Most bacteriocins are synthesised with an N-terminal leader peptide, which is cleaved during export from the cell. The leader peptide serves to hold the bacteriocin in an inactive state prior to release, and is often necessary for recognition by post-translational modification enzymes [178]. Producing strains are further protected from exported bacteriocin molecules through expression of a specific immunity protein [59]. The mechanism of most immunity proteins are still unknown, although the immunity proteins have been characterised for a few bacteriocins, [6,8,59].

1.2.1 The classes of bacteriocins

The current classification scheme for bacteriocins divides them into three main classes, as shown in Figure 1.5. **Class I** bacteriocins are small ($<$ 10 kDa) peptides, which are post-translationally modified, and **Class II** bacteriocins are small, non-modified peptides [8]. In addition to Figure 1.5, Class I has been recently extended to include the bottromycins and thiopeptides as members [4]. **Class III** bacteriocins are larger than 10 kDa in size, and have also been recently extended to include a third group, the tailocins [4].

Class I bacteriocins

Class I bacteriocins comprise small bacteriocins (less than 10 kDa) which contain post-translational modifications in addition to disulfide bridges (Figure 1.5). The first group of peptides belonging to this class that were characterised was the lanthipeptides (or lantibiotics), with the most extensively characterised member being nisin. Lanthipeptides are characterised by the presence of 'Lan' or 'Melan' rings, which are formed between a Cys residue enzymatically connected to a dehydrated form of Ser or Thr (dehydroalanine [Dha] or 2,3-dehydrobutyrine [Dhb], respectively) and which increases structural stability and resistance to proteases [96]. The

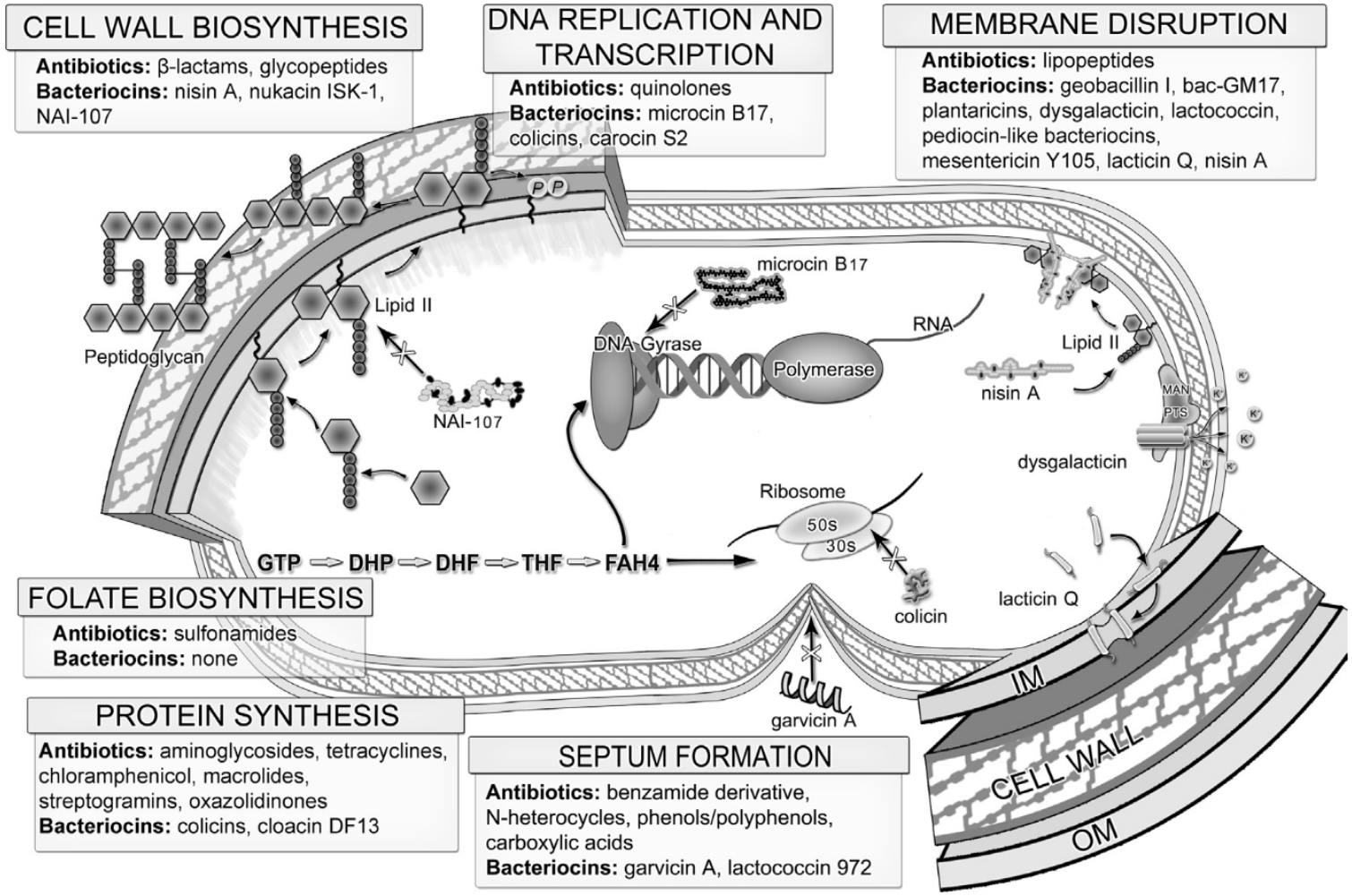


Figure 1.4: **Common mechanisms of antibiotics and bacteriocins**
The main targets of conventional antibiotics are shown, along with examples of bacteriocins that target similar systems. While there is overlap in all processes except folate biosynthesis, bacteriocins largely cause membrane disruption or disrupt cell wall biosynthesis. A large number disrupt some steps in the protein synthetic pathways. Taken from Cavera *et al.* 2014[34]. Reprinted with permission from Elsevier.

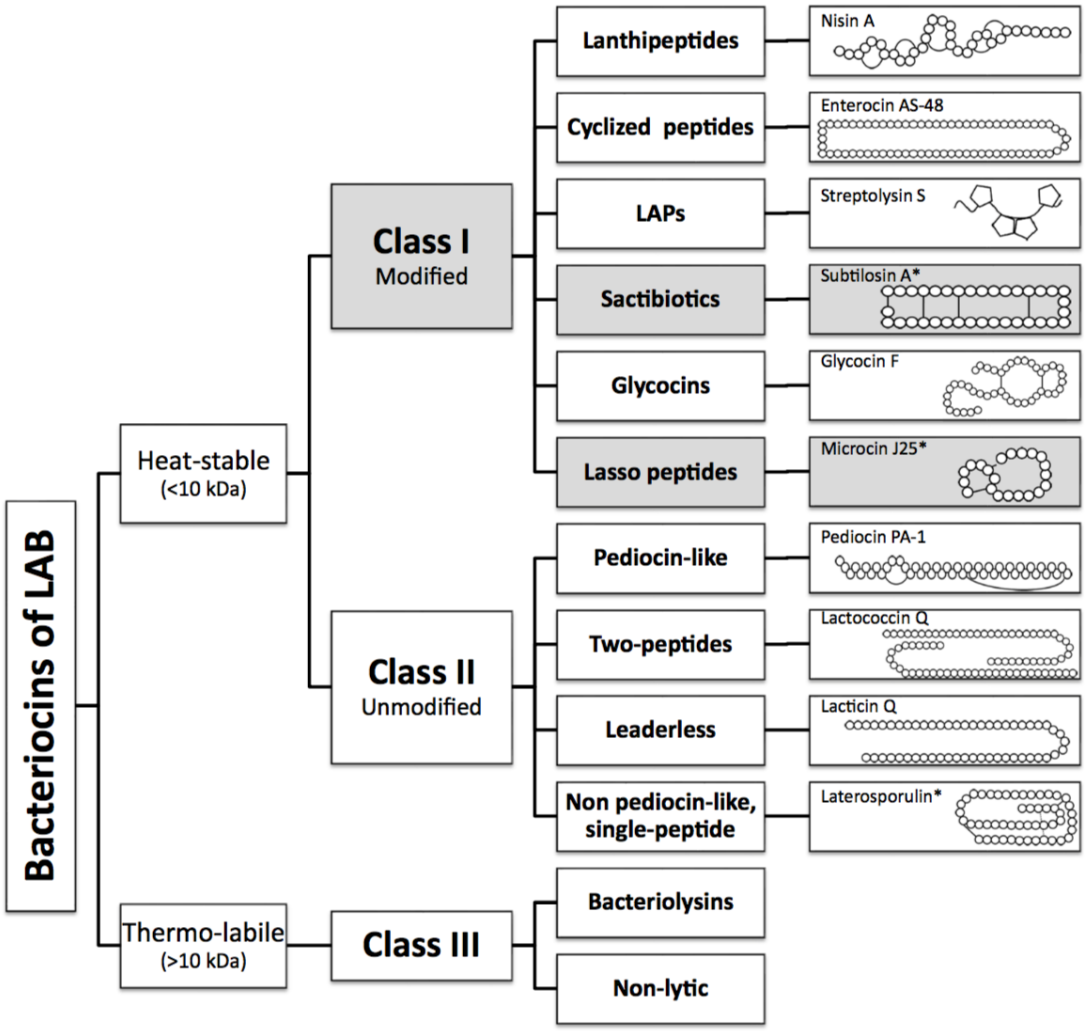


Figure 1.5: **The current classification scheme for bacteriocins**
Bacteriocins can be divided into three classes: Class I is small (< 10 kDa) peptides, which are post-translationally modified; Class II are small (< 10 kDa) unmodified peptides, and Class III are large (> 10 kDa) unmodified peptides. Grey shading indicates bacteriocins with no members produced by lactic acid bacteria, as of 2016. Recent work includes two additional groups to Class I (bottromycins and thiopeptides), and an additional member to Class III (tailocins)[4]. The focus of this work, glycocin F, is a member of the Class I bacteriocins: **Glycocins**. Figure taken from Alvarez-Sieiro *et al.* (2016).[8] Reprinted with permission from Springer.

lanthipeptides constitute a large class of peptides that has been further categorised into four sub-classes, and only two of these (Class I and II lanthipeptides) exhibit antimicrobial activity[96]. Other Class I bacteriocins which contain modified rings in their structure include the five-membered thiazole/oxazole-containing peptides, linear azol(in)e-containing peptides (LAPs), thiopeptides, bottromycins, and the sactibiotics. In addition to this, Class I also contains cyclised peptides (modified with an isopeptide bond between the N- and C-terminals), the lasso peptides (where the N-terminus is fused to the carboxyl moiety of a Glu or Asp, with the C-terminus passing through this ring (Figure 1.6)), and the sugar-modified glycocin class (see Section 1.2.3)[4,8].

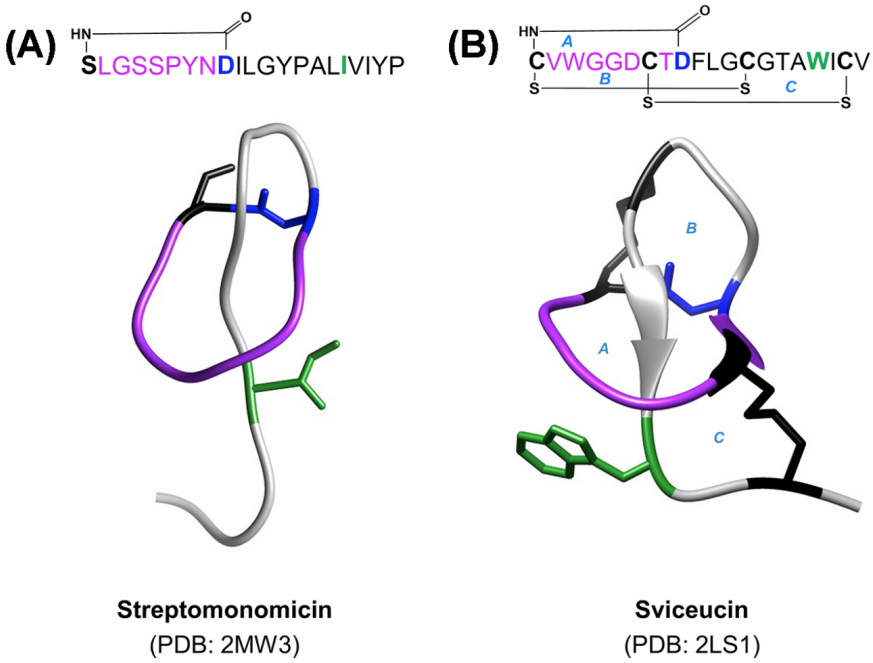


Figure 1.6: **Example structures of lasso peptides**
The NMR solution structures of streptonomonomicin (A) and sviceucin (B) are shown, along with their respective sequence above them showing the specific modifications. Figure taken from Acedo, J. Z. Chiorean, S., Vederas, J.C., and van Belkum, M. J. The exanding structural variety among bacteriocins from Gram-positive bacteria, *FEMS Microbiology Reviews*, 2018, 42, 805-828, by permission of Oxford University Press.

Of the Class I members with characterised mechanisms of action, the most common is the formation of pores in the cell membrane, leading to cell leakage and death (Table 1.2). Additional mechanisms used by Class I bacteriocins include inhibiting cell wall synthesis, inhibiting protein synthesis, or even targeting specific processes such as uncoupling ATPase activity[4,8]. An important distinction between different

bacteriocins with shared mechanisms is that they often utilise different targets. For example, while lantibiotics, sactibiotics and cyclised peptides all form pores in the cell membrane, the molecules they target differs. Lantibiotics target and sequester the cell wall synthesis component lipid II, preventing cell wall synthesis under low concentrations of the bacteriocin, or aggregating and forming pores in the cell membrane[38,81]. There is evidence for nisin also forming pores in the cell membrane even under low concentrations of lipid II, through a lipid II-independent mechanism[127]. Cyclised peptides, on the other hand, are recruited to and localised by the maltose ABC transporter before presumably diffusing into the cell membrane where they form pores[70]. Because of their structural characteristics and low MIC values it is unlikely that sactibiotics spontaneously enter cell membranes, and thus is expected that they also target an as yet unidentified molecule[151].

The targeting of specific bacterial proteins or molecules also dictates the phylogenetic range for each bacteriocin. A notable example is the lasso peptides, which have been found to target a specific genus of bacteria[4]. For example, the lasso peptide streptomonomicin, produced by *Streptomonospora alba*, has been found to be most active against *Bacillus* species[113], while the peptides lassomycin and lariatins A and B, produced by *Lentzea kentuckyensis* and *Rhodococcus sp.* respectively, target *Mycobacterium tuberculosis*[76,86]. This high specificity employed by the lasso peptides is thought to originate from the peptides folding in such a way as to mimic specific molecules in target cells[4].

Class II bacteriocins

Class II bacteriocins are small peptides that are not post-translationally modified, with the exception of having disulfide bonds[8]. This class is divided into four groups: the pediocin-like or YGNG-motif containing peptides, the two-peptide bacteriocins, the leaderless peptides, and other (linear) peptides (Figure 1.5, Table 1.2). Of this class, perhaps the best understood group is the YGNG-motif containing peptides, which are characterised by the presence of an antiparallel N-terminal β-sheet containing region, which is connected to an α-helix by a hinge, followed by a disordered 'random coil' C-terminus (Figure 1.7)[21,47]. These bacteriocins contain a conserved 'YGNGV' sequence in the N-terminus, and generally contain between one and three disulfide bonds to stabilise the structure[21,4]. The primary target of the YNGN-motif containing peptides is the mannose PTS system (Man-PTS)[94], although the bac-

Table 1.2: Bacteriocins subclasses with characterised mechanisms of action

Class	**Group**†	**Main target/ mechanism of action**	**References**
Class I	Lantibiotics (nisin)	Lipid II/ Inhibition of cell wall synthesis, membrane pore formation	29,38,81
	Thiopeptides	50S ribosomal subunit/ Inhibition of protein synthesis	20
	Bottromycins	50S ribosomal subunit/ Inhibition of protein synthesis	123
	Sactibiotics (subtilosin A)	Unknown/ Membrane pore formation	151,16
	Lasso peptides	Variable/ Inhibits cell wall synthesis (streptomonomycin), uncouples ATPase activity from ClpC1 ATPase (lassomycin)	4
	Cyclised peptides (garvicin ML)	Maltose ABC transporter/ Membrane pore formation	70
Class II	YGNG-motif containing peptides (pediocin PA-1)	Mannose PTS EIIC domain/ Membrane pore formation	94
	Two-peptide bacteriocins	Variable/ Membrane pore formation	4
	Leaderless peptides	Unknown/ Membrane pore formation	4
	Other/ linear bacteriocins	Mannose PTS (some)/ Membrane pore formation, inhibition of cell wall synthesis	59,4
Class III	Bacteriolysins	Cell wall / Peptidoglycan degradation	4
	Non-lytic (dysgalacticin)	Glucose and Mannose PTS/ Inhibition of sugar uptake	150
	Tailocins	Specific lipopolysaccharides/ Cell lysis	141

†Parentheses indicate the main bacteriocin characterised

teriocins are believed to kill cells through formation of membrane pores[47]. While it is generally believed that the Man-PTS may be utilised as a 'docking' molecule, the specific mechanism by which the membrane disrupted is not yet understood. In particular, there is evidence that the conserved N-terminal residues are involved in interacting with the Man-PTS, and the variable C-terminal region is involved in dictating the antimicrobial spectrum[47].

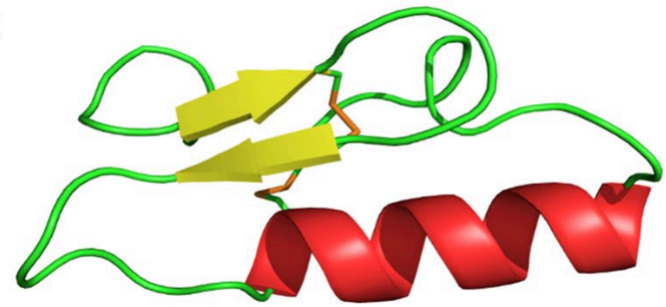


Figure 1.7: **Structure of pediocin PA-1**
The structure of a member of the YGNG-motif containing peptides, pediocin PA-1, is shown. The antiparallel, N-terminal β-sheets are shown in yellow, and the α-helix shown in red. Figure taken from Bédard *et al.* (2018)[21], under the Creative Commons Attribution 4.0 International License (https://creativecommons.org/licenses/by/4.0/).

The remaining members of Class II bacteriocins mostly kill target cells through membrane pore formation (Table 1.2). The two-peptide bacteriocins are unique in that they are characterised by the production of two linear peptides, which are either inactive or weak individually, but exhibit enhanced bactericidal activity when present together[8,118]. Both peptides are unstructured in water, but form α-helices under membrane or membrane-mimicking conditions[4]. Two-peptide bacteriocins have also been shown to target cell-surface receptors. Examples are lactococcin G, which targets membrane-bound undecaprenyl pyrophosphate phosphatase[93], while an amino acid transporter is involved in the susceptibility of cells to plantaricin JK[122]. Leaderless peptides, which are synthesised without a leader sequence, have been found to target other membrane proteins, such as a Zn-dependent metalloprotease[156]. There are also examples of leaderless bacteriocins that require no specific targets such as lacticin Q[173,174]. Lastly, the remaining members of the Class II bacteriocins encompass all other small, unmodified, leader-containing peptides. Of these, at least one member (lactococcin A) has been found to target the Man-PTS and form membrane pores, much like the YNGN-motif containing bacteriocins[59],

while another has been found to function by inhibiting septum formation through binding to lipid II[110].

Class III bacteriocins

Class III bacteriocins are larger than 10 kDa in size, generally heat-sensitive, and are grouped into three main sub-classes: the bacteriolysins, which possess an enzymatic function; non-lytic large bacteriocins, which lack enzymatic function; and the recently included class, the tailocins, which resemble phage tail particles (Figure 1.5; Table 1.2)[4]. The bacteriolysins generally function by enzymatically degrading the peptidoglycan crosslinks of cell walls, thus leading to cell lysis. Examples include lysostaphin, which cleaves the pentaglycine bridge of peptidoglycan, and zoocin A, which recognises the D-Ala-L-Ala crosslink[4,8]. Interestingly, the immunity factor of zoocin A, *zif*, encodes an enzyme that installs L-Ala in place of D-Ala in the cell wall, thus altering the recognition of its associated bacteriocin[74]. The non-lytic bacteriocins, on the other hand, function differently to the bacteriolysins. For example, dysgalactin, produced by *Streptococcus dysgalactiae*, has been found to inhibit uptake of glucose by target bacteria, presumably through blocking the glucose and mannose-specific PTS transporters. This irreversible binding results in cell starvation[150]. Finally, the tailocins are the largest class of bacteriocins, over 1 MDa in size[4]. They are composed of multiple proteins, and resemble the tail section of bacteriophage, and kill target cells by lysing the membrane. It is estimated that as little as one molecule is enough to kill a target cell, although the production of these molecules comes at a cost: the bacteria producing these bacteriocins are themselves lysed upon the synthesis and release of these bacteriocins, akin to the release of new bacteriophage from a host[141].

1.2.2 Immunity to bacteriocins

Because bacteriocins generally target specific receptors or proteins, that are widely found in strains that are closely-related to the producer, the bacteriocin producing cells need a way of protecting themselves from their own peptides. This is most commonly achieved through the expression of a so-called 'immunity protein', which forms part of the bacteriocin-encoding operon[49]. The first immunity protein mechanism was determined for the Class II d bacteriocin lactococcin A, which functions

by targeting the mannose PTS (Man-PTS) and entering the cell membrane to form pores (see Section 1.2.1). The immunity protein was found to associate with the Man-PTS only when the bacteriocin was present, and to form a strong complex, effectively locking the bacteriocin and transporter together[59]. In fact, growth time on glucose or mannose as the sole carbon source was seen to increase when immune protein-expressing cells were grown in the presence of lactococcin A. This mechanism of immunity is also thought to be used by Class IIa bacteriocins, which also target the Man-PTS[47]. A similar mechanism exists for the nisin immunity protein, NisI, which binds to lipid II and prevents the lethal nisin·lipid II complex from forming[6]. Another resistance mechanism for nisin exists, where cells express a serine peptidase which cleaves 6 amino acids off the C-terminus off the bacteriocin, decreasing the bacteriocins affinity for the cell membrane and rendering it inactive[149]. In contrast, zoocin A, alters specific amino acids in the peptidoglycan crosslinker of the producing cell, preventing the bacteriocin from cleaving the crosslinks[74]. However, for numerous bacteriocins the immunity protein and mechanism remains unknown. Knowledge of these would undoubtably provide valuable insights into the mechanism of action of their associated bacteriocins.

1.2.3 The glycocin class of bacteriocins

The glycocin class of bacteriocins are characterised by the post translational addition of a sugar unit (typically glucose or GlcNAc) to a serine or cystine residue[119]. The first glycocins (glycocin F and sublancin 168) were reported in 2011, and approximately 8 more glycocins have been characterised since then (see also Table 1.3)[119,121,132,145]. Aside from sharing the trait of being glycosylated, glycocins possess a helix-loop-helix motif, where the short antiparallel helices are locked in position by two nested disulfide bonds[119] (see also figures of glycocin F and sublancin in Figure 1.8 and Figure 1.9). The length of both the α-helices and the interhelical loop varies between glycocins, but in general can be divided into four 'scaffolds': a GccF-type (based off of glycocin F), a SunA-type (based off of sublancin and thurandacin), and enterocin 96- and enterocin F4-9-types, based of their foundation members[119,132]. The operons encoding glycocins contain five crucial genes, with an optional sixth: a gene encoding the pre-bacteriocin scaffold; a glycosyltransferase, to attach the required sugar moiety to the pre-peptide; one or two thioredoxins (or thioredoxin-like genes), presumably to facilitate correct formation of the disul-

fides; and an ABC transporter and associated peptidase, which have the dual role of cleaving the leader peptide from the pre-peptide before exporting the mature glycocin[5,90,132,161]. The sixth gene is the immunity factor, which provides the producer organism with immunity to its own bacteriocins[5,64]. The dedicated glycosyltransferase of each glycocin has been found to be specific for the sugar that gets added to the peptide scaffold *in vivo*. The strict requirement of these five genes in the operon has enabled genome mining. As a result, the heterologous expression of bacteriocin genes from various bacteria has been reported[90,132,161], using data obtained from programs such as BAGEL (Bacteriocin Genome Mining Tool)[157] and RODEO (Rapid ORF Description and Evaluation Online)[152].

No complete mechanism of action has yet been determined for any of the glycocins. Data indicates that the glycocins can exhibit one of two effects on target and susceptible cells: a bactericidal effect, where cells are killed outright (suspected for sublancin[121,124]), or a bacteriostatic effect, where cells are not killed outright, and can be recovered from stasis (confirmed in glycocin F[145], tentatively confirmed in enterocin F4-9[107]). A possible reason for the observed differences between these two glycocin mechanisms is not yet understood. Interestingly, both glycocin F and enterocin F4-9 are glycosylated with two GlcNAcs[119]. All native glycocin-producing strains are Gram-positive bacteria, and only appear to exhibit activity on other Gram-positive bacteria. The only exception is *E. coli* JM109, which contains a leaky secretin and thus would allow peptides to pass through the outer membrane[171]. The inability of glycocins to cross membranes unassisted has led to speculation that these peptides might target specific receptors on the outer membrane of Gram-positive bacteria, although there is no conclusive evidence supporting this[119,24]. Efforts have only been made to properly characterise the mechanism of action of two glycocins to date: sublancin 168 and glycocin F. The evidence to determine the mechanism of action of these glycocins will be discussed in the following sections.

1.2.3.1 Sublancin 168

Sublancin 168 is currently the most researched glycocin, with around 19 publications based on structural and genetic characterisation and chemical synthesis, at the time of writing. Numerous other papers have focussed on its use in animal and cell culture models. It was originally isolated by accident from *Bacillus subtilis* 168, where it was discovered as a contaminant in a strain of *B. subtilis* engineered to

Table 1.3: Current list of known glycocins

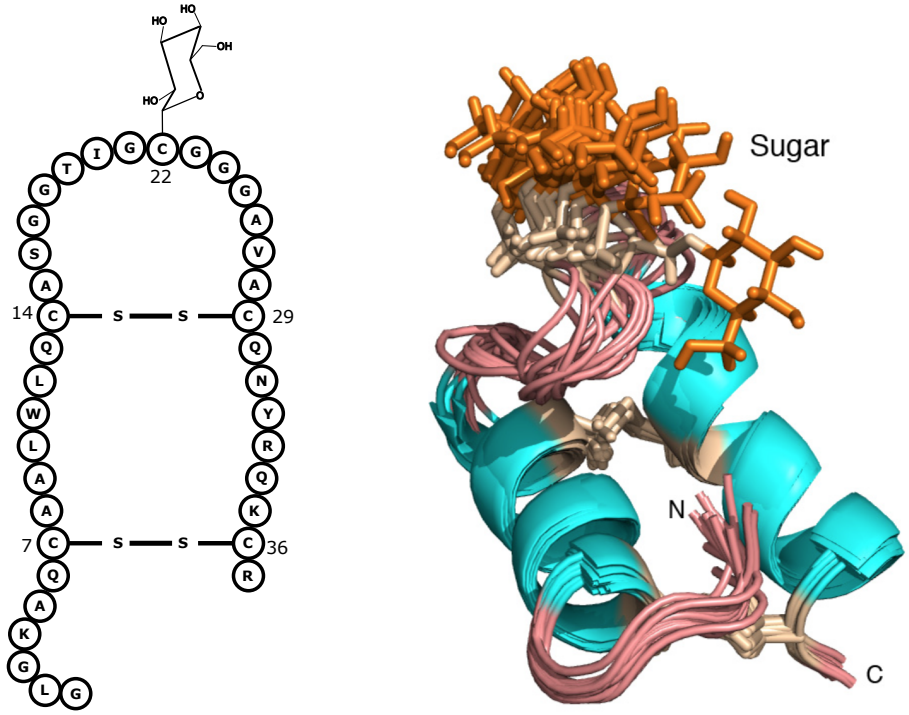
Glycocin	**Producer**	**Year discovered**	**Reference**
Glycocin F	*Lactobacillus plantarum* KW30	1996	[91]
Sublancin 168	*Bacillus subtilis* 168	1998	[124]
Enterocin 96‡	*Enterococcus faecalis* WHE 96	2009	[88]
ASM1	*Lb. plantarum* ASM1	2010	[83]
Thurandacin†	*B. thuringiensis serovar andalousiensis* BGSC 4AW1	2013	[161]
Enterocin F4-9	*E. faecalis* F4-9	2015	[107]
Bacillicin CER074†	*B. cereus* CER074	2018	[132]
Bacillicin BAG2O†	*B. cereus* BAG2O	2018	[132]
Geocillicin†/ Pallidocin†	*Geobacillus* sp. 8/ *Aeribacillus pallidus* 8	2018	[132,90]
Listeriocytocin†	*Listeria monocytogenes* SLCC2540	2018	[132]

†Discovered through genome mining, heterologously expressed in *E. coli*
‡Speculated to be a glycocin, but experimentally unconfirmed

produce subtilin[124]. Sublancin was initially thought to be a lanthipeptide based on the similarity of the leader peptide to other characterised lanthipeptides, the requirement of thio-disulfide oxidoreductases for full activity, and the lack of any free Cys residues following alkylation[60,124]. However, the results of Edman sequencing and mass spectrometry did not completely agree with the predicted structure. The true structure of sublancin 168 was not properly confirmed for over a decade, where it was revealed to be monoglycosylated with a single glucose on Cys22[121]. The NMR structure of sublancin confirmed its classification as a glycocin in 2014[73].

The genes that encode sublancin are located on the SPβ prophage of *B. subtilis* 168[56,124]. The structure of sublancin is fundamentally the same as the other glycocins, with two alpha-helices tethered together by two nested disulfide bonds, between Cys7-36 and Cys13-29 (Figure 1.8). It is monoglycosylated with a glucose on Cys22, which is situated in the interhelical loop joining the two α-helices. The interhelical loop that appears between residues Cys13 and Cys29 is described to exhibits a relatively flexible but defined conformation[73]. An overlay of 15 lowest-energy NMR conformations indicates there is some movement of the loop within a fixed region, with hydrogen bonds occurring between the amide proton of Gly17 and Gly18 with the carbonyl moiety of Gln13, and between the amide proton of Ile20 and side chain oxygen of Thr19[73]. The glucose, on the other hand, appears to possess conformational flexibility. The glycosyltransferse of sublancin, SunS, has been shown to be somewhat promiscuous with regards to the sugar subunit provided *in vitro*, although in all *in vivo* productions of sublancin SunS shows preference for glucose[121].

Sublancin has been shown to be bactericidal against a wide range of Gram-positive bacteria, including *S. aureus*, *Streptococcus pyogenes* and *Bacillus* species, as well as being active against spores[121,124]. However, the mechanism of action of this bacteriocin has yet to be determined. Genetic knockout evidence suggests that both the large mechanosensitive channel (MscL) and glucose-specific PTS transporters are involved in susceptibility to sublancin, but the precise role that each of these transporters plays is not yet understood[72,99]. Additionally, high concentrations of NaCl have been shown to be somewhat protective for susceptible strains of *B. subtilis* or *S. aureus*, which has been be linked to the role of the MscL in susceptibility[99]. Studies with propidium iodide indicated that sublancin kills its target cells in a non-lytic manner, although OD_{600} values for cells decrease following treatment[72]. In addition, a study monitoring the uptake of components of DNA, RNA, protein



(a) Schematic figure (b) NMR structure

Figure 1.8: **Structure of sublancin 168**
(a) Amino acid sequence of sublancin, showing glycosylation at Cys22. (b) NMR structure of sublancin, showing nested disulfides between both α-helices. NMR image taken from Garcia De Gonzalo *et al.* (2014)[73] (https://pubs.acs.org/doi/10.1021/cb4008106). Figure may not be redistributed further without explicit permission from ACS.

and peptidoglycan in *B. subtilis* following sublancin treatment indicated that DNA, RNA and protein synthesis pathways were all being repressed, with only a slight repression of the peptidoglycan pathway, suggesting that sublancin may be affecting an upstream pathway reliant on DNA replication[168].

The glycoactivity of sublancin has been the source of some debate. The initial publication regarding its status as a glycocin stated that the bacteriocin requires Cys22 to be glucosylated for activity[121], however subsequent research studying the glycosyltransferase SunS suggested that the aglycone exhibited comparable activity to the glycosylated form[162]. The experiments which showed this were solid agar indicator plate assays against a susceptible strain of *B. subtilis*, treated with sublancin pre-peptide which had been subjected to leader peptide removal and oxidative folding *in vitro*, but had not been subjected to glycosylation by SunS. Sublancin was once again deemed to be glycoactive in 2017, with new evidence suggesting that the C22S (non-glucosylated) mutant is inactive compared to the wild-type[25]. The reason for the initial plate clearing of the aglycones was attributed to high (1 mM) EDTA concentrations in the *in vitro* solution. Additional evidence in support of sublancin being glycoactive came from Wu *et al.* (2018), who showed that the presence of glucose and galactose at Cys22 resulted in similar MIC values, but placement of mannose or GlcNAc at that position resulted in decreased activity[168]. Additionally, susceptible cells can be protected from sublancin by the presence of excess glucose, and possibly even recovered from sublancin treatment after a short time following the addition of PTS sugars[72,132].

1.2.3.2 Geocillicin/ Pallidocin

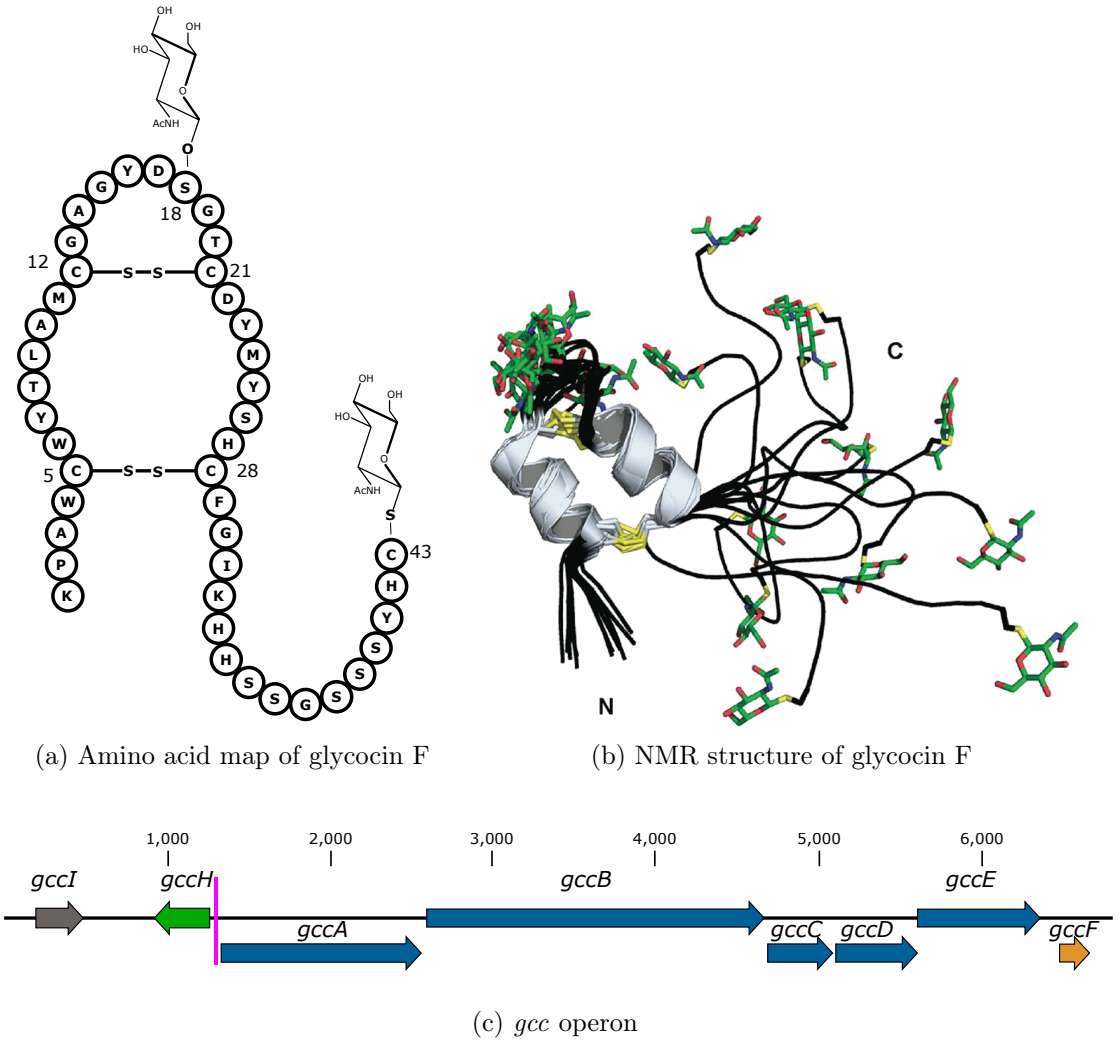
One curious glycocin which has recently been characterised by two different groups simultaneously is geocillicin/ pallidocin, which was identified in the genome of thermophile *Geobacillus* sp. 8, also referred to as *Aeribacillus pallidus* 8 (Table 1.3)[90,132]. In both cases, the gene cluster encoding the glycocin was cloned and refactored in a strain of *E. coli*, where it was expressed and post-translationally modified. The mature peptide is made up of 38 amino acids, and is glucosylated on Cys25 of the 16-residue interhelical loop, which is around the same size as the interhelical loop of sublancin[73]. Kaunietis *et al.* found that the dedicated glycosyltransferase for this glycocin (dubbed PalS) was able to successfully install a glucose on the sublancin pre-peptide *in vitro*[90]. This is not surprising, as the process that Ren *et al.* under-

took to discover the gene cluster for geocillicin/ pallidocin was to search for glycocin gene clusters using RODEO, with the sublancin glycosyltransferase gene *sunS* as a query sequence[132], suggesting there is an evolutionary relationship between these two glycocins.

The antimicrobial spectrum of geocillicin/ pallidocin was investigated by both groups, and weak activity was found by both against *B. cereus* ATCC 14579, with glucose providing a protective effect against this glycocin as was reported for sublancin, which would suggest a similar mode of action to sublancin[132]. Unsurprisingly, geocillicin/ pallidocin was found to be most active against other thermophiles, with the highest activity being reported against *Parageobacillus genomospecies* 1 NUB36187, with an MIC value of 2.4 pM[90]. The ability of glucose to protect thermophiles from the activity of this glycocin was not tested. In particular, the possibility that geocillicin/ pallidocin has a mechanism of action that is the same as sublancin could indicate a common mechanism or mechanisms used by glycocins across a range of bacteria. If this was the case, the possibility this could be exploited and tuned to target specific strains or species of bacteria may lead to the development of a new class of antimicrobials.

1.2.3.3 Glycocin F

Glycocin F (GccF) is a 43 amino acid bacteriocin that is produced by *Lb. plantarum* KW30, which was isolated from fermented corn[91,145]. It was one of the first glycocins to be characterised, along with sublancin 168. It is diglycosylated with two GlcNAc moieties, one on Ser18 (located on the interhelical loop), and the other at Cys43, at the C-terminus of the peptide (Figure 1.9). However it is only the Ser18-linked GlcNAc that is essential for activity[145,159]. The cysteine linked GlcNAc, while not absolutely necessary, enhances antimicrobial activity[145]. GccF imparts a unique effect on susceptible cells: rather than killing them outright, it induces a strong rapid bacteriostasis, from which cells are still viable and can be recovered. This is characterised by a halt in the increase of the OD_{600} of treated cells, but with no drastic decrease in OD_{600} or colony-forming units (CFU)[145]. In particular, recovery from this stasis can be immediately induced upon the addition of free GlcNAc[145]; likewise, cells are protected from the effect of GccF as long as GlcNAc is present, similar to the effect reported for glucose and sublancin[132]. No other sugar has been reported to be able to recover cells from GccF-induced stasis (unpublished results).



(a) Amino acid map of glycocin F

(b) NMR structure of glycocin F

(c) *gcc* operon

Figure 1.9: **Structure of glycocin F**
(a) Amino acid map of glycocin F, with the N-terminus on the left and C-terminus on the right. The O- and S-linked GlcNAc molecules can be seen, linked to serine 18 and cysteine 43, respectively. (b) NMR structure of GccF (PBD accession number 2KUY). The comparative disorder of the C-terminal tail can be seen. (c) The operon encoding GccF. The gene encoding the pre-peptide is highlighted in yellow, genes required for production in blue, and the immunity protein in green. The magenta line indicates the divergent promoter. NMR figure reprinted with permission from Venugopal *et al.* 2011.[159] Copyright © 2011 American Chemical Society.

Compared to sublancin, which has 14 amino acids between the cysteines of the topmost disulfide bridge (Figure 1.8 (a)), glycocin F has a smaller interhelical loop, with only 8 amino acids between the alpha-helices (Figure 1.9 (a)). The interhelical loop of GccF displays limited flexibility, as does the Ser18-linked GlcNAc[119,159]. Interestingly, there are no obvious hydrogen bonds between the interhelical loop and the rest of the protein structure[119]. Both helices are composed of 8 residues, with the N-terminal helix occurring between Cys5-Cys12, and the C-terminal helic occurring between Cys21-Cys28 (Figure 1.9). The N-terminal helix is amphipathic, sitting at a 25^o angle to the C-terminal helix[119].

Perhaps the most unique physical aspect of GccF is the glycosylation on the C-terminus. Aside from the closely-related ASM1, GccF is the only glycocin characterised to date which contains a glycosylated 'tail'. It was thought that thurandacin could be be glycosylated at two sites, though both sites are located on the interhelical loop[161], while studies of enterocin 96 have shown that it can be disglycosylated *in vitro* on a single residue by the glycosyltransferase EntS[117]. Moreover, the role of the C-terminal GlcNAc is still unknown, as the presence of the tail is not required for activity; removal of regions His42-Cys43GlcNAc by chymotrypsin or His33-Cys43GlcNAc by trypsin still results in active peptide, although the activity is reduced 200- and 75-fold, respectively[145]. This is in contrast to removal of the Ser18-linked GlcNAc by N-acetyl-β-D-glucosaminidase, which has been shown to completely abolish activity of GccF[145]. The operon encoding GccF only contains one glycosyltransferase, *gccA*[5,145], which would suggest that this enzyme installs GlcNAcs to both positions, despite the recipient amino acids being different.

The operon encoding GccF and accessory proteins is comprised of 7 essential genes, with six of these (*gccA-F*) being transcribed as a single unit: *gccA* encodes the glycosyltransferase, *gccB* the ABC transporter, *gccC* and *gccD* two thioredoxins, *gccE* a putative LytTR domain-containing response regulator, and *gccF*, the prepeptide; the seventh gene, *gccH*, encodes the immunity protein (Figure 1.9 (c))[5]. Genes *gccA-E* are constitutively transcribed throughout GccF production, while *gccF* shows significantly higher transcript levels, suggesting the presence of a second promoter directly upstream of this gene[5]. The conditions which initiate production of GccF are currently unknown, as is the role of *gccE*, although there is suspicion that this gene may play a role in regulating bacteriocin production. The role of the immunity gene was initially given to the prematurely-named upstream ORF *gccI*, as this was transcribed in the same direction as *gccA-F* even though it was

not a part of that same transcript. However, a study of the transcription profile of the *gcc* operon during GccF production showed that *gccI* was not expressed with the other genes of the operon, although *gccH* was expressed constitutively, which was surprising given that it faces the opposite direction of the rest of the operon (Figure 1.9 (c))[5]. The role of GccH as the immunity protein was later confirmed by recombinant expression, and shown to provide protection from GccF even when expressed in low quantities[17].

Current evidence for a possible target of GccF suggests the GlcNAc-specific PTS (GlcNAc-PTS) as the most likely candidate[62]. This is based off of the whole genome sequencing of a number of GccF-resistant mutants of *Lb. plantarum* ATCC 8014 and ATCC 14917, which revealed mutations restricted largely to the membrane-bound EIIC domain of the GlcNAc-PTS (*pts18CBA* gene in *Lb. plantarum*). Subsequent knockouts of *pts18CBA* in *Lb. plantarum* NC8 resulted in some degree of resistance to high concentrations of GccF (up to 1 μM), but cell growth was still slowed, indicating that there may still be another target of GccF[62]. This is supported by evidence of GccH providing a boost to cell growth in *pts18CBA* knockout strains of *Lb. plantarum* NC8 that were treated with GccF[17]. One suggestion is the second GlcNAc-PTS transporter found in *Lb. plantarum*, *pts22CBA*, which may be behaving as a secondary receptor for GccF[62]. Another is that there is an unknown second target involved in the mechanism.

The antimicrobial spectrum of GccF extends beyond other strains of *Lb. plantarum*. Kerr (2013) found that GccF also affects other Gram-positive bacteria including *Bacillus megaterium*, *Enterococcus faecalis*, *E. faecium*, and some strains of *Streptococcus agalactiae*, *S. mutans*, *S. pneumoniae* and *S. pyogenes*[92]. Moreover, free GlcNAc displayed different effects on susceptible bacteria when coupled with GccF treatment, such as protecting cells from the bacteriocin (*Lb. plantarum* and *E. faecium*), increasing sensitivity of cells to GccF (*E. faecalis* and *B. megaterium*), or having no effect (*Strepococcus* sp.)[92]. As with other glycocins, leaky secretin mutants of *E. coli* were also found to be susceptible, with GlcNAc also providing a protective effect. It is also interesting to note that the bacterial species which demonstrated sensitivity to GccF all possess GlcNAc-specific PTS transporters with sequence similarity to PTS18CBA[92].

Recently, a chemical synthesis for GccF was established. The resulting peptide was correctly folded and with the correct disulfide bridges in place, as confirmed by mass

spectrometry, circular dichroism, and NMR[30]. The chemically-synthesised GccF exhibited close to wild-type activity, although the presence of an amidated C-terminus was believed to contribute to the slight decrease in activity seen (IC_{50} of native GccF was $\sim$ 0.9 nM, IC_{50} of amidated, synthetic variant was $\sim$ 1.6 nM). The development of this synthetic route opens up the possibility of exploring other variants or analogues of GccF that are otherwise inaccessibly by traditional, recombinant methods.

1.3 Aims and objectives

The glycocins as a subset of bacteriocins are a growing class of antimicrobials, but they are still largely not well understood. In addition to this, glycocin F appears to convey a unique mode of action that is possibly only shared by enterocin F4-9, which makes it of particular interest with regard to finding new bacterial combating mechanisms. On top of this, GccF has unique structural characteristics compared to other glycocins, which warrant further study of its activity and properties. The focus of this research was to shed light on glycocin F's mechanism of action, and was carried out by trying to achieve the following objectives:

1. **Identify the role of different structural features of GccF**
   The successful chemical synthesis of glycocin F opened up a new avenue for studying structure-function relationships of this peptide, and allowed for design and testing of variants of GccF which are not easily accessible through recombinant DNA methods. The first objective was to design different structural variants of GccF, and assess the effect of each modification on activity against the indicator strain *Lb. plantarum* ATCC 8014.

2. **Study the transcriptional response of susceptible bacterial cells to GccF**
   In order to monitor the internal response of bacteria to GccF, RNA-seq was carried out on *E. faecalis* JH2-2 treated with GccF in two independent experiments. The first experiment involved monitoring how the response changed from early (0 minutes) to mid (10 minutes) or late (30 minutes) stages of treatment, while the second experiment involved studying what genes are affected by smaller concentrations of GccF.

3. **Identify bacterial proteins interacting with GccF's immunity protein, GccH**
   The identification of the immunity protein of GccF, GccH, provided a new avenue for trying to identify the primary target, by attempting to co-precipitate the immunity protein and complex with GccF and the target. This was based on a method employed by other groups to identify bacteriocin targets. Additionally, the synthesis of GccF molecules conjugated with photo active crosslinkers also allowed for a direct GccF-target pulldown to be attempted, although the immunity protein was not used for this experiment.

1.4 Roadmap of thesis

A number of different approaches were employed during this project, so the results have been divided into separate chapters for each approach. **Chapter 2** contains general materials, methods, bacterial strains and plasmids used throughout this research, with more specific methods being included at the beginning of each relevant chapter. **Chapters 3 - 5** contain the results: **Chapter 3** focuses on the chemically synthesised, modified analogues of glycocin F (Objective 1); **Chapter 4** details the transcriptomic response of *Enterococcus faecalis* JH2-2 when subjected to GccF, and compares this response to that of other published antibiotic responses (Objective 2); and **Chapter 5** includes efforts to isolate the primary binding target of GccF in *Lb. plantarum* NC8, through a series of immunoprecipitation reactions of tagged GccF immunity protein, and tagged GccF (Objective 3). Note that the digital version of this document contains embedded links within all in-text citations, to make referring to sections and figures/ tables easier.

2 | General materials and methods

2.1 Bacterial methods

2.1.1 Bacterial strains, plasmids, and growth conditions

Strains of bacteria used in this work can be found in Table 2.1. Unless otherwise specified, strains of *Lactobacillus plantarum* were cultured in MRS media (Merck), and strains of *Enterococcus faecalis* and *E. faecium* were cultured in TSB media (Oxoid) at 30 ℃ without aeration. Strains of *Escherichia coli* were cultured in TSB or LB (Sigma Aldrich) media at 37 ℃ with aeration. Agar plates were prepared by supplementing the relevant media with 1% (w/v) bacteriological agar (Acumedia). Bacterial indicator plates were prepared by mixing 100 μL of an overnight culture of cells with molten 1% agar with the appropriate media. Plasmids used and generated in this work can be found in Table 5.1 in Chapter 5.

Table 2.1: Bacterial strains used in this work

Species	**Strain**	**Description**	**Reference**
E. faecalis	JH2-2	Plasmid-less strain	Gift from Professor Gregory Cook[89]
	V583	Primary reference strain	[136]
	VI01	V583 with *rpoN* deleted	Gift from Professor Lyn Hancock[87]
	VI40	VI01 with *rpoN* recomplemented	Gift from Professor Lyn Hancock[87]
E. faecium	64/3	Plasmid-less strain	[164]
E. coli	JM109	Leaky secretin-containing variant	[171]
	BL21	Protein expression strain	[167]
	XL-1 Blue	Cloning strain	Promega[32]
Lb. plantarum	KW30	Wild-type GccF-producing strain	[91]
	ATCC 8014	GccF indicator strain	[48]
	NC8	Plasmid-less strain	Gift from Dr Lars Axelsson[14]

2.1.2 Purification of glycocin F

4 L of *Lb. plantarum* KW30 was grown at 25 ℃ for three days without aeration. Cells were pelleted at 5000 x *g* for 15 minutes, and supernatant collected, then mixed with 100 mL of pre-equilibrated (100 mM sodium formate, pH 4.6) SP C-25 Sephadex™ resin (Sigma Aldrich) at room temperature and with gentle overhead stirring overnight. The resin was then packed into a glass column and washed with the equilibration buffer followed by 1 L 50 mM MOPS (pH 7.0). GccF was eluted from column in three 100 mL washes of 50 mM ammonium bicarbonate (ABC)

in 70% acetonitrile. Eluted fractions were tested for activity, active fractions were pooled and lyophilised, before being resuspended in 0.1% trifluoroacetic acid (TFA).

Resuspended fractions were further purified by reverse phase HPLC (Jupiter 5 μm C_{18} 300 A, 250 x 10 mm, Phenomenex; 1 mL injection, 5 mL min^{-1}, 25 min 0-50% B linear gradient. A: H_2O, 0.1 % trifluoroacetic acid (TFA); B: acetonitrile, 0.08% TFA). Eluted peaks corresponding to GccF were freeze-dried and resuspended in milliQ water. GccF was quantified using UV (ultraviolet) absorbance at 205 and 280 nm.

Quantification of glycocin F

GccF was quantified using spectroscopy. The UV absorbance of GccF was measured over the wavelength range of 300 - 200 nm in a multi-wavelength spectrophotometer (Varian Cary-300 Bio, Agilent), in a 1 mL quartz cuvette (Hellma Analytics; 1 cm pathlength). Concentration of GccF was determined from absorbance at 205 and 280 nm, using molar absorptivity values of 220,030 L mol^{-1} cm^{-1} and 18,700 L mol^{-1} cm^{-1} for wavelengths of 205 and 280 nm, respectively.

2.1.3 Growth curve analysis

2.1.3.1 96-well plate reader

Appropriate concentrations of native or synthetic glycocin F were prepared in MRS media. 150 μL volumes of each GccF dilution were added to the wells of a flat-bottomed 96 well plate, followed by 150 μL of *Lb. plantarum* 8014 indicator cells prepared to an optical density (OD_{600}) of approximately 0.01 (approximately 2.0 x 10^6 CFU/mL). Each plate contained a negative control (300 μL of media without cells) and positive control (cells with no GccF) for comparison. All samples, including controls, were prepared and analysed in triplicate. The plate was analysed using a MultiSkan™ GO plate reader using the supplied SkanIt software. The plate chamber was incubated at 30 ℃ and OD_{600} measured every 15 minutes with a five-second medium shake implemented every hour. Growth curves were measured over 15 hours.

Data was analysed using Microsoft Excel and the statistical software R[129]. Due

to the lack of a 'zeroing' function within the SkanIt software, each growth curve was normalised by subtracting the absorbance data of the negative control from the growth curve data points. $I.C._{50}$ values were determined by plotting the reciprocal of the differences between each growth curve data point and the positive control data point, effectively providing a plot of 'differences' in growth with time. The maximum peak of each plot was taken to be the percentage inhibition.

2.1.3.2 Cuvette reader

To monitor the growth of larger cultures of cells than can be grown in a 96-well plate, or to allow aeration of cultures, a cuvette reader capable of simultaneously monitoring up to 12 cuvettes was used (Varian Cary-300 Bio, Agilent). Cell cultures were prepared in 4 mL plastic cuvettes, in a final volume of 3 mL each. Absorbance at 600 nm was measured every 5 minutes for the required period of time, generally 15 hours. The internal temperature of the chamber was held constant for the cells being incubated (see Section 2.1.1). When necessary, magnetic stirrer bars were included within the cuvettes to allow for aeration of the cultures. Data was saved as comma-separated value (CSV) files and plotted in Microsoft Excel.

2.2 Cloning and DNA manipulations

2.2.1 Competent cell preparation and transformation

Preparation of chemically-competent *E. coli* cells

An overnight culture of *E. coli* cells was diluted 100-fold in 100 mL of TSB media, and incubated at 37 ℃ with aeration until an OD_{600} of 0.15 - 0.2 was reached. Following this, cells were pelleted by centrifugation at 3,000 x *g* for 10 minutes at 4 ℃, and the supernatant removed. Cells were resuspended in 20 mL of ice-cold 0.1 M $CaCl_2$ centrifuged for 5 minutes at 3,000 x *g* at 4 ℃. After removal of the supernatant cells were resuspended in 1 mL of ice-cold 0.1 M $CaCl_2$, 10% glycerol, and divided into 50 μL aliquots before being stored at -80 ℃.

Preparation of electrocompetent *Lb. plantarum* cells

An overnight culture of *Lb. plantarum* NC8 cells was diluted 100-fold into fresh MRS media. Cells were grown with mild shaking (∼120 rpm) for approximately 3 hours at 30 ℃ until an OD_{600} of 0.2 - 0.3 was reached. They were then pelleted by centrifugation for 15 minutes at 3,000 x *g* at 4 ℃, then sequentially washed with 20 mL 0.1 M $MgCl_2$, 4 mL 30% PEG 1500, and finally resuspended in 200 μL 30% PEG 1500. All wash steps were separated by centrifugation steps at 3,000 x *g* and 4 ℃, for 15 minutes. 100 μL aliquots were then immediately used for electroporation.

2.2.2 Bacterial transformation of plasmids

Transformation of *E. coli* strains *via* heat-shock

50 μL aliquots of chemically-competent *E. coli* cells were thawed on ice, and 5 μL of plasmid DNA added. After incubating on ice for 30 minutes cells were heat-shocked at 42 ℃ for 60 seconds, and 1 mL of SOC broth (2% [w/v] tryptone, 0.5% [w/v] yeast extract, 10 mM NaCl, 2.5 mM KCl, 10 mM $MgCl_2$, 10 mM $MgSO_4$) added. Cells were then incubated with shaking at 37 ℃ for 1 - 2 hours, before being plated on TSB 1% (w/v) agar plates supplemented with the appropriate antibiotics.

Transformation of *Lb. plantarum* NC8 *via* electroporation

10 μL of concentrated plasmid DNA (around 100-300 ng/μL) was added to an electroporation cuvette with 2 mm gap (Bio-Rad). 100 μL of freshly-prepared competent cells were added to the electroporation cuvette, which was then placed into an electroporator (Gene Pulser™, Bio-Rad) and pulsed at 2.0 kV, 400 Ω resistance, and 25 μF capacitance. Cells were recovered using 1 mL of MRS media supplemented with 0.3 M sucrose and 0.1 M $MgCl_2$ at 30 ℃ for two hours with no agitation. 100 μL of recovered cells were plated on 1% agar MRS plates supplemented with appropriate antibiotics.

2.2.3 PCR conditions

PCR was carried out using either the high-fidelity Phusion DNA polymerase (Thermo) for cloning, or BioTaq (Bioline) for analysis and sequencing of small regions. PCR reactions were prepared to a final volume of 25 - 50 μL, with a 1x final concentration of the requisite buffer, 200 nM of each primer, 200 μM dNTPs, 5 - 20 ng of template DNA, and typically 1 unit of DNA polymerase. Typical PCR conditions were comprised of an initial denaturing step at 98 ℃ for 2 minutes, followed by 25 to 30 cycles of a standard three-step amplification cycle (denaturing for 10 seconds at 98 ℃, annealing for 30 seconds at the lowest melting temperature of the two primers in the reaction, and amplifying for 30 seconds per kb of amplicon size at 72 ℃), with a final extension step of 5 minutes at 72 ℃.

2.2.4 Plasmid isolation from *E. coli*

Plasmids were isolated from *E. coli* using a Roche High Pure Miniprep kit. Extraction was carried out as described in the manual.

2.3 DNA and protein gel conditions

2.3.1 DNA agarose gel preparation and running conditions

Agarose gels were prepared by dissolving 0.8 - 1.2% (w/v) agarose (Hyagarose, HydraGene) in TAE buffer (40 mM Tris, pH 8.3, 20 mM acetic acid, 1 mM EDTA). Gels were typically run at 100 V for 45-60 minutes before being stained in 0.5 μg/mL ethidium bromide in TAE buffer. Ethidium bromide-stained gels were imaged on a Gel Doc XR (Bio-Rad).

2.3.2 SDS-PAGE preparation and running conditions

12 or 15% acrylamide SDS-PAGE gels were prepared according to Table 2.2, with the 5% gel incorporated as the 'stacking' gel. All components aside from the APS and TEMED were mixed together and degassed for 15 minutes. APS and TEMED

were added, and solution mixed before being pipetted in between two glass plates held within a mini-PROTEAN III casting frame (Bio-Rad). Once the separating gel had set (30 - 60 minutes) the stacking gel was prepared and pipetted on top, with a well-forming comb inserted in the top.

Protein samples were denatured by mixing with 2x SDS loading dye (100 mM Tris pH 6.8, 20% [w/v] glycerol, 0.2% [w/v] bromophenol blue, 4% [w/v] SDS, 200 mM DTT [optional]) and heating samples at 95 ℃ for 2 minutes. Gels were typically run at 120 V for 60 - 90 minutes or until the dye front reached the bottom of the gel, in 1 x SDS running buffer (0.025 M Tris, 0.192 M glycine, 0.1% SDS, pH 8.5).

Table 2.2: SDS-PAGE composition

Component	**Gel percentage**		
	12%	15%	5%
1.5 M Tris (pH 8.8)	2.5 mL	2.5 mL	-
1.5 M Tris (pH 6.8)	-	-	0.248 mL
40% Acrylamide/ bisacrylamide	3 mL	3.75 mL	0.248 mL
10% SDS	0.1 mL	0.1 mL	20 μL
10% APS	0.1 mL	0.1 mL	20 μL
TEMED	4 μL	4 μL	2 μL
H_2O	4.3 mL	3.55 mL	1.482 mL
Total volume	10 mL	10 mL	2 mL

Colloidal Coomassie staining

SDS-PAGE gels were fixed by soaking in fixing solution (10% acetic acid, 40% ethanol) for 15-30 minutes. Following this, gels were incubated with colloidal coomassie stain (0.08% [w/v] Coomassie Brilliant Blue (G250), 1.6% [w/v] ortho-phosphoric acid, 8% [w/v] ammonium sulfate, 20% [v/v] methanol) overnight at room temperature with gentle shaking before being destained with water.

Silver staining

Colloidal Coomassie-stained gels were destained by incubating with 50% methanol overnight at RT, with gentle shaking. Gels were then washed three times with water (10 minutes each) before being soaked with sensitising solution (12 mM $Na_2S_2O_3.5H_2O$, 4.5 mM $K_3Fe(CN)_6$, 4.5 mM Na_2CO_3) for 1 minute. Gels were then washed three times (10 minutes each) before being incubated with 0.1% (w/v) $AgNO_3$ for 30 minutes at RT. Gels were then washed three more times with water, before being developed with developing solution (0.05% [v/v] formaldehyde, 0.24 M Na_2CO_3) until dark bands appeared. Once the desired level of staining was developed, the reaction was stopped by washing the gels in 5% (v/v) acetic acid.

3 | Chemically-synthesised analogues of glycocin F

3.1 Introduction

Previous attempts to clone and recombinantly express glycocin F had been unsuccessful, which made studying the role of different structural components of GccF difficult outside of what could be achieved through enzymatic processing (*e.g.* proteolytic and glucosaminidase treatments)[145]. However, in 2015 Brimble *et al.* published a successful strategy for synthesising GccF *via* a combination of solid-phase peptide synthesis (SPPS) and native chemical ligation (NCL). The resulting peptide exhibited close to wild-type activity[30]. Access to this synthetic process allowed the synthesis of structural analogues of GccF designed to investigate specific structural aspects of the molecule with respect to its function. The analogues were divided into four groups based on the part of GccF modified. These groups included modifications to the glycosidic bond (attaching both GlcNAcs to the peptide *via* cysteine or serine linkages), tail modifications (decreasing the length of the tail, introducing a charged amino acid, or substituting or removing of the C-terminal sugar), interhelical loop modifications (decreasing the number of amino acids in the interhelical loop, or modifying the glycosylated serine 18), and modifications to the nested disulfide bonds (disrupting each disulfide bond in isolation). The results of this chapter of work, along with the protocol for the chemical synthesis of GccF, have been published: Amso *et al.*, 2018, and Bisset *et al.*, 2018[11,24]. This chapter will focus on the effect of these modifications on GccF's antimicrobial activity, and the interpretation of the results with respect to a possible mechanism of action.

3.2 Experimental procedures

3.2.1 Synthesis of glycocin F analogues

Chemical synthesis of the glycocin F analogues was based on the original publication of GccF synthesis by Brimble *et al.*[30]. All syntheses were carried out by the University of Auckland using a modified method: see Amso *et al.* (2018) and Bisset *et al.* (2018)[11,24].

3.2.2 Characterisation of glycocin F analogues

3.2.2.1 Determination of IC_{50}

All GccF analogues were supplied in lyophilised form. Analogues were resuspended in milliQ H_2O, to a final concentration of either 200 or 500 μM. The concentration of each analogue was confirmed using its absorbance at 280 nm (see Table 3.1; Section 3.2.2.3). GccF analogues were further diluted in MRS media (Merck) and 150 μL volumes of each GccF dilution were added to the wells of a flat-bottomed 96-well plate, followed by 150 μL of *Lb. plantarum* 8014 indicator cells prepared to an optical density (OD_{600}) of approximately 0.05 ($\sim$ 1.0 x 10^7 CFU/mL). Each plate contained two controls for comparison: 300 μL of media without cells, and cells with no GccF. All samples, including controls, were prepared and analysed in triplicate. The plate was analysed using a MultiSkan™ GO plate reader using the supplied SkanIt software. The plate chamber was incubated at 30 ℃ and OD_{600} measured every 15 minutes with a five-second medium shake implemented every hour. Growth curves were measured over 15 hours.

The resulting data were processed and analysed using Microsoft Excel and the statistical software R. Due to the lack of a 'zeroing' function within the SkanIt software, each growth curve was normalised by subtracting the absorbance data of the negative control from the growth curve data points. IC_{50} values were determined by plotting the reciprocal of the differences between each growth curve data point and the positive control data point, effectively providing a plot of 'differences' in growth with time. The maximum peak of each plot was taken to be the percentage inhibition (see Section 3.3.2 for further explanation).

3.2.2.2 Circular dichroism characterisation

A far ultraviolet (UV) circular dichroism (CD) spectrum of each analogue was collected on a Chirascan spectrometer (Applied Photophysics, UK). Samples were diluted in degassed milliQ water to final concentrations between 0.5 and 1 mg mL^{-1}. Approximately 100 μL of degassed sample in a Quartz SUPRASIL precision cell (Hellma, Germany) with a 0.1 mm path length was used to measure the far UV spectra, using the following parameters: wavelength range, 185-250 nm; time per point, 0.5 seconds; bandwidth, 1nm, step size, 0.5 nm; number of repeats 20. The

Table 3.1: Predicted and measured monoisotopic masses of GccF analogues

Analogue	**Predicted monoisotopic mass (Da)**	**Measured monoisotopic mass (Da)**	**Predicted A_{280} (1 mg/mL)[†]**
$GccF_{Syn}$	5199.0472	5199.0488	3.844
$GccF_{C,C}$	5215.0633	5215.0623	3.831
$GccF_{S,S}$	5183.0676	5183.1117	3.856
$GccF_{Cys43Ser\Delta GlcNAc}$	4979.988	4980.039	3.856
$GccF_{\Delta Ser38-40}$	4937.948	4937.995	4.065
$GccF_{Cys43Man}$	5158.114	5158.061	3.855
$GccF_{Cys43Glc}$	5158.114	5158.050	3.855
$GccF_{Ser38Lys}$	5240.108	5240.426	3.826
$GccF_{\Delta GlyI}$	5142.028	5142.076	3.890
$GccF_{\Delta GlyII}$	5085.008	5085.051	3.990
$GccF_{\Delta GlyIII}$	5027.988	5028.028	4.040
$GccF_{Ser18MethGlcNAc}$	5213.0633	5213.0937	3.844[‡]
$GccF_{Cys12/21}$	5315.150	5315.144	3.844
$GccF_{Cys5/28}$	5343.203	5343.183	3.844

Monoisotopic masses and predicted A_{280} values calculated using ExPASy ProtParam (https://web.expasy.org/protparam/)
[†]Based on amino acid sequence only.
[‡]Estimated value only - methylated serine could not be included in the estimation.

molar ellipticity of each CD spectra was normalised based on concentration and molecular weight of the analogue.

3.2.2.3 Structure confirmation using mass spectrometry

Monoisotopic masses of synthetic GccF analogues were confirmed using an LC-coupled QExactive Focus mass spectrometer equipped with a high-energy collision-induced dissociation (HCD) collision cell and an Orbitrap mass analyser using a HESI ion source. The flow rate was 0.1 mL min^{-1} and the solvent 50% Acetonitrile/water in 0.1% formic acid. The sample was introduced into the mass spectrometer using a syringe pump on bypass mode. Full MS1 scans were acquired over a mass range of 500-3,000 m/z with detection in the Orbitrap mass analyser at a resolution setting of 70,000. The raw spectra collected were visualised and deconvoluted using FreeStyle 1.1 SP1 (ThermoFisher Scientific). Monoisotopic masses were compared to the predicted monoisotopic masses (Table 3.1).

3.3 Results and discussion

3.3.1 Structural characterisation of GccF analogues

Before activity studies were carried out on the GccF analogues, the compounds had to be characterised to confirm that they had been correctly synthesised and the disulfide bonds had been formed. Two methods were employed to characterise the compounds: mass spectrometry, to confirm the modifications, and circular dichroism (CD) spectroscopy, to study the effect of the given modifications on the secondary structure. Details of these tests are described in Section 3.2.2.3. A summary of the expected monoisotopic masses, along with the measured masses, is given in Table 3.1. From the table it can be seen that the masses were as expected, indicating that the nested disulfide bonds had formed, and all the correct modifications were present.

The far UV CD spectra of GccF has a positive peak at around 190-195 nm, and a negative peak at $\sim$ 210 nm, indicative of alpha-helical content[145], which has been confirmed in the native protein using nuclear magnetic resonance (NMR)[159]. CD spectra was measured for each of the analogues as a way of confirming that each analogue had the appropriate secondary structure, and that this had not been affected by the modification (Figure 3.1). All but two of the synthetic analogues ($GccF_{Cys12/21}$ and $GccF_{Cys5/28}$) displayed CD spectra peaks similar to those of the unmodified GccF, with some variation in the peak amplitude. See Section 3.3.3 for more discussion on this.

3.3.2 Determining IC_{50} values of GccF against *Lb. plantarum* ATCC 8014

The antimicrobial activity of a bacteriocin is normally reported as the minimum inhibitory concentration (MIC), which is the concentration of compound at which cell growth is not observed after a specific period of time, generally overnight. Typically this involves incubating cells (in 96- or 384-well plates) with a concentration range of antimicrobial agent overnight, followed by visually determining the lowest concentration at which growth was inhibited. However, when attempting to determine the activity of glycocin F this same technique cannot be used, as the bacteriostatic

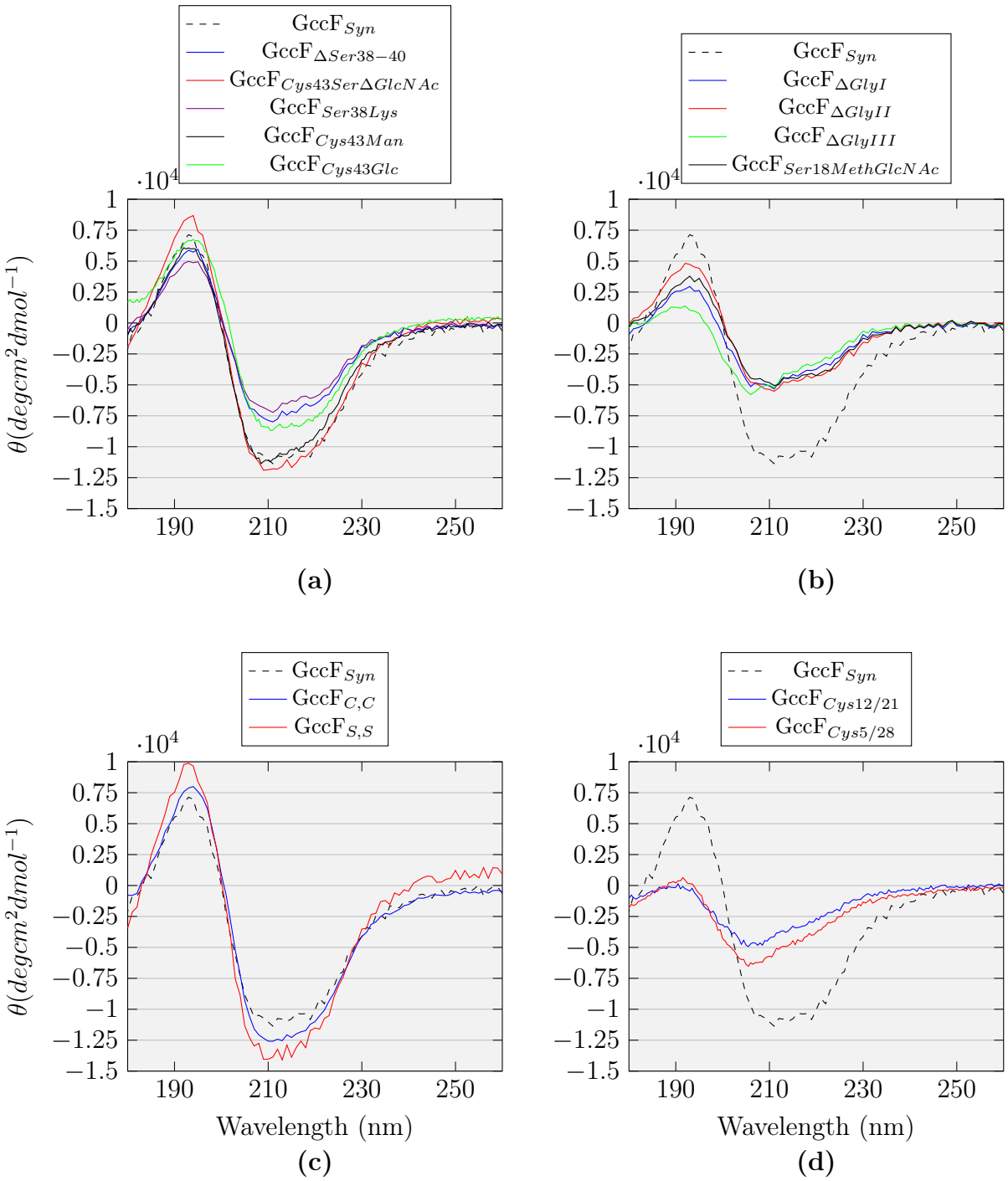


Figure 3.1: **CD spectra of GccF analogues**
The plots are as follows: (a) tail modified analogues, (b) loop-modified analogues, (c) sugar linkage-modified analogues, and (d) disulfide disrupted analogues. GccF$_{Syn}$ (synthetic, unmodified GccF) was included as a reference within each of the plots (dotted black line).

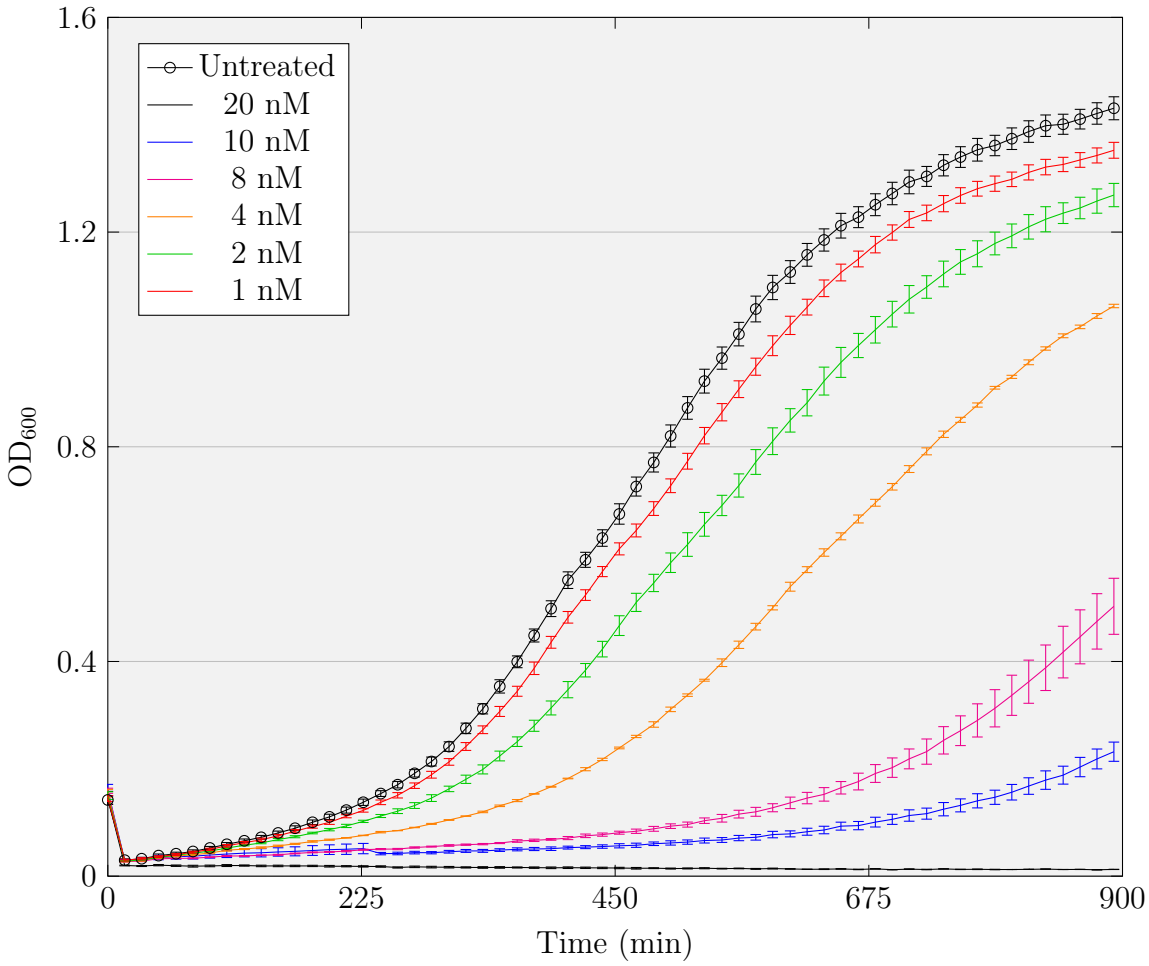


Figure 3.2: **Growth plots of *Lb. plantarum* ATCC 8014 cells treated with different concentrations of GccF**
The growth of equal amounts of *Lb. plantarum* ATCC 8014 treated with different concentrations of $GccF_{Native}$ was monitored in a 96-well plate over 900 minutes (15 hours). Readings were taken every 15 minutes. Cultures were prepared in triplicate; error bars show standard error of mean (SEM) for each triplicate at each time point.

nature means that cells eventually recover from even high concentrations[145]. In fact, when the growth of cells treated with glycocin F is monitored, the resulting growth curves appear to be almost 'delayed', with almost normal growth rates occurring once bacteriostasis has lifted (Figure 3.2). For this reason, all activities of glycocin F and synthetic analogues are reported as IC_{50} (the concentration that results in 50% growth inhibition) values. A description of the determination of these values is as follows.

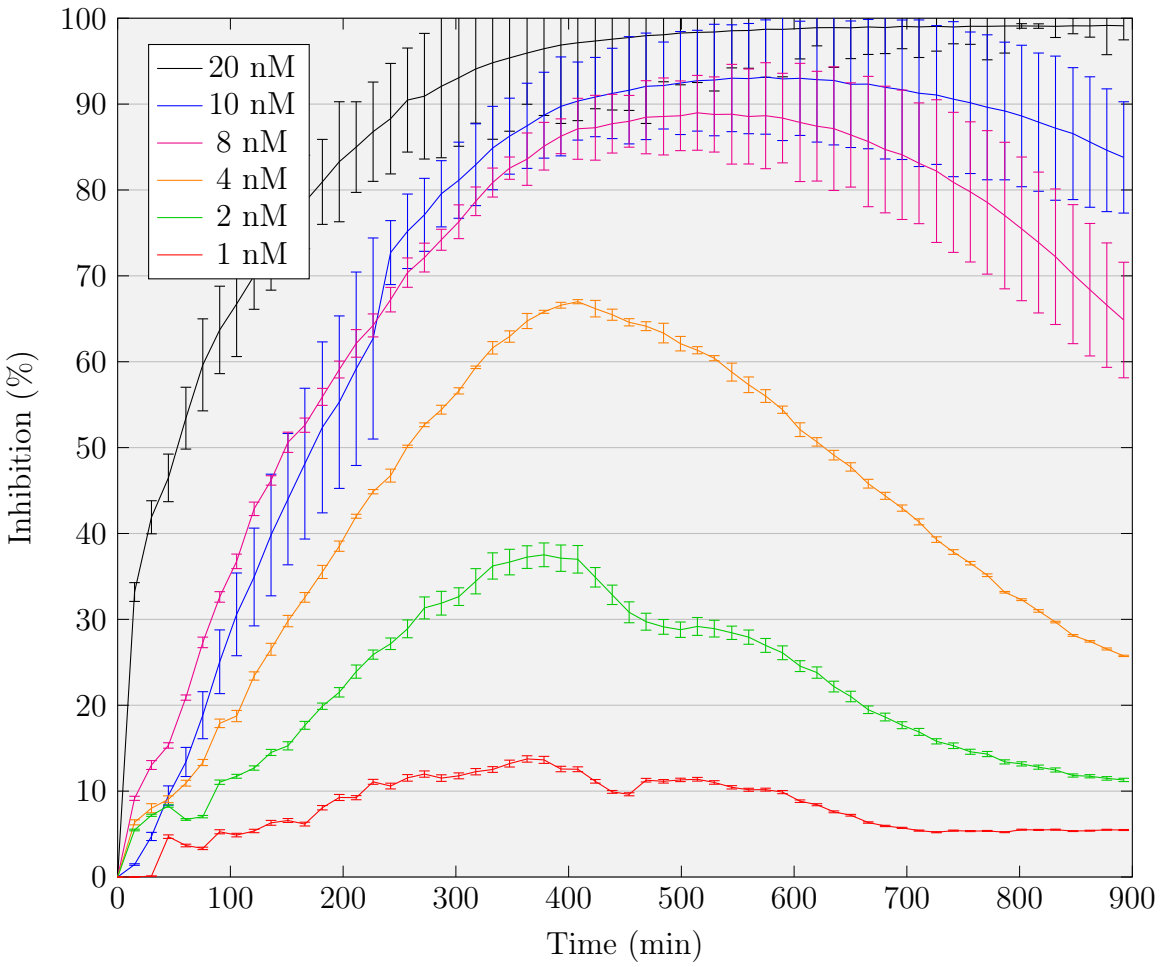


Figure 3.3: **Inhibition plots of GccF**
Inhibition plots of $GccF_{Native}$ against *Lb. plantarum* ATCC 8014 were generated from the growth curves of Figure 3.2. Error bars represent the standard error of mean (SEM) at each time point.

A concentration range of the GccF analogue to be tested was prepared in MRS media (Merck), and a 96-well microtitre plate was prepared by mixing equal volumes of diluted analogue with *Lb. plantarum* ATCC 8014 indicator strain, diluted to an

OD_{600} of approximately 0.05 (see Section 3.2.2.1). All dilutions were prepared in triplicate. Cells were grown at 30 ℃ for 15 hours, with OD_{600} readings taken every 15 minutes. When the growth curves are plotted, it can be considered that the density of each 'treated' sample is at some percentage of the unrestricted ('untreated') sample (for example, the density of the 20 nM treatment of Figure 3.2 could be considered to be growing at approximately 0% the density of the untreated samples at each reading). Taken inversely, this could be represented as a percentage inhibition (*i.e.* that same '0%' growth for the 20 nM sample could be considered as 100% inhibition at each reading). Therefore, the percentage inhibition was calculated for each time point (see following equation), and inhibition curves plotted (Figure 3.3).

$$Inhibition = (1 - \frac{OD_{treated}}{OD_{untreated}}) \times 100\%$$

The peaks of inhibition for each GccF concentration was then plotted against the concentration, and the concentration at which 50% inhibition occurs interpolated (Figure 3.4). Regarding the calculation of error, all of the standard error values were calculated across all OD_{600} measurements in a run (accounting for each triplicate) and the largest overall percentage error was associated to the resulting IC_{50} in a highly conservative approach.

3.3.3 Chemically synthesised GccF analogue activities

The synthetic glycocin F analogues that were tested were divided into four main groups based on their modifications: analogues with modified GlcNAc linkages (studying the effect of replacing *O*-GlcNAc with *S*-GlcNAc and *vice versa*), analogues with modified C-terminal tails (decreasing the length, adding a positive charge, and removing or substituting the C-terminal GlcNAc), analogues with modified interhelical loop regions (decreasing size, adding in steric modifiers), and analogues with single disulfide bonds broken. The effect of each modification on the biological activity of GccF was monitored by determining the IC_{50} values of each analogue (see Section 3.2.2.1) and comparing it to the synthetic, unmodified GccF ($GccF_{Syn}$) which has been shown to have an IC_{50} value of 1.48 ± 0.20 nM. The activities of all GccF analogues are summarised in Table 3.2.

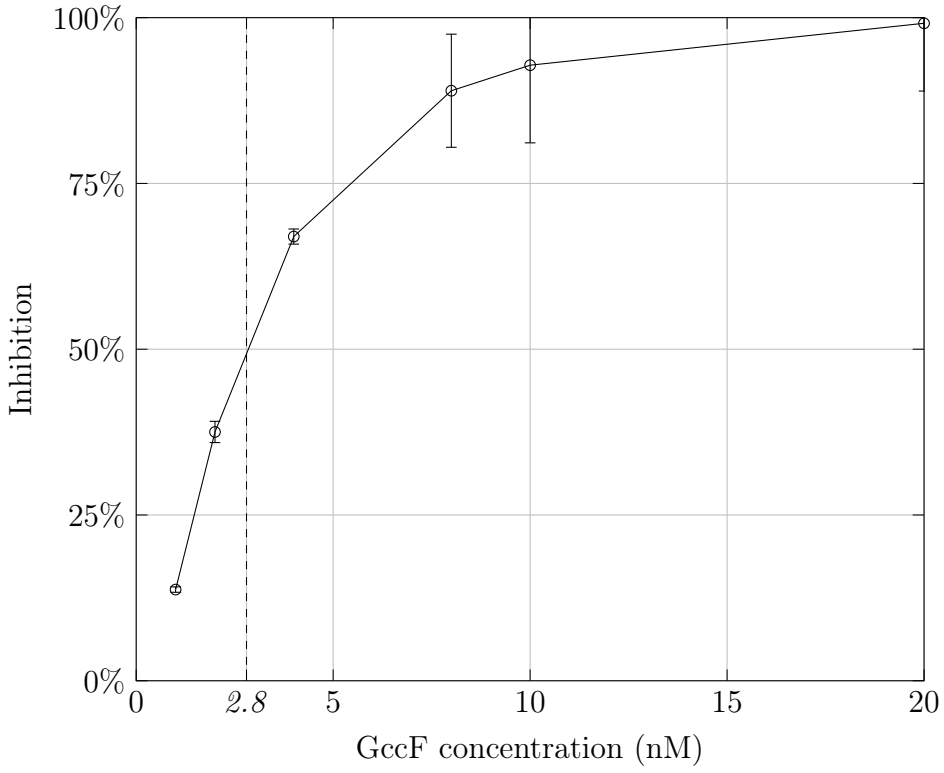


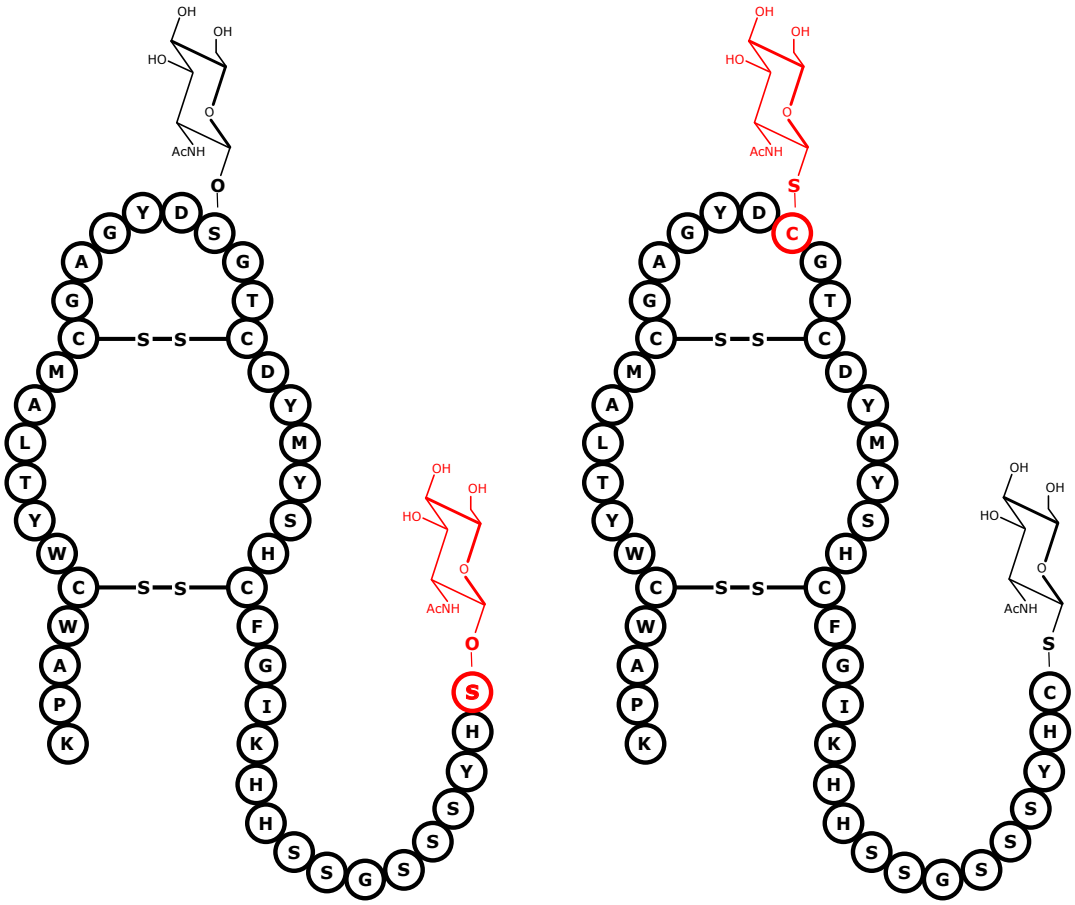
Figure 3.4: **Inhibition vs. concentration for GccF$_{Native}$**
Maximum inhibition peaks from Figure 3.3 were plotted against GccF$_{Native}$ concentration. Error bars represent the maximum standard error generated for each concentration range. Reading from the curve results in an IC_{50} value of 2.8 nM.

Table 3.2: Activity of synthetic and enzymatically prepared GccF derivatives

Modification	**Analogue**	**Description**	**IC_{50} (nM)**
Unmodified	$GccF_{Native}$	Native peptide purified from *Lb. plantarum* KW30	2.8 ± 0.3
	$GccF_{Syn}$	Chemically synthesised, unmodified GccF	1.48 ± 0.20
Glycosidic linkage	$GccF_{S,S}$	Cys43-GlcNAc replaced with Ser43-GlcNAc	12.1 ± 2.4
	$GccF_{C,C}$	Ser18-GlcNAc replaced with Cys18-GlcNAc	0.60 ± 0.1
C-terminal tail	$GccF_{Cys43Ser\Delta GlcNAc}$	Cys43-GlcNAc replaced with unmodified Ser	107 ± 21
	$GccF_{\Delta Ser38-40}$	Ser38, Ser39, and Ser40 removed	50 ± 8.4
	$GccF_{Cys43Man}$	Cys43-GlcNAc replaced with Cys43-Mannose	1015 ± 203
	$GccF_{Cys43Glc}$	Cys43-GlcNAc replaced with Cys43-Glucose	661 ± 132
	$GccF_{Ser38Lys}$	Ser38 replaced with Lys	1.71 ± 0.19
	$GccF_{1-32\Delta GlcNAc}$†	Residues 33 - 43 enzymatically removed with trypsin	∼ 400
	$GccF_{1-41\Delta GlcNAc}$†	Residues 42 - 43 enzymatically removed with chymotrypsin	∼ 150
Interhelical loop	$GccF_{\Delta GlyI}$	Gly13 removed from loop	57.6 ± 6.8
	$GccF_{\Delta GlyII}$	Gly13 and Gly15 removed from loop	2480 ± 540
	$GccF_{\Delta GlyIII}$	Gly13, Gly15, and Gly19 removed from loop	2710 ± 370
	$GccF_{Ser18MethGlcNAc}$	Methyl group added to α-carbon of Ser18	1510 ± 250
	$GccF_{Ser18\Delta GlcNAc}$†	Ser18-GlcNAc removed using *N*-acetyl-β-D-glucosaminidase	No activity
	$GccF_{Ser18Glc}$	Ser18-GlcNAc replaced with Ser18-glucose and Cys43-GlcNAc replaced with Ser	No activity
	$GccF_{Ser18Ser43Glc}$	Ser18-GlcNAc and Cys43-GlcNAc both replaced with Ser-glucose	No activity
Disulfide bonds	$GccF_{Cys12/21}$	Disulfide bridge between Cys12 and Cys21 disrupted	No activity
	$GccF_{Cys5/28}$	Disulfide bridge between Cys5 and Cys28 disrupted	8560 ± 1710

† Enzymatically-prepared GccF analogues (Stepper *et al.* (2011)[145]).

3.3.3.1 GlcNAc linkages



(a) GccF$_{S,S}$ (b) GccF$_{C,C}$

Figure 3.5: **Structures of sugar linkage modified GccF analogues**
Modifications are highlighted in red. (a) GccF$_{S,S}$ contains a serine substituting a cysteine at position 43, resulting in both GlcNAcs being attached *via* *O*-linkages. (b) GccF$_{C,C}$ contains a cysteine substituting at serine at position 18, resulting in both GlcNAcs being attached *via* *S*-linkages.

Two analogues of GccF were synthesised which contained substitutions to the GlcNAc-bearing residues. GccF$_{S,S}$ had Cys43 substituted with a Ser, resulting in both GlcNAcs being *O*-linked, while GccF$_{C,C}$ had Ser18 substituted with Cys, resulting in both GlcNAcs being *S*-linked (Figure 3.5). These modifications were chosen to test the different role of the specific linkages, as GccF is the only glycocin characterised to date containing both an *O*-linked and *S*-linked sugar[119].

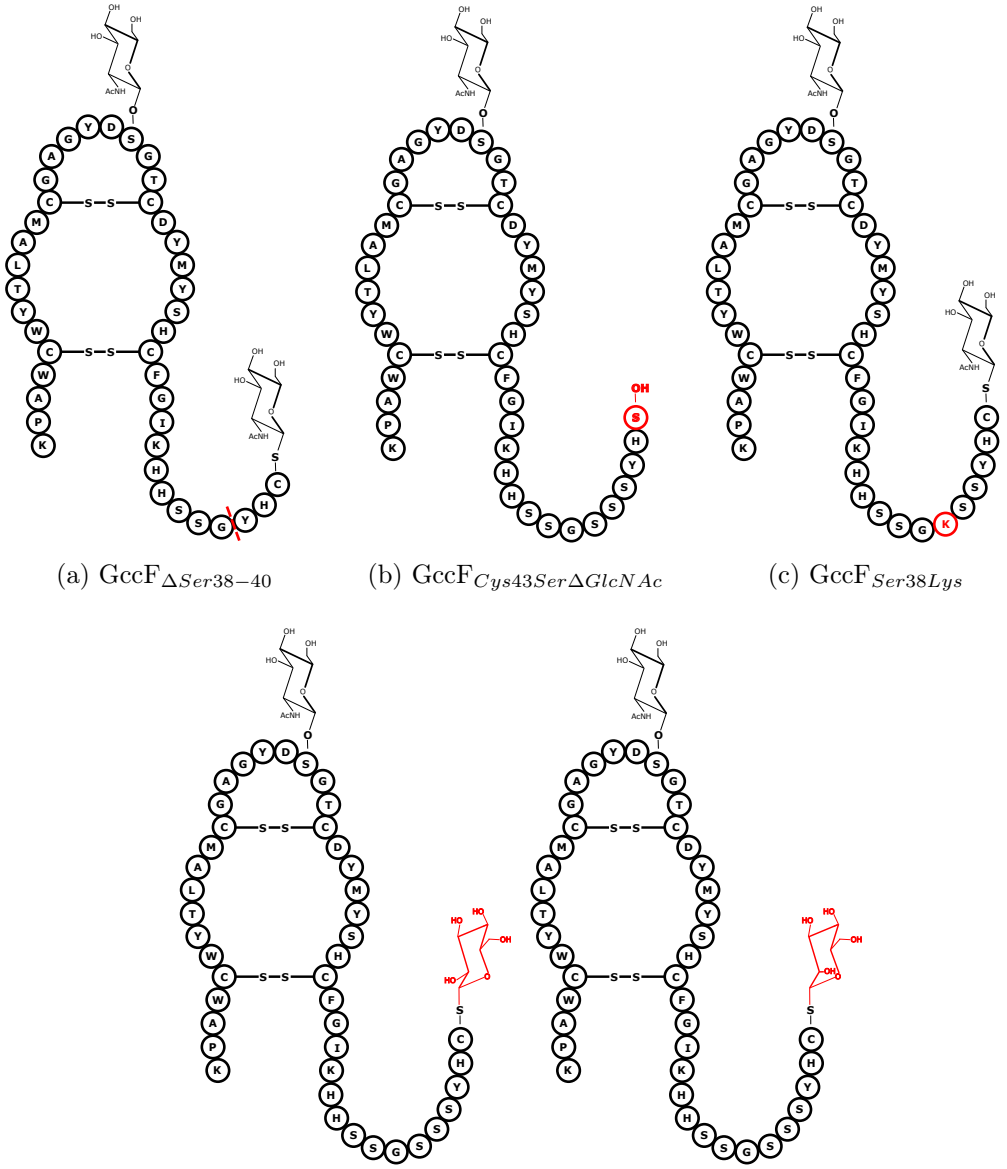
GccF$_{C,C}$ was approximately twice as active as unmodified GccF, with an IC$_{50}$ value

of 0.60 ± 0.10 nM compared to 1.48 ± 0.20 nM for GccF$_{Syn}$ (Table 3.2), making it approximately 2-fold more active. Conversely, GccF$_{S,S}$ was approximately 10-fold less active than GccF$_{Syn}$, with an IC$_{50}$ of 12.1 ± 2.4 nM. Stepper *et al.* had already shown that the GlcNAc at position 18 is vital for GccF's activity, as removal of this sugar by *N*-acetyl-β-D-glucosaminidase completely abolished activity (GccF$_{Ser18\Delta GlcNAc}$; Table 3.2)[145]. The substitution of Ser18 with Cys results in increased activity, indicating that the particular glycolinkage at this position does not itself play a role in activity.

Many bacteria secrete a variety of glycosidases into the environment to break down polysaccharide chains into nutrients, or into building blocks for the cell wall to facilitate growth[46,26]. It has been shown that human *O*-GlcNAcase is incapable of cleaving the cysteine-GlcNAc glycosidic bond[54]. The tail-linked GlcNAc is not critical for activity, as shown by residual antibacterial activity following tryptic and chymotryptic cleavage of the tail (GccF$_{1-32\Delta GlcNAc}$ and GccF$_{1-41\Delta GlcNAc}$; Table 3.2)[145], although its presence does increase activity. Presumably the sugar at the tail position would be easily accessible to glycosidases in the bacterial cell envelope, which would result in a number of GccF molecules losing the C-terminal GlcNAc when it is linked to a serine, hence the decreased activity for GccF$_{S,S}$. However, although the peptide-glycosidic bond at position 18 is probably less accessible to enzymatic attack, it will still be cleaved to a lesser degree. Substitution of the serine for cysteine and an *S*-linked GlcNAc would be less susceptible to enzymatic cleavage, resulting in the increased activity for GccF$_{C,C}$.

3.3.3.2 C-terminal tail

Five different GccF analogues featuring modifications to the C-terminal tail region were designed to test a number of hypotheses. Analogues GccF$_{\Delta Ser38-40}$ and GccF$_{Cys43Ser\Delta GlcNAc}$ (Figure 3.6 (a) and (b)) were designed to test one model of activity, in which both the loop- and tail-bound GlcNAcs bind to each sugar binding site of the homodimeric membrane-bound EIIC of the GlcNAc-PTS transporter[137]. GccF$_{Ser38Lys}$ (Figure 3.6 (c)) was designed to test two observations: firstly, does the inclusion of an additional positive charge increase the effectiveness of glycocin F, considering some class II bacteriocins have been speculated to function through cationic interactions with the more anionic bacterial membrane[15]? And secondly, was the decrease in activity seen with the original synthetic GccF due to the presence



(a) GccF$_{\Delta Ser38-40}$ (b) GccF$_{Cys43Ser\Delta GlcNAc}$ (c) GccF$_{Ser38Lys}$

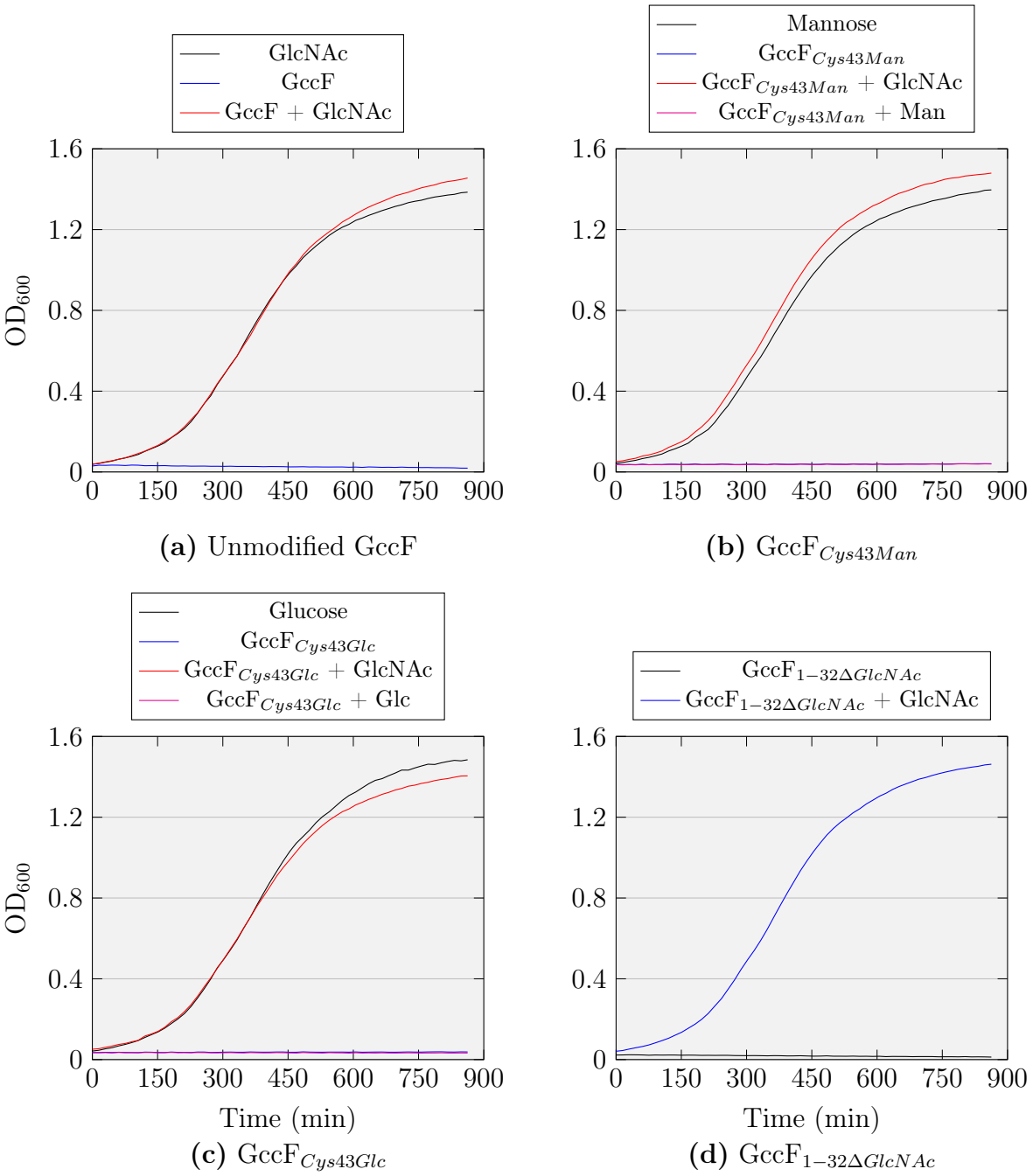
(d) GccF$_{Cys43Glc}$ (e) GccF$_{Cys43Man}$

Figure 3.6: **Structures of C-terminal tail-modified analogues**
Modifications are highlighted in red. (a) GccF$_{\Delta Ser38-40}$ has serines 38, 39, and 40 removed, resulting in a shorter tail structure. (b) GccF$_{Cys43Ser\Delta GlcNAc}$ has cysteine at position 43 substituted with a serine, and no GlcNAc attached. A serine was chosen over an unmodified cysteine to prevent possible disulfide interactions with the other four cysteine residues in the sequence. (c) GccF$_{Ser38Lys}$ has serine at position 43 replaced with lysine. (d) and (e) have the *S*-linked GlcNAcs at position 43 replaced with glucose (GccF$_{Cys43Glc}$) and mannose (GccF$_{Cys43Man}$), respectively.

of an amidated C-terminus[30]? Analogues GccF$_{Cys43Glc}$ and GccF$_{Cys43Man}$ (Figure 3.6 (d) and (e)), which contain the C-terminal GlcNAcs substituted with glucose and mannose, respectively, were created to test the role of the GlcNAc at position 43.

Both GccF$_{\Delta Ser38-40}$ and GccF$_{Cys43Ser\Delta GlcNAc}$ retained activity indicating that the presence of both the tail-bound and loop-bound GlcNAcs together are not required for full activity. Serines 38 through 40 were chosen for removal as this string of relatively small, polar residues was not expected to have a significant effect on the activity of the tail, and were close enough to the C-terminus to not play a role in protein folding. GccF$_{\Delta Ser38-40}$ had an IC$_{50}$ value of 50 ± 8.4 nM, and GccF$_{Cys43Ser\Delta GlcNAc}$ had an IC$_{50}$ of 107 ± 21 nM (Table 3.2), indicating that loss of the C-terminal GlcNAc is more detrimental than decreasing the tail length by three residues. The activity of both these analogues rules out the mechanism where both GlcNAcs must be bound to the sugar binding sites of the PTS, as GccF containing only a Ser18-linked GlcNAc still displayed activity. This is consistent with the observation by Stepper *et al.* (2011) that the loop-bound GlcNAc is the only sugar required for activity[145]. In particular, the difference in IC$_{50}$ values between GccF$_{\Delta Ser38-40}$ and GccF$_{Cys43Ser\Delta GlcNAc}$ suggests that the tail-bound GlcNAc may have more of a role in directing or localising the peptide, rather than contributing to activity directly. In fact, the IC$_{50}$ value for GccF$_{Cys43Ser\Delta GlcNAc}$ echoes a similar experiment carried out by Stepper *et al.*, where His42 and Cys43-GlcNAc were removed enzymatically using chymotrypsin[145]. The activity of chymotrypsin-digested GccF was approximately 150 nM (GccF$_{1-41\Delta GlcNAc}$, Table 3.2), compared to 107 ± 21 nM for GccF$_{Cys43Ser\Delta GlcNAc}$, indicating that the histidine residue at position 42, which would have a positive charge at pH 4-5, might also play a role in target recognition/ binding.

The original synthetic GccF produced by Brimble *et al.* in 2015 had an amidated C-terminus, and appeared to be half as active as wild-type GccF[30]. To test whether the inclusion of a positive charge in general would lead to a decrease in activity, serine 38 was substituted with a lysine, which is positively charged at neutral and acidic pHs (GccF$_{Ser38Lys}$; Figure 3.6 (c)). This position was chosen because it occurs within a string of neutral, relatively small residues. The substitution of lysine at position 38 did not have a significant effect, however, simply increasing the IC$_{50}$ to 1.71 ± 0.19 nM (Table 3.2).



(a) Unmodified GccF

(b) GccF$_{Cys43Man}$

(c) GccF$_{Cys43Glc}$

(d) GccF$_{1-32\Delta GlcNAc}$

Figure 3.7: **Effect of different sugars on GccF and GccF analogue activity**

Figure 3.7: *Lb. plantarum* ATCC 8014 cells were grown in the presence of 10x the IC_{50} value of (a) wild-type GccF, (b) $GccF_{Cys43Man}$, (c) $GccF_{Cys43Glc}$, or (d) $GccF_{1-32\Delta GlcNAc}$, supplemented with either GlcNAc, the C-terminal linked sugar, or nothing, as indicated by figure legends. Analogue concentrations: 20 nM wild-type GccF, 10 μM $GccF_{Cys43Man}$, 6 μM $GccF_{Cys43Glc}$, and 3.5 μM $GccF_{1-32\Delta GlcNAc}$. All sugars were added to a final concentration of 1 mM. From the growth curves, it can be seen that GlcNAc always provides a protective effect, regardless of the identity of the C-terminal sugar. The average of triplicate readings are shown. In panel (b), the blue line is overlapped by the magenta line. *Figure reproduced in Bisset* et al. *(2018)*[24].

The results from $GccF_{\Delta Ser38-40}$ and $GccF_{Cys43Ser\Delta GlcNAc}$ indicated that the Cys43-linked GlcNAc is not strictly required for activity. This led to the question of the exact role of the C-terminal GlcNAc, with one possibility being that the cysteine-linked sugar acted as a 'hook' and localised the bacteriocin to a specific region or receptor, and increasing the rate/ likelihood of the Ser18-linked GlcNAc binding to the main receptor. To test this, the Cys43 GlcNAc was substituted with either Cys43 glucose or Cys43 mannose ($GccF_{Cys43Glc}$ and $GccF_{Cys43Man}$, respectively; Figure 3.6 (d) and (e)). If the C-terminal GlcNAc was not playing a role in localising, then both analogues $GccF_{Cys43Glc}$ and $GccF_{Cys43Man}$ were expected to exhibit IC_{50} values close to that of $GccF_{Cys43Ser\Delta GlcNAc}$ (*i.e.* around 100 nM). However, the activity of these analogues was lower than this, with $GccF_{Cys43Glc}$ exhibiting an IC_{50} of 661 $\pm$ 132 nM, and $GccF_{Cys43Man}$ exhibiting an IC_{50} of 1015 $\pm$ 203 nM (Table 3.2). There are two critical and surprising observations from these activities: first, the substitution of the C-terminal GlcNAc with either a glucose or a mannose results in a lower activity than simply removing the GlcNAc; and second, the identity of the substituted sugar itself affects activity, with the mannose substitution decreasing the activity almost 2-fold more than the glucose substitution.

To determine if the presence of glucose or mannose at position 43 contributes directly to the activity of $GccF_{Cys43Glc}$ or $GccF_{Cys43Man}$, *Lb. plantarum* ATCC 8014 was incubated in the presence of excess (10x IC_{50} values) $GccF_{Native}$ (wild-type GccF), $GccF_{Cys43Glc}$, $GccF_{Cys43Man}$, or $GccF_{1-32\Delta GlcNAc}$ (trypsin-digested GccF), with additional GlcNAc, mannose or glucose (Figure 3.7), to assess the protective effect of these different sugars. From the growth curves it is apparent that only GlcNAc provides a protective effect on each analogue tested, regardless of the presence or identity of the C-terminal sugar. Additionally, an excess of glucose or mannose did not provide any protective effect when also placed in the presence of $GccF_{Cys43Glc}$

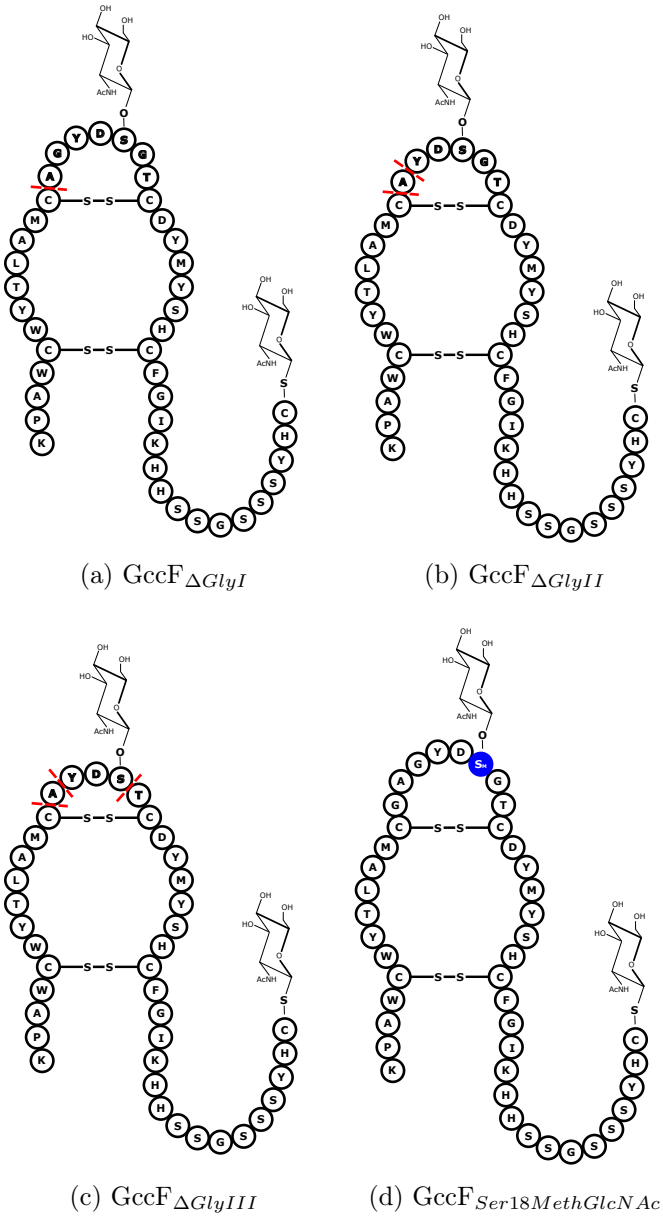
or GccF$_{Cys43Man}$, respectively. This confirms that the C-terminal sugar is not involved directly in the bacteriostatic mechanism. However, the increased reduction in activity that occurs with the substitution of the C-terminal GlcNAc compared to the removal of the sugar is still unexplained.

3.3.3.3 Interhelical loop

An NMR study of GccF, which contains 12 superimposed calculated structures, suggested that the glycosylated, interhelical loop exhibits limited flexibility[159]. To test whether the size of the loop was involved in the bacteriostatic mechanism, four analogues were designed: three contained a successively smaller number of residues achieved by the removal of one, two, or three glycines from the loop (GccF$_{\Delta GlyI}$, GccF$_{\Delta GlyII}$, and GccF$_{\Delta GlyIII}$; Figure 3.8 (a), (b) and (c), respectively). A fourth analogue contained the correct number of residues in the loop, but had the α-proton of Ser18 replaced with a methyl group (GccF$_{Ser18MethGlcNAc}$, Figure 3.8 (d)). Although the stereochemistry around the α-carbon of Ser18 of GccF$_{Ser18MethGlcNAc}$ is unchanged, the presence of the methyl group is designed to restrict the conformations sampled by the GlcNAc, and hence the serine sidechain. The activities of the interhelical loop-modified analogues is given in Table 3.2.

The removal of one glycine (Gly13) from the interhelical loop resulted in an IC$_{50}$ of 57.6 $\pm$ 6.8 nM (GccF$_{\Delta GlyI}$; Table 3.2), which is a decrease similar to that seen when the C-terminal tail was shortened by three residues. However, the removal of two or three glycines from the loop region drastically reduced the IC$_{50}$ of GccF, resulting in IC$_{50}$ values of 2480 $\pm$ 540 and 2710 $\pm$ 370 nM for GccF$_{\Delta GlyII}$ and GccF$_{\Delta GlyIII}$, respectively (Table 3.2). Additionally, the CD spectra of all three ΔGly analogues indicated a reduced α-helical content, suggesting that the shorter interhelical loops are placing strain on the structure, and as a result affecting the folding of the peptide (Figure 3.1 (b)).

The addition of a methyl group to the α-carbon of Ser18 had an unexpectedly strong effect on GccF activity, resulting in an IC$_{50}$ of 1510 $\pm$ 250 nM (GccF$_{Ser18MethGlcNAc}$; Table 3.2). The reduction in activity following placement of this methyl group on the α-carbon of the GlcNAc-carrying Ser18 suggests there is an important relationship between the sugar and the conformation of the amino acids surrounding it. Taken together with the reduced activity of GccF$_{\Delta GlyII}$ and GccF$_{\Delta GlyIII}$, it is clear that the recognition of GccF by its receptor does not only rely on the orientation

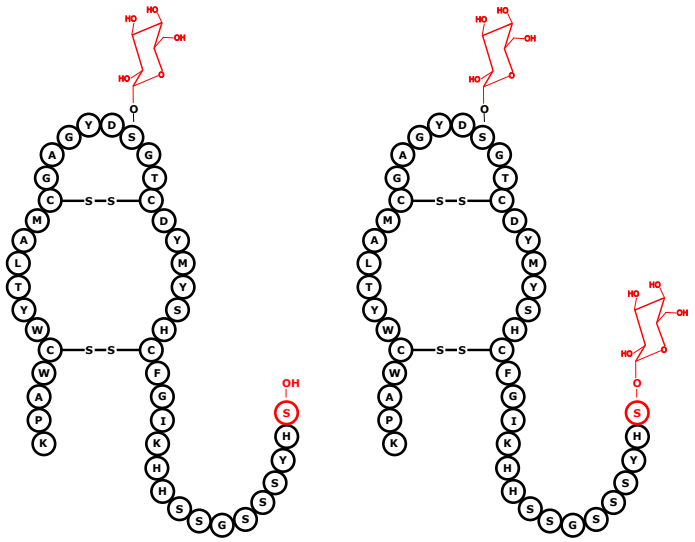


(a) GccF$_{\Delta GlyI}$ (b) GccF$_{\Delta GlyII}$

(c) GccF$_{\Delta GlyIII}$ (d) GccF$_{Ser18MethGlcNAc}$

Figure 3.8: **Structures of loop-modified analogues**
(a) A single glycine (Gly13) has been removed from the interhelical loop. (b) Both Gly13 and Gly15 have been removed from the loop. (c) Gly13, Gly15, and Gly19 have all been removed from the loop. (d) The serine at position 18 replaced with a variant containing a methyl group on the α-carbon, highlighted in blue. Red dashed lines indicate the positions from which glycines were removed from the peptide backbone.

of the GlcNAc, but also of the loop itself. Interestingly, the CD suggested that GccF$_{Ser18MethGlcNAc}$ has a slightly reduced alpha-helical content (Figure 3.1 (b)). However, owing to the sensitivity of the CD signal in this region, this could be due to a difference in concentration, even though great care was taken to measure the concentrations accurately.



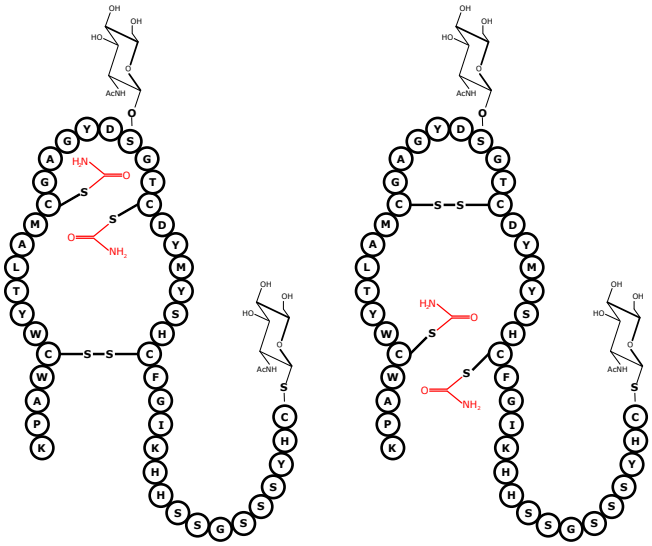
(a) GccF$_{Ser18Glc}$ (b) GccF$_{Ser18Ser43Glc}$

Figure 3.9: **Structures of loop-substituted glucose analogues**
(a) Ser18-GlcNAc replaced by glucose, Cys43-GlcNAc replaced by unmodified serine. (b) Ser18-GlcNAc and Cys43-GlcNAc both replaced by Ser-glucose.

To see if GccF could be engineered to target new cells, or exhibit a new effect on already susceptible cells (such as being bactericidal), two additional variants were designed and synthesised: GccF$_{Ser18Glc}$, which has the GlcNAc at position 18 substituted with a glucose, and Cys43-GlcNAc replaced with a Ser; and GccF$_{Ser18Ser43Glc}$, which contained serine-linked glucoses at both positions 18 and 43 (Figure 3.9). However, both compounds proved to be completely inactive when tested against indicator strain *L. plantarum* ATCC 8014 in MRS and minimal media solid plate tests. Additionally, the compounds were both inactive against *Enterococcus faecium* 64/3 (which has previously been found to be highly susceptible to wild-type GccF; data not shown) when tested on LB plates. Finally, to see if the substitution of GlcNAc with glucose allowed GccF to target new bacteria (such as those susceptible to sublancin), *Bacillus subtilis* 6633 was tested, which has previously been shown to be resistant to wild-type GccF[92]. However, the new analogues were also inactive

against this strain. Taken together, this new data indicates that the activity of GccF is sugar-specific, and suggests that the GlcNAc at position 18 is important for GccF's target interaction, and is not merely protecting the peptide from proteolytic cleavage as previously suggested[145].

3.3.3.4 Disulfide bonds



(a) $GccF_{Cys12/21}$ (b) $GccF_{Cys5/28}$

Figure 3.10: **Structure of disulfide-disrupted analogues**
Two analogues were prepared with disulfide bridges capped by iodoacetamide, shown in red. (a) Disulfide bridge between cysteines 12 and 21 are capped. (b) Disulfide bridge between cysteines 5 and 28 are capped.

The final two GccF derivatives synthesised were designed to test the contribution of the two disulfide bonds to GccF activity. The disulfide bonds are present between Cys5 and Cys28, and Cys12 and Cys21 and appear to hold the two alpha-helices in place. Previous work has shown that treatment of GccF with reducing agents completely destroys activity, which suggests that the spontaneous removal of both disulfide bridges renders GccF inactive. However, outside of controlled chemical synthesis it has been impossible to study the effect of reducing one disulfide bond in isolation. The analogues $GccF_{Cys12/21}$ and $GccF_{Cys5/28}$ were synthesised with the cysteines 12 and 21 or 5 and 28 capped with iodoacetamide, respectively (Figure 3.10).

GccF$_{Cys5/28}$, which has the disulfide bridge furtherest from the interhelical loop disrupted, had the highest IC_{50} value of all the analogues tested (IC_{50} = 8560 ± 1710 nM; Table 3.2), thus the lowest activity for a still-active analogue. GccF$_{Cys12/21}$ on the other hand, which has the disulfide closest to the interhelical loop disrupted, was the only synthetic analogue to show complete loss of activity. To confirm this, 5 μL volumes of undiluted ($\sim$ 500 μM) GccF$_{Cys12/21}$ and GccF$_{Cys5/28}$ were spotted onto an *Lb. plantarum* ATCC 8014 indicator plate, along with 100 nM of GccF$_{Native}$ (Figure 3.11). Visible clearing can be seen by the GccF$_{Native}$ spot, faint clearing by GccF$_{Cys5/28}$, and no clearing can be seen by GccF$_{Cys12/21}$. Additionally, CD spectra of both GccF$_{Cys12/21}$ and GccF$_{Cys5/28}$ show a significant reduction in the amplitude of the peaks, as well as a shifting of the positive peak to around 190 nm, and the negative peak to $\sim$ 205 nm (Figure 3.1 (d)), suggesting a loss of alpha-helical content in both structures. These findings further confirm the importance of the conformation of the loop holding the *O*-linked GlcNAc, and point to there being more than just recognition of the GlcNAc by the primary receptor, but rather a more complex recognition involving both peptide and sugar moieties.

3.4 Conclusion

From the IC_{50} values of the different analogues, a picture of the activity of GccF begins to emerge. Figure 3.12 summarises the effect that modifications to different parts of GccF have on GccF activity. To begin with, the roles of the loop-GlcNAc and the tail-GlcNAc appear to be independent. Although modifications to the tail regions have an effect on GccF activity, no single tail mutation completely abolishes antimicrobial activity. The greatest decrease on activity from a tail modification arises from substituting Cys43-GlcNAc with Cys43-Mannose (IC_{50} of GccF$_{Cys43Man}$ = 1015 ± 203 nM), which decreases the activity approximately 10 times more than removing the C-terminal GlcNAc (IC_{50} for GccF$_{Cys43Ser\Delta GlcNAc}$ = 107 ± 21 nM; Table 3.2). The increased IC_{50} values for GccF analogues containing different sugars at the C-terminus (GccF$_{Cys43Glc}$ and GccF$_{Cys43Man}$) compared to GccF$_{Cys43Ser\Delta GlcNAc}$ suggests that the role of the tail GlcNAc is to localise the peptide to its primary target on the cell, and not to initiate the bacteriostatic effect itself. This is further supported by the fact that, aside from the homologue ASM1, no other glycocins contain a glycosylated 'tail'[119,132].

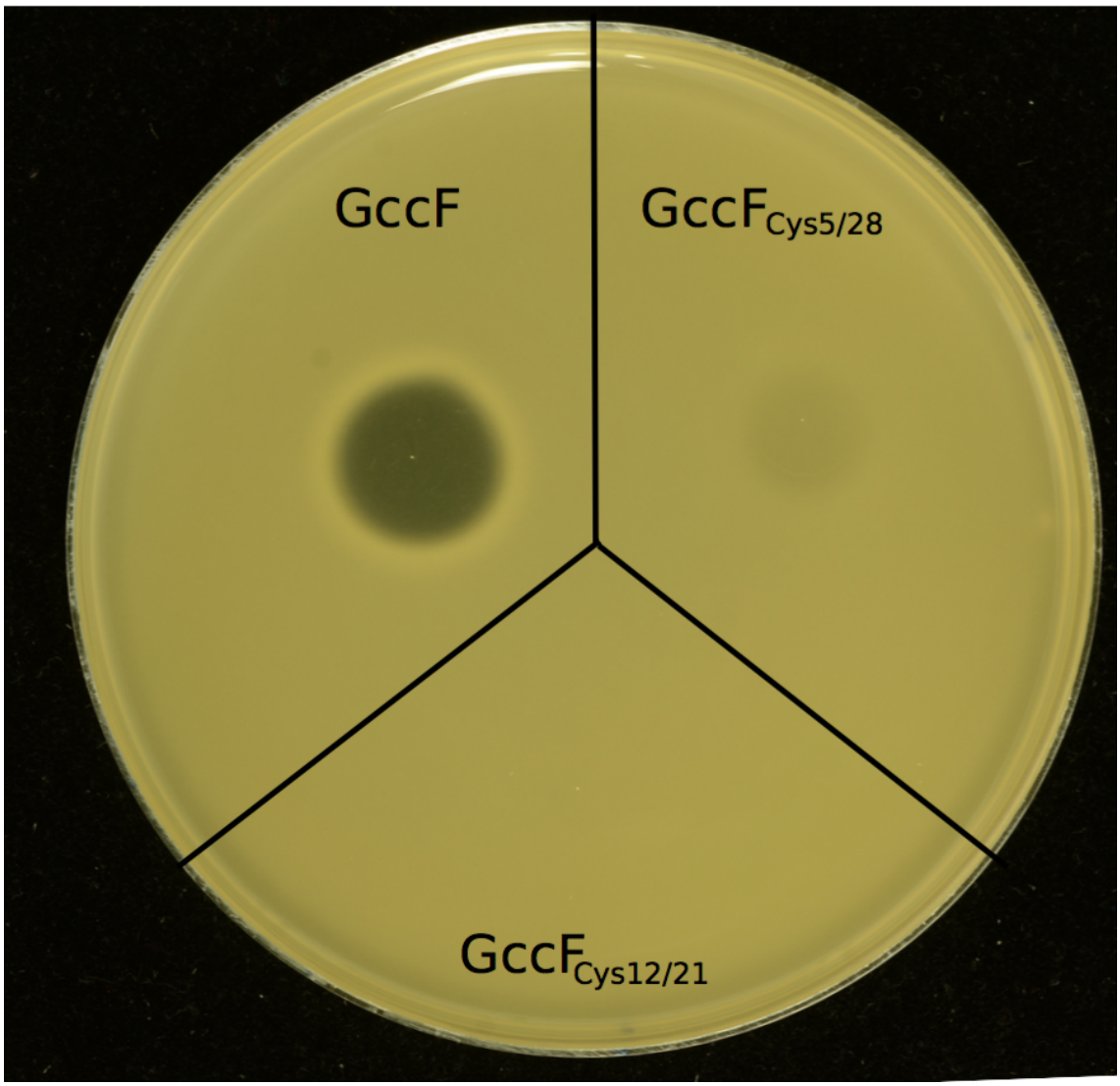


Figure 3.11: **GccF$_{Cys12/21}$ and GccF$_{Cys5/28}$ indicator plate**
500 μL volumes of GccF$_{Native}$, GccF$_{Cys12/21}$, and GccF$_{Cys5/28}$ were spotted on an *Lb. plantarum* ATCC 8014 indicator plate. Concentration of GccF$_{Native}$ spot was 100 nM; concentration of GccF$_{Cys12/21}$ and GccF$_{Cys5/28}$ were both $\sim$ 500 μM. Lack of clearing by GccF$_{12/21}$ indicates complete inactivity of this analogue. *Figure reproduced in Bisset* et al. *(2018)*[24].

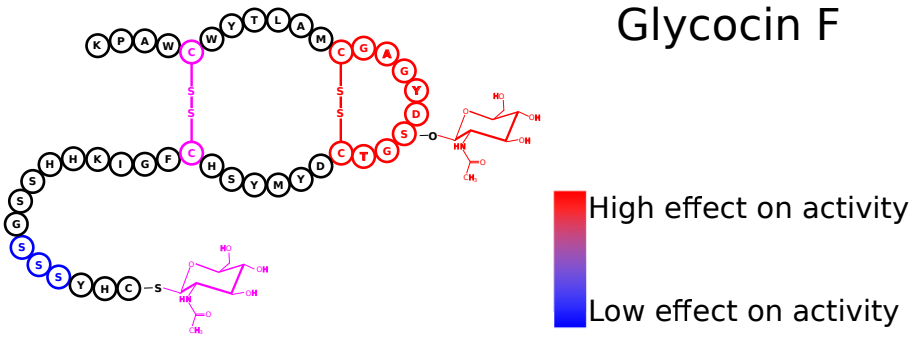


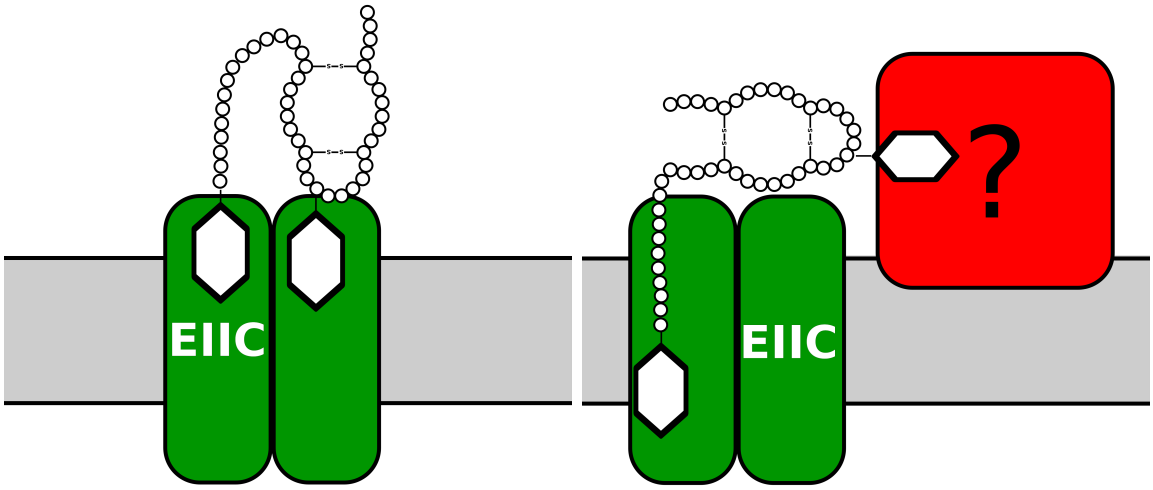
Figure 3.12: **Summary of the active sites of GccF**
Regions of glycocin F are coloured based on the effect modifications to the area had on activity. Regions coloured in red had high effects on activity (high increases in IC_{50}, or complete abolishment of activity), regions in blue had low effects (small changes in IC_{50}), and regions in magenta had intermediate or variable effects. *Figure reproduced in Bisset* et al. *(2018)*[24].

The modifications with the biggest effect on GccF activity (with the exception of the disulfide bridge modifications) were located on the interhelical loop. This is not surprising, as previous work has shown that removal of the Ser18-linked GlcNAc completely abolished activity[145]. Interestingly, the substitution of Ser18-GlcNAc with Ser18-Glc also abolished removed activity (Table 3.2). This indicates that not only is the glycosylation at position 18 of the peptide vital for activity, but the identity of the sugar is also crucial. One suggested role of the Ser18-GlcNAc was to protect the loop peptide from proteolytic cleavage[145], which would presumably work regardless of the specific glycan at that site. However, the inactivity of the Ser18-Glc analogues indicates that the loop peptide, along with the specific sugar, is required for interaction with GccF's primary target or receptor. This aligns with recent work done by Wu *et al.* on the significance of the glycan on sublancin. When the glucose at position 22 is substituted for a mannose or a GlcNAc, the bacteriocin is not rendered completely inactive, but there is a significant loss of activity (50 % and 75 % loss of activity, respectively)[168]. Additionally, the strong decrease in activity seen with the substitution of a methyl group for the alpha hydrogen of Ser18 (IC_{50} for $GccF_{Ser18MethGlcNAc}$ = 1510 ± 250 nM; Table 3.2) provides extra evidence that the orientation of the GlcNAc on the loop is also important for target/ receptor interaction.

A common feature of glycocins is the two nested disulfide bonds, and two thioredoxin-like proteins within the glycocin operon[119,132,60]. Treatment with reducing agents

typically inactivates glycocins, suggesting that the disulfide bonds are required for activity[145,132,119]. For the first time, we have been able to study the effect of removing one of the two disulfide bonds of GccF in isolation. The loss of either disulfide bond resulted in a loss of alpha-helical structure (Figure 3.1 (d)), and the loss of either bond had a significantly detrimental effect on activity. The loss of the bond closest to the glycosylated loop ($GccF_{Cys12/21}$) resulted in no activity, while the loss of the bond furtherest from the loop ($GccF_{Cys5/28}$) resulted in the analogue with the highest IC_{50} value while still remaining active (IC_{50} = 8560 ± 1710 nM). These results suggest that for the activity of GccF, the conformation and stability of the interhelical loop (and probably the shape of the alpha helices) is vital to the glycocin-target interaction. Similar to how the loss of three glycines from the interhelical loop would distort the peptide and alter the presentation of the GlcNAc, the loss of the Cys12-Cys21 bridge would destabilise the loop structure. Additionally, the loss of the Cys5-Cys28 bridge would allow the peptide to 'pivot' around the remaining Cys12-Cys21 bridge, increasing the flexibility of the glycosylated loop. Considering the strong degree of structural similarities among the glycocins, this may suggest a common glycan presentation method across this group of bacteriocins.

With the above evidence in mind, particularly concerning the apparent independent roles of the 'tail' and 'loop' GlcNAcs, a rough mechanism of action can be postulated. It seems most probable that the glycosylated tail of GccF plays a role in localising the peptide to a region of the cell surface through interaction with a surface receptor. Previous work on resistant mutants of GccF suggested that the GlcNAc-PTS is a target[62] (as well as additional transcriptomic evidence; see next Chapter), and work on sublancin has also highlighted the glucose-specific PTS as being of some importance[72]. Thus, with no current evidence of other transporters playing a role in GccF susceptibility, it is likely that the GlcNAc-PTS is recognising and binding the tail GlcNAc (Figure 3.13). Following this, with the peptide tethered to the cell, the glycosylated loop is brought closer to the primary target of GccF, and through interaction with this target, bacteriostasis is initiated. However, at this stage it is still uncertain whether the target interacting with the loop-GlcNAc is also the GlcNAc-PTS (Figure 3.13 (a)) or a second, as yet unidentified target, which is co-localised to the PTS (Figure 3.13 (b)). To help identify a mechanism of action for the bacteriostasis, and potentially identify the primary receptor of GccF, the transcriptional response of susceptible cells to GccF was studied, and the results detailed in the next chapter.



(a) Mechanism 1 (b) Mechanism 2

Figure 3.13: **Possible models of GccF-receptor interaction**
Based on the effects of the synthetic GccF analogues, two GccF mechanisms of binding are proposed. (a) Both GlcNAc moieties of GccF bind to the same receptor, the GlcNAc-specific PTS EIIC domain (shown in green). (b) The tail-bound GlcNAc binds to the GlcNAc-specific PTS EIIC domain, while the loop bound GlcNAc binds to a different membrane-bound receptor (red), which may or may not completely permeate the membrane. In both models only binding by the loop-bound GlcNAc results in bacteriostasis; the tail-bound GlcNAc is involved in localising the peptide. *Figure reproduced in Bisset* et al. *(2018)*[24].

4 | Transcriptional response to glycocin F

4.1 Introduction

In order to help determine the mechanism by which glycocin F exhibits its effect on cells, RNA sequencing (RNA-seq) was carried out on bacteria treated with the bacteriocin. Initial attempts to lyse *Lactobacillus plantarum* cells in a timely manner for RNA extraction were unsuccessful, with extraction times requiring pre-treatment with cell wall hydrolysing enzymes such as lysozyme and mutanolysin. Because the need to avoid any additional treatment steps beyond the GccF treatment was thought to be important, a different bacterium to *Lb. plantarum*, which was still susceptible to GccF but easier to disrupt, was chosen. Collaborators at Otago University had successfully used *Enterococcus faecalis* strain JH2-2 to extract RNA for sequencing. As this organism was shown to be sensitive to GccF, it was decided to use *E. faecalis* JH2-2 to investigate the transcriptional response to various concentrations of GccF. *E. faecalis* is a pathogen with known antibiotic resistance, and a number of microarray and RNA-seq experiments have been carried out on it providing data for comparison[2,3,158,144,131]. Interestingly, although *E. faecalis* JH2-2 was susceptible to GccF, growth was only slowed rather than completely inhibited upon initial treatment, and cells could be rescued by the addition of GlcNAc. However, protection was only temporary, being dependent on the concentration of GlcNAc used. Furthermore, a more significant stasis occurred following GlcNAc treatment, indicating that these cells were pre-sensitised to the effects of GccF by the amino sugar. This chapter describes the results of two independent RNA-seq experiments that were carried out on *E. faecalis* JH2-2: the initial tests were carried out over three time points (0, 10, and 30 minutes) while the second set of tests using a much lower concentration of GccF were done at one time point, 30 minutes.

4.2 Experimental procedures

4.2.1 Bacterial growth conditions and RNA isolation

Revival of received bacterial stocks

Bacterial strains of *E. faecalis* JH2-2, V583, VI01 and VI40 (Table 2.1) were received as glycerol stabs, and were subsequently plated on TSB 1.5% agar plates. Single

colonies were chosen and inoculated in TSB liquid media at 30 ℃ overnight, before being made into glycerol stocks with a final concentration of 10% gycerol.

RNA isolation

3x 1 mL of an overnight culture of *E. faecalis* JH2-2 or VI01 were diluted in 100 mL of TSB media and incubated at 37 ℃ with shaking at 130 rpm. Once an OD_{600} of 0.2 was reached, each sample was split into two 50 mL cultures (biological replicates), and each of which were further incubated at 37 ℃. Once an OD_{600} of 0.5 was reached, one biological replicate from each of the triplicates was treated with either GccF or GlcNAc, and cells were incubated under identical conditions for the required amount of time.

Cells were harvested by centrifugation at 7000 x g for 15 minutes at 4 ℃, the supernatant discarded and the cell pellets resuspended in 500 μL diethyl pyrocarbonate (DEPC)-treated milliQ water. The resuspended cells were transferred to a 2 mL screw-top cryotube containing 50 μL DEPC-treated glass beads (0.1 mm diameter, Sigma). Lysis was carried out in a Hybaid Ribolyser in three 30-second shaking steps (5000 rpm) with 60-second incubation steps on ice in between. Lysed cells were then centrifuged at 13,000 x g for 30 seconds, and the clear supernatant collected and mixed with three volumes of TRIzol reagent (Ambion). Phase separation of the TRIzol/ cell lysate mixture was induced by adding 200 μL chloroform and briefly mixing the tubes by inversion, then centrifuging for 15 minutes at 11,500 x g at 4 ℃. The upper aqueous layer was collected and purified using a Direct-Zol RNA-isolation column (Zymo Research) according to the manufacturer's instructions.

The quality of the isolated RNA was checked using agarose gel electrophoresis and/or LabChip analysis (Perkin Elmer), and quantified using a NanoDrop spectrophotometer (ThermoFisher Scientific). All samples were stored at -80 ℃ before being submitted to New Zealand Genomics Limited (NZGL) for sequencing. RNA sequencing libraries were prepared by NZGL using the Illumina TruSeq Stranded RNA Sample preparation kit with RiboZero bacterial depletion (for rRNA removal). Sequencing was carried out on an Illumina HiSeq 2000 platform, generating 100 bp single-ended reads.

4.2.2 Computational processing

Transcription read trimming, alignment, and quantification

RNA-seq data was provided by NZGL as raw reads in *fastq* format. These were trimmed using using **BBDuk** (part of the BBMap compilation, Joint Genome Institute) with the following settings: kmer length of 23, hamming distance (hdist) of 1 chosen, `tbo` flag selected. Trimmed reads were aligned to reference genomes using **Bowtie2** (default parameters)[102], against either the *E. faecalis* JH2-2 or V583 reference strains from the Ensembl Bacteria database (Taxonomy IDs 1320322 and 226185, respectively). Following alignment, the reads for each gene were counted using **HTSeq**, a Python-based package which counts the number of occurrences of each aligned gene against a gene list of the reference strain[12], using the 'Union' model of counting. The 'model of counting' parameter dictates how a read is assigned in the event that it overlaps two different open-reading frames - selecting 'Union' means that any reads that overlap two or more open-reading frames are binned as 'ambiguous'.

Statistical analysis

Statistical analysis was carried out on the gene counts using statistical software R. Most of the analysis was carried out using the package ***DESeq2***[105], although ***edgeR*** was also used to validate the analyses carried out by *DESeq2*. Both packages model the transcriptional data using a negative binomial model (effectively a binomial distribution where the null hypothesis is no change in transcription), and account for a false discovery rate (FDR) using the Benjamini-Hochberg method[22].

4.3 Results and discussion

4.3.1 Response of *Enterococcus faecalis* JH2-2 to glycocin F

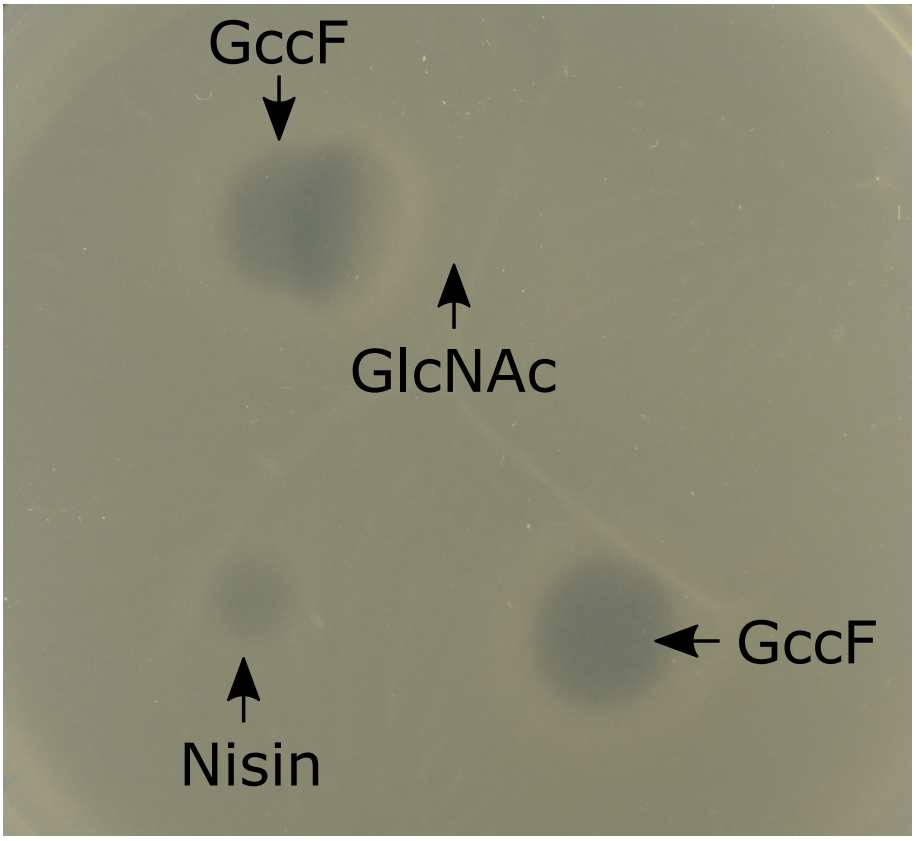


Figure 4.1: ***E. faecalis* JH2-2 GccF indicator plate**
A 1% agar TSB plate was inoculated with *E. faecalis* JH2-2, and titrated with 5 μL spots of 70 μM GccF, 1 mM GlcNAc and 25 ng/mL nisin. GccF can be seen to provide a moderate clearing effect on strain JH2-2. The interface between GccF and GlcNAc shows an enhanced degree of clearing, indicating that GlcNAc acts to pre-sensitise *E. faecalis* JH2-2 to GccF.

Enterococcus faecalis JH2-2 (Table 2.1) was a kind gift from Professor Gregory Cook, Otago University for the purpose of studying the transcriptional response of the organism to glycocin F. Before transcriptomic tests could be carried out, the cells had to be assessed for susceptibility to GccF. Indicator plates of *E. faecalis* JH2-2 were tested with 5 μL spots of 500 nM GccF and 500 μM GlcNAc, placed adjacently (Figure 4.1). From the clearing on the plate it can be seen that *E. faecalis* JH2-2 cells are susceptible to GccF. Additionally, the adjacent GlcNAc spot can be

seen to provide a protective effect against GccF, as the clearing zone stops on the border of the GccF and GlcNAc spots. Following this, susceptibility tests were repeated in liquid cultures, in order to determine an IC_{50} value for GccF against *E. faecalis* JH2-2. The method used was the same as that used to determine the IC_{50} of *Lb. plantarum* ATCC 8014 to GccF. However the *E. faecalis* cells did not grow well in 96 well plates. Therefore 1 mL cuvettes were used and growth monitored (unstirred) in a Varian 6x6 multicell Peltier in a Cary 300 Bio spectrophotometer (Varian) that had been pre-heated to 30 °C[30]. Regardless of concentration of GccF used, *E. faecalis* cells were never seen to enter complete stasis, making it impossible to determine as initial IC_{50} value (Figure 4.2).

The protective effect of GlcNAc on GccF-treated *E. faecalis* JH2-2 cells was also tested in liquid culture. An initial test showed that a final concentration of 500 μM GlcNAc was sufficient to protect JH2-2 cells from GccF for around 60 minutes for concentrations of GccF ranging from 50 nM to 500 nM (Figure 4.3 (a)). It was also observed that once this protection wore off the cells seemed to be more susceptible to the effects of GccF in that growth stopped completely when $\geq$ 250 μM GlcNAc was added (Figure 4.3 (b)). This increased susceptibility to GccF appeared to increase proportionally to the concentration of GlcNAc at which the *E. faecalis* cells were grown in. More specifically, similar degrees of stasis could be observed in cells treated with 500 μM GlcNAc and 100 nM GccF, or 500 nM GccF. Taken together, these results suggest there is synergy between the effects of GlcNAc and GccF on *E. faecalis* JH2-2 cells: firstly, the concentration of GlcNAc is directly related to the sensitivity of the cells to GccF (Figure 4.3 (a)); secondly, the extent of GlcNAc protection appears to be related to the amount of GlcNAc in the environment (inferred from Figure 4.3 (b)).

Given the protective effect of GlcNAc, the ability of GlcNAc to recover cells that are already under GccF-induced stasis was investigated. JH2-2 cells were grown in the presence of 1 mM GlcNAc and 100 nM GccF. Once GccF-induced stasis began, additional GlcNAc was added at a final concentration of 500 μM (approximately 360 minutes after measurement began), and recovery was seen almost immediately (Figure 4.4). Again, as expected, this recovery was only temporary, with increased bacteriostasis resuming presumably once the environmental GlcNAc was taken up by all the cells. The results of this experiment hint at a dynamic interplay occurring between the machinery recognising GccF, and the import of GlcNAc. In particular, it appears as though GlcNAc-mediated protection only occurs for as long as GlcNAc

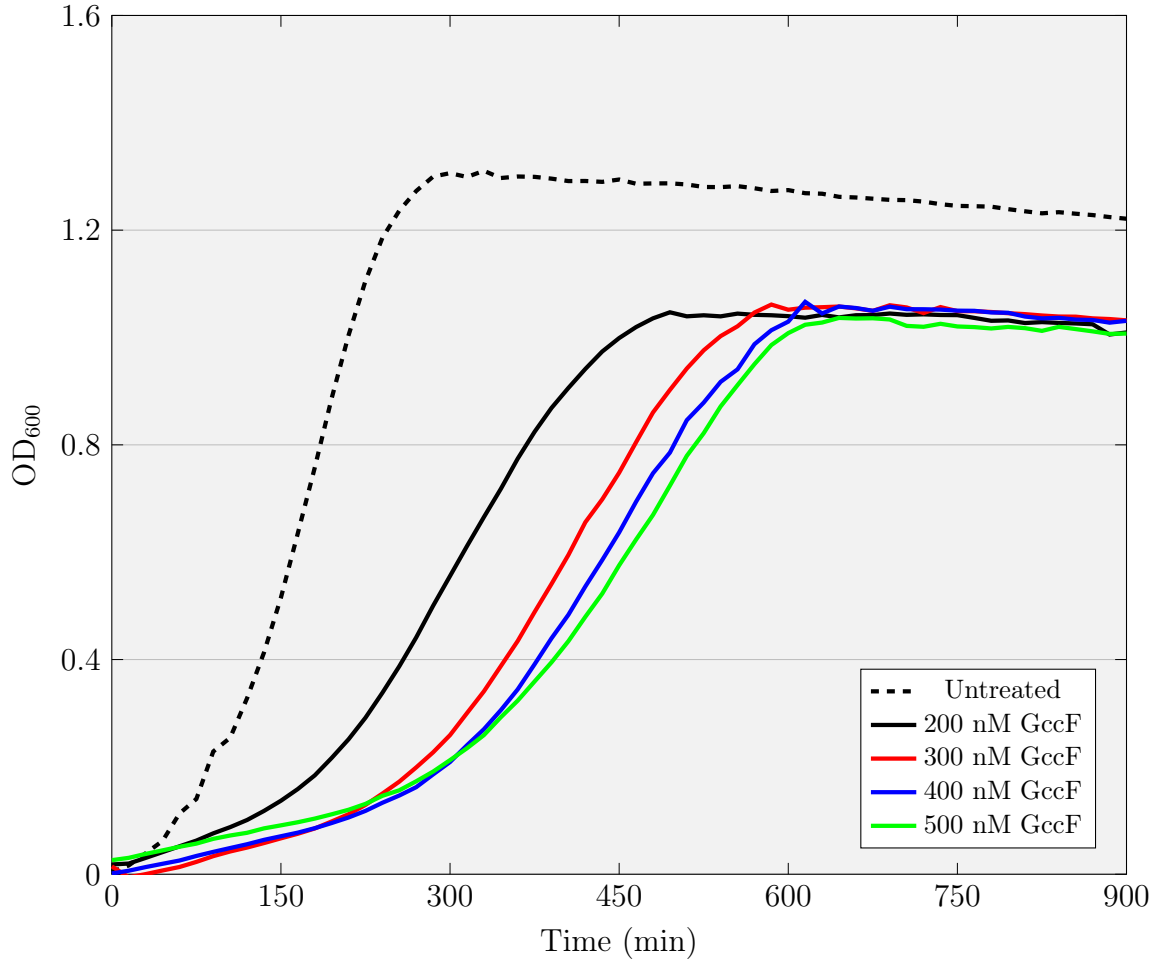


Figure 4.2: **JH2-2 cells treated with different concentrations of GccF**
E. faecalis JH2-2 cells were grown in the presence of different concentrations of GccF over 15 hours (900 minutes). GccF was added to all cultures at time 0. From the concentrations tested, it can be seen that cells never enter complete stasis. Plotted growth curves are the average of triplicate measurements. For clarity, averaged values are shown.

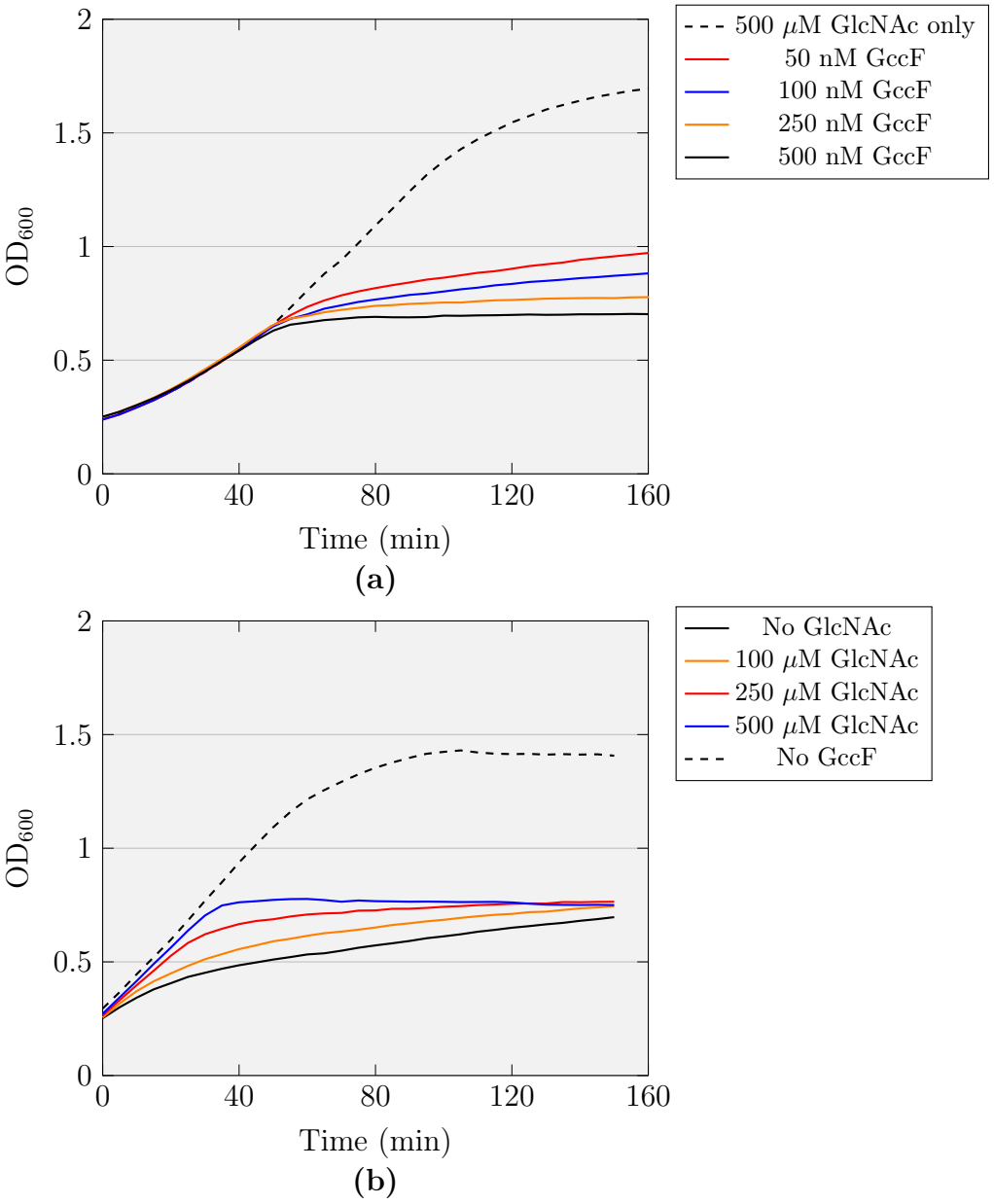


Figure 4.3: ***E. faecalis* JH2-2 GccF and GlcNAc protection tests**
(a) *E. faecalis* JH2-2 cells were grown in the presence of 500 μM GlcNAc and GccF (concentration indicated on legend). GlcNAc can be seen to provide a protective effect for the first $\sim$ 60 minutes. (b) JH2-2 cells were grown in the presence of 100 nM GccF and GlcNAc (concentration indicated on legend). The higher concentrations of GlcNAc can be seen to protect cells from stasis for longer. Additionally, the concentrations of GlcNAc appear to alter the degree of response to GccF, with higher concentrations of GlcNAc leading to more apparent stasis. GccF and GlcNAc all added at time 0, to final concentrations as indicated on the legends. All growth curves are averages of two measurements. For clarity, error bars are not shown.

is being imported into the cells. To better understand the effect GccF has on *E. faecalis* cells and to investigate the mechanism of action of this bacteriocin in general, RNA-seq was carried out on JH2-2 cells treated with GccF and GlcNAc.

4.3.2 RNA sequencing experiment 1

4.3.2.1 Experimental design

RNA-seq was employed to better understand the effects of both GccF and GlcNAc on *E. faecalis* JH2-2 cells, with the hope of identifying the mechanism of action of GccF against bacterial cells. Because treated cells have been observed to enter stasis indefinitely without dying, as shown by their rapid recovery upon exposure to GlcNAc (Figure 4.4) cells were sampled at a number of different time points informed by the growth curves of GccF treated cells. By doing this any immediate transcriptional responses to GccF exposure should be identified, and long term transcriptional changes would be targeted. Additionally, a GlcNAc-only treatment was included to evaluate the gene pathways being affected by GlcNAc, and possibly identify genes/ proteins contributing to the pre-sensitising and/ or protective effects.

Three time points were chosen for sampling following GccF exposure: 0 minutes (immediate harvesting of cells following GccF addition), 10 minutes, and 30 minutes. The 0 minute time point was chosen to sample any initial transcriptomic effects resulting from exposure to GccF, and the 30 minute time point was chosen to assess the transcriptome of cells that have entered and are maintaining bacteriostasis. The 10 minute time point was selected as a mid-point between 0 and 30 minutes. GccF was added to the cells at a final concentration of 250 nM because this is the lowest concentration that was observed to lead to complete stasis of *E. faecalis* JH2-2 cells following pre-sensitisation with 500 μM GlcNAc (Figure 4.3 (a)). GlcNAc pre-sensitisation was not utilised during this round of RNA-seq and consequently the cells did not enter complete stasis. For the GlcNAc-only test, cells were harvested 10 minutes after the addition of 1 mM GlcNAc. All cell cultures were prepared in triplicate, with each triplicate being split, one half being treated, and the other left as a negative control (see Section 4.2.1 for extensive description of methods). It is important to note here that during the 15 minute pelleting step, despite the cells being kept at 4 ℃ they would have remained transcriptionally active. As there was no way to prevent this, processing times were kept consistent across all time points

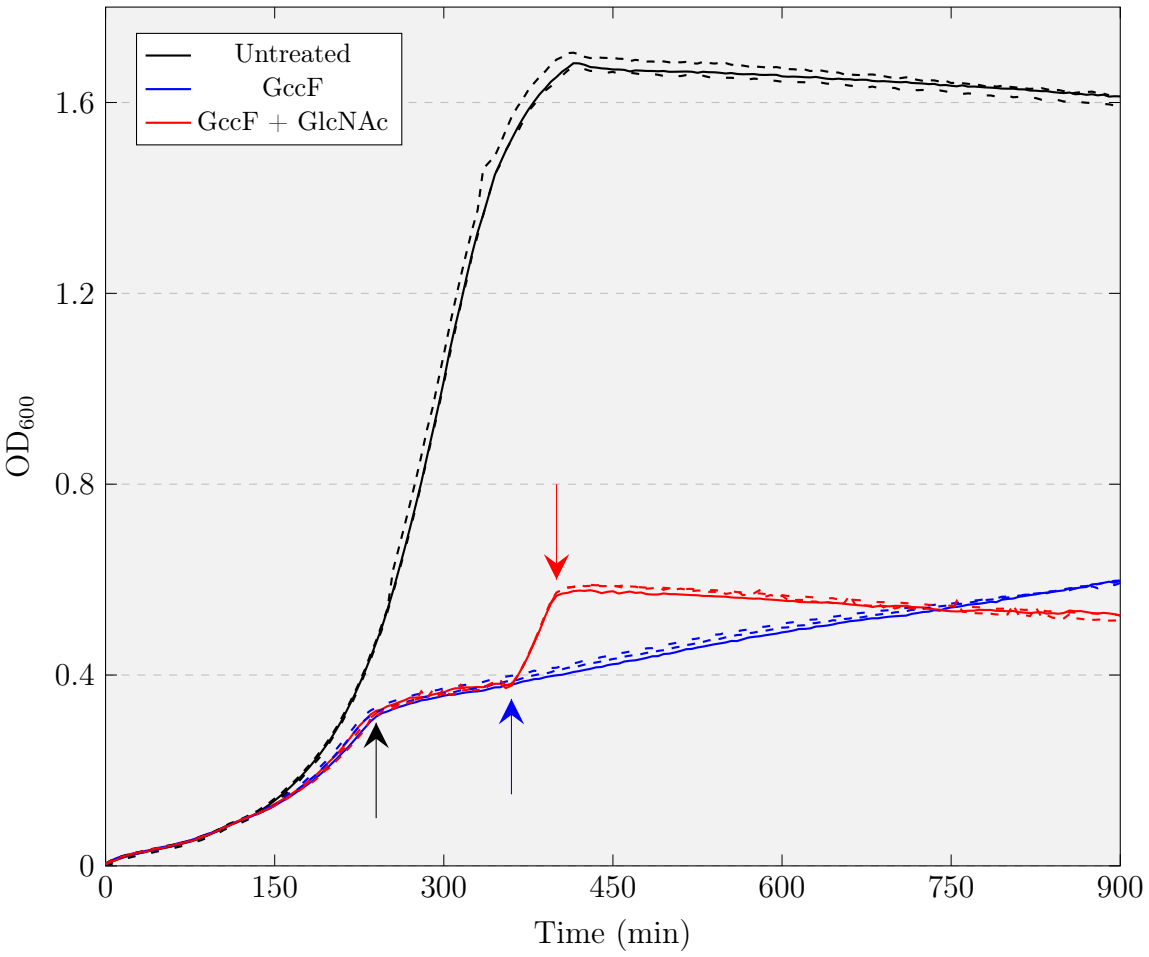


Figure 4.4: **Recovery of GccF-induced stasis by GlcNAc**
E. faecalis JH2-2 cells were grown in TSB media in the presence of 1 mM GlcNAc and 100 nM GccF (added at $t = 0$, red and blue lines), or no additive (black lines). GlcNAc can be seen to protect the cells from GccF for the initial $\sim$ 240 minutes of growth. Stasis can be seen to begin at around 240 minutes (black arrow). An additional 500 μM (final concentration) of GlcNAc was added to one triplicate at $t = 360$ (6 hours; blue arrow), and can be seen to rescue cells from stasis almost immediately (red lines). However, a stronger stasis can be seen to resume shortly after this ($t = 400$; red arrow). Cultures were repeated in triplicate; the legend indicates the colour of each growth curve treatment.

Table 4.1: Number of differentially expressed genes evaluated by *DESeq2*

Test condition	**$\log_2$-fold change < 1 (Down-regulated genes)**	**$\log_2$-fold change > 1 (Up-regulated genes)**
GlcNAc (10 min)	8	11
GccF (0 min)	108	100
GccF (10 min)	164	252
GccF (30 min)	348	373

to minimise any technical variations.

For each of the RNA samples sent for sequencing, around 8-9,000,000 reads were received. All sets of reads passed all quality control steps in *FastQC* version 0.11.3 aside from 'per base sequence content' and 'Kmer content'. Looking at the graphs of these two failed tests, it can be seen that the reason for these failures is most likely to do with the presence of adapters on the reads, as it is the first ~ 5 basepairs that fall outside the expected parameters (Supplementary Figure A.2). Otherwise the rest of the read was fine. The reads were aligned to the reference *E. faecalis* JH2-2 genome using *Bowtie2*, with over 99 % of reads being successfully aligned to the genome for each sample (unaligned percentages ranged from 0.23 % - 0.73 %; Supplementary Table A.8).

4.3.2.2 Statistical analysis using *DESeq2*

RNA-seq data was received from NZGL, and processed as described in Section 4.2.2. Analysis of the four experimental data sets (GccF treatment harvested at 0, 10, and 30 minutes, and GlcNAc treatment harvested at 10 minutes) was carried out using the *DESeq2* package in R[105]. The following filtering parameters were chosen: absolute $\log_2$-fold change greater than 1, and adjusted p-value less than 0.05. Looking at the numbers of up- and down-regulated genes from this output, it can be seen that there are significant changes in gene expression following exposure of JH2-2 cells to GccF. The results indicate that over 100 genes are either up- or down-regulated at time 0, and over 300 genes are up- or down-regulated at 30 minutes (Table 4.1, Figure 4.5). It is important to note that this high degree of gene expression is unlikely to be an experimental artifact, as the GlcNAc-only treatment resulted in only 19 genes in total passing the same parameters (Table 4.1).

4.3.2.3 Statistical analysis using *edgeR*

In order to validate the gene expression seen with *DESeq2*, the same set of genes was analysed using *edgeR*, another R-based differential expression package. Table 4.2 shows the overlap of genes that passed the differential expression criteria of absolute $\log_2$-fold change > 1 and p-value < 0.05. Interestingly, not only did *edgeR* allow a larger number of genes to pass the threshold in all experiments tested, but all genes that passed the filters for *DESeq2* also appeared in the *edgeR* list. In fact, this observation agrees with other published comparisons of *edgeR* and *DESeq2*, which suggests that *DESeq2* provides a more conservative approach to filtering[133]. Based on the relatively conservative approach of *DESeq2* compared to *edgeR*, and the large numbers of differentially-expressed genes overall, all further analysis was carried out using *DESeq2*.

Table 4.2: Comparison of differentially expressed genes as called by *DESeq2* and *edgeR*

	DESeq2		**edgeR**	
Test condition	**lfc <0**	**lfc >0**	**lfc <0**	**lfc >0**
GlcNAc 10 min	8	11	9	21
GccF 0 min	118	106	152	286
GccF 10 min	189	277	246	439
GccF 30 min	371	391	457	608

4.3.2.4 Gene responses

Response to GlcNAc treatment

The inclusion of a GlcNAc-only treatment was intended to highlight the possible up-regulation of a receptor or target for GccF, which would help to explain the observation that GlcNAc appears to pre-sensitise *E. faecalis* to the effects of GccF. The inclusion of GlcNAc in the media resulted in only 19 differentially expressed genes (DEGs): 11 of which were up-regulated, and 8 down-regulated (Table 4.1). When the DEGs from this treatment were annotated, they were consistent with what would be expected for a GlcNAc response (Table 4.3). There

Table 4.3: Genes differentially expressed in presence of GlcNAc

ORF	**Annotation**[†]	Log_2-fold change
ERT23544	BglG family transcriptional antiterminator (EF1515)	2.8
ERT23545	PTS system, N-acetylglucosamine-specific IIBC component (EF1516)	2.7
ERT21222	BglG family transcriptional antiterminator (EF2966)	2.4
ERT21384	Predicted BglG family transcriptional antiterminator	2.4
ERT21383	PTS system beta-glucoside-specific IIABC component (EF0270)	2.1
ERT21216	D-ribose pyranase RbsD (EF2960)	2.1
ERT21217	Ribokinase (rbsK)	2.0
ERT23907	Ribonucleoside hydrolase RihC (EF1921)	1.7
ERT25309	Galactose-6-phosphate isomerase subunit lacA	1.6
ERT23096	Mn^{2+}/Fe^{2+} transporter (EF1057)	1.6
ERT25492	Hypothetical protein (EF0769)	1.6
ERT21673	Glutamine synthetase, type I (glnA)	-2.9
ERT21665	Glucosamine-fructose-6-phosphate aminotransferase (glmS)	-2.9
ERT21674	Regulatory protein GlnR (glnR)	-2.8
ERT23150	Amino acid ABC transporter permease (EF1117)	-2.8
ERT23151	Amino acid ABC transporter permease (EF1118)	-2.5
ERT23153	Amino acid ABC transporter ATP-binding protein (EF1120)	-2.3
ERT23152	Amino acid ABC transporter amino acid-binding protein (EF1119)	-2.3
ERT25484	Amino acid ABC transporter ATP-binding protein (EF0760)	-1.6

[†]Names in brackets indicate V583 gene name

was a noticeable up-regulation of genes associated with the import of GlcNAc into the cell in the form of PTS transporters (ERT23545 and ERT21383), along with the upstream regulatory antiterminator gene of each transporter (ERT23544 and ERT21384, respectively). Interestingly, the third BglG family transcriptional antiterminator, ERT21222, was found to be directly upstream of a gene encoding an ascorbate-specific PTS EIIB subunit (ERT21221) which was not observed to be differentially expressed in this experiment. There was an up-regulation of genes involved in the inter-conversion of sugars acting as substrates for the enzymes involved in carbohydrate metabolism, such as ribokinase (*rbsK*, ERT21217 - interconverts D-ribose-5P and D-ribose) and D-ribose pyranase (*rbsD*, ERT21216 - interconverts D-ribopyranose and D-ribofuranose), both of which are involved in the D-ribose degradation pathway, and ribonucleoside hydrolase (*rihC*, ERT23907), which is involved in ribonucleoside metabolism[135].

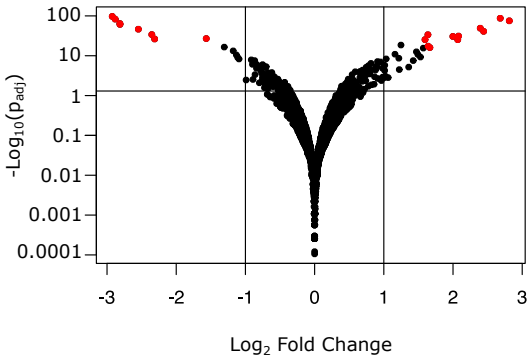
There is noticeable reduction of ATP-binding cassette (ABC) transporters associated with amino acid transport, such as the operon ERT23150-3, which belongs to Glu/ Asp/ Gln-specific transporters, and ERT25484, a Gln-specific transporter.

Additionally, there is down-regulation of glutamine synthetase (*glnA*, ERT21673), which is involved in the conversion of glutamine from glutamate and ammonia, and its upstream repressor (*glnR*, ERT21674). Interestingly, glucosamine-fructose-6-phosphate aminotransferase (GlmS, ERT21665) is also down-regulated, which is the first enzyme in the peptidoglycan synthesis pathway, which is involved in converting D-fructose-6-P to D-glucosamine-6-P by the addition of an amine group, with the amine being donated by a glutamine[19]. Analysis by STRING (https://string-db.org) shows that GlmS interacts to an extent with GlnA (Supplementary Figure A.3), and as such the down-regulation of both of these genes (along with GlnR) may be tied together with the import of GlcNAc. However, the mRNA of *glmS* is also known to form a ribozyme, which interacts with glucosamine-6-P (GlcN-6-P) to cleave itself, and thus down-regulate *glmS* expression[63,166]. As one of the first steps following import of GlcNAc as GlcNAc-6-P through the PTS transporter is interconversion to GlcN-6-P by GlcNAc-6-P deacetylase (Supplementary Figure A.4)[7], it is possible that the accumulation of this metabolite is causing down-regulation of *glmS* by triggering this ribozyme activity.

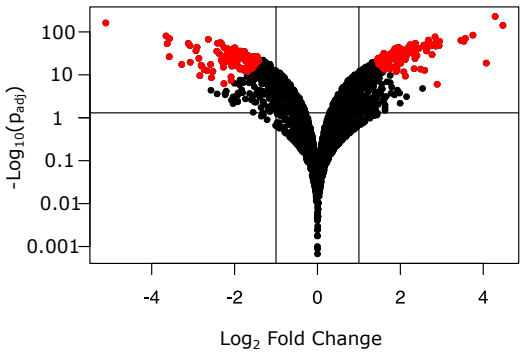
The most interesting observation from this list of GlcNAc-induced DEGs is the up-regulation of the GlcNAc-specific PTS transporter, ERT23545. Comparing the amino acid sequences of the hypothesised GccF receptor from *Lb. plantarum* ATCC 8014, PTS18CBA (NCBI gene ID: 1062723) against *E. faecalis* JH2-2 resulted in a top sequence similarity hit of ERT23545 (57.6 % identity, e-value = 0), suggesting that this transporter would be the most likely target of GccF in JH2-2 cells. The observations that GlcNAc pre-treatment increased sensitivity of JH2-2 cells towards GccF, and the gene encoding the PTS transporter is up-regulated by GlcNAc, strongly infers that GccF targets a GlcNAc-specific PTS transporter. However, what remains unclear at this point is whether the PTS is the primary receptor/ target of GccF, or merely a localising tool employed by the C-terminal GlcNAc (see Chapter 3).

Response to GccF time treatment

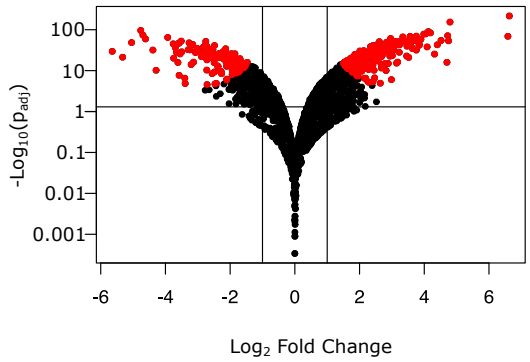
The time-course treatment of JH2-2 cells with GccF resulted in a significant number of DEGs, with roughly equal numbers of up- and down-regulated genes at each time point (Figure 4.5, Table 4.1). Principal component analysis (PCA) carried out on



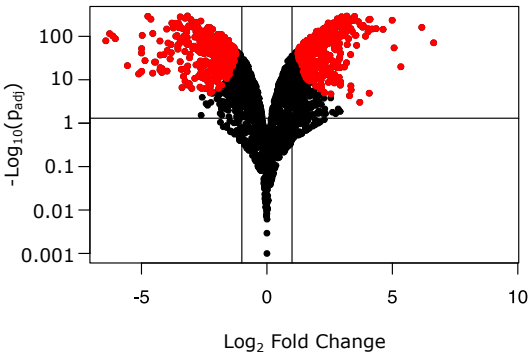
(a) 1 mM GlcNAc, 10 minutes



(b) 250 nM GccF, 0 minutes



(c) 250 nM GccF, 10 minutes



(d) 250 nM GccF, 30 minutes

Figure 4.5: **Volcano plots of GccF treatment time-course**
Each point in the above figures represents one gene from *E. faecalis* JH2-2. The $\log_2$-fold change (x- axis) is plotted against the -$\log_{10}$ of the adjusted p-value (p_{adj}; y-axis). The red spots indicate genes that passed the cut-off thresholds of $|\log_2\text{-fold change}| > 1$ and $p_{adj} < 0.05$, according to *DESeq2*'s parameters.

Table 4.4: GlcNAc response genes in GccF treatment

ORF	Annotation	Log_2-fold change			
			GccF addition		
		GlcNAc	0 min	10 min	30 min
ERT23544	BglG family transcriptional antiterminator (EF1515)	2.8	-	-1.9	-
ERT21384	Predicted BglG family transcriptional antiterminator	2.4	-3.6	-3.8	-2.8
ERT21383	PTS system beta-glucoside-specific IIABC component (EF0270)	2.1	-	-	-2.1
ERT21216	D-ribose pyranase (RbsD)	2.1	3.4	3.2	2.4
ERT21217	Ribokinase (RbsK)	2.0	2.0	-	-
ERT23907	Ribonucleoside hydrolase (RihC)	1.7	-	-3.7	-4.1
ERT25309	Galactose-6-phosphate isomerase subunit lacA	1.6	-	-3.5	-3.1
ERT23096	Mn^{2+}/Fe^{2+} transporter	1.6	2.2	-	-1.6
ERT21673	Glutamine synthetase, type I (glnA)	-2.9	-	-	-2.3
ERT21674	Regulatory protein GlnR (glnR)	-2.8	-	-	-2.2
ERT23150	Amino acid ABC transporter permease (EF1117)	-2.8	-	-	-2.3
ERT23151	Amino acid ABC transporter permease (EF1118)	-2.5	-	-	-2.4
ERT23153	Amino acid ABC transporter ATP-binding protein (EF1120)	-2.3	-	-2.5	-3.5
ERT23152	Amino acid ABC transporter amino acid-binding protein (EF1119)	-2.3	-	-	-3.1

the three time points indicated good statistical separation (Supplementary Figure A.5). The 10 most highly up- and down-regulated genes for each time point (based on log-fold change) are listed in Supplementary Tables A.2 and A.3, and the full lists of genes can be found in Tables S1, S2 and S3 of the Supplementary Data file. . At 10 and 30 minutes there is a strong repression of numerous spermidine/ putrescine ABC transporters, as well as up-regulation of other ABC transporters at all three time points. Interestingly, the specific DEGs associated with GlcNAc treatment alone were not all observed in the GccF treated samples, although there was some degree of overlap (Table 4.4). At one or more time points 12 genes that were differentially expressed by exposure to GccF were also affected by exposure to GlcNAc. Six of these DEGs were also down-regulated by GlcNAc (amino acid ABC transporters, GlnR, and glutamine synthetase) with the other six being up-regulated when exposed to the amino sugar (Table 4.4). The genes encoding ribokinase (*rbsK*, ERT21217) and D-ribose pyranase (*rbsD*, ERT21216) were up-regulated by GlcNAc, and also up-regulated at one or more of the time points sampled following GccF addition.

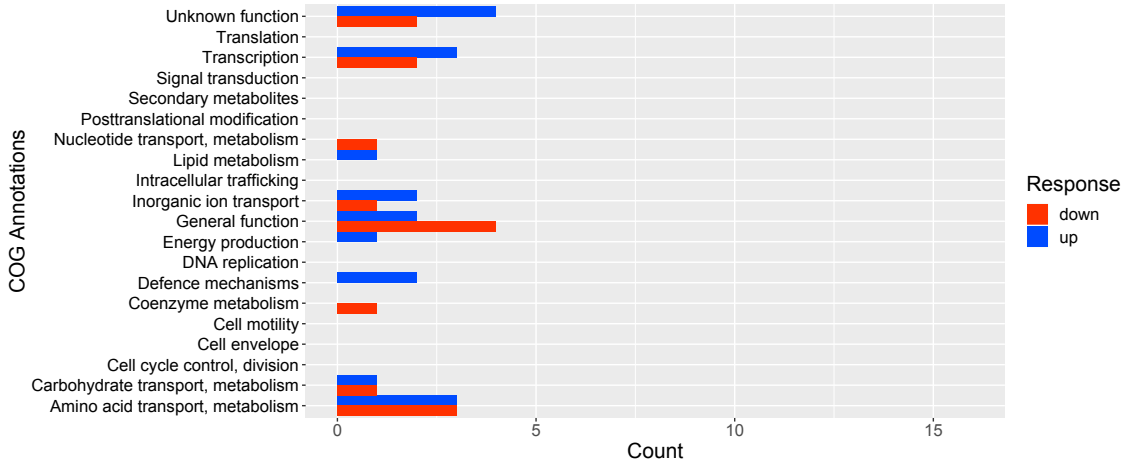
To better understand the processes and gene regulatory pathways being affect by JH2-2 cells, gene ontology (GO) enrichment was carried out by submitting Entrez gene IDs to the online database for annotation, visualisation and integrated discovery (DAVID; https://david.ncifcrf.gov). Entrez gene IDs were assigned to each open

reading frame (ORF) of JH2-2 by a reciprocal protein BLAST between the protein sequences translated from JH2-2 ORFs and the protein sequences translated from the ORFs of the reference strain *E. faecalis* V583. Lists of gene IDs for those genes up- or down-regulated at each time point were submitted to the online DAVID database, to allow for enrichment of GO terms belonging to genes being affected in response to GccF.

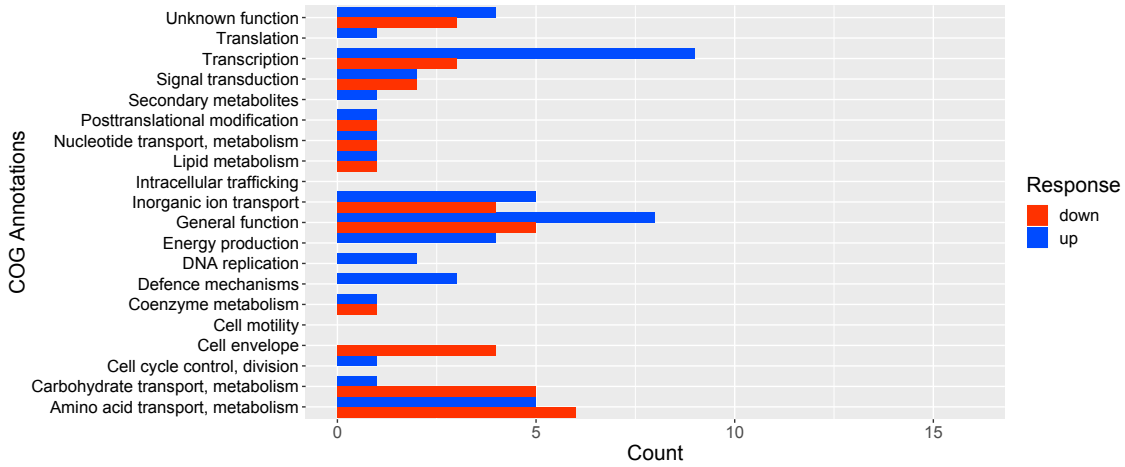
One of the tools provided by DAVID is annotation and analysis of COG terms (Clusters of Gene Ontologies) and KEGG (Kyoto Encyclopedia of Genes and Genomes) pathways. Counts of general COG terms enriched from each time point can be seen in Figure 4.6. Each count represents a term associated with a single gene in each list submitted to DAVID - however, note that only around 20% of genes in each list were actually assigned COG terms. Looking at the profile of down-regulated COG terms, 'amino acid transport and metabolism' and 'transcription' can be seen to be following a strongly time-dependent increase in transcription, followed by 'inorganic ion transport and metabolism'. These time-dependent changes indicate an increasing repression of genes associated with these functions as GccF-induced stasis continues. Likewise, looking at the up-regulated COG terms, 'transcription', 'defense mechanisms', and 'amino acid transport and metabolism' all appear to follow a time-dependent trend.

In addition to looking at the numbers of general COG terms that were enriched, specific GO terms enriched in each treatment were assessed. Tables 4.5 and 4.6 show the GO terms of up- or down-regulated genes that were enriched in each list of Entrez gene IDs supplied to DAVID, which had passed a p-value of ≤ 0.05. GO terms are generally divided into three main groups: biological process, molecular function, and cellular component. The only terms that are consistently down-regulated at each time point are 'PTS system' (biological process) and 'protein-N(PI)-phosphohistidine-sugar phosphotransferase activity' (molecular function) (Table 4.5), suggesting a global down-regulation of genes associated with PTS transporters. On the other hand, only one term is up-regulated at all three time points: 'integral component of membrane' (cellular component), although there are a number of GO terms associated with ATPase activity at 0 and 30 minutes (Table 4.6).

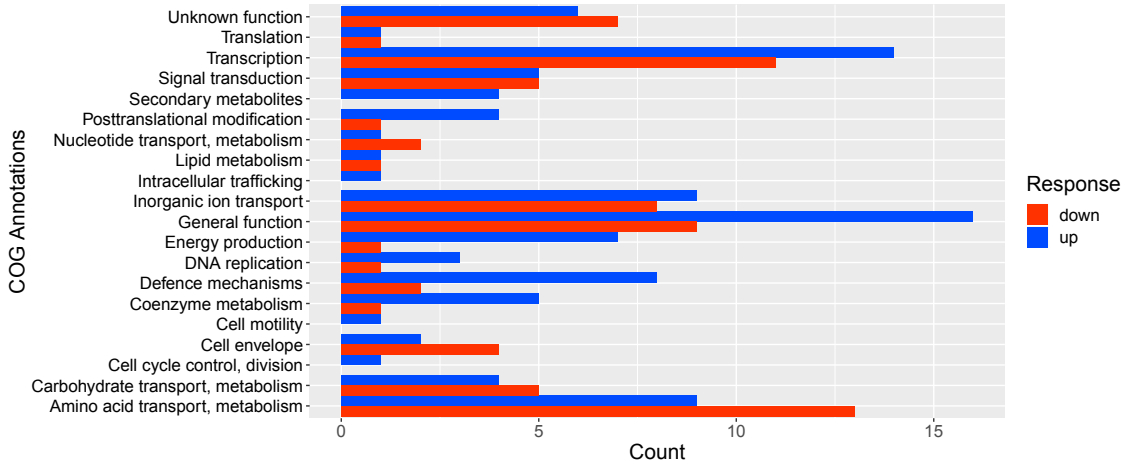
The presence of GO terms relating to PTS systems in the list of down-regulated genes at all three time points prompted investigation into which specific PTS-related genes



(a) 0 minutes



(b) 10 minutes



(c) 30 minutes

Figure 4.6: **COG Annotation counts for each time point**

Figure 4.6: General COG annotations were generated for each list of up- and down-regulated DEGs submitted to the online DAVID database. Graphs show the number of DEGs associated counted for each list supplied; red bars are up-regulated assignments, and blue bars are down-regulated assignments. (a) 0 minutes, (b) 10 minutes, and (c) 30 minutes. From these plots it appears that the largest number of COGs down-regulated at each time point are 'General function' and 'Amino acid transport and metabolism', while the largest number of up-regulated COGs are 'General function' and 'Transcription'. Note that only around 20% of each submitted gene list were given COG annotations to be counted.

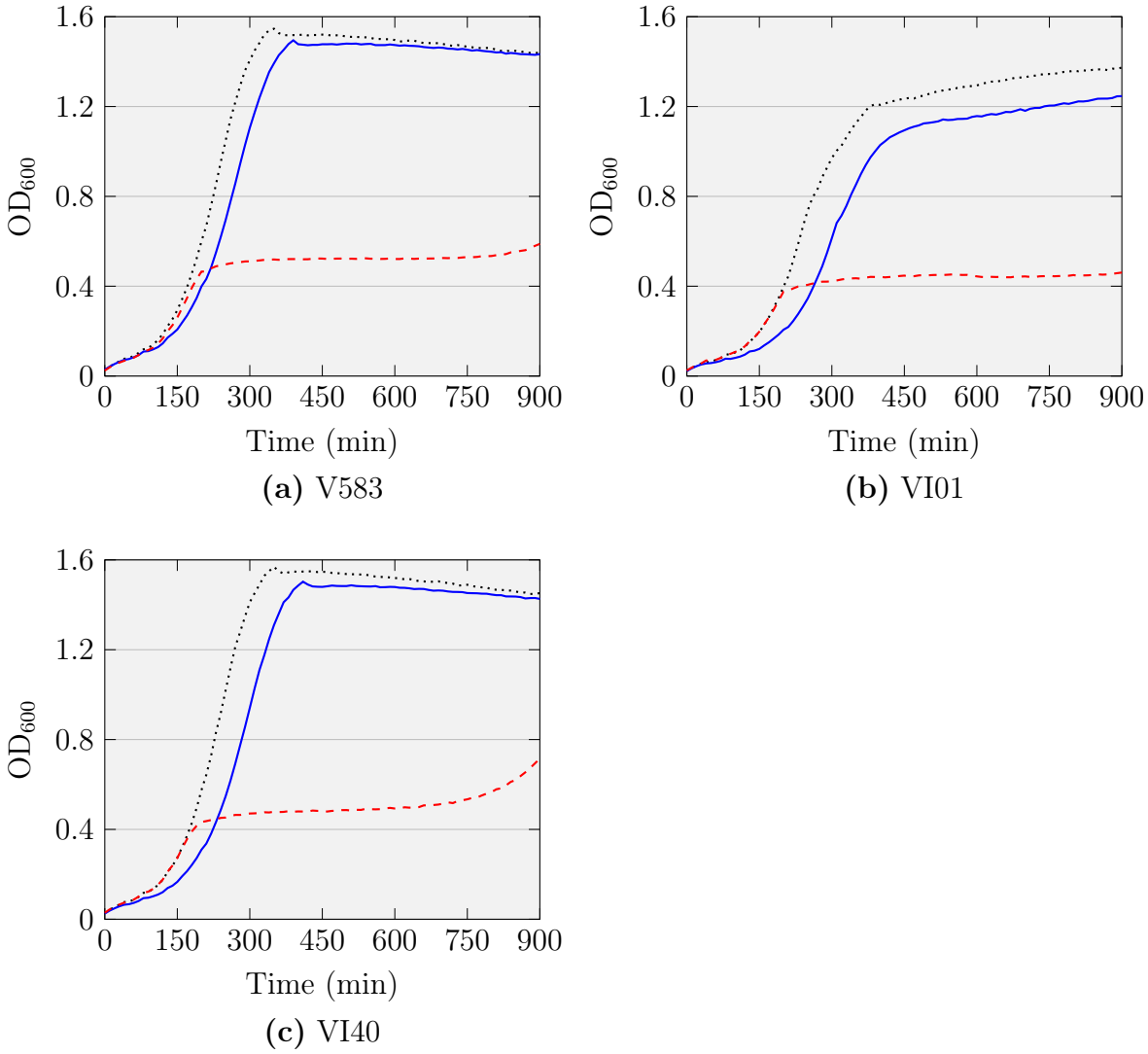
were being down-regulated. Approximately 13 genes encoding PTS subunits, located in four different operons, were down-regulated at one or more time points post GccF exposure, although the GlcNAc-specific PTS operons were not on this list (Table 4.7). A search through the annotated genome of *E. faecalis* JH2-2 revealed that there are approximately 70 genes encoding PTS components, with 54 genes falling into 18 operons, and 16 genes not falling into a multi-PTS-containing operon. Additionally, a search for PTS regulation domains (PRDs) within the genome showed there are a number of genes involved in regulation of individual PTS components such as BglI-type transcriptional antiterminators[57] and Sigma 54-specific enhancer-binding proteins (EBPs)[69].

Sigma factor 54 (σ^{54}) is an alternative sigma factor, which is found in a wide range of bacteria[69] and has been shown to modulate a number of cellular processes, including expression of a subset of PTS transporters in Gram-positive bacteria such as *Lb. plantarum*[146], and play a role in nitrogen metabolism[58]. It has also been associated with susceptibility of *E. faecalis* to Class IIa (YGNG-motif containing) bacteriocins, with knockout of its gene promoting resistance to these peptides[52]. The presence of PRD-containing sigma factor σ^{54} EBP genes in the genome of *E. faecalis* prompted a search for DNA binding consensus sequences of this protein, in order to account for the specific response of the four down-regulated PTS operons (Table 4.7). A search through the genome of *E. faecalis* JH2-2 for a σ^{54} EBP consensus DNA-binding site[69] using PATLOC (http://www.cmbl.uga.edu/software/patloc.html) showed that binding sites appeared upstream from all but one of the four PTS-encoding operons down-regulated in response to GccF treatment. The exception was the operon ERT25780-2, in which the predicted DNA-binding site occurred inside the gene encoding the EIIC domain of this ORF. To test the involvement of σ^{54} in the response of *E. faecalis* cells to GccF, a σ^{54} knockout strain of *E. faecalis* (one lacking the *rpoN*

Table 4.5: Down-regulated GO terms for time-course GccF treatment

Time	Classification	Gene Ontology	Count	*p*-value
0 minutes	Biological Process	PTS system	8	4.6×10^{-2}
		'De novo' UMP biosynthesis	3	3.4×10^{-2}
	Molecular Function	Hydrolase activity	7	1.8×10^{-2}
		Protein-N(PI)-phosphohistidine-sugar phosphotransferase activity	6	2.2×10^{-2}
		Transporter activity	6	1.3×10^{-2}
10 minutes	Biological Process	PTS system	15	3.8×10^{-3}
		Peptide transport	3	1.2×10^{-2}
	Cellular Component	Plasma membrane	18	2.8×10^{-2}
	Molecular Function	Protein-N(PI)-phosphohistidine-sugar phosphotransferase activity	11	3.6×10^{-4}
30 minutes	Biological Process	PTS system	20	2.3×10^{-2}
		Negative regulation of transcription	5	3.1×10^{-2}
		Peptide transport	3	3.8×10^{-2}
		Carbohydrate metabolic process	15	4.8×10^{-2}
		Transport	10	4.3×10^{-2}
	Cellular Component	Plasma membrane	32	7.6×10^{-4}
	Molecular Function	Transporter activity	13	1.5×10^{-3}
		Protein-N(PI)-phosphohistidine-sugar phosphotransferase activity	14	1.7×10^{-3}
		Hydrolase activity, hydrolyzing *O*-glycosyl compounds	6	2.2×10^{-3}

gene encoding σ^{54}, designated VI01) was obtained from the University of Kansas (kind gift from Dr Lynn Hancock), along with the unmodified parent strain (V583) and a recomplemented strain (VI40)[87]. These strains were tested for susceptibility to GccF using the culture methods described in Section 4.3.1 (Figure 4.7). While the absence of σ^{54} led to a decrease in the growth rate of the cells, there was still noticeable stasis induced by GccF against both strain VI01 and strain VI40, indicating that the bacteriostatic effect of GccF is not directly controlled by alternative sigma factor σ^{54}. Additionally, the 100 nM GccF-treated sample of strain VI01 appeared to recover from the effects of GccF, although more slowly than either the wild-type and complemented strains, suggesting a mildly increased susceptibility (Figure 4.7 (b)).



(a) V583

(b) VI01

(c) VI40

Figure 4.7: **Response of σ^{54} knockout strain to GccF**
Growth plots of *E. faecalis* strains (a) V583, (b) σ^{54} K/O strain VI01, and (c) recomplemented strain VI40, challenged with 100 nM GccF (blue), 100 nM GccF + 500 μM GlcNAc (red dashed), or unchallenged (black dotted), are shown. From the growth curves, it appears that absence of σ^{54} does not lead to resistance to GccF, although it does lead to a decreased growth rate. Addition of GlcNAc still pre-sensitises cells to the effects of GccF.

Table 4.6: Up-regulated GO terms for time-course GccF treatment

Time	Classification	Gene Ontology	Count	*p*-value
0 minutes	Biological Process	D-ribose catabolic process	2	5.0×10^{-2}
	Cellular Component	Integral component of membrane	32	3.4×10^{-2}
10 minutes	Biological Process	Response to oxidative stress	4	1.8×10^{-3}
	Cellular Component	Integral component of membrane	79	1.6×10^{-6}
	Molecular Function	Oxidoreductase activity	11	1.6×10^{-3}
		Nucleotide binding	6	9.9×10^{-3}
30 minutes	Biological Process	Plasma membrane ATP synthesis coupled proton transport	4	3.6×10^{-3}
		Response to oxidative stress	4	8.3×10^{-3}
		Transcription, DNA-templated	19	4.0×10^{-2}
		ATP hydrolysis coupled proton transport	4	2.5×10^{-2}
	Cellular Component	Integral component of membrane	104	4.3×10^{-3}
		Plasma membrane	26	2.3×10^{-2}
	Molecular Function	Oxidoreductase activity	14	4.7×10^{-3}
		Proton-transporting ATP synthase activity, rotational mechanism	6	3.1×10^{-3}
		Cation-transporting ATPase activity	4	1.3×10^{-2}
		Nucleotide binding	8	7.0×10^{-3}

4.3.3 RNA sequencing experiment 2

4.3.3.1 Experimental design

Treatment of *E. faecalis* JH2-2 cells with 250 nM GccF resulted in a large number of DEGs at each time point, with the largest number appearing at 30 minutes post-treatment (Figure 4.7; Table 4.1). However, some of these changes will be due to the cells shutting down, or going into stasis regardless of the cause of the stasis. To help reduce any secondary responses owing to a high concentration of GccF, RNA-seq was repeated using a lower concentration of GccF (20 nM). Cells were sampled at 30 minutes, to allow then to fully respond to GccF, and to provide a point of comparison to the first RNA-seq dataset. Two extra samples were included, both sampled 30 minutes post-treatment: GlcNAc pre-treated JH2-2 cells, which were subjected to 20 nM GccF, and VI01 (σ^{54} K/O) cells treated with 100 nM GccF. 100 nM was chosen for the sigma factor σ^{54}-deficient strain to increase the chances of detecting DEGs, as these cells had already been shown to be less susceptible to GccF.

Cells were pre-sensitised to GccF by adding GlcNAc to a final concentration of 100

Table 4.7: Down-regulated PTS genes

ORF	Description	Down-regulated?		
		0 minutes	10 minutes	30 minutes
ERT20558	PTS system mannose-specific IID component	No	No	**Yes**
ERT20559	PTS system mannose-specific IIC component	No	No	**Yes**
ERT20560	PTS system mannose-specific IIAB component	No	**Yes**	**Yes**
ERT20561	PTS system subunit IIB	No	**Yes**	**Yes**
ERT24743	PTS system subunit IIB	**Yes**	**Yes**	**Yes**
ERT24744	PTS system subunit IIC	No	**Yes**	**Yes**
ERT24748	PTS system subunit IIB	**Yes**	**Yes**	**Yes**
ERT24749	PTS system subunit IIA	**Yes**	**Yes**	**Yes**
ERT24750	PTS system cellobiose-specific IIB component	**Yes**	**Yes**	**Yes**
ERT25779	PTS system transporter subunit IID	No	No	**Yes**
ERT25780	PTS system transporter subunit IIC	No	**Yes**	**Yes**
ERT25781	PTS system transporter subunit IIB	**Yes**	**Yes**	**Yes**
ERT25782	PTS system subunit IIA	**Yes**	**Yes**	**Yes**

μM to cultures of JH2-2 cells that had reached an OD_{600} of 0.2, immediately before splitting. GccF was then added to one half of each half once the OD_{600} had reached 0.5, as in all other experiments (see Section 4.2.1). The reasonably low concentration of GlcNAc ensured that the amino sugar should be completely utilised by the cells before the addition of GccF, preventing the protective effect of GlcNAc. All methods for collecting RNA and processing the data were identical to RNA-seq experiment 1 (Section 4.3.2).

Around 10-11 million reads were received for each RNA sample submitted for sequencing. For the GlcNAc pre-treated samples, there was a similar percentage of reads aligned to the reference sequence as in the first RNA-seq experiment ($>$ 99 %). However, for the non pre-treated samples and the VI01 samples, this decreased to around 95 % (Supplementary Table A.9). The worst sample belonged to VI01 negative control 1 (sample 2776-16; Supplementary Table A.9), which had an unaligned percentage of 11 %. However, all of the reads passed the FastQC tests to the same degree as the reads from the previous experiment (Section 4.3.2), which was not indicative of any issue with the RNA integrity or sequencing. *DESeq2* was employed to statistically analyse the data as previous.

Table 4.8: Differentially expressed genes from second RNA-seq experiment

Test condition (GccF)	**$\log_2$-fold change < 0 (Down-regulated genes)**	**$\log_2$-fold change > 0 (Up-regulated genes)**
JH2-2 (20 nM)	52	43
GlcNAc pre-treated JH2-2 (20 nM)	102	109
VI01 (σ^{54} K/O) (100 nM)	92	186

4.3.3.2 Gene responses

Table 4.8 shows the number of DEGs for each experiment that passed the filter settings of *DESeq2* ($|\log_2\text{-fold change}| \geq 1$ and $p_{adj} < 0.05$) (see also Supplementary Tables A.4 and A.5 and Tables S4, S5 and S6 of the Supplementary Data file). For the non-GlcNAc pre-treated samples, it can be seen that there are considerably fewer DEGs in the cells treated with 20 nM GccF 250 nM GccF (Table 4.1). The GlcNAc pre-treated sample, on the other hand, shows approximately twice the number of DEGs compared to those in the non pre-treated sample, confirming the ability of GlcNAc to pre-sensitise *E. faecalis* cells to GccF. The sigma factor σ^{54} knockout mutant had more up-regulated and less down-regulated DEGs than the GlcNAc pre-treated sample. Additionally, principal component analysis on the three samples indicated good statistical separation (Supplementary Figure A.6).

Response to 20 nM GccF

All of the 43 genes up-regulated following treatment with 20 nM GccF were also up-regulated in the 250 nM 30 minute experiment, while 48 of the 52 down-regulated genes are also down-regulated in the 250 nM treatment. The four genes that were only down-regulated in the 20 nM treatment encode ERT23254 (chlorohydrolase), ERT23251 (spermidine/ putrescine ABC transporter ATP-binding protein), ERT24604 (cation transporter E1-E2 family ATPase), and ERT25522 (allantoinase). This overlap between the two experiments suggests a specific number of genes which are affected upon exposure to a low concentration of GccF. Genes that were down-regulated under this treatment that were also down-regulated under the 250 nM treatment included *glnA* (encoding glutamine synthetase), two genes encoding xanthine/ uracil permeases, a number of spermidine/ putrescine ABC transporters,

and numerous amino acid ABC transporters, predicted by KEGG to be involved in transport of glutamine. Genes that were up-regulated included numerous V-type ATPases, a handful of amino acid ABC transporters, and two Na^+/ H^+ antiporters (*napA* and *nhaC-2*). This raises questions as to what is triggering the larger set differentially expressed genes at higher (250 nM) concentrations of GccF: is this a general response to an antimicrobial substance, or is it specific to GccF?

Response to GlcNAc pre-treated 20 nM GccF

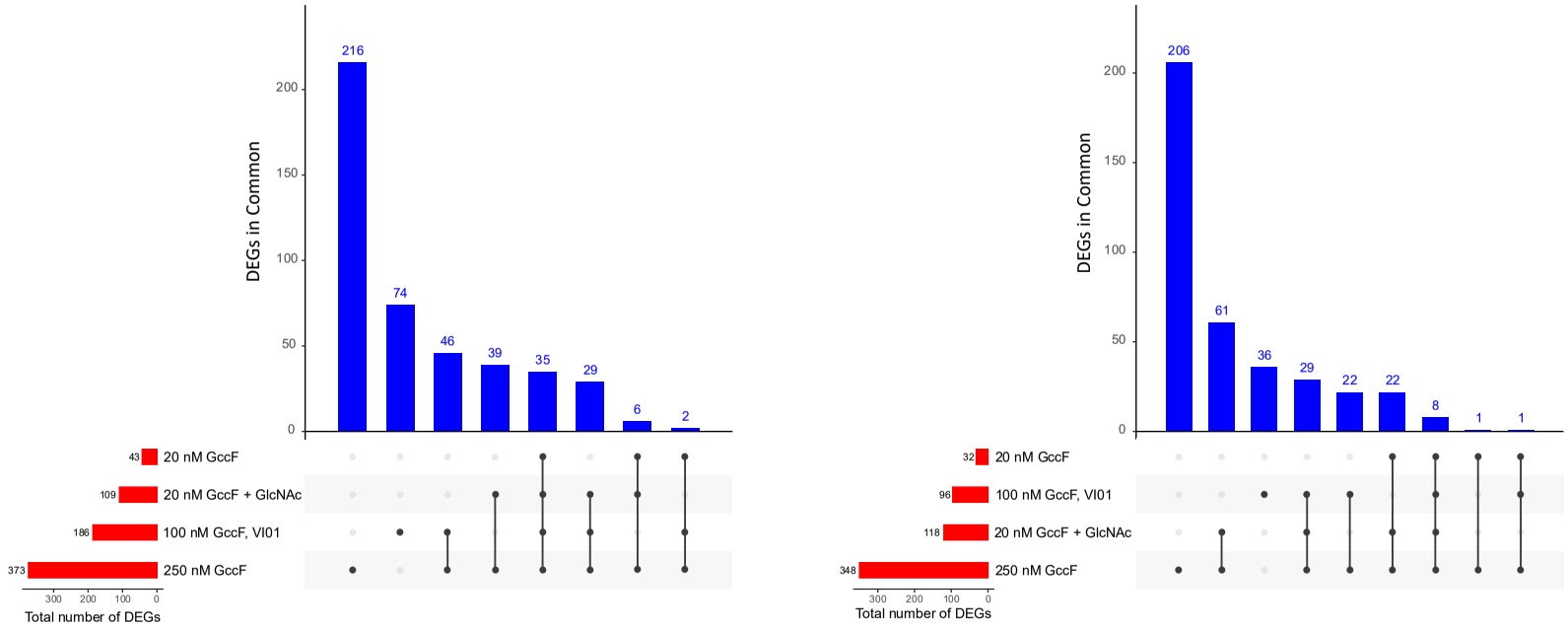
All but 12 genes (6 up and 6 down-regulated) of the GlcNAc pre-treated 20 nM GccF treatment were also differentially expressed in the 250 nM treatment. Exclusively up-regulated genes include those at loci ERT21274 (a LysM domain-containing protein), ERT23892 (a glycerophosphoryl diester phosphodiesterase family protein), ERT23695 (a permease), and ERT23691 (a MarR family transcriptional regulator). Exclusively down-regulated genes include loci ERT20801 (a formate/nitrite transporter family protein), ERT23251 (a spermidine/putrescine ABC transporter ATP-binding protein), ERT21232 (encoding 2-dehydropantoate 2-reductase), ERT24604 (a cation transporter E1-E2 family ATPase), and ERT24655 (a sulfate transporter family protein).

The GlcNAc pre-treated cells had double the DEGs compared to the non-pre-treated, and these DEGs all largely appear in the 250 nM treatment, indicating that increased expression of GccF receptors (such as GlcNAc-PTS) amplifies the GccF response. It is of particular interest to note that pre-treating JH2-2 cells with GlcNAc (which appears to lead to increased expression of GlcNAc importers (Table 4.3)) results in a transcriptional response akin to increasing the concentration of GccF. If the GlcNAc-PTS transporter is the primary receptor of GccF, then one would not expect an increase in external GccF concentration to necessarily lead to a proportional increase in transcriptional response, as presumably a GccF concentration would be reached which would saturate all GccF binding sites (see Figure 3.13 (a) from Chapter 3). However, such an effect would be seen if the GlcNAc-PTS is merely a secondary/ localising receptor (Figure 3.13 (b) from Chapter 3). These observations together indicate that the DEGs from the GlcNAc pre-treated 20 nM GccF condition reflect a strong response to GccF.

Response by Sigma 54 knockout mutant

The sigma factor σ^{54} knockout strain VI01 was included in the second set of RNA-seq to determine which genes are under the control of σ^{54}. Visualisation of the aligned VI01 reads to reference strain V583 confirmed that the strain is a σ^{54} (*rpoN*) knockout (Supplementary Figure A.7). Because this strain still exhibits sensitivity to GccF, the loss of differential expression for a set of genes would indicate that those specific pathways are not instrumental in initiating or maintaining a GccF response. σ^{54} is known to be involved in regulation expression of PTS components in a number of Gram-positive bacterial species[69]. The large down-regulation of PTS genes seen in all three time points in RNA-seq experiment one (Table 4.7), implied that σ^{54} could be involved in a response to GccF. When *E. faecalis* VI01 cells (lacking a functional copy of *rpoN*, which encodes σ^{54}) were treated with 100 nM GccF, there was complete loss of down-regulation of the PTS transporters, in contrast with what was seen at the 30 minute time point for *E. faecalis* JH2-2 cells. A heatmap of KEGG pathways affected by GccF treatment shows pathways containing PTS components are down-regulated in all cells exposed to GccF except the VI01 cells (Figure 4.9 (a); see next section). Additionally, it is interesting to note that the down-regulation of spermidine/ putrescine ABC transporters that was seen at 10 and 30 minutes in RNA-seq experiment 1, and was also seen at both the 20 nM GccF treatments, is absent in the VI01 sample, suggesting that these transporters too are under the regulation of σ^{54}.

Not all of the DEGs in the VI01 strain were seen in the 250 nM JH2-2 treatment: 34 genes were exclusively down-regulated, and 74 genes were exclusively up-regulated (Figure 4.8). The majority of exclusively up-regulated genes from the sigma 54 knockout experiment are hypothetical proteins, with no characterised function or homology to other proteins. They include a LysM domain-containing protein (EF0443), and glucosamine-6-phosphate isomerase (*nagB*) (Supplementary Table A.1). The role of NagB is the opposite of GlmS (described in Section 4.3.2.4) - rather than transfer an amine from glutamine to fructose-6-P, thus forming glucosamine-6-P, NagB removes the amino group from GlcN-6-P forming fructose-6-P. Looking at the KEGG pathway for amino sugar and nucleotide metabolism, it can be seen that the only fate of Fru-6-P following processing by this enzyme in glycolysis/ gluconeogenesis (Supplementary Figure A.8). The up-regulation of this enzyme would suggest that the cells are responding to a perceived shortage of carbon sources, and



(a) Up-regulated genes

(b) Down-regulated genes

Figure 4.8: **Gene expression overlap between samples harvested at 30 minutes**
Graphs showing the numbers of gene expression overlap between the four experiments harvested at 30 minutes for up-regulated genes (a) and down-regulated genes (b). The bar graphs on the left (red) indicate the total numbers of up- or down-regulated genes identified in each of the four experiments, while the top bar graphs (blue) show the number of genes overlapping the indicated experiments (solid black dots). This overlap representation is considered 'distinct'; that is, the numbers of genes highlighted in the blue graphs are the numbers that are *only* seen in the overlap regions. For another representation, please see Figure A.1.

are subsequently diverting GlcN-6-P to the glycolytic pathway rather than cell wall metabolism. Of the exclusive down-regulated genes, 5 encode proteins involved in citrate metabolism: *citC* ([citrate (pro-3S)-lyase] ligase), *citE* (citrate lyase subunit beta), *citX* (2-(5"-triphosphoribosyl)-3'-dephosphocoenzyme-A synthase), and *citF* (citrate lyase subunit alpha), which all form one operon; and EF1207, citrate carrier protein (Supplementary Table A.1). Interestingly, the upstream transcriptional antiterminator (ERT23544) of the GlcNAc-PTS identified in the GlcNAc-only treatment (ERT23545; Table 4.3) is down-regulated, although the PTS transporter directly under its control is not differentially expressed. The down-regulation of this transporter is interesting, as it is the only occurrence of a GlcNAc-PTS component being differentially expressed following GccF treatment. The regulatory mechanism behind the down-regulation of this antiterminator is not known, nor is its relationship to σ^{54}.

One thing worth noting is the role of σ^{54} in promoting susceptibility to Class IIa bacteriocins. It has been previously shown that the knockout of *rpoN* in *E. faecalis* JH2-2 leads to resistance to this class of bacteriocins, but does not affect the cells' susceptibility towards nisin[52]. The evidence here showing the loss of regulation of a subset of PTS components by strain VI01 following GccF treatment is in agreement with the expression of a subset of PTS components, including the Man-PTS, being regulated by σ^{54} EBPs[71]. Given the primary target of Class IIa bacteriocins is the Man-PTS, the loss of expression of this target following the deletion of *rpoN* would be expected to promote resistance. As the GlcNAc-specific PTS transporters are not regulated by σ^{54}, but rather individual antiterminators[71], the lack of resistance of strain VI01 to GccF in this experiment is explained.

4.3.4 Overview of the RNA-seq data

Comparing the results of all six GccF treatment experiments provides a good overview of how *E. faecalis* responds to GccF both over time, and at different bacteriocin concentrations. Figure 4.8 shows the large amount of gene expression overlap between the four experiments where cells were harvested 30 minutes post-GccF treatment. In particular, most of the DEGs seen in the 20 nM GccF treatment occur within the GlcNAc pre-treated sample, and all of the DEGs within these samples are seen in the 250 nM treatment. The σ^{54} knockout strain shares a number of DEGs with the other three 30 minute samples, but also contains a number of exclusive DEGs.

Additionally, the genes commonly expressed in all six experiments can be found in Supplementary Table A.6.

4.3.4.1 KEGG pathway analysis

The gene IDs from each experiment were submitted to the DAVID database, and their associated KEGG pathways were extracted. The KEGG pathways highlighted in each experiment are shown in Figure 4.9. The ABC transporters are strongly down-regulated in all experiments, and KEGG pathways involving PTS systems was strongly down-regulated in all experiments aside from the σ^{54} knockout strain, VI01 (Figure 4.9 (a)). Proteins involved in purine metabolism are also down-regulated at 30 minutes post-GccF addition. The core pathways that are down-regulated following addition of 20 nM GccF include those involved in the conversion of guanine to xanthine, adenine to hypoxanthine, and guanosine monophosphate (GMP) to inosine monophosphate (IMP). In the 250 nM GccF treatment, this list expands to include proteins involved in the conversion of guanosine to guanine, xanthosine to xanthine, and guanine/ GMP and xanthine/ XMP (xanthosine monophosphate). However, these down-regulated pathways all disappear in the VI01 treatment, again suggesting the involvement of sigma factor σ^{54}.

Among the up-regulated KEGG pathways, there is a strong response by the purine metabolism pathway, as well as oxidative phosphorylation and ABC transporters under all experimental conditions. The up-regulated purine metabolism genes remain fairly consistent over the 30 minute experiments, and do not show loss of regulation in sample VI01. These pathways include adenosine to inosine conversion, IMP to XMP conversion, and deoxyadenosine and deoxyinsoine interconversion. Likewise, the up-regulated ABC transporter and oxidative phosphorylation components do not lose their differential regulation in the VI01 strain, suggesting that the response by these pathways is not governed by σ^{54}. The up-regulated ABC transporters largely appear to be associated with export of toxins from the cell. At both low and high GccF concentrations, there is an up-regulation of almost all subunits of the V-type ATPases, and at high GccF concentration the F-type ATPases.

(a) Down-regulated pathways

(b) Up-regulated pathways

Figure 4.9: **Heatmap of KEGG pathways**
Heatmaps of number of genes represented in KEGG pathways for down-regulated lists (a) and up-regulated lists (b). From these heatmaps it can be seen that a number of pathways involving ABC transporter components are strongly up-regulated and down-regulated under all conditions. Components of PTS transporter pathways are strongly down-regulated in all treatment conditions apart from the σ^{54} knockout. Conversely, components involved in Oxidative Phosphorylation pathways are strongly up-regulated under all treatment conditions and time points tested.

4.3.4.2 Comparison to other transcriptional studies

A microarray analysis of *Bacillus halodurans* treated with sublancin, which is believed to utilise a glucose-specific PTS, highlighted a few genes associated with GlcNAc processing being up-regulated, including a GlcNAc-specific PTS transporter[72]. The analysis also revealed down-regulation of a number of glucose processing enzymes, and critically, the histidine-containing phosphocarrier protein, HPr, which is central to the PTS system[58]. However, none of the GccF-treated samples from either of the two RNA-seq experiments showed evidence of HPr or GlcNAc-specific PTS components being differentially expressed, suggesting that GccF and sublancin probably elicit different transcriptional responses to cells.

One of the reasons for selecting *E. faecalis* JH2-2 to study the transcriptomic response to GccF was to allow for comparison to the transcriptomic response when treated with other antibiotics, to see if GccF is targeting similar processes, and to rule out general stress response pathways. The RNA-seq data from the work was compared to transcriptomic profiles from two papers studying the response of *E. faecalis* to other antibiotics. Aakra *et al.* (2005)[2] studied the response of *E. faecalis* V583 to erythromycin over different time points, ranging from 0 to 90 minutes post-treatment, and Abranches *et al.* (2013)[3] studied the response of *E. faecalis* OG1RF to cell wall-targeting antibiotics ampicillin, cephalothin, bacitracin and vancomycin at 30 and 60 minutes post-treatment. Both papers assessed the transcriptome using microarrays. Erythromycin is considered a bacteriostatic antibiotic, due to its ability to block protein translation, and so provided useful comparison. Table 4.9 shows the number of DEGs overlapping between the RNA-seq samples collected from this experiment, and the response seen following antibiotic treatment from either of the two papers. Genes were considered to be expressed 'in common' if the gene in question was observed to be differentially expressed in the same way as found in this work at one or more time points in the paper.

From Table 4.9, it appears that the antibiotic which resulted in the lowest overlap of transcriptional response with GccF was ampicillin, while both bacitracin and vancomycin elicited responses with a large degree of overlap with GccF. The main genes that overlapped with bacitracin and vancomycin treatment involved down-regulated PTS components, and down-regulated amino acid ABC transporters. Interestingly, most overlapping down-regulated genes disappeared in the VI01 treatment, indicating that for bacitracin, vancomycin and erythromycin, a σ^{54}-mediated stress re-

Table 4.9: Antibiotic comparison table

GccF experiment	**Amp**†		**Bac**†		**Cep**†		**Van**†		**Ert**‡	
	↓	↑	↓	↑	↓	↑	↓	↑	↓	↑
250 nM, 0 min	2	10	17	23	5	21	13	19	9	6
250 nM, 10 min	3	12	25	43	9	30	32	30	26	14
250 nM, 30 min	7	12	52	46	15	34	68	43	49	25
20 nM, 30 min	1	7	10	10	2	17	11	11	12	5
20 nM (GlcNAc), 30 min	1	9	24	16	6	22	36	17	22	8
100 nM (VI01), 30 min	0	7	0	10	1	14	2	11	2	5

The number of genes commonly up- or down-regulated between GccF treatments (left) and antibiotic treatments (top) are highlighted
†Data taken from Abranches *et al.* (2013)[3].
‡Data taken from Aakra *et al.* (2005)[2].

sponse is being triggered as with GccF treatment (Table 4.9). It is particularly interesting that there was such a small overlap between the erythromycin treatment and the σ^{54} knockout, as erythromycin is generally considered a bacteriostatic agent itself. Common up-regulated genes between GccF treatment and antibiotic treatment included most hypothetical proteins and multidrug resistance proteins, as well as a few ABC transporters predicted to be involved in toxin export.

The RNA-seq gene lists were filtered to remove genes that were similarly affected (up- or down-regulated) in one or more of the antibiotic treatments. Table 4.10 shows the change in number of up- or down-regulated genes following removal of similar overlap. In some experiments, up to 50 % of the DEGs were removed, while the σ^{54} knockout strain was the least affected, most likely due to a large number of the genes being removed from other samples being regulated by σ^{54}. Additionally, this filtering was also carried out on the list of genes that were 'exclusively' affected in all six treatments (see Supplementary Table A.6). The number of genes down-regulated was reduced to 6, and up-regulated genes were reduced to 15 (Supplementary Table A.7). What is particularly interesting is the presence of 6 up-regulated V-type ABC transporters following this filtering. *E. faecalis* encodes 9 of these ABC transporters (ORFs EF1492 - 1500 in V583), which are involved in pumping ions across the cell membrane at the cost of ATP. They have been found to be up-regulated in *E. faecalis* under bile salt conditions[144], and to a lesser degree, under chloramphenicol

Table 4.10: Expression changes following removal of antibiotic overlap

Experiment	**Down-regulated genes**	**Up-regulated genes**
250 nM GccF, 0 min	107 → 76	99 → 65
250 nM GccF, 10 min	164 → 102	252 → 183
250 nM GccF, 30 min	346 → 219	371 → 274
20 nM GccF, 30 min	51 → 26	42 → 20
20 nM GccF + GlcNAc, 30 min	101 → 51	108 → 74
100 nM GccF, VI01, 30 min	92 → 88	186 → 143

Arrows indicate the decrease in gene number following removal of overlapping antibiotic responses

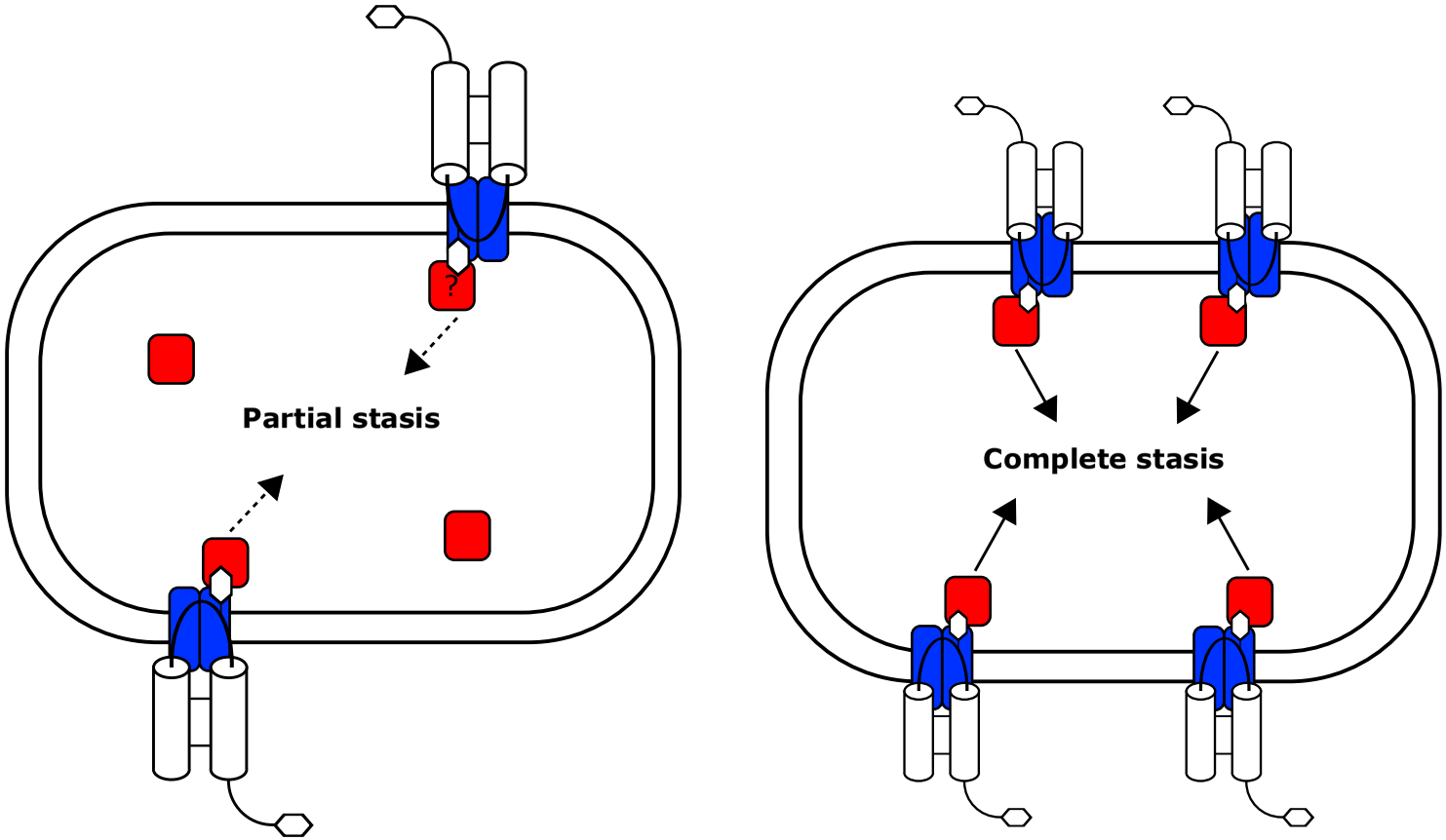
treatment[1]. In particular, the up-regulation of V-type ATPases has been attributed to maintaining the proton-motive force in cells, suggesting that cells are reacting to a loss or destabilisation of the proton motive force[144]. It is interesting that at least 8 of these 9 genes are up-regulated under low GccF concentration (20 nM), and rapidly following exposure to GccF (0 minute time point).

In addition to the V-type ATPases, a number of transporters involved in ion transport are up-regulated in all GccF treatment, including an Na^+/ H^+ antiporter, a sodium/ dicarboxylate symporter, magnesium translocating ATPase, and potassium uptake protein (Supplementary Table A.6). Under alkaline conditions, *E. faecalis* has been seen to down-regulate Na^+/ H^+ antiporters, presumably due to loss of a proton motive force due to increased extracellular pH[131]. Taken together, these observations suggest that the up-regulation of V-type ATPases, as well as other ion-transporting proteins, are part of a key response to GccF, and could suggest a slight permeabillisation or leakage of the cell membrane or another channel resulting in increased accumulation of Na^+ or H^+. Without a clearer understanding of the primary target of this bacteriocin, however, it is impossible to know if these responses are a critical clue to its mechanism of action or merely a secondary response.

4.4 Conclusion

The transcriptional response of *E. faecalis* JH2-2 to glycocin F sheds some light on the mechanism of action of this bacteriocin, although additional experimental evidence (such as identifying the primary target of GccF) is required to completely understand these changes. One interesting aspect about the relationship between *E. faecalis* and GccF is the propensity for free GlcNAc to both protect the cells from GccF, yet also amplify the affect of GccF (Figure 4.4). This effect has been previously reported for strains of *E. faecalis* treated with GccF[92], but the reason for it remained unknown. RNA-seq carried out on *E. faecalis* JH2-2 cells treated with 500 μM GlcNAc and harvested after 10 minutes resulted in only 8 genes being down-regulated and 11 genes being up-regulated (as calculated by *DESeq2* with an adjusted p-value < 0.05); of the up-regulated genes, one encoded a GlcNAc-specific PTS transporter (ERT23545) and one encoded a beta-glucoside-specific PTS transporter (ERT21383; Table 4.3). The up-regulation of a GlcNAc-PTS is compelling, and suggests that the GlcNAc-PTS is involved in the activity of GccF, particularly considering pre-existing evidence of a role in GccF activity against *Lb. plantarum*[62], and a suggested role of the glucose-specific PTS for sublancin[72]. However, it is curious that prior to GlcNAc pre-treatment, *E. faecalis* cells are not seen entering a dormant state despite high concentrations (up to 500 nM) GccF being supplied. The phenotypic effect of this GlcNAc pre-treatment, combined with the up-regulation of the GlcNAc-PTS following GlcNAc exposure, could suggest that the enhanced effect of GccF could be due to an increase in the ratio of PTS numbers compared to another, presumably intracellular, protein that GccF and the PTS interacts with (Figure 4.10).

Transcriptional analysis of *Bacillus halodurans* treated with sublancin showed a down-regulation of the central PTS component HPr, as well as proteins and transporters involved in GlcNAc and glucose metabolism[72]. Similar responses were not seen in *E. faecalis* treated with GccF, which suggests that both glycocins have different mechanisms of action. This could be due to both glycocins targeting different receptors, and may provide additional evidence regarding the different cellular effects each bacteriocin exhibits (*i.e.* bactericidal for sublancin and bacteriostatic for GccF). In fact, the regulation of beta-glucoside PTS components (such as GlcNAc-importer) and glucose-specific components are regulated differently within the cell. The up-regulation of beta-glucoside-specific PTS transporters is controlled by spe-



(a) Low numbers of GlcNAc-PTS transporters

(b) High numbers of GlcNAc-PTS transporters

Figure 4.10: **Possible mechanism of GlcNAc pre-sensitisation affecting GccF susceptibility**
The role of GlcNAc-PTS copy numbers on the strength of GccF-induced stasis may be dependent on a ratio between the number of membrane-bound EIIC domains (blue) and an internal cytosolic protein (red) which forms a complex with GccF and the GlcNAc-PTS. (a) Under basal conditions, when there is a low ratio of EIIC domains to internal protein, partial bacteriostasis is induced by the recruitment of only some of the cytosolic protein. (b) When there are at least equal numbers of PTS and cytosolic proteins, the binding of all GlcNAc-PTS copies by GccF, and subsequent recruitment of all copies of the mystery protein into complexes, results in complete bacteriostasis. EIIB and EIIA domains of the PTS not shown for simplicity.

cific, upstream antiterminator proteins, which are activated following dephosphorylation by the influx of their specific PTS sugar. The PTS components of glucose and mannose-type sugars, on the other hand, are regulated through the phosphorylation status of HPr[57].

The down-regulation of four different PTS encoding operons 30 minutes after GccF treatment suggested that repression of these transporters may indirectly play a role on the bacteriostatic effect of GccF. These operons were also found to be under the control of σ^{54}, which has been highlighted previously as important to susceptibility of Gram-positive bacteria to Class IIa bacteriocins[52]. Interestingly, the σ^{54} knockout strain, *E. faecalis* VI01, was not resistant to GccF (Figure 4.7), indicating that this sigma factor is not crucial for targeting by the bacteriocin. RNA-seq carried out on VI01 treated with GccF did reveal a loss of down-regulation of the previously-seen down-regulated PTS operons, which confirms the role σ^{54} in regulating the expression of this subset of PTS components[69]. One of the PTS components controlled by σ^{54} is the mannose-specific PTS, which is the confirmed target of Class IIa bacteriocins[94]; the loss of regulation of this transporter explains the inherited resistance to Class IIa bacteriocins that bacteria receive.

Genes associated with ABC transporters were strongly differentially regulated at all time points, and genes associated with oxidative phosphorylation were strongly up-regulated at all time points (Figure 4.9). Looking at the up-regulated genes corresponding to ABC transporters and oxidative phosphorylation, there is a number of transporters involved in responding to osmotic stress (such as ion-specific ABC transporters; see Supplementary Table A.6), and a number of V-type ATPases, which have been shown to be up-regulated in response to osmotic stress conditions[144]. The strong up-regulation of these transporters under even low concentrations of GccF suggest that their regulation may be part of a specific response to GccF, which may be either a result of the direct mechanism that GccF exhibits, or a side-effect. Additionally, this same pattern of regulation is not seen with any of the other antibiotic transcriptional profiles of *E. faecalis*, which suggests an effect that is somewhat unique to GccF.

There are a couple of factors that this experimental design did not take into account, which would serve to help deconvolute the data. One such factor is the lack of a naturally static control in any or all of the RNA-seq samples. Recent work on dormant forms of *B. subtilis*, which include transcriptomic studies, compared long-

term dormant forms of the bacteria with both log-phase growing cells, and lag phase 'naturally dormant' cells[80]. The inclusion of lag-phase cells in this study allowed the authors to differentiate transcriptional changes that occur as part of natural nutrient limitation-induced stasis, and transcriptional changes that were specific to the condition they were interested in[125]. Including such a control in any of these experiments would presumably have helped to filter similar responses in this experiment, and allow for more the specific GccF-induced transcriptional changes to be narrowed down. Another factor involves the use of spin columns during the RNA extraction step (Section 4.2.1), which potentially would not bind small RNAs. This could mean that small RNAs which were involved in regulation of specific genes would not have been sequenced and are thus missing from this data. While the lack of small RNAs should not change the number of genes that are identified in each experiment, such information would potentially be useful for identifying regulatory pathways.

Given the breadth of genes differentially expressed, particularly at higher concentrations of GccF, without further evidence of a specific cellular target for this bacteriocin it is impossible to precisely say what response is a result of this glycocin's effect on cells, and what is a secondary response. These RNA-seq experiments were intended to assist in understanding the precise mechanism that GccF employs in initiate bacteriostasis in cells, but thus far has provided information without context. It may be that once the primary target of GccF has been discovered, the responses detailed here may make more sense, and reveal the true mechanistic effect of this bacteriocin.

5 | Investigating the receptor of glycocin F

5.1 Introduction

The activities of the chemically synthesised analogues of GccF provided additional evidence of the key role played by the peptide structure of the interhelical loop in addition to the essential involvement of the GlcNAc on Ser18. They also hinted at the possible roles played by the glycosylated tail. As analysis of the transcriptional response of *E. faecalis* JH2-2 to GccF also strongly indicated a role for the GlcNAc-PTS transporter in target susceptibility, a number of questions about the specific processes involved in the mechanism used by GccF to induce stasis remain. One important aspect of GccF activity which had so far only been identified by genetic engineering was the specific cellular protein(s) that act as target(s) or receptor(s) for GccF. This was observed using immunoprecipitation combined with mass spectrometry, with two different approaches. The first approach involved implementing pulldowns using a tagged form of the GccF immunity protein, GccH, which had been identified from gene expression studies on the GccF-producing strain, *Lb. plantarum* KW30[5]. This approach echoed work done by Diep *et al.* who identified the primary target of the Class II bacteriocin, lactococcin A, using immunoprecipitation of the flag-tagged lactococcin A immunity protein[59]; the experiment here involved immunoprecipitation of the cell lysate from GccF-susceptible cells expressing tagged immunity protein under bacteriocin treated conditions. The second approach involved crosslinking chemically-synthesised, biotin-labelled GccF analogues to susceptible *Lb. plantarum* cells, and pulling down conjugated proteins using streptavidin-conjugated magnetic beads. The GccF analogues contained diazirine-conjugated Cys-GlcNAcs ('GlcNDAz') at either position 18 or 43, allowing for photo-crosslinking to nearby amino acids upon exposure to ultraviolet (UV) light. This approach echoed experiments done by Yu and others[176] who used *in vivo* addition of GlcNDaz to study the binding partners of O-GlcNAc-modified proteins in mammalian cells. The results of both approaches are detailed in this chapter.

Table 5.1: Plasmids used in this work

Plasmid	Description	Source
pNZ8148	Plasmid containing nisin-inducible promoter, for expression in *Lactobacillus* sp.	[115]
pNZ8148_HA-gccH	pNZ8148 modified to express HA-tagged GccH	This work
pNZ9530	Encodes two-component system *nisK* and *nisR* to enable expression of pNZ8148 in *Lactobacilli*	[95]
pProExHTb_gccH	*E. coli* expression vector, expression GccH	Kind gift from Mr Trevor Loo
pProExHTb_HA-*gccH*	pProExHTb_*gccH* with N-terminal HA-tag on GccH	This work
pRV613	Copper inducible promoter, for expression in *Lactobacilli*	[51]
pRV613_*gccH*	pRV613 expression GccH	Kind gift from Mr Trevor Loo
pRV613_flag-*gccH*	pRV613_*gccH* with N-terminal flag-tag	This work
pRV613_fL-*gccH*	pRV613_flag-*gccH* with 4 amino acid linker between flag-tag and GccH	This work
pRV613_HA-*gccH*	pRV613-*gccH* with N-terminal HA-tag	This work

5.2 Experimental procedures

5.2.1 Site-directed mutagenesis of plasmids

The principle behind site-directed, ligase-independent mutagenesis (SLIM) is outlined in Chiu *et al.* (2008)[40] and summarsed in Supplementary Figure B.1, along with the general procedure used in this work. Plasmids pRV613_flag-*gccH*, pRV613_fL-*gccH*, pRV613_HA-*gccH* and pProExHTb-_His_6-HA-*gccH* were all created using SLIM modification from parent plasmids pRV613_*gccH* or pProExHTb_*gccH* (Table 5.1). Briefly, the plasmid to be mutated was amplified in two PCR reactions (*Phusion* polymerase (ThermoFisher Scientific)) using pairs of primers that diverged from the point of mutation, with one primer of each pair having a 5´ 'tail' containing the mutation (insertion, deletion, or substitution) of interest (see Table 5.2 for primer list). Following this initial amplification two double-stranded, linearised versions of the plasmid were formed, each containing the mutation at opposing ends. These were then subjected to *Dpn*I digest at 37 ℃ for 60 minutes (10 units *Dpn*I in D-buffer [20 mM $MgCl_2$, 20 mM Tris pH 8.0, 5 mM DTT]) to remove any residual template DNA. Hybridisation reactions were then prepared, involving the addition of equal amounts of both linearised plasmids in the presence of H-buffer (150 mM NaCl, 25 mM Tris pH 9.0, 20 mM EDTA). Hybridisation was promoted by cycling through a series of heating and cooling steps (denatured at 99 ℃ followed by three cycles of 65 ℃ for 5 minutes and 30 ℃ for 40 minutes). The hybridisation products

were then transformed into a cloning strain of *E. coli* (such as XL-1 Blue; see Section 2.2.2) to repair nicks in the plasmid and amplify it, before plasmid extraction and sequencing.

Table 5.2: List of primers used in this work

Primer name	**Sequence**	**T_m**
pRV613_FLAG_F_T[a]	5′ GAT TAT AAG GAT GAC GAT GAC AAG TTG AAA CAC GAT GGG ATC CTT	51.1 ℃
pRV613_FLAG_R_T[a]	5′ CAA CTT GTC ATC GTC ATC CTT ATA ATC CAT CGT TAT TCC TCC TTT	46.3 ℃
pRV613_FLAG_F_nT[a,c]	5′ AAA CAC GAT GGG ATC CTT	51.1 ℃
pRV613_FLAG_R_nT[a,c]	5′ CAT CGT TAT TCC TCC TTT	46.3 ℃
FLAG_linker_F_T[b]	5′ GGT GCC GGA TCT AAA CAC GAT GGG ATC CTT ACT	53.6 ℃
FLAG_linker_R_T[b]	5′ AGA TCC GGC ACC CTT GTC ATC GTC ATC CTT ATA ATC C	53.2 ℃
FLAG_linker_F_nT[b]	5′ AAA CAC GAT GGG ATC CTT ACT	53.6 ℃
FLAG_linker_R_nT[b]	5′ CTT GTC ATC GTC ATC CTT ATA ATC C	53.2 ℃
pRV613_HA_F_T[c]	5′ TAC CCA TAC GAT GTT CCA GAT TAC GCT AAA CAC GAT GGG ATC CTT	51.1 ℃
pRV613_HA_R_T[c]	5′ AGC GTA ATC TGG AAC ATC GTA TGG GTA CAT CGT TAT TCC TCC TTT	46.3 ℃
pProEx-HA_F_T[d]	5′ TAC CCA TAC GAT GTT CCA GAT TAC GCT ACC GAA AAC CTG TAT TTT CAG GGC GC	61.7 ℃
pProEx-HA_R_T[d]	5′ AGC GTA ATC TGG AAC ATC GTA TGG GTA TGG GTA ATC GTG ATG GTG ATG GTG ATG GTA GTA	61.1 ℃
pProEx-HA_F_nT[d]	5′ ACC GAA AAC CTG TAT TTT CAG GGC GC	61.7 ℃
pProEx-HA_R_nT[d]	5′ TGG GTA ATC GTG ATG GTG ATG GTG ATG GTA GTA	61.1 ℃
HA-GccH_F_NcoI[e]	5′ AGC T **CC ATG G**TA TAC CCA TAC GAT GTT CCA GAT TAC	54 ℃
HA-GccH_R_HindIII[e]	5′ AGC T **AA GCT T**CT AAG CAT CAT ATT GAA ATA ATT TTA TTT TAA TTG CAT	54 ℃
pRV613_GccH_F[f]	5′ CAA CCG TTC CTG AGA TTG AAG CGA TG	59.8 ℃
pRV613_GccH_R[f]	5′ CTT ACT GTT AGG AAA CTT CAC GTA CTT AGC TTT GTT	59.6 ℃
pNZ8148_seq_F[g]	5′ CCG CTA TCT TTA CAG GTA CAT CAT	54.3 ℃
pNZ8148_seq_R[g]	5′ GGG TCG TAG TAC GGA AGC	54.9 ℃
pNZ9530_nisR_F[h]	5′ GAC AGC ATT AGA AAT GAG AAA CTA TG	52.2 ℃
pNZ9530_nisK_R[h]	5′ CCT TGA TTT AAA GCT ACA CC	52.2 ℃

The underlined regions represent 'tail' regions of SLIM primers
[a]Used in addition of FLAG-tag to GccH
[b]Used in addition of FLAG-tag linker to GccH
[c]Used in addition of N-terminal HA-tag to GccH in pRV613_*gccH*
[d]Used in addition of N-terminal HA-tag to GccH in pProExHTb_*gccH*
[e]Used in cloning of *HA-gccH* into pNZ8148
[f]Used for sequencing of *gccH* gene
[g]Used to confirm presence of pNZ8148
[h]Used to confirm presence of pNZ9530

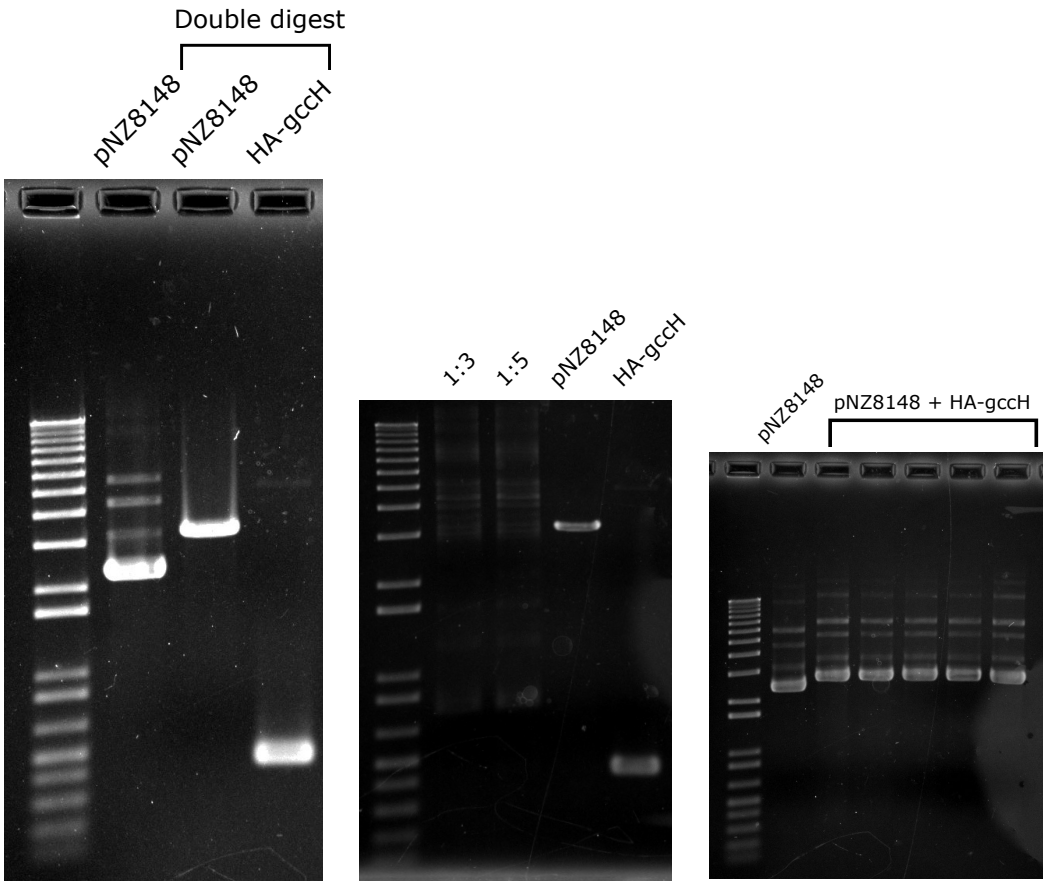
5.2.2 Cloning of plasmid pNZ8148

The *HA-gccH* gene was amplified from plasmid pRV613_HA-*gccH* using primer set *e* from Table 5.2, which introduced *Nco*I and *Hin*dIII restriction sites. Following this, plasmid pNZ8148 and amplified *HA-gccH* were double-digested with *Nco*I and *Hin*dIII (Figure 5.1 (a)), purified by ethanol precipitation, and ligated together in a vector:insert ration of 1:3 and 1:5. Based on the 'laddering pattern' observed following ligation (Figure 5.1 (b)), the 1:3 vector:insert ratio product was transformed into *E. coli* XL-1 Blue cells by heat-shock (see Section 2.2.2). Plasmid was extracted from five transformed colonies, and the resulting plasmid was confirmed to be larger than unmodified pNZ8148 (Figure 5.1 (c)). The correct insertion of the *HA-gccH* gene into pNZ8148 was confirmed by sequencing using primer set *g* from Table 5.2 (Supplementary Figure B.8).

5.2.3 Expression of tagged GccH

5.2.3.1 Copper-induced expression

Copper-induced expression of flag-tagged GccH (flag-GccH), flag-linker-tagged GccH (fL-GccH) and HA-tagged GccH (HA-GccH) was all carried out using plasmid pRV613, which contains a copper-inducible promoter[51]. Parent plasmid pRV613_*gccH* was modified using site-directed mutagenesis to add the three respective labels to the N-terminus of GccH (see Section 5.2.1), resulting in plasmid pRV613_flag-*gccH*, pRV613_fL-*gccH*, and pRV613_HA-*gccH*. Each respective plasmid was transformed into the plasmidless *Lb. plantarum* strain NC8 by electroporation using the transformation protocol outlined in Section 2.2.2. Expression was induced by diluting overnight cultures of transformed cells 100-fold in fresh MRS media containing a final concentration of 50 μM $CuSO_4$. Cells were grown to a final OD_{600} of at least 0.5 before being diluted for liquid protection assays, or being pelleted and lysed for western blotting and expression analysis. All cells transformed with pRV613_*gccH* or derivatives were cultured in the presence of 50 μg/mL erythromycin to select for the plasmid.



(a) Restriction digests (b) Ligation (c) Confirmed plasmid

Figure 5.1: **Cloning steps of pNZ8148_HA-*gccH***
(a) Plasmid pNZ8148 and *HA-gccH* amplicon double-digested with *Nco*I and *Hin*dIII. (b) Ligation reactions set up between double-digested vector (pNZ8148) and insert (*HA-gccH*) at respective ratios of 1:3 and 1:5. The laddering pattern in lanes 1:3 and 1:5 indicate activity of ligase. (c) Mobility shifts of empty plasmid pNZ8148 (labelled) and extracted ligation plasmids extracted from transformed XL-1 Blue cells. The decrease in mobility suggests that the plasmids all contain inserts.

5.2.3.2 Nisin-induced expression

Nisin-induced expression of HA-tagged GccH was carried out using the two plasmids pNZ8148 and pNZ9530. pNZ8148 (Supplementary Figure B.7) contains a nisin-inducible promoter which responds to the two-component system proteins encoded by *nisR* and *nisK* located on pNZ9530 (Supplementary Figure B.6). Following cloning of *HA-gccH* into pNZ8148 (resulting in pNZ8148_HA-*gccH*; see Section 5.2.2), *Lb. plantarum* NC8 was sequentially transformed with plasmids pNZ9530 and pNZ8148_HA-*gccH*. Overnight cultures of transformed cells were diluted 100-fold into fresh MRS supplemented with the required antibiotics (50 μg/mL erythromycin, 25 μg/mL chloramphenicol). When an appropriate OD_{600} was reached, expression of *HA-gccH* was induced by addition of 25 ng/mL nisin for at least 1 hour prior to addition of GccF or collection of cells for detection by western blotting.

5.2.4 Western blot procedures

Proteins were separated using SDS-PAGE on a 15 % acrylamide gel. Following this, they were transferred onto a 0.2 μm polyvinylidene difluoride (PVDF) membrane using a BioRad blotting module of the Mini Protean III electrophoresis apparatus (Bio-Rad). Transfer was carried out at 100V for 90 minutes in Towbin buffer (25 mM tris, 192 mM glycine, 20 % [v/v] methanol). Following transfer, the membrane was blocked by incubating in TBS-T (Tris-buffered saline + 0.1 % Tween-20) containing 5 % (w/v) skim milk for one hour at room temperature. The membrane was then washed five times in TBS-T (five minutes each wash), before being incubated with the primary antibody (anti-flag monoclonal M2 (Sigma-Aldrich), anti-flag polyclonal (Sigma-Aldrich), or anti-HA monoclonal (ThermoFisher Scientific)) at a dilution of 1:10,000 in TBS-T + 5 % skim milk for 1 - 6 hours at room temperature. The wash step was then repeated, followed by incubation of the membrane with secondary antibody (horse radish peroxidase (HRP)-conjugated anti-rabbit antibodies for flag-tag detection, or peroxidase-conjugate anti-mouse antibodies for HA-tag detection) at a dilution of 1:20,000 in TBS-T + 5 % skim milk overnight at 4 °C. The membrane was then washed five times in TBS-T, followed by a single wash in TBS, before being briefly blotted dry and developed using Clarity^{TM} Western ECL substrate (Bio-Rad) for approximately 90s. The membrane was imaged using an Azure c600 imager (ThermoFisher Scientific) in 10x 30s increments.

5.2.5 Protein immunoprecipitation

anti-Flag pulldown

Approximately 5 mL of *Lb. plantarum* NC8 cells expressing flag-tagged or fL-tagged GccH protein were briefly pelleted by centrifugation at 13,000 x g, and resuspended in 1 mL TBS-T. Cells were lysed by bead milling with 0.1 mm diameter zirconium beads (Sigma-Aldrich) in a Hybaid Ribolyser (3x 30s runs, 5,000 rpm, 4 ℃). 20 μL of anti-flag-conjugated magnetic beads (Sigma-Aldrich) were pre-washed in TBS, and added to the cell lysate, which was then incubated overnight at 4 ℃ on a rotating platform. Following this, the supernatant was removed, the beads were washed three times with three 400 μL volumes of TBS, and the protein eluted from beads by incubating twice (5 minutes each) with 100 μL volumes of 0.2 % (v/v) formic acid. The eluted fractions were neutralised (as confirmed by pH indicator strips) with 40 μL of 0.5 M ammonium bicarbonate (ABC) pH 7.9, before the volume was reduced to approximately 40 μL using vacuum centrifugation. Samples were stored at -80 ℃ prior to treatment with trypsin. Alternatively, elution was carried out by boiling the beads briefly in 2x SDS loading buffer.

Anti-HA pulldown

Approximately 5 mL of *Lb. plantarum* or *E. coli* cells expressing HA-tagged protein were pelleted and lysed as described above. 2 μL of biotin-conjugated, anti-HA antibodies (1 mg/mL concentration; ThermoFisher Scientific) were added to cell lysate, and incubated overnight at 4 ℃ with tumbling. Following this, 20 μL streptavidin-coated magnetic beads (ThermoFisher Scientific) were added to the cell lysate, and incubated at 4 ℃ for a further 6 hours. Streptavidin-coated beads were washed three times with 400 μL TBS, and protein eluted by incubating twice in 100 μL 1 % (v/v) formic acid, 5 minutes each. Acid was neutralised by addition of 160 μL freshly made 0.5 M ammonium bicarbonate, and volume reduced to $\sim$ 40 μL before being stored at -80 ℃ prior to trypsin digest.

Photo-crosslinking and pulldown using diazirine-conjugated GccF

Lb. plantarum NC8 was cultured overnight in chemically-defined media (CDM)[163] supplemented with 2 % (w/v) glucose. This strain was specifically chosen as the GccH pulldowns had been carried out in this plasmidless strain, and would allow for more direct comparisons between the two pulldown experiments. 1 mL of overnight culture was resuspended in 14 mL of fresh CDM, and left to recover for 1 hour at 30 ℃. GlcNDAz-containing GccF ($GccF_{Cys18GlcNDAz}$ or $GccF_{Cys43GlcNDAz}$) was then added to the cells at a final concentration of 250 nM, cells were incubated at 30 ℃ for 20 minutes, and cells were transferred to a standard 10 cm diameter petri dish. Cells were then irradiated with broad-beam 365 nm UV light for 30 minutes, before being pelleted by centrifugation (3000 x g, 10 minutes) and lysed in 1 mL of TBS buffer by bead milling using the same conditions as those used for the GccH pulldowns. Detergent was not added prior to bead milling, in an attempt to capture evidence of larger membrane-bound complexes. The cell lysate was then incubated overnight with 20 μL of pre-washed streptavidin-conjugated magnetic beads at 4 ℃ with tumbling. The beads were then washed with 3x 400 μL volumes of TBS-T and 1x 400 μL volume of TBS, before being resuspended in 40 μL of 50 mM ABC buffer and subjected to in-solution tryptic digest (see next section).

Photo-conjugated GccF analogues $GccF_{Cys18GlcNDAz}$ and $GccF_{Cys43GlcNDAz}$ were chemically synthesised by the University of Auckland Peptide Synthesis Facility, using GlcNDAz-Cys building blocks supplied by Dr. Matthew Pratt of the University of Souther California. The chemical synthesis involved a solid-phase peptide synthesis approach, followed by native chemical ligation, similar to that described in Bisset *et al.* 2018[24].

5.2.6 Proteomic analysis

In-solution tryptic digest

Protein-containing solutions were decreased to a volume of around 40 μL. Proteins were reduced by incubating with 4 μL of 200 mM dithiothreitol (DTT) for 1 hour at 37 ℃, alkylated with 4 μL of 100 mM iodoacetamide for 30 minutes at room temperature in the dark, and reduced again by adding 4 μL of 200 mM DTT. Trypsin (Sigma-Aldrich) was added to a final concentration of 1 μg, and incubated

overnight at 37 ℃. Following this, the volume was decreased to approximately 40 μL by vacuum centrifugation, and the remaining solution was clarified by centrifugation at 13,000 x g for 15 minutes prior to analysis by mass spectrometry.

Mass spectrometry analysis

LC-MS Analysis

The resulting peptides from in-gel digestion were separated by reversed-phase chromatography using a nanoLC system with an in-line reversed-phase peptide trap for desalting/ enrichment and a 50 cm reversed-phase analytical column for peptide separation. The liquid chromatography system was coupled to an Q Exactive Plus mass spectrometer equipped with a higher-energy collision-induced dissociation (HCD) collision cell, an Orbitrap mass analyser and a Nano Flex ion source (ThermoFisher Scientific, Waltham, MA USA). Details for the chromatographic and mass spectrometric settings are listed in Tables 5.3 and 5.4.

Table 5.3: Mass spectrometry instrument configuration

nanoLC system	Dionex Ultimate 3000 Binary RSLCnano LC system (ThermoFisher Scientific)
Mass spectrometer	QExactive Plus (ThermoFisher Scientific)
Ionization source	Nano Flex (ThermoFisher Scientific)
Trapping column	PepMap100 C18, 3 μm particle size, 75 μm inner diameter, 2 cm length (ThermoFisher Scientific)
Analytical column	PepMap C18, 2 μm particle size, 75 μm inner diameter, 50 cm length (ThermoFisher Scientific)
Flow rates	Trap: 15 μL/min Analytical: 300 nL/min
Column oven temperature	50 ℃
Gradient	3-30 % Acetonitrile in 0.1 % formic acid over 60 minutes + regen
Buffers	Trap loading: 0.1 % TFA/ 2 % MeCN/ water Analytical A: 0.1 % Formic acid/ 2 % MeCN/ water Analytical B: 0.1 % Formic acid/ 98 % MeCN/ water

The MS-Top10 method was used to collect data; a survey scan is performed and the top 10 most intense peptide ions within the specified range are fragmented to obtain sequence information. The process was repeated every scan cycle in the chromatographic run, and additional settings enabled repeating ions within a specified time window to be ignored, according to observed chromatographic peak width (typically 10-30 s). This allows more low abundance ions to be targeted and improve overall coverage.

Table 5.4: Mass spectrometer settings

Capillary temperature [°C]	250
S-Lens RF level [%]	50
Polarity	Positive
Source voltage [kV]	1.5
AGC target	Full MS 3e6, MS2 (HCD) 1e5
Max. injection times [ms]	Full MS 100 ms, MS2 (HCD) 110 ms
Full MS mass range	370 – 1,600 [m/z]
Resolution settings	Full MS 70,000, MS2 (HCD) 17,500
No. of microscans	1
Isolation width [m/z]	1.4
Loop count (TopN)	10
MSX count	1
Collision energy	27
Charge Exclusion	Unassigned, 1, >8
Peptide match	Preferred
Exclude isotopes	On
Dynamic Exclusion [s]	25
Spectrum data type	Centroid

Database Search

The raw data files were searched using the Proteome Discoverer™ v. 2.2 search engine (ThermoFisher Scientific, Waltham, MA USA). The search parameters used matched the specifications of the Q Exactive instrument, and variables resulting from chemical treatment and suspected natural modifications of proteins were listed in Table 5.5. The *Lactobacillus plantarum* NCBI database was used with the peptide sequence for GccH added, and in samples where the tryptic digest was carried out on the beads, the sequence for streptavidin from *Streptomyces avidinii*, accession number P22629, was included. The reverse database search option was enabled by default, and all protein matches were filtered to satisfy a false discovery rate (FDR) of 1 % or better and have at least 2 unique peptides.

Table 5.5: Proteome Discoverer search settings

Search engine	Proteome Discoverer v 2.2
Databases	Non-redundant compiled on May 2018 + contaminant + custom construct sequences
Taxonomy	*Lb. plantarum* NC8
Enzyme	Trypsin
Max # of missed cleavages	2
Min peptide length	6
Precursor mass tolerance	10 ppm
Fragment mass tolerance	0.02 Da
Decoy database search	Enabled (default)
Static modifications	Carbamidomethyl (C)
Variable modifications	Oxidation (M), N-terminal acetylation, Deamination (N/Q)
Experimental mass values	Monoisotopic
False Discovery Rate (FDR)	Set to 1 %

5.3 Results and discussion

5.3.1 GccH immunoprecipitation

The primary bacterial target for the bacteriocin lactococcin A, produced by *Lactococcus lactis*, was found by studying the interactions between the immunity protein for lactococcin A, LciA, and the bacteriocin itself. This was done by expressing a flag-tagged variant of the immunity protein in susceptible cells, and studying the complex that was formed in the presence of exogenous lactococcin A [59]. The isolation of the bacteriocin·immunity protein·receptor complex allowed the authors to identify the mannose-specific EIIC and EIID domains as being the primary receptors of lactococcin A. Curiously, the authors noted there was no LciA·receptor interaction in the absence of bacteriocin, and the mechanism used to promote cell lysis still remains unknown [59].

Gene expression analysis of the *gcc* gene cluster during glycocin F production in *Lb. plantarum* KW30 led to the identification of a candidate immunity protein, GccH [5]. GccH is a 118 amino acid (13.2 kDa) protein, with no predicted domains or similarity to any characterised protein. Interestingly, however, a BLASTP analysis of the *Lb. plantarum* genome using the GccH amino acid sequence as a search query revealed that the N-terminus of GccH shares some sequence similarity (38 %

identity) with the N-terminus of a hypothetical protein which appears downstream of the *pts18CBA* locus. This similarity only occurs over the first 42 amino acids of both proteins, and it is currently unknown whether this similarity has any impact on the activity of function of GccH, or if this is merely a coincidence.

More recently, cloning and expression of this gene in a susceptible strain of *Lb. plantarum* using a copper-inducible expression system confirmed that GccH did in fact provide immunity to GccF[17]. Using this expression vector, attempts were made to confirm the primary target or receptor of GccF using an approach similar to that used by Diep *et al*[59]. Initial experiments involved trialling an N-terminal flag-tagged GccH, but this was eventually abandoned in favour of the more specific HA-tag. Initial trials involved the copper-inducible plasmid, pRV613[51], but weak and leaky expression led to this being replaced by a nisin-induced expression system (NICE) vector[115], which provided much higher and controllable levels of protein expression.

5.3.1.1 Tagging GccH

Attachment of flag-tag to GccH

The immunity protein immunoprecipitation carried out by Diep *et al.* first involved adding the flag-tag sequence (DYKDDDDK) to the N-terminus of the immunity protein, LciA[59]. This same approach was trialled with GccH, using the already constructed *Lactobacillus*-specific copper inducible plasmid pRV613_*gccH* (Supplementary Figure B.2 (a))[17]. The peptide sequence used by Diep *et al.* (DYKDDDDKL) was inserted into pRV613_*gccH* between the codons for Met1 and Lys2 of the plasmid-bound *gccH* sequence using site-directed, ligase-independent mutagenesis (SLIM; see Section 5.2.1 for details) using primer set *a* from Table 5.2. The insertion of the correct sequence was confirmed by sequencing (Supplementary Figure B.2 (b)), and the new plasmid was named pRV613_flag-*gccH*.

Plasmid pRV613_flag-*gccH* was transformed into *Lb. plantarum* NC8 (see Section 2.2.2) for expression and further testing. This strain was chosen for all further expression tests as it contains no native plasmids, minimizing the chance of plasmid rejection or recombination with native plasmids[14]. More importantly, the strain is susceptible to GccF, with a reported IC_{50} of approximately 40 nM[17]. Culturing transformed NC8 in the presence of copper (50 μM) and 10x the IC_{50} concentration

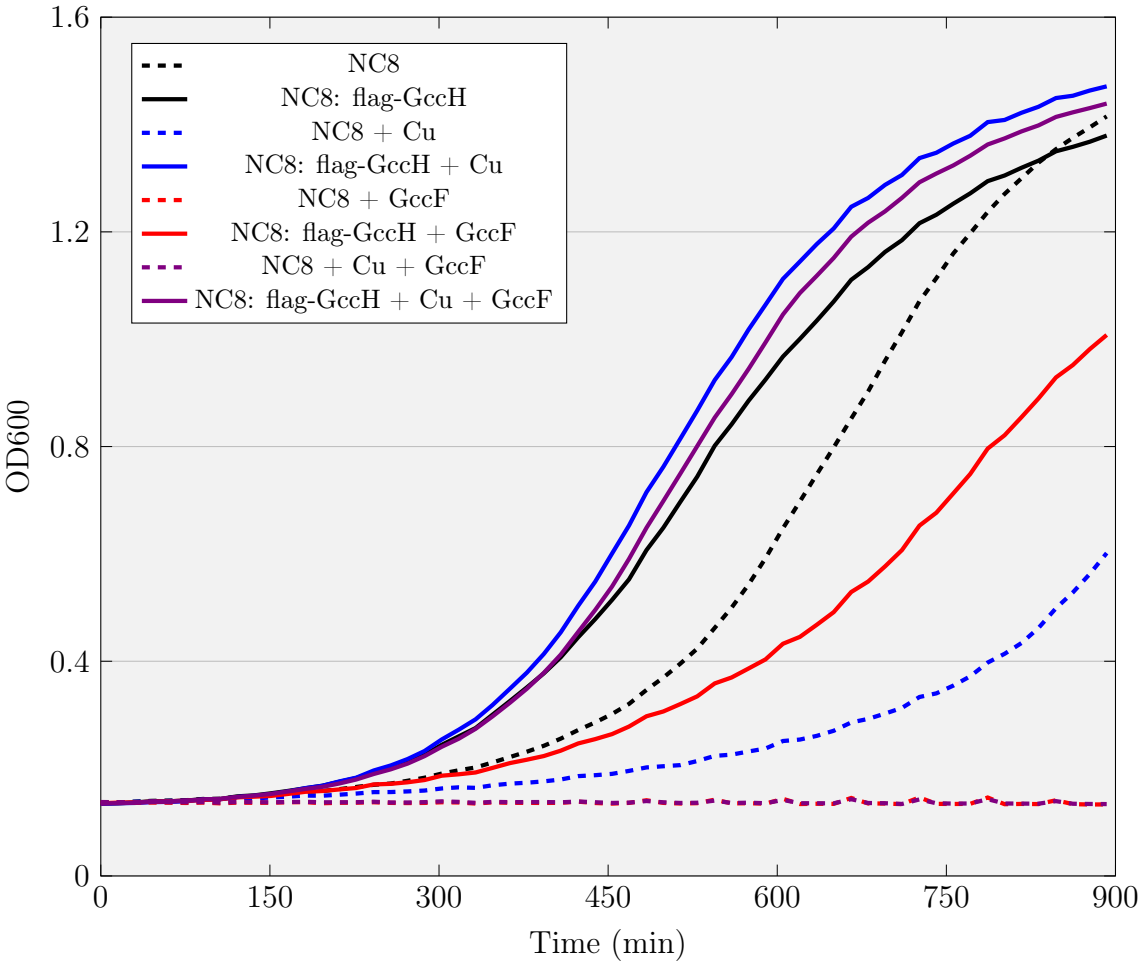
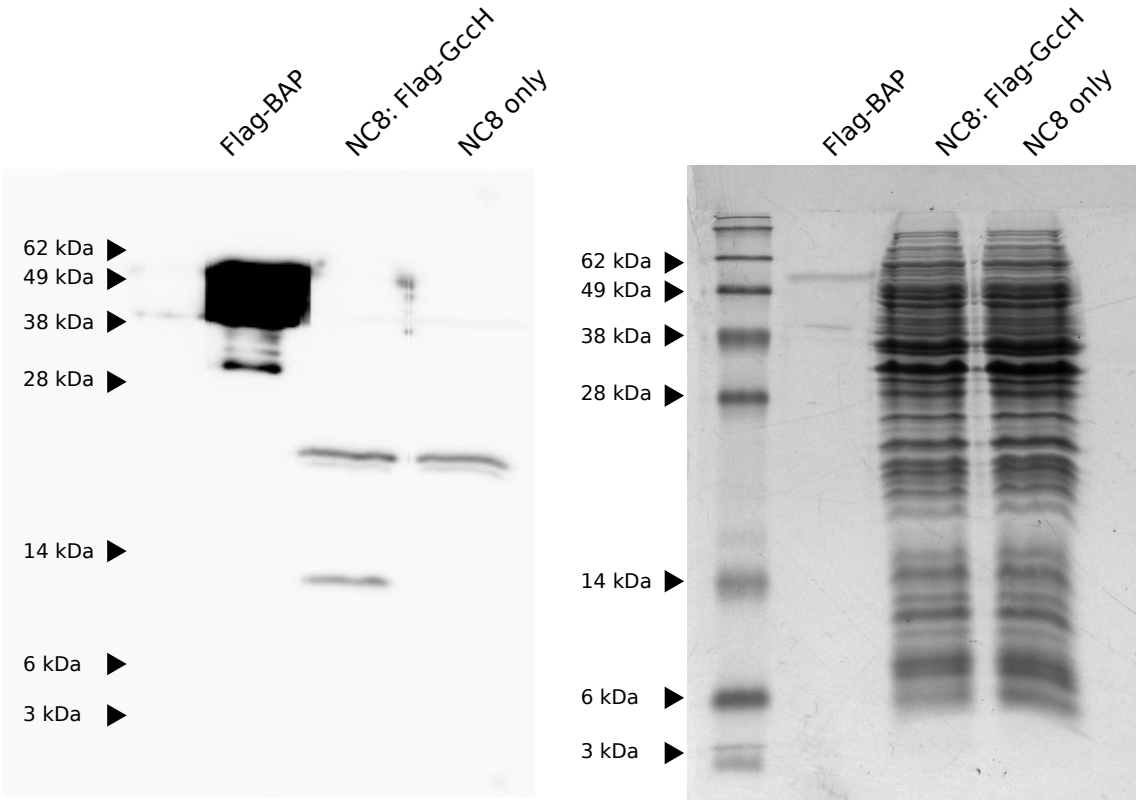


Figure 5.2: **NC8 flag-tagged GccH complementation**
Dashed lines show the growth of untransformed *Lb. plantarum* NC8 cells, while solid lines represent growth of NC8 cells harbouring plasmid pRV613_flag-*gccH*. The dashed red and purple lines ('NC8 + GccF' and 'NC8 + Cu + GccF') showed no growth, and both overlap at the bottom of the graph. Copper was added at a final concentration of 50 μM, and GccF was added at a final concentration of 400 nM (10x the IC_{50} value). Both copper and GccF were added at time 0. Plots shown are the average of three parallel runs. Also, note that the OD_{600} values are not absolute; lack of zeroing on the data puts the lowest reading at around 0.12 units regardless of cell density.



(a) Western blot

(b) Colloidal Coomassie stained gel

Figure 5.3: **Anti-flag western blot of Flag-tagged GccH**
Cell lysates of *Lb. plantarum* NC8 containing pRV613_flag-*gccH* or no vector were cultured overnight in MRS media supplemented with 50 μM $CuSO_4$, and cell lysates were run on a 15 % acrylamide SDS-PAGE gel. (a) Western blot of cell lysates probed with polyclonal anti-flag antibodies (Sigma-Aldrich). 0.5 μg N-terminal flag-BAP protein (Sigma-Aldrich) was included as a positive control. A lower band can be seen in the NC8: flag-GccH lane, which corresponds to roughly the expected molecular weight of flag-GccH (14.3 kDa). However, a non-specific band originating from the NC8 lysate can also be seen at around 20 kDa. (b) Colloidal coomassie stained gel of cell lysate, showing approximately equal loading of both untransformed and transformed NC8 lysate.

of GccF (400 nM) showed that the presence of an N-terminal tag on GccH did not adversely affect its activity, as its expression still protected cells from the effects of GccF (Figure 5.2).

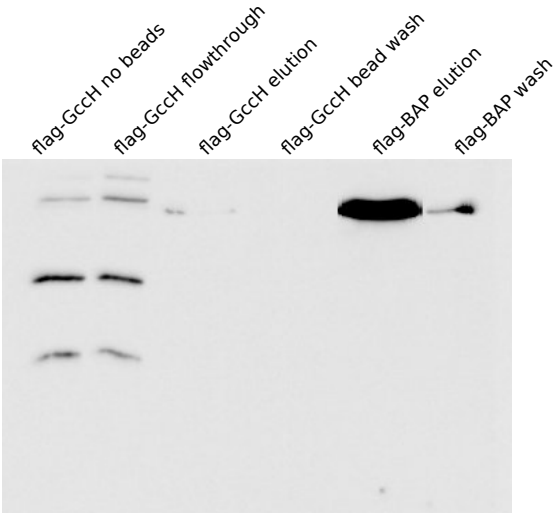


Figure 5.4: **Anti-flag magnetic bead pulldown**
Cell lysate of *Lb. plantarum* NC8 cells transformed with pRV613_flag-GccH were incubated overnight with magnetic beads conjugated with anti-flag M2 monoclonal antibodies (Sigma). Beads were washed with TBS, and bound protein eluted by boiling the beads in 2x SDS loading buffer. Samples were analysed by SDS-PAGE (15 % acrylamide gel), then blotted onto a PVDF membrane and probed with anti-flag polyclonal antibodies, which were subsequently detected using HRP-conjugated anti-rabbit secondary antibodies. From the western, it can be seen that there are no bands corresponding to flag-GccH in either the wash or eluted from the beads, in contrast to lysate alone ('flag-GccH no beads') and the flag-GccH flowthrough. N-terminally-tagged flag-BAP was successfully eluted off the beads, however, and a faint band can be seen in the flag-BAP wash lane.

Western blotting was used to confirm the expression of flag-tagged GccH, as well as ensure the tag was accessible to antibodies for downstream immunoprecipitation. Western blot detection carried out with polyclonal anti-flag antibodies confirmed the expression of flag-GccH, as a band was visible in the lysate of NC8 cells transformed with pRV613_flag-*gccH* that was not visible in the lysate of untransformed NC8 (Figure 5.3 (a)). It is also evident from this figure that the polyclonal antibodies react non-specifically with other proteins in *Lb. plantarum* NC8. Unfortunately, the expression of flag-GccH by pRV613_flag-*gccH* in *Lb. plantarum* NC8 cells was not very strong, as a corresponding band in the Coomassie-stained protein gel was not visible (Figure 5.3 (b)), preventing confirmation of band identity by mass spectrometry. It should also be noted here that the predicted molecular weight of all

GccH derivatives assessed throughout this work were expected to be around 14 kDa in size. The molecular weight markers used in the western blots were pre-stained, and were all confirmed to run at a slightly lower mobility than expected (data not shown).

Following on from this, anti-flag pulldowns were attempted. Cell lysate from NC8 cells transformed with pRV613_flag-*gccH* and grown in the presence of 50 μM $CuSO_4$ were incubated overnight with anti-flag M2-conjugated magnetic beads, washed and eluted by boiling in 2x SDS, as described in Section 5.2.5. This was then analysed by SDS-PAGE and western blot, with N-terminal flag-BAP protein (bacterial alkaline phosphatase, Sigma-Aldrich) included as a positive control. From Figure 5.4 it can be seen that there are no bands corresponding to flag-GccH in either of the wash or elute lanes of the transformed *Lb. plantarum* NC8 cell lysate, indicating that the flag-GccH pulldown was unsuccessful. As the flag-BAP was successfully pulled down, the reagents and the procedure used would appear to not be at fault. One possible reason for this lack of interaction is lack of recognition or binding by the flag-tag epitope of flag-GccH by the M2 antibodies. These antibodies are monoclonal, may not be able to recognise the tag due to its position on or within GccH, although without a structure, this is cannot be confirmed. To test this hypothesis a western blot of NC8: pRV613_flag-*gccH* cell lysate was carried out using M2 anti-flag antibodies (Figure 5.5). From the blot it is apparent that the lower band attributed to flag-GccH is no longer present, confirming that the flag-tag on the flag-GccH construct is not recognised by this antibody. It is possible that the flag epitope was not being recognised because it is too close to the protein. To test this hypothesis, a linker of four amino acids was introduced between the flag-tag and the GccH sequence, with care being taken not to change the charge or introduce any secondary structure. To this end, small neutral amino acids (Gly, Ala and Ser) were used, and the sequence 'GAGS' introduced.

Insertion of linker in Flag-GccH

Insertion of the linker was done using SLIM methodology, using primer set *b* from Table 5.2 (Supplementary Figure B.3). The new plasmid, dubbed pRV613_fL-*gccH*, retained its ability to protect *Lb. plantarum* NC8 cells from GccF (Figure 5.6). Western blotting was used to show that NC8 cells transformed with pRV613_fL-*gccH* were producing the correct protein (Figure 5.7 (a)). Additionally, incubation

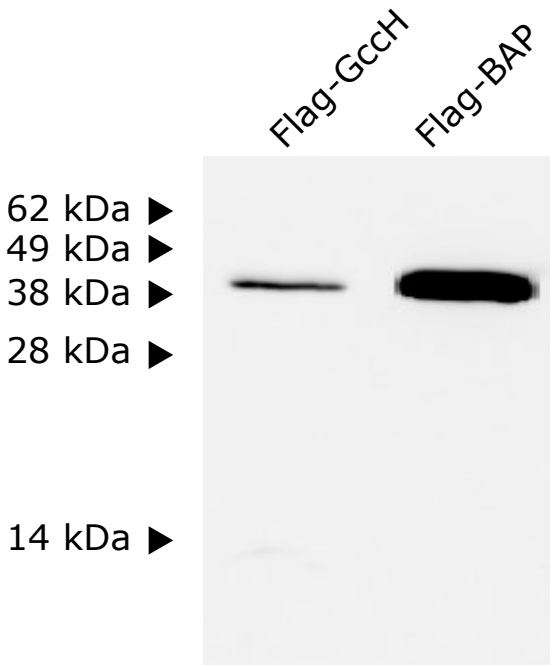


Figure 5.5: **Flag_GccH western blot with monoclonal antibodies**
Lb. plantarum NC8 cells transformed with pRV613_flag-*gccH* were cultured overnight in the presence of 50 μM $CuSO_4$. Cell lysate was analysed by SDS-PAGE, then blotted onto a PVDF membrane and probed with monoclonal anti-flag M2 antibodies. The lack of detection of lower-weight bands corresponding to flag-GccH (as observed in Figure 5.3) indicates that the M2 monoclonal antibodies are incapable of recognising the methionine-capped N-terminal flag fusion but not flag-GccH. A thin band corresponding to the same molecular weight as the flag-BAP protein is visible, and is most likely an accidental overflow of this protein during loading.

of the cell lysate with M2 conjugated magnetic beads resulted in GccH being pulled down as evidenced by the presence of the lower molecular weight band on the western blot (Figure 5.7 (b)). Unfortunately, two other NC8 proteins were detected by these antibodies. While it is strange these did not appear in previous experiments, it is most likely due to non-specific binding. Another possibility is that there is a specific interaction between this unknown protein and GccH, and non-specific binding of the antibody to it. Because detection of non-specific bands is not ideal, another tag was trialled in place of the flag-linker.

Attachment of HA-tag to GccH

The so-called 'HA-tag' (YPYDVPDYA), taken from amino acids 98 - 106 of the human influenza haemagglutinin protein, was inserted at the N-terminal end of *gccH* in the of pRV613_*gccH* plasmid. SLIM methodology was used, and there was no linker sequence between the tag and the *gccH* coding sequence (primer set *c*, Table 5.2; Supplementary Figure B.4). The resulting plasmid, dubbed pRV613_HA-*gccH*, produced an active protein when transformed into *Lb. plantarum* NC8 cells (Figure 5.8). Additionally, western blots carried out on NC8 cells producing this tagged protein using anti-HA monoclonal antibodies (ThermoFisher Scientific) showed only a single band, indicating no non-specific interactions (Figure 5.9 (a)). As a result, downstream work was simplified as the same antibody used for detection could also be used for pulldowns, because the monoclonal anti-HA antibody could also be obtained as a biotin conjugate. Following incubation of cell lysate with biotin-conjugated anti-HA antibodies, antibody-tagged proteins can be selectively removed from solution through incubation with magnetic beads conjugated to streptavidin (ThermoFisher Scientific). Figure 5.9 (b) shows the result of a pulldown using the combination of biotin-conjugating antibodies and streptavidin-conjugated magnetic beads, eluted under acid conditions.

HA-tagged GccH was immunoprecipitated from cell lysate, and the eluted proteins were digested using trypsin, and analysed using mass spectrometry (Section 5.2.6). Untransformed *Lb. plantarum* NC8 cells were processed identically and used as a negative control. However, there were few differences between the proteins identified in the untransformed control and the induced, HA-GccH-expressing cells, despite expression being confirmed by parallel western blots (data not shown). Almost all of the proteins identified by the mass spectrometry belonged to ribosome subunits

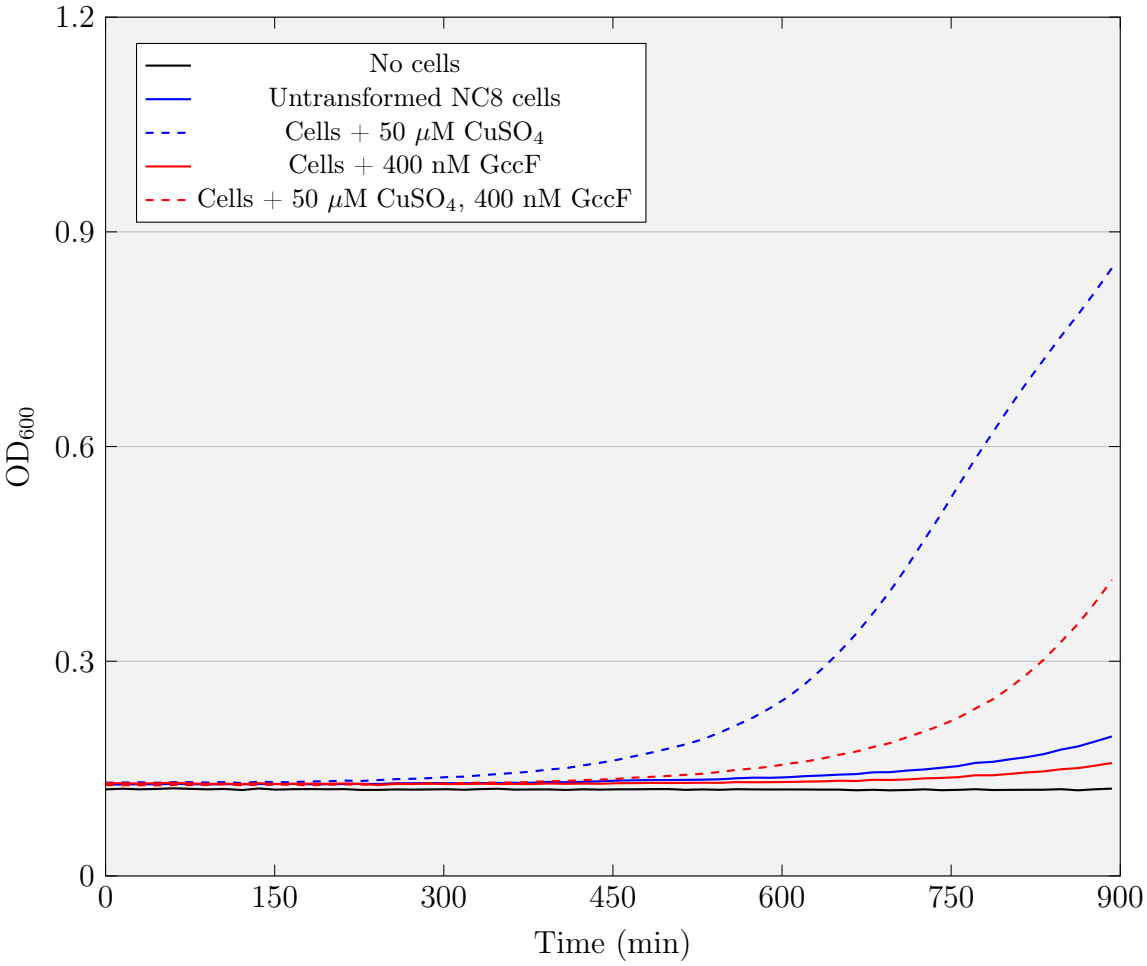
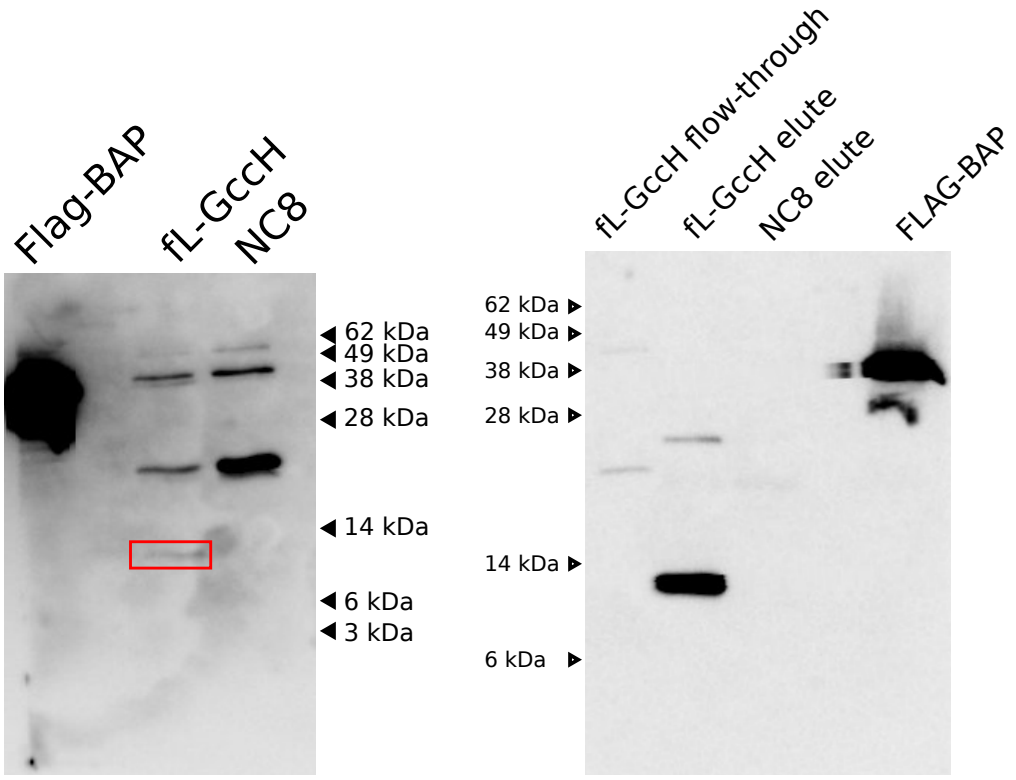


Figure 5.6: **Growth plots showing activity of fL-GccH**
Lb. plantarum NC8 cells transformed with pRV613_fL-*gccH* were tested for resistance to GccF. Cells were treated with 10x the IC_{50} concentration of GccF (400 nM) at time 0, with or without 50 μM $CuSO_4$. GccF can be seen to inhibit cell growth in the absence of copper, although the long lag-phase of these cells leads to negligible growth in the untransformed, untreated NC8 cells (solid blue trace). In the presence of copper, both GccF-treated and untreated cells exhibit cell growth. Plots shown are the average of three cultures grown in parallel. Note that the OD_{600} values are not absolute; lack of zeroing on the data puts the lowest reading at around 0.12 units regardless of cell density.



(a) (b)

Figure 5.7: **Anti-flag western blot and pulldown of fL-GccH**
(a) *Lb. plantarum* NC8 cells transformed with pRV613_fL-*gccH* were cultured overnight in MRS supplemented with 50 μM $CuSO_4$, and assessed for fL-GccH expression by western blot using polyclonal anti-flag antibodies. Flag-BAP was included as a positive control, and untransformed NC8 cells were included as a negative control. The presence of a faint band in the fL-GccH lane (highlighted with a red box) indicates the expression of fL-GccH, albeit at low levels. There are at least two non-specific bands from the NC8 proteins reacting with the polyclonal antibodies. (b) *Lb. plantarum* NC8 cells expressing fL-GccH were lysed and incubated overnight with anti-flag M2-conjugated magnetic beads (Sigma-Aldrich) at 4 ℃, before being washed and eluted from the beads by boiling in 2x SDS loading buffer. Two of the three bands from (a) were bound to the beads, and subsequently eluted. One of these bands was the low molecular weight band believed to belong to fL-GccH, indicating that the addition of the linker between the flag-tag and GccH peptide allowed the monoclonal antibodies to recognise and bind to the tag. The lower band also appears more concentrated after elution.

(data not shown). Additionally, no peptide fragments belonging to HA-GccH could be identified in the digest of proteins bound to the beads, an indication of low abundance expression of HA-GccH by pRV613 in NC8 cells. It is possible that because of the low abundance of the HA-GccH, peptides belonging to it were obscured by the more abundant intracellular proteins, which appeared to be non-specifically binding to the streptavidin magnetic beads. To confirm that the level of protein expression was an issue, the experiment was repeated using an *E. coli* expression system transformed with a modified expression vector.

Expression trial in *E. coli*

An inducible *E. coli* vector already containing the *gccH* gene was modified with an HA-tag and subjected to over-expression, followed by anti-HA pulldown. This vector, pProExHTb_*gccH*, was modified using SLIM (using primer set *d*, Table 5.2) to include an HA tag as an N-terminal fusion, directly downstream of the sequence encoding an N-terminal His_6 tag, and upstream of the sequence encoding a TEV protease cleavage site, and was named pProExHTb_His_6-HA-*gccH* (Supplementary Figure B.5).

Expression of HA-GccH was carried out by inducing transformed *E. coli* BL21 cells with 1 mM isopropyl β-D-1-thiogalactopyranoside (ITPG) for approximately 2 hours. Lysed cells were subjected to the same pulldown and elution procedure used previously using pRV613_HA-*gccH* expression as a positive control (Figure 5.10). From the western blot it is clear that HA-GccH was over-expressed in the *E. coli* cells, and the presence of prominent bands in both the flowthrough and wash lanes suggests that HA-GccH production exceeded the antibody/ streptavidin binding capacity (Figure 5.10 (a)). Furthermore, the presence of a visible band of the appropriate size in the post-transferred, coomassie-stained gel (Figure 5.10 (b)) indicated that expression was significantly stronger than that obtained from the pRV613 plasmids. This band was excised and the protein confirmed as His_6-HA-GccH using mass spectrometry (data not shown).

An anti-HA pulldown was repeated using the *E. coli* cells producing His_6-HA-GccH. The eluted proteins were digested with trypsin and identified using mass spectrometry. Uninduced cells were included as a negative control. Supplementary Table C.6 shows the proteins that were identified in both the induced and uninduced samples, while Table 5.6 shows the proteins that were exclusive to the induced sample.

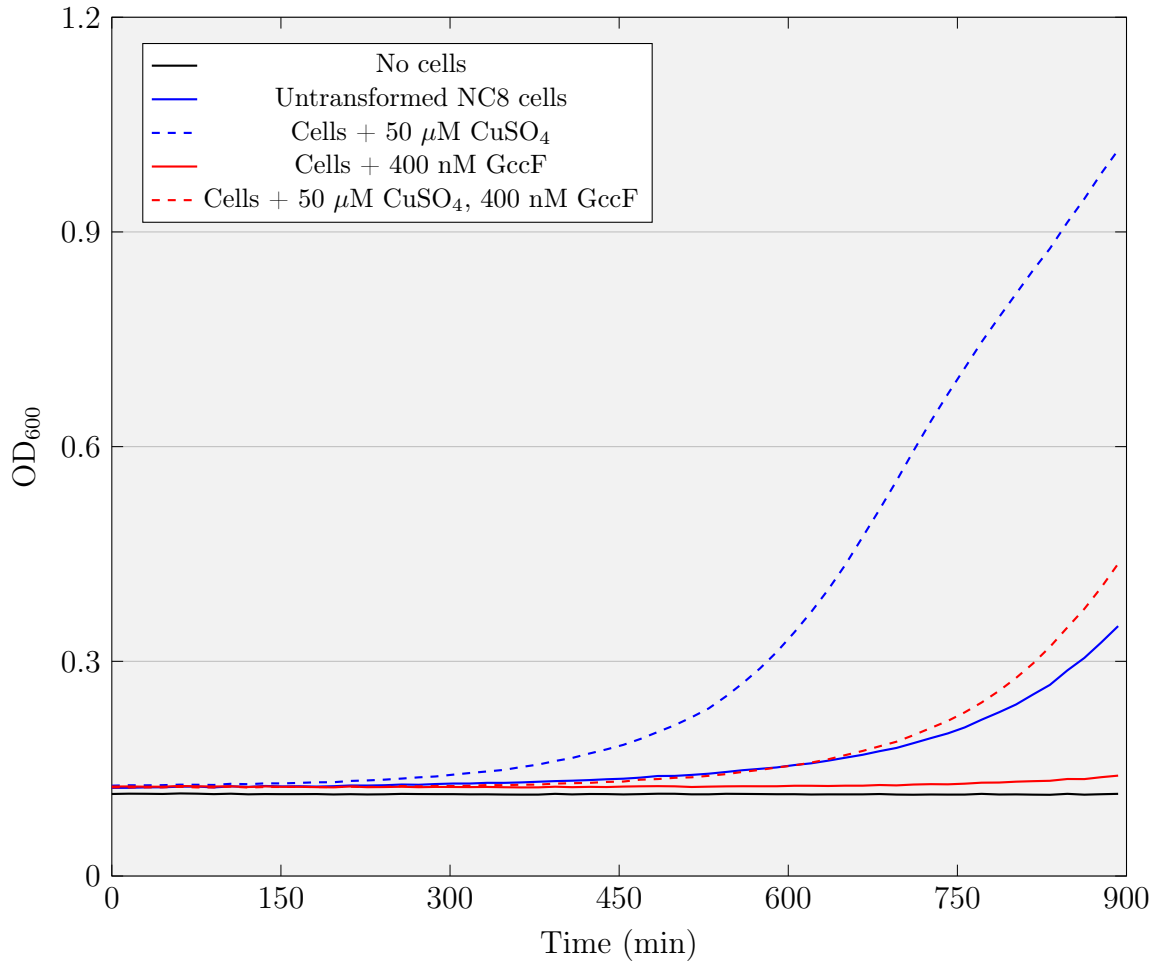
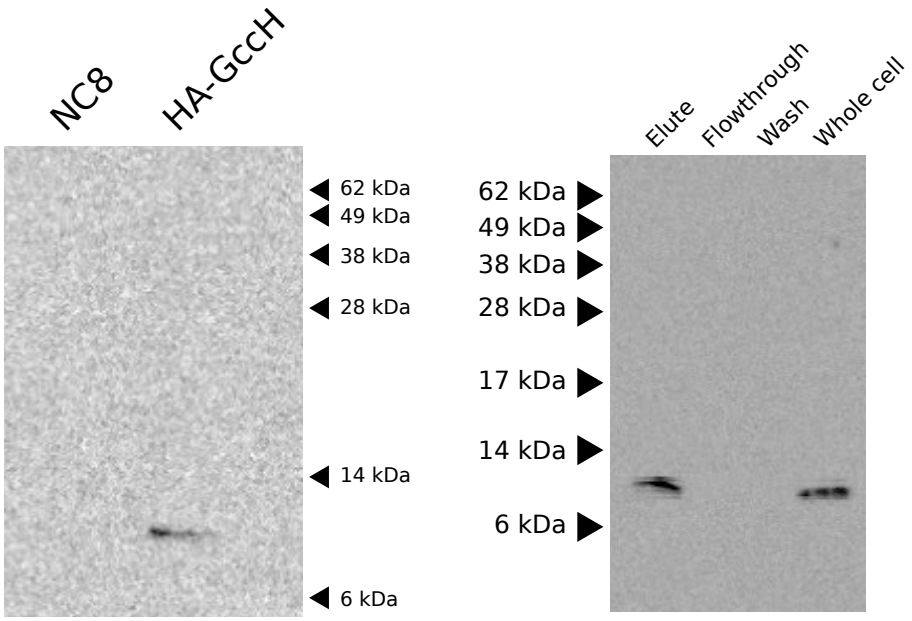
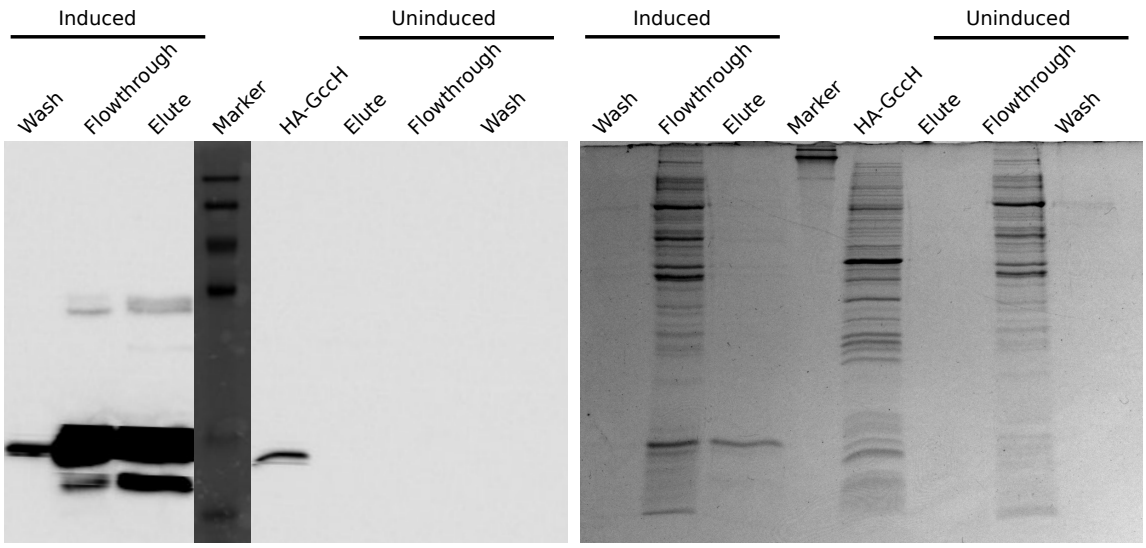


Figure 5.8: **Growth plots showing activity of HA-GccH**
Lb. plantarum NC8 cells transformed with pRV613_HA-GccH were tested for resistance to GccF. Cells were treated with 10x the IC_{50} concentration of GccF (400 nM) where indicated, with or without 50 μM $CuSO_4$. GccF can be seen to inhibit cell growth in the absence of copper, which the presence of copper in both GccF-treated and untreated increases cell growth. Plots shown are the average of three cultures grown in parallel. Note that the OD_{600} values are not absolute; lack of zeroing on the data puts the lowest reading at around 0.12 units regardless of cell density.



(a) HA-GccH western blot (b) HA-GccH pulldown

Figure 5.9: **Western blot and pulldowns of HA-tagged GccH**
(a) Cell lysates of *Lb. plantarum* NC8 wt and NC8 transformed with pRV613_HA-*gccH* plasmid were assessed by western blotting, using anti-HA monoclonal antibodies (ThermoFisher Scientific). From the presence of only a single band in the 'HA-GccH' lane, it appears that there is no non-specific protein detection using the antibodies. (b) HA-GccH-expressing cell lysate was incubated with biotin-conjugated anti-HA antibodies and streptavidin-conjugated magnetic beads. Following washing, proteins were eluted off beads by incubating in 1 % (v/v) formic acid. Whole cell lysate not incubated with antibodies and beads was also included as a control. The presence of a band in the 'elute' lane and no bands in the 'flowthrough' or 'wash' lanes indicates that the HA-GccH pulldown is successful.



(a) Western blot (b) Protein gel

Figure 5.10: **HA-GccH pulldown from *E. coli* BL21**
E. coli BL21 cells transformed with pProExHTb_His$_6$_HA_*gccH* were cultured, and induced and non-induced cells subject to anti-HA pulldowns. (a) anti-HA western blot of gel. Prominent bands can be seen in both the elute and flowthrough lanes of the induced cells, and a faint band of the appropriate size can be seen in the wash lane of the induced cells. The occurrence of second, lower molecular weight bands is unexplained. No bands of appropriate size are visible in the uninduced cells. Induced *Lb. plantarum* NC8 cells transformed with pRV613_HA-*gccH* were included as a positive control and reference ('HA-GccH'). (b) Post-transferred 15 % acrylamide SDS-PAGE gel from (a). A visible band can be seen in the elute lane of the induced cells.

Importantly, peptides corresponding to His$_6$-HA-GccH were identified in both the induced and uninduced pulldown samples. This confirmed that with a high-expression vector, tagged GccH can be identified in a pulldown profile. The identification of numerous ribosomal proteins and chaperones in the cell lysate of most of the pulldown experiments carried out could be significant (Supplementary Table C.6). It is more likely, however, that it is simply due to the high abundance of these proteins in the cells and their ability to bind non-specifically to both streptavidin and anti-flag-conjugated magnetic beads. While *E. coli* is not the native producer of GccF, and GccH cannot be expected to behave as it would in *Lb. plantarum*, it is still interesting to note the presence of at least one PTS component in the His$_6$-HA-GccH pulldown (ACT43846.1, Table 5.6), particularly given the prior evidence that strongly suggests a GlcNAc-PTS plays a significant role in GccF susceptibility[62].

Table 5.6: Exclusive protein matches from induced His_6-HA-GccH pulldown in *E. coli*

Protein	Peptides	Unique	AA Coverage	Description
E. coli His6-HA-GccH†	182	8	49.7%	**His_6 and HA-tagged GccH**
WP_000853869.1	81	24	44.0%	MULTISPECIES: tagatose-bisphosphate aldolase subunit GatZ
ACT41917.1	69	15	26.9%	chaperone Hsp40, DnaK co-chaperone
ACT45364.2	46	7	42.3%	heat shock chaperone
ACT44963.1	28	4	35.8%	50S ribosomal subunit protein L5
ACT43846.1	26	4	34.7%	**galactitol-specific enzyme IIA component of PTS**
ACT43843.1	25	4	11.6%	galactitol-1-phosphate dehydrogenase, Zn-dependent and NAD(P)-binding
Beta-galactosidase	21	4	5.1%	OS=Escherichia coli (strain K12) GN=lacZ PE=1 SV=2 Tax_Id=562
ACT44976.1	20	2	28.2%	30S ribosomal subunit protein S10
CAQ33548.1	13	4	48.5%	predicted IS600 orf-like protein
pdb|3DTM|A	13	4	6.8%	Chain A, Increased Folding Stability Of Tem-1 β-Lactamase
ACT44965.1	12	2	14.6%	50S ribosomal subunit protein L14
WP_000823266.1	11	2	46.8%	**MULTISPECIES: PTS galactitol transporter subunit IIB**
ACT44301.1	10	6	11.3%	protein disaggregation chaperone
ACT42291.1	10	3	38.9%	HU, DNA-binding transcriptional regulator, beta subunit
ACT44314.1	10	3	33.9%	50S ribosomal subunit protein L19
ACT42561.1	10	2	10.4%	dihydrolipoyltranssuccinase
WP_001140433.1	9	3	44.1%	MULTISPECIES: 50S ribosomal protein L30
ACT42558.1	9	2	6.0%	succinate dehydrogenase, flavoprotein subunit
pdb|1HQY|F	8	4	18.0%	Chain F, HEAT SHOCK LOCUS HSLU
WP_001700733.1	8	3	19.4%	MULTISPECIES: translation initiation factor IF-3
ACT45387.1	7	4	15.9%	tryptophanase/L-cysteine desulfhydrase, PLP-dependent
ACT42810.1	7	2	6.5%	30S ribosomal subunit protein S1
ACT43848.1	6	3	19.4%	D-tagatose 1,6-bisphosphate aldolase 2, catalytic subunit
ACT42452.1	5	3	23.0%	alkyl hydroperoxide reductase, C22 subunit
ACT43189.1	5	3	34.5%	lipid hydroperoxide peroxidase
ACT42560.1	5	2	4.4%	2-oxoglutarate decarboxylase, thiamine triphosphate-binding
ACT42980.1	5	2	36.8%	50S ribosomal subunit protein L32
ACT45413.1	5	2	11.3%	F1 sector of membrane-bound ATP synthase, alpha subunit
ACT45454.1	3	2	7.2%	transcription termination factor
WP_001157756.1	3	2	13.4%	MULTISPECIES: two-component system response regulator OmpR
ACT42583.1	2	2	11.6%	phosphoglyceromutase 1
ACT42802.1	2	2	4.9%	formate C-acetyltransferase 1, anaerobic; pyruvate formate-lyase 1

†Appeared in both lists, but was most abundant in the induced pulldown.
Proteins of interest highlighted in boldface.

Trialling the NICE expression system

With the evidence that a higher level of protein expression than that provided by the pRV613 plasmids was detectable by mass spectrometry following pulldown, attempts were made to find a tuneable expression system in *Lb. plantarum*. The nisin-controlled expression system (NICE), which makes use of the autoinducible element of nisin production from *L. lactis*, was chosen as it has been successfully used by other groups to produce recombinant Firmicute proteins[115,100]. The minimum element required for this expression system is a plasmid containing a nisin-inducible promoter (P_{nisA}), along with the two-component response regulator genes *nisK* (encoding a histidine-protein kinase) and *nisR* (encoding a response regulator), which can be either chromosomally-located or plasmid-bound. This system allows for high levels of protein expression when transformed cells are subjected to extremely low concentrations of nisin (in the order of ng/mL). A strain of *L. lactis* (NZ9000; NIZO, Netherlands) was available which contained *nisK* and *nisR* chromosomally located, however this strain proved to not be susceptible to GccF, and so was not used for further studies. Instead, the plasmid pNZ9530 (NIZO, Netherlands) was selected (Supplementary Figure B.6), which contains both the *nisK* and *nisR* genes, and is capable of replicating in *Lb. plantarum*. Likewise, the P_{nisA}-containing plasmid pNZ8148 (NIZO, Netherlands) was chosen for expression (Supplementary Figure B.7), as it is able to replicate in *Lb. plantarum* and does not share the same resistance marker as pNZ9530.

HA-gccH was amplified from plasmid pRV613_HA_gccH using primer set *e* from Table 5.2, which were engineered with *Nco*I and *Hin*dIII restriction sites, allowing for cloning into the multiple cloning site of pNZ8148 (Supplementary Table B.7). Following restriction digest and ligation, the correct insertion of HA-tagged *gccH* was confirmed (Supplementary Figure B.8), and the resulting plasmid was named pNZ8148_HA-*gccH*. Plasmids pNZ9530 and pNZ8148_HA-*gccH* were sequentially transformed into *Lb. plantarum* NC8, and the presence of both plasmids in the cell was confirmed by PCR (Supplementary Figure B.9) using primers designed to amplify each plasmid (primer sets *g* and *h* from Table 5.2). Expression of HA-GccH using the two-plasmid expression system was confirmed by western blot, with an induction by 50 ng/mL nisin appearing to provide GccH production for at least 10 hours (Figure 5.11). Additionally, HA-GccH produced using this system was shown to still provide immunity against GccF (Figure 5.12).

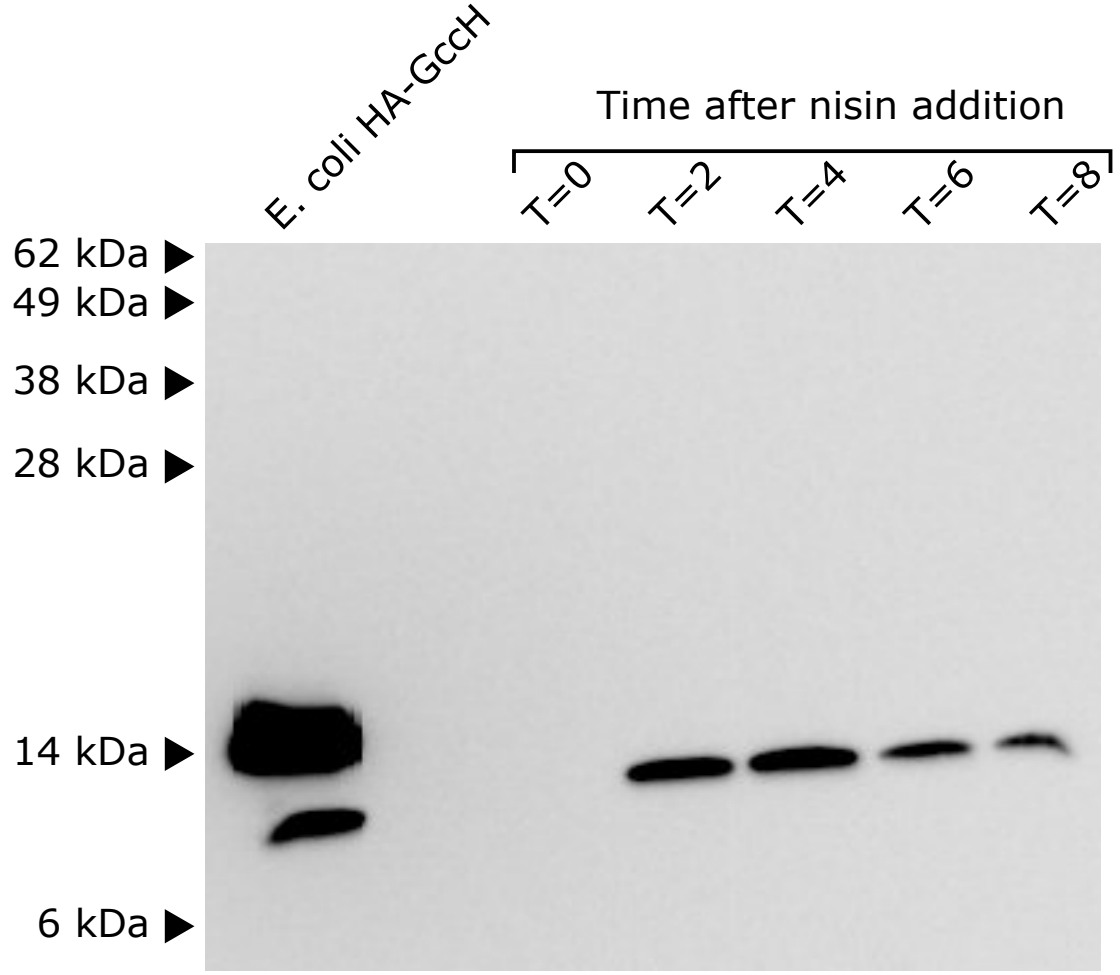


Figure 5.11: **Time course expression of HA-GccH using nisin-inducible system** *Lb. plantarum* NC8 cells transformed with with pNZ9530 and pNZ8148_HA-*gccH* were treated with 50 ng/mL nisin, and harvested after 0, 2, 4, 6 and 8 hours. No expression can be seen at time 0, while expression can be seen at all remaining time points, and maximum expression can be seen at around 2 and 4 hours.

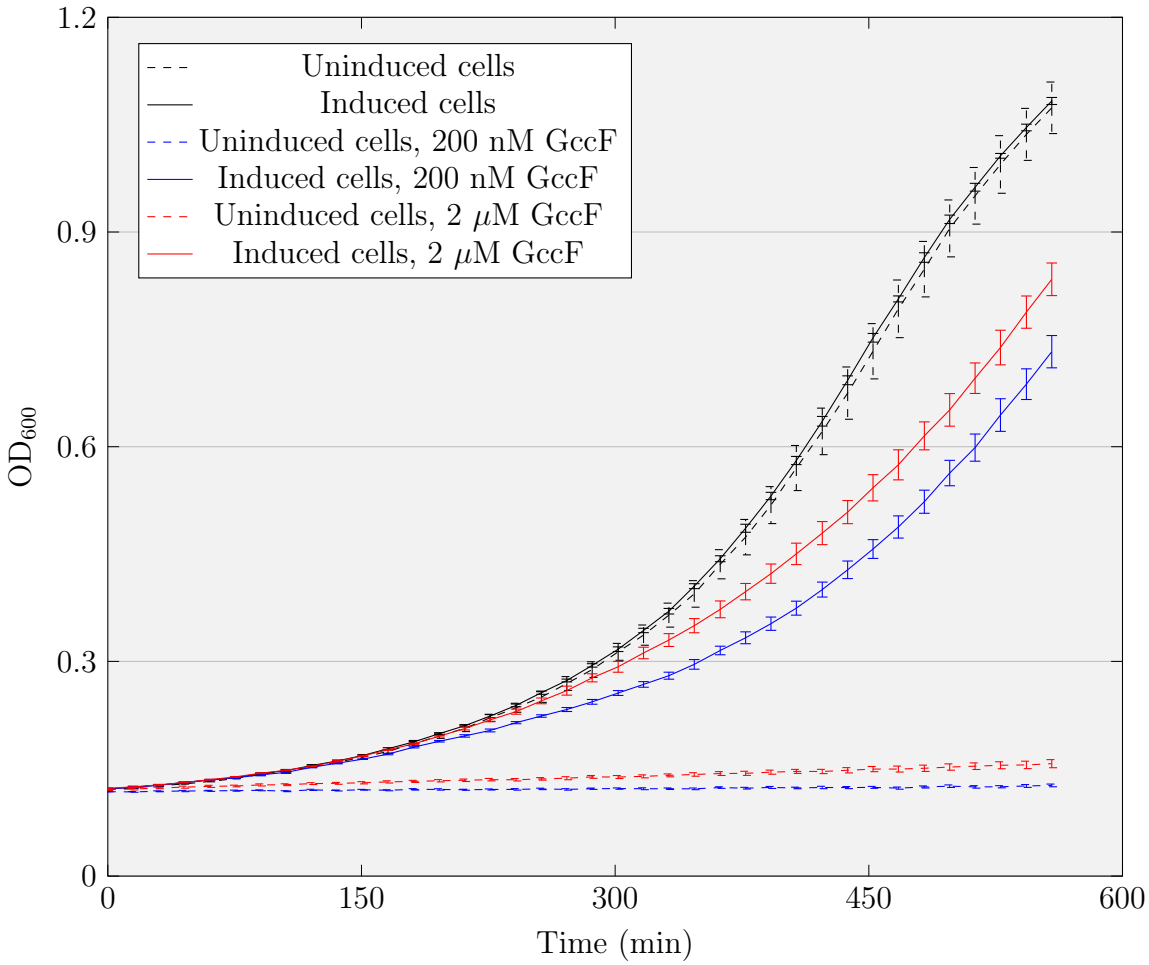


Figure 5.12: **Protective effect of nisin-induced HA-GccH**
Lb. plantarum NC8 cells transformed with both pNZ9530 and pNZ8148_HA-*gccH* were cultured in the presence of 200 nM or 2 μM GccF (added at 0 minutes). Dashed lines show the growth plots of uninduced cells, while solid lines show the growth of induced cells (50 ng/mL nisin, added one hour prior to addition of GccF). Growth curves are the average of three parallel runs. From the plot, it can be seen that induction of HA-GccH using nisin provides resistance against high concentrations of GccF for up to 10 hours (600 minutes).

5.3.1.2 HA-GccH pulldown results

An overnight culture of *Lb. plantarum* NC8 transformed with pNZ9530 and pNZ8148_HA-*gccH* was diluted 50-fold, and incubated at 30 ℃ until an OD_{600} of ~ 0.6 was reached. Following this, cells were split into three 10 mL volumes: nisin was added to two of these to a final concentration of 25 ng/mL, while the third fraction was left as an uninduced control. Cells were induced at 30 ℃ for a further 60 minutes, after which time GccF was added to one at a final concentration of 2 μM. All three cultures were left to incubate for a further 30 minutes, after which the cells were pelleted by centrifugation, lysed and immunoprecipitated using anti-HA antibodies. Streptavidin-conjugated magnetic beads were used to isolate the labelled proteins as described in Section 5.2.5. In order to increase the number of peptides detected, a tryptic digest was carried out directly on the streptavidin beads rather than on eluted proteins. A total of 2 μg of trypsin was used for each sample.

Figure 5.13 shows the numbers of proteins identified in each sample ('Control' [uninduced cells], 'HA-GccH' [cells expressing HA-GccH] and 'GccF' [cells expressing HA-GccH and treated with GccF]) following proteomic analysis, as well as the numbers of proteins found in common to each sample. A high degree of overlap between all three samples is immediately obvious. 502 proteins were common to the uninduced cell pulldown as well as the induced untreated and induced treated cells. This high level of non-specific binding could be indicative of an excess of streptavidin magnetic beads being used relative to the concentration of biotinylated anti-HA antibodies. Peptides belonging to streptavidin from *Streptomyces avidinii* were found in all three samples (accession number P22629, unique peptides = 7, coverage = 43 %).

The sample that was of particular interest initially was the 'GccF' sample which, if the GccF and GccH interaction behaved as that of lactococcin A and LciA[59], should reveal a specific interaction with a receptor. After filtering for protein IDs with only a single unique peptide, the 92 proteins exclusive to 'HA-GccH + GccF' were reduced to 15 (Table 5.7). Unfortunately, none of these protein matches belonged to any transporters, suggesting that strong complex between GccH, GccF and the receptor may not occur as expected. Moving on from this, the protein hits common to induced cells ('HA-GccH') and GccF-treated induced cells ('GccF') were evaluated. Again, after filtering for proteins IDs with only a single unique peptide, the number of overlapping IDs was reduced from 226 to 151 (Figure 5.13). The full list of these

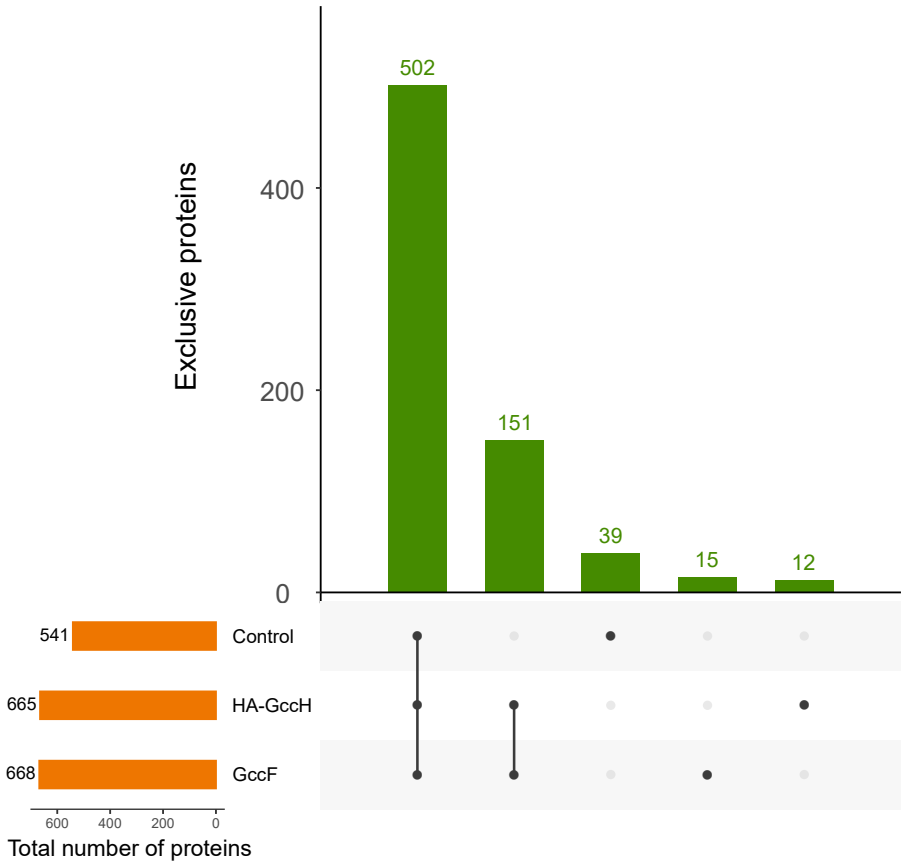


Figure 5.13: **Overlap of HA-GccH pulldown protein matches**
Bar graphs showing the overlap of proteins identified in HA-GccH pulldowns. The bars on the left (orange) represent the total number of proteins identified in each pulldown experiment, while the bars on the top (green) show the numbers of proteins that are identified in common to different experiments. 'Control' refers to pulldown carried out on uninduced cells, 'HA-GccH' refers to cells expressing HA-GccH, and 'GccF' refers to cells expressing HA-GccH that were also treated with 2 μM GccF. It can be seen that the greatest overlap occurs between all three samples, and the second greatest overlap occurs between HA-GccH and GccF samples, exclusively. For an alternative representation of this data, please see Figure B.10.

Table 5.7: Proteins unique to GccF-treated, HA-GccH produced *Lb. plantarum* pulldown

Accession	Unique Peptides	Coverage [%]	Description
HA-GccH†	**3**	**42**	**HA-tagged GccF immunity protein**
WP_054399571.1	3	9	DNA helicase RecQ
WP_033610374.1	2	48	ribitol-5-phosphate dehydrogenase
WP_003644250.1	2	14	response regulator transcription factor
WP_003642607.1	2	18	carbamoyl phosphate synthase small subunit
WP_003644144.1	2	18	glycosyltransferase
WP_003645420.1	2	14	peptide chain release factor N(5)-glutamine methyltransferase
WP_003640900.1	2	12	ammonia-dependent NAD(+) synthetase
WP_003644108.1	2	16	L-lactate dehydrogenase
WP_046947519.1	2	38	membrane protein
WP_015379544.1	2	7	LacI family DNA-binding transcriptional regulator
WP_033615425.1	2	12	hypothetical protein
WP_102124376.1	2	22	hypothetical protein
WP_003642898.1	2	10	choloylglycine hydrolase family protein
WP_003644445.1	2	13	tRNA (adenine-N(1))-methyltransferase
WP_003637970.1	2	7	type I pantothenate kinase

All proteins with only one unique peptide have been filtered out.
†HA-GccH appeared in both nisin-induced samples.

151 proteins are shown in Supplementary Table C.1. Peptide matches corresponding to HA-GccH were found in both samples where expression was induced (highlighted in Supplementary Table C.1), and not found in the uninduced sample, indicating that the expression system had worked as expected.

Five of the 151 proteins common to the GccF-treated and untreated induced samples belonged to transporters: two were ABC transporters (WP_003637950.1 and WP_016527128.1), one APC family transporter (WP_003637815.1), one uncharacterised transporter (WP_047672593.1), and a GlcNAc-specific PTS transporter (WP_003642736.1) (highlighted in Supplementary Table C.1). Searching this protein against the wider NCBI *Lb. plantarum* database revealed this to be PTS18CBA, the PTS transporter initially identified in GccF-resistant mutants of *Lb. plantarum*[62]. Additionally, another protein common to both HA-GccH induced samples was a GlcNAc-6-P deacetylase (WP_003643863.1; NagA), which is the next enzyme in the pathway directly after GlcNAc import by the PTS[126]. However, it remains unclear if this protein was precipitated because of a direct relationship with

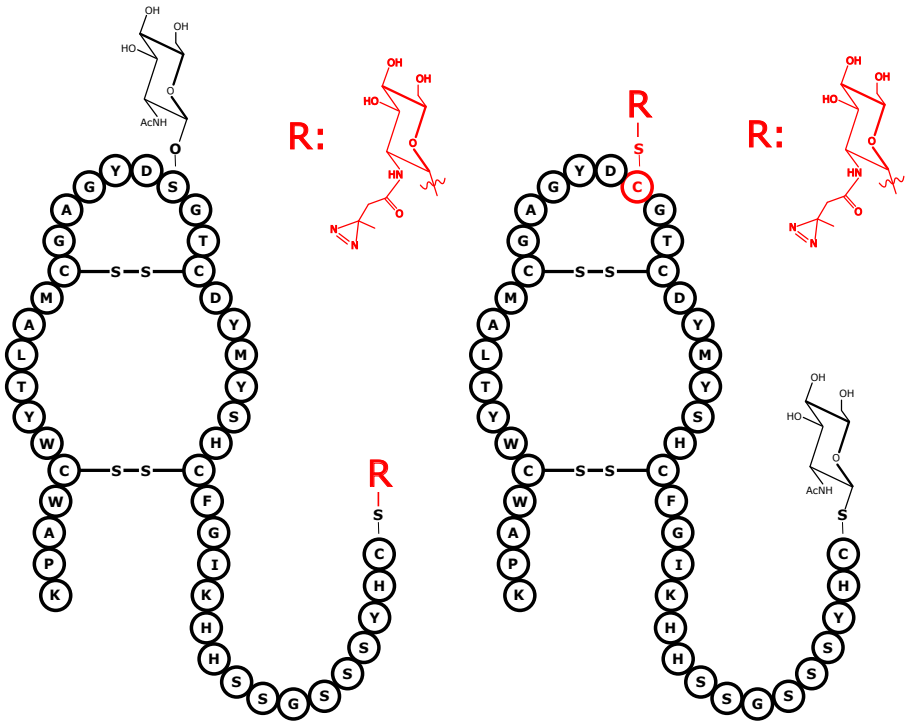
HA-GccH, or because it naturally forms a complex with PTS18CBA.

In addition to the identification of PTS18CBA, other proteins (7/151) involved in the cytoplasmic steps of peptidoglycan synthesis were identified, as well as multiple proteins involved in various other cellular process, including DNA replication and protein synthesis (Supplementary Table C.1). The sheer number of proteins identified by this method makes it difficult to draw any proper conclusions about the interacting partners of GccH, or about the mechanism used by this immunity protein. The identification of PTS18CBA in the samples where GccH is induced is a promising sign, but additional experiments to reduce the degree of non-specific binding are essential before any formal conclusions can be drawn. Such experiments would involve significantly increasing the volume/ density of cells lysed (to be in more of an accordance with Diep *et al.* (2007)[59]), and possibly treatment with low concentrations of some crosslinking agent, such as glutaraldehyde, prior to cell lysis[148]. An additional pulldown experiment using photo-active crosslinkable GccF was attempted in order to identify/ confirm the primary cellular target of GccF.

5.3.2 Photo-activated GccF analogue pulldowns

Biotin-tagged analogues of GccF containing diazirine-conjugated GlcNAc (**GlcNDAz**) were chemically synthesised, and photo-activated crosslinking was attempted. GlcNDAz-labelled proteins have been successfully used to identify enzymes and proteins that recognise GlcNAc modifications through photo-activated crosslinking[176]. Two analogues were prepared, to increase the probability of obtaining at least one cross-linked analogue-target molecule involving either the tail-GlcNAc or the loop-GlcNAc. These were $\text{GccF}_{Cys43GlcNDAz}$, which contained an unmodified GlcNAc at position 18 and Cys43-GlcNDAz, and $\text{GccF}_{Cys18GlcNDAz}$, which had an unmodified GlcNAc at position 43, and Cys18-GlcNDAz (Figure 5.14). These analogues were monobiotinylated after synthesis (presumably on the N-terminus or the sidechain of Lys1), to allow for precipitation of crosslinked products using streptavidin-conjugated magnetic beads. The activity of both compounds was tested using the indicator strain *Lb. plantarum* ATCC 8014, to see if the presence of diazirine on the GlcNAc affected activity. $\text{GccF}_{Cys18GlcNDAz}$ was inactive, and $\text{GccF}_{Cys43GlcNDAz}$ had an IC_{50} of 294 $\pm$ 59 nM, showing that the presence of the diazirine interferes with the GccF receptor interaction (Supplementary Figure B.11). Nevertheless, the modified peptides were still used for crosslinking. The rationale behind this was that with at least one

unmodified GlcNAc, the GccF should be located near one of the potential receptors and crosslinking may occur.



(a) GccF$_{Cys43GlcNDAz}$ (b) GccF$_{Cys18GlcNDAz}$

Figure 5.14: **Structures of GlcNDAz-conjugated GccF analogues**
(a) Structure of tail-conjugated photo-GccF through Cys43. (b) Structure of loop-conjugated photo-GccF through Cys18. GlcNDAz structure is shown in red. Note that both analogues were monobiotinylated following synthesis, but the specific location of the biotin has not been confirmed, although it is assumed to be attached to the N-terminus or sidechain of Lys1.

Diazirine can be activated with a long wavelength (330-370 nm) UV light source to form a reactive carbene, and thus interact with any organic compound containing a hydrogen bond. As such, this makes them preferable for crosslinking experiments compared to other chemical crosslinkers such as formaldehyde or glutaraldehyde, which are restricted to forming crosslinks with free amines of proteins. Diazirine, and specifically GlcNDAz, have been used successfully in the past to identify protein-protein interactions, with photocrosslinking being encouraged by irradiation with UV light over the range of 354 to 365 nm[176,160]. For this work, cells treated with 250 nM GccF$_{Cys18GlcNDAz}$, 250 nM GccF$_{Cys43GlcNDAz}$, or untreated were irradiated at 365 nM for 30 minutes in chemically-defined media[163], to minimise crosslinking

with non-cellular components. Cells were lysed, and proteins were extracted and analysed as described in Section 5.2.5.

Figure 5.15 shows the overlap in protein IDs from the *Lb. plantarum* NC8 database found in samples treated with tail- or loop-linked GlcNDAz, compared to the untreated control. As with the HA-GccH pulldown, streptavidin was found in all three samples (accession number = P22629, unique peptides = 8, coverage = 43 %). Lists of proteins were filtered to exclude proteins that only had one unique peptide hit. This reduced the number of proteins identified which were common to both GccF$_{Cys18GlcNDAz}$ and GccF$_{Cys43GlcNDAz}$ from 189 to 155, and the number of proteins exclusive to the GccF$_{Cys18GlcNDAz}$ or GccF$_{Cys43GlcNDAz}$ were reduced from 39 to 7 and 72 to 11, respectively. The protein matches common to both tail- and loop-crosslinked photo-GccF pulldowns are listed in Supplementary Table C.2, and protein hits exclusive to GccF$_{Cys18GlcNDAz}$ or GccF$_{Cys43GlcNDAz}$ are listed in Supplementary Tables C.4 and C.5, respectively. From the list of proteins exclusive to GccF$_{Cys43GlcNDAz}$ (tail-crosslink), there are no matches to any transporters. In fact, the proteins that are identified are all cytosolic proteins involved in various cytosolic processes, and thus appear to be the result of non-specific binding. The lack of membrane-bound proteins is most likely a result of the lack of detergent being added to the cells prior to lysis, and so is not a surprising result. Similarly, no proteins from the GccF$_{Cys18GlcNDAz}$ (loop-crosslink) exclusive list belong to any transporters, although one protein involved in peptidoglycan synthesis, MurE, is identified (Supplementary Table C.4). This enzyme is involved in one of the cytoplasmic steps of peptidoglycan synthesis, where it installs the addition of a third amino acid to the growing UDP-MurNAc-L-Ala-D-Glu dipeptide chain[19]. However, given that it does not recognise or use GlcNAc, it is unlikely that this protein is the target of GccF *in vivo*.

Despite the addition of diazirine to GlcNAc reducing the activity of GccF, it was hoped that having one unmodified GlcNAc on the peptide would allow at least partial localisation of the molecule to the native target, which would then allow the GlcNDAz to be close enough to crosslink with this protein. Because of this, it was possible that the primary receptor would be identified in both the loop-crosslinked and tail-crosslinked samples. Looking at the list of proteins common to both loop- and tail-crosslinked samples (Supplementary Table C.2), it can be seen that 7 of the 155 proteins belong to ABC transporter components. Meanwhile, 3 of the 155 proteins belong to cell division proteins (FtsA, MreB and GpsB), two

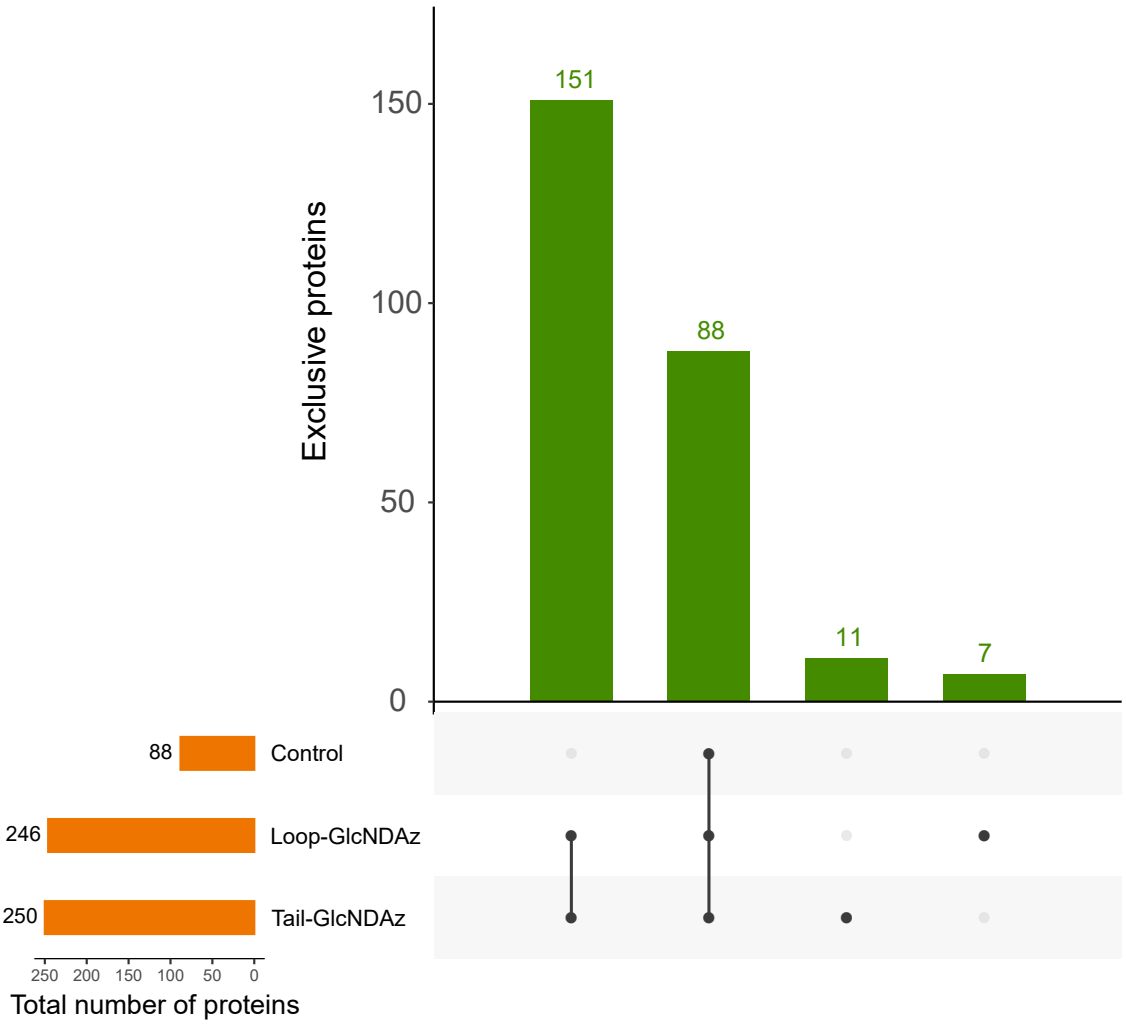


Figure 5.15: **Overlap of GccF-GlcNDAz pulldown proteins**
Bar graphs showing the overlap between proteins identified by pulldowns following crosslinking with GccF$_{Cys18GlcNDAz}$ ('Loop-GlcNDAz'), GccF$_{Cys43GlcNDAz}$ ('Tail-GlcNDAz'), or no crosslinking ('Control'). The bar graphs on the left (orange) indicate the total numbers of proteins identified in each pulldown, while the bar graphs on the top (green) show the numbers of proteins common or exclusive to each experiment. From these results, it appears that the largest degree of overlap occurs between GccF$_{Cys18GlcNDAz}$ and GccF$_{Cys43GlcNDAz}$ pulldowns, with the second largest overlap occurring between all three samples. For an alternative representation of this data, please see Figure B.12.

Table 5.8: Proteins common to HA-GccH and photo-activated GccF analogue pull-downs

Accession number	**Description**
WP_046039231.1	FAD-binding oxidoreductase
WP_003644631.1	glutamate racemase†
WP_033610823.1	D-alanyl-D-alanine carboxypeptidase†
WP_057717686.1	universal stress protein
WP_003642169.1	NADP oxidoreductase
WP_003638751.1	cell division regulator GpsB†
WP_003639329.1	aspartate-semialdehyde dehydrogenase

Accession numbers taken from *Lb. plantarum* database
†Proteins involved in cell division/ cell wall biosynthesis

belong to proteins involved in cell wall biosynthesis (LytF, and MurG), and one belonged to an autolysin (WP_003640832.1, unnamed protein). It is possible that the crosslinking with cell wall-active enzymes was the result of GccF interacting with proteins that recognise and bind to GlcNAc, such as MurG, which facilitates the addition of GlcNAc to the MurNAc pyrophosphate pentapeptide[27]. Notably, although no PTS transporter components were pulled down, it is interesting that the central regulatory component of the PTS system, HPr, was identified (EHS83685.1). However, the wealth of proteins pulled down is probably the result of non-specific interaction, considering the highly reactive nature of diazirine. It is interesting that many of these proteins are cytosolic, possibly due to the crosslinking reactions still occurring even once the cells have been lysed, or possibly due a number of glycocin F molecules entering the cells, which is something that had not previously been considered based on the fact that GccF is bacteriostatic. Additionally, it is possible that a lack of membrane solubilisation resulted in the majority of membrane-bound proteins being lost.

Lastly, proteins pulled down from the photoactive GccF crosslinking experiments were compared to those pulled down from the HA-GccH interaction (Section 5.3.1.2). There were no proteins in common between the induced HA-GccH pulldown samples ('HA-GccH' and 'GccF') and either the GccF$_{Cys18GlcNDAz}$ or GccF$_{Cys43GlcNDAz}$-exclusive lists, but there were 7 proteins in common between the HA-GccH samples and the 'loop and tail' common pulldown samples (Table 5.8). Of these seven pro-

teins, one was identified as the cell division regulator GpsB, and another was an enzyme involved in peptidoglycan biosynthesis (D-alanyl-D-alanine carboxypeptidase). GpsB is a cell division protein involved in the recruitment of peptidoglycan synthases to the divisome of actively dividing Gram-positive cells, and is though to be situated on the cytoplasmic side of the cell membrane, where it interacts with transmembrane domains of its recruited synthases[43]. As such, it is possible that it is capable of interacting with both GccH, which presumably does not exit the cell nor embed in the membrane, and photo-active GccF, if one of the associate proteins is crosslinked to the GlcNDAz. The D-alanyl-D-alanine carboxypeptidase (WP_033610823.1), on the other hand, contains a signal sequence and cleavage site (based on prediction by SignalP 5.0), suggesting that it is exported from the cell, and thus may be the situated in the cell membrane, and thus be accessible to GccF if it does not enter the cell.

5.4 Conclusion

The work carried out on the GccF immunity protein, GccH, gave some insights into how this protein may function. To begin with, it appears that the addition of tags to the N-terminus does not interfere with the function of GccH, although the inclusion of an untagged control within these experiments would allow this comparison to be confirmed. Initial attempts were made to modify the original pRV613_*gccH* plasmid to create a version with a C-terminal tag in the event that the N-terminal modification negated activity, but that region of the plasmid proved difficult to mutate using SLIM methodology[40], due to an AT-rich region downstream of the gene. Further attempts to modify the C-terminus were abandoned once the activity of pRV613_flag-*gccH* was confirmed.

The plasmid pRV613 is a low copy number plasmid created for the modification in *E. coli* and expression of proteins in *Lb. plantarum*, *Enterococcus faecalis* and *Bacillus subtilis*[51]. It contains a copper-inducible promoter which is reported to be inducible under copper concentrations as low as 30 μM, although in our hands it also displayed some leaky expression (data not shown), probably owing to residual copper in the cell culture media. The amount of protein expression observed using this plasmid system was too low to detect on a colloidal coomassie-stained SDS-PAGE gel, and expression had to be confirmed using western blotting. It is interesting

that expression of tagged GccH using this expression system still proved enough to protect *Lb. plantarum* NC8 cells from GccF over a 15 hour growth curve (Figure 5.2). This suggests that the number of copies of GccH required to protect the cells is not very high, and could suggest that the interacting partner of GccH is also not present in very high abundance, or that it functions by interacting with a central regulation system.

The idea of tagging GccH and using it as a bait for pulling down proteins that it interacted with in cells that had been exposed to GccF was based on the work carried out on the bacteriocin lactococcin A, which reported the formation of a complex between lactococcin A, its cognate immunity protein, and the bacteriocin target, the Man-PTS[59]. The mass spectrometry results of the pulldowns from this work indicate that a similar mechanism is unlikely, as there were no membrane-spanning proteins or transporters co-precipitated in this experiment when GccF was included (Table 5.7). However, there are clear differences between the methodology of this work and that of Diep *et al.*[59] For instance, the volume and concentration of cells used in the immunoprecipitation of lactococcin A (between 50 - 250 mL of overnight culture) was much higher than was used in these experiments (2 - 4 mL), the flag-tagged protein was purified by passing the cell lysate through anti-flag M2 affinity gel, as opposed to incubating with anti-HA antibodies and streptavidin-conjugated magnetic beads. Finally, the purified protein was run on an SDS-PAGE gel, and bands were excised and identified, rather than analysing all of the eluted proteins directly by mass spectrometry. This work would benefit from being repeated to more clearly identify the main protein(s) interacting with GccH and GccF, and to decrease the amount of non-specific interactions being observed.

The identification of PTS18CBA and GlcNAc-6-P deacetylase (NagA) in both the induced GccH samples is consistent with the previous evidence suggesting a role of the GlcNAc-PTS in GccF activity[62,17]. In particular, it is promising that this is the only PTS component seen in these pulldown results. The co-precipitation of both the PTS18CBA and the enzyme that acts on sugars directly after import by the GlcNAc-PTS suggests a possible co-localisation or complex formed by these and possibly other proteins. As such, it is possible that GccF functions by interacting with the GlcNAc-PTS, and causing an effect through an as-yet unidentified complex that forms around the cytoplasmic domains of the transporters.

The work carried out on the photoactive GccF analogues $\text{GccF}_{Cys18GlcNDAz}$ and

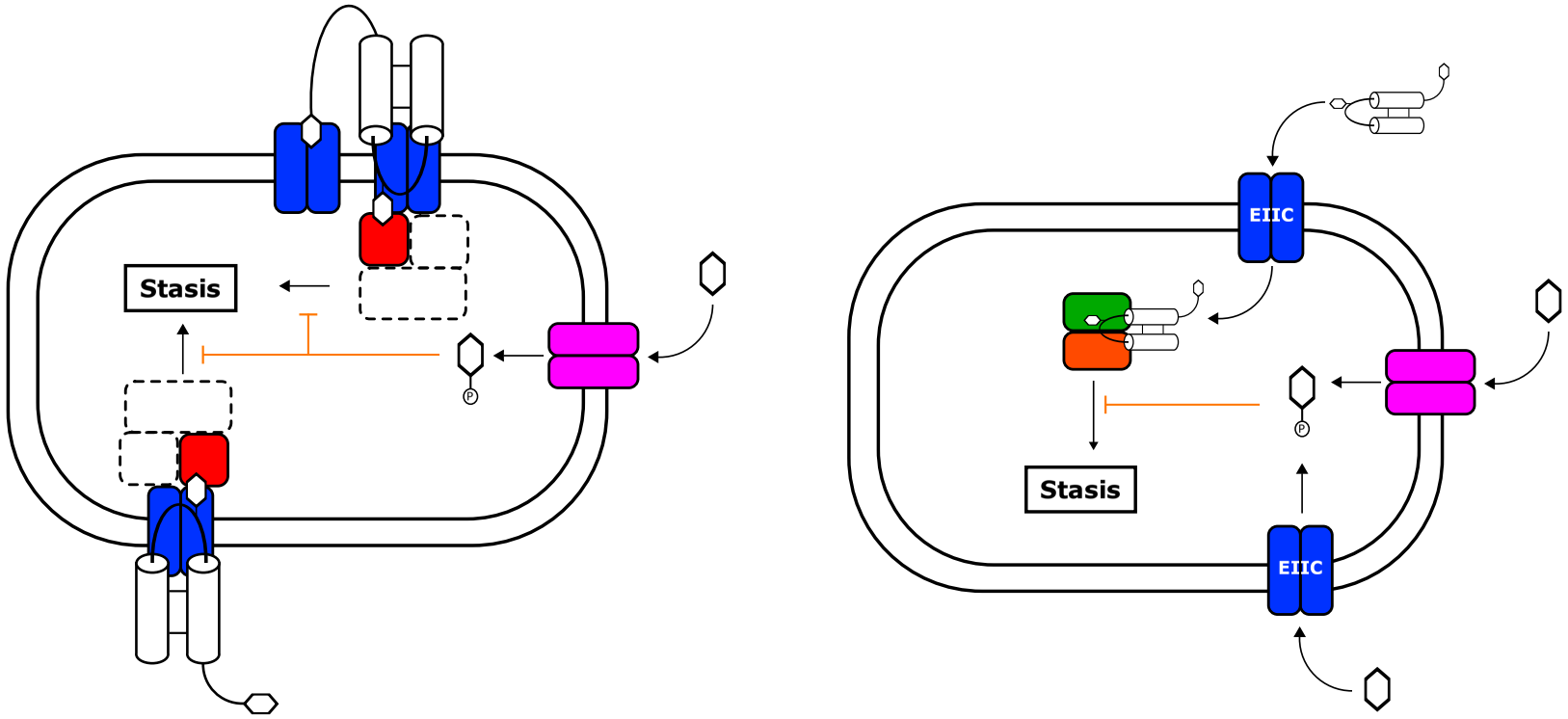
GccF$_{Cys43GlcNDAz}$ would benefit from being repeated, with additional optimisation steps carried out prior to data collection. These could include altering the amount of time for which treated cells are irradiated under UV light, altering the concentrations of GlcNDAz-containing GccF used, and properly solubilising or even isolating the membrane fractions of lysed cells. This last point is of particular importance, as more care could have been taken to solubilise the membrane fraction of the cells following lysis. This may explain why such a high number of cytosolic proteins were identified. However, it cannot be ruled out that the use of GlcNDAz-containing variants may never reveal the primary binding target of native GccF, due to the loss (or reduction) of activity seen when the diazirine group is placed on the GlcNAcs.

6 | Final conclusions and future directions

6.1 Summary

What was learnt from the inhibition studies carried out on the chemically synthesised chemical analogues? First and foremost, all the post-translational modifications (glycosylation and disulfide bonds) are involved in the activity of glycocin F. In other words the structure of the helix-loop-helix motif is specifically arranged to present the GlcNAc bound to Ser18 in a specific orientation, as reducing the size of the loop (deleting one, two or three Glys from the sequence), changing its structure (disruption of disulfide bond between Cys5 and Cys28), and even making small modifications to Ser18 itself (such as adding a methyl group at the alpha-position) all abolish activity (Table 3.2). The flexible tail itself and the GlcNAc bound to Cys43 are also important, although its sequence and the length are less involved. The presence and identity of the sugar bound to Cys43, however, does have a significant effect as seen by greater loss of activity when the GlcNAc is substituted with mannore or glucose compared to when it is removed altogether. This indicates that both sugars are involved in the mechanism.

The activities of these chemically synthesised variants of GccF led to two proposed mechanisms of action (Figure 3.13)[24]. Both mechanisms are based on the assumption that the C-terminal GlcNAc plays a role in localising/ concentrating the bacteriocin to a specific region of the cell surface through interaction with the GlcNAc-specific PTS18CBA. Because the truncated molecule $GccF_{1-32\Delta GlcNAc}$ remained active, it was assumed that bacteriostasis was due to the Ser18-loop GlcNAc interacting with a different site of the transporter (Figure 3.13 (a)), or with another membrane-bound protein localised near the GlcNAc-PTS (Figure 3.13 (b)). However, given the results following the transcriptomic and immunoprecipitation studies, the latter of these proposed mechanisms has become less likely than the former. The results of the work carried out for this thesis suggest the possibility of two additional mechanisms of action (Figure 6.1). In mechanism A, the the loop itself and Ser18-GlcNAc both interact with the transmembrane EIIC domain of the GlcNAc-specific PTS, effectively jamming the transporter (Figure 6.1 (a)). This jamming triggers an internal response, which causes the cells to enter bacteriostasis. In mechanism B, the GlcNAc-PTS acts as an entry point for GccF, and the protein is completely taken into the cells (Figure 6.1 (b)). From here it interacts with some as-yet undetermined protein(s) to trigger the response that ultimately results in bacteriostasis.



(a) Mechanism A

(b) Mechanism B

Figure 6.1: **Proposed mechanisms of GccF activity**
Two possible mechanisms of action of GccF-induced stasis are presented. (a) GccF binds to the GlcNAc-PTS transporter (blue) irreversibly, where a complex that forms on the intracellular side of the PTS comprised of at least one other 'key' protein (red) is stabilised, resulting in stasis (see also Figure 4.10) The C-terminal GlcNAc is also shown binding to another copy of the GlcNAc-PTS, highlighting its role in localising the bacteriocin. It should be noted the diagram is a model only, and does not indicate a specific mode of binding. (b) GccF enters the cell through the GlcNAc-PTS (blue), and initiates stasis by binding a protein or complex or protein inside the bacterium (denoted by green and orange boxes). In either mechanism, stasis is relieved by the import of GlcNAc as GlcNAc-6-P by another PTS transporter (magenta).

Evidence from other work supporting the GlcNAc PTS as being a molecular target of GccF began with the isolation of mutations in this transporter in GccF-resistant strains of *Lactobacillus plantarum*, supported by the GccF response of *Lb. plantarum* NC8 cells containing PTS18CBA deletions[62,17]. The observations from this work that for *Enterococcus faecalis* JH2-2 up-regulation of the GlcNAc-PTS correlates to increased susceptibility to GccF, and that PTS18CBA (along with GlcNAc-6-P deacetylase, NagA) was identified in pulldowns using the immunity protein GccH in *Lb. plantarum* NC8 (Supplementary Table C.1) also supports this. In addition to PTS18CBA and NagA, 148 other proteins were identified as interacting with GccH (Figure 5.13), which strongly suggests the need to revise the procedure to reduce non-specific binding. Nevertheless, the fact that PTS18CBA was identified in the HA-GccH pulldowns, despite the fact that the *Lb. plantarum* genome can encode as many as 40 different PTS transporters, suggests the pulldown has the potential to identify some proteins that may be involved in the formation of a complex between GccH, PTS18CBA, GccF and possibly other proteins. Additionally, reports of a glucose-specific PTS transporter conferring susceptibility to the glycocin sublancin 168[72] is in agreement with GccF targeting a PTS, and makes it likely that other glycocins also target this family of sugar transporters.

The onset of bacteriostasis by GccF is rapid, occurring in susceptible cells within minutes (the order of 1 - 15 minutes, depending on the bacterial species) of cells being exposed. Likewise, the recovery from stasis by free GlcNAc is also rapid (Figure 4.4), suggesting that the onset of and recovery from stasis is not initially dictated by a change in transcription of a set of genes, but rather by protein-protein interactions. The questions remains as to what protein GccF is binding to initiate this stasis, and how GlcNAc can act to reverse this stasis. In mechanism A, the binding of GccF to the GlcNAc-PTS stabilises an interaction between the PTS and another intracellular protein that is involved in GlcNAc import. The GlcNAc-PTS has been reported to be involved in cell wall/ peptidoglycan recycling[126]. It is possible that the GlcNAc-PTS acts as a kind of 'sensor', controlling the production of precursor molecules (e.g. GlcNAc and MurNAc) for peptidoglycan synthesis by detecting the presence or absence of GlcNAc. In the absence of external GlcNAc, both sugar precursors of peptidoglycan, UDP-GlcNAc and UDP-MurNAc, must be synthesised from glucosamine-6-phosphate (GlcN-6-P), which is produced from fructose-6-phosphate (Fru-6-P) by glucosamine-fructose-6-phosphate aminotransferase (GlmS) . The import of GlcNAc *via* the GlcNAc-PTS results in intracellular accumulation of GlcNAc-6-P, which is

immediately deactetylated by GlcNAc-6-P deacetylase (NagA) to form GlcN-6-P[98]. This intracellular accumulation of GlcN-6-P then acts to repress the expression of *glmS* through activation of this mRNA's native ribozyme activity[19], and prevent Fru-6-P being converted to GlcN-6-P. Transcriptionally, the repression of *glmS* was observed in this work when *E. faecalis* was grown in the presence of extra GlcNAc (Table 4.3).

One possibility of GccF's mechanism of action is that through binding the GlcNAc PTS, a protein complex formed on the cytosolic side of the PTS is held in a state mimicking the import of GlcNAc, and signalling the cell to not produce further GlcN-6-P from Fru-6-P. In the absence of extracellular GlcNAc, bacterial cells with GccF embedded in the PTS enter a static phase, where they lose the ability to synthesise peptidoglycan and are prevented from dividing, and thus growing. In fact, the sudden stasis that cells undergo following GccF exposure is very similar to that seen when cells are treated with a non-metabolisable variant of GlcN-6-P, which was shown to induce bacteriostasis in *B. subtilis* cells by causing the *glmS* ribozyme to continually cleave itself[142]. It has also been recently reported that the in the bacterium *Corynebacterium glutamicum*, the fructose and glucose-specific PTS transporters are localised to the cell division point when the cells are cultured in low concentrations of the respective sugars. However, once the cells are exposed to higher concentrations of the specific PTS sugars, the localisation of these PTS subunits changes, and the transporters become more spread out[111]. This raises interesting questions regarding possible localisation of the GlcNAc-PTS, and earlier work on *Lb. plantarum* treated with fluorescently-labelled GccF strongly indicated that the bacteriocin localised to the cell division point of these cells also (Unpublished results).

The model of GccF interacting with the GlcNAc-PTS and 'jamming' could also agree somewhat with one of the strongest, most consistently observed transcriptional responses of *E. faecalis* to GccF. At all time points tested, and at all concentrations tested, and even in the σ^{54} knockout strain, treated *E. faecalis* cells showed a strong up-regulation of the V-type ATPases, which is presumably occurring in response to minor leakage of protons and/or salt ions across the membrane. A situation where GccF is jamming the PTS transporters could result in minor, non-lethal permeations in the cell, resulting in loss of the proton motive force and thus the subsequent up-regulation of the V-type ATPases.

In mechanism B, GccF enters the bacterium and interacts directly with its targeted protein. Evidence supporting GccF entering the cell and interacting directly with internal molecules is based on the number of cytosolic proteins identified in the GlcNDAz-GccF pulldowns (Chapter 5) is weak wecause of the lack of membrane-bound proteins identified in the pulldowns. This is possibly due to the lack of solubilisation of the cell membrane fraction resulting in its loss during the washing steps prior to analysis by mass spectrometry, and so does not necessarily mean that the photo-active analogues did not crosslink to any membrane proteins. Other possibilities include the highly reactive diazirine crosslinking to other molecules on the cell surface, such as lipids and peptidoglycan, or that the extended UV irradiation led to some cells being lysed and releasing cytosolic proteins into the environment. Also, the loss of biological activity due to the GlcNDAz moieties cannot be ignored, and so it cannot be expected that the analogues would necessarily bind or be recruited to the correct membrane-bound molecule. Finally, if GccF is being completely internalised *via* the GlcNAc PTS, then photo-crosslinking should have allowed some peptides belonging to this transporter to be identified.

One reason for discarding the idea of GccF binding to another, external cell receptor other than the GlcNAc-PTS is the observed interplay between GccF and GlcNAc in *Enterococcus faecalis* JH2-2 (see Chapter 4). For this bacteria, complete bacteriostasis was never observed regardless of the amount of GccF provided unless cells were pre-sensitised by additional GlcNAc in the media (Figure 4.3). RNA-seq confirmed that the GlcNAc-specific PTS was up-regulated following growth with additional GlcNAc (Table 4.3). Initially, the idea of GccF targeting an extracellular protein other than the PTS suggested that the GlcNAc-PTS was localised near this unknown protein on the cell surface, and that interaction of the C-terminal GlcNAc of GccF with the GlcNAc-PTS acted to recruit GccF to this area, and increase the chance of it associating with the other protein. However, this is at odds with what is observed in *E. faecalis*, as even if there were lower numbers of the GlcNAc-PTS compared to this mystery protein under basal conditions, an excess of GccF should allow all of these protein binding sites to be occupied. Therefore, while it is still likely that the C-terminal GlcNAc of GccF is acting to localise GccF to its area of effect on the bacterial cell surface, it is unlikely that the Cys43-linked and Ser18-linked GlcNAcs are targeting different proteins.

In mechanism A, the question of how GlcNAc acts to reverse stasis remains. One possibility is that free GlcNAc competes with and excludes GccF from the PTS

transporter, and dislodges it for as long as there is enough GlcNAc entering the cells[119]. However, this mechanism would not necessarily explain the rapid onset/ resumption of stasis that is seen to occur when the additional GlcNAc in the environment is depleted (Figure 4.4). Another possibility is that following GccF binding and the onset of stasis, excess GlcNAc is imported into the cells through other PTS transporters. This could be through a β-glucoside PTS, such as the second PTS transporter seen up-regulated in the GlcNAc-only treatment (Table 4.3), or through the Man-PTS, which has been reported to be able to import GlcNAc[126]. If this is the case, it would be the presence of GlcNAc-6-P that relieves bacteriostasis, rather than competitive exclusion of GccF, through an as-yet unknown mechanism. It is unlikely to be the deacetylated product of GlcNAc-6-P deacetylase, GlcN-6-P, which relieves stasis as glucosamine (which is imported by the PTS as GlcN-6-P) does not relieve the cells from GccF-induced stasis (data not shown). This model would also explain the recovery by GlcNAc in the mechanism where GccF enters the cells, although in this situation GlcNAc would be primarily entering through the GlcNAc-PTS.

If GccF is targeting the GlcNAc-PTS, how does pre-sensitisation by GlcNAc treatment in *E. faecalis* JH2-2 work? Transcriptomic analysis clearly shows that in the presence of additional GlcNAc, the number of PTS-GlcNAc transcripts was up-regulated, leading to an increased response by the cell to GccF. Presumably, for other bacteria such as *Lb. plantarum*, there are enough copies of the GlcNAc-PTS to initiate full stasis. But how does this fit with either of the two mechanisms proposed here? To start with, it is difficult to rationalise this with the model of intact, folded GccF entering the cell *via* the GlcNAc-PTS (Figure 6.1 (b)), as presumably only a small number of GccF molecules entering the cell would be needed to induce stasis. Logically, however, a small number of transporters would result in a longer delay to stasis, but would not be a limiting factor. The only other explanation for GlcNAc sensitisation in an intracellular GccF mechanism is that GlcNAc pre-treatment leads to either an up- or down-regulation of some other critical component that GccF interacts with. However, looking at the genes differentially expressed upon GlcNAc-only exposure, there are none that might play this role (Table 4.3).

In contrast, up-regulation of GlcNAc-PTS components could explain the increase in susceptibility in mechanism A. This would occur through an increase in the number of membrane-bound EIIC components compared to some as-yet unknown intracellular protein, which is recruited to the GlcNAc-PTS during GlcNAc import, and

is responsible for the onset of stasis (Figure 4.10). It would then be this ratio of membrane-bound EIIC subunits compared to the internal unknown protein that would dictate the effectiveness of GccF, with strains of bacteria such as *Lb. plantarum* and *E. faecium* (which have been shown to be sensitive to GccF without need for presensitisation[92]) natively expressing as many as or more copies of the GlcNAc-PTS compared to this mystery protein. The identity of this mystery protein, and how a GccF:PTS complex would function to initiate a bacteriostatic response, remains to be determined.

GccH's mechanism of protecting cells could be explained for either proposed mechanisms of action of GccF, on the assumption that the protective effect is in fact due to GccH binding to small cytosolic domains of the PTS18CBA. In the event that GccF binds to the PTS irreversibly, GccH could prevent a signal for stasis from being propagated either by replacing or excluding a crucial protein or proteins, or by disrupting the complex completely. In the event that GccF enters the cell *via* the GlcNAc-PTS, then GccH would function by simply blocking import of GccF. If the latter situation were the case, then one would expect that the combined presence of GccF, GccH and the PTS18CBA would be identified as a strong complex, as in Diep *et al.* (2007)[59]. However, the finding from the pulldown table suggests that GccH may be recruiting to the PTS regardless of GccF presence. One way to test the mechanism by which GccH protects cells would be to see if GccH expression effects the growth rate of cells grown in minimal media supplemented with GlcNAc, with and without GccF added.

Finally, it is interesting to note the structural/ domain similarity exhibited by all glycocins characterised to date, specifically the helix-loop-helix structure with a glycosylated interhelical loop[119]. With such a unique shared structure, it seems unlikely that different glycocins would target different classes of bacterial transporters or receptors. Indeed, the mounting evidence of two glycocins, glycocin F and sublancin 168, targeting the sugar-specific PTS transporter system[72] seems too coincidental to not be indicative of a broadly shared mechanism. It seems more probable that the glycocins as a whole exploit the PTS system, likely with each glycocin targeting sugar-specific transporters denoted by their respective glycan. The recent suggestion of a grouping system within the glycocins is also of interest[132], and may indicate conserved targets or mechanisms. As research continues into the mechanism of action of this class of bacteriocins, it may pave the way for a new class of antimicrobials which make use of a relatively under-utilised group of transporters medically.

6.2 Future directions

Future work based on the evidence presented here can be divided into two groups: work that needs to be repeated/ optimised, and future experiments that can be designed and carried out. Some cases of work that needs to be better optimised (such as additional controls within the RNA-seq experiments) have already been discussed in respective conclusions, and so will not be focussed on here. However, it must be noted that the pulldown work presented in Chapter 5 was not properly optimised due to time constraints, and would benefit from being repeated with additional controls and optimisations. Specifically, the HA-tagged GccH immunity protein pulldown was carried out on a small scale, and as a result numerous co-precipitated proteins were identified. This would benefit from being repeated on a larger scale, with more care taken to solubilise the membrane fraction. Also, provision of larger quantities of pure GccH could be used to increase the number of stable protein complexes compared to weaker transient complexes between proteins. This may involve the addition of low concentrations of chemical crosslinkers (such as glutaraldehyde) to stabilise transient complexes formed. While these can be 'red herrings' in a search for binding proteins, there is always the possibility that the mechanism does involve the formation of such complexes. The diazirine-conjugated GccF photo-activated pulldowns should be repeated, as only initial trials were carried out and reaction conditions were not optimised. Due to the compromised activity of the modified GlcNAc moieties, caution is however recommended. As such, as far as using pulldowns to identify the binding partner(s) of GccF goes, it may be more beneficial to focus future efforts on GccH and chemically-crosslinked GccF pulldowns, rather than the diazirine-conjugated pulldowns.

An example of new experiments that should be carried out involved further characterising the relationship between number of the GlcNAc-PTS molecules and the strength of bacteriostasis, particularly as far as bacteria like *E. faecalis* are concerned. One such experiment would be to express GccH in *E. faecalis*. If it acted to protect these cells against the effect of GccF, it would provide a good opportunity to study the relationship between the relative number of GccH molecules produced compared to the numbers of GlcNAc-PTS transporters. One could either find a tunable expression vector for the production of GccH, or the pRV613 expression vector could be trialled, as this system has been reported to work in a wide range of cells including *E. faecalis*[51]. Assuming GccH expression could be tuned to allow

approximately equal copy numbers of the immunity protein as GlcNAc-PTS subunits in non-sensitised cells, the ability of GccH to protect the cells from the effects of GccF at equal ratios of GccH:GlcNAc-PTS could be compared to low GccH:PTS ratios. This would provide further evidence to indicate whether protective effect of GccH is dependent directly on its binding to the PTS, or whether it acts directly on the hypothesised 'key intracellular protein'.

6.3 Conclusion

The work presented in this thesis represents efforts undertaken to better understand the mechanism of action of the antimicrobial peptide glycocin F to impart bacteriostasis on susceptible strains and species of bacteria. This work has identified structural elements of the peptide that are import for activity, and has begun to identify transcriptional responses by a pathogenic bacteria to this molecule which may one day shed more light on the processes being targeted. Additionally, further evidence supporting a possible target of this peptide has been described. It is hoped that as work on this unique peptide continues, glycocin F may one day provide the blueprint for a new class of antimicrobial agents which will play a role in the ever-evolving battle against pathogenic bacteria.

7 | References

[1] A. Aakra, H. Vebo, U. Indahl, L. Snipen, O. Gjerstad, M. Lunde, and I. F. Nes. The response of *Enterococcus faecalis* V583 to chloramphenicol treatment. *International Journal of Microbiology*, 2010(483048), 2010.

[2] A. Aakra, H. Vebo, L. Snipen, H. Hirt, A. Aastveit, V. Kapur, G. Dunny, B. Murray, and I. F. Nes. Transcriptional response of *Enterococcus faecalis* V583 to erythromycin. *Antimicrobial Agents and Chemotherapy*, 49(6):2246–2259, 2005.

[3] J. Abranches, P. Tijerina, A. Aviles-Reyes, A. O. Gaca, J. K. Kajfasz, and J. A. Lemos. The cell wall-targeting antibiotic stimulon of *Enterococcus faecalis*. *PLoS One*, 8(6), 2013.

[4] J. Z. Acedo, S. Chiorean, J. C. Vederas, and M. J. van Belkum. The expanding structural variety among bacteriocins from Gram-positive bacteria. *FEMS Microbiology Reviews*, 42(6):805–828, 2018.

[5] S. Ahn, J. Stepper, T. S. Loo, S. W. Bisset, M. L. Patchett, and G. E. Norris. Expression of *Lactobacillus plantarum* KW30 *gcc* genes correlates with the production of glycocin F in late log phase. *FEMS Microbiology Letters*, 365(23), 2018.

[6] Z. Al Khatib, M. Lagedroste, I. Fey, D. Kleinschrodt, A. Abts, and S. H. J. Smits. Lantibiotic immunity: Inhibition of nisin mediated pore formation by NisI. *PLoS One*, 9(7), 2014.

[7] L. I. Alvarez-Anorve, M. L. Calcagno, and J. Plumbridge. Why does *Escherichia coli* grow more slowly on glucosamine than on N-acetylglucosamine? effects of enzyme levels and allosteric activation of GlcN6P deaminase (NagB) on growth rates. *Journal of Bacteriology*, 187(9):2974–2982, 2005.

[8] P. Alvarez-Sieiro, M. Montalban-Lopez, D. D. Mu, and O. P. Kuipers. Bacteriocins of lactic acid bacteria: Extending the family. *Applied Microbiology and Biotechnology*, 100(7):2939–2951, 2016.

[9] R. Aminov. History of antimicrobial drug discovery: Major classes and health impact. *Biochemical Pharmacology*, 133:4–19, 2017.

[10] R. I. Aminov. A brief history of the antibiotic era: Lessons learned and challenges for the future. *Frontiers in Microbiology*, 1(134), 2010.

[11] Z. Amso, S. W. Bisset, S. H. Yang, P. W. R. Harris, T. H. Wright, C. D. Navo, M. L. Patchett, G. E. Norris, and M. A. Brimble. Total chemical synthesis of glycocin F and analogues: S-glycosylation confers improved antimicrobial activity. *Chemical Science*, 9(6):1686–1691, 2018.

[12] S. Anders, P. T. Pyl, and W. Huber. HTSeq—a Python framework to work with high-throughput sequencing data. *Bioinformatics*, 31(2):166–169, 2015.

[13] P. G. Arnison, M. J. Bibb, G. Bierbaum, A. A. Bowers, T. S. Bugni, G. Bulaj, J. A. Camarero, D. J. Campopiano, G. L. Challis, J. Clardy, P. D. Cotter, D. J. Craik, M. Dawson, E. Dittmann, S. Donadio, P. C. Dorrestein, K. D. Entian, M. A. Fischbach, J. S. Garavelli, U. Goransson, C. W. Gruber, D. H. Haft, T. K. Hemscheidt, C. Hertweck, C. Hill, A. R. Horswill, M. Jaspars, W. L. Kelly, J. P. Klinman, O. P. Kuipers, A. J. Link, W. Liu, M. A. Marahiel, D. A. Mitchell, G. N. Moll, B. S. Moore, R. Muller, S. K. Nair, I. F. Nes, G. E. Norris, B. M. Olivera, H. Onaka, M. L. Patchett, J. Piel, M. J. T. Reaney, S. Rebuffat, R. P. Ross, H. G. Sahl, E. W. Schmidt, M. E. Selsted, K. Severinov, B. Shen, K. Sivonen, L. Smith, T. Stein, R. D. Sussmuth, J. R. Tagg, G. L. Tang, A. W. Truman, J. C. Vederas, C. T. Walsh, J. D. Walton, S. C. Wenzel, J. M. Willey, and W. A. van der Donk. Ribosomally synthesized and post-translationally modified peptide natural products: overview and recommendations for a universal nomenclature. *Natural Product Reports*, 30(1):108–160, 2013.

[14] T. Aukrust and H. Blom. Transformation of Lactobacillus strains used in meat and vegetable fermentations. *Food Research International*, 25(4):253–261, 1992.

[15] P. H. Axelsen. A chaotic pore model of polypeptide antibiotic action. *Biophysical Journal*, 94(5):1549–1550, 2008.

[16] K. Babasaki, T. Takao, Y. Shimonishi, and K. Kurahashi. Subtilosin A, a new antibiotic peptide produced by *Bacillus subtilis* 168 - isolation, structural analysis, and biogenesis. *Journal of Biochemistry*, 98(3):585–603, 1985.

[17] M. A. Bailie. *The role of the N-acetylglucosamine phosphoenolpyruvate phosphotransferase system from Lactobacillus plantarum 8014 in the mechanism of action of glycocin F : a thesis presented in partial fulfilment of the requirements*

for the degree of Master of Science in Biochemistry at Massey University, Manawatu, New Zealand. 2017.

[18] J. C. J. Barna and D. H. Williams. The structure and mode of action of glycopeptide antibiotics of the vancomycin group. *Annual Review of Microbiology*, 38:339–357, 1984.

[19] H. Barreteau, A. Kovac, A. Boniface, M. Sova, S. Gobec, and D. Blanot. Cytoplasmic steps of peptidoglycan biosynthesis. *FEMS Microbiology Reviews*, 32(2):168–207, 2008.

[20] S. L. Bausch, E. Poliakova, and D. E. Draper. Interactions of the N-terminal domain of ribosomal protein L11 with thiostrepton and rRNA. *Journal of Biological Chemistry*, 280(33):29956–29963, 2005.

[21] F. Bédard, R. Hammami, S. Zirah, S. Rebuffat, I. Fliss, and E. Biron. Synthesis, antimicrobial activity and conformational analysis of the class IIa bacteriocin pediocin PA-1 and analogs thereof. *Scientific Reports*, 8, 2018.

[22] Y. Benjamini and Y. Hochberg. Controlling the false discovery rate - a practical and powerful approach to multiple testing. *Journal of the Royal Statistical Society Series B-Methodological*, 57(1):289–300, 1995.

[23] A. A. Bhalodi, T. S. R. van Engelen, H. S. Virk, and W. J. Wiersinga. Impact of antimicrobial therapy on the gut microbiome. *Journal of Antimicrobial Chemotherapy*, 74:6–15, 2019.

[24] S. W. Bisset, S.-H. Yang, Z. Amso, P. W. R. Harris, M. L. Patchett, M. A. Brimble, and G. E. Norris. Using chemical synthesis to probe structure-activity relationships of the glycoactive bacteriocin glycocin F. *ACS Chemical Biology*, 13(5):1270–1278, 2018.

[25] S. Biswas, C. V. G. De Gonzalo, L. M. Repka, and W. A. van der Donk. Structure-activity relationships of the S-linked glycocin sublancin. *ACS Chemical Biology*, 12(12):2965–2969, 2017.

[26] L. A. Bohle, G. Mathiesen, G. Vaaje-Kolstad, and V. G. H. Eijsink. An endo-β-N-acetylglucosaminidase from *Enterococcus faecalis* V583 responsible for the hydrolysis of high-mannose and hybrid-type N-linked glycans. *FEMS Microbiology Letters*, 325(2):123–129, 2011.

[27] A. Bouhss, A. E. Trunkfield, T. D. H. Bugg, and D. Mengin-Lecreulx. The biosynthesis of peptidoglycan lipid-linked intermediates. *FEMS Microbiology Reviews*, 32(2):208–233, 2008.

[28] B. J. Bradbury and M. J. Pucci. Recent advances in bacterial topoisomerase inhibitors. *Current Opinion in Pharmacology*, 8(5):574–581, 2008.

[29] E. Breukink, I. Wiedemann, C. van Kraaij, O. P. Kuipers, H. G. Sahl, and B. de Kruijff. Use of the cell wall precursor lipid II by a pore-forming peptide antibiotic. *Science*, 286(5448):2361–2364, 1999.

[30] M. A. Brimble, P. J. Edwards, P. W. R. Harris, G. E. Norris, M. L. Patchett, T. H. Wright, S.-H. Yang, and S. E. Carley. Synthesis of the antimicrobial S-linked glycopeptide, glycocin F. *Chemistry – A European Journal*, 21(9):3556–3561, 2015.

[31] D. E. Brodersen, W. M. Clemons, A. P. Carter, R. J. Morgan-Warren, B. T. Wimberly, and V. Ramakrishnan. The structural basis for the action of the antibiotics tetracycline, pactamycin, and hygromycin B on the 30S ribosomal subunit. *Cell*, 103(7):1143–1154, 2000.

[32] W. O. Bullock, J. M. Fernandez, and J. M. Short. XL1-Blue - a high-efficiency plasmid transforming recA *Escherichia coli* strain with beta-galactosidase selection. *Biotechniques*, 5(4):376–, 1987.

[33] E. A. Campbell, N. Korzheva, A. Mustaev, K. Murakami, S. Nair, A. Goldfarb, and S. A. Darst. Structural mechanism for rifampicin inhibition of bacterial RNA polymerase. *Cell*, 104(6):901–912, 2001.

[34] V. L. Cavera, T. D. Arthur, D. Kashtanov, and M. L. Chikindas. Bacteriocins and their position in the next wave of conventional antibiotics. *International Journal of Antimicrobial Agents*, 46(5):494 – 501, 2015.

[35] Y. Cetinkaya, P. Falk, and C. G. Mayhall. Vancomycin-resistant Enterococci. *Clinical Microbiology Reviews*, 13(4):686–707, 2000.

[36] J. J. Champoux. DNA topoisomerases: Structure, function, and mechanism. *Annual Review of Biochemistry*, 70:369–413, 2001.

[37] J. Charney, W. P. Fisher, C. Curran, R. A. Machlowitz, and A. A. Tytell. Streptogramin, a new antibiotic. *Antibiotics and Chemotherapy*, 3(12):1283–1286, 1953.

[38] C. Chatterjee, M. Paul, L. L. Xie, and W. A. van der Donk. Biosynthesis and mode of action of lantibiotics. *Chemical Reviews*, 105(2):633–683, 2005.

[39] J. Z. Chen, S. Zhang, P. Cui, W. L. Shi, W. H. Zhang, and Y. Zhang. Identification of novel mutations associated with cycloserine resistance in *Mycobacterium tuberculosis*. *Journal of Antimicrobial Chemotherapy*, 72(12):3272–3276, 2017.

[40] J. Chiu, D. Tillett, I. W. Dawes, and P. E. March. Site-directed, ligase-independent mutagenesis (SLIM) for highly efficient mutagenesis of plasmids greater than 8kb. *Journal of Microbiological Methods*, 73(2):195–198, 2008.

[41] I. Chopra, L. Hesse, and A. J. O'Neill. Exploiting current understanding of antibiotic action for discovery of new drugs. *Journal of Applied Microbiology*, 92:4S–15S, 2002.

[42] J. Cleveland, T. J. Montville, I. F. Nes, and M. L. Chikindas. Bacteriocins: safe, natural antimicrobials for food preservation. *International Journal of Food Microbiology*, 71(1):1–20, 2001.

[43] R. M. Cleverley, Z. J. Rutter, J. Rismondo, F. Corona, H. C. T. Tsui, F. A. Alatawi, R. A. Daniel, S. Halbedel, O. Massidda, M. E. Winkler, and R. J. Lewis. The cell cycle regulator GpsB functions as cytosolic adaptor for multiple cell wall enzymes. *Nature Communications*, 10, 2019.

[44] A. Coates, Y. M. Hu, R. Bax, and C. Page. The future challenges facing the development of new antimicrobial drugs. *Nature Reviews Drug Discovery*, 1(11):895–910, 2002.

[45] F. Collin, S. Karkare, and A. Maxwell. Exploiting bacterial DNA gyrase as a drug target: current state and perspectives. *Applied Microbiology and Biotechnology*, 92(3):479–497, 2011.

[46] M. Collin and V. A. Fischetti. A novel secreted endoglycosidase from *Enterococcus faecalis* with activity on human immunoglobulin G and ribonuclease B. *Journal of Biological Chemistry*, 279(21):22558–22570, 2004.

[47] N. S. R. Colombo, M. C. Chalon, S. A. Navarro, and A. Bellomio. Pediocin-like bacteriocins: new perspectives on mechanism of action and immunity. *Current Genetics*, 64(2):345–351, 2018.

[48] R. N. Costilow and T. W. Humphreys. Nitrate reduction by certain strains of *Lactobacillus plantarum*. *Science*, 121(3136):168–168, 1955.

[49] P. D. Cotter, C. Hill, and R. P. Ross. Bacteriocins: Developing innate immunity for food. *Nature Reviews Microbiology*, 3(10):777–788, 2005.

[50] P. D. Cotter, R. P. Ross, and C. Hill. Bacteriocins - a viable alternative to antibiotics? *Nature Reviews Microbiology*, 11(2):95–105, 2013.

[51] A.-M. Crutz-Le Coq and M. Zagorec. Vectors for Lactobacilli and other Gram-positive bacteria based on the minimal replicon of pRV500 from *Lactobacillus sakei*. *Plasmid*, 60(3):212–220, 2008.

[52] K. Dalet, C. Briand, Y. Cenatiempo, and Y. Hechard. The rpoN gene of *Enterococcus faecalis* directs sensitivity to subclass IIa bacteriocins. *Current Microbiology*, 41(6):441–443, 2000.

[53] J. Davies and D. Davies. Origins and evolution of antibiotic resistance. *Microbiology and Molecular Biology Reviews*, 74(3):417–433, 2010.

[54] C. A. De Leon, P. M. Levine, T. W. Craven, and M. R. Pratt. The sulfur-linked analogue of O-GlcNAc (S-GlcNAc) is an enzymatically stable and reasonable structural surrogate for O-GlcNAc at the peptide and protein levels. *Biochemistry*, 56(27):3507–3517, 2017.

[55] W. H. Deitz, J. H. Bailey, and E. J. Froelich. *In vitro* antibacterial properties of nalidixic acid, a new drug active against Gram-negative organisms. *Antimicrobial Agents and Chemotherapy*, 161:583–7, 1963.

[56] E. L. Denham, S. Piersma, M. Rinket, E. Reilman, M. C. de Goffau, and J. M. van Dijl. Differential expression of a prophage-encoded glycocin and its immunity protein suggests a mutualistic strategy of a phage and its host. *Scientific Reports*, 9, 2019.

[57] J. Deutscher, F. M. D. Ake, M. Derkaoui, A. C. Zebre, T. N. Cao, H. Bouraoui, T. Kentache, A. Mokhtari, E. Milohanic, and P. Joyet. The bacterial phosphoenolpyruvate:carbohydrate phosphotransferase system: Regulation by protein

phosphorylation and phosphorylation-dependent protein-protein interactions. *Microbiology and Molecular Biology Reviews*, 78(2):231–256, 2014.

[58] J. Deutscher, C. Francke, and P. W. Postma. How phosphotransferase system-related protein phosphorylation regulates carbohydrate metabolism in bacteria. *Microbiology and Molecular Biology Reviews*, 70(4):939–+, 2006.

[59] D. B. Diep, M. Skaugen, Z. Salehian, H. Holo, and I. F. Nes. 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, 2007.

[60] R. Dorenbos, T. Stein, J. Kabel, C. Bruand, A. Bolhuis, S. Bron, W. J. Quax, and J. M. van Dijl. Thiol-disulfide oxidoreductases are essential for the production of the lantibiotic sublancin 168. *Journal of Biological Chemistry*, 277(19):16682–16688, 2002.

[61] S. M. Drawz and R. A. Bonomo. Three decades of β -lactamase inhibitors. *Clinical Microbiology Reviews*, 23(1):160–201, 2010.

[62] K. R. Drower. *The bacteriostatic diglycocylated bacteriocin glycocin F targets a sugar-specific transporter : a thesis presented in partial fulfilment of the requirements for the degree of Master of Science in Biochemistry at Massey University, Manawatu, New Zealand.* 2014.

[63] M. Dubecky, N. G. Walter, J. Sponer, M. Otyepka, and P. Banas. Chemical feasibility of the general acid/base mechanism of *glmS* ribozyme self-cleavage. *Biopolymers*, 103(10):550–562, 2015.

[64] J. Y. F. Dubois, T. Kouwen, A. K. C. Schurich, C. R. Reis, H. T. Ensing, E. N. Trip, J. C. Zweers, and J. M. van Dijl. Immunity to the bacteriocin sublancin 168 is determined by the SunI (YolF) protein of *Bacillus subtilis*. *Antimicrobial Agents and Chemotherapy*, 53(2):651–661, 2009.

[65] B. M. Duggar. Aureomycin - a product of the continuing search for new antibiotics. *Annals of the New York Academy of Sciences*, 51(2):177–, 1948.

[66] J. Ehrlich, Q. R. Bartz, R. M. Smith, D. A. Joslyn, and P. R. Burkholder. Chloromycetin, a new antibiotic from a soil Actinomycete. *Science*, 106(2757):417–417, 1947.

[67] M. E. Falagas, F. Athanasaki, G. L. Voulgaris, N. A. Triarides, and K. Z. Vardakas. Resistance to fosfomycin: Mechanisms, frequency and clinical consequences. *International Journal of Antimicrobial Agents*, 53(1):22–28, 2019.

[68] A. G. Fleming. Responsibilities and opportunities of the private practitioner in preventive medicine. *Canadian Medical Association Journal*, 20:11–13, 1929.

[69] C. Francke, T. G. Kormelink, Y. Hagemeijer, L. Overmars, V. Sluijter, R. Moezelaar, and R. J. Siezen. Comparative analyses imply that the enigmatic sigma factor 54 is a central controller of the bacterial exterior. *BMC Genomics*, 12, 2011.

[70] C. Gabrielsen, D. A. Brede, P. E. Hernández, I. F. Nes, and D. B. Diep. The maltose ABC transporter in *Lactococcus lactis* facilitates high-level sensitivity to the circular bacteriocin garvicin ML. *Antimicrobial Agents and Chemotherapy*, 56(6):2908–2915, 2012.

[71] A. Galinier and J. Deutscher. Sophisticated regulation of transcriptional factors by the bacterial phosphoenolpyruvate: Sugar phosphotransferase system. *Journal of Molecular Biology*, 429(6):773–789, 2017.

[72] C. V. Garcia De Gonzalo, E. L. Denham, R. A. T. Mars, J. Stülke, W. A. van der Donk, and J. M. van Dijl. The phosphoenolpyruvate:sugar phosphotransferase system is involved in sensitivity to the glucosylated bacteriocin sublancin. *Antimicrobial Agents and Chemotherapy*, 59(11):6844–6854, 2015.

[73] C. V. Garcia De Gonzalo, L. Zhu, T. J. Oman, and W. A. van der Donk. NMR structure of the S-linked glycopeptide sublancin 168. *ACS Chemical Biology*, 9(3):796–801, 2014.

[74] S. R. Gargis, A. S. Gargis, H. E. Heath, L. S. Heath, P. A. LeBlanc, M. M. Senn, B. Berger-Bachi, R. S. Simmonds, and G. L. Sloan. Zif, the zoocin a immunity factor, is a femabx-like immunity protein with a novel mode of action. *Applied and Environmental Microbiology*, 75(19):6205–6210, 2009.

[75] L. P. Garrod. Erythromycin group of antibiotics. *BMJ-British Medical Journal*, 2(13):57–63, 1957.

[76] E. Gavrish, C. S. Sit, S. G. Cao, O. Kandror, A. Spoering, A. Peoples, L. Ling, A. Fetterman, D. Hughes, A. Bissell, H. Torrey, T. Akopian, A. Mueller,

S. Epstein, A. Goldberg, J. Clardy, and K. Lewis. Lassomycin, a ribosomally synthesized cyclic peptide, kills *Mycobacterium tuberculosis* by targeting the ATP-dependent protease ClpC1P1P2. *Chemistry & Biology*, 21(4):509–518, 2014.

[77] M. Gaynor and A. S. Mankin. Macrolide antibiotics: Binding site, mechanism of action, resistance. *Current Topics in Medicinal Chemistry*, 3(9):949–960, 2003.

[78] H. Gelband, M. Miller-Petrie, S. Pant, S. Gandra, J. Levinson, D. Barter, A. White, and R. Laxminarayan. State of the world's antibiotics, 2015. Technical report, Center for Disease Dynamics, Economics & Policy, 2015.

[79] H. S. Gold and R. C. Moellering. Drug therapy - antimicrobial-drug resistance. *New England Journal of Medicine*, 335(19):1445–1453, 1996.

[80] D. A. Gray, G. Dugar, P. Gamba, H. Strahl, M. J. Jonker, and L. W. Hamoen. Extreme slow growth as alternative strategy to survive deep starvation in bacteria. *Nature Communications*, 10, 2019.

[81] H. E. Hasper, N. E. Kramer, J. L. Smith, J. D. Hillman, C. Zachariah, O. P. Kuipers, B. de Kruijff, and E. Breukink. An alternative bactericidal mechanism of action for lantibiotic peptides that target lipid II. *Science*, 313(5793):1636–1637, 2006.

[82] M. Hassan, M. Kjos, I. F. Nes, D. B. Diep, and F. Lotfipour. Natural antimicrobial peptides from bacteria: Characteristics and potential applications to fight against antibiotic resistance. *Journal of Applied Microbiology*, 113(4):723–736, 2012.

[83] T. Hata, R. Tanaka, and S. Ohmomo. Isolation and characterization of plantaricin ASM1: A new bacteriocin produced by *Lactobacillus plantarum* A-1. *International Journal of Food Microbiology*, 137(1):94–99, 2010.

[84] J. Heddle and A. Maxwell. Quinolone-binding pocket of DNA gyrase: Role of GyrB. *Antimicrobial Agents and Chemotherapy*, 46(6):1805–1815, 2002.

[85] D. Hendlin, E. O. Stapley, M. Jackson, H. Wallick, A. K. Miller, F. J. Wolf, T. W. Miller, L. Chaiet, F. M. Kahan, E. L. Foltz, and H. B. Woodruff. Phosphonomycin, a new antibiotic produced by strains of *Streptomyces*. *Science*, 166(3901):122–123, 1969.

[86] M. Iwatsuki, H. Tomoda, R. Uchida, H. Gouda, S. Hirono, and S. Omura. Lariatins, antimycobacterial peptides produced by *Rhodococcus sp* k01-b0171, have a lasso structure. *Journal of the American Chemical Society*, 128(23):7486–7491, 2006.

[87] V. S. Iyer and L. E. Hancock. Deletion of sigma(54) (*rpoN*) alters the rate of autolysis and biofilm formation in *Enterococcus faecalis*. *Journal of Bacteriology*, 194(2):368–375, 2012.

[88] E. Izquierdo, C. Wagner, E. Marchioni, D. Aoude-Werner, and S. Ennahar. Enterocin 96, a novel class II bacteriocin produced by *Enterococcus faecalis* WHE 96, isolated from munster cheese. *Applied and Environmental Microbiology*, 75(13):4273–4276, 2009.

[89] A. E. Jacob and S. J. Hobbs. Conjugal transfer of plasmid-borne multiple antibiotic-resistance in *Streptococcus faecalis* var zymogenes. *Journal of Bacteriology*, 117(2):360–372, 1974.

[90] A. Kaunietis, A. Buivydas, D. J. Citavicius, and O. P. Kuipers. Heterologous biosynthesis and characterization of a glycocin from a thermophilic bacterium. *Nature Communications*, 10, 2019.

[91] W. J. Kelly, R. V. Asmundson, and C. M. Huang. Characterization of plantaricin KW30, a bacteriocin produced by *Lactobacillus plantarum*. *Journal of Applied Bacteriology*, 81(6):657–662, 1996.

[92] A. P. Kerr. *The bacteriostatic spectrum and inhibitory mechanism of glycocin F, a bacteriocin from Lactobacillus plantarum KW30 : a thesis presented in partial fulfilment of the requirements for the degree of Master of Science in Microbiology at Massey University, Palmerston North, New Zealand / Andrew Philip Kerr.* 2013.

[93] M. Kjos, C. Oppegard, D. B. Diep, I. F. Nes, J. W. Veening, J. Nissen-Meyer, and T. Kristensen. Sensitivity to the two-peptide bacteriocin lactococcin G is dependent on UppP, an enzyme involved in cell-wall synthesis. *Molecular Microbiology*, 92(6):1177–1187, 2014.

[94] M. Kjos, Z. Salehian, I. F. Nes, and D. B. Diep. An extracellular loop of the mannose phosphotransferase system component IIC is responsible for specific

targeting by class IIa bacteriocins. *Journal of Bacteriology*, 192(22):5906–5913, 2010.

[95] M. Kleerebezem, M. M. Beerthuyzen, E. E. Vaughan, W. M. deVos, and O. P. Kuipers. Controlled gene expression systems for lactic acid bacteria: Transferable nisin-inducible expression cassettes for *Lactococcus*, *Leuconostoc*, and *Lactobacillus* spp. *Applied and Environmental Microbiology*, 63(11):4581–4584, 1997.

[96] P. J. Knerr and W. A. van der Donk. *Discovery, Biosynthesis, and Engineering of Lantipeptides*, volume 81 of *Annual Review of Biochemistry*, pages 479–505. 2012.

[97] M. A. Kohanski, D. J. Dwyer, and J. J. Collins. How antibiotics kill bacteria: From targets to networks. *Nature Reviews Microbiology*, 8(6):423–435, 2010.

[98] H. Komatsuzawa, T. Fujiwara, H. Nishi, S. Yamada, M. Ohara, N. McCallum, B. Berger-Bächi, and M. Sugai. The gate controlling cell wall synthesis in *Staphylococcus aureus*. *Molecular Microbiology*, 53(4):1221–1231, 2004.

[99] T. Kouwen, E. N. Trip, E. L. Denham, M. Sibbald, J. Y. F. Dubois, and J. M. van Dijl. The large mechanosensitive channel MscL determines bacterial susceptibility to the bacteriocin sublancin 168. *Antimicrobial Agents and Chemotherapy*, 53(11):4702–4711, 2009.

[100] O. P. Kuipers, M. M. Beerthuyzen, P. Deruyter, E. J. Luesink, and W. M. Devos. Autoregulation of nisin biosynthesis in *Lactococcus lactis* by signal transduction. *Journal of Biological Chemistry*, 270(45):27299–27304, 1995.

[101] A. Langdon, N. Crook, and G. Dantas. The effects of antibiotics on the microbiome throughout development and alternative approaches for therapeutic modulation. *Genome Medicine*, 8, 2016.

[102] B. Langmead and S. L. Salzberg. Fast gapped-read alignment with Bowtie 2. *Nature Methods*, 9(4):357–359, 2012.

[103] B. S. Laursen, H. P. Sorensen, K. K. Mortensen, and H. U. Sperling-Petersen. Initiation of protein synthesis in bacteria. *Microbiology and Molecular Biology Reviews*, 69(1):101–123, 2005.

[104] K. Lewis. Platforms for antibiotic discovery. *Nature Reviews Drug Discovery*, 12(5):371–387, 2013.

[105] M. I. Love, W. Huber, and S. Anders. Moderated estimation of fold change and dispersion for RNA-seq data with DESeq2. *Genome Biology*, 15(12), 2014.

[106] A. J. Macleod, G. Digout, R. L. Ozere, C. E. Vanrooyen, and H. B. Ross. Lincomycin - new antibiotic active against *Staphylococci* + other Gram-positive cocci - clinical + laboratory studies. *Canadian Medical Association Journal*, 91(20):1056–, 1964.

[107] M. A. Maky, N. Ishibashi, T. Zendo, R. H. Perez, J. R. Doud, M. Karmi, and K. Sonomoto. Enterocin F4-9, a novel O-linked glycosylated bacteriocin. *Applied and Environmental Microbiology*, 81(14):4819–4826, 2015.

[108] P. Margalith and G. Beretta. Rifomycin. xi. taxonomic study on *Streptomyces mediterranei* nov. sp. *Mycopathologia et Mycologia Applicata*, 13(4):321–330, 1960.

[109] K. Marild, W. M. Ye, B. Lebwohl, P. H. R. Green, M. J. Blaser, T. Card, and J. F. Ludvigsson. Antibiotic exposure and the development of coeliac disease: a nationwide case-control study. *BMC Gastroenterology*, 13, 2013.

[110] B. Martinez, A. Rodriguez, and J. E. Suarez. Lactococcin 972, a bacteriocin that inhibits septum formation in lactococci. *Microbiology-Uk*, 146:949–955, 2000.

[111] G. B. Martins, G. Giacomelli, O. Goldbeck, G. M. Seibold, and M. Bramkamp. Substrate-dependent cluster density dynamics of *Corynebacterium glutamicum* phosphotransferase system permeases. *Molecular Microbiology*, 111(5):1335–1354, 2019.

[112] J. M. McGuire, R. N. Wolfe, and D. W. Ziegler. Vancomycin, a new antibiotic. ii. *In vitro* antibacterial studies. *Antibiotics Annual*, 3:612–8, 1955.

[113] M. Metelev, J. I. Tietz, J. O. Melby, P. M. Blair, L. Y. Zhu, I. Livnat, K. Severinov, and D. A. Mitchell. Structure, bioactivity, and resistance mechanism of streptomonomicin, an unusual lasso peptide from an understudied halophilic actinomycete. *Chemistry & Biology*, 22(2):241–250, 2015.

[114] C. A. Michael, D. Dominey-Howes, and M. Labbate. The antimicrobial resistance crisis: Causes, consequences, and management. *Frontiers in Public Health*, 2, 2014.

[115] I. Mierau and M. Kleerebezem. 10 years of the nisin-controlled gene expression system (NICE) in *Lactococcus lactis*. *Applied Microbiology and Biotechnology*, 68(6):705–717, 2005.

[116] L. Minakhin, S. Bhagat, A. Brunning, E. A. Campbell, S. A. Darst, R. H. Ebright, and K. Severinov. Bacterial RNA polymerase subunit omega and eukaryotic RNA polymerase subunit RPB6 are sequence, structural, and functional homologs and promote RNA polymerase assembly. *Proceedings of the National Academy of Sciences of the United States of America*, 98(3):892–897, 2001.

[117] R. Nagar and A. Rao. An iterative glycosyltransferase EntS catalyzes transfer and extension of O- and S-linked monosaccharide in enterocin 96. *Glycobiology*, 27(8):766–776, 2017.

[118] J. Nissen-Meyer, C. Oppegard, P. Rogne, H. S. Haugen, and P. E. Kristiansen. Structure and mode-of-action of the two-peptide (Class-IIb) bacteriocins. *Probiotics and Antimicrobial Proteins*, 2(1):52–60, 2010.

[119] G. E. Norris and M. L. Patchett. The glycocins: In a class of their own. *Current Opinion in Structural Biology*, 40:112–119, 2016.

[120] J. M. Ogle, A. P. Carter, and V. Ramakrishnan. Insights into the decoding mechanism from recent ribosome structures. *Trends in Biochemical Sciences*, 28(5):259–266, 2003.

[121] T. J. Oman, J. M. Boettcher, H. A. Wang, X. N. Okalibe, and W. A. van der Donk. Sublancin is not a lantibiotic but an S-linked glycopeptide. *Nature Chemical Biology*, 7(2):78–80, 2011.

[122] C. Oppegard, M. Kjos, J. W. Veening, J. Nissen-Meyer, and T. Kristensen. A putative amino acid transporter determines sensitivity to the two-peptide bacteriocin plantaricin JK. *MicrobiologyOpen*, 5(4):700–708, 2016.

[123] T. Otaka and A. Kaji. Mode of action of bottromycin A2 - effect of bottromycin A2 on polysomes. *FEBS Letters*, 153(1):53–59, 1983.

[124] S. H. Paik, A. Chakicherla, and J. N. Hansen. 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, 1998.

[125] K. Papadimitriou, A. Alegria, P. A. Bron, M. de Angelis, M. Gobbetti, M. Kleerebezem, J. A. Lemos, D. M. Linares, P. Ross, C. Stanton, F. Turroni, D. van Sinderen, P. Varmanen, M. Ventura, M. Zuniga, E. Tsakalidou, and J. Kok. Stress physiology of lactic acid bacteria. *Microbiology and Molecular Biology Reviews*, 80(3):837–890, 2016.

[126] J. Plumbridge. An alternative route for recycling of N-acetylglucosamine from peptidoglycan involves the N-acetylglucosamine phosphotransferase system in *Escherichia coli*. *Journal of Bacteriology*, 191(18):5641–5647, 2009.

[127] A. Prince, P. Sandhu, P. Kumar, E. Dash, S. Sharma, M. Arakha, S. Jha, Y. Akhter, and M. Saleem. Lipid-II independent antimicrobial mechanism of nisin depends on its crowding and degree of oligomerization. *Scientific Reports*, 6, 2016.

[128] G. A. Prosser and L. P. S. de Carvalho. Reinterpreting the mechanism of inhibition of *Mycobacterium tuberculosis* D-alanine:D-alanine ligase by D-cycloserine. *Biochemistry*, 52(40):7145–7149, 2013.

[129] R Core Team. *R: A Language and Environment for Statistical Computing*. R Foundation for Statistical Computing, Vienna, Austria, 2019.

[130] V. Ramakrishnan. Ribosome structure and the mechanism of translation. *Cell*, 108(4):557–572, 2002.

[131] S. J. Ran, B. Liu, W. Jiang, Z. Sun, and J. P. Liang. Transcriptome analysis of *Enterococcus faecalis* in response to alkaline stress. *Frontiers in Microbiology*, 6, 2015.

[132] H. Ren, S. Biswas, S. Ho, W. A. van der Donk, and H. Zhao. Rapid discovery of glycocins through pathway refactoring in *Escherichia coli*. *ACS Chemical Biology*, 13(10):2966–2972, 2018.

[133] J. A. Robles, S. E. Qureshi, S. J. Stephen, S. R. Wilson, C. J. Burden, and J. M. Taylor. Efficient experimental design and analysis strategies for the

detection of differential expression using RNA-sequencing. *BMC Genomics*, 13, 2012.

[134] R. Roskoski, W. Gevers, H. Kleinkauf, and F. Lipmann. Tyrocidine biosynthesis by 3 complementary fractions from *Bacillus brevis* (ATCC 8185). *Biochemistry*, 9(25):4839–+, 1970.

[135] K. S. Ryu, C. Kim, I. Kim, S. Yoo, B. S. Choi, and C. Park. NMR application probes a novel and ubiquitous family of enzymes that alter monosaccharide configuration. *Journal of Biological Chemistry*, 279(24):25544–25548, 2004.

[136] D. F. Sahm, J. Kissinger, M. S. Gilmore, P. R. Murray, R. Mulder, J. Solliday, and B. Clarke. *In vitro* susceptibility studies of vancomycin-resistant *Enterococcus faecalis*. *Antimicrobial Agents and Chemotherapy*, 33(9):1588–1591, 1989.

[137] M. H. Saier. The bacterial phosphotransferase system: New frontiers 50 years after its discovery. *Journal of Molecular Microbiology and Biotechnology*, 25(2-3):73–78, 2015.

[138] A. Schatz, E. Bugie, and S. A. Waksman. Streptomycin, a substance exhibiting antibiotic activity against gram-positive and gram-negative bacteria. *Proceedings of the Society for Experimental Biology and Medicine*, 55(1):66–69, 1944.

[139] F. Schluenzen, C. Takemoto, D. N. Wilson, T. Kaminishi, J. M. Harms, K. Hanawa-Suetsugu, W. Szaflarski, M. Kawazoe, M. Shirouzo, K. H. Nierhaus, S. Yokoyama, and P. Fucini. The antibiotic kasugamycin mimics mRNA nucleotides to destabilize tRNA binding and inhibit canonical translation initiation. *Nature Structural & Molecular Biology*, 13(10):871–878, 2006.

[140] F. Schlunzen, R. Zarivach, J. Harms, A. Bashan, A. Tocilj, R. Albrecht, A. Yonath, and F. Franceschi. Structural basis for the interaction of antibiotics with the peptidyl transferase centre in eubacteria. *Nature*, 413(6858):814–821, 2001.

[141] D. Scholl. *Phage Tail-Like Bacteriocins*, volume 4 of *Annual Review of Virology*, pages 453–467. 2017.

[142] A. Schuller, D. Matzner, C. E. Lunse, V. Wittmann, C. Schumacher, S. Unsleber, H. Brotz-Oesterhelt, C. Mayer, G. Bierbaum, and G. Mayer. Activation

of the glmS ribozyme confers bacterial growth inhibition. *ChemBioChem*, 18(5):435–440, 2017.

[143] O. Skold. Sulfonamide resistance: mechanisms and trends. *Drug Resistance Updates*, 3(3):155–160, 2000.

[144] M. Solheim, A. Aakra, H. Vebo, L. Snipen, and I. F. Nes. Transcriptional responses of *Enterococcus faecalis* V583 to bovine bile and sodium dodecyl sulfate. *Applied and Environmental Microbiology*, 73(18):5767–5774, 2007.

[145] J. Stepper, S. Shastri, T. S. Loo, J. C. Preston, P. Novak, P. Man, C. H. Moore, V. Havlíček, M. L. Patchett, and G. E. Norris. Cysteine S-glycosylation, a new post-translational modification found in glycopeptide bacteriocins. *FEBS Letters*, 585(4):645–650, 2011.

[146] M. J. A. Stevens, D. Molenaar, A. de Jong, W. M. De Vos, and M. Kleerebezem. Sigma 54-mediated control of the mannose phosphotransferase sytem in *Lactobacillus plantarum* impacts on carbohydrate metabolism. *Microbiology*, 156:695–707, 2010.

[147] D. R. Storm, K. S. Rosenthal, and P. E. Swanson. Polymyxin and related peptide antibiotics. *Annual Review of Biochemistry*, 46:723–763, 1977.

[148] R. I. Subbotin and B. T. Chait. A pipeline for determining protein-protein interactions and proximities in the cellular milieu. *Molecular & Cellular Proteomics*, 13(11):2824–2835, 2014.

[149] Z. Z. Sun, J. Zhong, X. B. Liang, J. L. Liu, X. Z. Chen, and L. D. Huan. Novel mechanism for nisin resistance via proteolytic degradation of nisin by the nisin resistance protein NSR. *Antimicrobial Agents and Chemotherapy*, 53(5):1964–1973, 2009.

[150] P. M. Swe, G. M. Cook, J. R. Tagg, and R. W. Jack. Mode of action of dysgalacticin: A large heat-labile bacteriocin. *Journal of Antimicrobial Chemotherapy*, 63(4):679–686, 2009.

[151] S. Thennarasu, D. K. Lee, A. Poon, K. E. Kawulka, J. C. Vederas, and A. Ramamoorthy. Membrane permeabilization, orientation, and antimicrobial mechanism of subtilosin A. *Chemistry and Physics of Lipids*, 137(1-2):38–51, 2005.

[152] J. I. Tietz, C. J. Schwalen, P. S. Patel, T. Maxson, P. M. Blair, H. C. Tai, U. I. Zakai, and D. A. Mitchell. A new genome-mining tool redefines the lasso peptide biosynthetic landscape. *Nature Chemical Biology*, 13(5):470–478, 2017.

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

[154] F. T. F. Tsai, O. M. P. Singh, T. Skarzynski, A. J. Wonacott, S. Weston, A. Tucker, R. A. Pauptit, A. L. Breeze, J. P. Poyser, R. Obrien, J. E. Ladbury, and D. B. Wigley. The high-resolution crystal structure of a 24-kDa gyrase B fragment from *E. coli* complexed with one of the most potent coumarin inhibitors, clorobiocin. *Proteins-Structure Function and Bioinformatics*, 28(1):41–52, 1997.

[155] H. Tsuda, Y. Yamashita, Y. Shibata, Y. Nakano, and T. Koga. Genes involved in bacitracin resistance in *Streptococcus mutans*. *Antimicrobial Agents and Chemotherapy*, 46(12):3756–3764, 2002.

[156] G. Uzelac, M. Kojic, J. Lozo, T. Aleksandrzak-Piekarczyk, C. Gabrielsen, T. Kristensen, I. F. Nes, D. B. Diep, and L. Topisirovic. A Zn-dependent metallopeptidase is responsible for sensitivity to LsbB, a class II leaderless bacteriocin of *Lactococcus lactis* subsp lactis BGMN1-5. *Journal of Bacteriology*, 195(24):5614–5621, 2013.

[157] A. J. van Heel, A. de Jong, M. Montalban-Lopez, J. Kok, and O. P. Kuipers. BAGEL3: automated identification of genes encoding bacteriocins and (non)-bactericidal posttranslationally modified peptides. *Nucleic Acids Research*, 41(W1):W448–W453, 2013.

[158] H. C. Vebo, L. Snipen, I. F. Nes, and D. A. Brede. The transcriptome of the nosocomial pathogen *Enterococcus faecalis* V583 reveals adaptive responses to growth in blood. *PLoS One*, 4(11), 2009.

[159] H. Venugopal, P. J. B. Edwards, M. Schwalbe, J. K. Claridge, D. S. Libich, J. Stepper, T. Loo, M. L. Patchett, G. E. Norris, and S. M. Pascal. Structural, dynamic, and chemical characterization of a novel S-glycosylated bacteriocin. *Biochemistry*, 50(14):2748–2755, 2011.

[160] M. Vila-Perello, M. R. Pratt, F. Tulin, and T. W. Muir. Covalent capture of phospho-dependent protein oligomerization by site-specific incorporation of a diazirine photo-cross-linker. *Journal of the American Chemical Society*, 129(26):8068–8069, 2007.

[161] H. Wang, T. J. Oman, R. Zhang, C. V. Garcia De Gonzalo, Q. Zhang, and W. A. van der Donk. The glycosyltransferase involved in thurandacin biosynthesis catalyzes both O- and S-glycosylation. *Journal of the American Chemical Society*, 136(1):84–87, 01 2014.

[162] H. Wang and W. A. van der Donk. Substrate selectivity of the sublancin S-glycosyltransferase. *Journal of the American Chemical Society*, 133(41):16394–16397, 2011.

[163] A. Wegkamp, B. Teusink, W. M. de Vos, and E. J. Smid. Development of a minimal growth medium for *Lactobacillus plantarum*. *Letters in Applied Microbiology*, 50(1):57–64, 2010.

[164] G. Werner, R. J. L. Willems, B. Hildebrandt, I. Klare, and W. Witte. Influence of transferable genetic determinants on the outcome of typing methods commonly used for *Enterococcus faecium*. *Journal of Clinical Microbiology*, 41(4):1499–1506, 2003.

[165] D. N. Wilson. The A-Z of bacterial translation inhibitors. *Critical Reviews in Biochemistry and Molecular Biology*, 44(6):393–433, 2009.

[166] W. C. Winkler, A. Nahvi, A. Roth, J. A. Collins, and R. R. Breaker. Control of gene expression by a natural metabolite-responsive ribozyme. *Nature*, 428(6980):281–286, 2004.

[167] W. B. Wood. Host specificity of DNA produced by *Escherichia coli* - bacterial mutations affecting restriction and modification of DNA. *Journal of Molecular Biology*, 16(1):118–+, 1966.

[168] C. Wu, S. Biswas, C. V. Garcia De Gonzalo, and W. A. van der Donk. Investigations into the mechanism of action of sublancin. *ACS Infectious Diseases*, 5(3):454–459, 2018.

[169] J. Yamagishi, H. Yoshida, M. Yamayoshi, and S. Nakamura. Nalidixic acid-resistant mutations of the *gyrB* gene of *Escherichia coli*. *Molecular and General Genetics*, 204(3):367–373, 1986.

[170] J. I. Yamagishi, Y. Furutani, S. Inoue, T. Ohue, S. Nakamura, and M. Shimizu. New nalidixic-acid resistance mutations related to deoxyribonucleic-acid gyrase activity. *Journal of Bacteriology*, 148(2):450–458, 1981.

[171] C. Yanisch-Perron, J. Vieira, and J. Messing. Improved M13 phage cloning vectors and host strains - nucleotide-sequences of the m13MP18 and pUC19 vectors. *Gene*, 33(1):103–119, 1985.

[172] R. R. Yocum, J. R. Rasmussen, and J. L. Strominger. The mechanism of action of penicillin - penicillin acylates the active-site of *Bacillus stearothermophilus* D-alanine carboxypeptidase. *Journal of Biological Chemistry*, 255(9):3977–3986, 1980.

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

[174] F. Yoneyama, Y. Imura, K. Ohno, T. Zendo, J. Nakayama, K. Matsuzaki, and K. Sonomoto. Peptide-lipid huge toroidal pore, a new antimicrobial mechanism mediated by a lactococcal bacteriocin, lacticin Q. *Antimicrobial Agents and Chemotherapy*, 53(8):3211–3217, 2009.

[175] H. Yoshida, M. Bogaki, M. Nakamura, and S. Nakamura. Quinolone resistance-determining region in the DNA gyrase *gyrA* gene of *Escherichia coli*. *Antimicrobial Agents and Chemotherapy*, 34(6):1271–1272, 1990.

[176] S. H. Yu, M. Boyce, A. M. Wands, M. R. Bond, C. R. Bertozzi, and J. J. Kohler. Metabolic labeling enables selective photocrosslinking of O-GlcNAc-modified proteins to their binding partners. *Proceedings of the National Academy of Sciences of the United States of America*, 109(13):4834–4839, 2012.

[177] L. Zaffiri, J. Gardner, and L. H. Toledo-Pereyra. History of antibiotics. from salvarsan to cephalosporins. *Journal of Investigative Surgery*, 25(2):67–77, 2012.

[178] S. Zheng and K. Sonomoto. Diversified transporters and pathways for bacteriocin secretion in gram-positive bacteria. *Applied Microbiology and Biotechnology*, 102(10):4243–4253, 2018.

Appendices

A | Supplementary RNA-seq data

A.1 Supplementary tables

Table A.1: Differentially expressed genes exclusive to VI01 100 nM GccF treatment

ORF	Log_2-fold change	p_{adj}	Annotation
ERT20811	-5.42	7.89E-103	major facilitator family transporter(EF0082)
ERT21734	-2.75	3.39E-11	ABC transporter permease(EF2223)
ERT23742	-2.58	2.61E-05	bifunctional pyrimidine regulatory protein PyrR/uracil phosphoribosyltransferase(pyrR)
ERT23938	-2.50	3.05E-06	hypothetical protein(EF1959)
ERT25751	-2.46	2.25E-11	pheromone cAD1 lipoprotein(EF3256)
ERT21232	-2.45	2.02E-11	2-dehydropantoate 2-reductase(EF0517)
ERT23257	-2.43	1.01E-04	oxidoreductase(EF1226)
ERT20943	-2.39	3.99E-03	chlorohydrolase/aminohydrolase(EF2582)
ERT23256	-2.33	3.54E-05	thiamin biosynthesis ApbE(EF1225)
ERT23937	-2.25	3.09E-06	deoxyguanosinetriphosphate triphosphohydrolase-like protein(EF1958)
ERT23238	-2.01	1.90E-02	CCS family citrate carrier protein(EF1207)
ERT23104	-2.00	1.50E-04	hexapeptide repeat-containing acetyltransferase(EF1066)
ERT24588	-1.99	1.64E-02	signal peptidase I(EF0854)
ERT21733	-1.99	3.27E-05	ABC transporter permease(EF2222)
ERT21097	-1.94	5.95E-08	anaerobic ribonucleoside triphosphate reductase(nrdD)
ERT20918	-1.93	2.49E-06	fumarate reductase flavoprotein subunit(EF2556)
ERT23534	-1.92	5.25E-05	fructose-1;6-bisphosphatase(EF1503)
ERT23364	-1.86	6.86E-03	TetR family transcriptional regulator(EF1326)
ERT20439	-1.80	4.91E-05	sodium ion-translocating decarboxylase subunit beta(EF3324)
ERT20441	-1.79	4.65E-05	[citrate (pro-3S)-lyase] ligase(citC)

Continued on next page

Table A.1 – continued from previous page

ORF	Log_2-fold change	p_{adj}	Annotation
ERT20443	-1.78	1.35E-04	citrate lyase subunit beta(citE)
ERT27245	-1.73	3.01E-04	hypothetical protein(EF3281)
ERT23741	-1.70	5.11E-03	uracil permease(EF1720)
ERT20445	-1.63	1.07E-02	2-(5"-triphosphoribosyl)-3'-dephosphocoenzyme-A synthase(citX)
ERT23544	-1.59	4.48E-02	BglG family transcriptional antiterminator (EF1515)
ERT23899	-1.57	1.52E-02	ROK family protein(EF1912)
ERT20444	-1.53	3.94E-02	citrate lyase subunit alpha(citF)
ERT23728	-1.50	5.00E-02	glycosyl hydrolase(EF1707)
ERT21370	-1.48	1.53E-02	3-oxoacyl-ACP synthase(fabF-1)
ERT21165	-1.45	3.80E-02	acyl carrier protein(acpP)
ERT21359	1.43	1.72E-02	hypothetical protein(EF0294)
ERT24649	1.49	2.73E-02	50S ribosomal protein L35(rpmI)
ERT23296	1.57	1.76E-02	cation transporter E1-E2 family ATPase(EF1268)
ERT24614	1.61	1.45E-03	transcriptional regulator NrdR(nrdR)
ERT25923	1.64	1.13E-02	hypothetical protein(EF3055)
ERT24276	1.64	9.54E-04	hypothetical protein(EF1047)
ERT21293	1.64	3.94E-02	keto-hydroxyglutarate-aldolase/keto-deoxy-phosphogluconate aldolase(eda-1)
ERT24648	1.65	4.30E-04	translation initiation factor IF-3(infC)
ERT21240	1.66	1.90E-02	phage integrase family site specific recombinase(EF0479)
ERT24555	1.69	1.34E-02	hypothetical protein(EF0819)
ERT21855	1.74	1.45E-03	PadR family transcriptional regulator(EF2428)
ERT23360	1.74	4.55E-05	hypothetical protein(EF1322)
ERT21113	1.77	5.36E-06	hypothetical protein(EF2771)
ERT25356	1.78	9.94E-06	hypothetical protein(EF0637)
ERT21905	1.78	3.03E-06	hypothetical protein(EF2483)
ERT21291	1.83	1.28E-03	5-keto-4-deoxyuronate isomerase(kduI-1)
ERT23520	1.87	3.99E-03	hypothetical protein(EF1489)
ERT20460	1.88	2.36E-07	hypothetical protein(EF3301)
ERT21727	1.89	4.88E-06	hypothetical protein(EF2216)
ERT25955	1.91	2.26E-08	hypothetical protein(EF3021)
ERT24667	1.94	3.01E-09	hypothetical protein(EF0933)
ERT20800	1.96	1.86E-06	lipoprotein(EF0095)
ERT25471	2.00	1.06E-06	hypothetical protein(EF0747)

Continued on next page

Table A.1 – continued from previous page

ORF	Log_2-fold change	p_{adj}	Annotation
ERT25843	2.02	8.85E-09	hypothetical protein(EF3146)
ERT21274	2.05	7.85E-07	LysM domain-containing protein(EF0443)
ERT24693	2.07	2.33E-07	pyrroline-5-carboxylate reductase(EF0961)
ERT21190	2.09	2.78E-10	hypothetical protein(EF2913)
ERT25800	2.10	1.73E-13	lipase(EF3191)
ERT21292	2.13	1.06E-06	2-dehydro-3-deoxygluconokinase(EF0424)
ERT25920	2.14	5.33E-06	hypothetical protein(EF3057)
ERT21188	2.15	1.13E-13	LuxR family DNA-binding response regulator(EF2911)
ERT21722	2.22	8.64E-08	hypothetical protein(EF2211)
ERT21048	2.22	1.90E-11	tellurite resistance protein(EF2698)
ERT21251	2.29	1.09E-15	glucosamine-6-phosphate isomerase(nagB)
ERT21189	2.40	8.65E-17	sensor histidine kinase(EF2912)
ERT23958	2.44	3.46E-09	universal stress protein(EF1982)
ERT21249	2.49	2.17E-10	LemA family protein(EF0468)
ERT24616	2.50	2.34E-21	primosomal protein DnaI(dnaI)
ERT23709	2.55	4.26E-13	hypothetical protein(EF1686)
ERT23696	2.62	5.14E-03	ABC transporter ATP-binding protein(EF1673)
ERT24545	2.79	1.10E-27	hypothetical protein(EF0798)
ERT21215	2.86	2.51E-14	ribose uptake protein(EF2959)
ERT23772	2.87	7.47E-28	hypothetical protein(EF1752)
ERT23891	2.91	7.37E-34	hypothetical protein(EF1903)
ERT23771	2.95	4.45E-27	hypothetical protein(EF1751)
ERT24544	3.01	2.93E-35	hypothetical protein(EF0797)
ERT24594	3.02	4.26E-28	cation efflux family protein(EF0859)
ERT21047	3.24	1.13E-29	hypothetical protein(EF2697)
ERT21248	3.28	8.73E-27	hypothetical protein(EF0469)
ERT23695	3.33	7.43E-24	permease(EF1672)
ERT23892	3.56	2.27E-76	glycerophosphoryl diester phosphodiesterase family protein(EF1904)

Note: JH2-2 gene labels shown. Genes that have no match to JH2-2 labels have been omitted.

Table A.2: Top 10 up-regulated genes for each timepoint for experiment 1

Time	ORF	Annotation†	Log_2-fold change
0 minutes	ERT21738	ABC transporter ATP-binding protein/permease (EF2227)	4.5
	ERT21737	ABC transporter ATP-binding protein/permease (EF2226)	4.3
	ERT23105	Hypothetical protein (EF1067)	4.1
	ERT20555	Hypothetical protein (EF0026)	3.7
	ERT24666	Hypothetical protein (EF0932)	3.6
	ERT23442	Hypothetical protein (EF1412)	3.5
	ERT21216	D-ribose pyranase (EF2960)	3.5
	ERT24556	50S ribosomal protein L25/general stress protein Ctc (rplY)	2.9
	ERT23333	Magnesium-translocating P-type ATPase (EF1304)	2.9
	ERT23416	Hypothetical protein: aquaporin?	2.9
10 minutes	ERT21738	ABC transporter ATP-binding protein/permease (EF2227)	6.6
	ERT23105	Hypothetical protein (EF1067)	6.6
	ERT21737	ABC transporter ATP-binding protein/permease (EF2226)	4.8
	ERT25357	Hypothetical protein (EF0638)	4.8
	ERT21647	Multidrug resistance protein (EF2068)	4.7
	ERT23333	Magnesium-translocating P-type ATPase (EF1304)	4.7
	ERT25827	Methionine sulfoxide reductase B (EF3164)	4.5
	ERT24666	Hypothetical protein (EF0932)	4.4
	ERT23134	Collagen adhesin protein (EF1099)	4.3
	ERT23092	Hypothetical protein: no domains (tRNA encoding?)	4.2
30 minutes	ERT21738	ABC transporter ATP-binding protein/permease (EF2227)	9.4
	ERT23105	Hypothetical protein (EF1067)	6.6
	ERT20732	Iron ABC transporter substrate-binding protein (EF0188)	6.2
	ERT21737	ABC transporter ATP-binding protein/permease (EF2226)	5.7
	ERT23625	Catalase/peroxidase (katA)	5.4
	ERT25445	1-phosphofructokinase (fruK-2)	5.3
	ERT21647	Multidrug resistance protein (EF2068)	5.1
	ERT23134	Collagen adhesin protein (EF1099)	5.1
	ERT25393	Hypothetical protein (EF0673)	5.0
	ERT23619	N-acetyltransferase (EF1590)	5.0

† Names/ numbers in parenthese indicate gene labels for strain V583.

Table A.3: Top 10 down-regulated genes for each time point for experiment 1

Time	ORF	Annotation†	Log_2-fold change
0 minutes	ERT25488	Hypothetical protein (EF0764)	-5.1
	ERT23292	Hypothetical protein (EF1263)	-3.7
	ERT21384	Hypothetical protein: CAT domain; PRD domain	-3.6
	ERT24749	PTS system transporter subunit IIA (EF1018)	-3.6
	ERT24637	Mevalonate kinase (mvk)	-3.6
	ERT24743	PTS system transporter subunit IIB (EF1012)	-3.3
	ERT21274	LysM domain-containing protein (EF0443)	-3.1
	ERT24699	Hypothetical protein: ABC transporter	-3.1
	ERT24748	PTS system transporter subunit IIB (EF1017)	-3.1
	ERT23629	Glycosyl hydrolase (EF1602)	-2.9
10 minutes	ERT24749	PTS system transporter subunit IIA (EF1018)	-5.6
	ERT24743	PTS system transporter subunit IIB (EF1012)	-5.3
	ERT23249	Spermidine/putrescine ABC transporter permease (EF1218)	-5.1
	ERT23292	Hypothetical protein (EF1263)	-4.8
	ERT25488	Hypothetical protein (EF0764)	-4.7
	ERT24750	PTS system transporter subunit IIC (EF1019)	-4.6
	ERT23251	Spermidine/putrescine ABC transporter ATP-binding protein (EF1220)	-4.4
	ERT23250	Spermidine/putrescine ABC transporter permease (EF1219)	-4.3
	ERT24024	23S rRNA (adenine(2503)-C(2))-methyltransferase RlmN (EF2048) N	-3.9
	ERT21384	Hypothetical protein: CAT domain; PRD domain	-3.8
30 minutes	ERT23251	Spermidine/putrescine ABC transporter ATP-binding protein (EF1220)	-6.4
	ERT23249	Spermidine/putrescine ABC transporter permease (EF1218)	-6.3
	ERT23250	Spermidine/putrescine ABC transporter permease (EF1219)	-6.1
	ERT20811	Major facilitator family transporter (EF0082)	-6.0
	ERT25488	Hypothetical protein (EF0764)	-6.0
	ERT24750	PTS system transporter subunit IIC (EF1019)	-5.6
	ERT24751	Glycosyl hydrolase (EF1020)	-5.1
	ERT24749	PTS system transporter subunit IIA (EF1018)	-5.0
	ERT20810	Hypothetical protein (EF0083)	-5.0
	ERT23640	Formate acetyltransferase (pflB)	-5.0

† Names/ numbers in parenthese indicate gene labels for strain V583.

Table A.4: Top 10 up-regulated genes for each sample for experiment 2

Experiment	**ORF**	**Annotations**[†]	**Log_2-fold change**
20 nM GccF	ERT21738	ABC transporter ATP-binding protein/permease (EF2227)	3.9
	ERT20732	Iron ABC transporter substrate-binding protein (EF0188)	3.5
	ERT25354	Amino acid permease (EF0635)	3.4
	ERT25353	Decarboxylase (EF0634)	3.4
	ERT23532	Hypothetical protein (EF1501)	3.3
	ERT23531	V-type ATP synthase subunit D (EF1500)	3.3
	ERT24666	Hypothetical protein (EF0932)	3.2
	ERT23333	Magnesium-translocating P-type ATPase (EF1304)	3.1
	ERT23525	V-type ATP synthase subunit K (EF1494)	3.1
	ERT23529	V-type ATP synthase subunit A (EF1498)	3.1
20 nM GccF,	ERT21738	ABC transporter ATP-binding protein/permease (EF2227)	7.3
GlcNAc pre-treated	ERT21737	ABC transporter ATP-binding protein/permease (EF2226)	5.5
	ERT23105	Hypothetical protein (EF1067)	5.1
	ERT25357	Hypothetical protein (EF0638)	4.3
	ERT24666	Hypothetical protein (EF0932)	4.3
	ERT23333	Magnesium-translocating P-type ATPase (EF1304)	4.2
	ERT23134	Collagen adhesin protein (EF1099)	4.1
	ERT20732	Iron ABC transporter substrate-binding protein (EF0188)	4.0
	ERT21647	Multidrug resistance protein (EF2068)	3.8
	ERT25827	Methionine sulfoxide reductase B (EF3164)	3.8
100 nM, Sigma 54 K/O	ERT21738	ABC transporter ATP-binding protein/permease (EF2227)	5.7
	ERT23564	Hypothetical protein (EF1533)	5.3
	ERT24666	Hypothetical protein (EF0932)	5.1
	ERT25357	Hypothetical protein (EF0638)	5.0
	ERT20555	Hypothetical protein (EF0026)	4.6
	ERT20732	Iron ABC transporter substrate-binding protein (EF0188)	4.5
	ERT23773	Hypothetical protein (EF1753)	4.5
	ERT21066	ABC transporter ATP-binding protein (EF2720)	4.4
	ERT23333	Hagnesium-translocating P-type ATPase (EF1304)	4.4
	ERT23442	Hypothetical protein (EF1412)	4.3

[†] Names/ numbers in parenthese indicate gene labels for strain V583.

Table A.5: Top 10 down-regulated genes for each sample for experiment 2

Experiment	**ORF**	**Annotation**[†]	**Log_2-fold change**
20 nM GccF	ERT23254	Chlorohydrolase (EF1223)	-4.0
	ERT23252	Spermidine/putrescine ABC transporter (EF1221)	-3.9
	ERT23152	Amino acid ABC transporter amino acid-binding protein (EF1119)	-3.9
	ERT23253	Adenine deaminase (ade)	-3.7
	ERT23153	Amino acid ABC transporter ATP-binding protein (EF1120)	-3.5
	ERT23250	Spermidine/putrescine ABC transporter permease (EF1219)	-3.4
	ERT23151	Amino acid ABC transporter permease (EF1118)	-3.4
	ERT23251	Spermidine/putrescine ABC transporter ATP-binding protein (EF1220)	-3.4
	ERT23249	Spermidine/putrescine ABC transporter permease (EF1218)	-3.3
	ERT23150	Amino acid ABC transporter permease (EF1117)	-3.3
20 nM GccF,	ERT20801	Formate/nitrite transporter family protein (EF0094)	-4.7
GlcNAc pre-treated	ERT20811	Major facilitator family transporter (EF0082)	-4.6
	ERT23250-	Spermidine/putrescine ABC transporter permease (EF1219)	-4.5
	ERT24749	PTS system transporter subunit IIA (EF1018)	-4.5
	ERT24628	Amino acid ABC transporter amino acid-binding/permease (EF0893)	-4.4
	ERT23251	Spermidine/putrescine ABC transporter ATP-binding protein (EF1220)	-4.3
	ERT23249	Spermidine/putrescine ABC transporter permease (EF1218)	-4.2
	ERT24748	PTS system transporter subunit IIB (EF1017)	-4.2
	ERT23252	Spermidine/putrescine ABC transporter (EF1221)	-4.2
	ERT25488	Hypothetical protein (EF0764)	-3.9
100 nM, Sigma 54 K/O	ERT20811	Major facilitator family transporter (EF0082)	-5.4
	ERT20810	Hypothetical protein (EF0083)	-4.8
	ERT208	Formate/nitrite transporter family protein (EF0094)	-3.3
	ERT27246	U32 family peptidase (EF3280)	-3.2
	ERT27247	U32 family peptidase (EF3279)	-3.0
	ERT20437	Hypothetical protein (EF3326)	-2.9
	ERT20979	Cadmium-translocating P-type ATPase (cadA)	-2.8
	ERT21734	ABC transporter permease (EF2223)	-2.7
	ERT24024	23S rRNA (adenine(2503)-C(2))-methyltransferase RlmN (EF2048)	-2.7
	ERT20954	TetR family transcriptional regulator (EF2594)	-2.7

[†] Names/ numbers in parenthese indicate gene labels for strain V583.

Table A.6: DEGs common to all six *E. faecalis* RNA-seq experiments

JH2-2 ORF	Annotation	Up- or down-regulated
ERT23907	ribonucleoside hydrolase RihC(EF1921)	Down
ERT24629	glycerol dehydrogenase(EF0895)	Down
ERT24627	amino acid ABC transporter ATP-binding protein(EF0892)	Down
ERT23292	hypothetical protein(EF1263)	Down
ERT24628	amino acid ABC transporter amino acid-binding/permease(EF0893)	Down
ERT20954	TetR family transcriptional regulator(EF2594)	Down
ERT21868	hypothetical protein(EF2441)	Down
ERT23533	beta-lactamase(EF1502)	Down
ERT21357	Na+/H+ antiporter(napA)	Up
ERT25395	glycine betaine/carnitine/choline ABC transporter glycine betaine/carnitine/choline-binding protein(EF0675)	Up
ERT23773	hypothetical protein(EF1753)	Up
ERT21057	membrane protein(EF2708)	Up
ERT25357	hypothetical protein(EF0638)	Up
ERT23927	hypothetical protein(EF1947)	Up
ERT23381	magnesium-translocating P-type ATPase(EF1352)	Up
ERT23523	V-type ATPase subunit F(EF1492)	Up
ERT25761	hypothetical protein(EF3239)	Up
ERT20555	hypothetical protein(EF0026)	Up
ERT25958	hypothetical protein(EF3018)	Up
ERT25393	hypothetical protein(EF0673)	Up
ERT23625	catalase/peroxidase(katA)	Up
ERT21324	sodium/dicarboxylate symporter family protein(EF0387)	Up
ERT23619	protease synthase and sporulation negative regulatory protein pai 1(EF1590)	Up
ERT21737	ABC transporter ATP-binding protein/permease(EF2226)	Up
ERT21646	hypothetical protein(EF2067)	Up
ERT24606	potassium uptake protein(EF0872)	Up
ERT23105	hypothetical protein(EF1067)	Up
ERT20468	inosine 5'-monophosphate dehydrogenase(guaB)	Up
ERT23134	collagen adhesin protein(EF1099)	Up
ERT23442	hypothetical protein(EF1412)	Up
ERT21647	multidrug resistance protein(EF2068)	Up
ERT23526	V-type ATPase subunit E(EF1495)	Up
ERT23524	V-type ATP synthase subunit I(EF1493)	Up
ERT23527	V-type ATPase subunit C(EF1496)	Up
ERT23528	V-type ATP synthase subunit F(EF1497)	Up
ERT23530	V-type ATP synthase subunit B(EF1499)	Up
ERT23529	V-type ATP synthase subunit A(EF1498)	Up
ERT23333	magnesium-translocating P-type ATPase(EF1304)	Up
ERT24666	hypothetical protein(EF0932)	Up
ERT23531	V-type ATP synthase subunit D(EF1500)	Up
ERT23532	hypothetical protein(EF1501)	Up
ERT20732	iron ABC transporter substrate-binding protein(EF0188)	Up
ERT21738	ABC transporter ATP-binding protein/permease(EF2227)	Up

Table A.7: DEGs in common to all six *E. faecalis* RNA-seq experiments, minus antibiotic overlap

JH2-2 ORF	Annotation	Up- or down-regulated
ERT20954	TetR family transcriptional regulator(EF2594)	Down
ERT25488	hypothetical protein(EF0764)	Down
ERT24628	amino acid ABC transporter amino acid-binding/permease(EF0893)	Down
ERT23533	beta-lactamase(EF1502)	Down
ERT23292	hypothetical protein(EF1263)	Down
ERT24629	glycerol dehydrogenase(EF0895)	Down
ERT21357	Na+/H+ antiporter(napA)	Up
ERT20468	inosine 5'-monophosphate dehydrogenase(guaB)	Up
ERT23529	V-type ATP synthase subunit A(EF1498)	Up
ERT21057	membrane protein(EF2708)	Up
ERT23526	V-type ATPase subunit E(EF1495)	Up
ERT23532	hypothetical protein(EF1501)	Up
ERT25395	glycine betaine/carnitine/choline ABC transporter glycine betaine/carnitine/choline-binding protein(EF0675)	Up
ERT21324	sodium/dicarboxylate symporter family protein(EF0387)	Up
ERT23527	V-type ATPase subunit C(EF1496)	Up
ERT23528	V-type ATP synthase subunit F(EF1497)	Up
ERT23524	V-type ATP synthase subunit I(EF1493)	Up
ERT23523	V-type ATPase subunit F(EF1492)	Up
ERT25393	hypothetical protein(EF0673)	Up
ERT23134	collagen adhesin protein(EF1099)	Up
ERT21738	ABC transporter ATP-binding protein/permease(EF2227)	Up

Table A.8: Read alignment statistics for RNA-seq experiment 1

Sample number	Description	Total number of reads	# aligned reads	# unaligned reads
2203-01	10 min GccF 1	7846794	7806363	40431 (0.52 %)
2203-02	10 min GccF 2	8883154	8847370	35784 (0.40 %)
2203-03	10 min GccF 3	9198754	9169546	29208 (0.32 %)
2203-04	10 min Neg 1	9384034	9315468	68566 (0.73 %)
2203-05	10 min Neg 2	9719694	9652212	67483 (0.69 %)
2203-06	10 min Neg 3	9136884	9076276	60608 (0.66 %)
2203-07	10 min GlcNAc 1	9308701	9266143	42558 (0.45 %)
2203-08	10 min GlcNAc 2	8932805	8885330	47475 (0.53 %)
2203-09	10 min GlcNAc 3	9391305	9343852	47453 (0.51 %)
2203-10	0 min GccF 1	909 5501	9070175	25326 (0.28 %)
2203-11	0 min GccF 2	9141231	9120040	21191 (0.23 %)
2203-12	0 min GccF 3	9273525	9246057	27468 (0.30 %)
2203-13	0 min Neg 1	9266418	9221247	45171 (0.49 %)
2203-14	0 min Neg 2	8811669	8770750	40919 (0.46 %)
2203-15	0 min Neg 3	9173678	9120587	53091 (0.58 %)
2203-16	30 min GccF 1	9726553	9697354	29198 (0.30 %)
2203-17	30 min GccF 2	8971256	8943932	27324 (0.30 %)
2203-18	30 min GccF 3	9542632	9512148	30484 (0.32 %)
2203-19	30 min Neg 1	9240562	9178875	61687 (0.68 %)
2203-20	30 min Neg 2	8956664	8894711	61953 (0.69 %)
2203-21	30 min Neg 3	9513868	9449905	63963 (0.67 %)

Table A.9: Read alignment statistics for RNA-seq experiment 2

Sample number	Description	Total number of reads	# aligned reads	# unaligned reads
2776-01	JH2-2 GccF 1	9946078	9727601	218477 (2.2 %)
2776-02	JH2-2 GccF 2	9329434	8825024	504380 (5.4 %)
2776-03	JH2-2 GccF 3	10339414	9791762	547652 (5.3 %)
2776-04	JH2-2 Neg 1	10061787	9303765	758022 (7.3 %)
2776-05	JH2-2 Neg 2	10680841	10257368	423473 (4.4 %)
2776-06	JH2-2 Neg 3	9999587	9838147	161440 (1.6 %)
2776-07	GlcNAc GccF 1	10060562	10006320	54242 (0.54 %)
2776-08	GlcNAc GccF 2	9800853	9744201	56752 (0.58 %)
2776-09	GlcNAc GccF 3	9692217	9636261	55956 (0.58 %)
2776-10	GlcNAc Neg 1	10270471	10197539	72932 (0.71 %)
2776-11	GlcNAc Neg 2	11280577	11208266	72311 (0.64 %)
2776-12	GlcNAc Neg 3	11104709	11046658	58051 (0.52 %)
2776-13	VI01 GccF 1	9500508	9077390	423118 (4.5 %)
2776-14	VI01 GccF 2	9812131	9251336	560795 (5.7 %)
2776-15	VI01 GccF 3	10258882	9782098	476784 (4.6 %)
2776-16	VI01 Neg 1	9344792	8299828	1044972 (11 %)
2776-17	VI01 Neg 2	11715926	11229728	486198 (4.1 %)
2776-18	VI01 Neg 3	11380700	10946555	434145 (3.8 %)

A.2 Supplementary figures

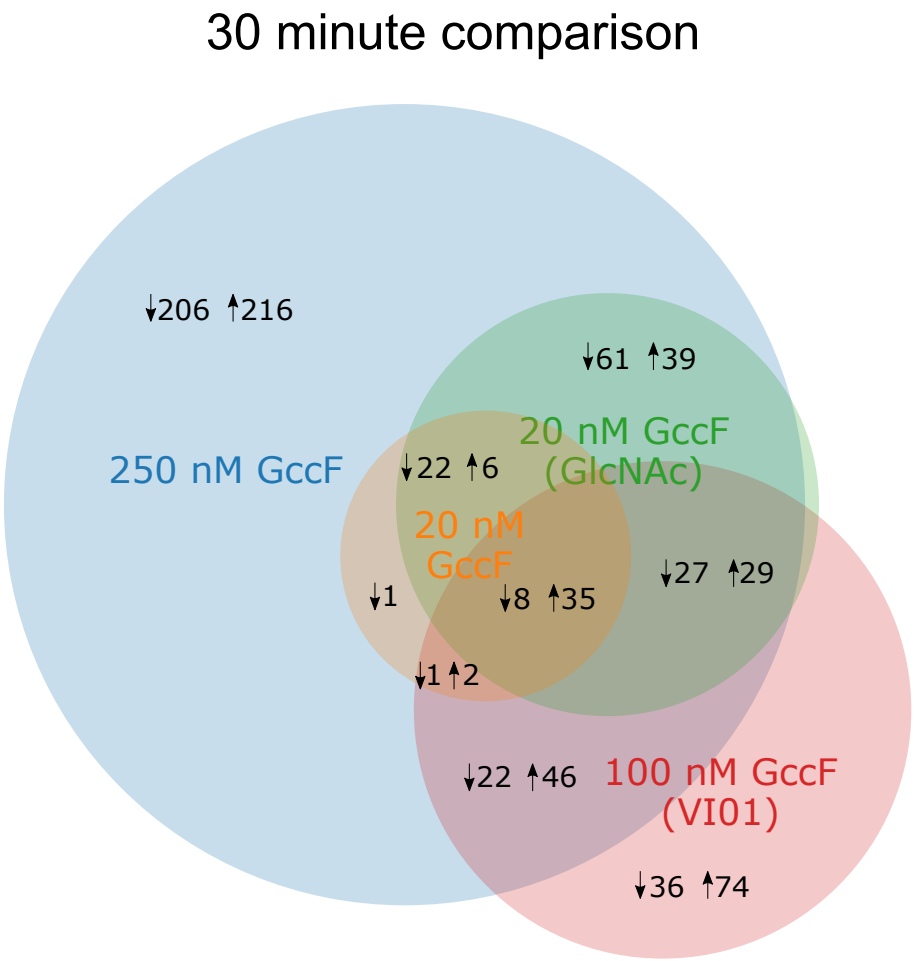
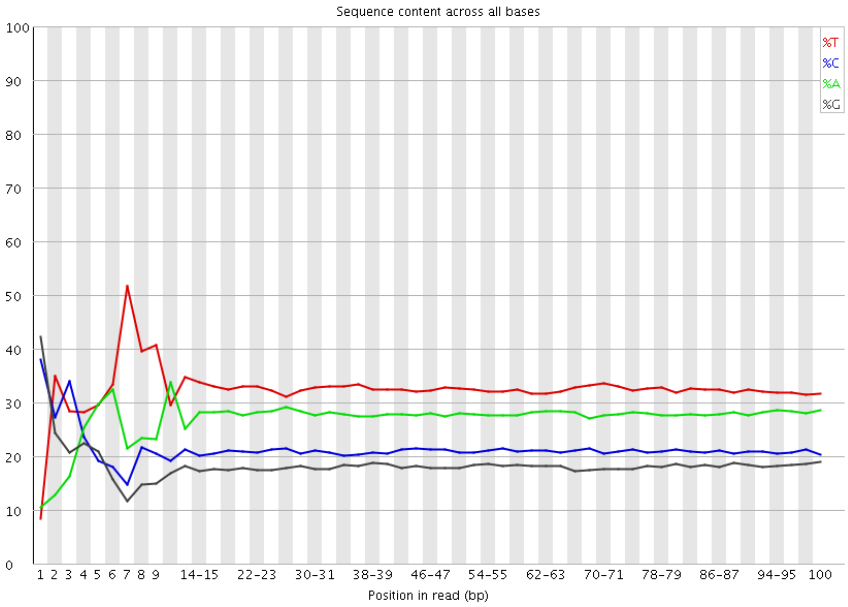
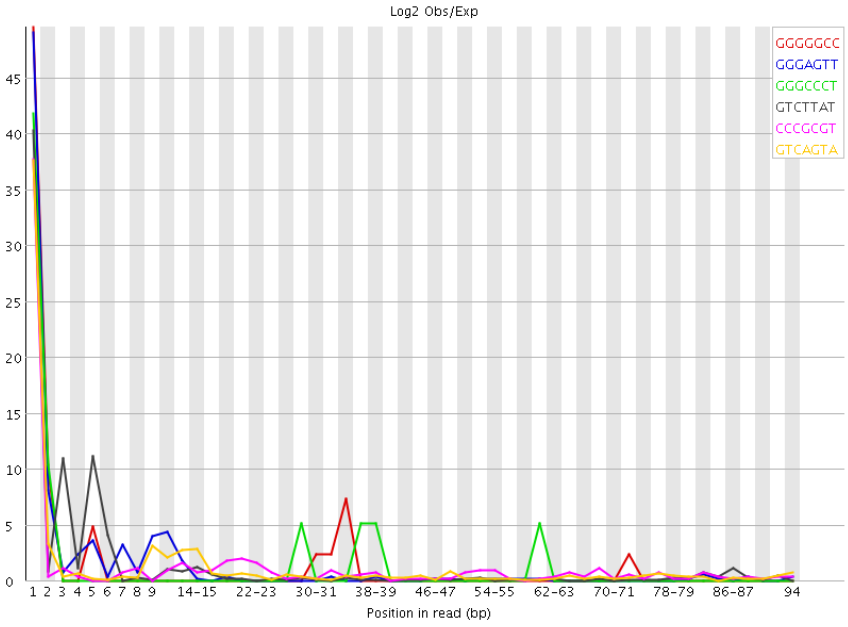


Figure A.1: **Venn diagram of gene expression overlap in 30 minute data sets** The numbers of genes up- and down-regulated in each experiment at $t = 30$ minutes are presented, with areas of overlapping gene expression shown. The up and down arrows refer to numbers of up- and down-regulated genes, respectively. From the figure, it can be seen that the DEGs from the 20 nM and 20 nM GlcNAc pre-treated experiments are all differentially expressed in the 250 nM treatment. However, the sigma 54 knockout strain (VI01, red) has a significant number of DEGs which do not overlap with the other experiments. The list of 43 genes common to all experiments can be found in Supplementary Table A.6



(a) Per base sequence content



(b) Kmer content graph

Figure A.2: **An example of a failed *FastQC* test for RNA-seq experiment 1** Representative plots for 'per base sequence content' (a) and 'Kmer content' (b), as produced by *FastQC* version 0.11.3 for RNA-seq experiment 1. These were the only two tests to be flagged as 'failed' for each read. These specific plots were taken from sample 2203-02.

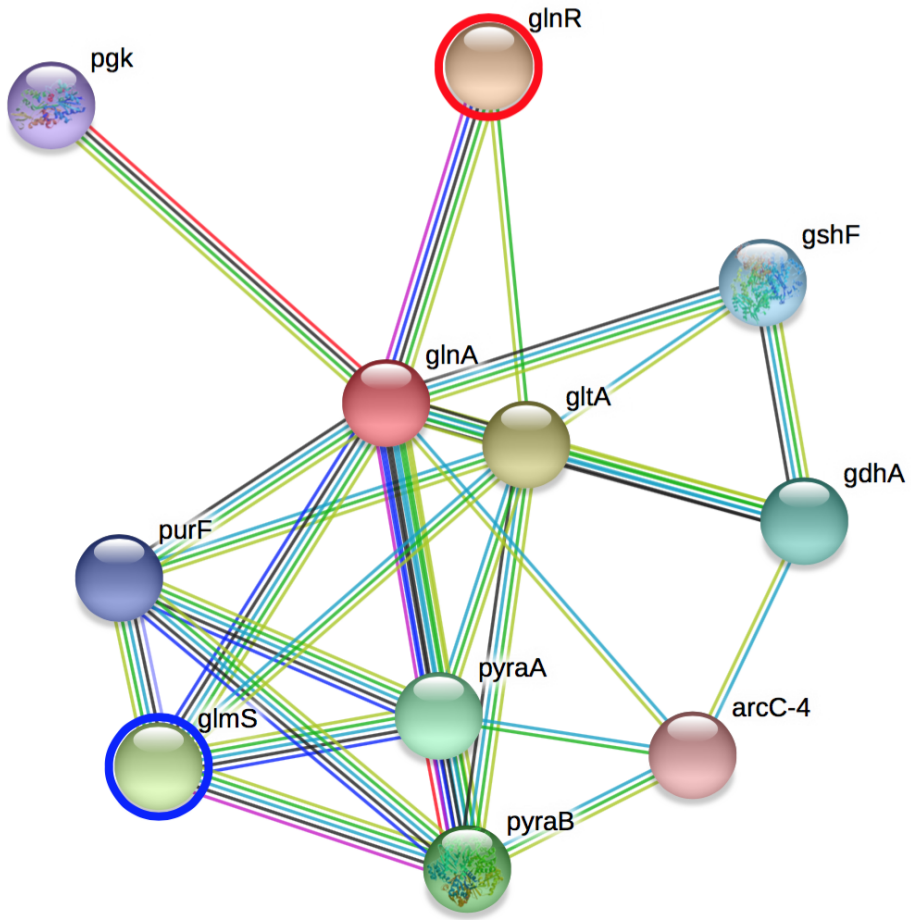


Figure A.3: **Proteins interaction with GlnA**
Protein interaction profile of glutamine synthetase (GlnA) from *Enterococcus faecalis*. Lines show interaction with GlnR (regulatory protein) and GlmS (glutamine/ glucosamine-fructose-6-phosphate aminotransferase). Image generated by https://string-db.org.

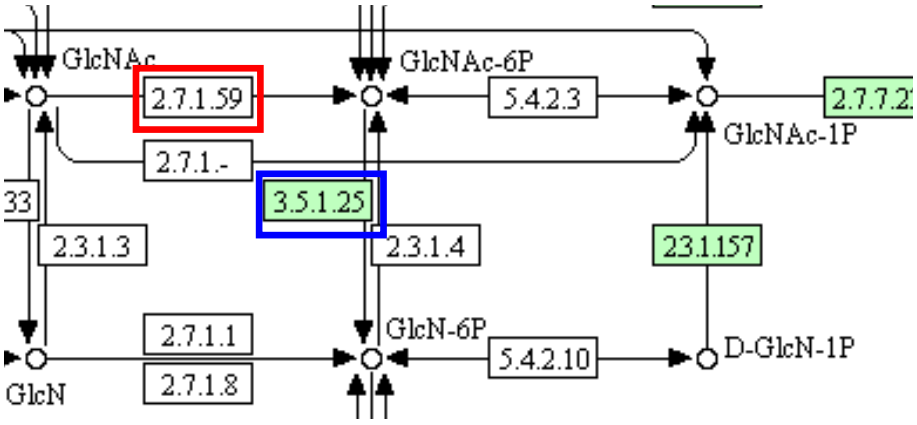
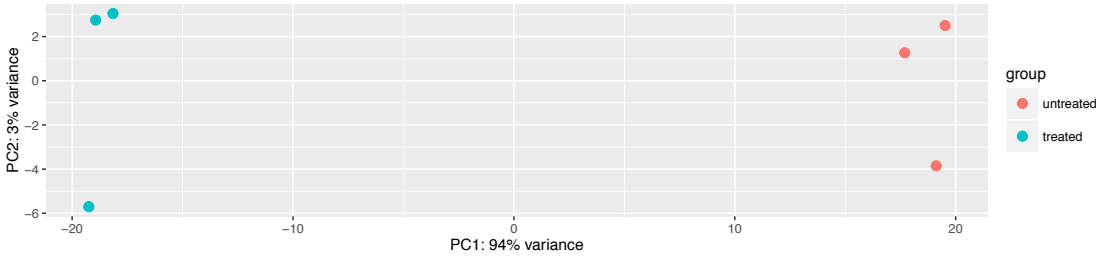
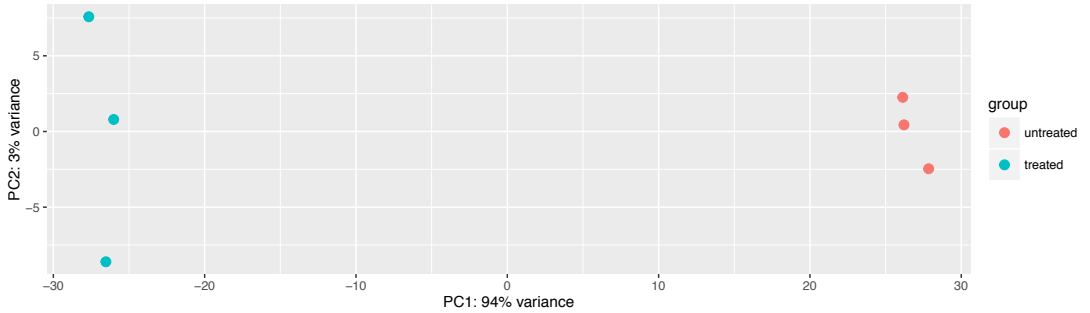


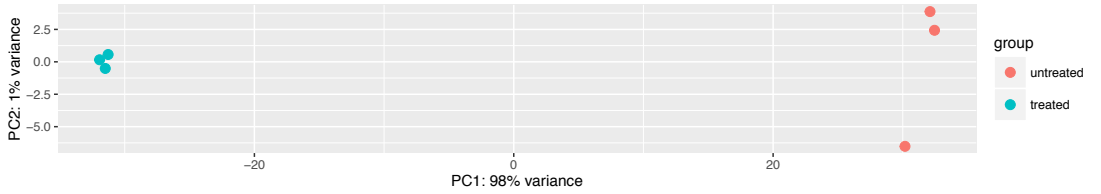
Figure A.4: **GlcNAc-6-P fate KEGG pathway**
GlcNAc is imported into bacterial cells as GlcNAc-6-P by the GlcNAc-specific PTS (red). From here, the main fate for GlcNAc-6-P in *E. faecalis* is deamination by NagA (blue), forming GlcN-6-P. From here, GlcN-6-P is a central precursor to numerous pathways, such as glycolysis or peptidoglycan synthesis.



(a) PCA plot: 0 minute treatment

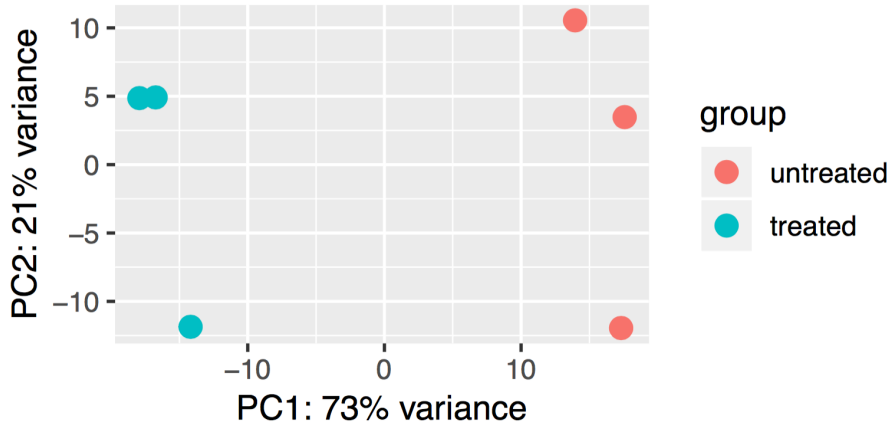


(b) PCA plot: 10 minute treatment

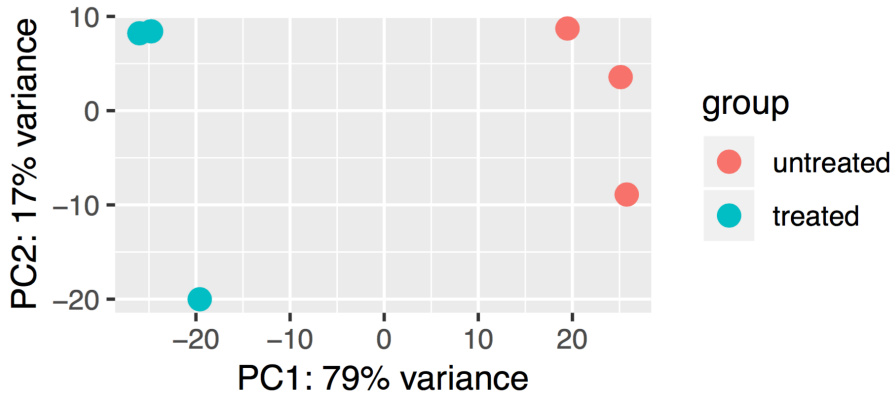


(c) PCA plot: 30 minute treatment

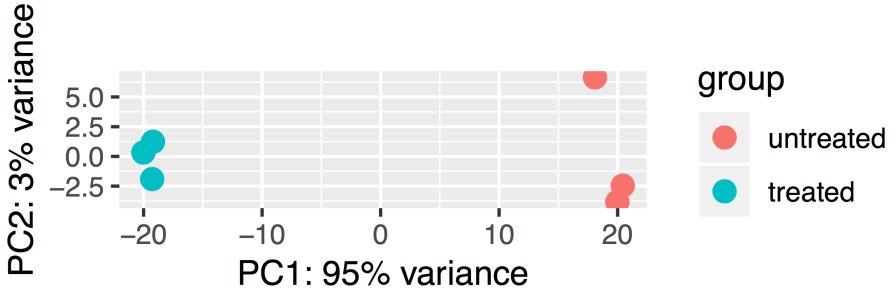
Figure A.5: **PCA plots of three GccF treatment time points**
Principal component analysis was carried out on variance-stabilised count data generated by *DESeq2*. Treatment-specific clustering can be observed for all three GccF treatment time points.



(a) PCA plot: 20 nM GccF



(b) PCA plot: 20 nM GccF, GlcNAc pre-treated



(c) PCA plot: 100 nM GccF, σ^{54} knockout

Figure A.6: **PCA plots of second RNA-seq experiment**
Principal component analysis was carried out on variance-stabilised count data generated by *DESeq2*. Treatment-specific clustering can be observed in all three samples.

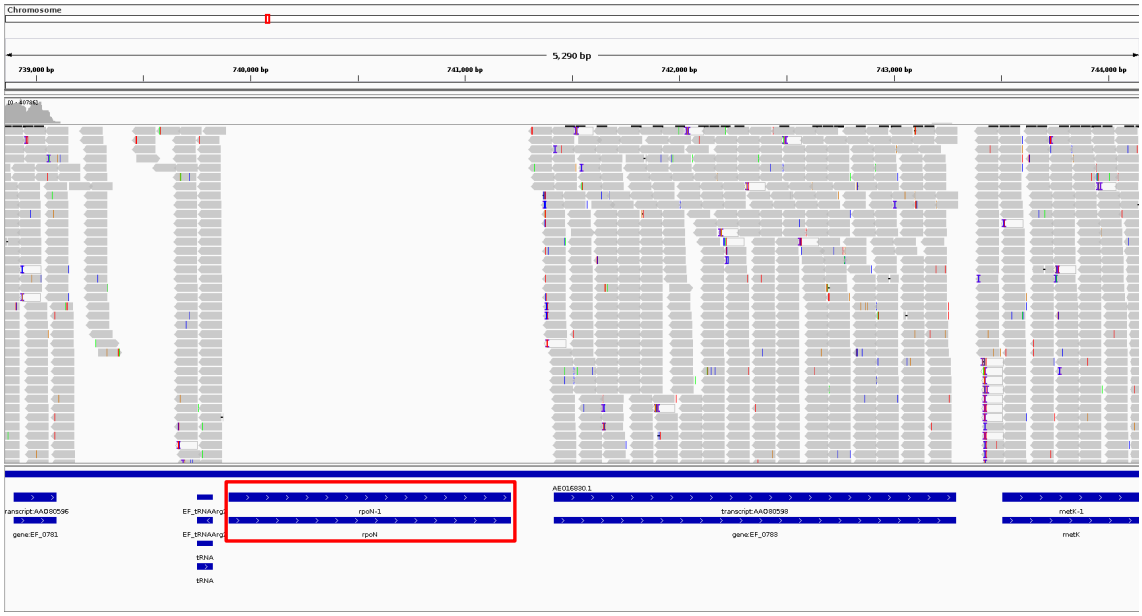


Figure A.7: **Confirmation of sigma 54 knockout**
A set of RNA-seq reads taken from *E. faecalis* VI01 were aligned to reference strain V583, and the mapping of the reads visualised using Integrated Genome Viewer version 2.3 (Broad Institute; software.broadinstitute.org/software/igv). The gene encoding sigma factor 54 (*rpoN*) is highlighted in red. The lack of any reads aligning to the region indicates that the strain is a successful knockout.

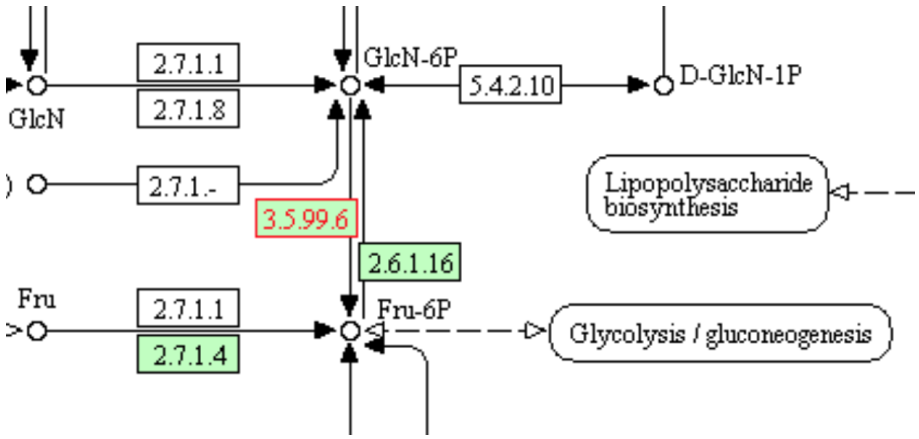


Figure A.8: **KEGG pathway of fructose-6-P**
NagB (glucosamine-6-phosphate isomerase) is highlighted in red. From the KEGG pathway, it can be seen that the fate of Fru-6-P following deamination of GlcN-6-P, other than being re-aminated in a futile cycle, is to enter the glycolytic pathway.

B | Supplementary figures

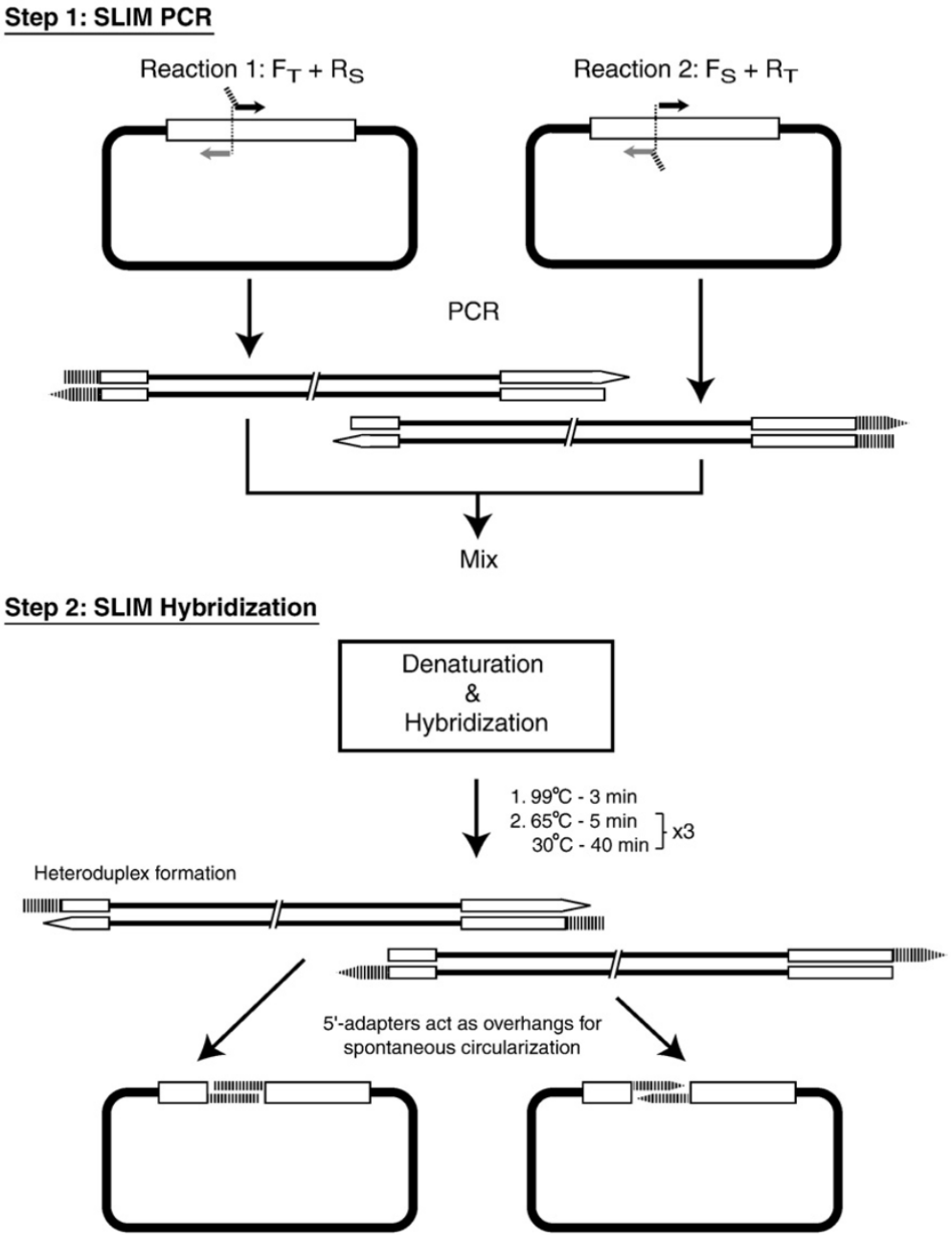
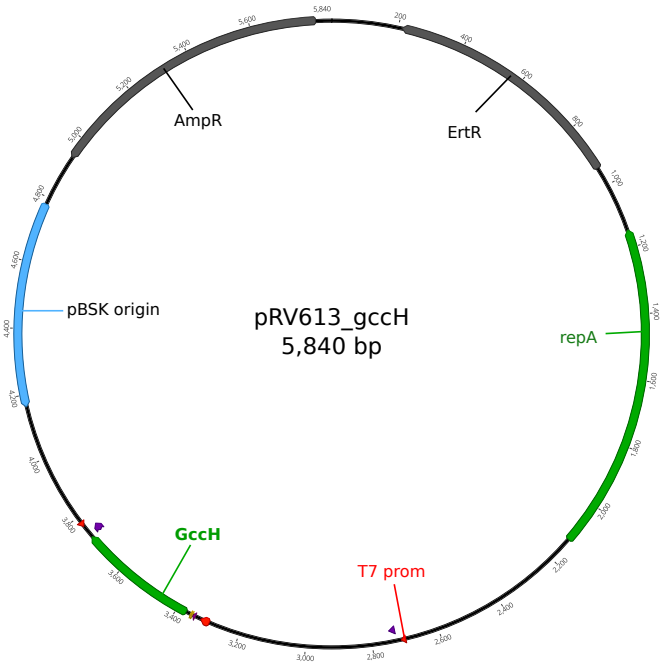
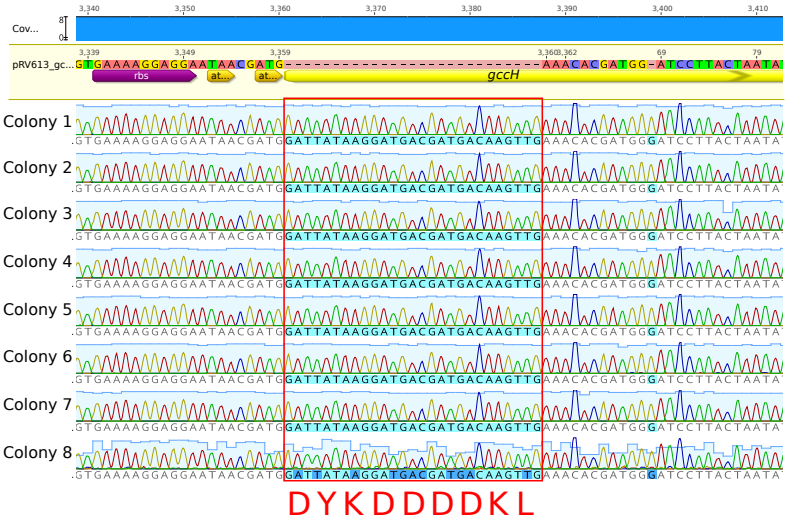


Figure B.1: **Outline of SLIM concept**
Two PCR amplifications are carried out, with one of each primer pair having a 5´ tail containing the modification of interest, resulting in two linear amplicons with modified DNA at opposing ends. These two amplicons are mixed together, subjected to heating and cooling steps to anneal the fragments together into a circular form with single-stranded nicks, before being transformed into *E. coli* to repair the nicks and propogate the plasmid. Figure taken from Chiu *et al.* (2008)[40], with permission from Elsevier.



(a) Plasmid map of pRV613_gccH



(b) Sequence alignment of flag-tagged pRV613_gccH mutants

Figure B.2: **pRV613_*gccH* plasmid map and flag-tag sequence alignment**
(a) Plasmid map of pRV613 modified to include *gccH* after the copper-inducible promoter. (b) Sequence alignment of the start of the *gccH* gene, showing sequence encoding the flag-tag.

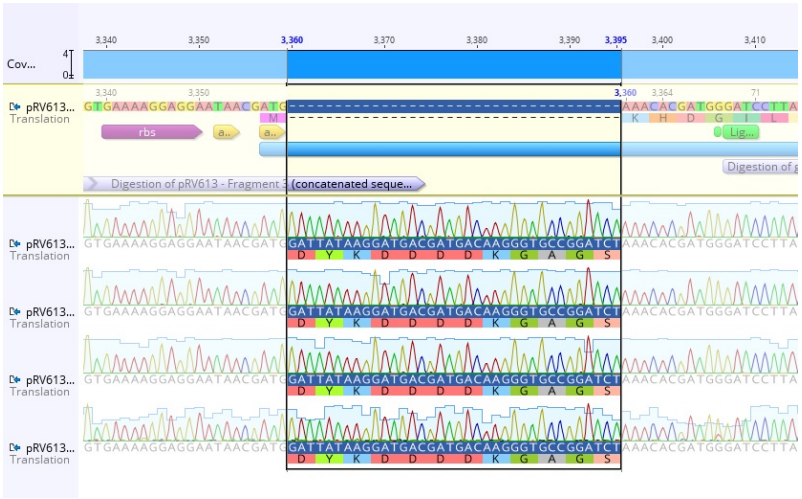
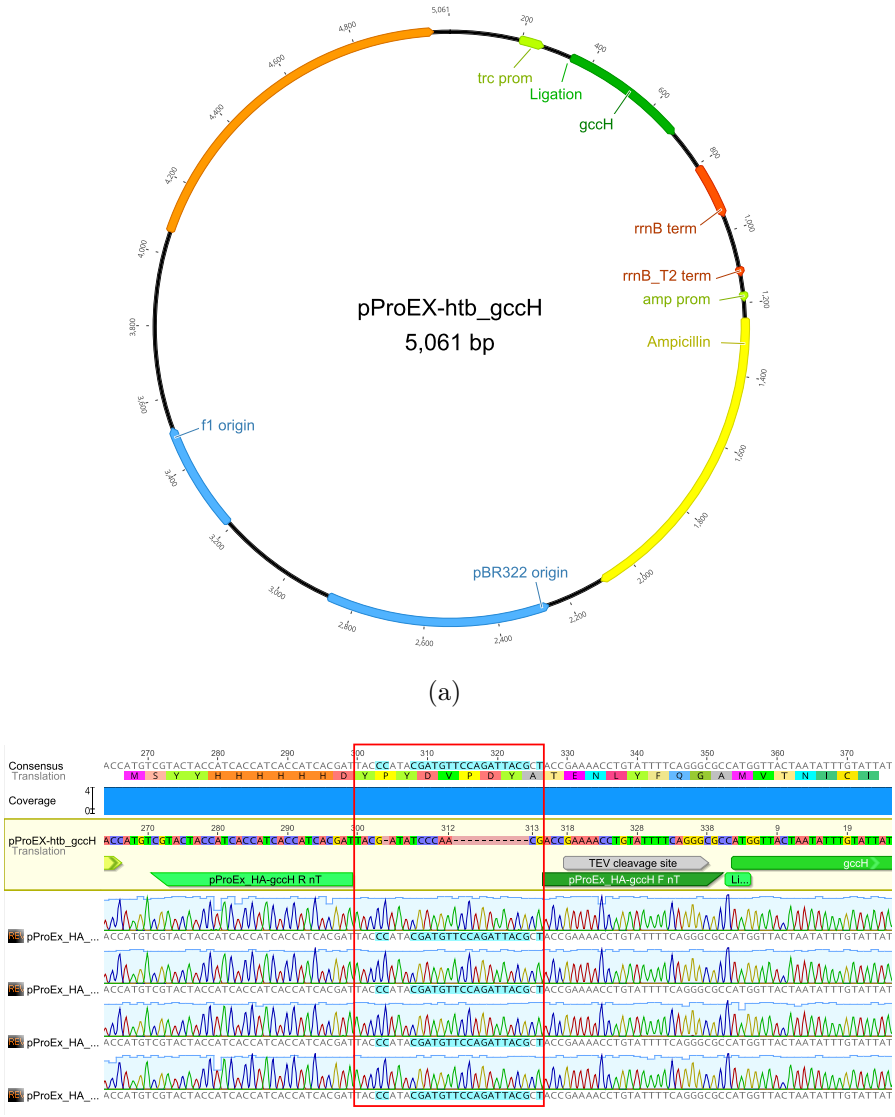


Figure B.3: **Sequence alignment of pRV613_fL-*gccH***
Sequence alignment of pRV613_flag_*gccH* modified to include GAGS linker after flag tag. Aligned against plasmid map of pRV613_*gccH*.

Figure B.4: **Sequence alignment of pRV613_HA-*gccH***
Sequence alignment against pRV613_*gccH*, showing correct addition of sequence encoding HA tag at start of *gccH* gene.



(a)

(b)

Figure B.5: **Plasmid map of pProExHTb_*gccH* and sequence alignment of *HA-gccH***
(a) Plasmid map of pProEx-htb_*gccH*. (b) Sequence alignment showing insertion of HA tag downstream of His_6 tag in pProEx-htb_*gccH*.

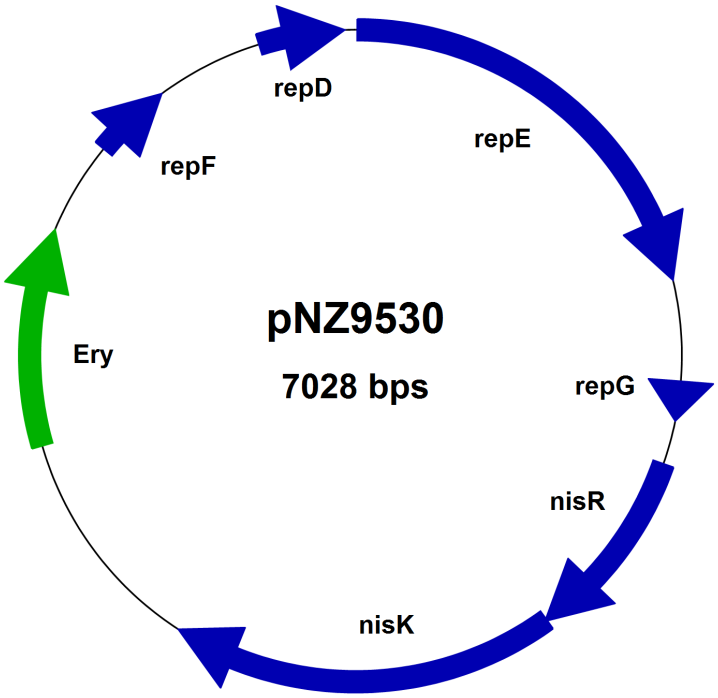


Figure B.6: **Plasmid map of pNZ9530**
Plasmid map of pNZ9530. Genes of interest are *nisK* and *nisR*, which encode nisin-detecting histidine kinase and response-regulator, respectively. *Ery* encode erythromycin resistance marker.

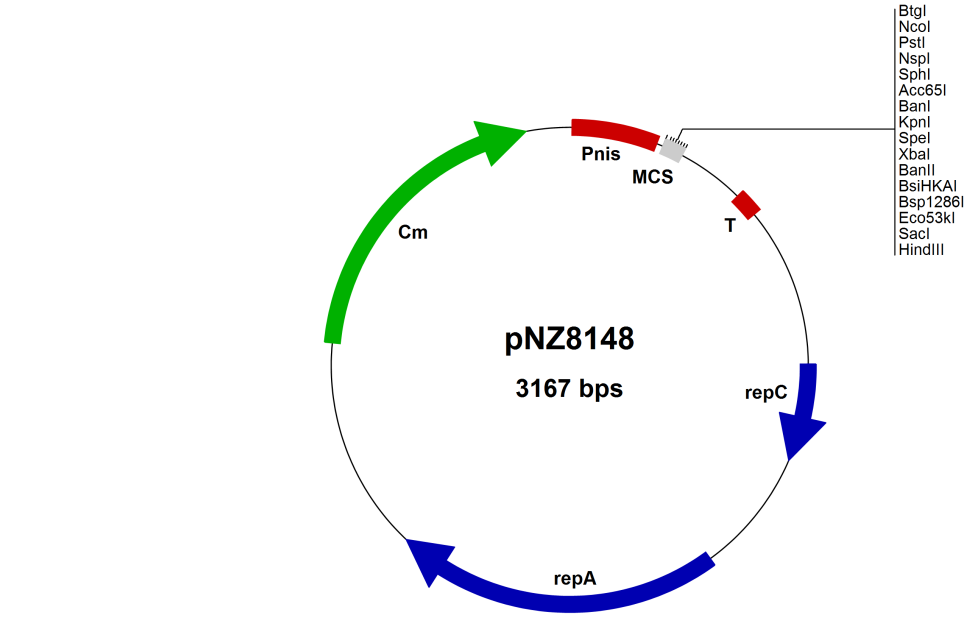


Figure B.7: **Plasmid map of pNZ8148**
Plasmid map of pNZ8148. *Cm* refers to chloramphenicol resistance marker.

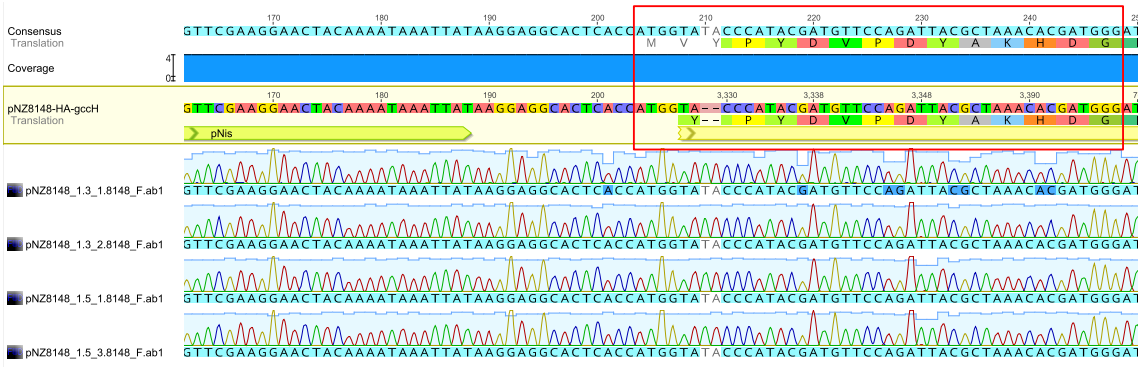


Figure B.8: **Sequence alignment of pNZ8148_HA-*gccH***
Forward sequencing reads four plasmids isolated from *E. coli* XL-1 Blue cells transformed with pNZ8148 + *HA-gccH* ligations were aligned against a map of pNZ8148_HA-*gccH*. All four reads can be shown to align with this coding sequence, indicating insertion of the gene into the plasmid (start of coding sequence of *HA-gccH* highlighted by a red box). Note that a Val was inserted between Met1 and the first Tyr of the HA-tag.

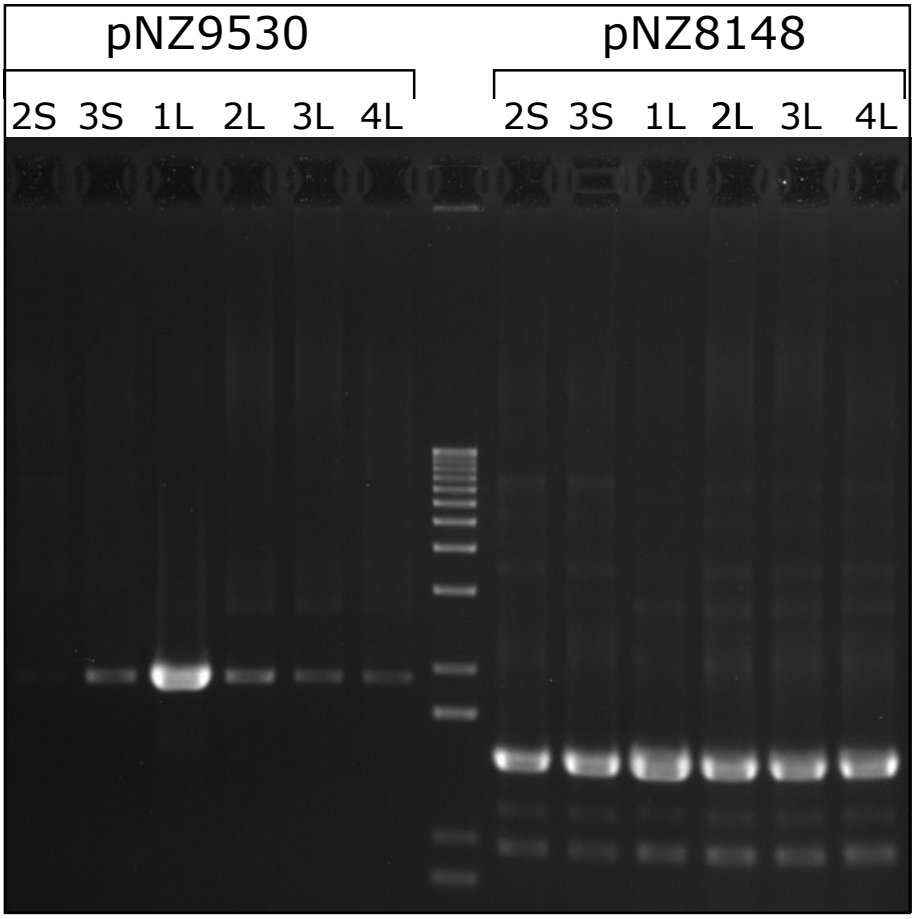


Figure B.9: **Dual-transformation confirmation gel**
DNA extracts of colonies of *Lb. plantarum* NC8 cells transformed with pNZ9530 and pNZ8148_HA-*gccH* were amplified with primer pairs designed to amplify regions of each plasmid. All six colonies were confirmed to contain plasmid pNZ8148_HA-*gccH*, and all colonies aside from 2S were confirmed to contain plasmid pNZ9530.

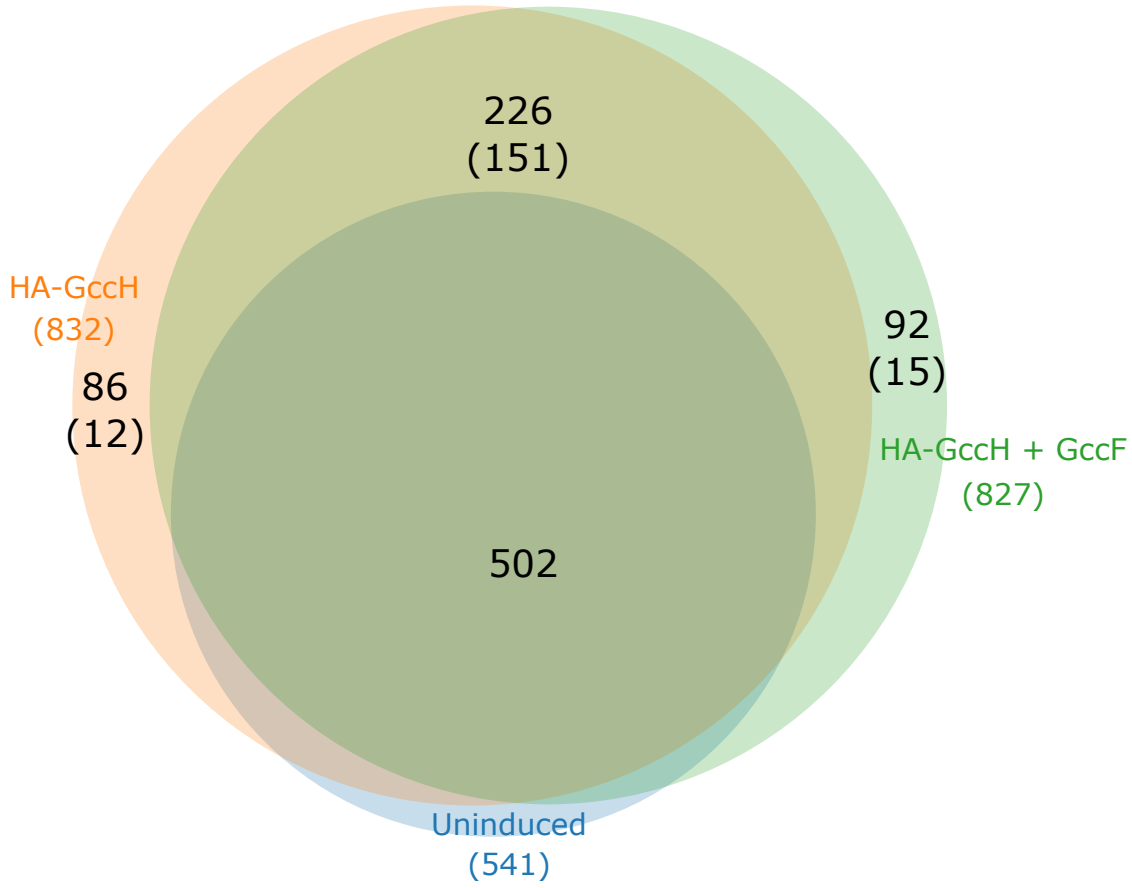


Figure B.10: **Venn diagram of HA-GccH pulldown protein matches**
The proteins immunoprecipitated from the nisin-induced expression system were identified by mass spectrometry, using the *Lb. plantarum* NCBI database. The total number of protein IDs for each sample are shown, prior to filtering for ≥ 1 unique peptides. Numbers in black parentheses show the numbers of peptides remaining after filtering for 2 or more unique peptides. The specific numbers of overlapped protein IDs are also shown.

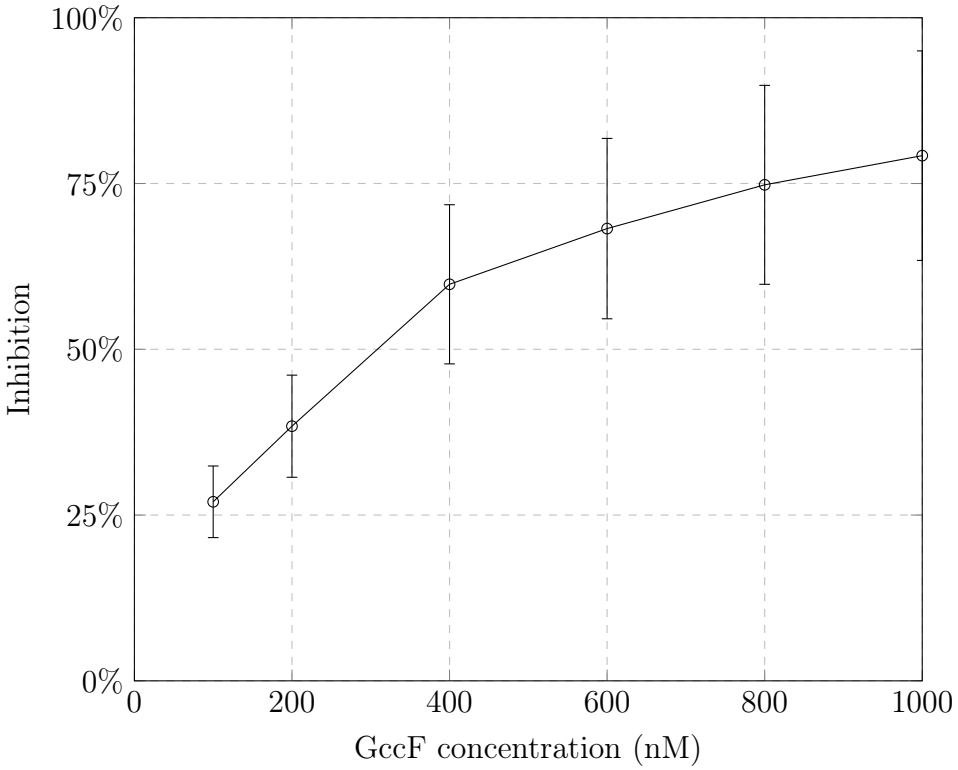


Figure B.11: **Inhibition plot for GccF**$_{Cys43GlcNDAz}$
Inhibition values of different concentrations of GccF$_{Cys43GlcNDAz}$ tested on *Lb. plantarum* ATCC 8014 are shown. 50 % inhibiton occurs at around 294 nM.

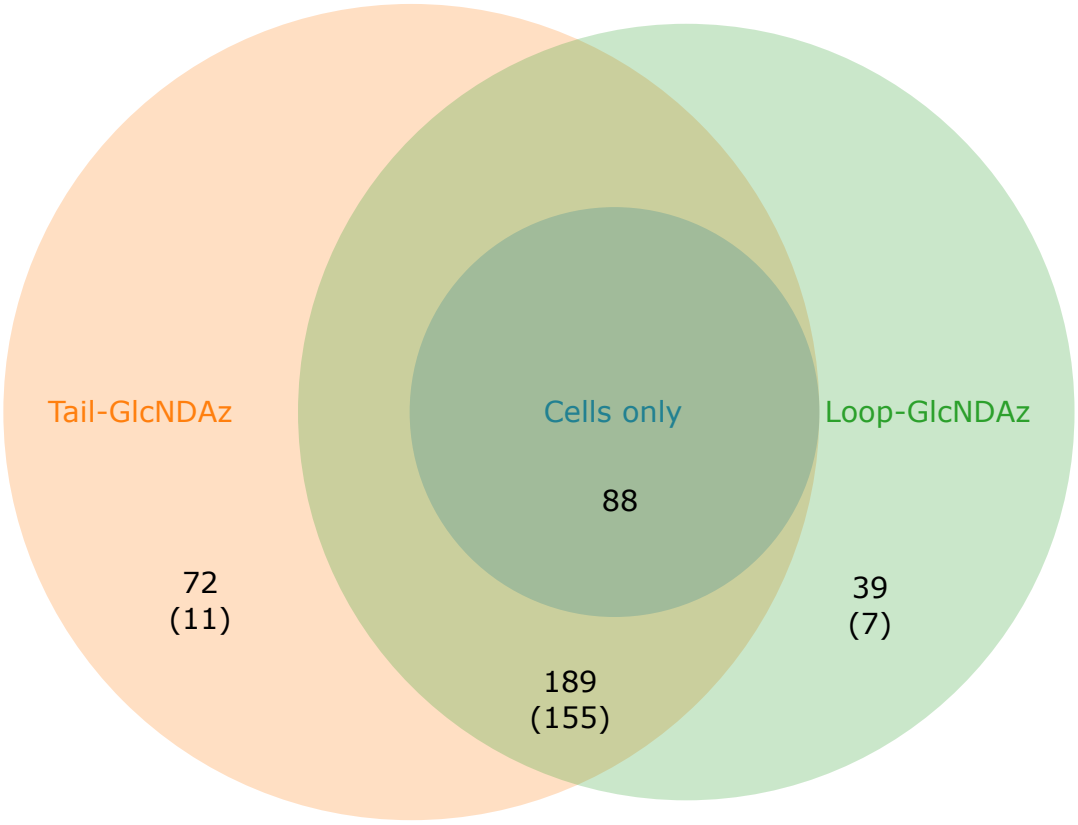


Figure B.12: **Venn diagram of GccF-GlcNDAz pulldown overlap**
Mass spectrometry analysis of the photo-crosslinked GccF pulldowns resulted in 316 unique peptide matches for cells treated with GccF$_{Cys18GlcNDAz}$, 349 unique matches for cells treated with GccF$_{Cys43GlcNDAz}$, and 88 unique matches for untreated cells. All 88 protein matches for the untreated cells were found in both photo-crosslinked samples, while 189 proteins overlapped from the GccF$_{Cys18GlcNDAz}$ and GccF$_{Cys43GlcNDAz}$ pulldowns. Numbers in parentheses indicate the numbers of peptides remaining after filtering for less than two unique peptides.

C | Supplementary tables

Table C.1: Protein matches common to 'HA-GccH' and 'HA-GccH + GccF' pull-downs

Accession Number	Unique Peptides	Coverage [%]	Description
WP_003640378.1	11	13	chromosome segregation protein SMC
WP_003641994.1	10	22	ATP-dependent DNA helicase
WP_003646514.1	8	26	transcription termination factor Rho
WP_003644442.1	7	14	ATP-dependent chaperone ClpB
WP_046039231.1	6	42	FAD-binding oxidoreductase
WP_003644631.1	6	27	glutamate racemase
WP_021356532.1	6	31	glycosyltransferase family 4 protein
WP_063206772.1	6	49	thymidine kinase
WP_013355564.1	6	32	thymidylate synthase
WP_003642622.1	6	22	glucose-6-phosphate dehydrogenase
WP_003642037.1	5	38	sugar-phosphatase
WP_003644320.1	5	39	ribosome small subunit-dependent GTPase A
WP_003644265.1	5	28	aminotransferase
WP_003644362.1	5	11	diguanylate cyclase
WP_003642052.1	5	28	hypothetical protein
WP_003637815.1	**5**	**10**	**APC family permease**
WP_003637294.1	4	39	response regulator transcription factor
WP_003638623.1	4	59	Asp23/Gls24 family envelope stress response protein
WP_027821856.1	**4**	**12**	**cell wall hydrolase**[†]
YP_004889713.1	4	36	oxidoreductase, medium chain dehydrogenases/reductase (MDR)/zinc-dependent alcohol dehydrogenase-like family
WP_015380700.1	4	22	lipoate–protein ligase
WP_003644313.1	4	28	aminopeptidase P family protein
WP_003643523.1	4	49	O-acetyl-ADP-ribose deacetylase

Continued on next page

Table C.1 Common protein hits – continued from previous page

Accession Number	Unique Peptides	Coverage [%]	Description
WP_003643583.1	4	13	catalase
WP_015380329.1	4	29	mevalonate kinase
WP_003642470.1	4	16	hypothetical protein
WP_003644074.1	4	30	histidine phosphatase family protein
WP_058011626.1	4	24	alpha/beta hydrolase
WP_003637770.1	4	40	transcription termination/antitermination protein NusG
WP_003642696.1	4	10	LTA synthase family protein
WP_013355364.1	4	18	MurR/RpiR family transcriptional regulator
WP_003641053.1	4	35	SsrA-binding protein SmpB
WP_047672593.1	**4**	**4**	**transporter**
WP_003640715.1	4	22	riboflavin biosynthesis protein RibF
WP_033610823.1	**3**	**40**	**D-alanyl-D-alanine carboxypeptidase**†
HA-GccH	**3**	**42**	**HA-tagged GccF immunity protein**
WP_064523202.1	3	39	SorC family transcriptional regulator
WP_064306082.1	3	29	2-deoxyuridine 5-triphosphate nucleotidohydrolase
WP_003645162.1	3	28	bifunctional methylenetetrahydrofolate dehydrogenase/methenyltetrahydrofolate cyclohydrolase
WP_003641476.1	3	54	GAF domain-containing protein
WP_003644231.1	3	23	16S rRNA pseudouridine(516) synthase
WP_003643445.1	3	19	16S rRNA (guanine(527)-N(7))-methyltransferase RsmG
WP_003644585.1	3	37	16S rRNA (guanine(966)-N(2))-methyltransferase RsmD
WP_011101248.1	3	13	FAD:protein FMN transferase
WP_016527128.1	**3**	**13**	**ABC transporter substrate-binding protein/permease**
WP_003644073.1	3	34	GNAT family N-acetyltransferase
WP_003644314.1	3	22	polyprenyl synthetase family protein
WP_003640353.1	3	15	TlyA family rRNA (cytidine-2'-O)-methyltransferase
WP_003642516.1	3	34	2-amino-4-hydroxy-6-hydroxymethyldihydropteridine diphosphokinase
WP_063206009.1	**3**	**30**	**glutamine amidotransferase**†
WP_015380398.1	3	19	peptidase T
WP_003642736.1	**3**	**11**	**PTS N-acetylglucosamine transporter subunit IIABC (PTS18CBA)**
WP_003643958.1	3	17	NADH oxidase
WP_027822501.1	3	17	histidine phosphatase family protein

Continued on next page

Table C.1 Common protein hits – continued from previous page

Accession Number	Unique Peptides	Coverage [%]	Description
WP_003644254.1	3	28	MarR family transcriptional regulator
WP_003644499.1	3	23	tRNA1(Val) (adenine(37)-N6)-methyltransferase
WP_003640460.1	3	12	23S rRNA methyltransferase
WP_003644792.1	3	16	tRNA-dihydrouridine synthase family protein
WP_003641832.1	3	9	DUF2252 domain-containing protein
WP_003643518.1	3	15	alpha/beta hydrolase
WP_003644004.1	3	17	proline-specific peptidase family protein
WP_103851679.1	3	19	polysaccharide deacetylase family protein
WP_033620034.1	3	19	Cys-tRNA(Pro) deacylase
WP_044431095.1	3	11	alpha-ketoacid dehydrogenase subunit beta
WP_003643674.1	3	23	SDR family oxidoreductase
WP_057717686.1	3	46	universal stress protein
WP_003640605.1	3	5	DNA polymerase III subunit alpha
WP_003642169.1	3	22	NADP oxidoreductase
WP_003641944.1	3	11	type 1 glycerol-3-phosphate oxidase
WP_015379741.1	3	14	ABC transporter ATP-binding protein
WP_003644709.1	3	27	hypothetical protein
WP_003642988.1	3	8	bifunctional metallophosphatase/5'-nucleotidase
WP_003641226.1	3	24	metal-dependent transcriptional regulator
WP_003642484.1	3	7	phytoene desaturase
WP_011101142.1	3	19	ribosomal-protein-alanine N-acetyltransferase
WP_003640897.1	3	23	GntR family transcriptional regulator
WP_003639204.1	**2**	**54**	**D-alanine–D-alanine ligase**[†]
WP_076642757.1	2	76	ACP S-malonyltransferase
WP_003643690.1	2	60	oxidoreductase
WP_003638977.1	2	32	transcriptional repressor LexA
WP_003640698.1	2	22	16S rRNA (uracil(1498)-N(3))-methyltransferase
WP_003641119.1	2	21	HAD family hydrolase
WP_003643875.1	2	18	hypothetical protein
WP_013355799.1	2	9	M3 family oligoendopeptidase
WP_003644645.1	2	12	cysteine desulfurase
WP_054399000.1	2	12	cell surface protein
WP_003643844.1	**2**	**9**	**UDP-N-acetylmuramoyl-tripeptide–D-alanyl-D-alanine ligase**[†]
WP_003641701.1	2	24	pyridoxamine 5'-phosphate oxidase
WP_003638600.1	2	12	exodeoxyribonuclease VII large subunit
WP_003640582.1	2	12	CCA tRNA nucleotidyltransferase
WP_003643581.1	2	34	hypothetical protein

Continued on next page

Table C.1 Common protein hits – continued from previous page

Accession Number	Unique Peptides	Coverage [%]	Description
WP_003643702.1	2	12	aminotransferase class V-fold PLP-dependent enzyme
WP_021337058.1	2	4	sigma-54-dependent transcriptional regulator
WP_003641485.1	2	8	bifunctional folylpolyglutamate synthase/dihydrofolate synthase
WP_003642298.1	2	11	SDR family oxidoreductase
WP_064578370.1	2	21	16S rRNA (cytidine(1402)-2'-O)-methyltransferase
WP_063851703.1	2	13	deoxyribonuclease IV
WP_003640423.1	2	23	acyl-CoA thioesterase
WP_003638559.1	2	28	acylphosphatase
WP_003645944.1	2	19	methyltransferase domain-containing protein
WP_003641566.1	2	29	hypothetical protein
WP_003640479.1	2	23	universal stress protein
WP_054398655.1	2	14	peptide chain release factor 1
WP_003638751.1	**2**	**37**	**cell division regulator GpsB**[†]
WP_015381013.1	2	10	GMP reductase
WP_003640918.1	2	13	acetyl-CoA carboxylase carboxyltransferase subunit beta
WP_015380932.1	2	27	TetR/AcrR family transcriptional regulator
WP_003642609.1	2	11	dihydroorotate dehydrogenase
WP_003642605.1	2	8	aspartate carbamoyltransferase catalytic subunit
WP_097037156.1	2	6	ABC-F family ATP-binding cassette domain-containing protein
WP_027821813.1	2	41	hypothetical protein
WP_033615579.1	2	17	TetR family transcriptional regulator
WP_003642021.1	2	17	4-(cytidine 5'-diphospho)-2-C-methyl-D-erythritol kinase
WP_041153477.1	2	13	signal peptidase I
WP_003640917.1	2	8	acetyl-CoA carboxylase biotin carboxylase subunit
WP_003642953.1	2	27	hypothetical protein
WP_003641941.1	2	14	ABC transporter ATP-binding protein
WP_060683972.1	2	10	ArgE/DapE family deacylase
WP_134901445.1	2	12	Fe-S cluster assembly ATPase SufC
WP_063491142.1	2	14	50S ribosomal protein L9
WP_003644791.1	2	10	FAD:protein FMN transferase
WP_003646316.1	2	17	hypothetical protein
WP_033610066.1	2	12	glycosyltransferase family 2 protein
WP_003646174.1	2	23	23S rRNA (pseudouridine(1915)-N(3))-methyltransferase RlmH

Continued on next page

Table C.1 Common protein hits – continued from previous page

Accession Number	Unique Peptides	Coverage [%]	Description
WP_003643863.1	**2**	**11**	**N-acetylglucosamine-6-phosphate deacetylase**
WP_003639045.1	2	10	pyruvate dehydrogenase (acetyl-transferring) E1 component subunit alpha
WP_003639329.1	2	11	aspartate-semialdehyde dehydrogenase
WP_003640825.1	2	17	peptide deformylase
WP_003642559.1	2	9	Cof-type HAD-IIB family hydrolase
WP_003643010.1	2	30	hypothetical protein
WP_024971625.1	2	9	Cof-type HAD-IIB family hydrolase
WP_003642145.1	2	14	LysR family transcriptional regulator
WP_003640264.1	2	8	Fe-S cluster assembly protein SufD
WP_003639405.1	2	8	HlyC/CorC family transporter
WP_061871762.1	2	12	HslU–HslV peptidase proteolytic subunit
WP_003641919.1	2	9	hypothetical protein
WP_003642038.1	2	9	HD domain-containing protein
WP_061871511.1	2	7	Stk1 family PASTA domain-containing Ser/Thr kinase
WP_003642139.1	2	27	hypothetical protein
WP_003638945.1	2	11	heat-inducible transcriptional repressor HrcA
WP_003640622.1	2	12	zinc-binding dehydrogenase
WP_003639099.1	2	21	GTP pyrophosphokinase family protein
WP_003645663.1	2	16	ComE operon protein 2
WP_003638050.1	2	18	CtsR family transcriptional regulator
WP_003637950.1	**2**	**14**	**amino acid ABC transporter permease**
WP_134901816.1	2	9	copper oxidase
WP_003642560.1	2	5	threonine synthase
WP_064619075.1	2	6	excinuclease ABC subunit UvrA
WP_003644361.1	2	10	diphosphomevalonate decarboxylase
WP_021337855.1	2	27	hypothetical protein
WP_003638657.1	2	15	acyl carrier protein

Note: All proteins taken from the NCBI *Lb. plantarum* database

Proteins of interest are highlighted in boldface

[†]Proteins inferred to be involved in peptidoglycan synthesis/ cell division

Table C.2: Protein matches common to GccF$_{Cys18GlcNDAz}$ and GccF$_{Cys43GlcNDAz}$ pulldowns

Accession Number	Unique Peptides	Coverage [%]	Description
WP_003641250.1	15	34	elongation factor G
EHS83294.1	14	21	excinuclease ABC, subunit A
WP_003644418.1	13	52	LysR family transcriptional regulator
WP_003644725.1	11	39	bifunctional phosphoribosylaminoimidazolecarboxamide formyltransferase/IMP cyclohydrolase
EHS83476.1	11	35	acetolactate synthase
WP_003643815.1	11	60	tryptophan–tRNA ligase
EHS83039.1	11	26	methionine-tRNA synthetase
WP_003640932.1	10	27	glutamate–tRNA ligase
WP_003641542.1	10	30	catabolite control protein A
EHS84181.1	9	53	phosphoglycerate mutase family protein
WP_003641045.1	9	51	phosphoglycerate kinase
EHS83833.1	9	29	linoleic acid isomerase
WP_003640556.1	9	20	DNA topoisomerase IV subunit A
WP_003641904.1	8	39	class II fructose-1,6-bisphosphate aldolase
EHS82886.1	8	22	leucine-tRNA synthetase
WP_003643128.1	8	28	arginine–tRNA ligase
WP_003644337.1	8	39	hypothetical protein
WP_003641436.1	8	31	serine hydroxymethyltransferase
WP_003639227.1	7	56	uracil phosphoribosyltransferase
WP_003644650.1	7	68	universal stress protein
WP_003641632.1	7	22	DNA topoisomerase (ATP-hydrolyzing) subunit B
EHS82612.1	7	21	cell division protein FtsA
WP_003640986.1	7	24	chaperonin GroEL
EHS83062.1	7	19	CTP synthase
EHS82065.1	7	20	glutamate-ammonia ligase
EHS83582.1	7	28	aspartyl/glutamyl-tRNA amidotransferase subunit A
EHS82433.1	7	15	translation initiation factor IF-2
WP_003643227.1	7	16	peptide chain release factor 3
CAJ75872.1	7	13	PyrAb1 protein
WP_003641520.1	7	21	DEAD/DEAH box helicase
EHS82097.1	7	18	16S rRNA m(5)C 967 methyltransferase
WP_003641056.1	6	50	amino acid ABC transporter ATP-binding protein
EHS83315.1	6	15	exoribonuclease II
EHS84251.1	6	32	DegV family protein

Continued on next page

Table C.2 Common protein hits – continued from previous page

Accession Number	Unique Peptides	Coverage [%]	Protein description
WP_003643438.1	6	18	class I SAM-dependent methyltransferase
EHS82445.1	6	36	translation elongation factor Ts
WP_003642090.1	6	16	lysine–tRNA ligase
EHS83517.1	6	41	DNA-directed RNA polymerase, alpha subunit
EHS84261.1	6	16	adenylosuccinate synthase
WP_003644496.1	6	23	transcription termination/antitermination protein NusA
EHS82722.1	6	39	30S ribosomal protein S4
EHS83680.1	6	25	oligopeptide ABC transporter, ATP-binding protein
WP_003641343.1	6	28	Asp-tRNA(Asn)/Glu-tRNA(Gln) amidotransferase subunit GatB
EHS82875.1	6	13	phenylalanyl-tRNA synthetase, beta subunit
EHS82555.1	6	17	metallo-beta-lactamase superfamily protein
EHS83488.1	6	29	30S ribosomal protein S7
EHS83249.1	6	40	redox-sensitive transcription regulator
EHS83500.1	5	39	30S ribosomal protein S17
WP_003640381.1	5	20	signal recognition particle protein
WP_003644566.1	5	20	GTPase ObgE
WP_003640691.1	5	27	histidine–tRNA ligase
EHS83309.1	5	24	triosephosphate isomerase
EHS83471.1	5	26	transcriptional attenuator, cell envelope-related, LytR family
WP_003643733.1	5	17	NAD(P)/FAD-dependent oxidoreductase
WP_003642093.1	5	21	tyrosine–tRNA ligase
EHS83394.1	5	18	GTPase, HflX subfamily
WP_003640679.1	5	29	GTPase Era
WP_003642640.1	5	11	phosphoketolase family protein
EHS83515.1	4	34	30S ribosomal protein S13
WP_003642703.1	4	28	aspartate-semialdehyde dehydrogenase
EHS81778.1	4	52	pyrimidine operon regulator
WP_003642577.1	4	32	ABC transporter ATP-binding protein
EHS82750.1	4	24	cell shape determining protein MreB
WP_003638594.1	4	33	50S ribosomal protein L27
EHS83217.1	4	66	50S ribosomal protein L12/L7
WP_003644091.1	4	25	50S ribosomal protein L15
WP_003640604.1	4	22	ATP-dependent 6-phosphofructokinase
WP_003641094.1	4	22	exopolyphosphatase
EHS84214.1	4	30	cystine ABC transporter, substrate binding protein
EHS82245.1	4	23	DegV family protein
WP_003640710.1	4	15	molecular chaperone DnaJ

Continued on next page

Table C.2 Common protein hits – continued from previous page

Accession Number	Unique Peptides	Coverage [%]	Protein description
EHS82278.1	4	11	topoisomerase IV, subunit B
WP_003640711.1	4	11	molecular chaperone DnaK
WP_003644206.1	4	11	ATP-dependent Clp protease ATP-binding subunit
EHS83202.1	4	16	DNA repair protein RadA
WP_003640588.1	4	17	ribosome biogenesis GTPase Der
EHS82682.1	4	10	alanine-tRNA synthetase
EHS84196.1	4	14	serine-type D-Ala-D-Ala carboxypeptidase
EHS82225.1	4	22	oxidoreductase, FAD-dependent
EHS82590.1	4	21	ribose-phosphate pyrophosphokinase
WP_003640749.1	4	21	hydroxymethylglutaryl-CoA synthase
EHS82552.1	4	15	trigger factor / peptidylprolyl isomerase
AAS21888.1	4	26	response regulator PlnD
WP_003644422.1	4	20	FADH(2)-oxidizing methylenetetrahydrofolate–tRNA-(uracil(54)-C(5))-methyltransferase TrmFO
EHS82876.1	4	15	aminodeoxychorismate lyase
WP_003642739.1	4	13	ABC-F family ATP-binding cassette domain-containing protein
EHS84203.1	4	14	GTP-binding protein/GTPase, Obg family, YchF subfamily
WP_003641555.1	4	18	acetate kinase
WP_003641629.1	4	16	DNA polymerase III subunit beta
WP_003641545.1	3	26	YebC/PmpR family DNA-binding transcriptional regulator
WP_003644315.1	3	19	ATP-binding cassette domain-containing protein
EHS81996.1	3	12	NADH peroxidase
WP_003640554.1	3	12	manganese-dependent inorganic pyrophosphatase
EHS81946.1	3	10	ABC transporter, ATP-binding protein
EHS83507.1	3	40	50S ribosomal protein L18
EHS82849.1	3	35	hydrolase, HAD superfamily
EHS82380.1	3	11	DNA-directed RNA polymerase, sigma factor 42
EHS81927.1	3	34	nucleotide-binding protein, universal stress protein UspA family
EHS84359.1	3	16	**extracellular protein, gamma-D-glutamate-meso-diaminopimelate muropeptidase (LytF)**
WP_003644321.1	3	8	DAK2 domain-containing protein
EHS83406.1	3	8	GMP synthase (glutamine-hydrolysing)
WP_003644323.1	3	16	phosphate acyltransferase PlsX
WP_003644319.1	3	15	methionyl-tRNA formyltransferase

Continued on next page

Table C.2 Common protein hits – continued from previous page

Accession Number	Unique Peptides	Coverage [%]	Protein description
WP_003642169.1	3	30	NADP oxidoreductase
WP_003642030.1	3	17	pur operon repressor
EHS82910.1	3	18	hypothetical protein nc8_1173
EHS82412.1	3	8	GTP-binding translation elongation factor LepA
EHS83223.1	3	8	ribonucleoside-diphosphate reductase, alpha chain
EHS81816.1	3	30	phosphohydrolase, MutT/nudix family
WP_003644345.1	3	44	universal stress protein
EHS81849.1	3	7	hypothetical protein nc8_2372
WP_003642490.1	3	9	adenylosuccinate lyase
WP_003644649.1	3	6	bifunctional glutamate–cysteine ligase GshA/glutathione synthetase GshB
EHS83479.1	3	23	deoxyadenosine kinase / deoxyguanosine kinase
EHS82285.1	3	6	DNA topoisomerase
WP_003644608.1	3	6	isoleucine–tRNA ligase
EHS82402.1	3	6	GTP pyrophosphokinase / pyrophosphohydrolase
WP_003641537.1	2	32	DUF948 domain-containing protein
EHS83881.1	2	9	ABC transporter, substrate binding protein
EHS83509.1	2	50	50S ribosomal protein L30
EHS82585.1	2	8	metallo-beta-lactamase superfamily protein
EHS82758.1	2	25	H(+)-transporting two-sector ATPase, B subunit
EHS81796.1	2	8	phosphoribosylaminoimidazole carboxylase, ATPase subunit
WP_003640832.1	2	10	**peptidoglycan endopeptidase/ autolysin**
EHS82413.1	2	8	poly(glycerophosphate chain) D-alanine transfer protein DltD
WP_003641258.1	2	37	50S ribosomal protein L24
EHS83468.1	2	61	cold shock protein CspC
WP_003644730.1	2	13	tyrosine-protein phosphatase
WP_003642592.1	2	11	phosphoribosylformylglycinamidine cyclo-ligase
EHS82080.1	2	17	bifunctional protein: methylenetetrahydrofolate dehydrogenase/methenyltetrahydrofolate cyclohydrolase
WP_003644631.1	2	14	glutamate racemase
WP_003642648.1	2	29	universal stress protein
EHS84049.1	2	31	nucleotide-binding protein, universal stress protein UspA family
WP_003638751.1	2	37	cell division regulator GpsB
WP_003640274.1	2	24	transcriptional repressor NrdR
EHS82058.1	2	12	bifunctional protein: transcription regulator/sugar kinase, ROK family

Continued on next page

Table C.2 Common protein hits – continued from previous page

Accession Number	Unique Peptides	Coverage [%]	Protein description
WP_003641075.1	2	6	**glutamine–fructose-6-phosphate transaminase (isomerizing) (GlmS)**
WP_011101015.1	2	10	16S rRNA (adenine(1518)-N(6)/adenine(1519)-N(6))-dimethyltransferase RsmA
WP_003641444.1	2	35	F0F1 ATP synthase subunit epsilon
WP_003644211.1	2	16	adenine phosphoribosyltransferase
WP_003640905.1	2	12	homoserine dehydrogenase
EHS83760.1	2	8	serine protease HtrA
WP_003640954.1	2	23	nucleoside deaminase
EHS83685.1	2	19	**phosphocarrier protein Hpr**
EHS82444.1	2	15	uridylate kinase
EHS82742.1	2	9	L-2-hydroxyisocaproate dehydrogenase
WP_003641340.1	2	9	CamS family sex pheromone protein
EHS83716.1	2	45	30S ribosomal protein S18
WP_003644354.1	2	9	Cof-type HAD-IIB family hydrolase
EHS83299.1	2	9	P-loop ATPase family protein
WP_003641770.1	2	7	extracellular solute-binding protein
EHS84100.1	2	11	transcription regulator, AraC family
EHS82695.1	2	3	DNA mismatch repair protein MutS/HexA
WP_003644638.1	2	8	ribonuclease Y
EHS82614.1	2	8	**UDP-N-Acglucosamine-N-Acmuramyl-(pentapeptide) pyrophosphoryl-undecaprenol N-Acglucosamine transferase**

Note: All proteins taken from *Lb. plantarum* NC8 database
Proteins of interest are highlighted in boldface

Table C.3: Protein matches unique to HA-GccH induced pulldown

Accession	Unique Peptides	Coverage [%]	Description
WP_003641196.1	3	15	MULTISPECIES: aspartate–ammonia ligase
WP_104795666.1	2	71	UDP-N-acetylmuramate–L-alanine ligase (MurC)
WP_003640869.1	2	21	MULTISPECIES: NAD-dependent epimerase/dehydratase family protein
WP_003643660.1	2	16	MULTISPECIES: bifunctional hydroxymethylpyrimidine kinase/phosphomethylpyrimidine kinase
WP_128383862.1	2	34	dihydroneopterin aldolase
WP_003644624.1	2	30	MULTISPECIES: S1 RNA-binding domain-containing protein
WP_003644674.1	2	18	MULTISPECIES: VOC family protein
WP_003639823.1	2	10	MULTISPECIES: HAMP domain-containing histidine kinase
WP_003640420.1	2	14	MULTISPECIES: 4'-phosphopantetheinyl transferase superfamily protein
WP_003641082.1	2	7	NADPH-dependent oxidoreductase
WP_003639152.1	2	35	MULTISPECIES: IreB family regulatory phosphoprotein
WP_021337811.1	2	8	serine hydrolase

Table C.4: Protein matches exclusive to $\mathrm{GccF}_{Cys18GlcNDAz}$ pulldown

Accession Number	Unique Peptides	Coverage [%]	Protein Description
WP_003644284.1	3	6	DNA polymerase I
EHS83030.1	2	16	tRNA-dihydrouridine synthase
EHS82281.1	2	7	ATP-dependent Hsl protease, ATP-binding subunit HslU
WP_003642022.1	2	11	alpha/beta hydrolase
WP_003640886.1	2	24	hypothetical protein
EHS83578.1	2	4	ATP-dependent DNA helicase PcrA
EHS82768.1	2	4	UDP-N-acetylmuramyl tripeptide synthase (MurE)

Note: All proteins taken from *Lb. plantarum* NC8 database

Table C.5: Protein matches exclusive to GccF$_{Cys43GlcNDAz}$ pulldown

Accession Number	**Unique Peptides**	**Coverage [%]**	**Protein Description**
WP_003644051.1	6	13	M1 family peptidase
WP_003637271.1	3	69	30S ribosomal protein S6
WP_003641442.1	3	19	F0F1 ATP synthase subunit gamma
WP_003644669.1	2	10	glucose-6-phosphate isomerase
WP_003644157.1	2	21	dTDP-4-dehydrorhamnose reductase
WP_003640313.1	2	18	uridine kinase
EHS83251.1	2	50	GroES co-chaperonin
WP_003641196.1	2	8	aspartate–ammonia ligase
EHS81723.1	2	5	pyruvate oxidase
EHS83554.1	2	6	citrate lyase, alpha chain
EHS83212.1	1	21	transcription antitermination protein NusG

Note: All proteins taken from *Lb. plantarum* NC8 database

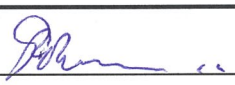
Table C.6: Protein matches common to both induced and uninduced *E. coli* pull-downs

Protein	**Description**
WP_001138117.1	MULTISPECIES: 30S ribosomal protein S19 [Bacteria]
ACT45365.1	heat shock chaperone [Escherichia coli BL21(DE3)]
ACT41916.1	chaperone Hsp70, with co-chaperone DnaJ [Escherichia coli BL21(DE3)]
ACT44996.1	30S ribosomal subunit protein S7 [Escherichia coli BL21(DE3)]
WP_000301864.1	MULTISPECIES: 50S ribosomal protein L2 [Bacteria]
ACT44952.1	30S ribosomal subunit protein S11 [Escherichia coli BL21(DE3)]
sp|CON_P02769|Serum	albumin OS=Bos taurus GN=ALB PE=1 SV=4 Tax_Id=9913 [Bos taurus]
ACT44970.1	50S ribosomal subunit protein L22 [Escherichia coli BL21(DE3)]
WP_000358960.1	MULTISPECIES: 50S ribosomal protein L18 [Bacteria]
ACT45802.1	Cpn60 chaperonin GroEL, large subunit of GroESL [Escherichia coli BL21(DE3)]
ACT44951.1	30S ribosomal subunit protein S4 [Escherichia coli BL21(DE3)]
WP_000805902.1	MULTISPECIES: lactose operon repressor [Bacteria] [Archaea]
ACT44317.1	30S ribosomal subunit protein S16 [Escherichia coli BL21(DE3)]
ACT44962.1	30S ribosomal subunit protein S14 [Escherichia coli BL21(DE3)]
sp|CON_P00761|Trypsin	OS=Sus scrofa PE=1 SV=1 Tax_Id=9823 [Sus scrofa]
WP_000529945.1	MULTISPECIES: 30S ribosomal protein S3 [Bacteria]
ACT44739.1	30S ribosomal subunit protein S21 [Escherichia coli BL21(DE3)]
WP_000246815.1	MULTISPECIES: 30S ribosomal protein S12 [Bacteria]
ACT44895.1	50S ribosomal subunit protein L13 [Escherichia coli BL21(DE3)]
ACT44958.1	30S ribosomal subunit protein S5 [Escherichia coli BL21(DE3)]
pdb|4P3Y|A	Chain A, Elongation factor Tu
WP_001216368.1	MULTISPECIES: 50S ribosomal protein L17 [Bacteria]
ACT44953.1	30S ribosomal subunit protein S13 [Escherichia coli BL21(DE3)]
ACT44836.1	30S ribosomal subunit protein S15 [Escherichia coli BL21(DE3)]
ACT41929.1	30S ribosomal subunit protein S20 [Escherichia coli BL21(DE3)]
ACT44894.1	30S ribosomal subunit protein S9 [Escherichia coli BL21(DE3)]
ACT43602.1	glyceraldehyde-3-phosphate dehydrogenase A [Escherichia coli BL21(DE3)]
WP_000135199.1	MULTISPECIES: 30S ribosomal protein S18 [Bacteria]
ACT44956.1	50S ribosomal subunit protein L15 [Escherichia coli BL21(DE3)]
WP_000579833.1	MULTISPECIES: 50S ribosomal protein L3 [Bacteria]
WP_001026276.1	MULTISPECIES: co-chaperone GroES [Proteobacteria]
WP_000424395.1	MULTISPECIES: 50S ribosomal protein L4 [Bacteria]
ACT44917.1	acetyl CoA carboxylase, BCCP subunit [Escherichia coli BL21(DE3)]
ACT44855.1	50S ribosomal subunit protein L21 [Escherichia coli BL21(DE3)]
WP_000729185.1	MULTISPECIES: 50S ribosomal protein L24 [Bacteria]
ACT44854.1	50S ribosomal subunit protein L27 [Escherichia coli BL21(DE3)]
ACT44968.1	50S ribosomal subunit protein L16 [Escherichia coli BL21(DE3)]
ACT45292.1	50S ribosomal subunit protein L28 [Escherichia coli BL21(DE3)]

MASSEY UNIVERSITY
GRADUATE RESEARCH SCHOOL

STATEMENT OF CONTRIBUTION
DOCTORATE WITH PUBLICATIONS/MANUSCRIPTS

We, the candidate and the candidate's Primary Supervisor, certify that all co-authors have consented to their work being included in the thesis and they have accepted the candidate's contribution as indicated below in the *Statement of Originality*.

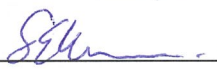
Name of candidate:	Sean William Bisset
Name/title of Primary Supervisor:	Associate Professor Gillian Norris
Name of Research Output and full reference:	
Total chemical synthesis of glycocin F and analogues: S-glycosylation confers improved antimicrobial activity. Chem. Sci., 2018, 9, 1686	
In which Chapter is the Manuscript /Published work:	Chapter 3
Please indicate:	
• The percentage of the manuscript/Published Work that was contributed by the candidate:	50%
and	
• Describe the contribution that the candidate has made to the Manuscript/Published Work:	
Carried out antimicrobial analysis on the glycocin F analogues provided.	
For manuscripts intended for publication please indicate target journal:	
Candidate's Signature:	
Date:	20/6/2019
Primary Supervisor's Signature:	
Date:	20/6/2019

(This form should appear at the end of each thesis chapter/section/appendix submitted as a manuscript/ publication or collected as an appendix at the end of the thesis)

MASSEY UNIVERSITY
GRADUATE RESEARCH SCHOOL

STATEMENT OF CONTRIBUTION
DOCTORATE WITH PUBLICATIONS/MANUSCRIPTS

We, the candidate and the candidate's Primary Supervisor, certify that all co-authors have consented to their work being included in the thesis and they have accepted the candidate's contribution as indicated below in the *Statement of Originality*.

Name of candidate:	Sean William Bisset
Name/title of Primary Supervisor:	Associate Professor Gillian Norris
Name of Research Output and full reference:	
Using chemical synthesis to probe structure-activity relationships of the glycoactive bacteriocin glycocin F. ACS Chem. Biol., 2018, 13, 1270-1278	
In which Chapter is the Manuscript /Published work:	Chapter 3
Please indicate:	
• The percentage of the manuscript/Published Work that was contributed by the candidate:	50%
and	
• Describe the contribution that the candidate has made to the Manuscript/Published Work:	
Carried out antimicrobial activity analysis on the glycocin F analogues provided, wrote paper.	
For manuscripts intended for publication please indicate target journal:	
Candidate's Signature:	
Date:	20/6/2019
Primary Supervisor's Signature:	
Date:	20/6/2019

(This form should appear at the end of each thesis chapter/section/appendix submitted as a manuscript/ publication or collected as an appendix at the end of the thesis)