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, Manawatū, New Zealand.

Marc Alex Bailie

2017

Abstract

The rise in antibiotic-resistant bacteria is becoming a severe public health problem because of the shortage of new antibiotics to combat existing resistant bacterial pathogens. Should this trend of increasing bacterial drug resistance continue, the previously treatable conditions may once again become fatal. Using broad-spectrum antibiotics causes collateral damage to the commensal microbiota of the host leading to complications and a greater susceptibility to opportunistic pathogenic infection. As a result, narrow spectrum antibacterials effective against specific pathogens, are becoming increasingly sought after. Among the many alternative classes of narrow-spectrum antibiotics, is a diverse group of ribosomally-synthesised antimicrobial peptides known as bacteriocins. Glycocin F (GccF), a rare and uniquely diglycosylated bacteriocin produced by *Lactobacillus plantarum* KW80, appears to target a specific N-acetylglucosamine (GlcNAc) phosphotransferase system (PTS) and causes almost instant bacteriostasis by an as yet unknown mechanism. This thesis demonstrates how the GlcNAc-PTS is involved in the GccF mechanism of action and that the *gccH* gene provides immunity to GccF. Using transgenic and gene editing techniques, regions of the GlcNAc-PTS were either removed or altered to prevent normal function before being tested *in vivo*. The results demonstrated that only the EIIC domain of the GlcNAc-PTS is required in the GccF mechanism of action and that it acts like a "lure" that attracts the bacteriocin to the main target that is as yet unknown. Furthermore, the immunity gene was discovered, and using PTS knockout cell lines the immunity mechanism was shown to act independently of the GlcNAc-PTS. This work will form the foundation for the work needed to unravel the bacteriostatic mechanism of action of GccF, which may lead to the development a novel antimicrobial agent.

My loving wife

A guiding light in the darkness, and a place of solitude and shelter through the tempest of life. The go-to authority on me and my work.

Mom and Dad

A pair of rare, irreplaceable models of excellence and support, who made all this possible.

In memory of

Aunty Hester Wallace

You gave me an addiction to stationery that would lead to literacy and ultimately science. There is no greater gift than intelligence, and with every caressing pen stroke, you are missed.

1957 - 2015

"dubito, ergo cogito, ergo sum"
Antoine Léonard Thomas, praise of Descartes, 1765.

Acknowledgements

Associate Professor Gill Norris and Dr Mark Patchett.
For your patience, guidance and supervision.

Mr Trevor Loo
For your technical expertise, and advice.

To all my colleagues and friends at X-Labs
A dear thank-you for being a good distraction, a source of debate, and allowing me to bounce ideas off you.

Institute of Fundamental Sciences
For housing and funding my research.

Contents

List of Figures

List of Tables

List of Abbreviations

Å	Angstrom
ADP	Adenosine diphosphate
Amp	Ampicillin
ATP	Adenosine triphosphate
bp	Base pair
cm	Centimeter
Chl	Chloramphenicol
Da	Dalton
DNA	Deoxyribonucleic acid
dNTP	Deoxyribonucleotide triphosphate
EAT	Empirical antibiotic therapy
EDTA	Ethylenediaminetetraacetic acid
EI	Enzyme I
EII	Enzyme II
EIIA	Enzyme IIA
EIIB	Enzyme IIB
EIIC	Enzyme IIC
EIID	Enzyme IID
$EIICBA^{GlcNAc}$	The GlcNAc specific PTS with all domains also known as the PTS18CBA
Ert	Erythromycin
FDA	Food and Drug Administration

g	Gram
GAS	Group A Streptococcal
gDNA	Genomic DNA
GccF	Glycocin F
GlcNAc	N-Acetylglucosamine
HP	Hairpins
HPr	Histidine-Phosphorylation protein
HPrK/P	HPr kinase/Phosphatase
kbp	Kilobase pair
kDa	Kilodalton
kPa	Kilopascal
L	Litre
LAB	Lactic acid bacteria
Lac	Lactose
MDR	multi-drug resistant
MDRO	multi-drug resistant organisms
M	Molar
MIC	Minimum inhibitory concentration
MCS	Multiple cloning site
mg	Milligram
ms	Millisecond
μL	Microlitre
μM	Micromolar
mL	Millilitre
mM	Millimolar
MRS	De Man, Rogosa and Sharpe medium
MW	Molecular weight
NaCl	Sodium chloride
NCBI	National Center for Biotechnology Information
NGS	Next generation sequencing

nL	Nanolitre
OD_{600}	Optical density at 600 nm
OF	Outward facing
PCR	Polymerase chain reaction
PDB	Protein data bank
PEG	Polyethylene glycol
PEP	Phosphoenopyruvate
PH	periplasmic helices
PMF	Proton motive force
PRD	PTS regulatory domain
PTM	Post translational modification
PTS	Phosphoenopyruvate phosphotransferase system
RBS	Ribosome binding site
SDS	Sodium dodecyl sulfate
SDS-PAGE	SDS-polyacrylamide gel electrophoresis
SLS	Streptolysin S
TBE	Tris-Boric Acid-EDTA
TEMED	N,N,N′,N′-Tetramethylethane-1,2-diamine
Tm	Melting temperature
TH	Transmembrane helix
V	Volts
v/v	Volume/Volume
w/v	Weight/Volume
WT	Wild-type
× g	Multiple of earth's gravitational force
°C	Degree Celsius

Introduction

1.1 A world defined by antibiotics

The discovery of antibiotics is arguably one of the defining features of the 20th century, where previously fatal injuries are now considered minor. Antibiotic treatment of infectious disease also decreases the morbidity and mortality of not only humans but also the plants and animals that sustain us. Our antibiotic-dependant way of life is currently under threat from multi-drug resistant (MDR) bacterial pathogens which is becoming a severe global public health problem[1,2]. Bacterial drug resistance is a naturally occurring phenomenon that is magnified by antibiotic overuse which increases selection pressure on bacterial cells, combined with a failure to prevent the spread of MDR organisms (MDRO) in both healthcare and community settings[3]. The New Zealand government reported that there are over 700,000 deaths globally each year caused by MDR bacteria and that if action is not taken this number could grow as large as 10 million by 2050[4].

Among these escalating issues are nosocomial infections, also known as hospital-acquired infections which are either contracted from the environment or staff of healthcare facilities[5]. Hospitals are plagued with pathogenic species of ***E****nterococcus faecium*, ***S****taphylococcus aureus*, ***K****lebsiella pneumonia*, ***A****cinetobacter baumanii*, ***P****seudomonas aeruginosa*, and ***E****nterobacter* (**ESKAPE**) that are resistant to several, in some cases, all known antibiotics in current use[6]. ESKAPE and other pathogens are leading us back to a pre-antibiotic era with devastating consequences[6]. A systematic review of nosocomial spontaneous bacterial peritonitis (N-SBP) revealed that MDROs are responsible for 22 to 73% of reported cases and that the survival from these infections is only 20% due to a lack of empirical antibiotic therapy (EAT)[7]. Over the last six years, only eight new antibiotic formulations

were approved by the Food and Drug Administration (FDA), which is very low in comparison to the 40 new anticancer agents approved during the same period[8].

The answer to the MDRO problem might be found in allelopathy, the production of toxic compounds by one organism that inhibits the growth of other organisms to which the producer is resistant thus manufacturing an environmental niche. Niches are not pre-defined and waiting to be filled but instead are partially engineered or constructed by the effects of an organism within a given environment[9,10]. For microorganisms, allelopathy has been widely demonstrated as a mediator of intra- and inter-specific interactions and therefore remains very necessary for the maintenance of microbial biodiversity[11,12].

The arsenal used by bacteria to suppress competition is extensive, and includes a range of compounds from metabolic by-products and traditional antibiotics such as bacitracin, to proteins such as bacteriolytic enzymes[12]. Among the proteinaceous weapons is a diverse class of ribosomally synthesised antimicrobial peptides known as bacteriocins, which when exported to the extracellular medium exhibit antagonistic activity against phylogenetically related bacteria[13].

1.2 Bacteriocins

Bacteriocins are more traditionally thought of as naturally produced, odourless, colourless, tasteless food preservatives, with low toxicity derived from lactic acid bacteria (LAB) that are used as starter cultures or in fermentation processes[14–16]. These LABs are generally regarded as safe (GRAS) because of their long history as food preservatives and commensal bacteria that have continually provided a rich source of bacteriocins[17–19]. These next generation antibiotics are not limited to only LABs but are common to almost all bacteria. It has been estimated that 30-99% of archaea and bacteria produce one or more bacteriocins[20], leaving a rich cache of unexploited novel antibiotics that are likely to be amiable to genetic engineering[21–23].

1.2.1 Classifications and mechanisms of action

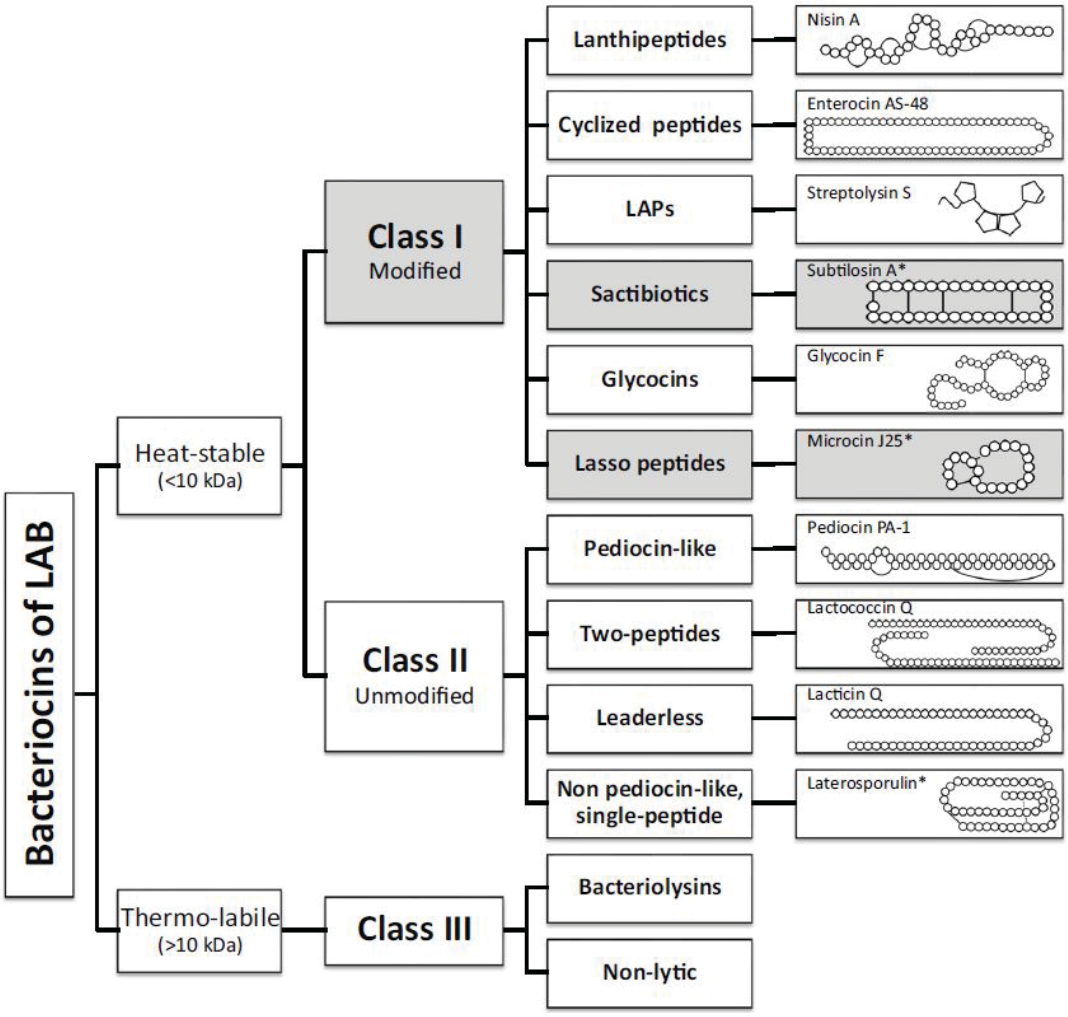


Figure 1: **Alvarez-Sieiro *et al.* classification of LAB bacteriocins**
The three class Alvarez-Sieiro *et al.* classification of LAB bacteriocins. **Class I Modified:** <10 kDa peptides that are heat-stable and have various post-translational modifications. **Class II Unmodified:** <10 kDa peptides that are heat-stable and do not have any post-translational modifications. **Class III:** large (>10 kDa) thermo-labile bacteriocins, that include both lytic and non-lytic variations[24]. The image was taken with permission from Springer Berlin Heidelberg as part of an open-access article distributed under the terms of the Creative Commons Attribution License. Source: Alvarez-Sieiro, P., Montalbán-López, M., Mu, D. et al. Appl Microbiol Biotechnol (2016) 100:2939. https://doi.org/10.1007/s00253-016-7343-9[24].

Due to the vast array of bacteriocins, and their producer strains, classifications have been controversial and under constant review as new bacteriocins are discovered. Typical classifications rank bacteriocins according to their chemical properties, proteolytic enzyme stability, genetic characteristics, presence or absence of post-translational modification, antimicrobial action, and producer organisms[19,25]. Alvarez-Sieiro *et al.* recently proposed an update on the well established Cotter *et al.* two class classifications[13], but is entirely focused on LAB bacteriocins and differs by reintroducing the previously excluded bacteriolysins and non-lytic bacteriocins as a thermo-labile third class (Figure 1)[24].

LAB bacteriocins initially divided into two groups based on heat stability. The first of these two groups (Heat-stable) is further divided into a Class I and Class II design that is commonly found in other classification schemes. The second division of the bacteriocin family is the large (>10 kDa) thermo-labile bacteriocins form Class III which contains two subgroups, bacteriolysins and large non-lytic bacteriocins.

1.2.1.1 Class I Bacteriocins

Class I modified bacteriocins can be subdivided into six subgroups: **Ia**) Lanthipeptides, **Ib**) head-to-tail cyclized peptides, **Ic**) sactibiotics, **Id**) Linear azol(in)e-containing peptides (LAPs), Class **Ie**) glycocins, Class **If**) lasso peptides.

Class Ia or Lanthipeptides undergo several enzymatic maturation reactions that cyclise the peptide *via* unusual lanthionine and/or (methyl)lanthionine amino acids to form (methyl)-lanthionine rings[26–28]. Lanthipeptides are a diverse family of bacteriocins with distinctly different maturation processes that have lead to further sub-classifications and controversy[29]. The most well studied of the Lanthipeptides is nisin , which has a two-step bactericidal mechanism of action. At low concentrations, nisin specifically targets Lipid II removing it from its natural location and preventing peptidoglycan synthesis (Figure 2). Peptidoglycan is a major component of bacterial cell walls,[30] and is part of an ATP dependent synthetic pathway that includes Lipid I which is a precursor molecule to Lipid

II during biosynthesises[31]. Nisin binding causes a reduction of membrane capacitance and prevents the lateral diffusion of lipid I and II[32] stopping cell wall synthesis, ultimately resulting in cell death[33]. At higher concentrations, the nisin that is bound to lipid II facilitates the binding of other nisin molecules leading to pore formation allowing the rapid efflux of small cytosolic biological molecules such as amino acids, cations and ATP from the cell, leading to cell death[28,31,34,35].

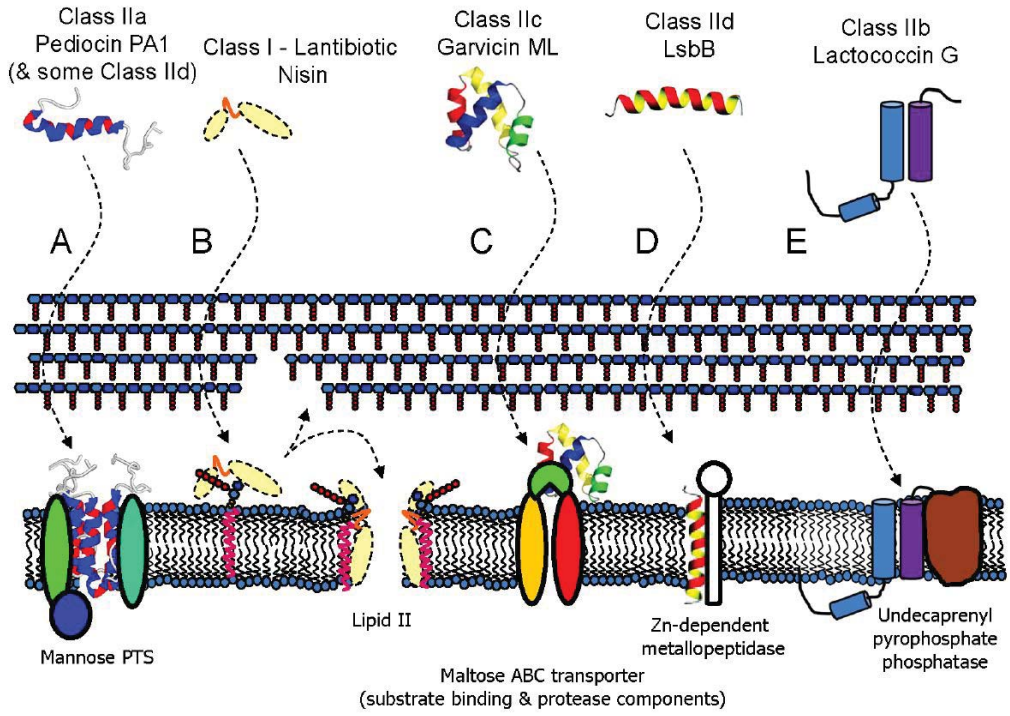


Figure 2: **Mechanisms of action of Class I and Class II bacteriocins**
Models for the mechanisms by which different bacteriocins target Gram-positive bacteria. **(A)** Mannose phosphoenolpyruvate phosphotransferase system targeted by class IIa and some class IId bacteriocins; **(B)** class I (lantibiotics) targeting lipid II and related peptidoglycan precursors to form pores in the membrane and allowing the efflux of small biological molecules; **(C)** the class IIc bacteriocin target involving the maltose ABC transporter (garvicin ML); **(D)** The class IId bacteriocin, LsbB (involving a Zn-dependent metallopeptidase) and **(E)** undecaprenyl pyrophosphate phosphatase (UppP) targeted by lactococcin G-like class IIb bacteriocins. Image from Cotter, P. D. (2014), An Upp-turn in bacteriocin receptor identification. Molecular Microbiology, 92: 1159–1163. doi:10.1111/mmi.12645[36] taken with permission from Elsevier `http://www.elsevier.com`.

Class Ib are head-to-tail cyclised peptides created by a peptide bond that links the N- and C-termini. The cyclisation allows the bacteriocin to form four to five α-helices that have a similar structure to a saposin fold[37–40]. The saposin-like fold was first noticed in enterocin AS-48 which is also one of the most studied bacteriocins of this class[41,42]. Class Ib

bacteriocins are typically active against a broad-spectrum of gram-positive bacteria, and some gram-negative bacteria when EDTA is added or at significantly higher concentrations than those used against gram-positive bacteria[43,44]. Cyclized bacteriocins have two dominant mechanisms of action, both of which generate pores in the bacterial cell membrane. Carnocyclin A (CclA) produced by *Carnobacterium maltaromaticum* UAL307 forms anion-selective ion channels that are activated in a voltage-dependent manner[45]. Enterocin AS-48 forms pH-dependant dimers in the solution which are the most dominant species at physiological conditions. On binding to a target bacterial cell membrane, these dimers undergo a conformational change from a closed to an open state resulting in the accumulation of positively charged molecules at the membrane surface which leads to a disruption of the membrane potential[46–48].

Class Ic, the Sactibiotics or Sactipeptides are very diverse in structure and are characterised by universal sulfur-to-α-carbon PTMs that cause hairpin structures[26]. Sactipeptides are typically not from LAB but have been included in the LAB classification because putative clusters have been identified in LAB using *in silico* methods[49]. Sactibiotics like thuricin CD from *Bacillus thuringiensis* DPC 6431 have a narrow, potent bactericidal range of activity against sporulating gram-positive bacteria like *Clostridium difficile*, the most notable target of thuricin CD[50]. The mechanism of action and cognate receptor for this class of bacteriocin has yet to be discovered. They function by insertion of the bacteriocin into the membrane to allow pore formation, membrane depolarisation and ultimately cell death.[51–53].

Class Id contains all **l**inear **a**zol(in)e-containing **p**eptides (LAPs). These peptides contain heterocyclic rings that are derived from cysteine, serine, and threonine residues that are altered by ATP-dependent cyclodehydration and subsequent flavin mononucleotide dependent dehydrogenation to form thiazole and (methyl)oxazole rings[54]. The most notable of the LAPs is Streptolysin S (SLS), a potent cytolytic toxin and virulence factor that is produced by nearly all *Streptococcus pyogenes* strains that cause **G**roup **A** **S**treptococcal (GAS) infections[55]. Although SLS has been rigorously studied for over a hundred years,

little is known about its mode of action and obtaining a structure has remained elusive[55]. SLS functions as a signalling molecule, aiding the movement of GAS cells across the epithelial barrier, and is involved in soft tissue damage, the elimination of neutrophils at infection sites and negatively impacts phagocytic cells[56–58].

Class Ie comprises the glycocins, a small group in which the polypeptide side chain is modified by a single monosaccharide that can be either O-linked, through a serine or threonine residue, or S-linked through a cysteine[26,59]. Little is known about their mechanism of action, apart from the fact that some appear to be bactericidal in mode, while others are bacteriostatic[60–64].

Glycocin F (GccF) is a 43-amino acid bacteriocin produced by *Lactobacillus plantarum* KW30 with extensive post-translational modifications that are essential for its antibiotic activity. GccF is modified by covalently bound N-acetylglucosamine moieties and (C-X6-C)2 fold in which two nested disulphide bonds stabilise a helix loop helix (Figure 3)[26,59].

GccF has been found to have a bacteriostatic effect against *Enterococcus*, *Streptococcus*, *Bacillus*, and *Lactobacillus*[60,65,66]. This activity is very dependant upon the two GlcNAc moieties, as enzymatic cleavage of the O-linked glycan renders the bacteriocin inactive, while activity is dramatically decreased with the removal of the S-linked glycan[60]. Interestingly bacteria under the effects of GccF induced bacteriostasis can be revived by adding GlcNAc to the media or sensitised to GccF activity by adding GlcNAc before the GccF treatment[60,66]. Both these events provide some indication that the cognate receptor of GccF may use GlcNAc as a substrate. In 2014, genetic screening was conducted on natural GccF resistant *Lactobacillus* mutants which universally showed mutations in the GlcNAc specific phosphoenolpyruvate phosphotransferase system (PTS^{GlcNAc}) and resistance could be fostered in susceptible cells by removing PTS^{GlcNAc} which provided validation that the first molecular interaction in the GccF mechanism of action is likely with the PTS^{GlcNAc}[67]. This interaction between GccF and PTS^{GlcNAc} is likely to occur at the outer leaflet of the membrane because ATP efflux using bioluminescence assay indicated

that neither small molecule efflux nor cell death was occurring when susceptible cells are treated with GccF[66]. Beyond this interaction, little else is known about the mechanism of action of glycocins or how the producer strain remains immune to GccF.

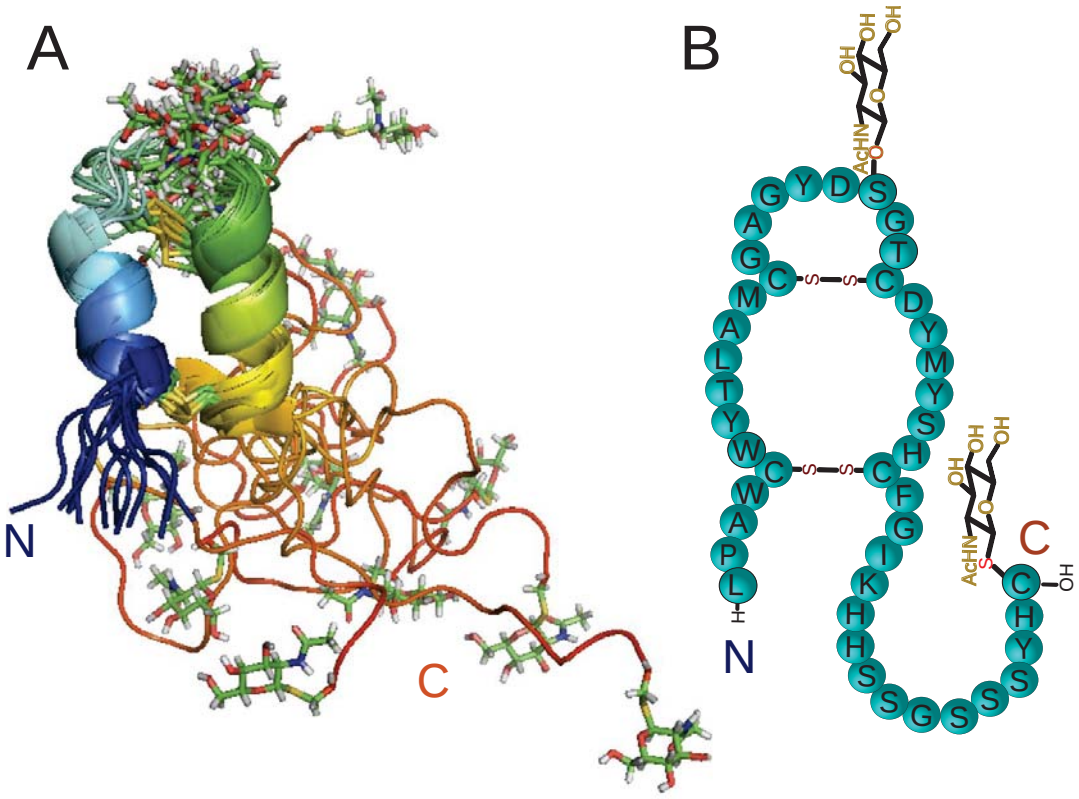


Figure 3: **Glycocin F structure**
(A) Superposition of 12 glycocin F NMR structures as a cartoon. Displaying the main backbone (N-term Blue to C-term Red), disulphide bonds (Yellow), and S- and O- linked N-acetylglucosamine as sticks coloured CHNOS (PDB: 2kuy)[59]. **(B)** A 43-amino acid stylised representation of the glycocin F structure including the amino acids sequence, disulphide bonds and S- and O- linked N-acetylglucosamine. N-terminal (Dark Blue) and C-terminal (Orange) are indicated by N and C respectively on both structures.

(Figure 3A) Image from Gillian E Norris, Mark L Patchett, "The glycocins: in a class of their own", In Current Opinion in Structural Biology, Volume 40, 2016, Pages 112-119, ISSN 0959-440X, `https://doi.org/10.1016/j.sbi.2016.09.003`[68] taken with permission from Elsevier `http://www.elsevier.com`.

Class If or lasso peptides named so because of their characteristic lasso shape caused by a macrolactam ring comprised of seven to nine residues that is formed between the N-terminal α-amino group and an aspartate or glutamate side chain. The remaining C-terminal peptide tail remains linear and is occasionally threaded through the loop forming a structure reminiscent of a lariat knot[26,69]. Lasso bacteriocins are very diverse in modes and mechanisms of action and in very few cases their molecular targets have been identified[70–78]. Unfortunately lasso peptides have only been bioinformatically identified rather than physically isolated from LAB[24]. Some notable examples include Microcin J25 which enters the cell by targeting siderophore transporter FhuA and temporarily prevents RNA transcription. Lassomycin, on the other hand, is a protease inhibitor. Other lasso peptides have anti-viral and antifungal activities[79–81].

1.2.1.2 Class II and III Bacteriocins

Class II contains unmodified peptides subdivided into pediocin-like bacteriocins, two-peptide bacteriocins, leaderless bacteriocins, and non-pediocin-like bacteriocins. Each of these subclasses can be even further subdivided into multiple independent categories[24,26,36]. Class II bacteriocins, like those of class I are mechanistically diverse, but appear to be largely bactericidal in mode. The class II targets are also generally better understood than class I bacteriocins (Figure 2).

Class IIa (pediocin-like) and some **class IId** (non-pediocin-like) have been shown to target the mannose phosphotransferase system (PTS^{Man}) to facilitate membrane insertion and pore formation. Zn-dependent metallopeptidases are the targets of other class IId bacteriocins such as LsbB from *Lactococcus lactis* subsp. *lactis* BGMN1-5[82–84].

Class IIb or two-peptide bacteriocins like Lactococcin G from *Lactococcus lactis* subsp. *lactis* (*Streptococcus lactis*) require both peptides to be fully active[85,86]. These two peptides form a membrane-penetrating helix-helix structure that targets undecaprenyl pyrophosphate phosphatase (UppP), causing membrane leakage and cell death[87,88].

Class IIc bacteriocins are synthesised without out a leader peptide that is usually present in bacteriocins to ensure that they are inactive while inside the producer strain, but are instead exported by a dedicated ABC transporter[89]. Enterocin 7A (Ent7A) and enterocin 7B (Ent7B) are both broad-spectrum antibiotic agents isolated from *Enterococcus faecalis* 710C which is found in beef products[89]. Enterocin 7A (Ent7A) and enterocin 7B (Ent7B) are circular and have a synergistic activity without significant binding in solution, a feature of many class IIc bacteriocins[38,90]. Lacticin Q is another example of a circular leaderless bacteriocin that causes membrane leakage without a specific membrane docking receptor[91,92]. A receptor-free mechanism of action may be commonly found among class IIc peptides but is not universal[93]. Recent evidence shows that garvicin ML, a bacteriocidal bacteriocin produced by *Lactococcus garvieae* DCC43 targets the maltose ABC transporter in *Lactococcus lactis* at low concentrations, but still possesses a target independent killing mechanism at high concentrations[93–95].

Class III bacteriocins are large proteins (>10 kDa), heat-labile bacteriocins that usually have multiple domains and two main mechanisms of bactericidal activity: lytic and non-lytic. For instance zoocin A from *Streptococcus equi* subsp. *zooepidemicus* has both a catalytic[96] and a recognition domain[97] and causes cell lysis by cleaving the peptidoglycan cross-links of the target (*streptococci*) cell wall[98]. Non-lytic class III bacteriocins cause cell death without lysing the cells, instead enter the cell and prevent DNA biosynthesis like caseicin from *Lactobacillus casei* or by inhibiting glucose and/or mannose by targeting their respective PTS transporters seen in the dysgalacticin mechanism of action, which also causes membrane leakage[99,100].

1.2.2 Bacteriocin gene clusters

Bacteriocin gene clusters may be chromosomally or plasmid-encoded, and contain the genes for the bacteriocin structure, maturation, export, and immunity which are usually arranged in close proximity to one another (Figure 4)[13,20,24,101]. For example glycocin F contains 8 genes in the cluster with gccA-E containing no intergenic spaces between them[60]. *GccA* was found to have sequence similarity to a type 2 glycosyl transferase

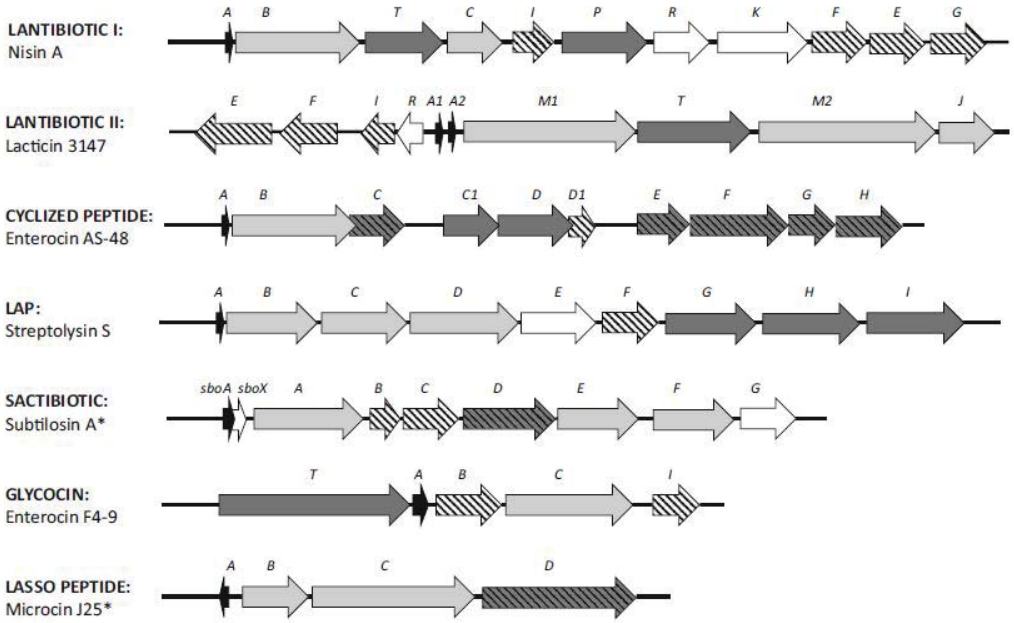
which is predicted to glycosylate both Ser18 and/or Cys43 of preglycocin F, *gccB* encodes a dedicated ABC transporter, while the *gccB* gene product cleaves the leader peptide and exports GccF[60]. Further maturation is thought to be carried out by the *gccC* and/or *gccD* genes that encode for exoplasmic membrane-anchored proteins containing thioredoxin-like domains predicted to facilitate disulfide bond formation[60]. The last gene in the cluster is *gccE* which includes a LytTR response domain and a second domain with no homology to any known protein, which sits upstream of preglycocin F gene *gccF*[60]. Two additional genes, *gccH* and *gccI*, sit outside the gene cluster and their functions are not known.

Zhao & Kuipers recently reported that genome mining 328 strains of 57 species of Bacillales identified 583 putative bacteriocin gene clusters, while another study found 785 putative bacteriocin gene clusters from 238 complete LAB genomes deposited in public databases[24,102]. These new results indicate that there may be between 24 - 29 glycocins that have yet to be identified, and a huge cache of other uncharacterised bacteriocins that could provide blueprints for new antimicrobial compounds.

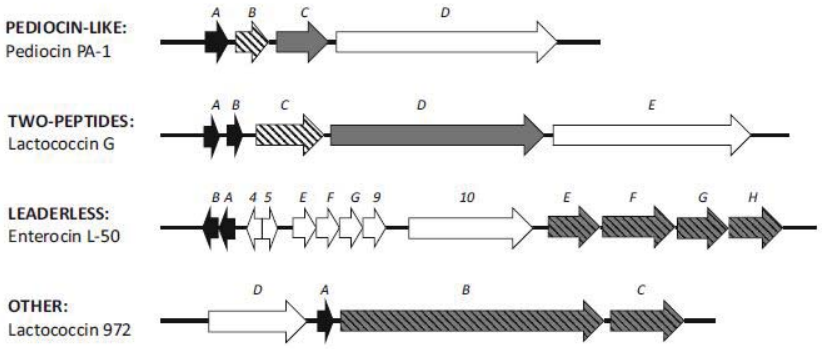
Figure 4: **Bacteriocin gene clusters separated by class**

Bacteriocin gene clusters represented as directional schematics (not drawn to scale). Bars represent the immunity genes; black arrows indicate the structural genes; dark grey arrows show transporter genes; light grey arrows represent modification and maturation genes; white shows other genes that are either unidentified or are specialised to that specific bacteriocin. *Bacteriocins from non-lactic acid bacteria. Source: Alvarez-Sieiro, P., Montalbán-López, M., Mu, D. et al. Appl Microbiol Biotechnol (2016) 100: 2939. https://doi.org/10.1007/s00253-016-7343-9[24].

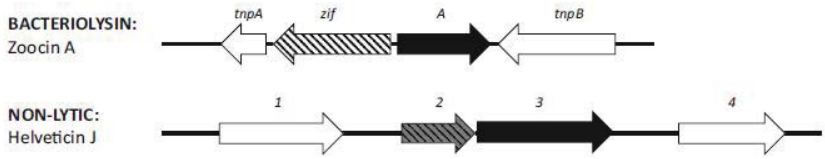
CLASS I



CLASS II



CLASS III



1.3 Phosphoenolpyruvate Phosphotransferase Systems

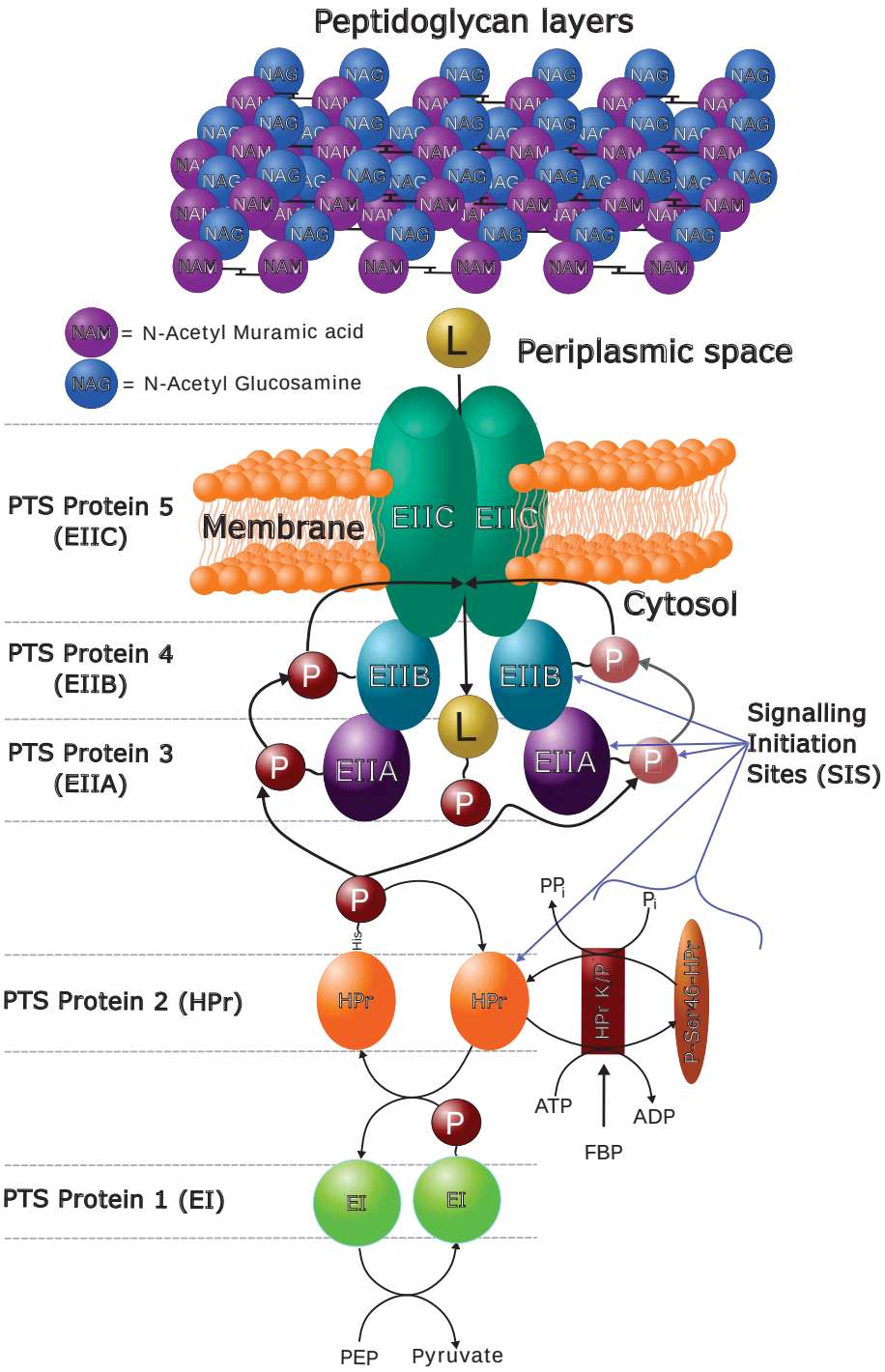


Figure 5: **PTS-mediated sugar transport and phosphorylation**

The PTS-Mediated Sugar (L: Ligand) transport and phosphorelay. The five PTS proteins are labelled 1 - 5: EI, HPr, EIIA, EIIB and EIIC respectively. The phosphorelay functions in bimolecular interactions starting at PEP which phosphorylates EI (Red). The phosphate is then transferred up the relay from EI to HPr which spontaneously phosphorylates EIIA. EIIA spontaneously phosphorylates EIIB which then transfers the phosphate to the ligand (L: Gold) as it enters the cell through EIIC. Signal initiation sites (SIS) are indicated with blue arrows and represent molecular interaction points that give rise to signal transduction, quorum sensing and various forms of metabolic regulation and control[102–104].

The bacterial phosphoenolpyruvate (PEP) carbohydrate phosphotransferase systems (PTS) are a family of sugar transporters found in all *bacteria* and some *archaea*[104]. The PEP-PTS were discovered more than 50 years ago[105–107] and are implicated in the transport and phopsphorylation of a wide variety of molecules like amino sugars such as glucoselysine and fructoselysine, as well as gluconic acids, glucosaminate, hexoses and more[103,108,109].

1.3.1 PEP-PTS order and phosphorelay

PEP-PTS is composed of five proteins, four soluble and one membrane spanning protein or protein-complex that relay phosphates *via* bimolecular interactions that facilitate active transport of the substrate (Figure 5)[110].

Enzyme I (EI) and the histidine phosphocarrier protein (HPr) are the most universal of the cytosolic proteins and provide the initial phosphates used by the phospho-relay system to most PEP-PTS. Enzyme I (EI) is phosphorylated by phosphoenolpyruvate which in-turn transfers the phosphoryl group to HPr. HPr directly phosphorylates Enzyme II (EII) complex *via* a bimolecular interaction with EIIA which donates the phosphoryl group to EIIB where the membrane spanning EIIC (and sometimes EIID e.g. fructose PTS) catalyses the coupled translocation and phosphoryl transfer from EIIB to the incoming sugar (Figure 5)[110,111]. These domains can appear as individual proteins or connected together by long flexible linkers in a single polypeptide chain[111]. The domain order of EII is similar in all PTS transporters, being comprised of a membrane bound EIIC, followed by EIIB and EIIA domains. Despite this common organisation of Enzyme II and the conserved phospho-carrying histidine and cysteine residues in EIIA and EIIB respectively, there is no structural or sequence homology shared between any of them[112–126]. Enzyme II is involved in specific carbohydrate transport, to avoid confusion a three letter superscripted abbreviation will be used to indicate the PTS substrate or regulatory function, for example $EII^{\mathbf{Mtl}}$ for EII Mannitol, or $EII^{\mathbf{Glc}}$ for EII Glucose.

1.3.2 EIIC and substrate import

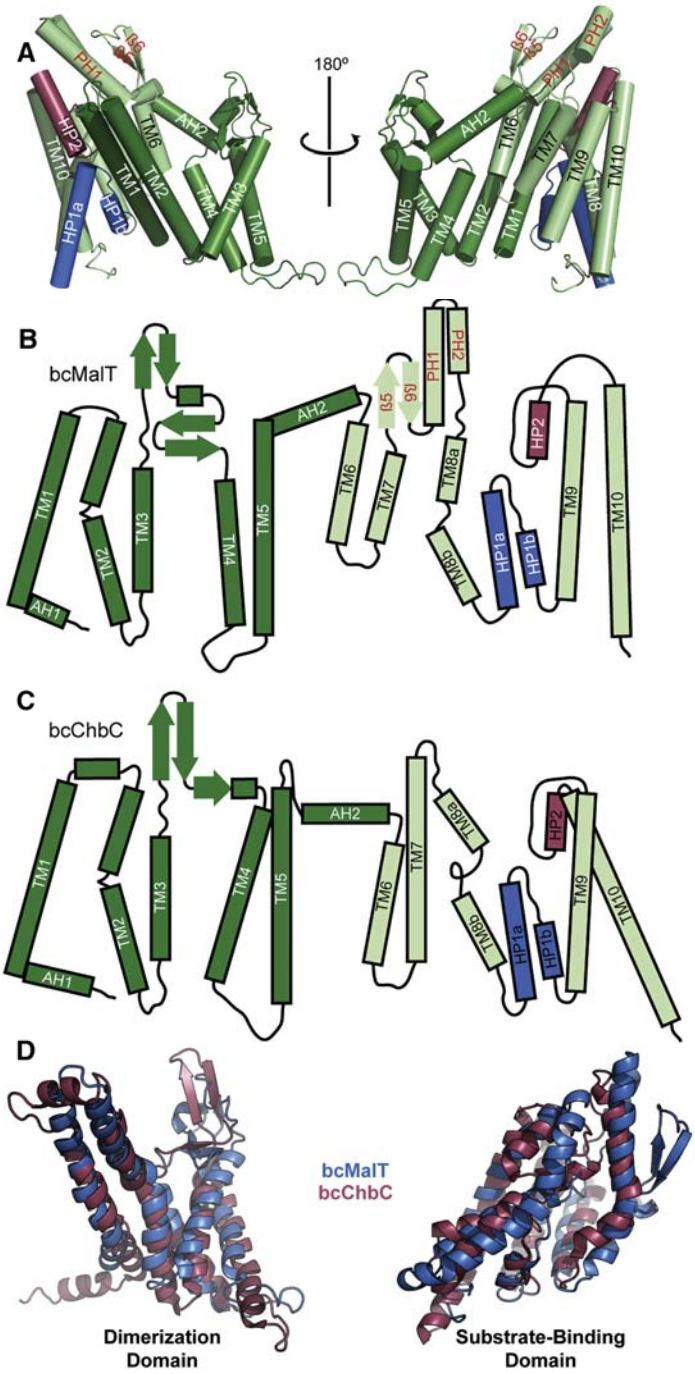


Figure 6: **EIIC topology diagrams**

(A and B) Topology diagrams of $EIIC^{MalT}$ oriented with the periplasmic side up. **(C)** Topology diagram of $EIIC^{ChbC}$ coloured according to the scheme in (A). **(D)** Superposition of the dimerisation regions (TM1-5, left) and substrate-binding regions (TM6-TM10, right) of $EIIC^{MalT}$ (Raspberry) and $EIIC^{ChbC}$ (Blue). **Key**: Dimerisation region (Dark Green), Transport region (Light Green), HP1 (Blue), HP2 (Raspberry), structural elements not shared with $EIIC^{ChbC}$ (Red labels) including TM7, PH1, PH2, two antiparallel β-strands b5 and b6. Image from McCoy, J. G. et al. "The structure of a sugar transporter of the glucose EIIC superfamily provides insight into the elevator mechanism of membrane transport." Structure 24, 956–964 (2016).[127] taken with permission from Elsevier `http://www.elsevier.com`.

The glucose-fructose-lactose (GFL) superfamily is the largest of the PEP-PTS superfamilies and is considered to be the most physiologically important for prokaryotes and mechanistic understanding of the PEP-PTS transport mechanism[128–130]. Crystal structures of the EIIC domains of two members of this PEP-PTS family, the $EIIC^{ChbC}$ and $EIIC^{MalT}$ structures, have provided some insight into how they transport PTS specific substrates into the cell using an elevator mechanism[127,131]. The maltose PTS from *Bacillus cereus* (PTS^{MalT}) does not have its own EIIA domain, but the EIIC and EIIB are expressed as a single polypeptide much like the GlcNAc PTS ($EIIABC^{GlcNAc}$ are all joined as a single polypeptide), while the $EIIC^{ChbC}$ is not linked to any of the cytosolic proteins at all. The crystal structure of PTS^{MalT} was determined by using a trypsinized PTS^{MalT} containing residues 8 to 450 in which the EIIB domain was removed. This truncated PTS remained capable of substrate binding and facilitated diffusion of PTS specific ligands in the absence of the other PTS proteins. In this state the EIIC functions as a weak transporter and is capable of only facilitated diffusion[127,129].

The EIIC forms a symmetrical homodimer within the membrane leaflet where most of the EIIC domain is found (Figure 7A black lines). Very little of the domain is solvent accessible[127]. The structure shows that there are 10 transmembrane helices (TH), 2 periplasmic helices (PH), a couple of hairpin bends (HP), two amphipathic -helices, and -two strands (Figure 6). These secondary structural elements form two distinct domains within this monomer, a dimerisation domain, and a substrate domain. TM1, TM2, TM3, and TM5 (Figure 7A) are the major components of the dimerisation region and most of the interface between two protomers of $EIIC^{MalT}$. The dimerisation interface is mostly hydrophobic with a buried surface area of $\sim$2,700 $Å^2$ (Figure 7A dark numbered helices)[127]. The substrate binding region is formed by TM6 - TM10, β-strands 5 and 6, and PH1 and PH2 (TM6 and TM7 do not extend to the cytoplasmic surface) (Figure 6B). Although each substrate binding site is in close proximity of the dimerisation interface, the residues involved in substrate binding come from the substrate binding domain, and are conserved in other PTS transporters (Figure 7A)[127]. The PTS^{MalT} dimerisation and substrate binding domains compare well with those of $EIIC^{ChbC}$ when they are separated

and superimposed (Figure 6D). However due to differences in the relative orientations of these regions the alignment is poor for the entire monomer. Despite this difference in orientation the similarities between the individual EIIC regions and the universal involvement of HP1a, HP1b and HP2 in substrate binding provides enough support that these two EIIC domains are likely to have similar import mechanics[127].

The EIIC is the only PEP-PTS domain that binds and imports PTS specific substrates, and does so as a homodimer. Interestingly, each of the EIIC monomers remains capable of independent substrate binding and import using an elevator car substrate import mechanism (Figure 8). In the $EIIC^{MalT}$ crystal structure the substrate binding site is located on the periplasmic side but is too small to allow substrate access[127,129]. Based on three structures in different states, the Chbc, UlaA and Malt structures, it was determined that the latter structure is captured in an outward-facing (OF) occluded state of the transfer cycle. Modelling shows that the cavity is opened when TM7 rotates towards AH2 which moves the moderately conserved tyr249 located on TM7 away enlarging the opening to the cavity (Figure 7)[127,129]. Simulations using surface models showed that a small conformational movement of AH7 moved the OF occluded state to the OF open state providing evidence for the first two stages of the elevator car import mechanism that had not been shown by the $EIIC^{ChbC}$ crystal structure[127,129,131]. The substrate binding domain can be moved as a rigid body to the inward-occluded state without disrupting the substrate-protein interactions during movement towards the cytosolic side of the membrane which results in an inward-facing occuluded state that matches the crystal structure of ChbC providing evidence for the validity of the model[127,129,131]. The full model proposed by McCoy *et al.* uses all three crystal structures (Chbc, UlaA and Malt) to verify the elevator model for movement of PTS specific ligands from the OF to the inner face (IF) of the membrane (Figure 8). The movement of TM7 alters the conformation of the monomer to that of the OF occluded state (Figure 8B) based on the MalT crystal structure and modelling[127]. The rigid body movement of HP1, HP2 and TM7 with the bound substrate assumes the IF occluded state (Figure 8C) seen in the ChbC crystal structure which can then be modelled to show the IF open state (Figure 8D).[127,129,131,132].

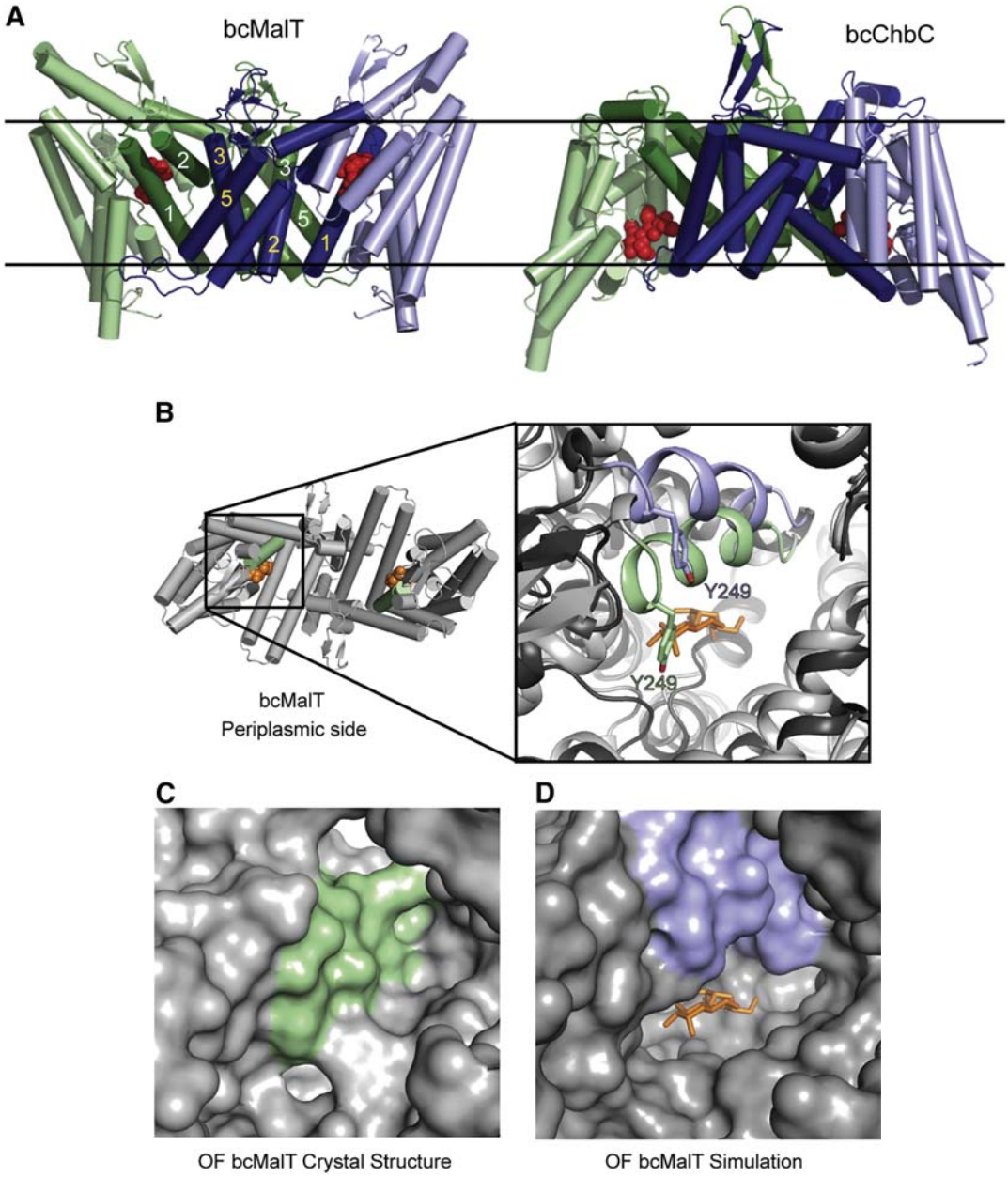


Figure 7: **EIIC dimerisation and substrate binding**

(A) Dimeric arrangement of EII^{MalT} (left) and EII^{ChbC} (right). The N-terminal dimerisation regions are coloured dark green and dark blue. The C-terminal substrate-binding regions are coloured light green and light blue. Maltose or $(GlcNAc)_2$ is shown as red spheres. The approximate location of the membrane is represented by black lines. Transmembrane helices in the dimerisation domains of both protomers of EII^{MalT} are labelled. **(B)** The EII^{MalT} crystal structure viewed from the periplasmic side. The bound maltose are shown in orange and the inset on right shows a close-up view of TM7 (Green) in the EII^{MalT} crystal. In the frame an extended MD simulation showing the release of the maltose (Blue) by moving TM7 causes Y249 shift to an open state. The maltose molecule from the EII^{MalT} crystal structure is shown as orange sticks. **(C and D)** Surface representation of the EII^{MalT} crystal structure (C) "closed" and the simulation model (D) "open" shown from the same orientation and colour to match the in the frame from B. Image from McCoy, J. G. et al. "The structure of a sugar transporter of the glucose EIIC superfamily provides insight into the elevator mechanism of membrane transport." Structure 24, 956–964 (2016).[127] taken with permission from Elsevier `http://www.elsevier.com`.

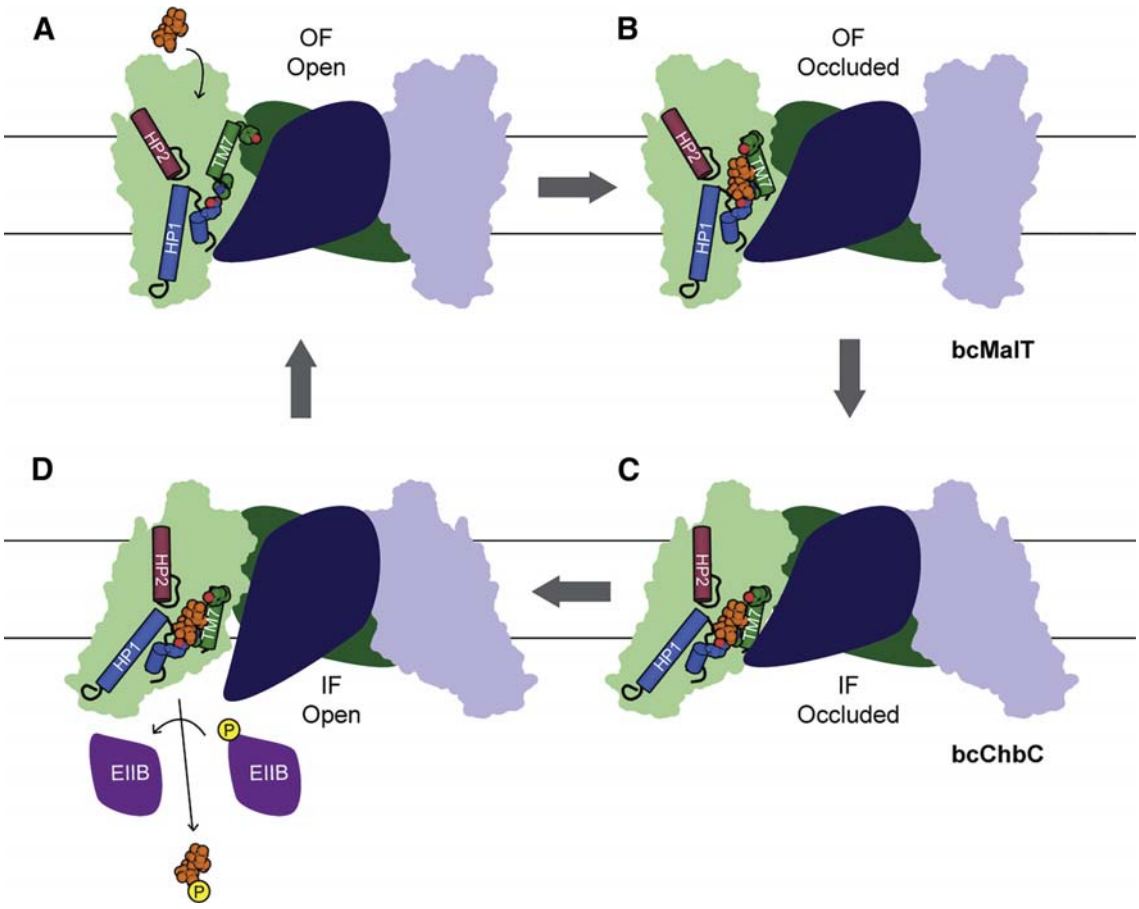


Figure 8: **PEP-PTS elevator car substrate import mechanism**

Diagram showing the model proposed by McCoy *et al.* for the glucose superfamily EIIC transport cycle. **(A–D)** The substrate binds to the outward-facing open structure **(A)**, which is then converted to the outward-facing occluded state **(B)** by a movement of TM7. The substrate-binding domain moves toward the cytoplasmic side of the membrane to form the inward-facing occluded state **(C)**, and final movement of the TM4-5 loop allows phosphorylation by EIIB and release of the phosphorylated substrate **(D)**. Image from McCoy, J. G. et al. "The structure of a sugar transporter of the glucose EIIC superfamily provides insight into the elevator mechanism of membrane transport." Structure 24, 956–964 (2016).[127] taken with permission from Elsevier `http://www.elsevier.com`.

1.3.3 Transduction and quorum sensing

PEP-PTS systems not only transport ligands and phosphorylate them, they also carry out regulatory roles[133]. They play roles in nutrient sensing, chemotaxis, quorum sensing, and other regulatory functions[104,110,134–136]. The involvement of the PTS systems and cellular functions are extensively reviewed by Deutscher *et al.*[110] and Galinier & Deutscher[104].

The involvement of PEP-PTS in regulation is highly variable. In some species incomplete transporters have exclusive regulatory roles while in other species PEP-PTS transporters are used to import specific substrates as well as play a regulatory role. The EIIB, EIIA and HPr components of the system can either interact with or phosphorylate histidine kinases of two-component signal transduction systems, catabolic enzymes, and transporters[104,137–139]. These interactions usually depend on the phosphorylation state of the PTS signalled through HPr or *via* a specific domain called the PTS regulation domain (PRD) found in PTS interacting proteins. These PRDs contain conserved histidine residues that are reversibly phosphorylated by PTS components.

Past research has revealed that it is likely that GccF binds to the EIICBA^{GlcNAc} (PTS18CBA)[67] during the GccF mechanism of action, however how these events lead to bacteriostasis or how the signal is transmitted when GccF interacts with the PTS remains a mystery. Signal transduction is possible through PTS transporters *via* either their phosphorelay or PRD binding proteins[104,110]. Three models for the involvement of PTS18CBA in the GccF mechanism of action were designed based on substrate import, and the various signal transduction pathways used by PTS transporters (Figure 9). The first model makes use of the phosphorelay as a signalling method when GccF binds to the EIIC^{GlcNAc} domain (Figure 9 Left). The second model assumes that signal transduction occurs through EIIA and EIIB domains *via* PRD containing proteins (Figure 9 Centre) and the third model accounts for alternative signalling methods that do not use the PTS for signalling but still include it in the mechanism of action (Figure 9 Right).

1.4 Hypothetical PTS^{GlcNAc}:GccF interaction models

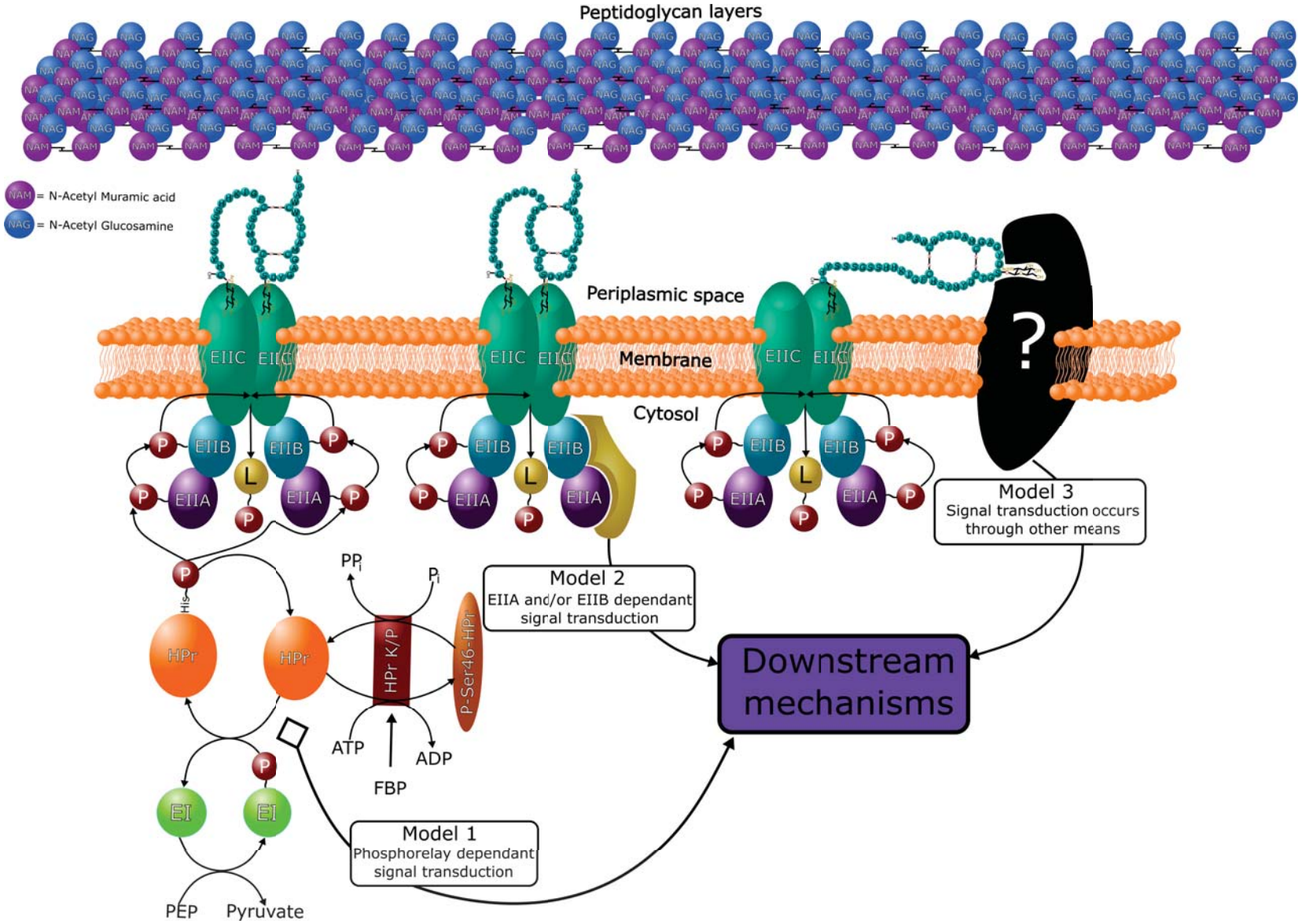


Figure 9: **Hypothetical PTS^{GlcNAc}:GccF interaction models**

The involvement of PTS18CBA in GccF induced bacteriostasis. **Model 1**: Signal transduction occurs through phosphorylation of HPr and/or other downstream mechanisms as a result of the phosphorelay system caused by the PTS18CBA:GccF interaction. **Model 2**: The EIIA and/or EIIB interact with PTS regulation domain (PRD) proteins to induce bacteriostasis in *L. plantarum* cells. **Model 3**: The PTS18CBA attracts and interacts with GccF, but transduction does not occur through PTS18CBA. The cell wall is shown as a full peptidoglycan layer, N-acetylmuramic acid and N-acetylglucosamine are represented by pink and blue spheres respectively. GccF is represented by a stylised amino acid sequence (marine). EIIA, EIIB and EIIC are coloured purple, blue and green respectively. Phosphate (Red) is shown to move along the phosphorelay by black arrows that originate from EI (light green) and move through the HPr (orange) phosphorylation cycle before spontaneously phosphorylating EIIA or an interaction with a secondary regulatory pathway that links to carbon metabolism through Fructose 1,6-bisphosphate (FBP).

1.5 Research goals

Aim 1: Optimize *L. plantarum* transformations.

L. plantarum transformations had been very difficult in our laboratory and was time consuming and have a low level of success. Therefore, the first aim was to improve the reliability and efficiency of established transformation procedures to confirm involvement of PTS18CBA in the GccF mechanism of action. This involved the following objectives:

1. Optimize the transformation process.

2. Design methods to determine when cells have been successfully transformed.

Aim 2: Determine the involvement of the PTS^{GlcNAc} in the GccF mechanism of action.
The involvement of the PTS in the GccF mechanism of action was previously shown using naturally occurring GccF resistant mutants and genome sequencing[67]. Support for this was found by generating PTS knockout mutant strains that proved more resistant to GccF than the wild type. The second aim was to reintroduce the PTS^{GlcNAc} back into knockout (KO) mutant strains on a pre-developed inducible plasmid and characterise the transformed cells to determine if the reappearance of the PTS can re-sensitise the cells to GccF by complementation. To achieve this aim involved the following objectives:

1. Generate high quality pRV-613:*pts18cba-Flag* plasmid stocks.
2. Using the optimized transformation method, transform *L. plantarum* NC8 KO cells with the pRV-613:*pts18cba-Flag*.
3. Characterise the GccF sensitivity of transformants along side control mutant strains to determine if the cells are able to be re-sensitised to GccF.

Aim 3: Establish how the PTS^{GlcNAc} is involved in the GccF mechanism of action.
Cells without the PTS^{GlcNAc} have increased resistance to GccF, which are also involved in quorum sensing and signal transduction through the cytosolic PTS domains and downstream molecular interactions that respond to the phosphorelay. PTS^{GlcNAc} is expressed as a single polypeptide that contains the EIIC, EIIB and EIIA domains allowing for genetic manipulation that will ultimately change the expressed protein. A two step approach will be used to test if these PTS regulatory functions were involved in the GccF mechanism of action. Firstly by removing the phosphorelay residues and stalling phosphate transfer and secondly by removing the internal EIIA and EIIB domains. This aim was achieve using the following objectives:

1. Use Quick change site direct mutagenesis to replace conserved the phospho-carrying His584 and Cys429 found in EIIA and EIIB respectively with Alanine and confirm the changes using DNA sequencing.
2. Develop a truncation method for altering pre-existing pRV-613:*pts18cba-Flag* plasmid stocks to exclude EIIA and EIIAB from the *pts18cba* gene insert.

3. Use the developed truncation method to remove the EIIA and EIIAB domains and confirm successful construct creation using DNA sequencing.

4. Transform site directed mutant and truncated plasmids into *L. plantarum* NC8 KO cells.

5. Characterise the newly transformed cells for sensitivity to GccF.

Aim 4: Determine if *gccH* provides immunity to GccF and if the PTS18CBA is required.

The GccF gene cluster conforms to the layout and structure of other bacteriocin gene clusters, However, the immunity gene remains unidentified. Two genes within the GccF cluster (*gccH* and *gccI*) have no identified function, but only *gccH* appears in both the *gccF* and its closest homologues (*asm1*) cluster. Aim 4 is to isolate the *gccH* gene from the *gccF* gene cluster and insert it into an inducible plasmid for transformation into GccF sensitive cells to determine if *gccH* can convey immunity and assess if the gene product requires the PTS18CBA to function. This aim was achieved using these objectives:

1. Isolate the *gccH* from the *gccF* gene cluster using PCR.

2. Replace the *LacZ* gene insert in pRV-613:*LacZ* with the GccH gene and confirm the change using DNA sequencing.

3. Transform the pRV-613:*gccH* plasmid into *L. plantarum* NC8 WT and *L. plantarum* NC8 KO cells.

4. Phenotype the newly transformed cells for sensitivity to GccF.

5. Compare *L. plantarum* NC8 WT and *L. plantarum* NC8 KO cells transformed with the pRV-613:*gccH* plasmid to determine if the PTS18CBA is involved in *gccH* generated immunity.

Material and Methods

2.1 Materials

2.1.1 Water

Deionised (dH_2O) or "Millpore" or "MilliQ" water used for buffers and genetic manipulation was obtained from a Barnstead™Nanopure™system (Thermo Scientific; Wilmington, DE, USA). Tap water (H_2O) was used for growth media when chemically defined media was not being used.

2.1.2 Luria-Bertani Broth (Miller's)

LB was made by adding the following to 800 mL H_2O. Bacto-tryptone (10 g), yeast extract (5 g), NaCl (10 g) and was adjusted to pH 7.5 with 10M NaOH before adding dH_2O to 1 L. The broth was sterilised by autoclaving at 121 °C for 20 minutes.

Alternatively 25 g of pre-made Luria-Bertani Broth base (Invitrogen Wilmington. DE, USA) was made up to 1L with tap water and sterilised at 121 °C for 20 minutes.

2.1.3 Luria-Bertani (LB) agar plates

25 g of Luria broth base and 1% (w/v) of Agar Bacteriological (Agar No. 1) (Oxoid Ltd; Basingstoke, England) was made up to 1 L with tap water and sterilised by autoclaving. The molten LB agar was equilibrated to 50 °C in a water bath and then poured into sterile Petri dishes.

2.1.4 de Man, Rogosa, & Sharpe broth

10 g of casein peptone (tryptic digest) was dissolved along with meat extract (10 g), yeast extract (5 g), D-glucose (20 g), Tween-80 (1 g), K_2HPO_4 (2 g), Na-acetate (5 g), $(NH4)_2$ citrate (2 g), $MgSO_4$-$7H_2O$ (0.2 g), and $MnSO_4$-H_2O (0.05 g), in 850 mL of distilled. Then Adjusted to pH 6.2 - 6.5 with acetic acid or sodium hydroxide (if required) before making the broth up to 1 L with H_2O and sterilised by autoclaving for no more than 15 minutes at 121 °C to give a golden liquid.

Alternatively 55 g of Lactobacilli de Man, Rogosa & Sharpe (MRS) broth (Acumedia Neogen Corporation, Lansing, Michigan, USA) was made up to 1 L with tap water and sterilised by autoclaving as above. MRS supplemented with copper or GccF was done so using high concentration filter sterilised $CuSO_4$ or GccF solutions. Typically 1 µL of these solution was required per 10 mL of MRS broth to reach required concentrations.

2.1.5 de Man, Rogosa, & Sharpe agar plates

1% (w/v) of Agar Bacteriological was added to the MRS media before being sterilised by autoclaving. The molten MRS agar was equilibrated to 50 °C in a water bath and then poured into sterile Petri dishes.

2.1.6 de Man, Rogosa, & Sharpe recovery media

0.4 M sucrose, 0.1 M $MgCl_2$, 0.5% w/v GlcNAc and 55 g of Lactobacilli de Man, Rogosa & Sharpe (MRS) broth (Acumedia Neogen Corporation, Lansing, Michigan, USA) was made up to 1 L with H_2O and sterilised by autoclaving.

2.1.7 RF_1 buffer

RbCl (100 mM), $MnCl_2$ (50 mM), potassium acetate (30 mM) and $CaCl_2$ (10 mM) were dissolved in 50 mL of dH_2O (these are final concentrations). The pH was adjusted to 5.8 using 10% glacial acetic acid and made up to 100 mL before being sterilised by passing the solution through a sterile 0.22 µm syringe filter into pre-autoclaved bottles.

2.1.8 RF_2 buffer

RbCl (10 mM), MOPS (10 mM), $CaCl_2$ and glycerol 15% (v/v) were dissolved in 50 mL of dH_2O volume (these are final concentrations). The pH was adjusted to 5.8 using 100 mM NaOH before being made up to 100 mL (total volume). The solution was sterilised by passing the solution through a sterile 0.22 µm syringe filter into pre-autoclaved bottles.

2.1.9 SOB media

5 g Bacto yeast extract and 2 mL of 5 M NaCl were dissolved along with 20 g Bacto tryptone in 900 mL. 2.5 mL of 1M KCl, 10 mL of 1M $MgCl_2$ and 10 mL of 1 M $MgSO_4$ were added to the Bacto solution which was then made up to 1 L with H_2O and sterilised by autoclaving.

2.1.10 SOC media

5 g Bacto yeast extract and 2 mL of 5 M NaCl were dissolved along with 20 g Bacto tryptone in 900 mL H_2O. 2.5 mL of 1M KCl, 10 mL of 1 M $MgCl_2$ and 10 mL of 1 M $MgSO_4$ were added to the Bacto solution which was then sterilised by autoclaving. 20 mL of filter sterilised 1 M glucose was then added to the sterile SOB media before adjusting it to 1 L with autoclaved H_2O using aseptic techniques.

2.1.11 2YT media

16 g Bacto tryptone was added to 10 g Bacto yeast extract and 5 g NaCl in 900 mL of H_2O. The pH was adjusted to 7.0 with 5 M NaOH and made up to 1 L with H_2O before sterilisation by autoclaving.

2.1.12 Polyethylene glycol 1500

Polyethylene glycol 1500 30% w/v was made fresh each day in dH_2O and sterilised by passing the solution through a sterile 0.22 µm syringe filter into pre-autoclaved bottles.

2.1.13 Antibiotics

Table 1: **Antibiotics stocks**

Antibiotic	Concentration ($mg \cdot mL^{-1}$)	Solvent
Ampicillin	100	dH_2O
Chloramphenicol	100	Absolute EtOH
Erythromycin	100	Absolute EtOH

The antibiotics (Table 1) were administered to both *E.coli* and *L. plantarum* cells. *E.coli* were treated with 100 - 150 $\mu g \cdot mL^{-1}$ ampicillin and Erythromycin. Dosages for chloramphenicol remained between 15 - 45 $\mu g \cdot mL^{-1}$ when treating *E.coli*. *L. plantarum* cells were treated with 15 - 30 $\mu g \cdot mL^{-1}$ of either chloramphenicol or erythromycin or both or 0.05 $\mu g \cdot mL^{-1}$ of ampicillin during the first stage of competent cell development. Ampicillin treatments did not significantly improve transformation results and were discontinued.

2.1.14 Bacterial strains

Table 2: **Base bacterial isolates**

Species	Genetic background	Abbreviation
E. coli EC100	ΔendA1 and ΔrecA1∗	EC100
L. plantarum NC8 KO	Δ*pts18cba*	NC8 KO
L. plantarum NC8	Wild type	NC8 WT
L. plantarum 8014	Wild type	8014
L. plantarum subsp. ATCC 14917	Wild type	14917

All cells were provided by the laboratory cell library. ∗*E.coli* EC100 cells are endonuclease minus (*endA1*) and recombination minus (*recA1*) with the genotype *F*– *mcr*A Δ(*mrr-hsd*RMS-*mcr*BC) φ80d*lacZ*ΔM15 Δ*lacX*74 *rec*A1 *end*A1 *ara*D139 Δ(*ara*, *leu*)7697 *gal*U *gal*K λ–*rps*L *nup*G.

2.1.15 Alkaline lysis buffer

3% SDS was dissolved in 0.2 M NaOH.

2.1.16 Lysozyme-sucrose solution

Lysozyme was added to 25% w/v sucrose in dH_2O, to a final concentration of 30 $mg \cdot mL^{-1}$.

2.1.17 DNA loading dyes

2.1.17.1 6X Orange G DNA Loading Dye

0.15% (w/v) Orange G was added to 10 mM Tris-HCl (pH 7.6), and 60% glycerol, 60 mM EDTA in dH_2O and stored at 4 °C. Alternatively 0.15 g of Orange G was dissolved in 2 mL 50X TAE, 60 mL glycerol (60% v/v), and made up to 100 mL with dH_2O and stored at 4 °C.

2.1.17.2 6X bromophenol blue DNA Loading Dye

30% (v/v) glycerol, 0.25% (w/v) bromophenol blue, and 0.25% (w/v) xylene cyanol FF in dH_2O, stored at 4 °C.

2.1.18 Protein staining

2.1.18.1 Coomassie stain

For 100 mL of stain, 0.25 g of coomassie brilliant blue R250 was dissolved in 45 mL of methanol and 45 mL of dH_2O and added 10 mL of glacial acetic acid.

2.1.18.2 Colloidal Coomassie stain

1 L of Colloidal Coomassie stain was made by dissolving 100 g of ammonium sulfate in 500 mL dH_2O followed by the addition 24 mL of a 5% coomassie blue G-250, 118 mL of 85% (Ortho-) Phosphoric acid and 200 mL of absolute EtOH. The solution was stirred then filtered through Whatman paper and stored out of direct sunlight at room temperature.

2.1.19 SDS-PAGE Electrode (Running) Buffer

15.0 g of Tris, 72.0 g of glycine, 5.0 g of SDS was dissolved in 900 mL dH_2O and then made up to 1 L with dH_2O.

2.2 Methods

2.2.1 Sterilisation

Sterilisation was accomplished by either pressure cooking or autoclaving at 121 °C at 103.4 kPa for 18 minutes, or by passing solutions through a sterile 0.2 µm membrane filter.

2.2.2 Restriction digests

Table 3: **Restriction digests**

Restriction enzyme	10 units is sufficient, generally 1 μL was used
DNA	1 μg
10X NEBuffer	5 μL (1X)
Total reaction volume	50 μL
Incubation time	1 hour
Incubation temperature	enzyme dependant

Enzymes were sourced from a variety of suppliers. The same enzymes from different suppliers were also used interchangeably. The most common suppliers were Invitrogen™, New England Biolabs® incorporated (NEB), Sigma-Aldrich Roche®, and Promega.

2.2.3 Gel electrophoresis

2.2.3.1 Agarose

Agarose gels were made by microwaving (∼2minutes on high) the appropriate amount of HyAgarose™ agarose powder (HydraGene Co, Ltd) in tris-base (10.8 $g{\cdot}L^{-1}$), boric acid (5.5 $g{\cdot}L^{-1}$), ethylenediaminetetraacetic acid (EDTA) (2 mM) buffer (TBE). The mixture was swirled every 30 - 45 seconds until all the agarose was uniformly dissolved and

allowed to cool to 50 °C. It was then poured into either a Mini-Sub Cell GT Tank (Bio-Rad Laboratories; USA) or a Wide Mini-Sub Cell GT Tank (Bio-Rad Laboratories; USA) containing a Bio-Rad gel tray and casting wedges, and allowed to set for 20 - 30 minutes. The DNA samples were then loaded with either orange G or bromophenol blue loading dye along with 100 ng of Invitrogen™ 1 kb Plus DNA Ladder and run at 80 volts for ∼35 minutes. 0.6% w/v agarose was used for genomic DNA and 1% w/v agarose was used for all other general electrophoresis.

2.2.3.2 SDS-PAGE

Proteins were separated on the basis of mass by **S**odium **D**odecyl **S**ulfate **P**olyacrylamide **G**el **E**lectrophoresis (SDS-PAGE) using the discontinuous buffer method and sample buffer described by Laemmli[140].

Wells were loaded with protein samples prepared using the Laemmli[140] method and run alongside 6 µL of Benchmark™ Protein Ladder (Invitrogen™; Wilmington, DE, USA) using the Mini-Protean® 3 system (BioRad Industries, Milan, Italy) at 180 V at room temperature (20 °C to 25 °C) for ∼50 minutes or until the dye front was within 1 cm of the bottom of the gel.

2.2.4 Gel staining and image capturing

2.2.4.1 Imaging

All gels were imaged using the appropriate predefined Image Lab™ software protocol and a Bio-Rad Molecular Imager Gel Doc™ XR+ System.

2.2.4.2 Agarose

Agarose gels were stained post run in 0.5 $\mu g \cdot mL^{-1}$ ethidium bromide solution for 30 minutes, followed by a brief wash with dH_2O before imaging.

2.2.4.3 SDS-PAGE

SDS-PAGE gels were stained for at least 1 hour in coomassie or colloidal coomassie with gentle agitation at room temperature or overnight. Excess stain was decanted and the gels rinsed with dH_2O and destained with dH_2O (Colloidal Coomassie) or with 40% (v/v) methanol and 10% (v/v) acetic acid (Coomassie blue) with gentle agitation at room temperature.

2.2.5 DNA extraction

2.2.5.1 Plasmid DNA extraction from *E. coli*

E. coli cells were chemically lysed using a Wizard® Plus SV Miniprep DNA Purification Systems (Catalog # A1465) according to the manufacturer's protocol.

2.2.5.2 Plasmid DNA extraction from *L. plantarum*

Plasmid DNA harboured by *L. plantarum* cell was harvested by chemical lysis. All cells were treated with a lysozyme-sucrose solution (2.1.14) for 30 - 60 minutes before using either gram-positive protocol for the Wizard® Plus SV Miniprep DNA Purification System (Catalog # A1465), or the extraction method designed and outlined by O'Sullivan & Klaenhammer[141].

2.2.5.3 Genomic DNA extraction from *L. plantarum*

All cells were treated with a lysozyme-sucrose solution (2.1.14) for 30 - 60 minutes before using the Wizard® Genomic DNA Purification system (Catalog # A1125).

2.2.6 Bacterial indicator plates

30 mL of MRS agar (2.1.5) was made up in a 100 mL glass Kimax® media bottle and sterilised by autoclaving. The MRS agar was allowed to cool in a water bath until it reached 40 °C before being inoculated with *L. plantarum* cells from an overnight culture to give an OD_{600} of 0.01 and gently mixed to uniformly spread the cells throughout the MRS agar. The *L. plantarum* contaminated MRS agar was then poured into sterile Petri

dishes and allowed to set before 1μL of GccF sample was dotted onto the test surface of the plate.

2.3 Bacterial manipulation techniques

2.3.1 Chemically competent *E. coli*

E. coli EC100 cells were grown at 37 °C in 50 mL of LB broth until the culture reached OD_{600} ~0.5. The cells were chilled for 15 minutes on ice and pelleted by centrifugation at 2,600 x g for 20 minutes at 4 °C. The cell pellet was resuspended in 18 mL of RF_1 buffer and incubated on ice for 30 minutes. The cells were re-pelleted as before and the pellet re-suspended in 4 mL of RF_2 buffer. The cell suspension was divided into sterile micro-centrifuge tubes in 50 μL aliquots and immediately snap-frozen in liquid nitrogen and stored at -80 °C.

2.3.2 *E. coli* heat shock

E. coli of transformations were done using heat shock as described by Inoue *et al.*[142]. A 50 μL aliquot of chemically competent *E. coli* (2.3.1) was mixed and incubated on ice with 50 ng to 250 ng of plasmid DNA. The cells were subjected to heat shock at 42 °C for 60 to 90 seconds and immediately cooled on ice for 5 minutes. 500 μL of LB broth or S.O.C media was added to the chilled transformed cells and mixed. The cells were then allowed to recover at 37 °C for 45 minutes. After the recovery period the cells were spread onto LB agar selection plates supplemented with appropriate antibiotics and incubated at 37 °C overnight.

2.3.3 Electrocompetent *L. plantarum*

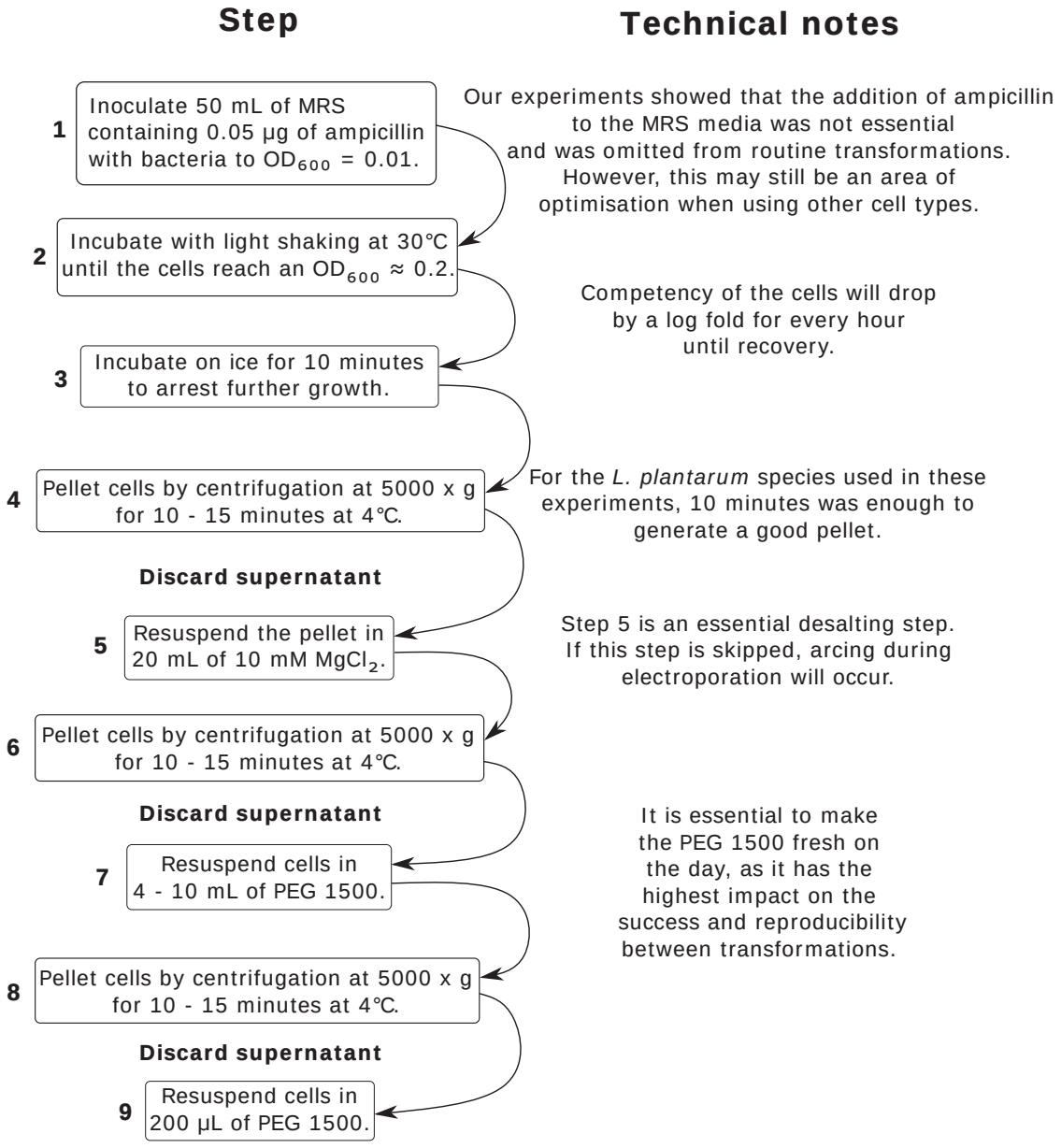


Figure 10: **Transformation flow chart: generating competent cells**
Method adapted from Drower[67] section 2.2.3.8 on page 36. A full discussion regarding the optimisations of this method can be found in 3.3.1.1.

2.3.4 *L. plantarum* electroporation

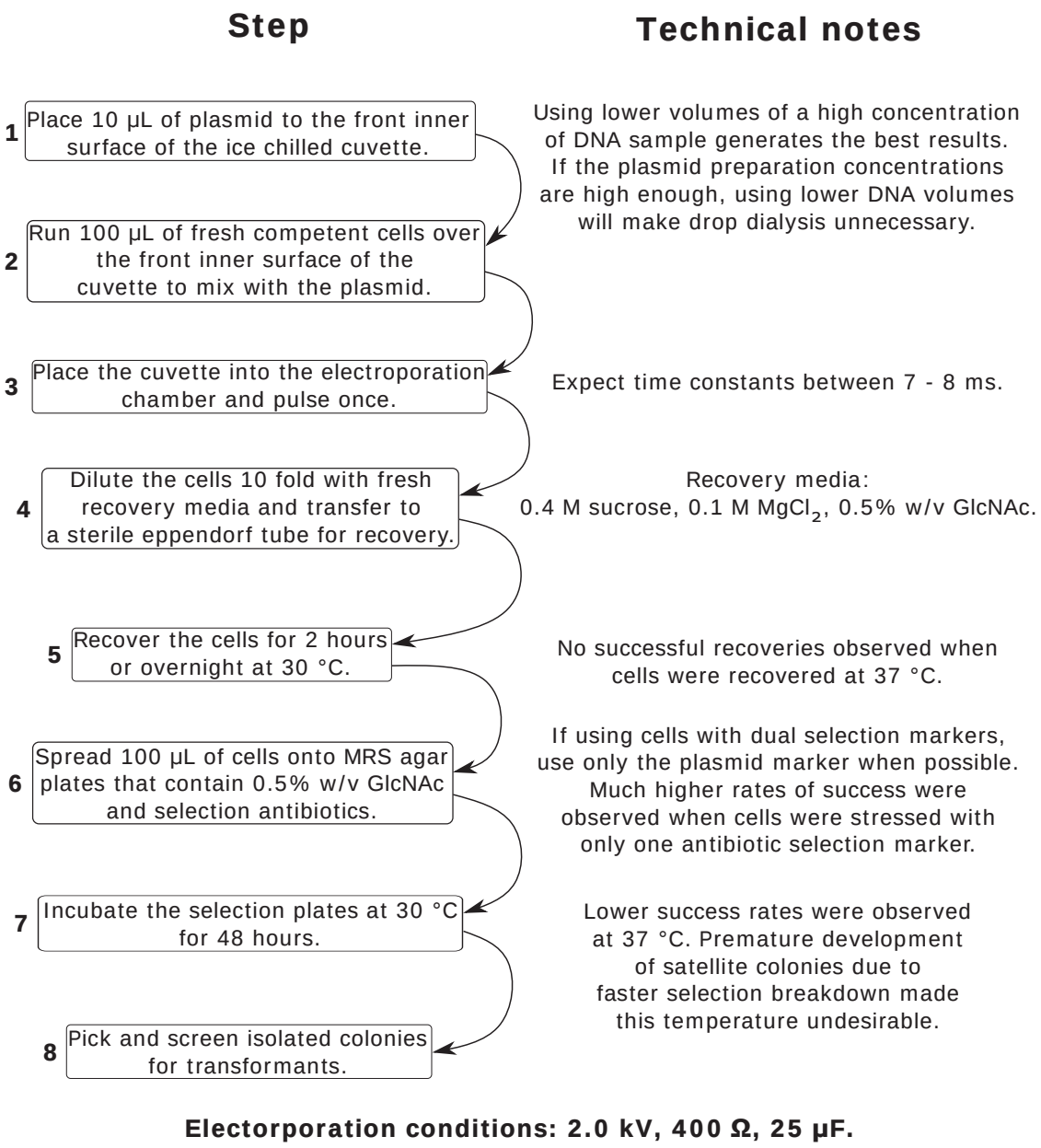


Figure 11: **Transformation flow chart: Electroporation procedure**
Method adapted from Drower[67] section 2.2.4.8 on page 44. A full discussion regarding the optimisations of this method can be found in 3.3.1.2.

2.3.5 Bacterial storage and revival

Bacterial stocks were made from overnight (10 - 13 hours) cell cultures grown with appropriate antibiotics. The cells were pelleted by centrifugation at 5000 x g, then re-suspended in fresh sterile media containing 20% glycerol (v/v) and transferred into a 1 mL CryoTube™ (Nunc™, Thermo Fisher Scientific) before snap freezing in liquid nitrogen and stored at -80 °C.

The cells were recovered from frozen stocks by either streaking onto the appropriate agar plates containing selection antibiotics or directly into 10 mL of antibiotic supplemented liquid media and incubated for 12 - 15 hours.

2.3.6 Growth conditions

2.3.6.1 General bacterial growth conditions

Bacterial cells were grown in liquid media or on solid agar. Liquid cultures were grown as 10 mL cultures in 40 mL autoclaved DURAN® borosilicate glass test tubes or as 50 mL cultures in 250 mL DURAN® wide neck Erlenmeyer flasks stoppered with sterile cotton bungs.

L. plantarum was usually grown at 30 °C. Incubation from frozen stocks in liquid media or on solid agar was always done overnight at 30 °C and recovery from transformations was done at 30°C for 48 hours on solid media. *L. plantarum* containing liquid growth media was shaken at no more than 150 RPM overnight or until the desired OD_{600} was reached.

E. coli containing liquid cell cultures were always grown at 37 °C with shaking at 200 RPM either overnight or until the desired OD_{600} was reached.

2.3.7 Bacterial screening

Transformed cells were first screened on agar plates under antibiotic selection. Resistant cells were picked and grown in liquid cultures overnight or until they reached an OD_{600} 0.8 in antibiotic supplemented liquid media. The cell cultures were pelleted at 5000 x g and the plasmid DNA extracted as described in section 2.2.5 before PCR and DNA sequencing.

2.4 DNA manipulation techniques

2.4.1 DNA quantification

Aqueous DNA concentrations were estimated using a NanoDrop™ ND-1000 Spectrophotometer v3.7 (Thermo Scientific; Wilmington, DE, USA) according to the manufacturer's instructions.

2.4.2 DNA desalting

DNA extractions were desalted using a piece of 0.025 µm 'V' series membrane (Millipore; Billerica, USA) shiny side up following the manufacturers specifications for 2 - 4 hours. Desalted DNA solutions were recovered and quantified using a NanoDrop™ ND-1000 Spectrophotometer v3.7 (Thermo Scientific; Wilmington, DE, USA).

2.4.3 Polymerase chain reaction

Table 4: **Standard PCR volumes for amplification of long and short products**

Component	Short	Long	KAPA*
10× reaction buffer**	5 µL	5 µL	25 µL
10 mM dNTPs mix	5 µL	10 µL	5 µL
10 µM Forward primer	2.5 µL	2.5 µL	-
10 µM Reverse primer	2.5 µL	2.5 µL	-
DNA template***	1 µL	1 µL	1 µL
DNA polymerase	0.5 µL	0.5 µL	-
dH_2O	33.5 µL	28.5 µL	25µL
Total Volume	50 µL	50 µL	56 µL

For a full PCR troubleshooting discussion see section 3.2.1.

Short PCR = standard PCR using Taq polymerase for products $\leq$ 3 kb. Long PCR differs from standard PCR by doubling the dNTP concentration to allow the formation of larger PCR products and the use of various high fidelity enzymes (3.2.1.).*KAPA HiFi HotStart Uracil+ ReadyMix (2X) (kit code: KK2801). **KAPA HiFi HotStart premixed was diluted two fold before use and directly supplemented with 5 µL of dNTP to double the concentration. 25 mM $MgCl_2$ concentrations were already included in all the buffers. *** DNA concentrations were made up to be between 1 ng·µL^{-1} - 100 ng·µL^{-1}.

Table 5: **Thermocycling profile**

Step	Temperature	Time	Number of cycles
Initial denaturation	95 °C	1-3 min	1
Denaturation	95 °C	30 s	25 - 30
Annealing*	Tm - 5°C	30 s	25 - 30
Extension**	72 °C	1 min/kb	25 - 30
Final Extension	72 °C	5-15 min	1

*The melting temperature (Tm) is primer dependant, temperatures used were usually 5 °C under predicted annealing temperature. **Polymerase extension times are enzyme specific and should be adapted according to manufacture specifications.

2.4.4 Site directed mutagenesis

Primer design for QuickChange™ site directed mutagenesis was done according to the method described by Zheng *et al.*[143]. The PCR was optimised to allow the amplification of very large amplicons which were treated as discussed in section 3.2.3.

2.4.5 DNA truncation method

The full method and discussion can be found in section 3.2 and 3.2.2.

2.4.6 DNA cloning *gccH*

pRV613:LacZ plasmid stock was purified using the GeneJET Plasmid Miniprep Kit (ThermoFisher Scientific, Waltham, MA USA) and double digested with *BamHI*/*Xba*I (NEB, Ipswich, MA USA) overnight. The reactions were checked on a 1% agarose gel stained with EtBr (2.2.4.1) and the appropriate gel bands were excised and purified using the GeneJET gel extraction kit (ThermoFisher Scientific, Waltham, MA USA).

The *gccH* gene was amplified from *L. plantarum* KW30 genomic DNA using PCR. The primers contained *Bam*HI and *Xba*R restriction sites for directional cloning.

Table 6: ***gcc*H amplification primers**

Primer name	Primer sequence	Restriction site
GccH_BamHI_EM	GATGGATCCTAACTAATATTTGTATTATTCCTTCTC	*Bam*HI
GccH_XbaR_EM	GTACGTCTAGACTAAGCATCATATTGAAATAATTT	*Xba*R

The amplified gene was ligated into the pRV613 plasmid at molar ratio of 1:3 with T4 ligase (NEB) at 4°C overnight. The next day, EC100 cells were transformed with pRV613:GccH ligation mixture using heatshock (2.3.4), then spread on LB erythromycin plates and incubated overnight at 37°C. Colonies were grown in erythromycin supplemented LB media for 8 hours and then divided. One half was stored as glycerol stocks and the other half was lysed and the plasmid DNA purified. The purified DNA was double digested with *BamHI*/*Xba*I for comparison with the linear pRV613 stocks and amplified

gccH gene stock. A small amount of the successful culture was revived from glycerol stock by growing the cells to the static phase overnight (8 - 14 hours) in the presence of the appropriate antibiotics. The cells were pelleted by centrifugation at 5000 x g, lysed and the plasmid DNA was purified and sent to the Massey Genome Service for DNA sequencing to confirm the identity of the insert. Confirmed pRV613:GccH stocks were used to transform *L. plantarum* NC8 KO mutants.

2.5 Data collection and handling

2.5.1 Cell growth measurements

2.5.1.1 General optical density measurements

Bacterial culture optical density (OD_{600}) was measured at 600 nm using either the Bio-Rad Smartspec™ Plus spectrophotometer (Bio-Rad Laboratories; USA) or a 12 cuvette Cary 300 UV-Visible spectrophotometer (Agilent Technologies Inc; Santa Clara, CA, USA). Bacterial cultures were routinely diluted in sterile media so that measurements were between OD_{600} 0.1 and 0.5.

2.5.1.2 Titre plate optical density measurements

Titre plates were used to gather cell growth data for up to 19 hours. Overnight culture concentrations were first measured using either the Bio-Rad Smartspec™ plus spectrophotometer (Bio-Rad Laboratories; USA) or a 12 cuvette Cary 300 UV-Visible spectrophotometer (Agilent Technologies Inc; Santa Clara, CA, USA) before serial dilution to OD_{600} of about 0.1 in 1.2 mL of the appropriate MRS media for testing in a Eppendorf tube (Eppendorf®, 022363204 (USA)) and mixed. The solutions of cells were then loaded into ethanol sterilised 96 well titre plates (Greiner Bio-One International: 655801) that were air dried by lamina flow under UV exposure. The sterile plates containing cell solutions were measured using a Microplate Spectrophotometer (Multiskan™ GO Microplate Spectrophotometer, Thermo Scientific™) in a 15 minute kinetic loop for 16 - 19 hours at 35°C.

2.5.2 R scripts and statistics

2.5.2.1 Dealing with data sheets

The Microplate Spectrophotometer (Multiskan™ GO Microplate Spectrophotometer, Thermo Scientific™) generated excel spreadsheets containing the recorded data as columns with the well identification number as the first cell of each column that represents the well number on the titre plate. All other information above the identification number was deleted. A column called "Time" was created and a formula was introduced to convert the time measurements that were in minutes to hours.

A plate map containing three columns titled "Well", "Strain", and "Environment" was created to track the well numbers ("Well") and the corresponding treatment or media ("Environment") along with what cells were added (Strain).

2.5.2.2 R packages and environment

The package **reshape** and **dplyr** are needed to "melt" and re-map the data frames so that they are less wide and are able to be used in R. The Dplyr package may cause a mask "global" stats warning, this does not interfere with the data output. Loading **ggplot2** enables better plotting, nicer colours and easier coding.

library(reshape2)
library(dplyr)
library(ggplot2)

2.5.2.3 Data conversion

This set of code used "reshape" and "dplyr" environments to reshape the spreadsheets containing the OD^{600} readings from 96 well titre plates, which were then matched to a predefined "platemap" so that the readings could be assigned to grouping variables that represent the experimental parameters.

*reshaped <- melt(**"Spreadsheet name"**, id=c("Time"),*

variable.name="Well")

value.name="OD600")

*annotated <- inner_join(reshaped, **"Platemap name"**, by="Well")*

Saves annotated data

write.csv(annotated, "data-annotated.csv")

2.5.2.4 Statistical functions

The statistical functions were all loaded into the R environment as follows.

The 95% confidence interval.

conf_int95 <- function(data) { n <- length(data)

*error <- qt(0.975, df=n-1) * sd(data)/sqrt(n) return(error) }*

Statistical data frame, summary, and standard error.

The file name myData is a place holder used to aggregate the data stored in the file "annotated" as a function of the OD_{600} readings by strain and environment.

myData <- aggregate(annotated$OD600,

by = list(strain = annotated$Strain, environment = annotated$Environment),

FUN = function(x) c(mean = mean(x), sd = sd(x),

n = length(x)))

myData <- do.call(data.frame, myData)

myData$se <- myData$x.sd / sqrt(myData$x.n)

colnames(myData) <- c("Strain", "Environment", "mean", "sd", "n", "se")

par(mar = c(5, 6, 4, 5) + 0.1)

2.5.2.5 Plotting growth curves

stats <- annotated %>%

group_by(Environment, Strain, Time) %>%

summarise(N=length(OD600),

Average=mean(OD600),

CI95=conf_int95(OD600)) %>%

filter(!is.na(Strain))

Group the data by the different experimental variables and calculate the sample size, average OD_{600}, and 95% confidence limits around the mean and among the replicates. Also remove all records where the Strain is NA.

Plotting growth curves by strain using a 95% confidence interval

theme(plot.title = element_text(hjust = -0.45, vjust=2.12)))

ggplot(data= stats, aes(x=Time, y=Average, color= Environment)) +

geom_ribbon(aes(ymin= Average-CI95, ymax= Average+CI95, fill=Environment),

color=NA, alpha=0.3) +

geom_line() + theme_bw()+

#scale_y_log10() **This function is needed to change the scale to $\log_{10}$.**

facet_grid(Strain ~ .) +

labs(x="Time (Hours)", y=" Absorbance at 600 nm ") +

theme(axis.text=element_text(size=16), strip.text=element_text(size=12),+

,legend.title=element_text(size=16), axis.title=element_text(size=16,face="bold")) +

theme(legend.text=element_text(size=16)) +

*ggtitle("**Enter title here**")+ theme(plot.title = element_text(size = 20, face = "bold"))*

Plotting growth curves by environment using a 95% confidence interval

ggplot(data=stats, aes(x=Time, y=Average, color=Strain)) +

geom_ribbon(aes(ymin=Average-CI95, ymax=Average+CI95, fill=Strain),

color=NA, alpha=0.3) +

geom_line() +

scale_y_log10() + theme_bw()+

 # This line adds a dotted line at the set intercept in this case "8"

"geom_vline(xintercept=8, linetype="dashed", color = "red")+"

facet_grid(Environment ∼.) +

labs(x="Time (Hours)", y="Absorbance at 600 nm") +

theme(axis.text=element_text(size=16), strip.text=element_text(size=12),+

,legend.title=element_text(size=16), axis.title=element_text(size=16,face="bold")) +

*theme(legend.text=element_text(size=16)) +ggtitle("**Title here**")+*

theme(plot.title = element_text(size = 20, face = "bold"))

2.5.2.6 Bar charts

limits <- aes(ymax = myData$mean + myData$se,

ymin = myData$mean - myData$se)

p <- ggplot(data = myData, aes(x = factor(Environment), y = mean,

fill = factor(Strain)))

p + geom_bar(stat = "identity",

position = position_dodge(0.9)) + theme_bw()+

geom_errorbar(limits, position = position_dodge(0.9),

width = 0.25) + theme(axis.text=element_text(size=16), strip.text=element_text(size=16),

,legend.title=element_text(size=16),

axis.title=element_text(size=16,face="bold")) +

labs(x = "Environment", y = "OD600") + theme(legend.text=element_text(size=16)) +

*ggtitle("**Enter title here**") + theme(plot.title = element_text(size = 20, face = "bold")) +*

scale_fill_discrete(name = "Strains")

2.5.2.7 Code sources

The code and methods were taken from open source blog posts. Growth curves were adapted from Brian Connelly's "analyzing microbial growth with R 2014"[144], while the bar charts were adapted from Chris Wetherill's "Building Barplots with Error Bars"[145]. Both sources have excellent annotation and should be used for empirical troubleshooting, plate_maps and examples.

Results and discussion

3.1 PTS18CBA protein properties and structure

3.1.1 Primary structure and predicted protein parameters

To ascertain the most appropriate regions to cut in the truncation analysis it was necessary to bioinformatically model the structural characteristics of PTS18CBA. Outside domain order and the use of conserved phospho-carrier residues there is no sequence or structural homology between PTS transports. The following bioinformatic analysis was done to gain an understanding of the PTS^{GlcNAc} so that practical experiments could be planned and carried out. The *pts18cba* nucleic acid sequence from *L. plantarum* 8014 was aligned against *pts18cba* sequences obtained from *L. plantarum* NC8 and *L. plantarum* subsp. ATCC 14917. T-COFFEE alignments produced a score of 99 with excellent alignment across all the sequences[146–148]. A Clustal Omega alignment confirmed the T-Coffee results showing there was little difference between the sequences except for a few single-nucleotide polymorphisms (SNPs) that were not consistent between all three sequences (Appendix 5.2.4)[147–149]. The PTS18CBA amino acid sequence was generated by *in silico* reverse translation of the nucleic acid sequences downloaded from NCBI. Only one of the six possible reading frames produced an uninterrupted protein sequence that was retrospectively confirmed by sequencing the entire *pts18cba* gene from plasmid pRV613:*pts18cba-Flag* made by K. Drower[67].

The amino acid sequence was then analysed using two programs from the ExPASy website, ProtParam[150,151] (5.3.2) and Compute pI/Mw[151–153], which predicted that the gene

product contains 662 amino acids, has a mass of 70.32 kDa, and a pI of ≈ pH 6.4. The protein contains 44 negatively charged (Asp + Glu), 40 positively charged (Arg + Lys) amino acids, and has a stability index of 28.20, suggesting that it is stable. The aliphatic index for PTS18CBA was calculated to be 106.71, predicting that the protein is thermostable. A grand average of hydrophobicity (GRAVY) was also calculated for PTS18CBA and found to be 0.540 indicating that more than > 50% of the amino acids were hydrophobic. Interestingly there is also only one cysteine residue (Cys 429) in this large protein which is somewhat unusual for a protein of this size. The mathematical occurrence for a two condon amino acid is 3.28%[154,155]. The relative frequency of cysteines in prokaryotic proteins has been found to average 1.03% - 1.13% of total protein composition[155]. A minimum of 6.82 cysteines should therefore be expected for a protein containing 662 amino acids. Cys 429 is conserved among EIIB domains found in PEP-PTS transporters and thought to be a phosphorylation site along with His 584, and may function in a phosphorelay system associated with this class of protein.

3.1.2 Secondary structure of PTS18CBA by *in silico* analysis

Regardless of the developments in the reliability of predictive software, the availability of only a few structures make modelling difficult. The differences between predictive software packages, the algorithms they use, and their varying dependence on pre-existing databases, will generate variation in the structural models they produce. Multiple bioinformatics software packages were used to combat this innate difference, and to produce the most reliable prediction of the secondary structure of this molecule. Predicting the presence of secondary structure is necessary to allow informed decisions about modifying or removing blocks of residues from EII^{GlcNAc} without compromising the folding of the resulting protein.

As this is an integral membrane protein, TMpred[156] was used first to predict all possible transmembrane secondary structures. It uses the amino acid hydrophobicity score and respective orientation in the membrane to predict helices without threading or modelling the sequence to known structures. The results predicted 14 inside to outside, and 13 outside

to inside α-helices across the entire sequence, with no transmembrane β-strands (Appendix 5.3.3.1). The hydrophobicity plot suggested that the α-helices were more likely to be located near the N-terminal domain with the likelihood of any transmembrane helices near the C-terminus being low (5.3.3.1 C). The TOPCONS[157] web server provided additional support for these predictions. Using the default settings produced a structural model in which the N-terminal half of the protein formed α-helices that are likely to be able to interact with membrane lipids (Appendix 5.3.3.2A). The suite of packages used during a TOPCONS analysis showed a strong consensus for the likelihood of only 12 transmembrane α-helices opposed to the 13 - 14 helices predicted by TMpred. All but one of the TOPCONS algorithms predicted that the C-terminal region was most likely to be solvent accessible and orientated towards the cytosolic side of the membrane. In summary TMpred and TOPCONS predicted that the N-terminal half of the protein is likely to contain 12 - 14 transmembrane α-helices and that the C-terminal domains are cytosolic.

To gather more information about the other secondary structural elements, and provide further support for previous predictions made by TMpred, TOPCONS, MINNOU[158], Jpred[159], Phyre^2[160] and I-TASSER[161–163] analyses were also carried out using default settings (5.3.3.3 - 5.3.3.5). These all indicated that there are likely to be three highly ordered regions, separated by two flexible linking regions. The N-terminal region is predicted to contain 12 - 14 α-helices that are for the most part lipid accessible. This transmembrane region encompasses nearly half of the protein sequence, and starts roughly at residue 9 and ending at residue 366. This region is followed by a 37 residue long flexible region that is predicted to have little to no structure. The second domain is predicted to consist of $\sim$70 residues between 404 and 472, which contains a mixture of β-strands and α-helices that are likely to be cytosolic. Domain two is separated from the putative cytosolic third domain by a solvent accessible linker region of approximately 31 residues. The third domain is dominated by β-strands and a few small helices.

All the bioinformatic analyses showed a good alignment of the secondary structural elements and identified three structured domains linked by solvent accessible linkers. In

an attempt to characterise these domains, a NCBI conserved domain database search was conducted, and cross-referenced with Pfam to try to identify the domain structure (5.3.3.2 B). The three structures were identified to contain conserved domains of enzyme II belonging to the PEP-PTS superfamily. The domains were identified as EIIC, EIIB and EIIA respectively and their approximate domain boundaries are shown in Table 7 and Figure 13.

Table 7: **Domain boundary summarised by predictive web server**

Web server	EIIC	Linker 1	EIIB	Linker 2	EIIA
Jpred	3 - 365	366 - 403	404 - 472	473 - 504	505 - 662
MINNOU	2 - 368	369 - 404	405 - 483	484 - 514	515 - 662
NCBI	9 - 311	NA	411 - 444	NA	515 - 635
Phyre	3 - 370	371 - 405	406 - 485	486 - 513	514 - 660
I-TASSER	2 - 369	370 - 405	406 - 484	485 - 513	514 - 660

The regions represented above are all amino acid positions in a single PTS18CBA polypeptide. The domain boundaries represented by each web server were determined using the default settings and only the algorithms provided by each independent system and were not user trained.

3.1.3 Tertiary structure of PTS18CBA

Predictive protein 3D modelling is done by threading the amino acids sequence onto the structures of homologous proteins. Two attempts, separated by a year, were made to characterise the PTS18CBA sequence further and build predictive models that could inform experimental design using the I-TASSER and Phyre^2 web servers. Both systems identified the same EIIC, EIIB, EIIA conserved domains, but did not use the same homologous structures for threading templates. Both web servers identified that the EIIA domain had a highly stable β-barrel sandwich hybrid fold found in the EIIA domain of glucose permeases (PDB 2GPR)[164]. This protein is cytosolic and its fold appears to be reasonably well conserved among EIIA domains of other PTS transporters. The PTS18CBA EIIB fold was predicted to have a structure most similar to that of a homing endonuclease fold found in the glucose permease domain IIB (PDB 3BP8)[165]. The fold of the EIIC domain was initially found to be most similar to that of the EIIC of the N,N′-diacetylchitobiose-

specific phosphotransferase system (PDB 3QNQ)[131] by both web servers during the first attempt and the maltose-specific PTS EIIC^{Malt} domain (PDB 5IWS)[127] during the second attempt, which had a higher sequence similarity to the EIIC of PTS18CBA.

However, during the first threading attempt using Phyre^2, an unknown error occurred and the system was unable to thread the linker region between the EIIA and EIIB domains causing only the EIIA to be displayed in the resulting model (Figure 12B). I-TASSER has different tolerances to Phyre^2. Instead of using the homologous PTS structures without the linker regions, I-TASSER used the Archaeoglobus fulgidus oligosaccharyltransferase (PDB 3WAJ)[166] that included linker regions. Threading proceeded without errors and successfully generated a model of the PTS18CBA threaded onto this alternative transferase structure albeit with a low C-score (which is a measure of reliability of the protein prediction, -5 to 2) (Figure 12D).

The modelling exercise was repeated and a second attempt was made to thread new predictive models using both Phyre^2 and I-TASSER. During this attempt Phyre^2 successfully threaded the entire sequence to the homologous crystal structures by omitting the linker regions from the threading process which produced a model that used 93% of the residues threaded with >90% confidence (Figure 12A). According to the reliability scores used by Phyre^2, this is the most accurate model generated by either server. Interestingly, I-TASSER produced a model identical to the initial one (Figure 12C and D). The inability of I-TASSER to thread smaller structures is likely due to the LOMETS prediction, the first step in the process, which predicted similar secondary structural elements for both the PTS18CBA and the oligosaccharyltransferase sequences. Regardless of the different algorithms used by these two systems, good agreement regarding the homologous structure identification and the domain order was found. Because the homologous proteins that were identified during this process are not expressed as a single polypeptide chain like the PTS18CBA it is likely that the PTS18CBA domains are stable and fold independently of each other.

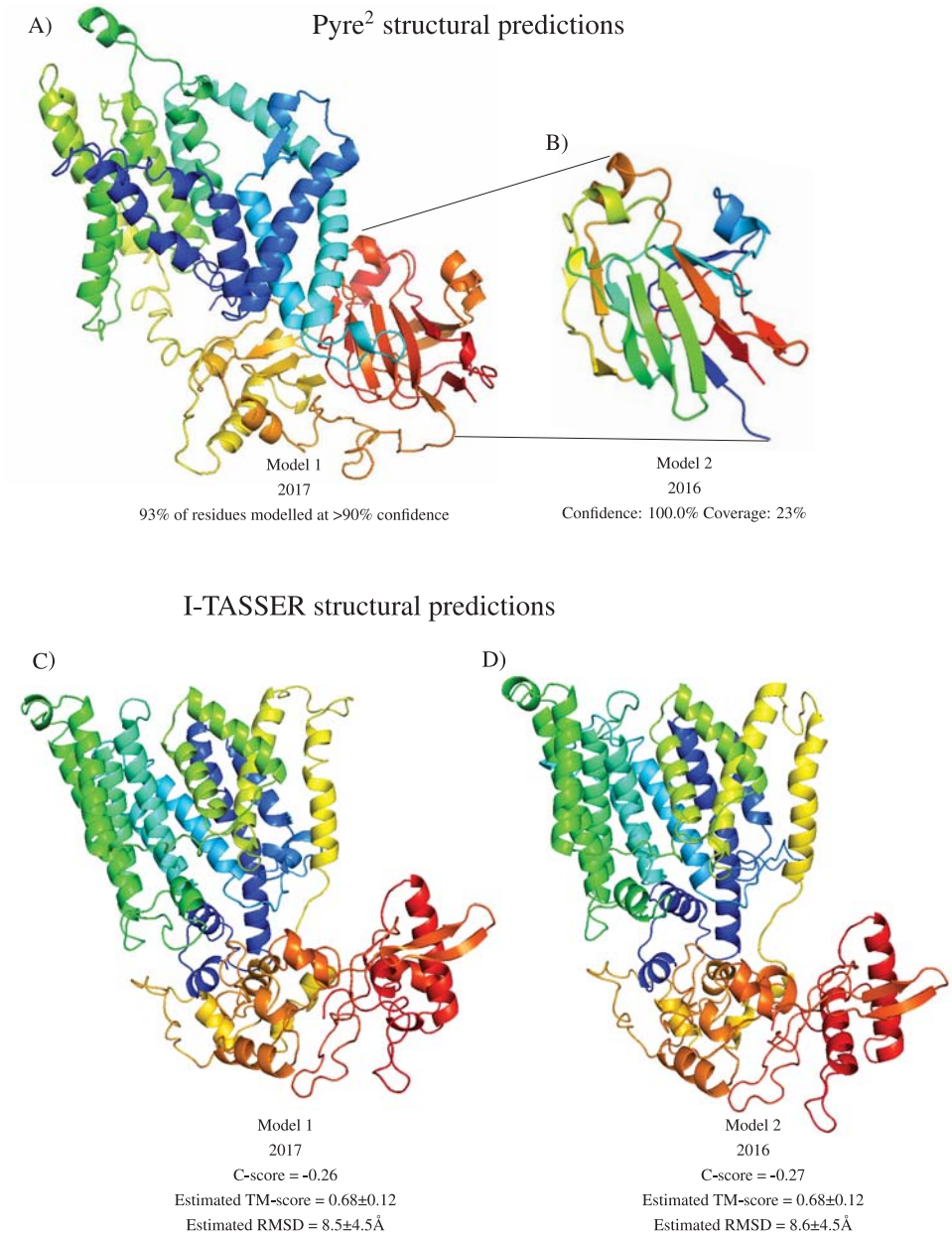


Figure 12: **Phyre^2 and I-TASSER tertiary predictions of the pts18cba structure**
(A) Phyre^2 Model 1: predicted in 2017 threaded 93% of the residues onto known homologous crystal structures with an improved degree of accuracy compared to earlier models. **(B)** Phyre^2 Model 2: predicted in 2016 failed to thread the amino acids sequence beyond the EIIA domain, due to an unknown error when dealing with unstructured regions that do not match the templates like the linker regions. **(C)** I-TASSER Model 1: predicted in 2017 was threaded against the same oligosaccharyltransferase as the 2016 model. **(D)** I-TASSER Model 2: was processed and threaded by I-TASSER in 2016 and was the first model to be used to inform the experimental design for the truncation analysis. C-score is a confidence score for estimating the quality of predicted models by I-TASSER with a range of -5 to 2 (higher scores are better). Both estimated TM-score and Estimated RMSD scores are based on the C-score and are known standards for measuring the structural similarity between two structures. All models coloured from N-terminus (Blue) to C-terminus (Red) and orientated in the frontal (coronal) plane with the EIIC domain at the top and EIIA at the bottom.

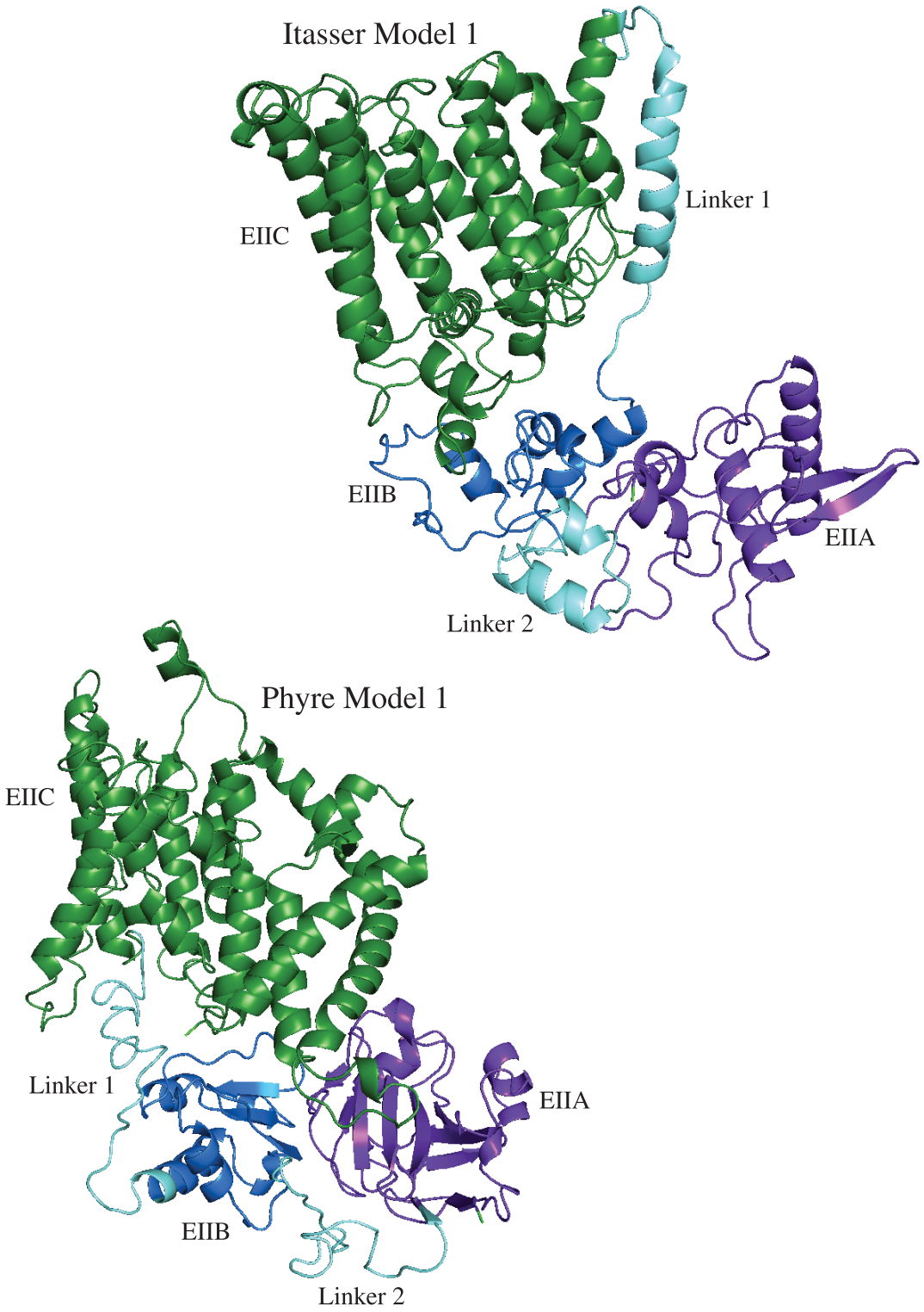


Figure 13: **PTS18CBA domain order displayed as I-TASSER and Phyre[2] models** **(Top)** I-TASSER Model 1 coloured to show the domains and linker regions. **(Bottom)** Phyre[2] Model 1 coloured to show the domains and linker regions. All models were arbitrarily coloured to highlight the domains predicted by previous structural analyses (Table 7): EIIC (Green), EIIB (Dark Blue), EIIA (Deep Purple), linkers (Aqua). The models are vertically orientated with the EIIC up and viewed in the frontal (coronal) plane.

3.1.4 EIIC^{Malt} homodimer and membrane interaction

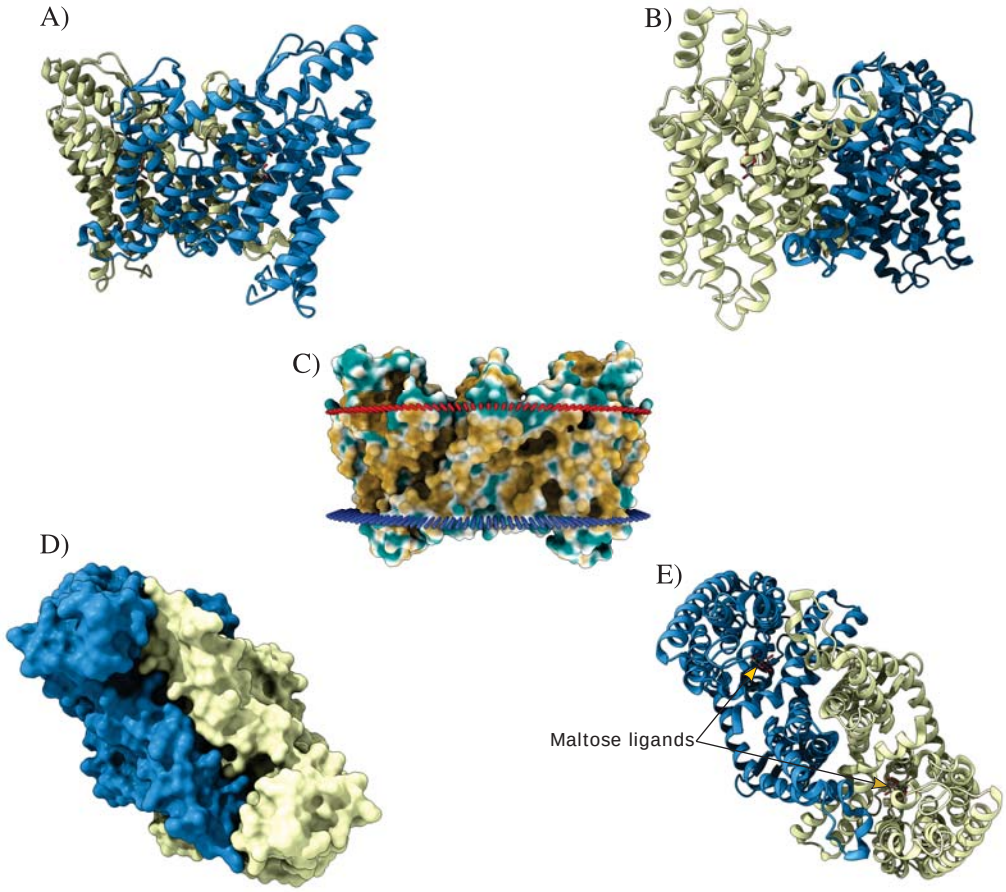


Figure 14: **EIIC^{Malt} homodimer and membrane interaction**
(A) Cartoon showing a horizontal view of the PTS EIIC^{Malt} chain A (Blue) and chain B (Cream) presented as a cartoon. **(B)** 180° lateral rotation of the PTS EIIC^{Malt} homodimer. **(C)** Horizontal front on view of the PTS EIIC^{Malt} homodimer, viewed along the plane of the membrane and coloured to show the hydrophobic (Gold), hydrophilic (Blue) and amphipathic (White) surface area. The membrane interface is represented as rods, the extra-cellular membrane interface (Red), and the cytosolic membrane interface (Blue). **(D)** Space filled model of chain A (Blue) and chain B (Cream) as a biological homodimer, viewed from the extracellular side. **(E)** The top-down extracellular view of the dimer as a cartoon. Each monomer appears to contain a binding pocket for an individual nested centred maltose ligand (Gold headed arrows). The images were downloaded from RCSB PDB: 5IWS[127].

The recent publication of the EIIC^{Malt} crystal structure from *Bacillus cereus* was found to have some homology to the EIIC^{GlcNAc} and also provided some additional information regarding how these systems dimerise and function. The glucose PTS (PTS^{Glc}), PTS^{GlcNAc}, and PTS^{MalT} are all considered to be in the PTS Glucose-Glucoside (Glc) Family (TC 4.A.1) of bacterial PTS transporters (NCBI: cl21492) which function as membrane-bound

homodimers with varying levels of domain linkage. The PTS^{GlcNAc} transporter is transcribed as single polypeptide chain, in contrast to the PTS^{Glc} that is expressed as three individual peptides, while PTS^{MalT} contains a linker region that joins the EIIC and EIIB domains. The linkers are believed to increase the efficiency of the PTS transporter by increasing the local concentration of the respective domains[111,167], but no evidence could be found regarding their involvement in any other function. Considering that the PTS^{Glc} operates by folding its domains as independent proteins and the $EIIC^{Malt}$ was biochemically characterised and crystallised after removing the $EIIB^{Malt}$ domain using trypsin. These events suggest that the domains are stable when separated and are likely folding independently as well.

Interestingly there appears to be little change in the substrate specificity and activity of truncated $EIIC^{Malt}$ when compared to the full-length protein, furthermore the truncated $EIIC^{Malt}$ was also able to dimerise[127]. Closer examination of the structure showed that it has two functional domains, a dimerisation domain that contains a buried interface with a surface area of $\sim$2,700 $Å^2$ and a substrate binding domain[127]. When dimerised the substrate binding sites are in close proximity to the binding domains, however, all the residues interacting with the substrate are conserved to the substrate binding domain of the EIIC and are not involved in dimerisation domain[127]. Despite limited information regarding any additional duties of the linker regions between the domains, they are the most rational points to target for domain removal in the truncation analysis.

This analysis found that the PTS^{GlcNAc} contained three stable independent domains linked by flexible polypepides. In addition to domain order and stability, the conserved phospho-carrying residues were found to be His584 and Cys429. These data were used to inform experimental designs aimed at removing whole domains without disrupting the remaining protein. These bioinformatic studies also informed the choice of residues to disrupt the phosphorelay system of this PTS.

3.2 Construct development

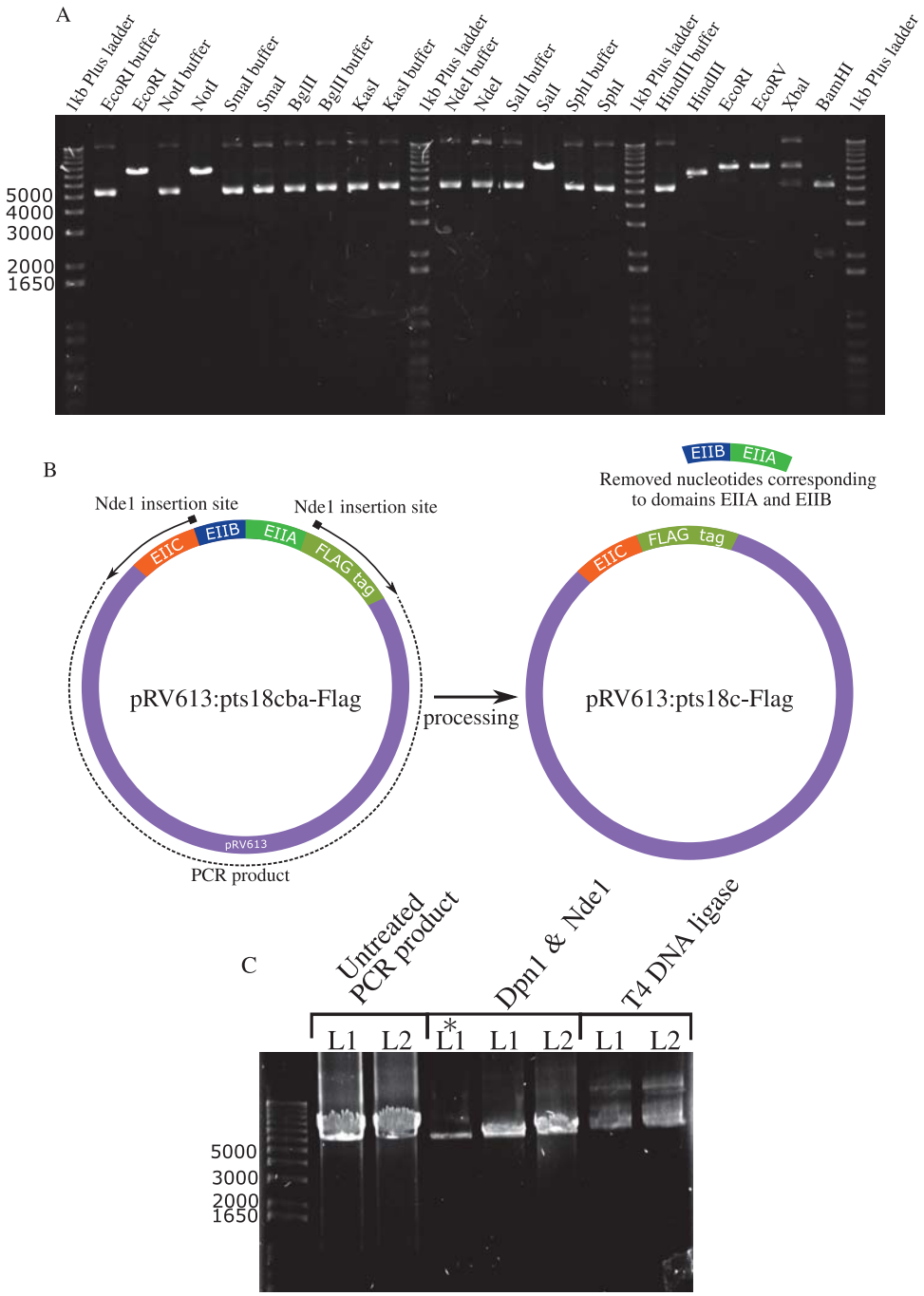


Figure 15: **pRV613-*pts18cba-FLAG* restriction digest map and long PCR diagram**
(A) 1% agarose gel of a restriction digest of pRV613-pts18cba-FLAG using only buffer or enzyme and buffer. 7 enzymes cut the vector: *Eco*RI, *Eco*RV, *Sal*I, *Not*I, *Xba*l, *Hind*III and *BamH*I. 5 enzymes that did not cut the vector: *Sma*I, *Bgl*II, *Kas*I, *Nde*I, and *Sph*I. **(B)** Schematic of the whole plasmid PCR strategy that removed the domains from *pts18cba*. The fragments are re-circularised by the restriction sites introduced on the primers. **(C)** Untreated PCR products: 5 µL samples from PCR. L1: *Nde*1 site introduced into linker 1 removing EIIA and EIIB, L2: restriction site introduced into linker 2 removing EIIA. *Dpn*1 and *Nde*1 endonuclease: treatments post heat inactivation and PCR clean up. T4 DNA Ligase: results post clean up #2 and the successfully re-circularised plasmid using T4 DNA ligase. *The sample was diluted 3 fold before loading.

The EIIC domain of PTS18CBA should be able to fold and dimerise without the aid of either the EIIA or EIIB domains. There is also little sequence homology or substrate specificity conserved between the various PTS transporters within this protein superfamily making the functional rescue of truncated PTS18CBA protein unlikely. The domains of PTS18CBA are expressed together as an individual polypeptide chain allowing the removal of unwanted regions, without the need for multiple plasmids or complex vectors. As part of this work, the polymerase chain reaction was used to remove whole domains of the PTS18CBA protein and change specific residues to test their roles in the mechanism of action of GccF (Figure 15).

Table 8: **Primers used for construct development**

#	Primer name	Primer Sequence
1	L1.2-pts18CB	TCCGCATGCCATATGCTTTGCTTGGCGCATATACTTATC
2	L2-pts18BA	TCCGCATGCCATATGGGCCTTTGATTCAGCATCAGA
3	L3-pts18-FLAG	GGAGCATGCCATATGGACTACAAAGACGACGACGAC
4	F-Cys:Ala-pts18cba	CAATGCTACGACGCGGCTGC
5	R-Cys:Ala-pts18cba	CGCGTCGTAGCATTGTTGATCACAC
6	F-His:Ala-pts18cba	CTTATTAGCCATGGGGATCAATACCGTG
7	R-His:Ala-pts18cba	CCCATGGCTAATAAGACTTCTAAACCACTCG

Primers used for construct development. L3-pts18-FLAG is a universal primer that can be paired with both L1 and L2 to produce truncation constructs. Cys:Ala and His:Ala are specifically paired for site directed mutagenesis.

3.2.1 PCR troubleshooting

Initial PCR experiments were done using Sigma-Aldrich Taq DNA Polymerase (TAQ-RO ROCHE) or BIOTAQ^{TM} DNA Polymerases. Despite thorough troubleshooting these polymerases failed to produce any PCR products using the pRV613-*pts18cba-FLAG* template DNA and the appropriate primers despite the positive controls (OPH6 and pET) showing the system was functional (Figure 16A and G). To rule out template inhibition, a series of template concentrations of positive control template DNA showed that the reaction mix

was tolerant of DNA concentrations exceeding 9 ng·μL^{-1}. However, 1 - 2 ng·μL^{-1} of DNA was optimal, and was used for future experiments.

A possible reason for the lack of product could have been poor pRV613:*pts18cba-FLAG* primer design. New sequencing primers designed to bind within the *pts18cba* gene and function at the same temperature were used to check the forward primer function using the same PCR protocol as before. This approach generated smaller fragments that were consistent with the expected PCR product sizes of a truncated *pts18cba* gene, showing that the forward primer was functioning as intended. This method was also used to develop a PCR method to test *L. plantarum* cells for successful transformations (Figure 17, page 70). The sequence of the pRV613:*pts18cba-FLAG* plasmid was obtained using the reverse primer showed that this primer also bound the vector.

The same tests conducted using Platinum™ Taq DNA Polymerase High Fidelity (ThermoFisher #: 11304011), Phusion® High-Fidelity DNA Polymerase (NEB #: M0530S), and Q5® High-Fidelity DNA Polymerase (NEB #: M0491S) generated a single high-intensity band for the positive control, but only multiple low-intensity bands and background degradation with pRV613-*pts18cba-FLAG* DNA (Figure 16B).

To rule out template quality and concentration as sources of failure, multiple *E. coli* strains were transformed with the pRV613-*pts18cba-FLAG* stock plasmid. Multiple colonies from each transformation attempt were picked and cultured overnight in 50 mL of media. 500 μL of each of the overnight cultures was used to create glycerol stock for storage at -80 °C, and plasmid DNA was harvested from 10 mL aliquots of remaining culture. *E. coli* EC100 cells produced the best results (Figure 16C), and the plasmid extract was confirmed by analytical digest using *BamH*I (Figure 16D).

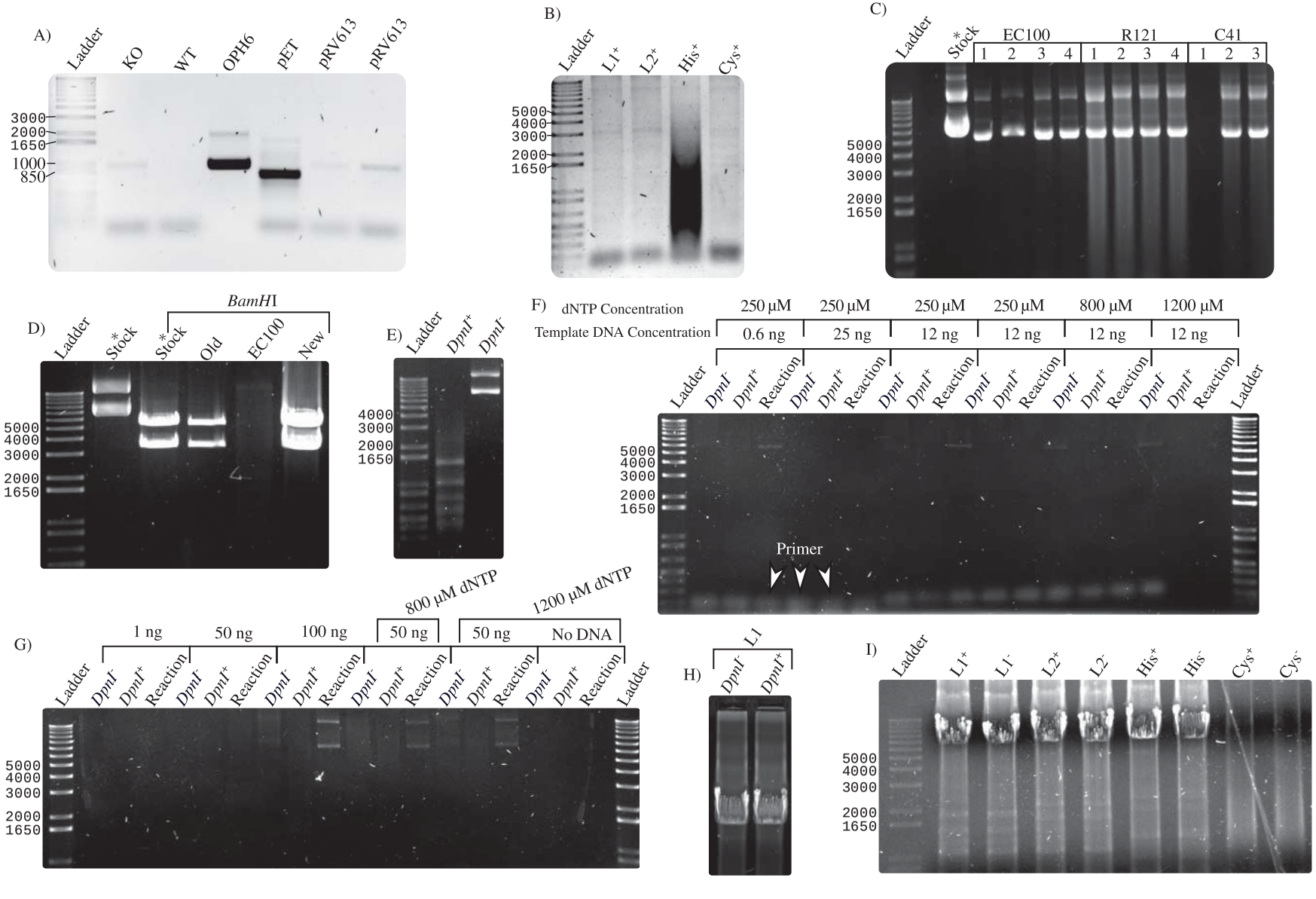

A)
Ladder
KO
WT
OPH6
pET
pRV613
pRV613
3000
2000
1650
1000
850
B)
Ladder
L1+
L2+
His+
Cys+
5000
4000
3000
2000
1650
C)
Ladder
*Stock
EC100
R121
C41
1 2 3 4 1 2 3 4 1 2 3
5000
4000
3000
2000
1650
D)
BamHI
Ladder
*Stock
*Stock
Old
EC100
New
5000
4000
3000
2000
1650
E)
Ladder
DpnI+
DpnI-
4000
3000
2000
1650
F)
dNTP Concentration
250 µM 250 µM 250 µM 250 µM 800 µM 1200 µM
Template DNA Concentration
0.6 ng 25 ng 12 ng 12 ng 12 ng 12 ng
Ladder
DpnI- DpnI+ Reaction
Ladder
5000
4000
3000
2000
1650
Primer
G)
800 µM dNTP
1200 µM dNTP
1 ng 50 ng 100 ng 50 ng 50 ng No DNA
Ladder
DpnI- DpnI+ Reaction
Ladder
5000
4000
3000
2000
1650
H)
L1
DpnI- DpnI+
I)
Ladder
L1+ L1- L2+ L2- His+ His- Cys+ Cys-
5000
4000
3000
2000
1650

Figure 16: **PCR troubleshooting**

(A) Initial attempts to generate large PCR products (>6.5 kb) from pRV613-*pts18cba-flag* template DNA using standard *Taq* polymerase and reaction mix. KO, *L. plantarum* NC8 Δ*pts18cba*; WT, *L. plantarum* NC8 with no modifications (KO and WT do not harbour any plasmids and were used as a negative control). OPH6, a pBluescript plasmid containing a PNGase F insert (PCR product ∼1 kb); pET, control vector (PCR product ∼850 kb) are provided positive controls. pRV613, pRV613-*pts18cba-flag* plasmid with Sigma-Aldrich Taq DNA Polymerase (TAQ-RO ROCHE) in the first lane and BIOTAQ^{TM} DNA Polymerase in the second. **(B)** Background degradation experienced using Phusion® High-Fidelity DNA Polymerase (NEB #: M0530S) with 250µM dNTPs. pRV613-*pts18cba-FLAG* template DNA and primers designed for site directed mutagenesis of His584 (His), Cys429 (Cys) or remove domains via linker region 1 (L1) or linker region 2 (L2) **(C)** *E. coli* strains transformed with pRV613-*pts18cba-FLAG*. Normalised cell volumes were lysed using a commercial kit to find the best pRV613-*pts18cba-FLAG* plasmid producer strain. **(D)** *BamH*I analytical restriction digests of pRV613-*pts18cba-FLAG*. *Stock, original stocks created during previous projects. Old, the plasmid stock used during PCR experiments. EC100, Untransformed EC100 cells exposed to the full plasmid preparation procedure in parallel with EC100 cells newly transformed with pRV613-*pts18cba-FLAG* (New). **(E)** *Dpn*I control experiment using 400 ng·µL^{-1} of newly extracted pRV613-*pts18cba-FLAG*. **(F)** Template concentration trials against standard dNTP concentrations using BIOTAQ^{TM} DNA Polymerase. **(G)** Template trials using BIOTAQ^{TM} DNA Polymerase and literature recommended template DNA concentrations[143]. Total concentrations per 50 µL reactions are shown (100 ng, 2 ng·µL^{-1}) **(H)** 1 ng·µL^{-1} of template using 2 mM of dNTPs and Phusion® High-Fidelity DNA Polymerase (M0530) along with L1 primers. 50 µL reaction was split into two 20 µL aliquots and either treated with *Dnp*I or left untreated. **(I)** The same experiment as in H using KAPA HiFi HotStart Uracil+ ReadyMix (2X) supplement with ∼1.2 mM of dNTPs (5 µL of 10 mM stock per 50 µL reaction). The experiment was repeated for each primer pair design for construct developement and either treated with *Dpn*I of left untreated. All images were 1% agarose gels stained with EtBr post electrophoresis.

Doubling the deoxynucleotide (dNTP) concentration of standard PCR reaction mixes (2.4.3, Table 4) was enough to produce the expected ∼8 kb band reliably, albeit at a low intensity (Figure 16H). It is likely that the high fidelity polymerases generated the background degradation by editing the ends of the large PCR products when dNTP concentrations ran low after 25 - 30 cycles. BIOTAQ^{TM} DNA Polymerase did not amplify template DNA regardless of the dNTP or template concentrations which can be seen by the line of unincorporated primers (Figure 16F).

KAPA HiFi HotStart Uracil+ ReadyMix (2X) (kit code: KK2801) supplemented with additional dNTPs generated the highest yield of the correct product (Figure 16I) and allowed the appropriate diagnostics for each step of PCR and subsequent reactions that needed to be carried out (Figure 15C). All future experiments were done using this optimised method and reaction mix (2.4.3).

3.2.2 Truncation construct development

3.2.2.1 Domain removal strategy

The truncation analysis was based on a series of constructs derived from pRV613:*pts18cba-FLAG* and developed to explore how the removal of EIIA and EIIAB domains would affect the binding and activity of GccF. Template pRV613-pts18cba-FLAG was purified from transformed EC100 cells using the Wizard® Plus SV Minipreps DNA Purification System (Catalog # A1465). PCR was then used to alter the gene insert without making any changes to the surrounding vector by amplifying a linear fragment from 5′ end of the FLAG-tag to the 3′ end of the last included domain (Figure 15B). To ensure that the resulting fragments could be religated with the exclusion of unwanted domain(s) a *Nde*1 restriction site was added to the ends of the primers (Figure 15B). The pRV613-*pts18cba-FLAG* has not been completely sequenced, although sequence has been confirmed for regions immediately surrounding the *pts18cba* insertion site. The sequence of the parent plasmid (pRV610) from which pRV613 was derived is complete and available. Even with a known pRV610 sequence, there was doubt sur-

rounding the multiple cloning sites identified on the pRV610 plasmid, as they may have changed during the construction of pRV613. To find enzymes that met the experimental parameters, a full screen of available restriction enzymes was done (Figure 15A) and mapped back to the pRV610 sequences. *Nde*1 did not digest the pRV613:*pts18cba-FLAG* vector nor did it introduce a frameshift while also having the advantage of being heat inactivated at 65 °C. Additionally, the enzyme was readily available, used the same buffer as *Dpn*1, had no star activity, and could cut well if it was 3 - 5 nucleotides from the end of a DNA fragment. *Sph*I also meets the same conditions but can cut at the end of a DNA molecule, and although less common than *Nde*I, provides an ideal backup restriction site when added to the very ends of the primer sequences. Placing the *Sph*I sequence at the end of the primer provided the necessary nucleotide spacer so that *Nde*1 could cut efficiently, and if *Nde*1 failed, *Sph*I could be used without causing a frameshift but at the cost of having four extra amino acids in the resulting protein opposed to two.

Early PCR experiments generated products of approximately 7 kb and 6.7 kb in length which was consistent with expectations of either an EIIA or double EIIAB deletion respectively. The PCR products were treated with *Dpn*1 and *Nde*1 with the same buffer and incubated at 37 °C to remove the methylated template and to trim the edges off completed PCR products. The reaction mix was purified using a GeneJET PCR Purification Kit (ThermoFisher: K0701). A sample of purified PCR product was recircularized using T4 DNA ligase in the appropriate buffer for two hours at room temperature. Upon gel verification that the oligonucleotides were re-circularised, the plasmids were used to transform *Escherichia coli* strain EC100 by heat shock and stored in 30% glycerol at -80 °C.

3.2.2.2 Removal of EIIA or EIIAB via the linker regions

Plasmid DNA harvested from EC100 cells were digested with *BamH*I, which cuts the pRV613 plasmid twice at sites that flank the gene insertion site allowing any cloned gene to be completely removed with a single enzyme. The insert sizes proved to be consistent with expectations for the truncated *pts18cba* genes and were sent for DNA sequencing to confirm the modifications. Unfortunately, the sequencing results indicated

that many of the transformed EC100 strains had random frameshifts at the position of the *Nde*1 restriction site. The results were consistent when the experiments were repeated, and PCR troubleshooting did little to change the outcome. Changes to *Dpn*I and *Nde*I digestion and subsequent heat inactivation processes appeared to have the largest effect. It is possible that during heat inactivation at 65 °C, without a cleanup step, the polymerase may have been adding nucleotides to the freshly cut ends. Introducing a cleanup step after the PCR reaction was treated with *Dpn*1 to remove the template DNA, and another after the *Nde*1 digests, before the ligation step produced the best results. Unfortunately not all the abnormalities that occurred at the *Nde*1 site were removed and as a result DNA concentrations were much lower than expected. Increasing the PCR reaction volumes was necessary to allow gel verification, but did little to increase the abundance of the final product. Further investigation regarding the origins of these unexpected changes was abandoned due to a lack of time and resources.

Because a population of plasmids containing frameshifts was present regardless of optimisation, 30 - 40 colonies of transformed EC100 cells were used to find experimentally ideal products. The sequencing results showed that a number of the EIIA deletion constructs had no frameshifts, and that the FLAG-tag was within the reading frame. Cell lines containing these truncated plasmids were grown in bulk. An aliquot of the bulk cultures were stored in 30% glycerol at -80 °C while the rest of the plasmids were harvested and stored as construct "L2" at -20 °C.

An exhaustive effort was taken to generate a truncation within linker 1 for the removal of the EIIA and EIIB domains. However, sequencing results showed that all the products had reading frameshifts of varying severity after the *Nde*1 restriction site. Two dominant species were evident; in the first species most of the vectors contained a stop codon nine nucleotides downstream of the expected stop condon changing only the flag sequence. In the second species, the stop codon was far beyond the expected stop codon, running into the multiple cloning site generating a long tail. *In silico* translation of both sequences indicated that the EIIC domains and the remaining linker regions would contain an extra

25 or 37 residues at the C-terminal end of the sequence, but no FLAG-tag. This is 3-16 residues more than the insert would have had if the expected stop codon had been read. Cell lines containing these two constructs were grown in mass, a portion of these cultures were stored in 30% glycerol at -80 °C while the rest of the plasmids were harvested and stored as "L1C3-6" and "2L1".

3.2.3 Site directed mutagenesis

A modified QuickChange™ site-directed and site-saturation mutagenesis protocol[143] was used to replace cysteine 429 and histidine 584 with alanine preventing them from functioning in the phosphorelay. This site-directed mutagenesis (SDM) method makes use of partially overlapping PCR primers to introduce nucleic acids changes to an expression plasmid, that upon induction will generate a protein containing the desired modifications.

The same template DNA, PCR conditions and reaction mix ratios used during the truncation experiments was also used for the SDM, only differing in the primers added. His 584 Ala mutagenesis produced a vector that showed no other changes to the *pts18cba-FLAG* gene insert other than the two nucleic acid changes intended to convert the histidine to an alanine on the first attempt. Cysteine was successfully replaced with alanine, although sequencing showed that none of the 16 transformed EC100 colonies tested contained a useful plasmids.

All attempts to mutate cysteine 429 produced a sequence containing a deletion of 20-90 nucleotides immediately after the mutation site, followed by an insertion of what seemed to be random DNA not related to the *pts18cba* gene or any known organism. Upon closer examination of the inserted DNA it appeared to contain various overlapping and repeating sequences related to the primers used to generate the mutation, which appeared to be annealing randomly downstream of the mutation site resulting in deletions. There is an expectation that while primer dimers can form during SDM, there will always be

a proportion that binds to the template DNA to produce the desired product. Primer dimers that can anneal and interact with regions surrounding the mutation site and downstream elements were completely unexpected and could be related to the sequence of the template. These aberrant modifications were incorporated into plasmids because final PCR products were used to directly transform EC100 cells as non-covalently closed circular DNA with the expectation that the host cells will repair the nicks, making these modifications permanent. Continued efforts into resolving this issue by optimising the PCR strategy, which included increasing the annealing temperature and changing the salt concentrations of the reaction mix had no observed effect on the PCR reactions and did not resolve this issue. The primers were redesigned but due to time constraints further work developing the Cys429Ala and double mutants had to be abandoned.

3.2.4 GccH

GccH was cloned into linear pRV613 that had the *LacZ* insert enzymatically removed, the full method can be found in the materials and methods section 2.4.6.

3.3 Extrachromosomal transformations of *L. plantarum* NC8

Experience within our group and others has shown that transformations of species of *L. plantarum* have either been difficult or bordering upon unachievable. The most favourable results generated only 1-2 colonies out of 3 or more attempts using the *L. plantarum* optimised vector pRV613, which harbours a β-galactosidase reporter gene under the control of a copper-inducible promoter. A modified electroporation method developed within our laboratory group showed improved results[67]. Optimisation of this method, would not only make transformations more convenient but also massively improve the chances of successfully transforming this species of bacteria. To achieve this goal, optimisations were explored using the 8.6 kb pRV613 vector. While these improvements are specifically for *L. plantarum* NC8, understanding which steps are vital may improve the likelihood of

successful extrachromosomal transformations of other *L. plantarum* species.

3.3.1 pRV613 transformation of pts18 deficient *L. plantarum* NC8

To further improve the chances of a successful transformation, *Lactobacillus plantarum* NC8 was selected as a model because it has adequate susceptibility to GccF and had no native plasmids. This *Lactobacillus* species does contain a natural *pts18cba* gene which was removed to generate the knockout mutants (KO) by replacing the *pts18cba* gene with a chloramphenicol (Chl) resistance gene. Furthermore, using EC100 cells as an intermediate step allows the production of high concentration and high quality methylated plasmid preparations, which is vital for making routine transformations of *Lactobacillus plantarum* more reliable. The transformations were done by electroporation using the pRV613 vector or constructs derived from modifications to it to modify *L. plantarum* NC8 wild-type (WT) or a *L. plantarum* NC8 KO strain.

3.3.1.1 Generating competent cell stocks

Generating competent cells was the most influential component in optimising the transformation process and the final method used is outlined in section 2.3.1. Competent cells must be generated fresh on the day before electroporation; some species are tolerant enough to be stored at -80°C, however all attempts using frozen stock failed. When generating fresh competent cells, an overnight culture is usually inoculated and allowed to reach the stationary phase. This overgrown culture is then used to inoculate 50 mL of MRS in a sterile 250 mL flat bottom flask to an OD_{600} of $\sim$0.01-0.05 and then allowed to reach an OD_{600} of $\sim$0.2 with mild shaking at 1000 RPM, which for *L. plantarum* is just before the start of the log phase. The desired OD_{600} was usually reached within $\sim$4 - 5 hours. Early experiments using 10 mL solutions in borosilicate glass had variable results and often failed altogether. Using 50 mL of media in a 250 mL flask increased the reliability of not only transformations using *L. plantarum* cells but also the quality and quantity of plasmid extracted from aerobic bacteria like *E. coli*. The improvements observed when the size of the vessel and volume of the media were increased is probably

due to aeration of the broth and potentially longer logarithmic growth phase.

Selection antibiotics were always added to the media used for overnight cultures for any cells containing resistance markers. However, the addition of selection antibiotics to the first step of generating competent cells significantly impacted the growth rate of the cells and did not seem to change the rate of false positives occurring on agar plates after electroporation. Using small amounts of ampicillin during this initial step is, however, routine and is believed to create small non-fatal holes in the membranes of the cells to improve the transfer of DNA across the membrane. As the addition of ampicillin did not seem to influence the success rate of *L. plantarum* NC8 transformation by electroporation, this practice was stopped.

Chilling bacteria before the washing phase is an important step[168], despite the fact that a recent report showed that leaving bacteria at room temperature improved efficiency[169]. Although room temperature washing and transformations were not done during this work, it was observed that once the cells had been cooled, their competency appeared to drop by approximately a log fold for every hour the cells were left cold. It was also observed that *L. plantarum* 8014 produces loose pellets and will need to be centrifuged for 15 - 20 minutes while *L. plantarum* NC8 and *E. coli* produce tight pellets after 5 - 10-minute centrifugation cycles. Optimising the centrifuge cycles can significantly impact success by reducing the time cells remain chilled.

The most significant change was with the use of fresh polyethylene glycol 1500 (PEG 1500). No successful transformation attempts were recorded with stocks older than a week, while the most successful attempts were experienced when the PEG1500 was made on the day.

3.3.1.2 Electroporation

Electroporation was carried out exactly as outlined in section 2.3.2 and is unchanged from the original method. Standard electroporation practices insist that smaller DNA concen-

trations generate the best results, which is most accurate when electroporating or heat-shocking DNA into *E.coli*, However, the best results when transforming *L. plantarum* NC8 were observed using 1 - 5 µg of DNA. Using large volumes of untreated plasmid preparations would result in arcing. Drop dialysis was commonly used to remove additional salt form DNA preparations, However, this resulted in dilution of very salty DNA stocks. Drop dialysis and arcing were both avoided by using a small volume of very concentrated DNA.

Time constants around 7 - 8 ms were common during successful *L. plantarum* NC8 transformations, and post electroporation treatment of the cells proved to be very important. When cells were recovered at 37°C immediately after electroporation transformations were unsuccessful. Temperature was also very important for the growth of transformed cells on plates. When recovery plates were incubated at 37 °C, the premature development of untransformed satellite colonies was observed causing a larger number of false positives and in some cases overwhelmed the plate altogether.

3.3.1.3 Methods of confirming transformations in *L. plantarum*

Confirming transformation success is a multi-step process that starts with inspecting selection plates for colonies of cells (Figures 23A and 24A). This first step is usually followed by physically confirming the presence of plasmid DNA in the cells along with its identity either by analytical restriction digestion or sequencing. Although restriction digests are robust, and do not need specialised equipment, analytical digests of pRV613 plasmid proved problematic when extracted from *L. plantarum* NC8 due to the low copy number origin of replication for the plasmid which lead to low plasmid extraction yields. As a result analytical restriction digests of DNA from *L. plantarum* were typically inconclusive and DNA concentrations were often close to the limit of detection. Bacterial lysis methods also clouded detection even further by introducing background contamination from sheared DNA (Figure 17A). The source of this shearing was thought to be due to the aggressive chemical and physical lysis methods used to extract DNA from Firmicutes. Control lanes were created by using pRV613 plasmid that was extracted from the ampli-

fication strain *E.coli* EC100 to produce plasmid stock with concentrations ~450 ng·μL^{-1}. 100 ng of the DNA stock was digested with either *Xba*I which cuts the plasmids once or with *BamH*I which cuts the plasmid twice, once on either side of the gene insertion site. The two bands generated by the *BamH*I digest were ~4.9 kb (remaining plasmid) and ~3.6 kb (LacZ gene insert), which are additively close to the expected 8.69 kb of the pRV613 plasmid (Figure 17A lane 3). Analytical digests were used to test 14 transformed *L. plantarum* NC8 KO colonies that were cultured overnight in selection antibiotics and found that 12 were successfully transformed and 2 were false positives. A 500 μL sample from each overnight culture was taken and stored in 30% glycerol in MRS labelled KEV 1 - 14 (**K**nockout **E**mpty **V**ector) at -80 °C.

Polymerase chain reaction (PCR) was used as a routine diagnostic method, the resulting amplicons could also be sent for sequencing as an additional benefit (Figure 17B). However, genomic DNA contaminated results with three to four bands that amplified in a cycle dependent manner (Figure 17C). Changing the temperature had no significant effect on amplification of contaminating bands nor did it modify the intensity of plasmid dependant bands caused by the 10 ng of pRV613-pts18cba-FLAG DNA that was spiked into B1 genomic DNA preparations (Figure 17D). The contaminating bands were less visible and were often obscured by one of the intense bands caused by the pRV613 gene insert. Because the primers were designed to bind the plasmid 100 bp above and below the gene insertion site, the size of the PCR product roughly matches the expected sizes of the gene inserts (figure 17B).

The best results appeared to be those that had higher plasmids to genomic DNA ratios (data not shown). A commercial plasmid preparation kit, genomic DNA preparation kit and a chemical plasmid extraction protocol[141] were used to find the best routine plasmid extraction method for Firmicutes PCR diagnostic assays. The chemical extraction method produced the best PCR results while there appeared to be no significant difference between either of the kits (2.2.5.2-2.2.5.3). As the use of PCR to confirm transformations was faster and more reliable than analytical digests, this method was used to judge the

success of all future transformations, along with subsequent sequencing of the PCR products to confirm their identity.

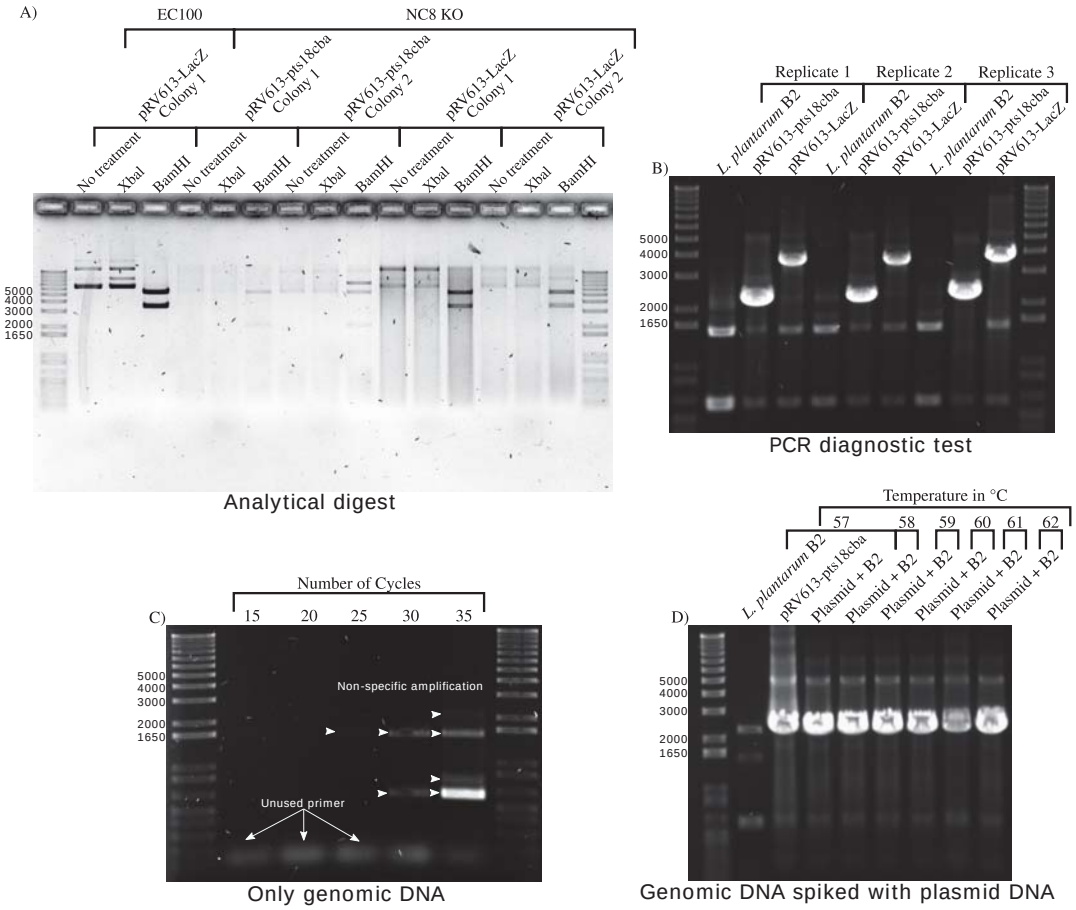


Figure 17: **Methods of transformation confirmation and optimisation**
(A) Analytical plasmid digest of transformed cells: EC100 pRV613-LacZ, the stock used for transformations in *L. plantarum* NC8 KO. pRV613-pts18cba and pRV613-LacZ plasmid stocks extracted from transformed *L. plantarum* NC8 KO overnight cultures were treated with either nothing (No treatment), *Xba*I or *BamH*I. **(B)** Using PCR to confirm transformation: overnight cultures of *L. plantarum* B2, *L. plantarum* NC8 containing pRV613-LacZ or pRV613-pts18cba were grown and treated with commercial plasmid isolation kits, followed by PCR amplification and detection on EtBr stain 1% agarose gels. **(C)** Effect of PCR cycles on non-specific amplification generated by *L. plantarum* B2 genomic DNA used in B, no plasmid DNA was added to any of the reactions. **(D)** Effect of temperature on non-specific amplification when using *L. plantarum* B2 stocks generated using the Klienhammer extraction method. The stock was then spiked with pRV613-pts18cba plasmid (B2 + plasmid) and subjected to PCR at a range of temperatures.

3.3.2 Other factors influencing transformation efficiency

The conditions that had the most impact on the success of transformations in *L. plantarum* NC8 using pRV613 were recovery temperature, media composition, selection antibiotics,

the freshness of the polyethylene glycol 1500 (PEG 1500), the purity and concentration of the DNA stock and the cell type. *L. plantarum* NC8 wild-type generated many more false positives than *L. plantarum* NC8 KO, which also appeared to grow slower and was not as tolerant of changes to growth conditions.

The most optimal recovery temperature for *L. plantarum* transformations was 30 °C. Optimisation of plate recovery temperature made use of the same two-hour recovery stock, which was plated onto six selection plates and grown at three different temperatures, 25 °C, 30 °C and 37 °C. No cells grew at 37 °C, while plates incubated at 30 °C had excellent growth. The cells on recovery plates incubated at 25 °C grew very slowly and were mainly false positives.

Alterations to the culture media included manipulation of the selection antibiotics, the addition of selected salts, and the addition of N-acetylglucosamine (GlcNAc) to the recovery media. In the case of *L. plantarum* NC8 KO, the *pts18cba* gene has been replaced with the chloramphenicol (Chl) selection marker, and when transformed with the pRV613 plasmid these cells acquire an erythromycin (Ert) selection marker as well. Cells grown in the presence of both selection markers grew very slowly probably due to the additional stress, and the number of successful transformants was significantly reduced. When only the plasmid selection marker was used cell growth improved. Transformed cells that were picked from plates made with a single selection marker were grown overnight in media containing both selection antibiotics to show that they were indeed *L. plantarum* NC8 KO cells before glycerol stocks were prepared. While the addition of $MgCl_2$ and sucrose to the recovery media did improve the viability of transformed cells, the addition of GlcNAc had a much bigger impact on success rate, although it seem to depend on cell type. Using GlcNAc in the recovery media produced many false positives when using *L. plantarum* NC8 wild-type cells. If the concentration of the cells were too high when plating onto GlcNAc supplemented agar plates, a lawn would grow that completely obscured the transformants. The addition of GlcNAc did however impact the recovery of *L. plantarum* NC8 KO cells without the growth of satellite colonies or lawns.

3.3.2.1 Vector size and cell type considerations

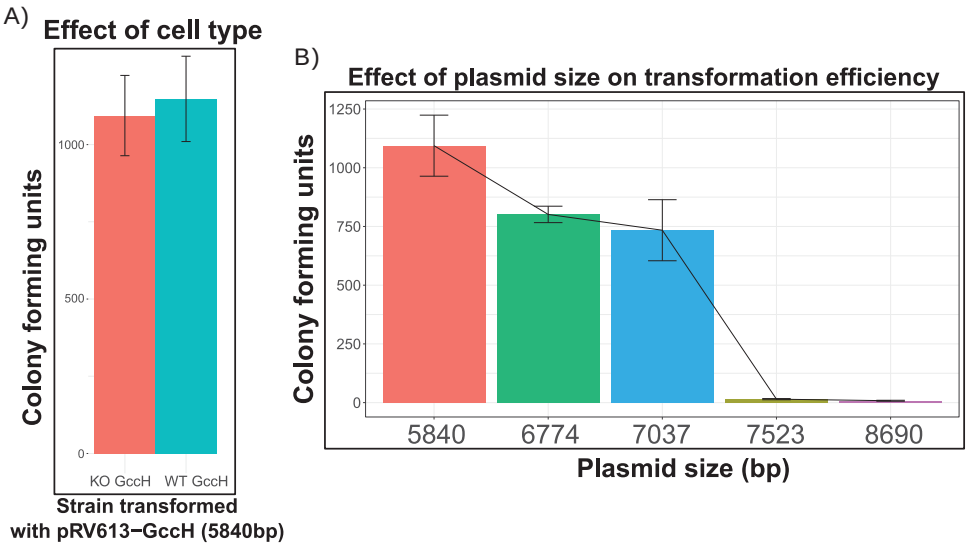


Figure 18: **Apparent effect of plasmid size on transformation efficiency**
(A) The average cell counts for independent transformation events using the pRV613-GccH (5840 bp) vector used to transform both *L. plantarum* NC8 WT and KO cells (n = 3). The cells were recovered on GlcNAc free media and MRS agar. **(B)** A comparison of cell counts vs plasmid size. Counts are the average of 3 - 4 independent *L. plantarum* NC8 KO transformation events using pRV613 plasmid variants and recovered on GlcNAc supplemented MRS agar. Each of the plasmids variants differ only by the gene insert. Each transformation event was carried out independently from start to finish using either *L. plantarum* NC8 KO, or *L. plantarum* NC8 wild cells. Error bars represent the standard error.

The apparent transformation efficiency for both NC8 KO and NC8 WT cells were similar regardless of whether GlcNAc was in the recovery media or not when transforming cells with smaller plasmids (Figure 18A). However, the size of the plasmid did have an effect. The larger the plasmid, the lower the apparent transformation efficiency with a sharp decline occurring between 7 kb and 7.5 kb (Figure 18B). The term "apparent" was used because the cells were not tested to confirm that they were indeed transformed, but were just counted without considering false positives.

3.3.2.2 Antibiotic and temperature toxicity

The KO mutant is very resistant to chloramphenicol as expected (Figure 19A) but sensitive to erythromycin, with concentrations of 30 $\mu g \cdot mL^{-1}$ was enough to almost entirely stop growth. At 30 °C, a 15-hour time course showed that the confidence interval shifted

near to the end of the 15-hour mark. Understanding this time delay is important because it takes approximately 48 hours for cells to become visible to the naked eye, at this stage single colonies are picked and cultured. If selection is ineffective before that time, the risk of contamination will be high. A closer look at the behaviour of the WT cells containing no resistance markers showed that these cells recover from chloramphenicol selection very early, making this antibiotic less effective over the longer time frames needed for transformations (Figure 19B). Using both selection antibiotics appeared to be more robust over the 15 hours required for transformed cells to grow. As erythromycin was more efficient than chloramphenicol for this species, it is possible that selection for only the plasmid can be effectively used.

The liquid culture experiments were confirmed on agar plates over the full time frame required to produce visible colonies. Using the same conditions with a wider selection of cell types. Full growth was seen for all the cell stocks used on plates without antibiotics (Figure 20A), and all untransformed cells failed to grow in the presence of 30 $\mu g \cdot mL^{-1}$ erythromycin although there was some recovery of the wild type strain in the presence of 30 $\mu g \cdot mL^{-1}$ of chloramphenicol media (Figure 20B and C). Dual selection at 30 °C showed similar results to single erythromycin selection, although with much weaker cell growth supporting the hypothesis that dual selection may be too stressful for newly recovering cells (Figure 12D). Repeating the double antibiotic selection at 37 °C showed that while the KO cells remained inhibited, the WT cells were capable of recovering from inhibition over this time frame at high temperatures (Figure 20E) and generally appear fitter than the KO mutants. Thus 30 °C remains the most optimum temperature for recovery and growth of cells exposed to selection conditions.

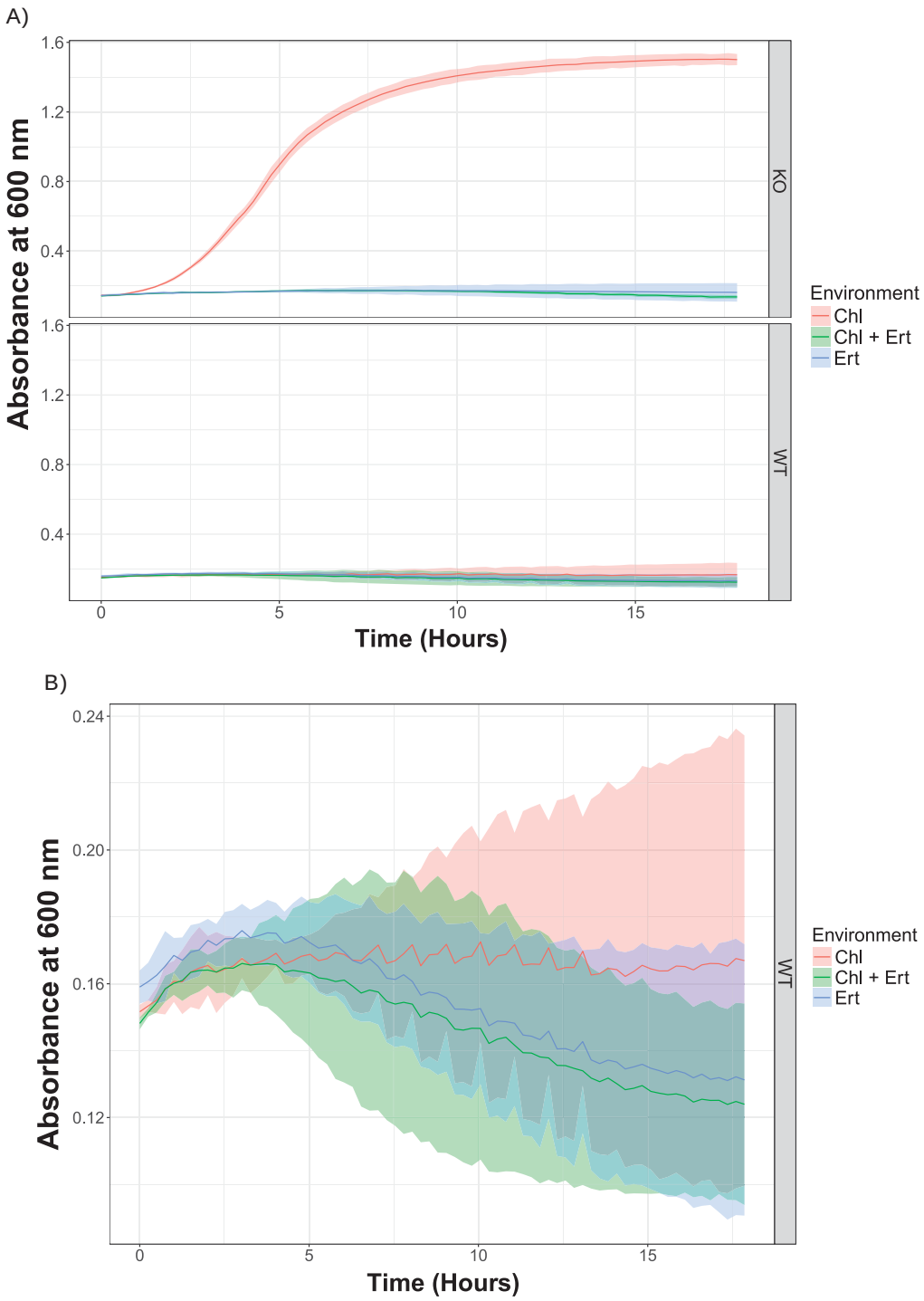


Figure 19: **Antibiotic effects on selected *L. plantarum* NC8 cell types**
(A)*Lactobacillus plantarum* NC8 wild type (WT) and *Lactobacillus plantarum* NC8 *pts18cba* knockout (KO) cell stocks treated with either 30 $\mu g \cdot mL^{-1}$ of chloramphenicol (Chl), 30 $\mu g \cdot mL^{-1}$ of erythromycin (Ert) or both for 17 hours. **(B)** Re-plot of the *Lactobacillus plantarum* NC8 WT graph from A to improve the resolution for confidence interval assessment. Sample size for each condition from independent overnight cultures: n = 4.

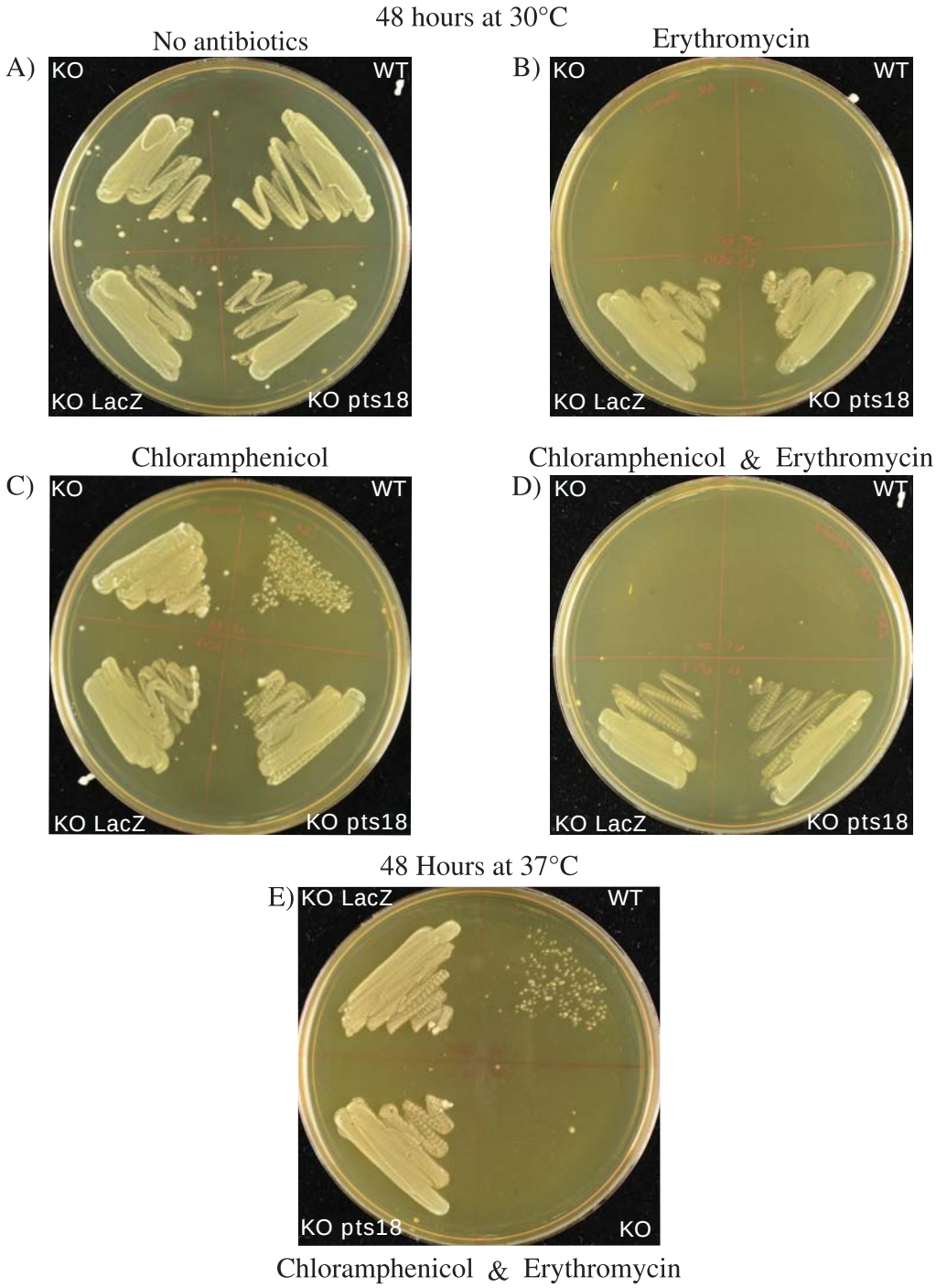


Figure 20: **Antibiotic and temperature effects on *L. plantarum* NC8**
L. plantarum NC8: Δ*pts18cba* (KO), wild type (WT), KO pRV613-*pts18cba-flag* (KO *pts18cba*), and KO pRV613-*LacZ* (KO *LacZ*) glycerol stock recovery on MRS plates containing 1% agar. The plates contained either no antibiotics, 30 μg·mL^{-1} of chloramphenicol, 30 μg·mL^{-1} of erythromycin or 30 μg·mL^{-1} of both antibiotics and were incubated at 30 °C or 37 °C for 48 hours before photography. **(A)** 30 °C with no antibiotics, all the strains show strong abundant growth. **(B)** 30 °C on MRS containing erythromycin. Both KO and WT failed to grow, KO *LacZ* and KO *pts18cba* are uninhibited. **(C)** 30 °C on MRS containing chloramphenicol. WT produces weak growth after 48 hours while all other strains grow unimpeded **(D)** 30 °C on MRS containing both antibiotics. WT and KO show no growth, while KO *pts18cba-flag* and KO *LacZ* show weak growth. **(E)** 37 °C on MRS containing both antibiotics. Both KO *LacZ* and KO *pts18cba-flag* are uninhibited, WT shows very weak growth, KO failed to grow.

3.3.2.3 Copper toxicity

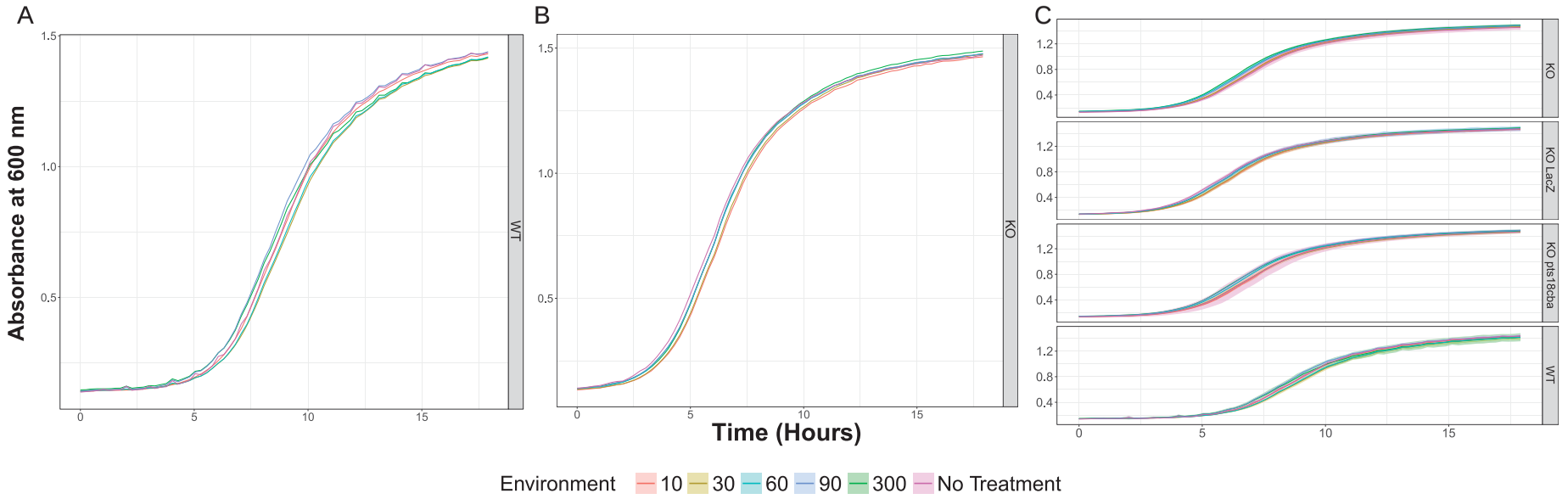


Figure 21: ***L. plantarum* NC8 cell types: Cu^{2+} toxicity**
Three cell cultures were grown to stationary phase before being diluted into the appropriate copper supplemented MRS media and measured in triplicate. **(A)** *L. plantarum* NC8 wild type,**(B)** *L. plantarum* NC8 KO cells grown in MRS media containing various concentrations of copper. **(C)** *L. plantarum* NC8 WT, KO, and KO cells containing pRV613:pts18cba-FLAG (KO-pts18cba) or pRV613:LacZ (KO-LacZ). The shading represents the 95% confidence interval for averaged biological replicates (n = 3).

L. plantarum NC8 cells appeared to be very tolerant of copper. Adding copper to media that already contained trace amounts of copper did not appear to harm the cells (Figures 21A and B). At higher concentrations, copper appeared to marginally increase growth although this was not consistent with all MRS media stocks. The biggest effect was an increase in the 95% confidence interval for samples that exceeded 300 μM copper concentrations, but was not significant at the resolution required for future experiments.

3.3.2.4 Expected results for successful transformations

The success of a transformation varies depending on a number of variables that for the most part are species specific. When conducting *L. plantarum* transformations, smooth, waxy, circular colonies that appear solid and beige in colour are considered normal (Figure 22). During most attempts, a slight translucent lawn made up of dead cells caused a matte like appearance on the surface of the agar (Figures 23C and 24C). The lawn is likely the result of high cell concentrations that are typical for this type of transformation method, but should never appear thick as seen in Figures 23D and 24D.

Another species specific variable is recovery time, which is as a rule of thumb, based around the doubling time of the cells. However using the doubling time is often misleading because cells need additional time to repair their cell walls, and suffer from stress caused by the sudden invasion of foreign DNA. The doubling time for *L. plantarum* is around 40 - 51 minutes under optimal conditions although a 2 hour recovery time is generally recommended (Figure 23A). *L. plantarum* NC8 KO cells appear to be less fit than *L. plantarum* WT and generally recover a lot slower than the WT. It is in this situation that higher levels of success were seen when allowing the cells to recover overnight in antibiotic free recovery media. This practice is not without its risks, and contamination can be a real issue. The most common form of contamination is the development of satellite colonies (Figure 22A). These untransformed cells appear on the selection plates near the end of the incubation cycle and are morphologically identical to the transformed cells with the exception of their size (Figure 22A).

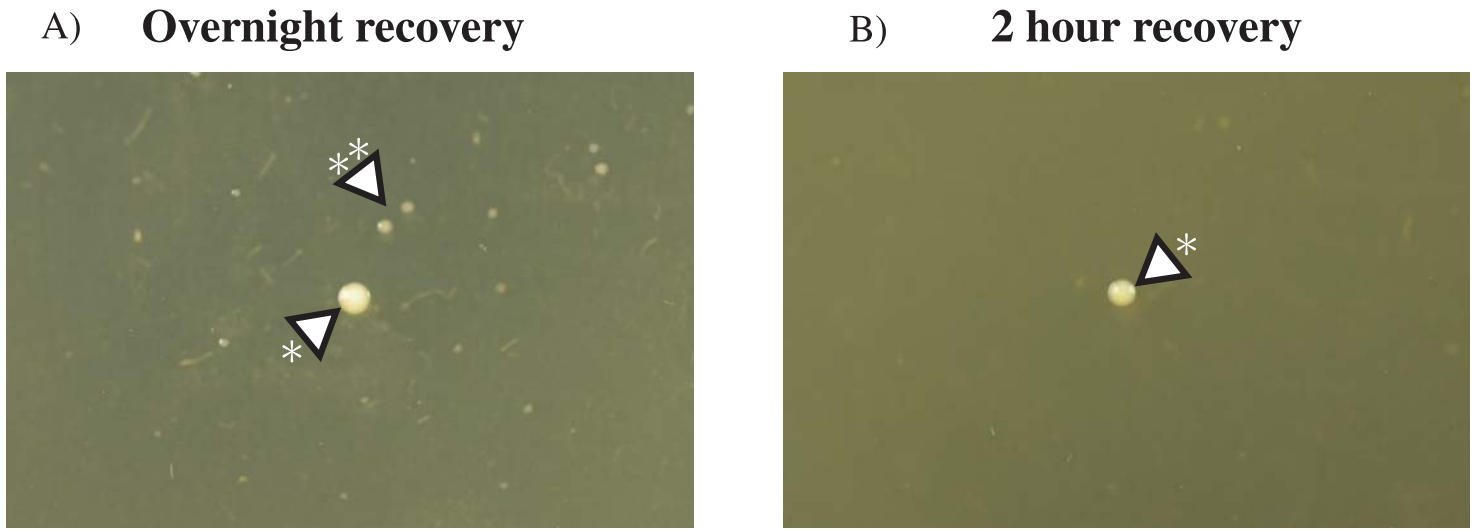


Figure 22: **Successful transformations from overnight or 2 hour recovery times**
(A) Successful transformation: sample colonies from an overnight recovery. **(B)** Successful transformation: sample colonies from a 2-hour recovery. **Cell morphology**: Smooth, circular and beige in colour. Cells should be large, circular and uniform (A - bottom arrow and B). Contamination from satellite colonies (A - top arrow) are common for long recovery times, or when outgrowth plates are incubated for extended periods. *Successfully transformed colonies. **Satellite colonies that proved to be false positives.

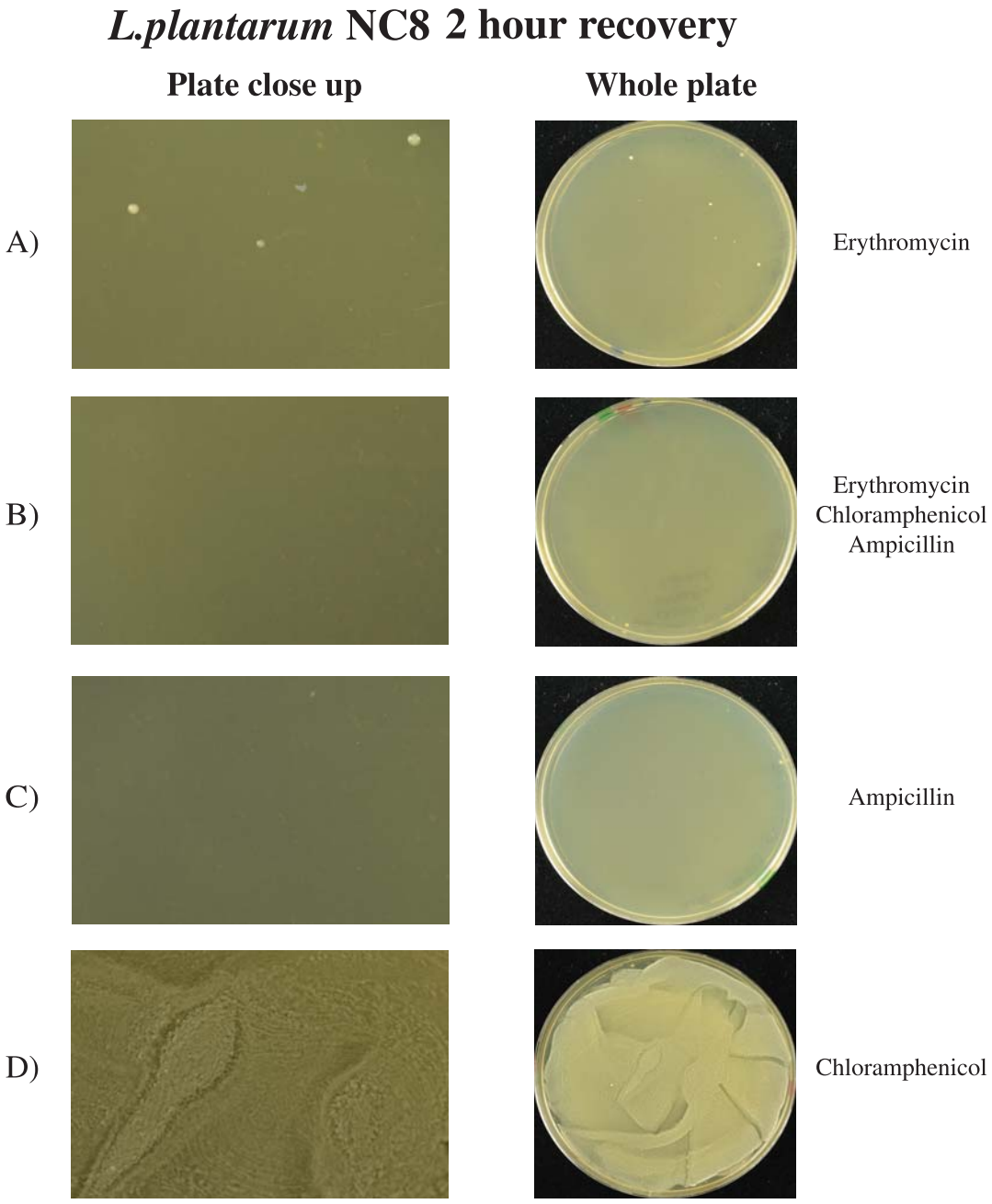


Figure 23: **Complete plate results from a 2 hour liquid recovery before plating** *Lactobacillus plantarum* NC8 pts18cba knockout (KO) strains transformed with the 8.69 kb pRV613-LacZ vector and allowed to recover for 2 hours in MRS media supplemented with 0.5% w/v GlcNAc, 0.4 M sucrose and 0.1 M $MgCl_2$ without antibiotic selection. Post recovery, cells were plated on MRS recovery agar plates containing $\sim$1% agar and supplemented with 0.5% GlcNAc, and containing either 30 µM of ampicillin, or erythromycin, or chloramphenicol, or all three antibiotics. The recovery plates were then incubated for 48 hours at 30 °C before photography. These cells were labelled KO-LacZ.

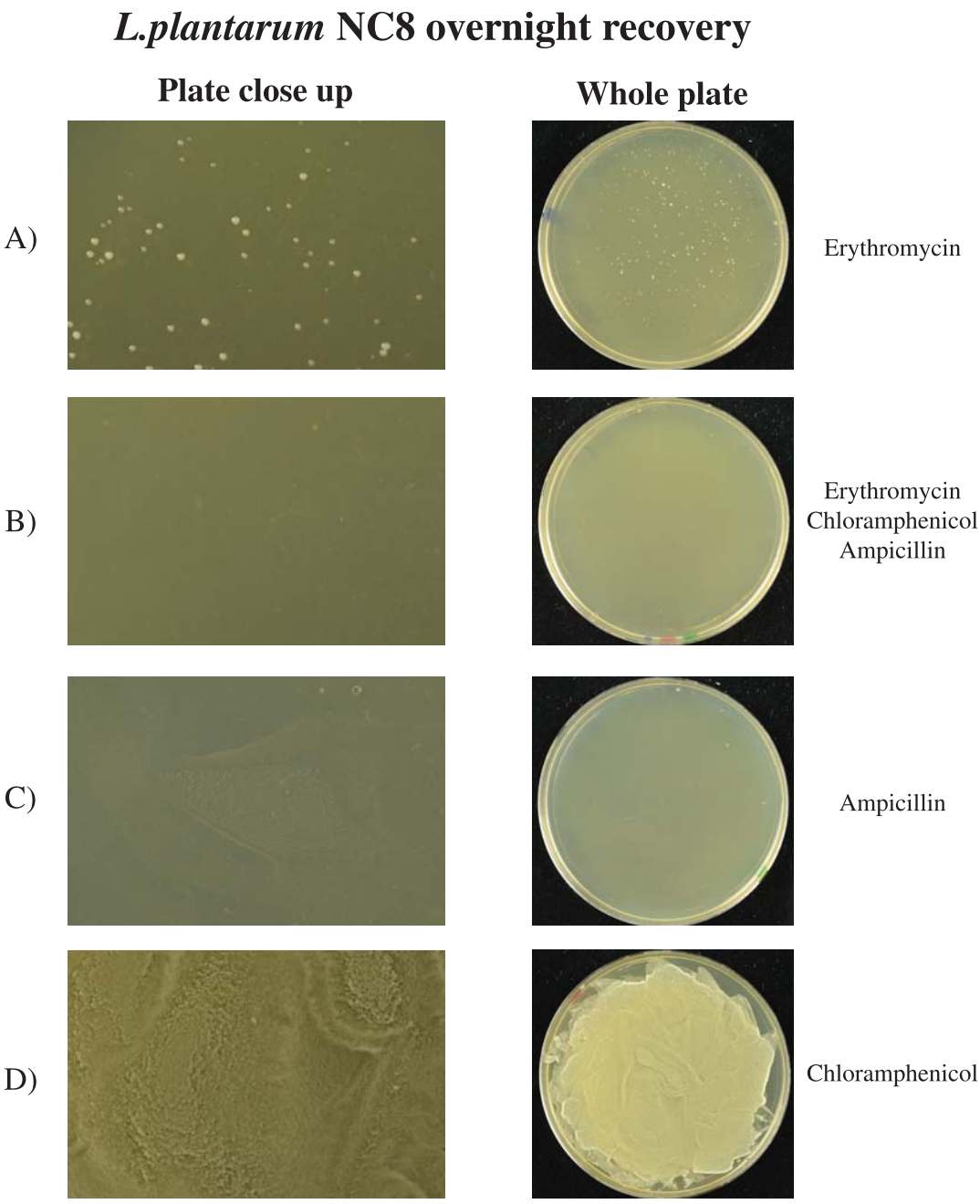


Figure 24: **Complete plate results from an overnight liquid recovery before plating** *Lactobacillus plantarum* NC8 pts18cba knockout (KO) strains transformed with the 8.69 kb pRV613-LacZ vector and allowed to recover overnight in MRS media supplemented with 0.5% w/v GlcNAc, 0.4 M sucrose and 0.1 M $MgCl_2$ without antibiotic selection. Post recovery, cells were plated on MRS recovery agar plates containing $\sim$1% agar and supplemented with 0.5% GlcNAc, and containing either 30 µM of ampicillin, or erythromycin, or chloramphenicol, or all three antibiotics. The recovery plates were then incubated for 48 hours at 30 °C before photography. These cells were labelled KO-LacZ.

Table 9: ***Lactobacillus plantarum*** **NC8 strains**

#	Foundation strain	Modification shorthand	Genetic background	Plasmid	Origin
1	*L. plantarum* NC8 pts knockout	KO	Δ*pts18cba*	None	K. Drower[67]
2	*L. plantarum* NC8 pts knockout	KO LacZ	Δ*pts18cba*	pRV613:*LacZ*	This study
3	*L. plantarum* NC8 pts knockout	KO pts18cba	Δ*pts18cba*	pRV613:*pts18cba-FLAG*	K. Drower[67]
4	*L. plantarum* NC8 pts knockout	KO pts18cb	Δ*pts18cba*	pRV613:*pts18cb-FLAG* ΔEIIA	This study
5	*L. plantarum* NC8 pts knockout	KO pts18c	Δ*pts18cba*	pRV613:*pts18c* (S) ΔEIIA & EIIB	This study
6	*L. plantarum* NC8 pts knockout	KO pts18c	Δ*pts18cba*	pRV613:*pts18c* (L) ΔEIIA & EIIB	This study
7	*L. plantarum* NC8 pts knockout	KO H584A	Δ*pts18cba*	pRV613:*pts18cba-FLAG* His584Ala	This study
8	*L. plantarum* NC8 pts knockout	KO GccH	Δ*pts18cba*	pRV613:*GccH*	This study
9	*L. plantarum* NC8 Wild Type	WT GccH	Wild type	pRV613:*GccH*	This study
10	*L. plantarum* NC8 Wild Type	WT NC8	Wild type	None	Aukrust & Blom[170]
10	*L. plantarum* 8014	B2	*pts18cba*Δ762-831	None	K. Drower[67]

Lactobacillus plantarum NC8 strains used or developed for use during this study. The *lactobacillus plantarum* NC8 knockout (KO) mutants or wild type (WT) were used to produce all the other bacterial isolates and are represented with the short hand "KO" for knockout or "WT" for wild type. The untransformed KO and WT isolates do not contain any natural plasmids. The KO isolates were generated by removing the *pts18cba* gene from the genetic background of *Lactobacillus plantarum* NC8 WT cells[67]. All other strains were created by introducing extrachromosomal plasmid harbouring the gene of interest by electroporation and named according to the foundation strain and the gene insert found harboured on plasmid. For instance KO GccH contains the *gccH* gene cloned into the pRV613 vector and transformed into the *L. plantarum* NC8 KO mutant. Each "foundation strain" is represented by no less that three cell stocks isolated from separate transformation events.

3.4 The effect of plasmids on bacterial growth

3.4.1 Complementation of *pts18cba* into *L. plantarum* NC8 cells

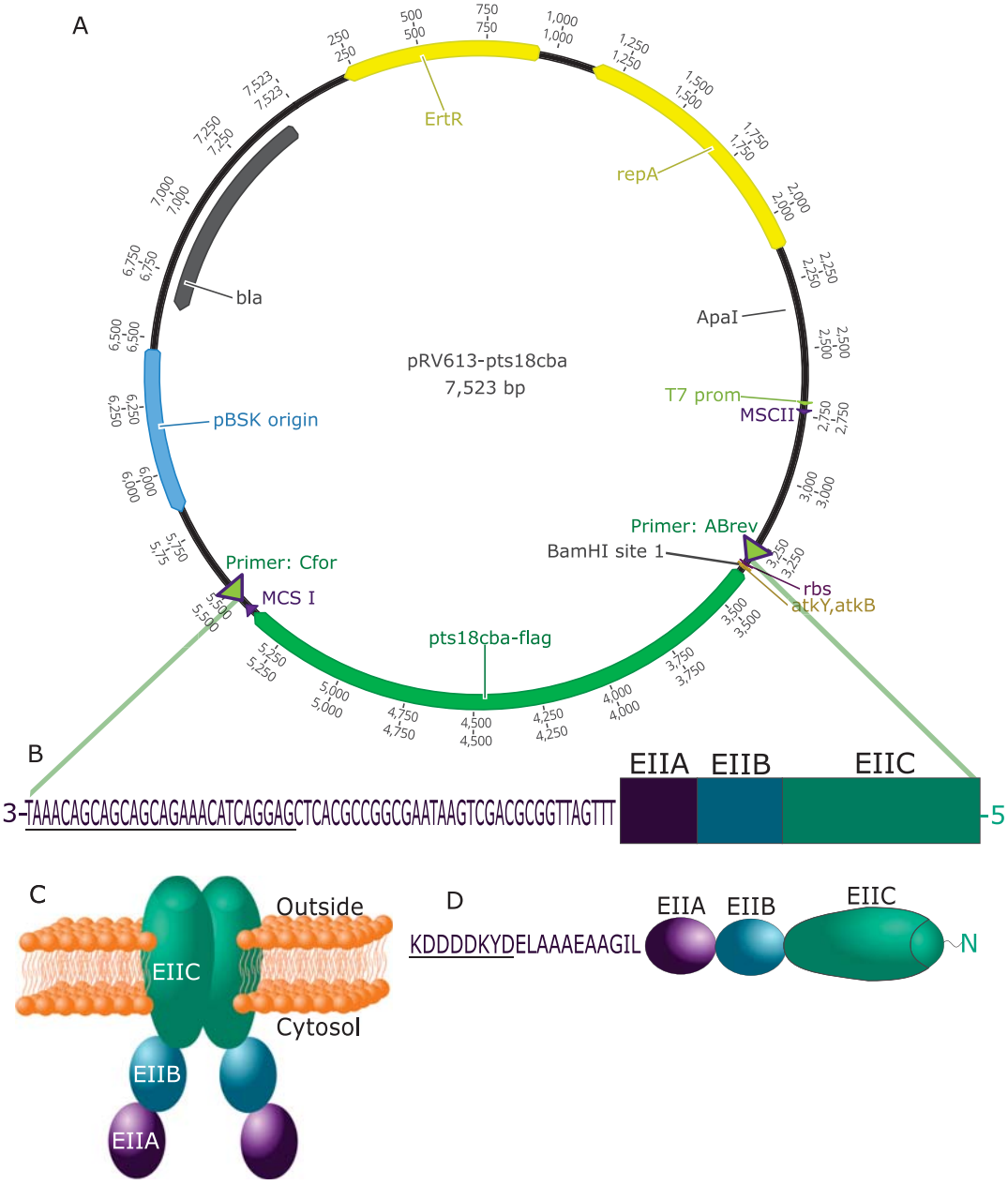


Figure 25: **Complementation: pRV613-*pts18cba-FLAG* vector map**
(A) Plasmid map of pRV613-*pts18cba-FLAG*: atk and atkY (Gold) are copper inducible promoters, pBSK (Blue) *l. plantarum* origin of replication, bla, repA replication factors (Yellow), rbs the ribosomal binding site (Purple), MSCI and MCSII (Purple arrowheads) are multiple cloning sites. Purple border green arrows represent the binding position of the primers used to confirm transformation success. **(B)** *Pts18cba* insert orientation and domain order. The EIIC (Green), EIIB (Blue), and EIIA (Purple), along with the 3′ nucleic acid sequence containing the FLAG-tag (underlined). **(C)** A model for protein dimerisation, orientation and insertion into the membrane (orange). **(D)** Model of expected gene product with C terminal FLAG-tag amino acid sequence (underlined) represented C- to N-terminus for consistency with gene orientation from the plasmid map.

Previous natural mutation studies showed that the PTS18CBA might be involved in the GccF mechanism of action. Complementation experiments were done to explore this result more directly by removing the *pts18cba* gene from the genetic background of a *L. plantarum* NC8 WT cells, then reintroducing the gene back into this knockout mutant (KO) using a plasmid. The pRV613-*pts18cba-FLAG* plasmid used for these experiments was a kind gift from K. Drower and was originally designed for FLAG-tag specific protein precipitation assays but proved ideal for complementation experiments and was used as a template for the construction of all the other vectors. This *lactobacillus* optimised vector places the *pts18cba* gene under the direct control of a copper-inducible promoter and has a low copy number *in vivo*.

The phenotyping experiments were done alongside an untransformed *L. plantarum* knockout cells (KO), the *L. plantarum* WT strain that was used to make the KO mutant strain, and a KO cells transformed with pRV613-*LacZ*, the parent plasmid that was used to construct the pRV613-*pts18cba-FLAG*. The pRV613-*pts18cba-FLAG* and derivative plasmids are identical to the pRV613-*LacZ* plasmid except that the pRV613-*LacZ* has a *LacZ* gene under the control of the copper inducible promoter instead of the *pts18cba* gene. Data was collected using a 96 well titre plate that measured growth over 15 hours, opposed to the traditional 6.5 hours so that additional information regarding the growth trends could be gathered (2.5.1.2). The growth data were collated into a single database so that all the controls would be the same.

The cells were either treated with 30 µM Cu^{2+}, or 60 nM of GccF or both. Early experiments made use of a GccF treatment in Cu^{2+} free media to show that the cells were resistant when the *pts18cba* gene expression was not induced. These results proved to be highly variable and inconsistent over time and were not reproducible. Inspection into the cause of the variation and loss of induction control appeared to be natural fluctuations of the copper concentration in the commercial growth media. As the media contains animal extracts with variable copper concentrations it was outside the control of the manufacturers. The difference proved significant between brands and time of purchase, and was

reflected in differences in cell growth. Attempts to make media from individual components also failed to allow selective induction of the gene. Using chelating agents capable of removing the divalent cations and reintroducing them as needed is another method that would allow selective control of the promoter. While such media should result in control of the gene induction, it may also introduce unwanted bias. Instead of defined media treated with chelating agents, it was decided to use a mock plasmid to mimic an uninduced phenotype.

This mock strain was made from the same KO stock but contained the *LacZ* instead of *pts18cba*. Because *LacZ* gene is unlikely to be involved in GccF mechanism of action, and the gene is larger than the *pts18cba* gene its induction should cause similar behavioural changes in the expressing cells, making it a good experimental control. Once the strains had been generated, an overnight culture was grown, and plasmid specific PCR was carried out to check that the cells were transformed, and the products were sent for DNA sequencing. Analysis by gel electrophoresis showed the presence of bands that matched the expected sizes for both pts18cba (3484 bp) and LacZ (2317 bp) (Figure 26A). The sequencing results showed that neither gene insert contained any unexpected changes and matched their respective master sequences.

The growth profiles for each of the strain showed no significant differences between the no treatment controls and the Cu^{2+} (copper) induced samples (Figure 26D). When treated with GccF all the strains showed some growth inhibition. Cell lines like KO that were expected to be completely resistant only began to grow after approximately 4 hours, whereas cells that were complemented with the *pts18cba* gene and should have been completely inhibited also showed growth after $\sim$4 hours (Figures 26D and 27A). The WT cells were far more sensitive to GccF than any of the knockout strains as expected. Indicator plates confirmed that regardless of the presence or absence of the PTS, all cells were susceptible to GccF (Figure 26B). However, the presence of the PTS increased the bacteriostatic response to GccF (Figure 26B).

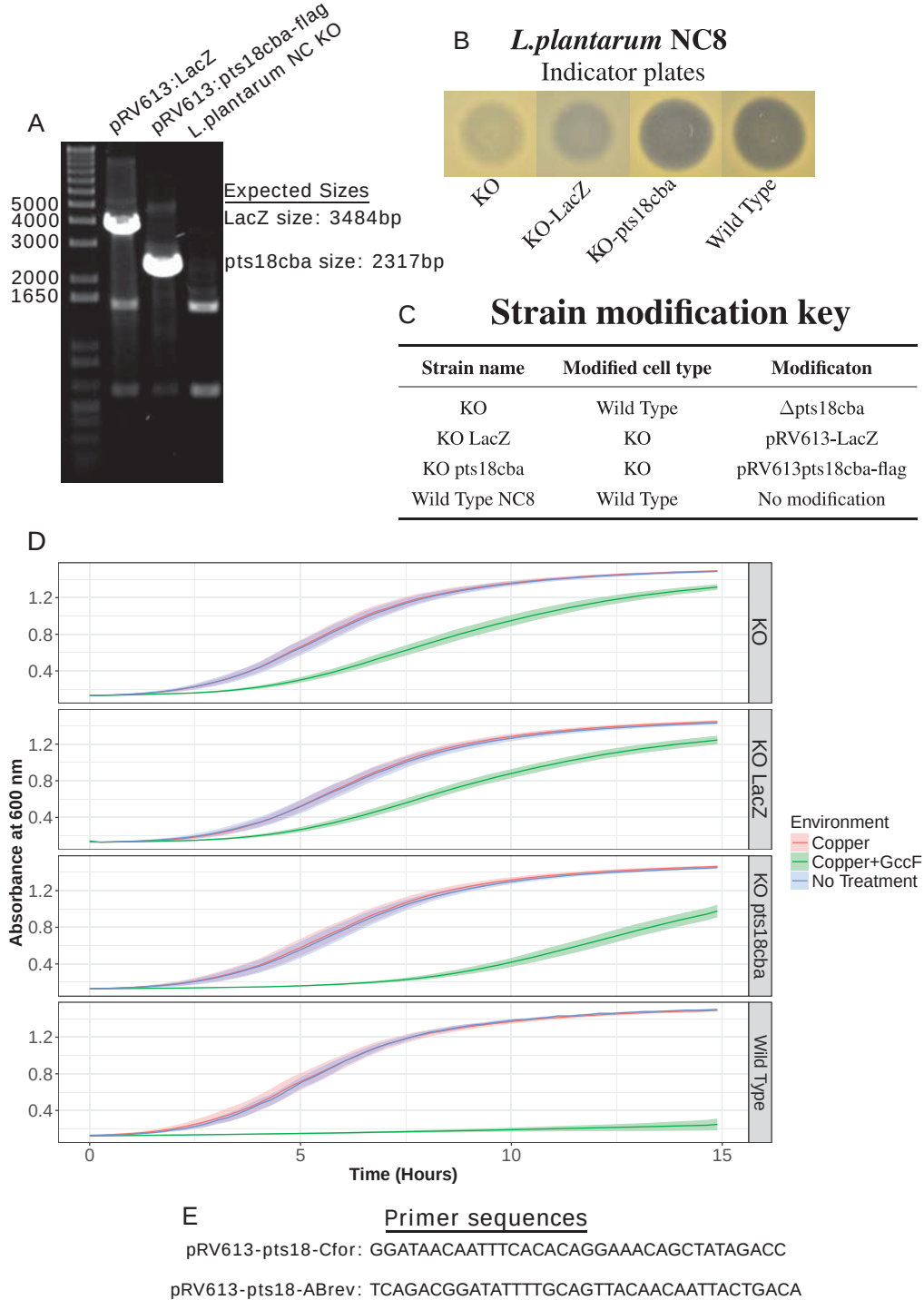


Figure 26: **Complementation growth data grouped by strain**
(A) Plasmid inserts amplified by PCR from positively transformed *L. plantarum* NC8 cells. **(B)** Clearing zones on indicator plates caused by 5 µL of undiluted GccF ($\sim$ 900 µM). The axis scales are linear. **(C)** KO: *L. plantarum* NC8 *pts18cba* knockout mutant; KO-LacZ: *L. plantarum* NC8 KO mutants transformed with the pRV613-*LacZ* plasmid; KO-pts18cba-flag: *L. plantarum* NC8 KO mutant containing the pRV613-*pts18cba-FLAG* plasmid; Wild Type NC8: An unmodified *L. plantarum* NC8 strain. **(D)** Growth data collected by micro titre plate, and grouped by strain. Experiments were conducted without antibiotic selection. Copper (Red): MRS supplemented with 30 µM of Cu^{2+}. Copper+GccF (Green): MRS media supplemented with 60 nM of Glycocin F and 30 µM Cu^{2+}. No treatment (Blue): Only sterilised MRS was used. Shading represents a 95% confidence interval. **(E)** PCR primers used for transformation confirmation. Expected insert sizes: LacZ (3484bp) and pts18cba (2317bp).

The WT and KO strains behave almost identically in the presence or absence of induction. They do differ slightly however from both strains that contain plasmids, which also behave the same as each other. The most likely explanation for this minor difference in growth between transformed and untransformed cells is stress to the cells caused by the presence of the plasmid and its insert (Figure 27A). GccF is most effective against the WT strain and causes a complete halt of bacterial growth for 7 hours.

The strains harbouring the pRV613-*pts18cba-FLAG* plasmids (Figure 25) were significantly more sensitive to GccF than both the KO and KO-LacZ cells (Figure 27A). The lack of a significant difference in growth of KO-LacZ cells shows that the vector itself has minimal impact on cell viability. However, complemented cells do not return the same magnitude of inhibition as the WT cells do, which could be because the *pts18cba* insert is constitutively expressed by a low copy number plasmid, and therefore may not reach the same level of abundance as the WT cells. Usually when cells experience stress or when necessary proteins are at an undesirable level, the genes that control their expression can be regulated to compensate for any changes. Plasmid constructs do not have the ability to naturally regulate their gene expression which is usually constitutive in the presence of an inducer.

An IC_{50} is a measurement that determines the minimum amount of GccF that is required to reduce growth by 50%. The IC_{50} for GccF treated *L. plantarum* NC8 wild type cells was 40 nM of GccF. The IC_{50} measurements could not be determined for the resistant mutants because of the large amount of peptide required. An estimate of cell growth at eight hours was therefore used to determine the phenotype. This time point was selected because there was enough difference to clearly determine differences in respect to GccF for each of the tested strains (Figure 27).

At eight hours there was no difference between untransformed strains that had not been induced with copper. WT cells have a larger error in growth rate with the addition of copper which is consistent with the toxicity study that found copper increases growth without

altering the growth profile. Both transformed cell types performed worse than untransformed cells, but were not significantly different from each other. The KO-pts18cba strain did appear to be slightly fitter than the KO-LacZ during transformations and phenotyping at eight hours. However, this observation proved insignificant when errors where calculated or when a 95% confidence interval was fitted over the entire 15-hour growth profile.

Harbouring a plasmid places the cells under additional stress and causes a reduction in growth (plasmid pressure). A comparison was made between KO and KO-LacZ cells to determine the effect of the plasmid on growth. When cells were treated with copper, the plasmid reduces growth of the transformed strains by $\sim$10%, which was also reflected by $\sim$9% loss when treated with GccF (Figure 27). Furthermore, the cells containing the *pts18cba* gene appear to experience less plasmid stress than those that contain the *LacZ* gene. This difference may be due to the smaller size of the *pts18cba* gene or that it may provide the cells with access to an additional carbon source that *LacZ* does not. Although noticeable, this difference is not statistically relevant for either of the control groups.

When treated with GccF the KO-LacZ and KO cells were expected to remain resistant, however 60 nM of GccF reduced KO cell growth $\sim$57% of the copper treated control (Figure 27). Despite the plasmid pressure, a comparable result was seen in the KO-LacZ strain, ruling out the plasmid influence beyond the gene insert for phenotypes expressed by the wild type and complemented strains. Because both the KO and KO-LacZ strains show susceptibility to GccF without PTS18CBA, this could be an indication that there is a second mechanism or that the PTS18CBA is not entirely responsible for the bacteriostatic phenotype when exposed to GccF. The WT cells are the most sensitive to GccF as expected, which reduces cell growth by $\sim$76% of the GccF treated KO growth. Interestingly the KO mutants can be re-sensitised to GccF when transformed with the pRV613-*pts18cba-FLAG* plasmid (KO-pts18cba).

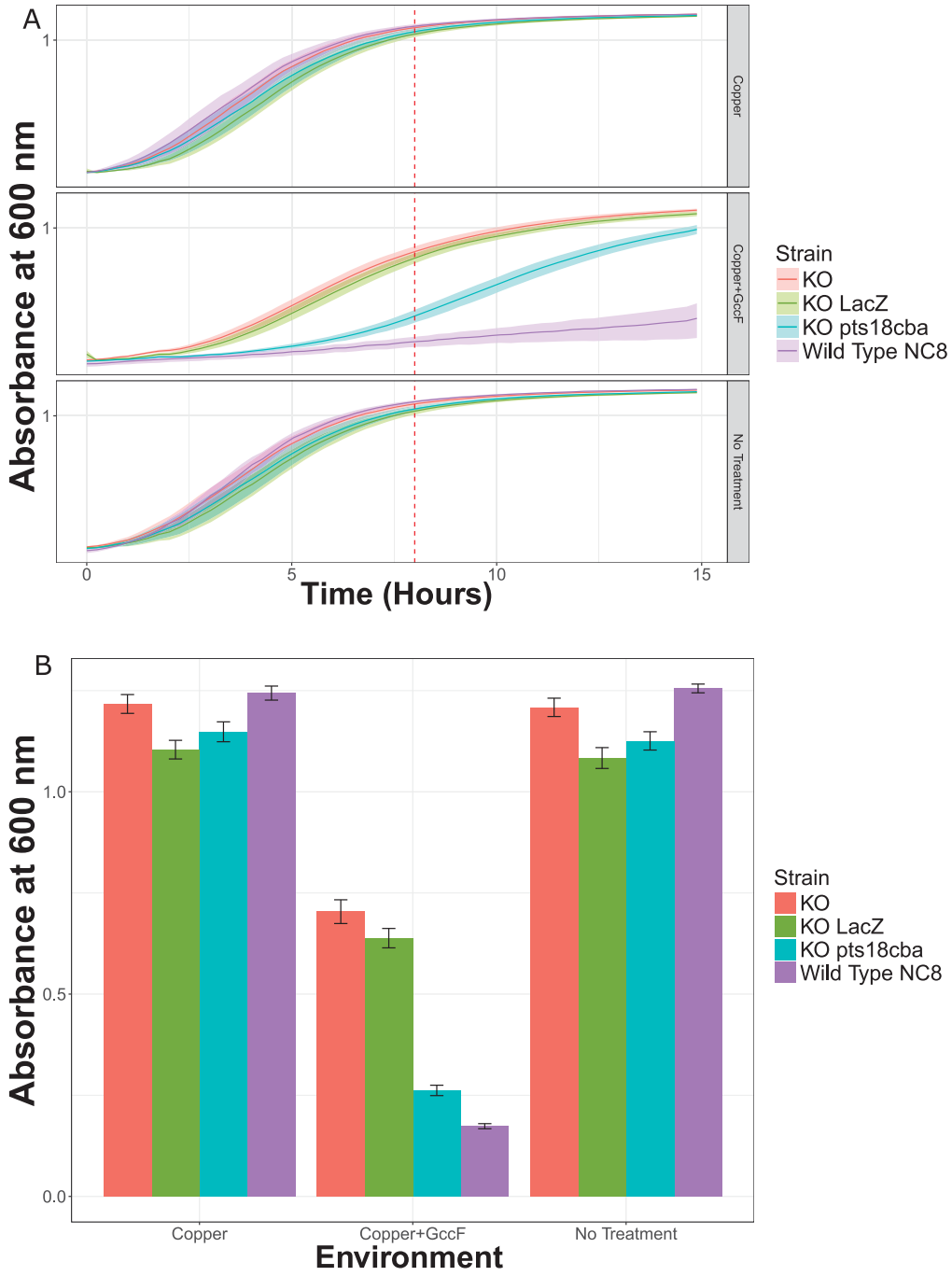


Figure 27: **Complementation: Strain analysis grouped by environmental condition** **(A)** Complementation growth data grouped by environment, where shading represents a 95% confidence interval, the red dotted line indicates the 8 hour interval used for data quantification in B below. **(B)** Bar chart showing quantitative data of cell growth at 8 hours as indicated by the dotted line in growth data above A. Error bars represent the standard error (Table 16). **Key:** KO (Red) pts18cba deficient *L. plantarum* NC8 cells. KO-LacZ (Green) KO cells transformed with the pRV613-LacZ expression vector. KO-pts18cba (Blue) KO cells transformed with the pRV613-pts18cba-flag expression vector; Wild type (Lavender) unmodified *L. plantarum* NC8 cells. **Environmental conditions:** (Copper) MRS media with a final concentration of 30 µM of Cu^{2+}. (Copper+GccF) MRS media with a final concentration of 60 nM of Glycocin F and 30 µM of Cu^{2+}. The No treatment contained only sterilised MRS.

When treated with 60 nM of GccF KO-pts18cba growth is reduced by ∼63% when compared to GccF treated KO growth. The *pts18cba* gene appears to increase sensitivity to GccF by an additional ∼36% when compared to GccF treated KO cells that were normalised against their respective copper control. There is also ∼8% difference between KO-pts18cba and WT strains when compared to GccF treated normalised KO growth, which can be explained by the WT cells natural ability to upregulate the *pts18cba* gene thereby increasing the local concentration of the EII enzyme, a function that the pRV613-*pts18cba-FLAG* plasmid can not mimic. It is clear that the *pts18cba* is involved in the GccF mechanism of action. However only ∼36% of the observed phenotype can be accounted for by the presence of the *pts18cba* gene, and ∼43% of the phenotype appears to be due to an unknown mechanism that affects all the *L. plantarum* NC8 cells when growth is measured at eight hours regardless of any modifications that were made to the plasmid. This result does provide additional support that the PTS18CBA transporter is involved in the GccF mechanism of action, possibly as a sensitisation agent.

3.4.2 The gradual loss of GccF induced bacteriostasis

The potency of GccF drops over time; this phenomenon was first thought to be mainly due to Darwinian selection against the plasmids transformed into these strains. All the construct phenotyping trials were conducted without antibiotic selection because antibiotics could not be used for all the controls and selective use would have introduced bias. However, growth tests using the KO-pts18cba and KO-LacZ cells did not show any changes in their growth profiles when treated with GccF over 15 hours regardless of exposure to the plasmid selection marker erythromycin (Figure 28). These results are consistent with published findings that 95% of plasmids will be retained after 40 generations when exposed to negative selection (∼26 hours)[171]. A drop in growth rate was observed for erythromycin treated samples which was likely the result of additional stress the antibiotics placed on the cells. The initial lag phase remained roughly the same regardless of antibiotic selection indicating that alleviation from bacteriostasis was independent of the antibiotics, and plasmid loss was not significant over a 15-hour time frame.

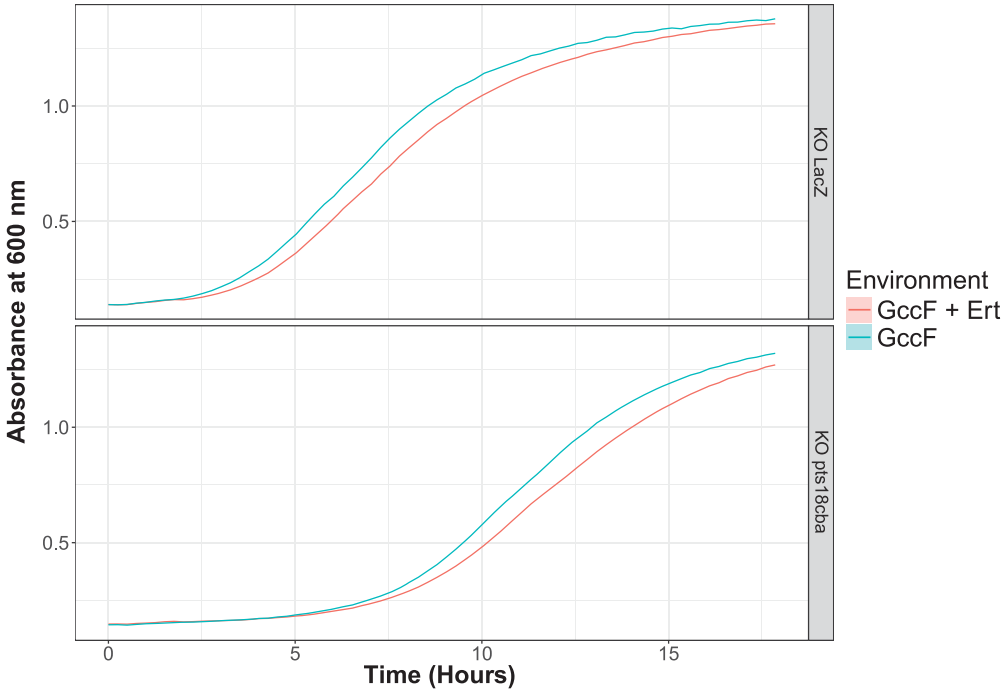


Figure 28: **Complementation: Effect of selection on vectors**
To assess the effect of natural selection against pRV613 *-LacZ* and *-pts18cba-FLAG* plasmids transformed into KO-LacZ and KO-pts18cba strains over 17.5 hours, and the associated alleviation from GccF inhibition. The cells were all grown in sterile MRS containing 60 nM of GccF and 30 µM of Cu^{2+}, and either treated with erythromycin (Ert) or without. n = 4 for each condition.

The stability of GccF was also considered, and was tested by adding GccF to unmodified MRS, MRS containing an additional 30 µM of copper, water, and water containing an extra 30 µM of copper and left at room temperature for a week. A fresh sample containing GccF in water was prepared from frozen GccF stock on the day to serve as a positive control. The activity of these samples were measured by soaking pieces of sterile filter paper in each sample, or GccF free MRS, or water and placing them on MRS agar plates containing *L. plantarum* NC8 cells that were applied by spreading, followed by overnight incubation. The GccF containing samples all showed comparable areas of clearing in contrast to the negative controls which showed no growth inhibition (data not shown). Due to the long time frame of these assays, stability was an unlikely cause for the rapid loss of activity during growth assays. The cells were therefore the most likely cause of the loss of GccF activity during growth assays. It is clear that GccF is bacteriostatic, but by these data, the cells are not likely to be biologically inactivated by GccF.

3.4.3 Removal of domain EIIA from *pts18cba*

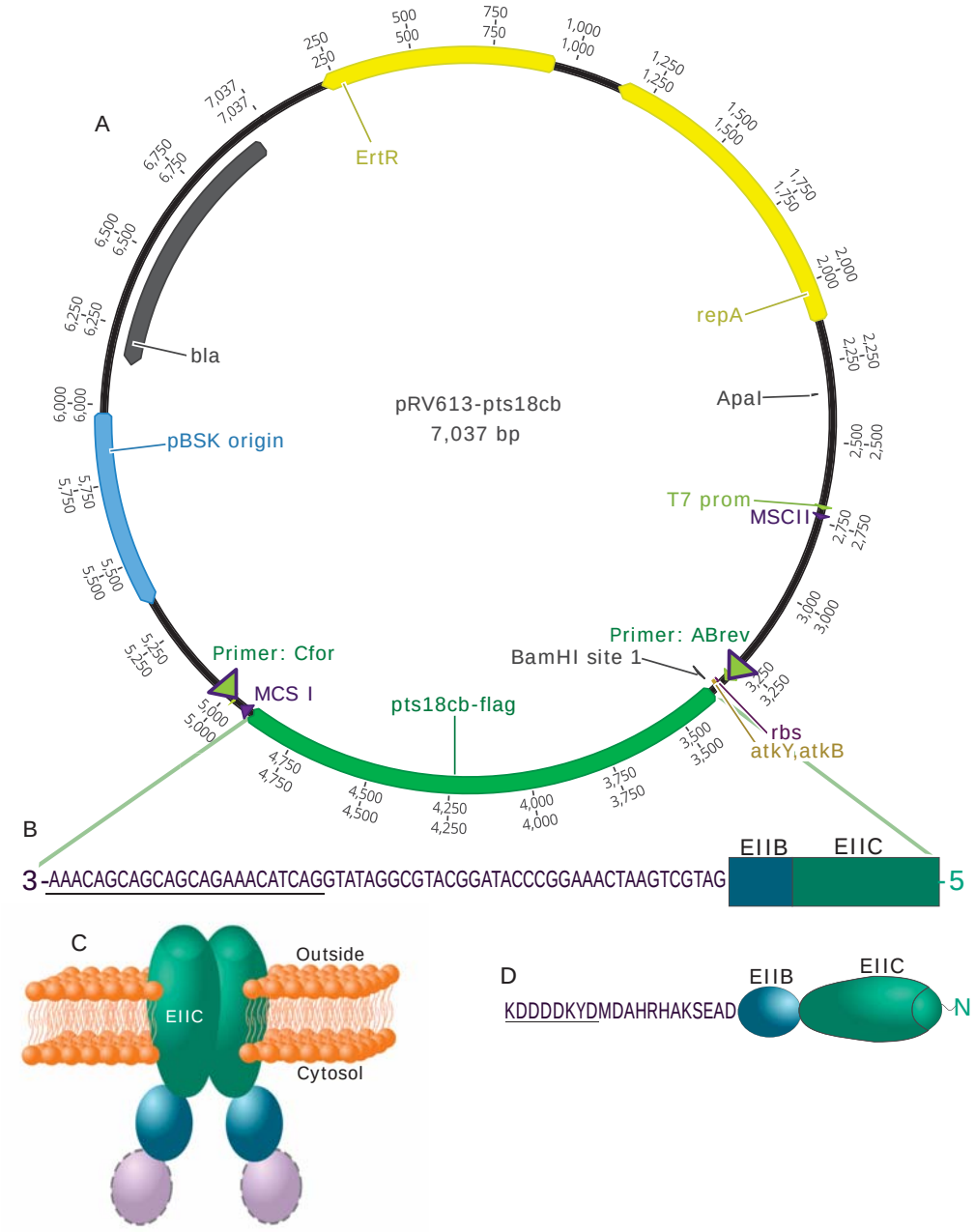


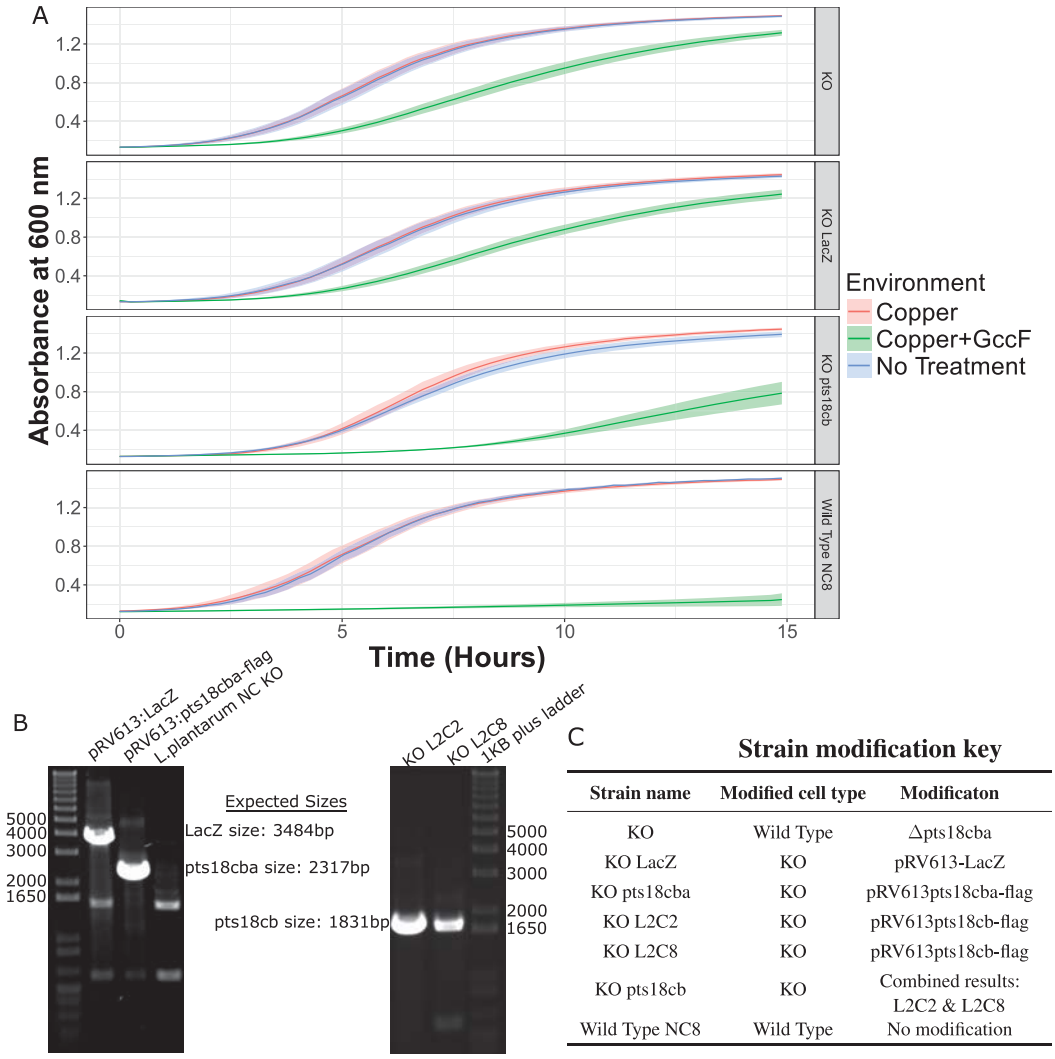
Figure 29: **EIIA deletion: *pts18cb-FLAG* vector map**
(A) Plasmid map of *pRV613-pts18cb-FLAG*: atk and atkY (Gold) are copper inducible promoters, pBSK (Blue) *l. plantarum* origin of replication, bla, repA replication factors (Yellow), rbs the ribosomal binding site (Purple), MSCI and MCSII (Purple arrowheads) are multiple cloning sites. Purple border green arrows represent the binding position of the primers used to confirm transformation success. **(B)** *Pts18cb-FLAG* insert orientation and domain order. The EIIC (Green) and EIIB (Blue) along with the 3′ nucleic acid sequence containing the FLAG-tag (underlined). **(C)** A model for protein dimerisation, orientation and insertion into the membrane (orange) and dotted lines indicating the removal of domain EIIA. **(D)** Model of expected gene product with C terminal FLAG-tag amino acid sequence (underlined).

The PTS18CBA is involved in the GccF mechanism of action, however its function in the mechanism remains unclear. The PTS EII family of enzymes have diverse roles that include ligand transport, and can operate as environmental sensors. The EII^{GlcNAc} could be involved through binding the GlcNAc moieties on GccF that are essential for its activity or *via* the cytosolic components that participate in signal transduction using the phosphorelay or by directly interacting with the downstream signalling proteins. There is sufficient evidence from reports of other EII systems to support that the EII^{GlcNAc} domains can be separated without interrupting independent folding or cause a significant loss of activity. Removing any of the cytosolic domains would disrupt the set-wise bimolecular interactions used by the phosphorelay effectively stalling it. Furthermore, transduction networks involving the EIIA domain would no longer be able to operate and contribute to GccF induced bacteriostasis thereby eliciting a more resistant phenotype.

The transformation of *L. Plantarum* NC8 KO was confirmed by PCR (Figure 30) and sequencing. Deleting the EIIA (Figure 29) domain did not appear to affect the overall growth profile of the *L. plantarum* NC8 cells, nor did the cell exhibit unusual behaviour (Figure 30A). The plasmid containing the truncated gene appear to generate slightly more stress than what was seen for the plasmids containing a full *pts18cba* gene (Figure 31). The KO-pts18cb cells also grew significantly slower than its mock counterpart (pRV613-LacZ) as well. The KO-pts18cb was less fit than the other strains, but had marginal improvement with the addition of copper (Figure 31B). KO-pts18cb still responded to GccF and displayed a change in the cell density that was similar to that seen for plasmids that contained the full gene. This change in growth profiles was not seen in the mock control indicating that the gene inserts are responsible for the change in phenotype.

At eight hours GccF treated KO-pts18cb cells had $\sim$65% less growth than GccF treated KO-LacZ cells (Figure 31B). The truncated gene insert causes an additional $\sim$37% reduction in cell density over normalised GccF treated KO cells which is comparable to the $\sim$36% caused by the full gene. Even while KO-LacZ cells retain some sensitivity to GccF like the KO strain, the presence of at least the EIICB confers greater sensitivity to GccF

than without the truncated protein. Total re-sensitisation back to a wild type phenotype was once again not achieved by this construct, and the EIICBA sensitivity phenotype was not altered with the removal of the EIIA domain, therefore it is unlikely to be involved in the GccF mechanism of action. Removal the EIIA reduces cell density at eight hours by about $\sim$3% when comparing the growth of KO-ptscb cells against KO-pts18cba in all conditions, a result that falls within the experimental error and is therefore not significant.



Strain name	Modified cell type	Modificaton
KO	Wild Type	Δpts18cba
KO LacZ	KO	pRV613-LacZ
KO pts18cba	KO	pRV613pts18cba-flag
KO L2C2	KO	pRV613pts18cb-flag
KO L2C8	KO	pRV613pts18cb-flag
KO pts18cb	KO	Combined results: L2C2 & L2C8
Wild Type NC8	Wild Type	No modification

Figure 30: **EIIA deletion data grouped by strain**
(A) Domain EIIA deletion growth data grouped by strain. Copper (Red): 30 µM of Cu^{2+} was added to MRS media; Copper+GccF (Green): 60 nM of GccF and 30 µM of Cu^{2+} was added to MRS media. No treatment (Blue): Only sterilised MRS was used. Shading represents a 95% confidence interval. **(B)** PCR results amplified from positively transformed *L. plantarum* NC8 strains. Expected insert sizes: LacZ (3484bp) and pts18cba (2317bp), pts18cb (1831bp) using primers Cfor and ABrev. **(C)** Strain modification key: The strains used to generate the data. KO pts18cb is the average of L2C2 and L2C8 *L. plantarum* NC8 KO strains which both contain a *pts18CB-FLAG* gene insert.

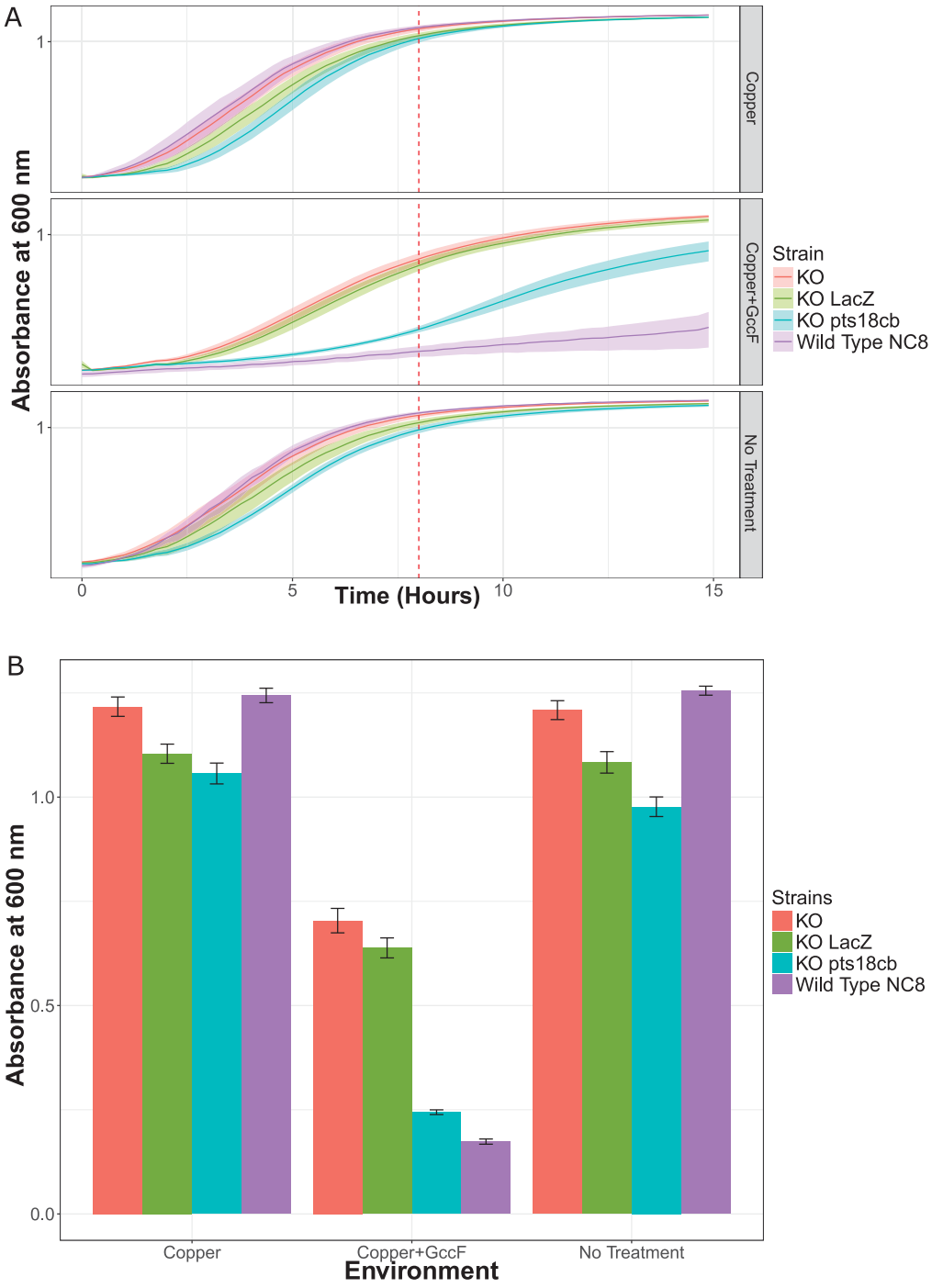


Figure 31: **EIIA deletion data grouped by environment**
(A) EIIA domain deletion growth data grouped by environment, shading represents a 95% confidence interval, the red dotted line indicates the 8 hour interval used for data quantification in B below. **(B)** Bar chart showing quantitative data of cell growth at 8 hours as indicated by the dotted line in growth data above A. Error bars represent the standard error (Table 16). **Key:** KO (Red) pts18cba knockout *L. plantarum* NC8 cells; KO-LacZ (Green) KO cells transformed with the pRV613-LacZ expression vector; KO-pts18cb (Blue) KO cells transformed with the pRV613-pts18cb-flag expression vector; Wild type (Lavender) unmodified *L. plantarum* NC8 cells. unmodified *L. plantarum* NC8 cells. **Growth conditions:** (Copper) MRS media containing a final concentration of 30 µM of Cu^{2+}; (Copper+GccF) MRS media containing final concentration of 60 nM of GccF and 30 µM of Cu^{2+}; (No treatment) only sterilised MRS.

3.4.4 Removal of EIIAB from *pts18cba*

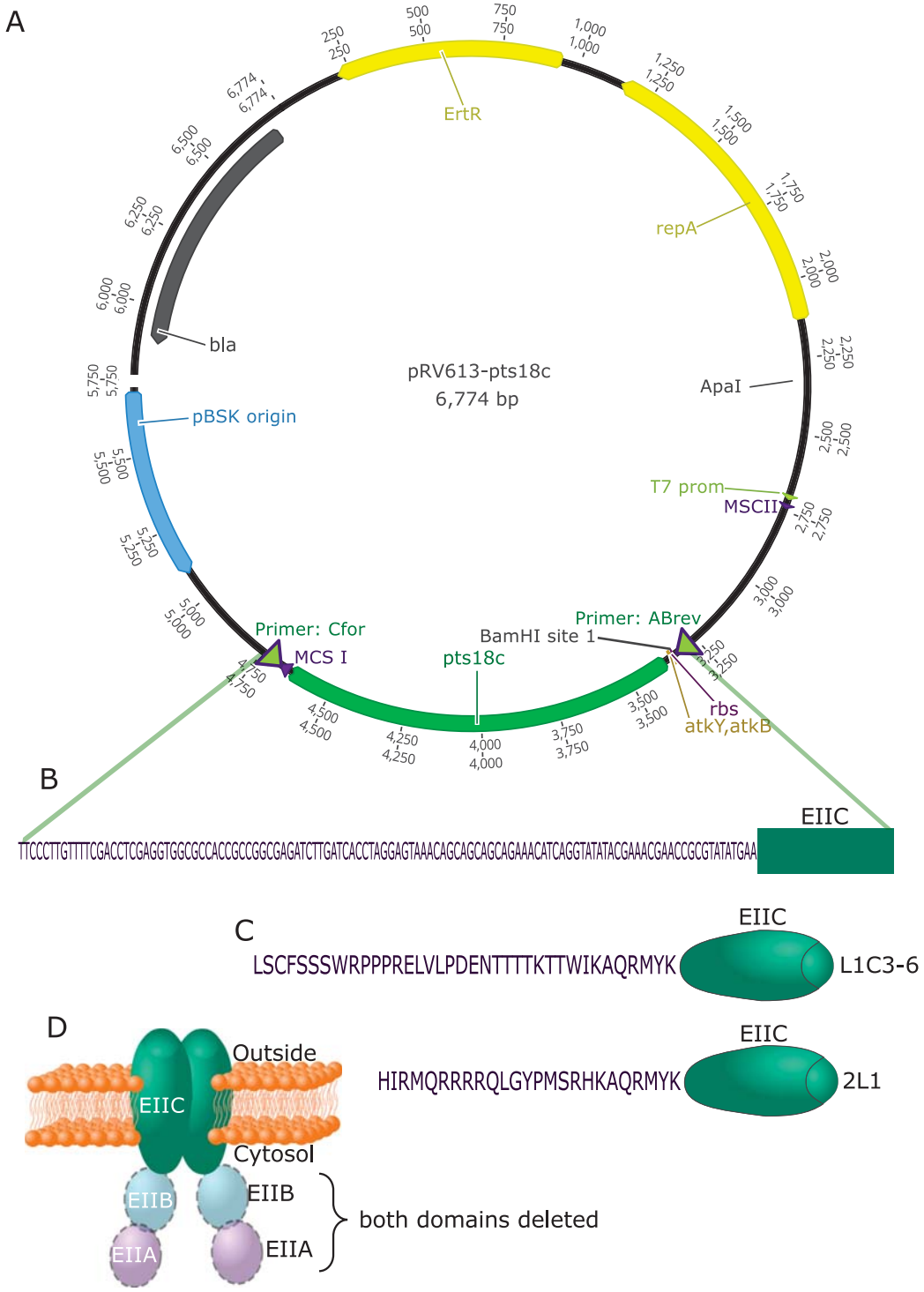
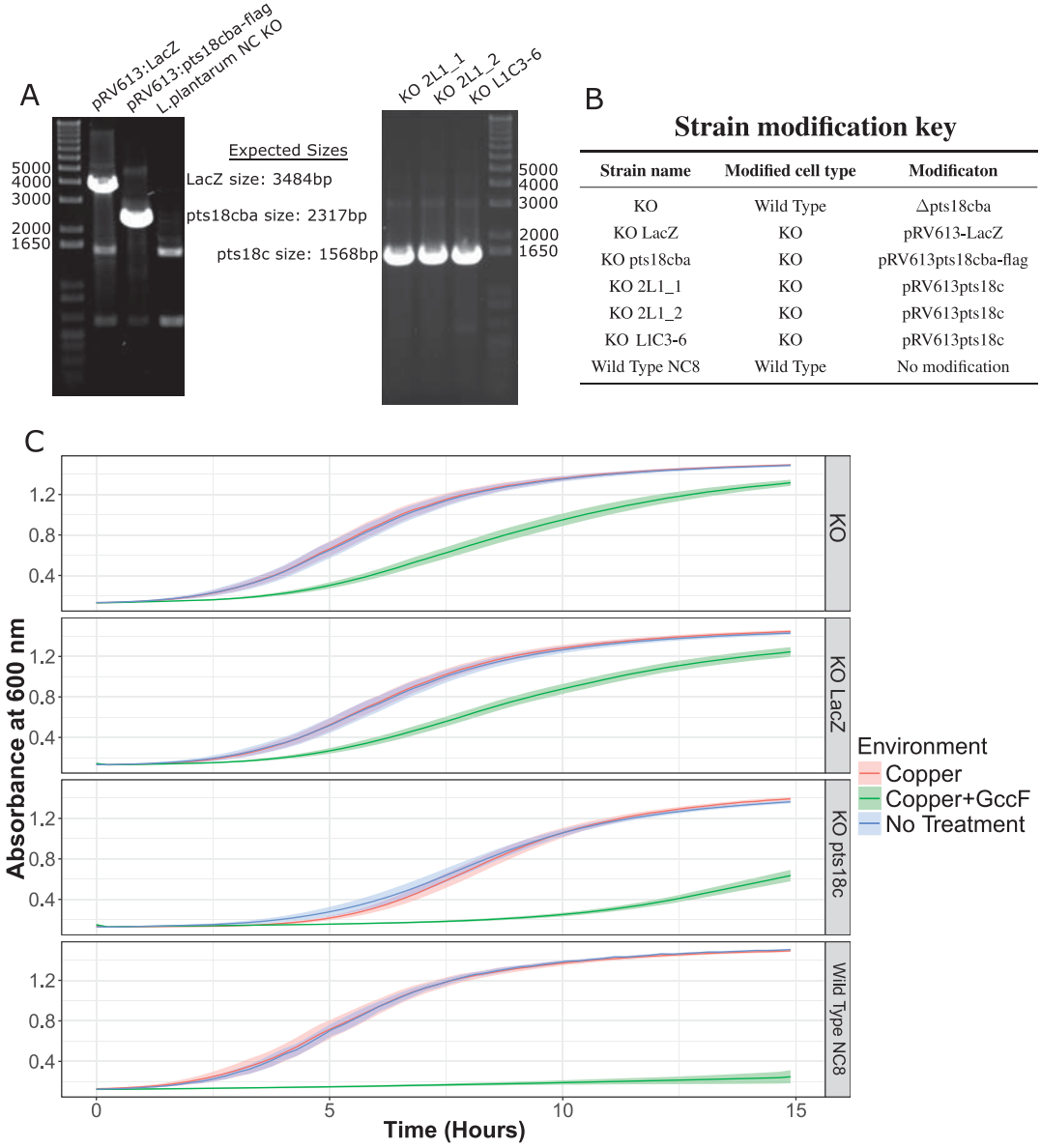


Figure 32: **EIIAB deletion: *pts18c* vector map**
(A) Plasmid map of *pRV613-pts18c*: atk and atkY (Gold) are copper inducible promoters, pBSK (Blue) *L. plantarum* origin of replication, bla, repA replication factors (Yellow), rbs the ribosomal binding site (Purple), MSCI and MCSII (Purple arrowheads) are multiple cloning sites. Purple border green arrows represent the binding position of the primers used to confirm transformation success. **(B)** *Pts18c* insert orientation and EIIC domain (Green). **(C)** Model of the expected gene products from L1C3 and 2L1 constructs both of which contain frameshifts at the *Nde*I site. **(D)** A model for protein dimerisation, orientation and insertion into the membrane (orange) along deleted domains EIIB (Blue), EIIA (Purple).

Transformations were confirmed by PCR and DNA sequencing that revealed no additional changes to the any of the plasmids (Figure 33A). The three constructs were tested independently under the assumption that the tails would affect the growth phenotypes. Two of these constructs contained similar mutations that caused the same frameshift that resulted in the mistranslation of the FLAG-tag. The third construct had a significantly longer tail (Figure 32) which also contained a single cysteine. All three constructs were found to have identical growth in *L. plantarum* NC8 and were averaged. The averaged data representing the three strains were collectively called KO-ptsc.

KO-ptsc growth profiles were similar for both copper and no treatment controls and did not vary from expected phenotype (Figure 33C). However, the removal of both EIIA and EIIB domains appeared to be more toxic to the cells than when only the EIIA domain was removed. Although KO-pts18c cells experienced significantly more stress (loss of cell growth) than the other strains, they still grew. The KO-pts18c cells did not show improved resistance to GccF and experience loss of cell density comparable to the WT cells at eight hours. The significant difference between KO-LacZ and the KO-pts18c growth profiles also indicated that the modified *pts18cba* gene was still sensitising the cells to GccF without either the EIIA or EIIB domains (Figure 34).

At eight hours there was a significant decrease in growth from KO-pts18c cells in both control groups, where this cell type grew roughly half the rate of the wild-type cells (Figure 34). The addition of GccF caused an additional $\sim$41% reduction of KO-pts18c cell density when compared to the normalised GccF treated KO control which was not that dissimilar to the reduction cause by either the full gene or when the EIIA domain was removed. The toxicity of the modified gene was calculated to have reduced growth by $\sim$10% more than the plasmid containing the full gene. While the toxicity of this construct exceeded experimental error, it is not high enough to generate the observed re-sensitisation phenotype when cells are treated with GccF. Despite the additional toxicity, neither the EIIA nor the EIIB appears to be involved in the GccF mechanism of action, and only the EIIC domain is needed to enhance GccF susceptibility.



Strain name	Modified cell type	Modificaton
KO	Wild Type	Δpts18cba
KO LacZ	KO	pRV613-LacZ
KO pts18cba	KO	pRV613pts18cba-flag
KO 2L1_1	KO	pRV613pts18c
KO 2L1_2	KO	pRV613pts18c
KO L1C3-6	KO	pRV613pts18c
Wild Type NC8	Wild Type	No modification

Figure 33: **EIIAB deletion data grouped by strain**
(A) PCR amplified plasmid inserts from positively transformed *L. plantarum* NC8 strains. Expected insert sizes: LacZ (3484bp) and pts18cba (2317bp), pts18c (1568bp). **(B)** Strain modification key: KO pts18c is the average of 2L1_1, 2L2_2 and L1C3-6 which are *L. plantarum* NC8 strains containing plasmids harbouring a modified *pts18cba* gene from which the sequence that encodes for the EIIA and EIIB domains had been removed. **(C)** EIIAB deletion growth data grouped by strain; Copper (Red): MRS containing 30 µM of Cu^{2+}; Copper+GccF (Green): MRS made up with 60 nM of Glycocin F and 30 µM of Cu^{2+}; No treatment (Blue): Only sterilised MRS. Shading represents a 95% confidence interval.

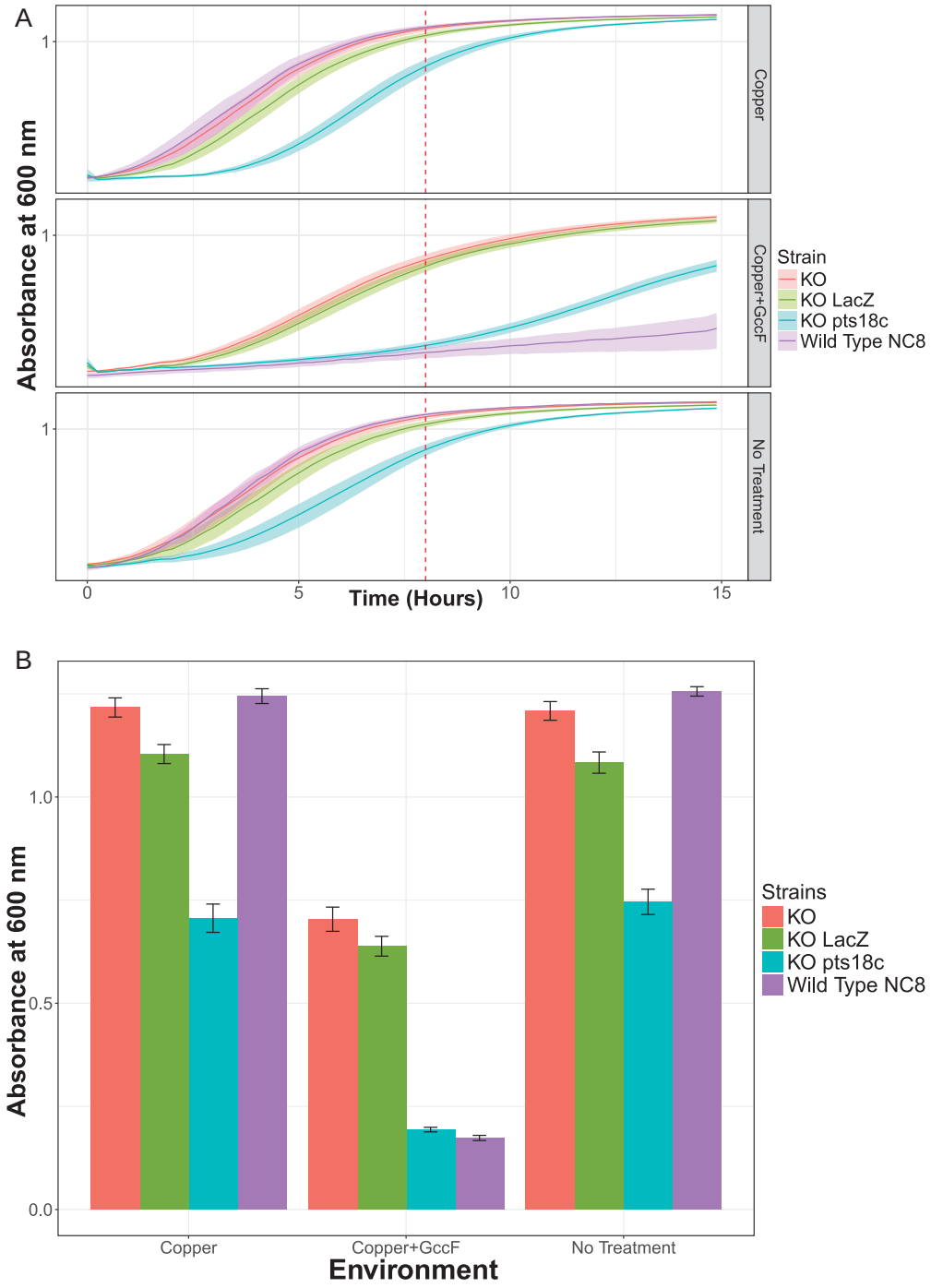


Figure 34: **EIIAB deletion data grouped by environment**
(A) EIIAB deletion growth data grouped by environment, where shading represents a 95% confidence interval, the red dotted line indicates the 8 hour interval used for data quantification in B below. **(B)** Bar chart showing quantitative data of cell growth at 8 hours as indicated by the dotted line in growth data above A. Error bars represent the standard error (Table 16). **Key:** KO (Red) *L. plantarum* NC8 *pts18cba* knockout cells without a plasmid; KO-LacZ (Green) KO cells transformed with the pRV613-*LacZ* expression plasmid; KO-pts18c (Blue) KO cells transformed with one of the pRV613-*pts18c* expression plasmids (2L1_1, 2L1_2 and L1C3-6 growth data); Wild type (Lavender) unmodified *L. plantarum* NC8 cells. **Growth conditions:** (Copper) MRS media containing a final concentration of 30 µM of Cu^{2+}; (Copper+GccF) MRS media containing a final concentration of 60 nM of Glycocin F and 30 µM of Cu^{2+}; (No treatment) only sterilised MRS.

3.4.5 Site directed mutagenesis: His584Ala

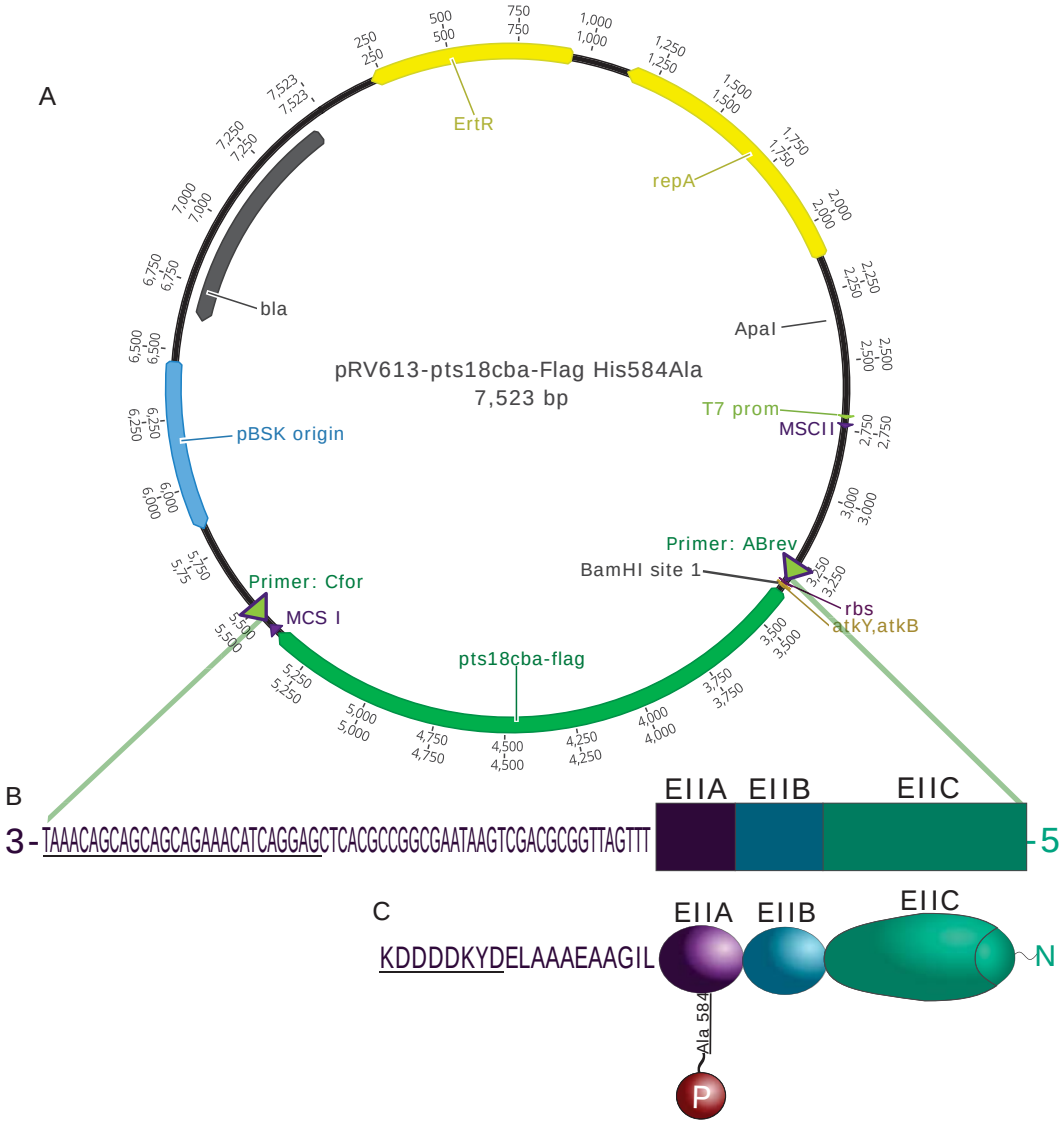


Figure 35: **His584Ala plasmid map**
(A) Vector map of the 7.5 kb pRV613-*pts18cba-FLAG*: His584Ala plasmid. This vector contains two adjacent SNPs, C→G and A→C that change Histidine 584 to an Alanine shown in Figure 37. Copper inducible promoters atk and atkY (Gold), pBSK (Blue) *L. plantarum* origin of replication, bla, repA replication factors (Yellow), rbs the ribosomal binding site (Purple), MSCI and MCSII (Purple arrowheads) are multiple cloning sites. Purple border green arrows represent the binding position of the primers used to confirm transformation success. **(B)** Graphical representation of the gene layout, domain order and orientation in the plasmid along with 3′ FLAG sequence (underlined). **(C)** Diagrammatic representation of the expected protein product of induced expression including the amino acid tail and FLAG-tag (underlined). The model is presented C-terminal to N-terminal to be consistent with the gene layout in the plasmid. Modified residue H584A (underlined) was arbitrarily placed within the EIIA domain with attached phosphate (Red).

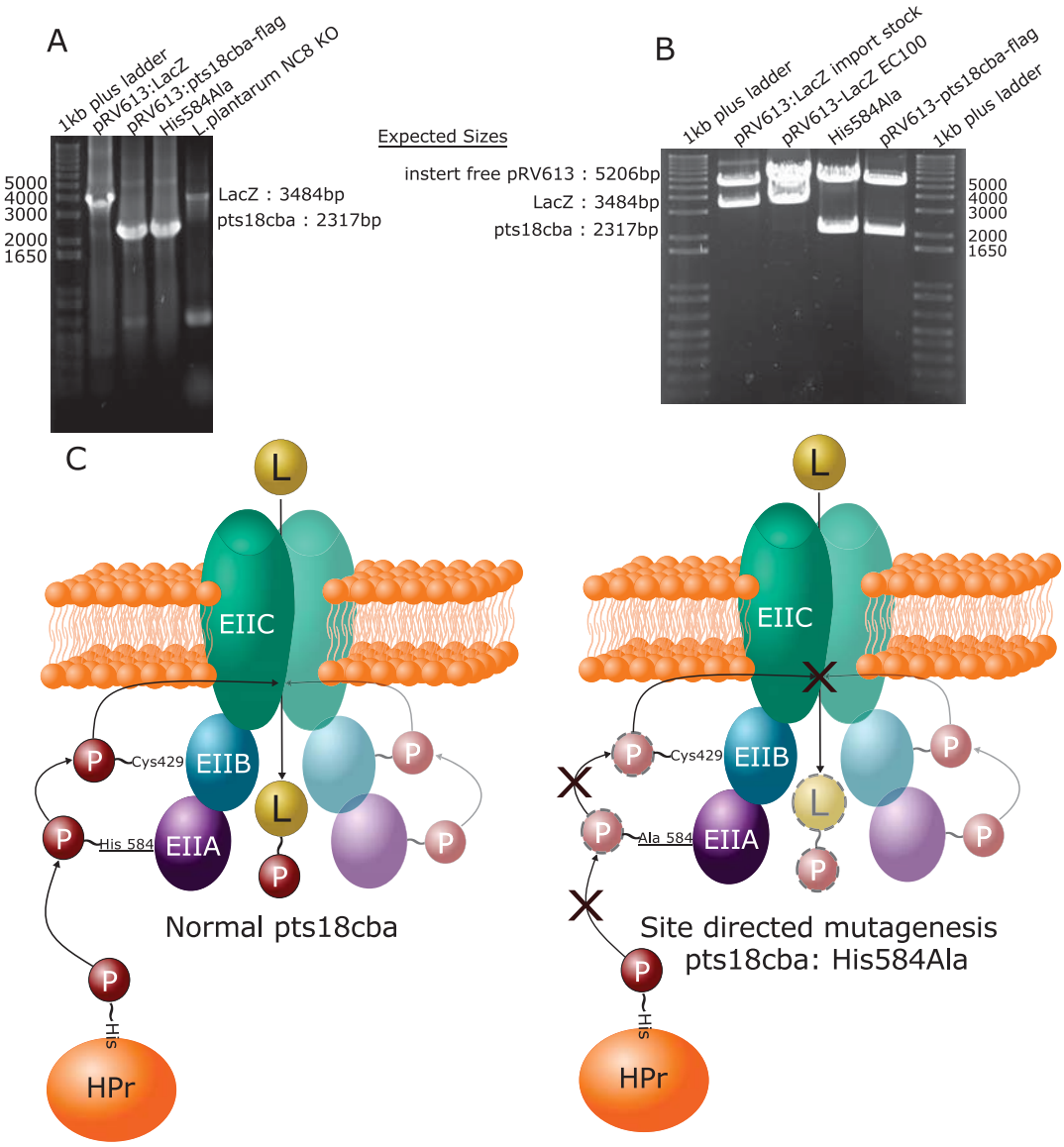


Figure 36: **His584Ala transformation confirmation and Phospho-inhibition model** **(A)** Transformation confirmation using PCR and expected product sizes. PCR product were sequenced for the final identity (Figure 37 vi). **(B)** *BamH*I analytical restriction digest of plasmids extracted from stored stocks or transformed EC100 cells that were sequenced and used as plasmid stocks for site directed mutagenesis of Histidine 584. pRV613:LacZ import stock were plasmid stocks from previous studies (LacZ positive control). PRV613-LacZ EC100 was plasmid preparation from transformed EC100 cells containing pRV613-LacZ import stocks for storage and further plasmid production. His584Ala is the pRV613-*pts18cba-FLAG* after site directed mutagenesis of His584. **(C)** Diagrammatic representation of the hypothesised outcome of altering histidine 584 to an alanine 584 causing the disruption of the phosphorelay (phosphate: Red), along with suspected membrane orientation, dimerisation, and downstream HPr → EIIA interaction.

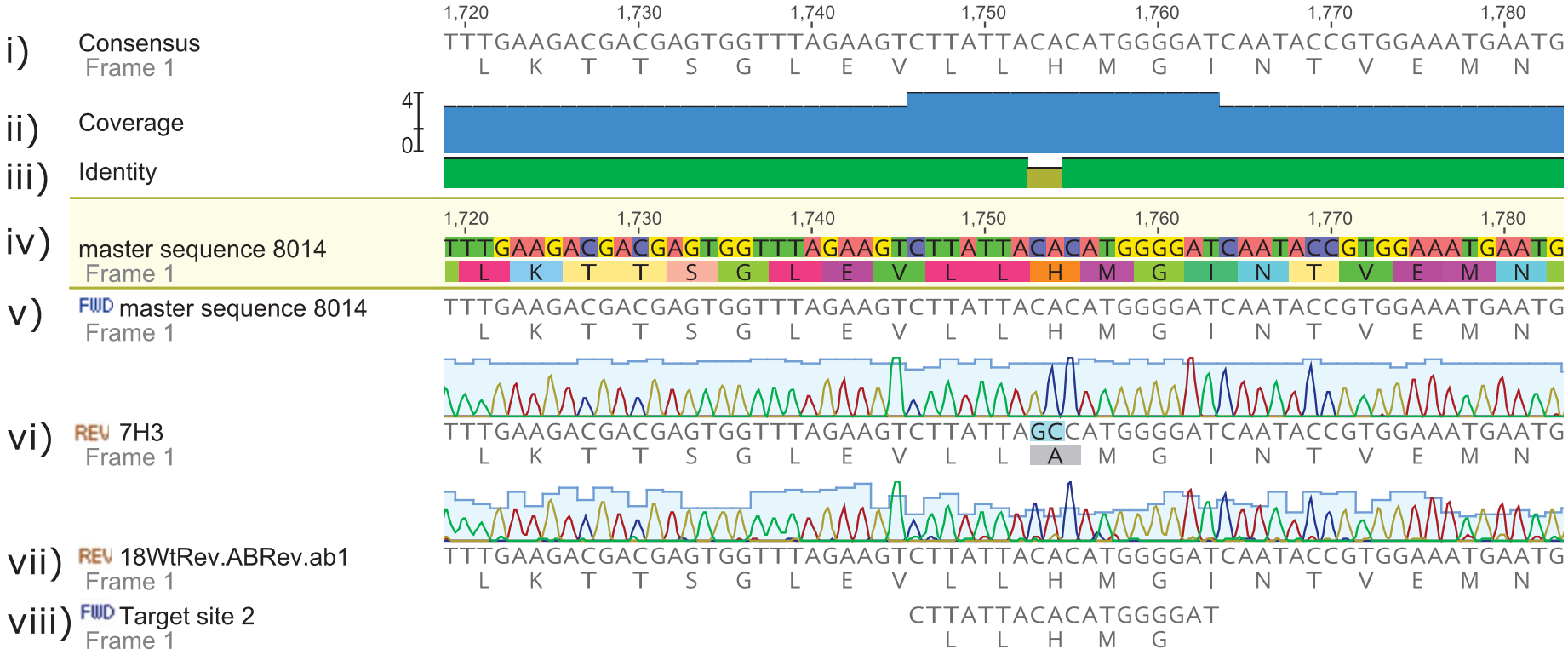


Figure 37: **His584Ala sequencing results and alignment**
Sequence alignment results of PCR products from confirmed *L. plantarum* NC8 KO cells transformed with His584Ala mutant pRV613-*pts18cba-FLAG* plasmids. Shown as nucleic acids sequences, sequencing chromatographs and the first open reading frame amino acid translation. **i)** Consensus sequences: the average identity nucleic acid and first frame amino acid identities. **ii)** Total coverage from overlapping contigs. **iii)** Identity confidence, lower confidence regions indicate changes in nucleic acid sequences. **iv)** Master sequence: *L. plantarum* 8014 master sequence downloaded from NCBI focused on the EIIA region that contains His584. **v)** Sequencing results of *L. plantarum* 8014 wild type cells using the EIIA forward primer. **vi)** Sequencing results of pRV613-pts18cba-FLAG His584Ala PCR transformation check using the EIIA reverse primer. **vii)** Sequencing results of *L. plantarum* 8014 wild type cells using the EIIA reverse primer. **viii)** Target site taken from primer design to show mutation sites using the wild type sequence.

Histidine 584 is predicted to be a highly conserved amino acid that forms part of the PTS^{GlcNAc} phosphorelay. Due to the nature of the bimolecular interactions and their specificity, it is unlikely that the relay could function without this step or that other PTS transporters would be able to rescue the phosphorelay making histidine 584 an ideal target for stalling the PTS^{GlcNAc} phosphorelay. Overlapping primers were used to change the CAC codon at 1753-1754 to GCC altering the two least variable nucleotides (C→G and A→C) to avoid translation confusion when expressing the changes that will create the alanine instead of the histidine (Figure 35). The genetic equivalent of chemically removing the imidazole group from the histidine by using site directed mutagenesis (SDM). Alanine is a good choice because it cannot be phosphorylated and has a neutral charge.

Analytical restriction digests were used to confirm plasmid stocks before SDM (Figure 36B). Successful *L. plantarum* NC8 transformations of mutation containing plasmids was determined by PCR (Figure 36A) and the identity of the plasmids confirmed by sequencing (Figure 37). The sequencing results showed that the two nucleotide changes to the *pts18cba-FLAG* gene were the only detectable changes to the pRV-613:pts18cba-FLAG plasmid (Figure 37).

The growth profiles showed no difference between copper and no treatment conditions and did not deviate from the expected growth behaviour seen before. KO-H584A appears to be the weakest of the tested cells, and performs slightly worse than KO-pts18cba. This subtle difference in performance was caused by a two nucleic acid substitution and was taken as a sign that the PEP-PTS phosphorelay was successfully stalled. Both KO-pts18cba and KO-H584A appear to be susceptible to GccF with similar profiles for the entire 15-hour period of recorded growth (Figure 38).

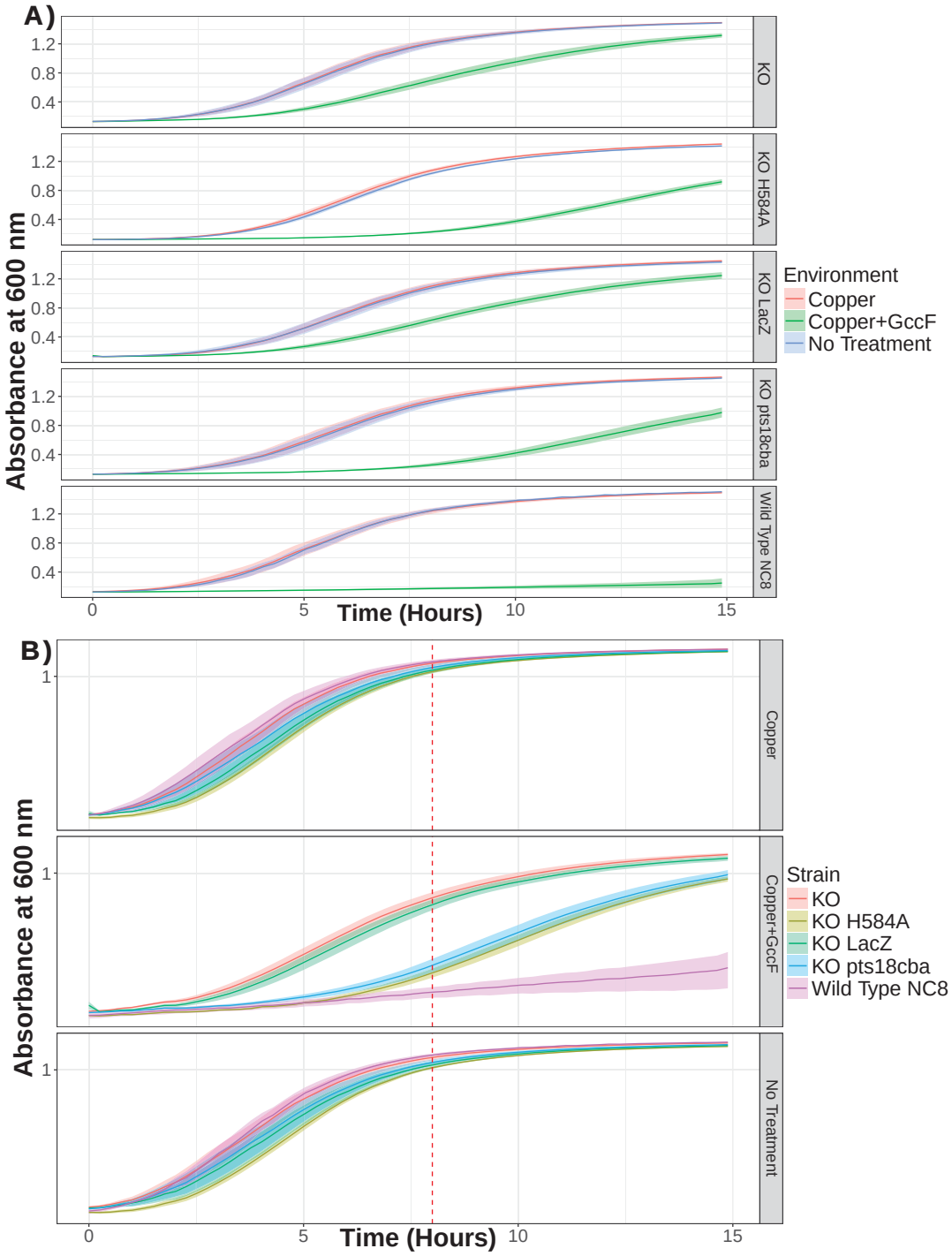


Figure 38: **H584A: growth curves grouped by environment**
(A) Growth data grouped by strain. **(B)** Growth data grouped by environment. The red dotted line indicates the eight hour mark used for qualitative analysis (Figure 39). Shading represents a 95% confidence interval. **Environmental conditions:** (No treatment) Sterile MRS media; (Copper) MRS media supplemented with 30µM of Cu^{2+}; (Copper+GccF) MRS supplemented with 30µM of Cu^{2+} and 60 nM of GccF. These conditions were tested against ***L. plantarum*** **NC8 cells lines:** KO - *pts18cba* knockout (Red); KO-H584A - KO mutants transformed with pRV613-pts18cba-FLAG (H584A) (khaki); KO-LacZ - KO mutants transformed with pRV613-LacZ (Green); KO-pts18cba - KO mutants transformed with pRV613-pts18cba-FLAG (Blue); and Wild type - unmodified *L. plantarum* NC8 (Lavender).

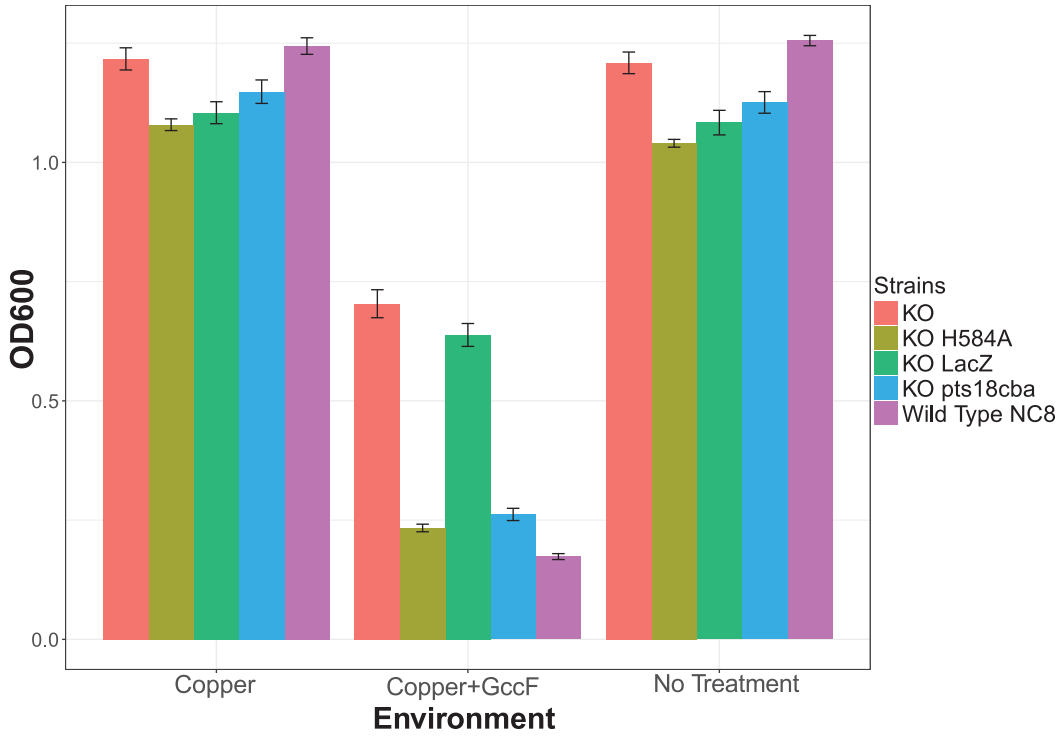


Figure 39: **His584Ala quantitative phenotype analysis at eight hours**
Growth data taken at eight hours. **Environmental conditions:** (No treatment) Sterile MRS media; (Copper) MRS media supplemented with 30µM of Cu^{2+}; (Copper+GccF) MRS supplemented with 30µM of Cu^{2+} and 60 nM of GccF. These conditions were tested against ***L. plantarum*** **NC8 cells lines:** KO - *pts18cba* knockout (Red); KO-H584A - KO mutants transformed with pRV613-pts18cba-FLAG (H584A) (khaki); KO-LacZ - KO mutants transformed with pRV613-LacZ (Green); KO-pts18cba - KO mutants transformed with pRV613-pts18cba-FLAG (Blue); and Wild type - unmodified *L. plantarum* NC8 (Lavender). Error bars represent the standard error (Table 16).

GccF treated KO-H584A cells were significantly more sensitive to GccF than the KO and KO-lacZ strains (Figure 39). When treated with GccF the KO-His584Ala cells experienced a ∼78% drop in cell density at eight hours when compared to the KO-His584Ala copper control. This reduction in growth exceeded the mock control, and could not be explained by experimental error, the plasmid stress, or any toxicity generated by the insert. This demonstrated that the gene was still functioning in the GccF mechanism of action and thus providing additional evidence ruling out the phosphorelay as a transduction tool. The His584Ala gene appeared to contribute ∼38% of the observed phenotype at eight hours, which was far more than the plasmid strain which was calculated to be ∼3% leaving the normalised His584Ala value closely comparable to the unmodified gene (36%).

3.4.6 Direct construct comparison

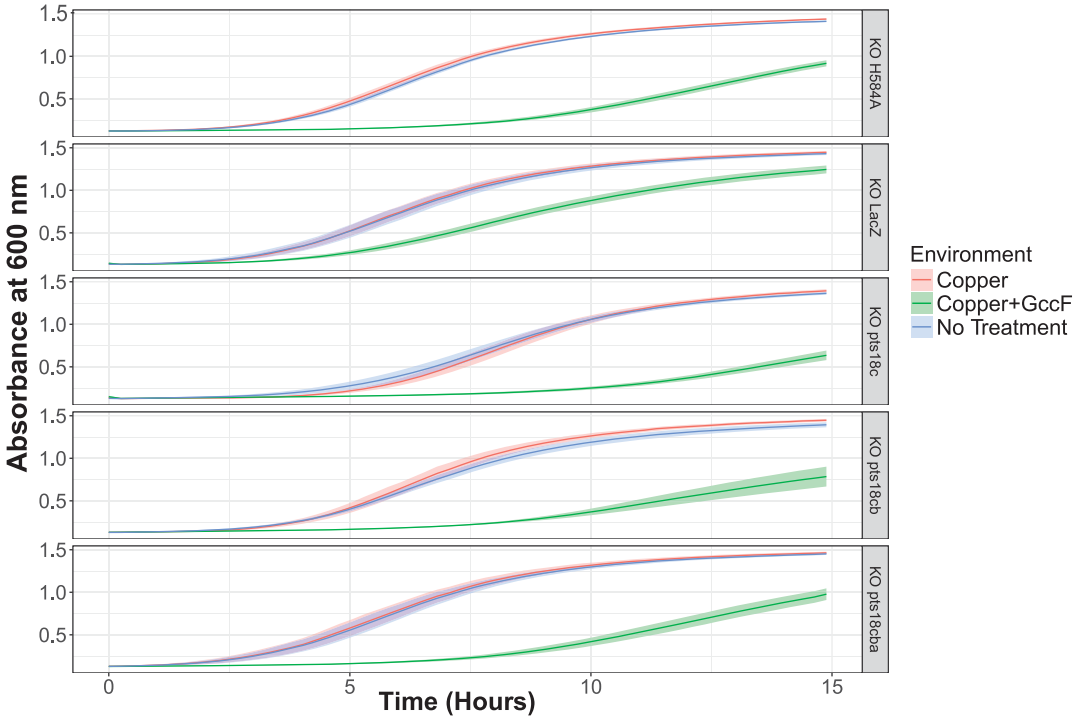


Figure 40: **Direct construct comparison by strain**
Growth data grouped by strain. Shading represents a 95% confidence interval. **Environmental conditions:** (No treatment) sterile MRS media; (Copper) MRS supplemented with 30µM of Cu^{2+}; (Copper+GccF) MRS supplemented 30µM of Cu^{2+} and 60 nM of GccF. **Bacterial strains**: The strains were generated using the *L. plantarum* NC8 *pts18cba* knockout mutant and transformed with the relevant plasmid. **KO-pts18cba** contains the pRV613-*pts18cba-FLAG*, **KO-H584A** was transformed with pRV613-*pts18cba-FLAG* (H584A), **KO-LacZ** harbours the pRV613-*LacZ*, **KO-pts18cb** contains the EIIA deletion construct pRV613-*pts18cb-FLAG* and **KO-ptsc** is the averaged results from three isolates containing various version of the EIIAB deletion construct pRV613-*pts18c.* from separate transformation events.

Both KO-pts18cb and KO-H584A strains grew slightly better with the addition of copper, however none of the strains appeared to have significant difference between the copper and no treatment control over the entire growth period (Figure 40). All the cells did appear to be susceptible to GccF regardless of the gene insert, albeit the KO-LacZ strain was significantly more resistant to GccF (Figure 40). All the cells that contained a *pts18* related gene insert except *KO-pts18c* had very comparable growth profiles, with a long initial lag phase and late alleviation of inhibition.

Grouping the data by environment shows that there is no statistical difference between

the growth profiles of KO-H584, KO-pts18cba, and KO-pts18cb for most of the growth period for all conditions (Figure 41A). pRV613-*ptsc* plasmid had a significantly stronger effect on the cells than the other plasmids preventing the harbouring cells from reaching comparable cell densities as the other strains. A reduction in cell fitness caused by plasmids harbouring *pts18* variants appeared to increase linearly with the removal of the domains at the eight hours time point (Figure 41B).

Directly comparing growth data from all the pts18 constructs demonstrated that the inclusion of any gene variants that contained the *pts18* EIIC domain sequence when exposed to GccF lost on average $\sim$38% more cell density than the GccF treated KO-LacZ control regardless of the modifications to the EIIA or EIIB domain sequence. However replacing the *pts18cba* gene with *LacZ* provides a more resistant phenotype, indicating that only the EIIC domain is required for increased sensitivity to GccF (Figure 41B).

Mass spectroscopy provided some evidence that the *pts18cba* gene product reached the protein level during earlier work[67]. ATP efflux using bioluminescence assay indicated that small biological molecules were not escaping the cells when exposed to GccF[66]. Microscopy of *L. plantarum* 8014 cells treated with Fluorescein isothiocyanate (FITC) labelled GccF showed that it did not enter the cells, and that the cells did not die (Appendix: Figure 54). Combined with these data it is likely that GccF is interacting with the folded $EIIC^{GlcNAc}$ domain at the outer leaflet of the membrane. The low copy number of the pRV613-*pts18cba-FLAG* plasmid and low concentrations of the PTS^{GlcNAc} this plasmid produces makes it doubtful that GccF acts like a molecular staple. The disruption of the phosphorelay and domain deletions both indicated that transduction does not occur through the PTS^{GlcNAc} using any known means. The source of $\sim$43% reduction in growth of KO and KO-LacZ strains when exposed to GccF has yet to be discovered. It appears that the GccF does not use duel mechanisms of action at 60 nM, but rather this mystery mechanism and the PTS18 EIIC domain work in symphony to generate a bacteriostatic phenotype.

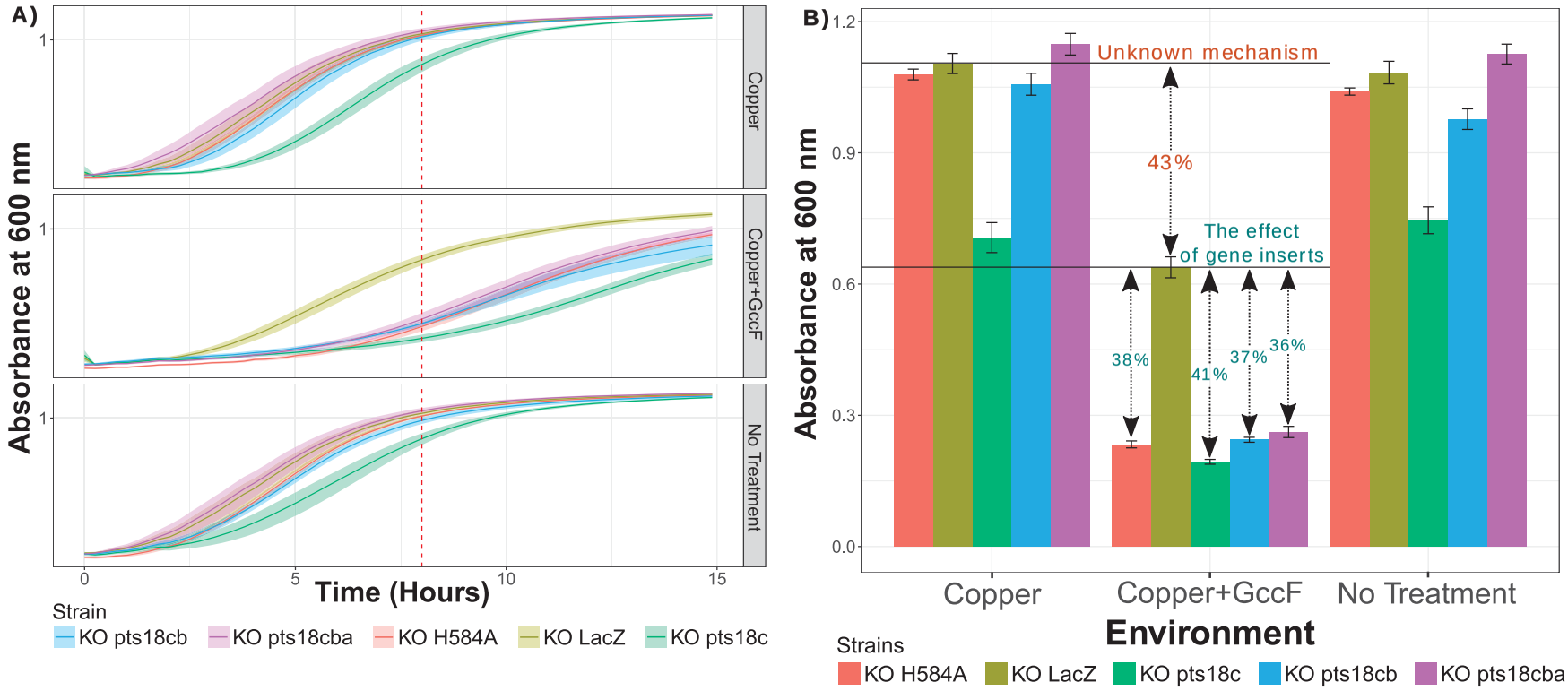


Figure 41: **The effect of the gene inserts grouped by Environment**

(A) Growth data grouped by environment, shading represents the 95% confidence interval. The red dotted line indicates the 8 hour mark used for qualitative analysis in bar chart below. **(B)** Growth data grouped by environment at 8 hours. **(Key)** Environmental conditions: (No treatment) sterile MRS media; (Copper) MRS supplemented with 30µM of Cu^{2+}; (Copper+GccF) MRS supplemented 30µM of Cu^{2+} and 60 nM of GccF. **Cell strains**: The strains were generated using the *L. plantarum* NC8 knockout mutant and transformed with the relevant plasmid. **KO-pts18cba** contains the pRV613-*pts18cba-FLAG*, **KO-H584A** was transformed with pRV613-*pts18cba-FLAG* (H584A), **KO-LacZ** harbours the pRV613-*LacZ*, **KO-pts18cb** contains the EIIA deletion construct pRV613-*pts18cb-FLAG* and **KO-ptsc** is the averaged results from three isolates containing various version of the EIIAB deletion construct pRV613-*pts18c*. Percentages indicate the relative contribution of each construct towards the phenotype with respect to the KO-LacZ copper control.

3.5 GccH as the immunity gene for the GccF cluster

3.5.1 GccH *in silico* analysis

Bacteriocin immunity genes are commonly found within the gene cluster that control bacteriocin production. None of the identified genes in the *gcc* clusters are likely to produce resistance to GccF, however the cluster does have two unidentified genes, *gccI* and *gccH*. NCBI database searches revealed that the only organism that contains the *gccH* gene is *L. plantarum* KW30 and its closest homologue was *asmH* (GenBank: AOF43519.1 AsmH) found in *L. plantarum* A-1 and only had a 66% similarity. Both the *asm1* and *gcc* gene clusters contain a version of the *gccH* gene whereas *gccI* only appears in the *gccF* cluster making *gccH* the most likely candidate for the glycocin F immunity gene. No known conserved domains, leader peptide, export sequences or protein structures associated to the *gccH* could be found. However, a broader NCBI search revealed that there are 24 similar hypothetical genes that are unidentified, and interestingly are all limited to *L. plantarum*. *GccH* appears to be novel, unstudied, and its only homologue is also located in a gene cluster that produces a di-glycosylated bacteriocin that has close structural homology to GccF.

A)

TMPred: all possible inside to outside

From	To	Score	Center
5 (5)	25 (25)	888	16

TMPred: all possible outside to inside

From	To	Score	Center
5 (12)	37 (28)	529	20

B) GccH: TMpred hydrophobicity plot

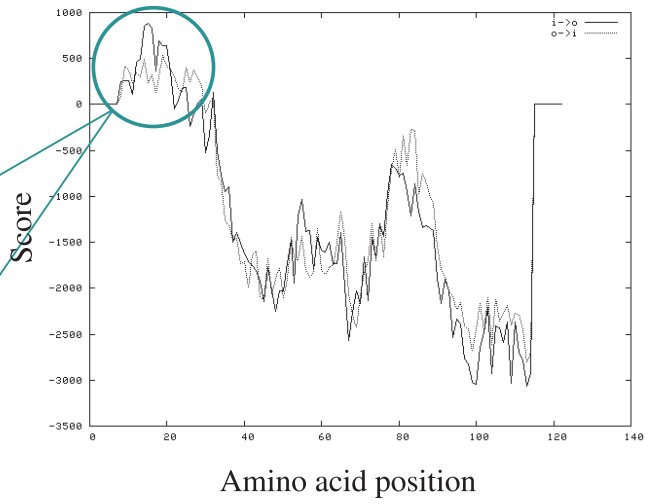


Figure 42: **GccH: Transmembrane prediction and hydrophobicity plot**
(A) TMpred prediction of all possible transmembrane helices and their orientations with respect to a membrane. **(B)** A TMpred hydrophobicity plot of the amino acid position against its relative hydrophobicity score used to determine the reliability of helical formations.

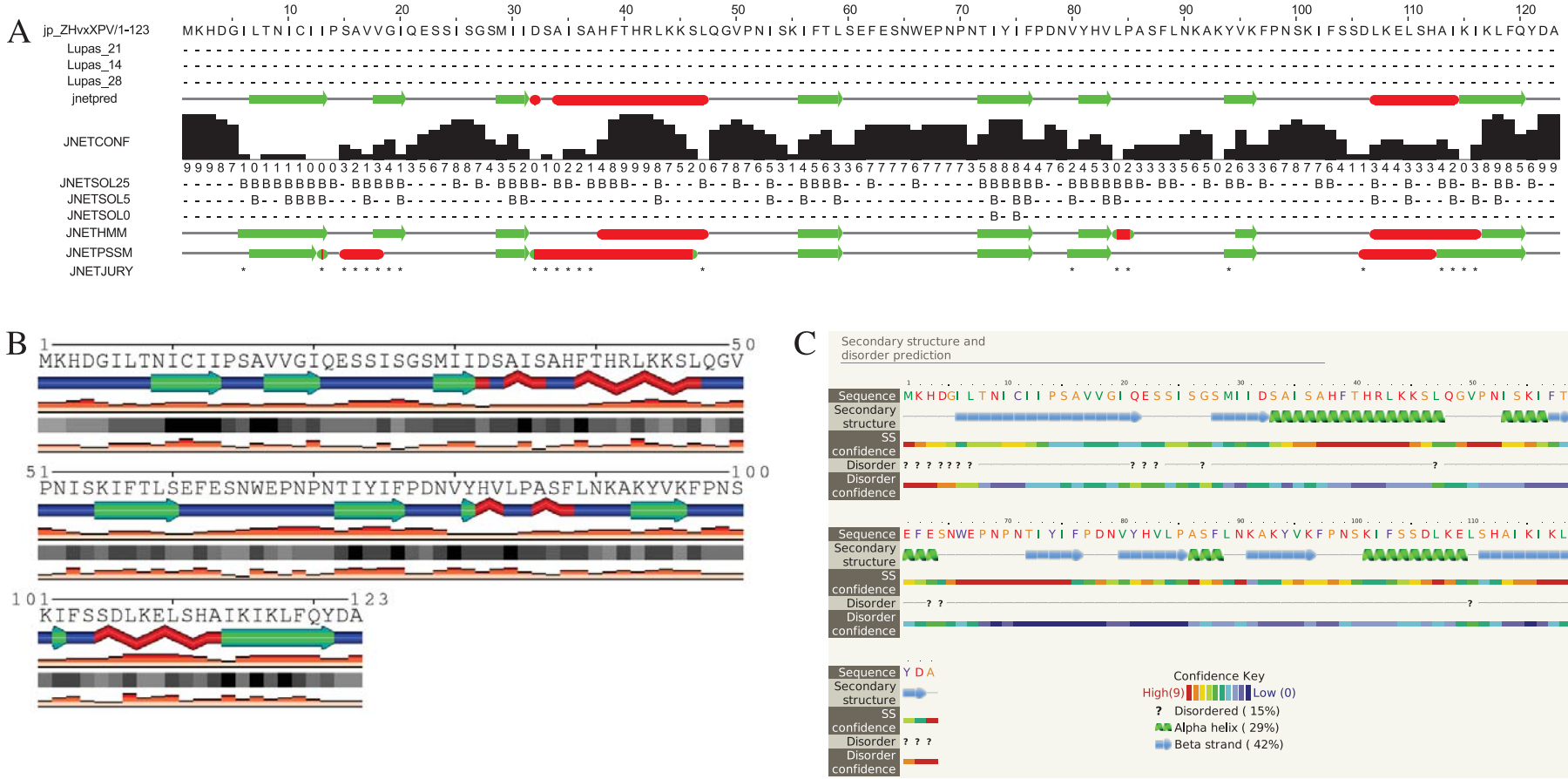


Figure 43: **GccH: various secondary structural predictions**

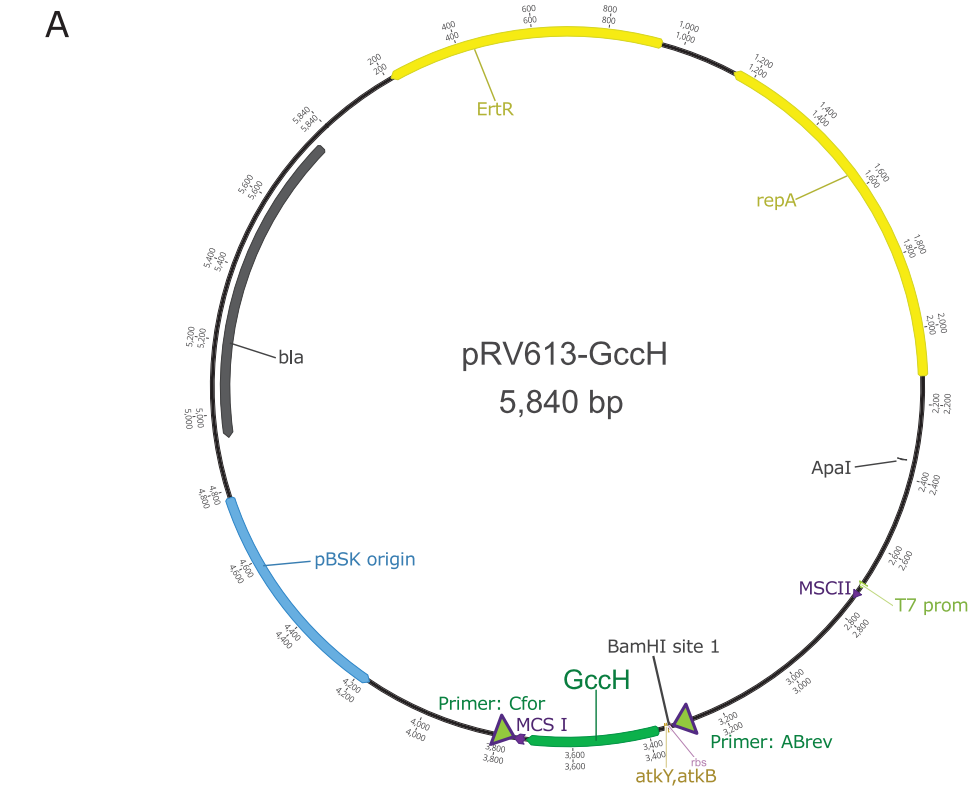
(A) Jpred prediction output red bars indicate α-Helices, green arrows represent β-strands, JNETCONF is the confidence by which each call was made, and B represents buried residues. **(B)** MINNOU prediction output red bars indicate α-Helices, green arrows represent β-strands. The amino acid sequence was programmed to highly lipid accessible in yellow. **(C)** Phyre^2 prediction output Green indicate α-Helices, blue arrows represent β-strands. The disorder is represented by ?.

The *in silico* translated sequence was analysed using two programs from the ExPASy website, ProtParam[150,151] and Compute pI/Mw[151–153]. The results showed that the protein would consist of 123 amino acids, with a molecular weight of 13.79 kDa and a theoretical pI of 8.69. The stability index of GccH was calculated to be 33.63, which is on the border of instability (>40), and the GRAVY score was calculated to be 0.011 which is an indicator that GccH has few hydrophobic amino acids. The *gccH* sequence indicates that it is likely a cytosolic protein, which is supported by TMpred[156] (Figure 42) that was only able to predict one possible transmembrane helix, with a poor score, while all the TOPCONS[157] servers found no transmembrane helices or any homology to known transmembrane proteins. The sequence was submitted to MINNOU[158], Jpred[159] and Phyre2[160] for secondary structure predictions (Figure 43). These packages had good consensus, and show that the protein is likely to be ordered. The web servers predict a structure made up of between seven and nine β-strands and two to three helices.

3.5.2 GccH *in vivo* phenotyping

The *gccH* gene was isolated from *L. plantarum* KW30 and cloned into the pRV613 vector as described in section 2.4.6 (Figure 44A). The successful transformation of *L. plantarum* wild type (WT) and *pts18cba* knockout cells (KO) were confirmed by PCR (Figure 44B). The PCR products were sent for sequencing, and the results aligned perfectly with the *gccH* sequence download from NCBI.

Both the copper supplemented and the no treatment control had similar growth profiles for all tested strains (Figure 45A). The *L. plantarum* NC8 WT strain transformed with pRV613:*GccH* (WT-GccH) experienced some minor growth improvement with the addition of copper. However, confidence intervals indicated that the cells did not behave abnormally and were in line with exceptions from the copper toxicity study (Section 3.3.2.3). When exposed to 60 nM of GccF KO, WT and KO-LacZ cell growth was initially attenuated in contrast to the cells containing plasmids harbouring the *GccH* gene (Figure 45). Interestingly, there is no significant difference in the growth profile of KO-GccH regardless of treatment, this cell line is completely immune to GccF (Figure 45A).



C

Strain modification key

Strain name	Modified cell type	Modificaton
KO	Wild Type	Δpts18cba
KO LacZ	KO	pRV613-LacZ
KO pts18cba	KO	pRV613pts18cba-flag
KO GccH	KO	pRV613:GccH
WT GccH	WT	pRV613:GccH
WT	Wild Type	No modification

D

Primer sequences

pRV613-pts18-Cfor: GGATAACAATTTCACACAGGAAACAGCTATAGACC

pRV613-pts18-ABrev: TCAGACGGATATTTTGCAGTTACAACAATTACTGACA

Figure 44: **GccH plasmid map and confirmation of transformation**
(A) Vector map for pRV613-GccH **(B)** PCR confirmation that *L. plantarum* NC8 wild type and pts18cba knockout cell were successfully transformed with pRV613-GccH (634 bp). These PCR products were sequenced to confirm the identity of the inserts. **(C)** Strain modification key for all the cell used to test the GccH phenotype including controls. **(D)** The PCR and sequencing primers that were used to confirm that the cells were successfully transformed with pRV613-GccH plasmids.

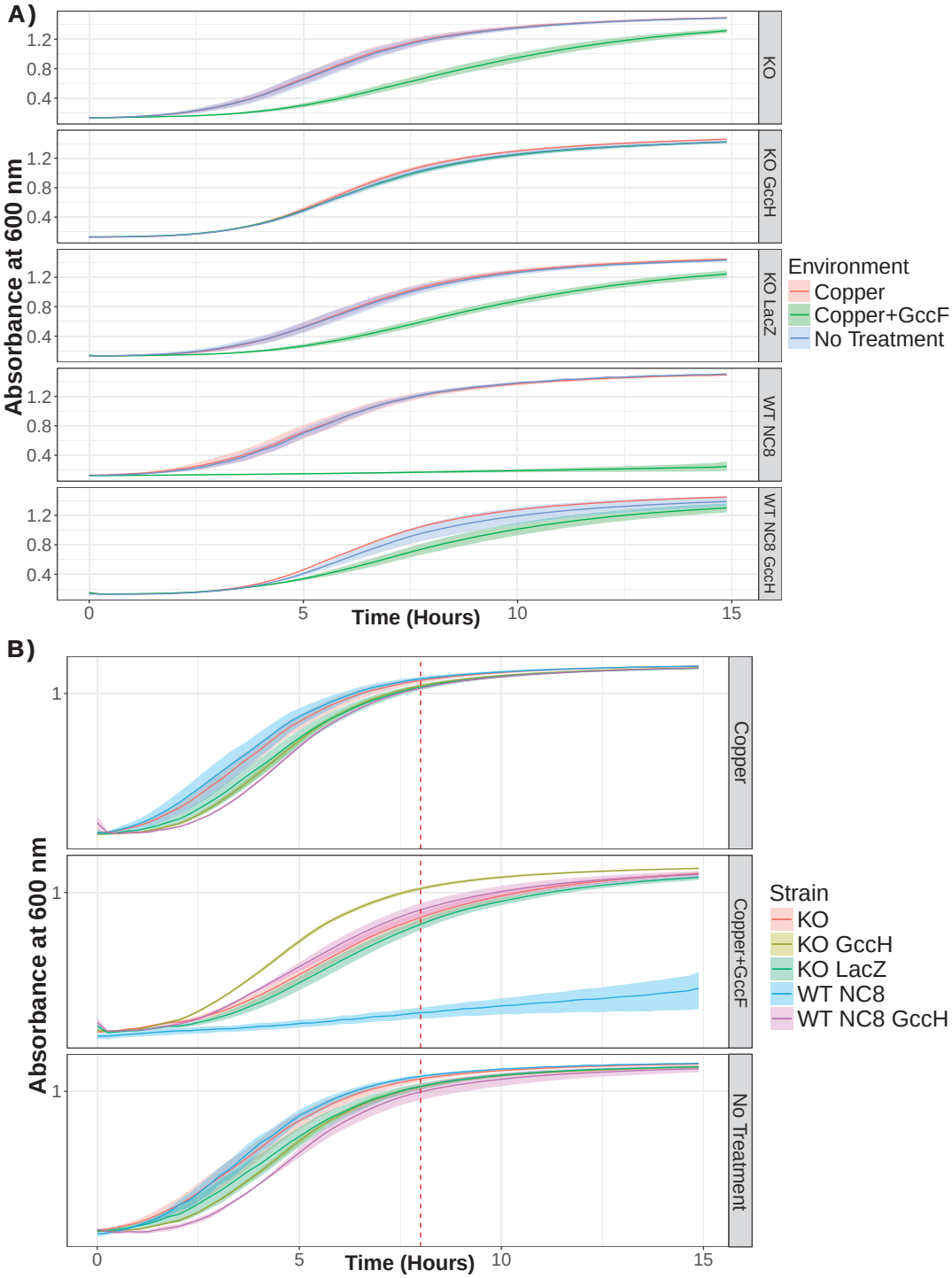


Figure 45: **GccH: Growth curve data**

(A) KO-GccH growth data grouped by strain. **(B)** GccH growth data grouped by environment, the red dotted line indicates the 8 hour interval used for data quantification in Figure 46. **Key:** KO (Red) *L. plantarum* NC8 *pts18cba* knockout cells without a plasmid; KO-GccH (khaki) the KO mutant transformed with pRV613-*GccH*; KO-LacZ (Green) KO cells transformed with the pRV613-*LacZ* expression plasmid; *L. plantarum* NC8 wild type (Blue) unmodified NC8 cells; WT-GccH (Lavender) *L. plantarum* NC8 wild type cells transformed with the pRV613-*GccH* plasmid. **Growth conditions:** (Copper) MRS media containing a final concentration of 30 µM of Cu^{2+}; (Copper+GccF) MRS media containing a final concentration of 60 nM of Glycocin F and 30 µM of Cu^{2+}; (No treatment) only sterilised MRS. Shading represents a 95% confidence interval.

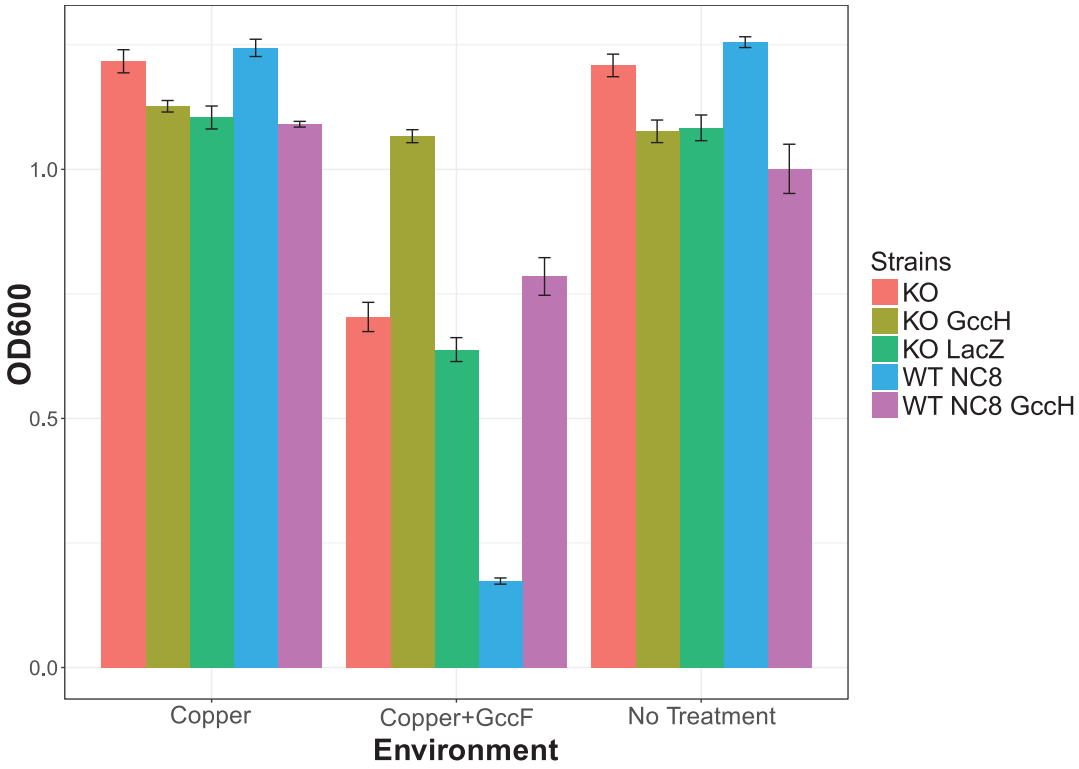


Figure 46: **GccH: quantitative phenotype analysis at ∼8 hours**
Growth data taken at eight hours. **Key:** KO (Red) *L. plantarum* NC8 *pts18cba* knockout cells without a plasmid; KO-GccH (khaki) the KO mutant transformed with pRV613-*GccH*; KO-LacZ (Green) KO cells transformed with the pRV613-*LacZ* expression plasmid; *L. plantarum* NC8 wild type (Blue) unmodified NC8 cells; WT-GccH (Lavender) *L. plantarum* NC8 wild type cells transformed with the pRV613-*GccH* plasmid. **Growth conditions:** (Copper) MRS media containing a final concentration of 30 µM of Cu^{2+}; (Copper+GccF) MRS media containing a final concentration of 60 nM of Glycocin F and 30 µM of Cu^{2+}; (No treatment) only sterilised MRS. Error bars represent the standard error (Table 16).

The WT-GccH expresses a stronger resistance phenotype to GccF than the KO mutant and untransformed WT cells but still remains susceptible to GccF. GccF reduces the WT-GccH cell density ∼37% at eight hours which is roughly the calculated induced influence of PTS EIIC domain variants at the same time point. The WT cells have always been more sensitive to GccF than any other tested strains and usually took 7 hours to recover from GccF induced bacteriostasis. At eight hours the GccF treated WT cells have ∼86% less cell density than the copper control, however this loss of growth is significantly reduced when the cells harbour plasmids containing the *GccH* gene (Figure 46).

The rise in GccF resistance in both tested strains containing a *gccH* insert is strong evidence that this is the GccF immunity gene. Furthermore, it is likely to function at the protein level because it is under the control of a copper-inducible promoter housed on the pRV613 plasmid which contains the ribosomal binding site and start codon. This assumption is also supported by the *in silico* analysis of the *gccH* gene. Bioinformatic predictions suggested that if the *gccH* gene generated a protein, it would be cytosolic. There is no evidence that this gene product is secreted, due to a lack of a leader peptide or signal motif. It is clear that the *gccH* gene does provides resistance to GccF, but may not be the sole immunity mechanism as some bacterial strains appear to have two mechanisms of action as seen for *L. plantarum* NC8.

The GccH immunity mechanism does not involve the EII^{GlcNAc} because a resistance phenotype was still expressed in EII^{GlcNAc} knockout cells harbouring the pRV613-*GccH* plasmid. Had the GccH immunity mechanism been dependant on the EII^{GlcNAc} transporter, the resistance phenotype would have been absent from the EII^{GlcNAc} knockout cells. However EII^{GlcNAc} is involved in the GccF mechanism of action because WT cells transformed with GccH are more responsive to GccF than the GccH containing KO cells. The secondary target of GccH remains a mystery. This finding provides some new tests, which will be invaluable in discovering the mechanism used by GccF to shut down cell growth.

Conclusion

Optimising the transformation processes and using PCR to amplify gene inserts from plasmids used to transform *L .plantarum* NC8 cells for sequencing were both instrumental in the success of these experiments.

Bioinformatic analysis of the *pts18cba* gene and its possible protein product along with the literature surrounding other EII enzymes of the phosphoenolpyruvate phosphotransferase system (PTS) suggested the active form was a dimer. Each contributing monomer is comprised of one membrane bound domain (EIIC) and two cytosolic domains (EIIA, EIIB). Results from other labs on other members of the family suggested that the cytosolic domains could be removed without compromising the integrity of the EIIC domain.

A truncation analysis designed around the stability of EIIC domain revealed that the GccF mechanism of action could induce bacteriostasis with equal effectiveness in cells in which both the EIIA and EIIB domains had been removed. However removal of these two domains results in lower cell fitness indicating that the whole transporter is necessary for cell growth, or that the removal of the domains interferes with metabolic pathways generating toxicity. Despite the additional toxicity, the response to GccF for all EIIC containing cells lines was of equal magnitude. In contrast, the removal of the intact *pts18cba* gene from cells conveyed partial resistance.

The ability to recover the biological activity of the EII^{GlcNAc} by recruiting EIIA and EIIB domains from other PTS transporters is unlikely because the domains bear no sequence homology to other PTS transporters, and are unlikely to share the same structural features. Further empirical support was found using site-direct mutagenesis to disrupt the PTS18CBA

phosphorelay removing its natural potential as an active GlcNAc transporter without deleting the cytosolic domains thus preventing the rescue of the biological activity of EII^{GlcNAc}. This is based on the findings that assuming His 584 is involved in the phosphorelay and signalling, its mutation to Ala does seem to result in slower growth of the cells harbouring this plasmid.

Signal transduction is also unlikely to occur through the $EIICBA^{GlcNAc}$ beyond the EIIC. Removal of the phosphorelay through SDM and domain removal would impede chemical signalling via the phosphorelay or structural changes to the protein when a ligand binds. Transduction possibly occurs at the small surface of the EIIC domains that are exposed to the cytosolic side of the membrane. However, there is still a significant reduction in the growth of cells that do not contain the *pts* gene at all. It is from this evidence that while the EIIC domain of the PTS18CBA is clearly the initial target of GccF, any further involvement through potential signal transduction is unlikely. Previous work showed that GccF does not enter the cells (Figure 54) and the cells can still detect GccF without the PTS, at least in *Lactobacillus plantarum* NC8 cells. It stands to reason that the EIIC is not a sensor but a lure that increases the local concentration of GccF near the real target increasing the cell sensitivity. This is supported by the discovery of the GccF immunity gene *gccH*. GccH protein does not appear to interact with PTS18CBA transporter, because less resistance is seen when expressing *gccH* in the presence of the PTS18CBA, however complete resistance is seen when PTS18CBA is removed.

Figure 47 shows the previously proposed models in light of this body of work, which supports Model 3: "The Lure" as best description of the behaviour of the PTS^{GlcNAc} in the GccF Mechanism of action.

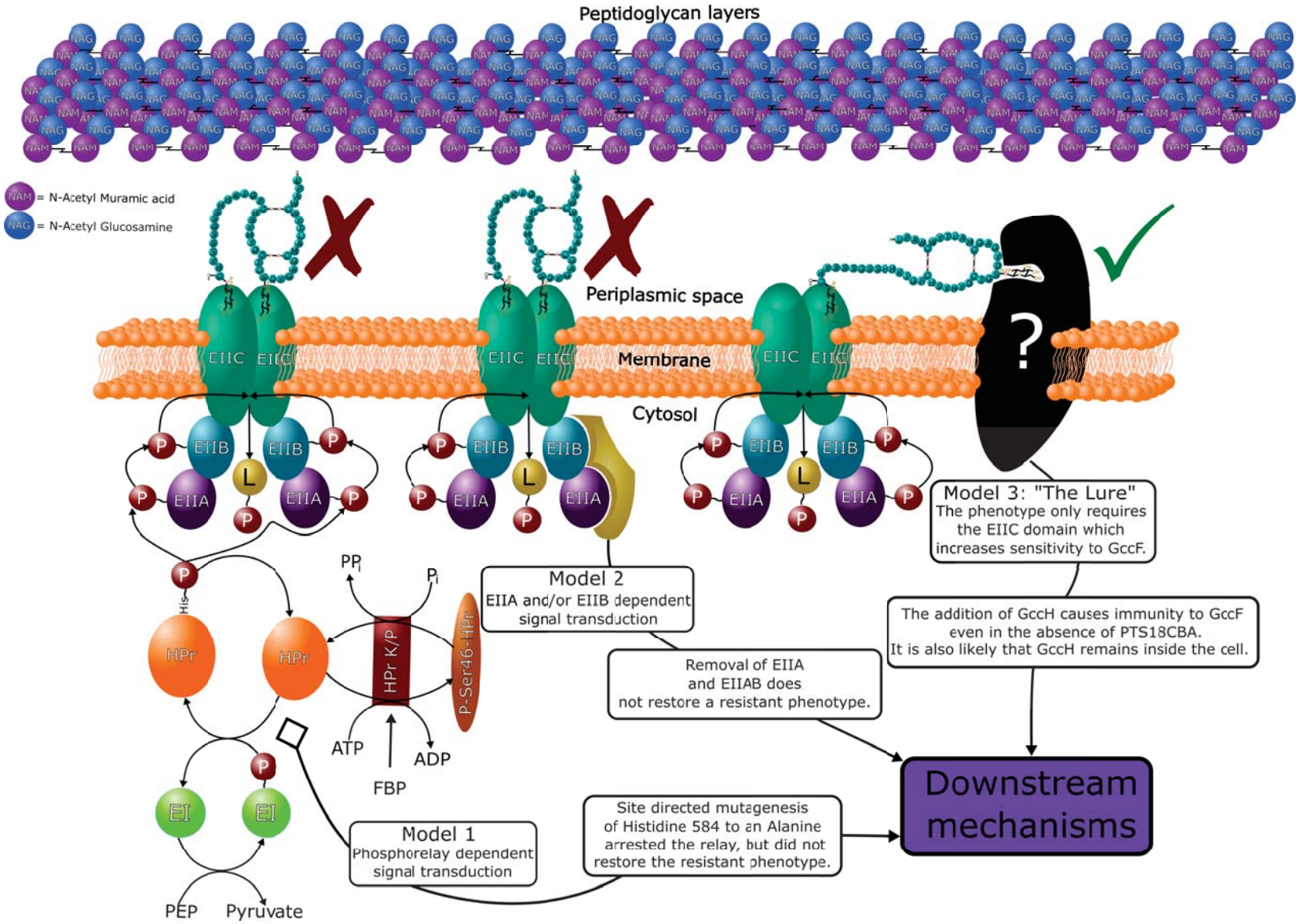

Peptidoglycan layers
NAM = N-Acetyl Muramic acid
NAG = N-Acetyl Glucosamine
Periplasmic space
Membrane
Cytosol
EIIC
EIIB
EIIA
P
L
?
Model 3: "The Lure"
The phenotype only requires the EIIC domain which increases sensitivity to GccF.
The addition of GccH causes immunity to GccF even in the absence of PTS18CBA.
It is also likely that GccH remains inside the cell.
Model 2
EIIA and/or EIIB dependent signal transduction
Removal of EIIA and EIIAB does not restore a resistant phenotype.
His
HPr
HPr K/P
P-Ser46-HPr
PPi
Pi
ATP
ADP
FBP
EI
PEP
Pyruvate
Model 1
Phosphorelay dependent signal transduction
Site directed mutagenesis of Histidine 584 to an Alanine arrested the relay, but did not restore the resistant phenotype.
Downstream mechanisms

Figure 47: **Summary: Proposed interaction models in light of the evidence**

PTS18CBA involvement in the GccF mechanism of action as molecular target. **Model 1**: Transduction occurs through HPr in a highly phosphorylated form or other downstream mechanisms as a result of the phosphorelay and PTS18CBA:GccF interaction. **Model 2** The EIIA and/or EIIB interact with PTS regulation domain (PRD) proteins to induce bacteriostatis in *L. plantarum* cells. **Model 3**: The PTS18CBA attracts and interacts with GccF, but transduction does not occur through PTS18CBA.

GccH is very likely to interact with the actual transduction target directly. This target is expected to be essential because it did not appear during natural mutation screening and largely remains unidentified. The PEP-PTS transporters are capable of facilitated diffusion when decoupled from their cytosolic proteins. The GccF mechanism of action is believed to make use of this function allowing the GlcNAc bound to the tail of GccF to act as bait which is lured by the PTS18CBA, allowing the essential O-linked GlcNAc to interact with the unknown signalling protein, that is suspected of being in the membrane based on the immunity function and cytosolic nature of GccH. A final model of how the PTS^{GlcNAc} is suspected to function in the GccF mechanism of action that includes insights provided by the study of *gccH* is shown in Figure 48.

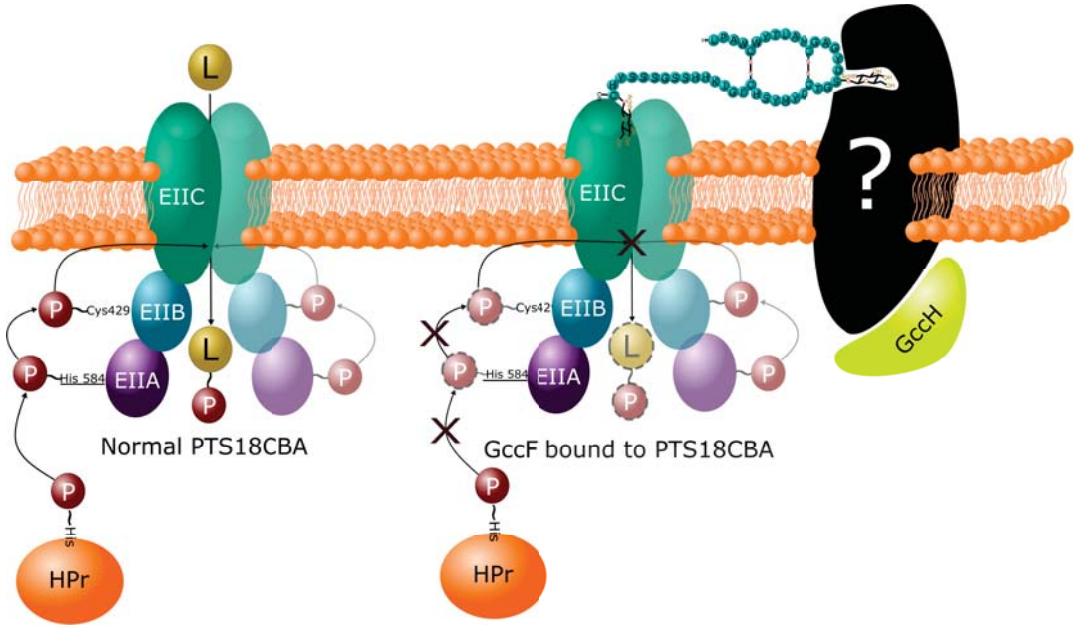


Figure 48: **Summary: Proposed behaviour of PTS18CBA, GccF and GccH**
(Left) A model depicting the normal function of the PTS18CBA. The phosphorelay functions in bimolecular interactions starting at PEP which phosphorylates EI. The phosphate is then transferred up the relay from EI to HPr which spontaneously phosphorylates EIIA. EIIA spontaneously phosphorylates EIIB which then transfers the phosphate to the ligand (L: Gold) as it enters the cell through EIIC. **(Right)** A model depicting the suspected behaviour of the PTS18CBA in the GccF mechanism of action. The PTS lures GccF *via* the attached GlcNAc moieties, increasing the local concentration and thus impacting the GccF sensitivity. The unknown signalling protein is shown as a black shape labelled with a question mark and bound to the immunity factor GccH (Lime).

Future directions

An immunoprecipitation study of the GccH protein expressed from a pRV613 or alternative plasmids could lead to the discovery of the secondary target of GccF by studying the binding partners through mass spectroscopy. The crystal and NMR structures of GccH will provide a good understanding of how this protein works and provide a measure of its stability. Combining the structure and immunoprecipitation studies a binding model can be developed and may even lead to a method for purifying the secondary target and allowing its structure to be determined. Understanding the GccH and its target(s) will provide some new insight into not only this novel protein but will unravel the mechanism of action for GccF and potentially other glycocins. *GccI* phenotyping should also be done to find the function of this mysterious gene which could be achieved in the same way *gccH* was done during this study.

Increasing GccF production and making it more convenient is essential for streamlining research and making this bacteriocin commercially viable. With new insights into the mechanism of action and the *gcc* cluster, it may be now possible to produce cells that are capable of artificially induced GccF production. While rates of expression are low using the pRV613 vector in *L. plantarum* cells, alternative plasmids are available that are capable of higher copy numbers and increased protein production rates. If this is done in the *L. plantarum* NC8 KO cells, GccF will have little effect and may work well as a GccF producer strain.

References

1. Organization, W. H. *et al. Antimicrobial resistance: global report on surveillance* (World Health Organization, 2014).

2. Carlet, J., Pulcini, C. & Piddock, L. J. Antibiotic resistance: a geopolitical issue. *Clinical Microbiology and Infection* **20,** 949–953 (2014).

3. Xia, J., Gao, J. & Tang, W. Nosocomial infection and its molecular mechanisms of antibiotic resistance. *Bioscience trends* **10,** 14–21 (2016).

4. MPI. *Antibiotics and resistance* 2017. `https://www.mpi.govt.nz/food-safety/whats-in-our-food/chemicals-and-food/agricultural-compounds-and-residues/antibiotics-and-resistance/`.

5. D Clair, J. & Colatrella, S. Opening Pandora's (tool) Box: health care construction and associated risk for nosocomial infection. *Infectious Disorders-Drug Targets (Formerly Current Drug Targets-Infectious Disorders)* **13,** 177–183 (2013).

6. Boucher, H. W. *et al.* Bad bugs, no drugs: no ESKAPE! An update from the Infectious Diseases Society of America. *Clinical Infectious Diseases* **48,** 1–12 (2009).

7. Fiore, M. *et al.* Nosocomial spontaneous bacterial peritonitis antibiotic treatment in the era of multi-drug resistance pathogens: A systematic review. *World Journal of Gastroenterology* **23,** 4654 (2017).

8. Dos Santos, D. F. K., Istvan, P., Quirino, B. F. & Kruger, R. H. Functional metagenomics as a tool for identification of new antibiotic resistance genes from natural environments. *Microbial ecology* **73,** 479–491 (2017).

9. Laland, K. N., Odling-Smee, J. & Feldman, M. W. Niche construction earns its keep. *Behavioral and brain sciences* **23,** 164–172 (2000).

10. Odling-Smee, F. J., Laland, K. N. & Feldman, M. W. Niche construction. *American Naturalist,* 641–648 (1996).

11. Laland, K. N., Odling-Smee, F. J. & Feldman, M. W. Evolutionary consequences of niche construction and their implications for ecology. *Proceedings of the National Academy of Sciences* **96,** 10242–10247 (1999).

12. Riley, M. A. & Chavan, M. A. *Bacteriocins* 7–127 (Springer, 2007).

13. Cotter, P. D., Hill, C. & Ross, R. P. Bacteriocins: developing innate immunity for food. *Nature Reviews Microbiology* **3,** 777–788 (2005).

14. Delves-Broughton, J., Blackburn, P., Evans, R. & Hugenholtz, J. Applications of the bacteriocin, nisin. *Antonie van Leeuwenhoek* **69,** 193–202 (1996).

15. Messaoudi, S. *et al.* Lactobacillus salivarius: bacteriocin and probiotic activity. *Food microbiology* **36,** 296–304 (2013).

16. Perez, R. H., Zendo, T. & Sonomoto, K. Novel bacteriocins from lactic acid bacteria (LAB): various structures and applications. *Microb Cell Fact* **13,** S3 (2014).

17. Nes, I. F. *et al.* Biosynthesis of bacteriocins in lactic acid bacteria. *Antonie van Leeuwenhoek* **70,** 113–128 (1996).

18. O'Flaherty, S. & Klaenhammer, T. R. The role and potential of probiotic bacteria in the gut, and the communication between gut microflora and gut/host. *International Dairy Journal* **20,** 262–268 (2010).

19. Kaškonienė, V. *et al.* Current state of purification, isolation and analysis of bacteriocins produced by lactic acid bacteria. *Applied microbiology and biotechnology* **101,** 1323–1335 (2017).

20. Klaenhammer, T. R. Genetics of bacteriocins produced by lactic acid bacteria. *FEMS Microbiol. Rev* **12,** 39–85 (1993).

21. Nami, Y. *et al.* Probiotics or antibiotics: future challenges in medicine. *Journal of medical microbiology* **64,** 137–146 (2015).

22. Gillor, O., Nigro, L. M. & Riley, M. A. Genetically engineered bacteriocins and their potential as the next generation of antimicrobials. *Current pharmaceutical design* **11,** 1067–1075 (2005).

23. Lien, S. & Lowman, H. B. Therapeutic peptides. *Trends in biotechnology* **21,** 556–562 (2003).

24. Alvarez-Sieiro, P., Montalbán-López, M., Mu, D. & Kuipers, O. P. Bacteriocins of lactic acid bacteria: extending the family. *Applied microbiology and biotechnology* **100,** 2939–2951 (2016).

25. Abee, T., Klaenhammer, T. & Letellier, L. Kinetic studies of the action of lactacin F, a bacteriocin produced by Lactobacillus johnsonii that forms poration complexes in the cytoplasmic membrane. *Applied and Environmental Microbiology* **60,** 1006–1013 (1994).

26. Arnison, P. G. *et al.* Ribosomally synthesized and post-translationally modified peptide natural products: overview and recommendations for a universal nomenclature. *Natural product reports* **30,** 108–160 (2013).

27. Yang, X. & Van Der Donk, W. A. Michael-type cyclizations in lantibiotic biosynthesis are reversible. *ACS chemical biology* **10,** 1234–1238 (2015).

28. Lubelski, J., Rink, R., Khusainov, R., Moll, G. & Kuipers, O. Biosynthesis, immunity, regulation, mode of action and engineering of the model lantibiotic nisin. *Cellular and Molecular Life Sciences* **65,** 455–476 (2008).

29. Knerr, P. J. & Van Der Donk, W. A. Discovery, biosynthesis, and engineering of lantipeptides. *Annual review of biochemistry* **81,** 479–505 (2012).

30. Reisinger, P., Seidel, H., Tschesche, H. & Hammes, W. P. The effect of nisin on murein synthesis. *Archives of microbiology* **127,** 187–193 (1980).

31. Breukink, E. & de Kruijff, B. The lantibiotic nisin, a special case or not? *Biochimica et Biophysica Acta (BBA)-Biomembranes* **1462,** 223–234 (1999).

32. Giffard, C. J. *et al.* Structure-function relations of variant and fragment nisins studied with model membrane systems. *Biochemistry* **36,** 3802–3810 (1997).

33. Brötz, H. *et al.* Role of lipid-bound peptidoglycan precursors in the formation of pores by nisin, epidermin and other lantibiotics. *Molecular microbiology* **30,** 317–327 (1998).

34. Ruhr, E. & Sahl, H.-G. Mode of action of the peptide antibiotic nisin and influence on the membrane potential of whole cells and on cytoplasmic and artificial membrane vesicles. *Antimicrobial Agents and Chemotherapy* **27,** 841–845 (1985).

35. Hasper, H. E. *et al.* An alternative bactericidal mechanism of action for lantibiotic peptides that target lipid II. *Science* **313,** 1636–1637 (2006).

36. Cotter, P. D. An Upp-turn in bacteriocin receptor identification. *Molecular microbiology* **92,** 1159–1163 (2014).

37. Montalbán-López, M., Sánchez-Hidalgo, M., Cebrián, R. & Maqueda, M. Discovering the bacterial circular proteins: bacteriocins, cyanobactins, and pilins. *Journal of Biological Chemistry* **287,** 27007–27013 (2012).

38. Lohans, C. T. *et al.* Solution structures of the linear leaderless bacteriocins enterocin 7A and 7B resemble carnocyclin A, a circular antimicrobial peptide. *Biochemistry* **52,** 3987–3994 (2013).

39. Acedo, J. Z. *et al.* Solution structure of acidocin B, a circular bacteriocin produced by Lactobacillus acidophilus M46. *Applied and environmental microbiology* **81,** 2910–2918 (2015).

40. Himeno, K. *et al.* Identification, characterization, and three-dimensional structure of the novel circular bacteriocin, enterocin NKR-5-3B, from Enterococcus faecium. *Biochemistry* **54,** 4863–4876 (2015).

41. González, C. *et al.* Bacteriocin AS-48, a microbial cyclic polypeptide structurally and functionally related to mammalian NK-lysin. *Proceedings of the National Academy of Sciences* **97,** 11221–11226 (2000).

42. Towle, K. & Vederas, J. Structural features of many circular and leaderless bacteriocins are similar to those in saposins and saposin-like peptides. *MedChemComm* **8,** 276–285 (2017).

43. Gálvez, A., Valdivia, E., Martínez, M. & Maqueda, M. Bactericidal action of peptide antibiotic AS-48 against Escherichia coli K-12. *Canadian journal of microbiology* **35,** 318–321 (1989).

44. Gálvez, A., Maqueda, M., Martínez-Bueno, M. & Valdivia, E. Bactericidal and bacteriolytic action of peptide antibiotic AS-48 against Gram-positive and Gram-negative bacteria and other organisms. *Research in microbiology* **140,** 57–68 (1989).

45. Gong, X., Martin-Visscher, L. A., Nahirney, D., Vederas, J. C. & Duszyk, M. The circular bacteriocin, carnocyclin A, forms anion-selective channels in lipid bilayers. *Biochimica et Biophysica Acta (BBA)-Biomembranes* **1788,** 1797–1803 (2009).

46. Abriouel, H., Valdivia, E., Gálvez, A. & Maqueda, M. Influence of physico-chemical factors on the oligomerization and biological activity of bacteriocin AS-48. *Current microbiology* **42,** 89–95 (2001).

47. Sánchez-Barrena, M. *et al.* Structure of bacteriocin AS-48: from soluble state to membrane bound state. *Journal of molecular biology* **334,** 541–549 (2003).

48. Cebrián, R. *et al.* The bacteriocin AS-48 requires dimer dissociation followed by hydrophobic interactions with the membrane for antibacterial activity. *Journal of structural biology* **190,** 162–172 (2015).

49. Murphy, K. *et al.* Genome mining for radical SAM protein determinants reveals multiple sactibiotic-like gene clusters. *PloS one* **6,** e20852 (2011).

50. Rea, M. C. *et al.* Thuricin CD, a posttranslationally modified bacteriocin with a narrow spectrum of activity against Clostridium difficile. *Proceedings of the National Academy of Sciences* **107,** 9352–9357 (2010).

51. Noll, K. S., Sinko, P. J. & Chikindas, M. L. Elucidation of the molecular mechanisms of action of the natural antimicrobial peptide subtilosin against the bacterial vaginosis-associated pathogen Gardnerella vaginalis. *Probiotics and antimicrobial proteins* **3,** 41–47 (2011).

52. Sit, C. S., McKay, R. T., Hill, C., Ross, R. P. & Vederas, J. C. The 3D structure of thuricin CD, a two-component bacteriocin with cysteine sulfur to α-carbon cross-links. *Journal of the American Chemical Society* **133,** 7680–7683 (2011).

53. Mathur, H. *et al.* Insights into the Mode of Action of the Sactibiotic Thuricin CD. *Frontiers in microbiology* **8** (2017).

54. Melby, J. O., Nard, N. J. & Mitchell, D. A. Thiazole/oxazole-modified microcins: complex natural products from ribosomal templates. *Current opinion in chemical biology* **15,** 369–378 (2011).

55. Molloy, E. M., Cotter, P. D., Hill, C., Mitchell, D. A. & Ross, R. P. Streptolysin S-like virulence factors: the continuing sagA. *Nature reviews. Microbiology* **9,** 670 (2011).

56. Miyoshi-Akiyama, T. *et al.* Cytocidal effect of Streptococcus pyogenes on mouse neutrophils in vivo and the critical role of streptolysin S. *The Journal of infectious diseases* **192,** 107–116 (2005).

57. Bates, C. S., Montanez, G. E., Woods, C. R., Vincent, R. M. & Eichenbaum, Z. Identification and characterization of a Streptococcus pyogenes operon involved in binding of hemoproteins and acquisition of iron. *Infection and immunity* **71,** 1042–1055 (2003).

58. Eichenbaum, Z., Muller, E., Morse, S. A. & Scott, J. R. Acquisition of iron from host proteins by the group A streptococcus. *Infection and immunity* **64,** 5428–5429 (1996).

59. Venugopal, H. *et al.* Structural, dynamic, and chemical characterization of a novel S-glycosylated bacteriocin. *Biochemistry* **50,** 2748–2755 (2011).

60. Stepper, J. *et al.* Cysteine S-glycosylation, a new post-translational modification found in glycopeptide bacteriocins. *FEBS letters* **585,** 645–650 (2011).

61. Maky, M. A. *et al.* Enterocin F4-9, a novel O-linked glycosylated bacteriocin. *Applied and environmental microbiology* **81,** 4819–4826 (2015).

62. Oman, T. J., Boettcher, J. M., Wang, H., Okalibe, X. N. & van der Donk, W. A. Sublancin is not a lantibiotic but an S-linked glycopeptide. *Nature chemical biology* **7,** 78–80 (2011).

63. Paik, S. H., Chakicherla, A. & Hansen, J. N. 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,** 23134–23142 (1998).

64. Stein, T. Bacillus subtilis antibiotics: structures, syntheses and specific functions. *Molecular microbiology* **56,** 845–857 (2005).

65. Kelly, W., Asmundson, R. & Huang, C. Characterization of plantaricin KW30, a bacteriocin produced by Lactobacillus plantarum. *Journal of Applied Bacteriology* **81,** 657–662 (1996).

66. Kerr, A. P. *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* (2013).

67. Drower, K. R. *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 [EMBARGOED until 1 July 2015]* (2014).

68. Norris, G. E. & Patchett, M. L. The glycocins: in a class of their own. *Current opinion in structural biology* **40,** 112–119 (2016).

69. Hegemann, J. D., Zimmermann, M., Xie, X. & Marahiel, M. A. Lasso peptides: an intriguing class of bacterial natural products. *Accounts of chemical research* **48,** 1909–1919 (2015).

70. Salomon, R. & Farías, R. N. Microcin 25, a novel antimicrobial peptide produced by Escherichia coli. *Journal of bacteriology* **174,** 7428–7435 (1992).

71. Potterat, O. *et al.* Aborycin–A Tricyclic 21-Peptide Antibiotic Isolated from Streptomyces griseoflavus. *European Journal of Organic Chemistry* **1994,** 741–743 (1994).

72. Tsunakawa, M. *et al.* Siamycins I and II, new anti-HIV peptides: I. Fermentation, isolation, biological activity and initial characterization. *The Journal of antibiotics* **48,** 433–434 (1995).

73. Katahira, R., Yamasaki, M., Matsuda, Y. & Yoshida, M. MS-271, A novel inhibitor of calmodulin-activated myosin light chain kinase from Streptomyces sp.—II. Solution structure of MS-271: Characteristic features of the 'lasso'structure. *Bioorganic & medicinal chemistry* **4,** 121–129 (1996).

74. Yano, K. *et al.* MS-271, a novel inhibitor of calmodulin-activated myosin light chain kinase from Streptomyces sp.—I. Isolation, structural determination and biological properties of MS-271. *Bioorganic & medicinal chemistry* **4,** 115–120 (1996).

75. Kimura, K.-i. *et al.* Propeptin, a New Inhibitor of Prolyl Endopeptidase Produced by Microbispora. *The Journal of antibiotics* **50,** 373–378 (1997).

76. Delgado, M. A., Rintoul, M. R., Farías, R. N. & Salomón, R. A. Escherichia coli RNA polymerase is the target of the cyclopeptide antibiotic microcin J25. *Journal of bacteriology* **183,** 4543–4550 (2001).

77. Iwatsuki, M. *et al.* Lariatins, antimycobacterial peptides produced by Rhodococcus sp. K01- B0171, have a lasso structure. *Journal of the American Chemical Society* **128,** 7486–7491 (2006).

78. Knappe, T. A. *et al.* Isolation and structural characterization of capistruin, a lasso peptide predicted from the genome sequence of Burkholderia thailandensis E264. *Journal of the American Chemical Society* **130,** 11446–11454 (2008).

79. Mathavan, I. *et al.* Structural basis for hijacking siderophore receptors by antimicrobial lasso peptides. *Nature chemical biology* **10,** 340–342 (2014).

80. Gavrish, E. *et al.* Lassomycin, a ribosomally synthesized cyclic peptide, kills Mycobacterium tuberculosis by targeting the ATP-dependent protease ClpC1P1P2. *Chemistry & biology* **21,** 509–518 (2014).

81. Maksimov, M. O., Pan, S. J. & Link, A. J. Lasso peptides: structure, function, biosynthesis, and engineering. *Natural product reports* **29,** 996–1006 (2012).

82. Diep, D. B., Skaugen, M., Salehian, Z., Holo, H. & Nes, I. F. Common mechanisms of target cell recognition and immunity for class II bacteriocins. *Proceedings of the National Academy of Sciences* **104,** 2384–2389 (2007).

83. Kjos, M. *et al.* Target recognition, resistance, immunity and genome mining of class II bacteriocins from Gram-positive bacteria. *Microbiology* **157,** 3256–3267 (2011).

84. Uzelac, G. *et al.* 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,** 5614–5621 (2013).

85. Nissen-Meyer, J., Holo, H., Håvarstein, L., Sletten, K. & Nes, I. A novel lactococcal bacteriocin whose activity depends on the complementary action of two peptides. *Journal of bacteriology* **174,** 5686–5692 (1992).

86. Nissen-Meyer, J., Oppegård, C., Rogne, P., Haugen, H. S. & Kristiansen, P. E. Structure and mode-of-action of the two-peptide (class-IIb) bacteriocins. *Probiotics and antimicrobial proteins* **2,** 52–60 (2010).

87. Rogne, P., Fimland, G., Nissen-Meyer, J. & Kristiansen, P. E. Three-dimensional structure of the two peptides that constitute the two-peptide bacteriocin lactococcin G. *Biochimica et Biophysica Acta (BBA)-Proteins and Proteomics* **1784,** 543–554 (2008).

88. Kjos, M. *et al.* Sensitivity to the two-peptide bacteriocin lactococcin G is dependent on UppP, an enzyme involved in cell-wall synthesis. *Molecular microbiology* **92,** 1177–1187 (2014).

89. Liu, X. *et al.* Identification of an N-terminal formylated, two-peptide bacteriocin from Enterococcus faecalis 710C. *Journal of agricultural and food chemistry* **59,** 5602–5608 (2011).

90. Masuda, Y., Zendo, T. & Sonomoto, K. New type non-lantibiotic bacteriocins: circular and leaderless bacteriocins. *Beneficial microbes* **3,** 3–12 (2012).

91. Yoneyama, F. *et al.* Lacticin Q, a lactococcal bacteriocin, causes high-level membrane permeability in the absence of specific receptors. *Applied and environmental microbiology* **75,** 538–541 (2009).

92. Acedo, J. Z. *et al.* Nuclear magnetic resonance solution structures of lacticin Q and aureocin A53 reveal a structural motif conserved among leaderless bacteriocins with broad-spectrum activity. *Biochemistry* **55,** 733–742 (2016).

93. Gabrielsen, C., Brede, D. A., Nes, I. F. & Diep, D. B. Circular bacteriocins: biosynthesis and mode of action. *Applied and environmental microbiology* **80,** 6854–6862 (2014).

94. Gabrielsen, C., Brede, D. A., Hernández, P. E., Nes, I. F. & Diep, D. B. The maltose ABC transporter in Lactococcus lactis facilitates high-level sensitivity to the circular bacteriocin garvicin ML. *Antimicrobial agents and chemotherapy* **56,** 2908–2915 (2012).

95. Maldonado-Barragán, A. *et al.* Garvicin A, a novel class IId bacteriocin from Lactococcus garvieae that inhibits septum formation in L. garvieae strains. *Applied and environmental microbiology* **79,** 4336–4346 (2013).

96. Xing, M., Simmonds, R. S. & Timkovich, R. Solution structure of the Cys74 to Ala74 mutant of the recombinant catalytic domain of Zoocin A. *Proteins: Structure, Function, and Bioinformatics* **85,** 177–181 (2017).

97. Chen, Y., Simmonds, R. S., Young, J. K. & Timkovich, R. Solution structure of the recombinant target recognition domain of zoocin A. *Proteins: Structure, Function, and Bioinformatics* **81,** 722–727 (2013).

98. Simmonds, R., Pearson, L., Kennedy, R. & Tagg, J. Mode of action of a lysostaphin-like bacteriolytic agent produced by Streptococcus zooepidemicus 4881. *Applied and environmental microbiology* **62,** 4536–4541 (1996).

99. Swe, P. M., Cook, G. M., Tagg, J. R. & Jack, R. W. Mode of action of dysgalacticin: a large heat-labile bacteriocin. *Journal of antimicrobial chemotherapy* **63,** 679–686 (2009).

100. Müller, E. & Radler, F. Caseicin, a bacteriocin from Lactobacillus casei. *Folia microbiologica* **38,** 441–446 (1993).

101. Riley, M. A. & Wertz, J. E. Bacteriocins: evolution, ecology, and application. *Annual Reviews in Microbiology* **56,** 117–137 (2002).

102. Zhao, X. & Kuipers, O. P. Identification and classification of known and putative antimicrobial compounds produced by a wide variety of Bacillales species. *BMC genomics* **17,** 882 (2016).

103. Deutscher, J., Francke, C. & Postma, P. W. How phosphotransferase system-related protein phosphorylation regulates carbohydrate metabolism in bacteria. *Microbiology and Molecular Biology Reviews* **70,** 939–1031 (2006).

104. Galinier, A. & Deutscher, J. Sophisticated regulation of transcriptional factors by the bacterial phosphoenolpyruvate: Sugar phosphotransferase system. *Journal of Molecular Biology* (2017).

105. Hoffee, P., Englesberg, E. & Lamy, F. The glucose permease system in bacteria. *Biochimica et Biophysica Acta (BBA)-Specialized Section on Biophysical Subjects* **79,** 337–350 (1964).

106. Kundig, W., Ghosh, S. & Roseman, S. Phosphate bound to histidine in a protein as an intermediate in a novel phospho-transferase system. *Proceedings of the National Academy of Sciences* **52,** 1067–1074 (1964).

107. Lengeler, J. W. PTS 50: Past, Present and Future, or Diauxie Revisited. *Journal of molecular microbiology and biotechnology* **25,** 79–93 (2015).

108. Reizer, J. *et al.* The phosphoenolpyruvate: sugar phosphotransferase system in gram-positive bacteria: properties, mechanism, and regulation. *CRC Critical reviews in microbiology* **15,** 297–338 (1988).

109. Postma, P., Lengeler, J. & Jacobson, G. Phosphoenolpyruvate: carbohydrate phosphotransferase systems of bacteria. *Microbiological reviews* **57,** 543–594 (1993).

110. Deutscher, J. *et al.* The bacterial phosphoenolpyruvate: carbohydrate phosphotransferase system: regulation by protein phosphorylation and phosphorylation-dependent protein-protein interactions. *Microbiology and Molecular Biology Reviews* **78,** 231–256 (2014).

111. Clore, G. M. & Venditti, V. Structure, dynamics and biophysics of the cytoplasmic protein–protein complexes of the bacterial phosphoenolpyruvate: sugar phosphotransferase system. *Trends in biochemical sciences* **38,** 515–530 (2013).

112. Feese, M. D., Comolli, L., Meadow, N. D., Roseman, S. & Remington, S. J. Structural studies of the Escherichia coli signal transducing protein IIAGlc: implications for target recognition. *Biochemistry* **36,** 16087–16096 (1997).

113. Worthylake, D. *et al.* Three-dimensional structure of the Escherichia coli phosphocarrier protein IIIglc. *Proceedings of the National Academy of Sciences* **88,** 10382–10386 (1991).

114. Liao, D.-I. *et al.* Structure of the IIA domain of the glucose permease of Bacillus subtilis at 2.2-A resolution. *Biochemistry* **30,** 9583–9594 (1991).

115. Van Montfort, R. L. *et al.* The structure of the Escherichia coli phosphotransferase IIA mannitol reveals a novel fold with two conformations of the active site. *Structure* **6,** 377–388 (1998).

116. RS, N., JC, G.-T., PJ, R., *et al.* Structure of the IIA domain of the mannose transporter from Escherichia coli at 1.7 angstroms resolution. *Journal of molecular biology* **259,** 502–511 (1996).

117. Sliz, P., Engelmann, R., Hengstenberg, W. & Pai, E. F. The structure of enzyme IIA lactose from Lactococcus lactis reveals a new fold and points to possible interactions of a multicomponent system. *Structure* **5,** 775–788 (1997).

118. Tang, C., Williams, D. C., Ghirlando, R. & Clore, G. M. Solution structure of enzyme IIAChitobiose from the N, N′-diacetylchitobiose branch of the Escherichia coli phosphotransferase system. *Journal of Biological Chemistry* **280,** 11770–11780 (2005).

119. Legler, P. M., Cai, M., Peterkofsky, A. & Clore, G. M. Three-dimensional solution structure of the cytoplasmic B domain of the mannitol transporter IIMannitol of the Escherichia coli phosphotransferase system. *Journal of Biological Chemistry* **279,** 39115–39121 (2004).

120. Suh, J.-Y., Tang, C., Cai, M. & Clore, G. M. Visualization of the phosphorylated active site loop of the cytoplasmic B domain of the mannitol transporter II Mannitol of the Escherichia coli phosphotransferase system by NMR spectroscopy and residual dipolar couplings. *Journal of molecular biology* **353,** 1129–1136 (2005).

121. Van Montfort, R. L. *et al.* The structure of an energy-coupling protein from bacteria, IIB cellobiose, reveals similarity to eukaryotic protein tyrosine phosphatases. *Structure* **5,** 217–225 (1997).

122. Eiso, A. *et al.* NMR structure of cysteinyl-phosphorylated enzyme IIB of the N, N′-diacetylchitobiose-specific phosphoenolpyruvate-dependent phosphotransferase system of Escherichia coli. *Journal of molecular biology* **308,** 993–1009 (2001).

123. Cai, M. *et al.* Solution structure of the phosphoryl transfer complex between the signal-transducing protein IIAGlucose and the cytoplasmic domain of the glucose transporter IICBGlucose of the Escherichia coli glucose phosphotransferase system. *Journal of Biological Chemistry* **278,** 25191–25206 (2003).

124. Hu, J. *et al.* Solution NMR structures of productive and non-productive complexes between the A and B domains of the cytoplasmic subunit of the mannose transporter of the Escherichia coli phosphotransferase system. *Journal of Biological Chemistry* **283,** 11024–11037 (2008).

125. Schauder, S., Nunn, R. S., Lanz, R., Erni, B. & Schirmer, T. Crystal structure of the IIB subunit of a fructose permease (IIB Lev) from Bacillus subtilis. *Journal of molecular biology* **276,** 591–602 (1998).

126. Orriss, G. L., Erni, B. & Schirmer, T. Crystal structure of the IIB Sor domain of the sorbose permease from Klebsiella pneumoniae solved to 1.75 Å resolution. *Journal of molecular biology* **327,** 1111–1119 (2003).

127. McCoy, J. G. *et al.* The structure of a sugar transporter of the glucose EIIC superfamily provides insight into the elevator mechanism of membrane transport. *Structure* **24,** 956–964 (2016).

128. Nguyen, T. X., Yen, M.-R., Barabote, R. D. & Saier Jr, M. H. Topological predictions for integral membrane permeases of the phosphoenolpyruvate: sugar phosphotransferase system. *Journal of molecular microbiology and biotechnology* **11,** 345–360 (2006).

129. Vastermark, A. & Saier, M. H. Time to stop holding the elevator: a new piece of the transport protein mechanism puzzle. *Structure* **24,** 845–846 (2016).

130. Saier, M., Hvorup, R. & Barabote, R. *Evolution of the bacterial phosphotransferase system: from carriers and enzymes to group translocators* 2005.

131. Cao, Y. *et al.* Crystal structure of a phosphorylation-coupled saccharide transporter. *Nature* **473,** 50 (2011).

132. Luo, P. *et al.* Crystal structure of a phosphorylation-coupled vitamin C transporter. *Nature structural & molecular biology* **22,** 238–241 (2015).

133. Deutscher, J. *et al.* Regulation of carbon metabolism in Gram-positive bacteria by protein phosphorylation. *Folia microbiologica* **42,** 171–178 (1997).

134. Paul, K., Brunstetter, D., Titen, S. & Blair, D. F. A molecular mechanism of direction switching in the flagellar motor of Escherichia coli. *Proceedings of the National Academy of Sciences* **108,** 17171–17176 (2011).

135. Deutscher, J. The mechanisms of carbon catabolite repression in bacteria. *Current opinion in microbiology* **11,** 87–93 (2008).

136. Hazelbauer, G. L. Bacterial chemotaxis: the early years of molecular studies. *Annual review of microbiology* **66,** 285–303 (2012).

137. Joyet, P. *et al.* Transcription regulators controlled by interaction with enzyme IIB components of the phosphoenolpyruvate: sugar phosphotransferase system. *Biochimica et Biophysica Acta (BBA)-Proteins and Proteomics* **1834,** 1415–1424 (2013).

138. Stülke, J., Arnaud, M., Rapoport, G. & Martin-Verstraete, I. PRD—a protein domain involved in PTS-dependent induction and carbon catabolite repression of catabolic operons in bacteria. *Molecular microbiology* **28,** 865–874 (1998).

139. Deutscher, J. *et al.* Loss of protein kinase-catalyzed phosphorylation of HPr, a phosphocarrier protein of the phosphotransferase system, by mutation of the ptsH gene confers catabolite repression resistance to several catabolic genes of Bacillus subtilis. *Journal of bacteriology* **176,** 3336–3344 (1994).

140. Laemmli, U. K. Cleavage of structural proteins during the assembly of the head of bacteriophage T4. *nature* **227,** 680–685 (1970).

141. O'Sullivan, D. J. & Klaenhammer, T. R. Rapid mini-prep isolation of high-quality plasmid DNA from Lactococcus and Lactobacillus spp. *Applied and environmental microbiology* **59,** 2730–2733 (1993).

142. Inoue, H., Nojima, H. & Okayama, H. High efficiency transformation of Escherichia coli with plasmids. *Gene* **96,** 23–28 (1990).

143. Zheng, L., Baumann, U. & Reymond, J.-L. An efficient one-step site-directed and site-saturation mutagenesis protocol. *Nucleic acids research* **32,** e115–e115 (2004).

144. Connelly, B. *analyzing microbial growth with R 2014* 2014. `http://bconnelly.net/2014/04/analyzing-microbial-growth-with-r/`.

145. Wetherill, C. *Building Barplots with Error Bars* 2015. `https://www.r-bloggers.com/building-barplots-with-error-bars/`.

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

147. Li, W. *et al.* The EMBL-EBI bioinformatics web and programmatic tools framework. *Nucleic acids research* **43,** W580–W584 (2015).

148. McWilliam, H. *et al.* Analysis tool web services from the EMBL-EBI. *Nucleic acids research* **41,** W597–W600 (2013).

149. Sievers, F. *et al.* Fast, scalable generation of high-quality protein multiple sequence alignments using Clustal Omega. *Molecular systems biology* **7,** 539 (2011).

150. Walker, J. M. *The proteomics protocols handbook* (Springer, 2005).

151. Gasteiger, E. *et al. Protein identification and analysis tools on the ExPASy server* (Springer, 2005).

152. Bjellqvist, B., Basse, B., Olsen, E. & Celis, J. E. Reference points for comparisons of two-dimensional maps of proteins from different human cell types defined in a pH scale where isoelectric points correlate with polypeptide compositions. *Electrophoresis* **15,** 529–539 (1994).

153. Bjellqvist, B. *et al.* The focusing positions of polypeptides in immobilized pH gradients can be predicted from their amino acid sequences. *Electrophoresis* **14,** 1023–1031 (1993).

154. Dyer, K. The quiet revolution: a new synthesis of biological knowledge. *Journal of Biological Education* **5,** 15–24 (1971).

155. Miseta, A. & Csutora, P. Relationship between the occurrence of cysteine in proteins and the complexity of organisms. *Molecular biology and evolution* **17,** 1232–1239 (2000).

156. Hofmann, K. TMbase-A database of membrane spanning proteins segments. *Biol. Chem. Hoppe-Seyler* **374,** 166 (1993).

157. Tsirigos, K. D., Peters, C., Shu, N., Käll, L. & Elofsson, A. The TOPCONS web server for consensus prediction of membrane protein topology and signal peptides. *Nucleic acids research* **43,** W401–W407 (2015).

158. Cao, B., Porollo, A., Adamczak, R., Jarrell, M. & Meller, J. Enhanced recognition of protein transmembrane domains with prediction-based structural profiles. *Bioinformatics* **22,** 303–309 (2005).

159. Drozdetskiy, A., Cole, C., Procter, J. & Barton, G. J. JPred4: a protein secondary structure prediction server. *Nucleic acids research* **43,** W389–W394 (2015).

160. Kelley, L. A., Mezulis, S., Yates, C. M., Wass, M. N. & Sternberg, M. J. The Phyre2 web portal for protein modeling, prediction and analysis. *Nature protocols* **10,** 845–858 (2015).

161. Zhang, Y. I-TASSER server for protein 3D structure prediction. *BMC bioinformatics* **9,** 40 (2008).

162. Roy, A., Kucukural, A. & Zhang, Y. I-TASSER: a unified platform for automated protein structure and function prediction. *Nature protocols* **5,** 725 (2010).

163. Yang, J. *et al.* The I-TASSER Suite: protein structure and function prediction. *Nature methods* **12,** 7–8 (2015).

164. Huang, K., Kapadia, G., Zhu, P.-P., Peterkofsky, A. & Herzberg, O. A promiscuous binding surface: crystal structure of the IIA domain of the glucose-specific permease from Mycoplasma capricolum. *Structure* **6,** 697–710 (1998).

165. Nam, T.-W. *et al.* Analyses of Mlc–IIBGlc interaction and a plausible molecular mechanism of Mlc inactivation by membrane sequestration. *Proceedings of the National Academy of Sciences* **105,** 3751–3756 (2008).

166. Matsumoto, S. *et al.* Crystal structures of an archaeal oligosaccharyltransferase provide insights into the catalytic cycle of N-linked protein glycosylation. *Proceedings of the National Academy of Sciences* **110,** 17868–17873 (2013).

167. Suh, J.-Y., Cai, M., Williams, D. C. & Clore, G. M. Solution structure of a post-transition state analog of the phosphotransfer reaction between the A and B cytoplasmic domains of the mannitol transporter IIMannitol of the Escherichia coli phosphotransferase system. *Journal of Biological Chemistry* **281,** 8939–8949 (2006).

168. Chassy, B. M. & Flickinger, J. L. Transformation of Lactobacillus casei by electroporation. *FEMS Microbiology Letters* **44,** 173–177 (1987).

169. Tu, Q. *et al.* Room temperature electrocompetent bacterial cells improve DNA transformation and recombineering efficiency. *Scientific reports* **6** (2016).

170. Aukrust, T. & Blom, H. Transformation of Lactobacillus strains used in meat and vegetable fermentations. *Food Research International* **25,** 253–261 (1992).

171. Crutz-Le Coq, A.-M. & Zagorec, M. Vectors for lactobacilli and other Gram-positive bacteria based on the minimal replicon of pRV500 from Lactobacillus sakei. *Plasmid* **60,** 212–220 (2008).

Appendix

5.1 Statistical data

Table 15: **Statistical data for GccH growth data at ∼8 hours**

#	Strain	Environment	mean	sd	n	se
1	KO	Cu^{2+}	1.2170	0.1161	25	0.0232
2	KO-GccH	Cu^{2+}	1.1267	0.0490	18	0.0115
3	KO-LacZ	Cu^{2+}	1.1041	0.1261	30	0.0230
4	WT	Cu^{2+}	1.2439	0.0716	17	0.0174
5	WT-GccH	Cu^{2+}	1.0906	0.0173	9	0.0058
6	KO	Cu^{2+} & GccF	0.7037	0.2116	52	0.0293
7	KO-GccH	Cu^{2+} & GccF	1.0666	0.0866	44	0.0130
8	KO-LacZ	Cu^{2+} & GccF	0.6382	0.1746	53	0.0240
9	WT	Cu^{2+} & GccF	0.1735	0.0206	11	0.0062
10	WT-GccH	Cu^{2+} & GccF	0.7850	0.1927	26	0.0378
11	KO	NT	1.2087	0.1199	28	0.0227
12	KO-GccH	NT	1.0764	0.1038	21	0.0226
13	KO-LacZ	NT	1.0834	0.1389	29	0.0258
14	WT	NT	1.2554	0.0390	13	0.0108
15	WT-GccH	NT	1.0010	0.1563	10	0.0494

Summary of statistical outputs from titre plate experiments post sample averaging. The bar chart and standard error were calculated by taking all data at 8 hours. Conditions: Strain and Environment were used as grouping variable; mean: the mean of averaged biological at 8 hours. sd: standard deviation, n: number of biological replicates, each biological replicate was measurements in triplicate and averaged. Each biological replicate for transformed strains was an independent overnight culture grown from one of three glycerol stocks containing a pure cell culture created using independent transformation events. WT and KO cell biological replicates were independent overnight cultures.

Table 16: **Statistical summary of growth data at ∼8 hours**

#	Strain	Environment	mean	sd	n	se
1	KO	Cu^{2+}	1.220	0.120	25	0.023
2	KO LacZ	Cu^{2+}	1.100	0.130	30	0.023
3	KO pts18cba	Cu^{2+}	1.148	0.121	24	0.025
4	KO pts18cb	Cu^{2+}	1.057	0.079	10	0.025
5	KO pts18c	Cu^{2+}	0.706	0.124	13	0.034
6	KO H584A	Cu^{2+}	1.079	0.070	33	0.012
7	Wild Type NC8	Cu^{2+}	1.240	0.070	17	0.017
8	KO	Cu^{2+} & GccF	0.700	0.210	25	0.023
9	KO LacZ	Cu^{2+} & GccF	0.640	0.170	53	0.024
10	KO pts18cba	Cu^{2+} & GccF	0.262	0.085	44	0.013
11	KO pts18cb	Cu^{2+} & GccF	0.244	0.034	36	0.006
12	KO pts18c	Cu^{2+} & GccF	0.194	0.044	63	0.006
13	KO H584A	Cu^{2+} & GccF	0.234	0.065	66	0.008
14	Wild Type NC8	Cu^{2+} & GccF	0.170	0.020	11	0.006
15	KO	NT	1.210	0.120	28	0.023
16	KO LacZ	NT	1.100	0.140	29	0.026
17	KO pts18cba	NT	1.126	0.104	21	0.023
18	KO pts18cb	NT	0.977	0.097	17	0.024
19	KO pts18c	NT	0.746	0.168	30	0.031
20	KO H584A	NT	1.040	0.047	33	0.008
21	Wild Type NC8	NT	1.260	0.040	13	0.011

Summary of statistical outputs from titre plate experiments post sample averaging. Sample data was take at 8 hours for the creation bar charts and to calculate the standard error at that time point. Conditions: Strain and Environment were used as grouping variable; mean: the mean of averaged biological at 8 hours. sd: standard deviation, n: number of biological replicates, each biological replicate was measurements in triplicate and averaged. Each biological replicate for transformed strains was an independent overnight culture grown from one of three glycerol stocks containing a pure cell culture created using independent transformation events. WT and KO cell biological replicates were independent overnight cultures.

5.2 Gene sequences

5.2.1 *Lactobacillus plantarum* 8014

>8014_pts18_500_300_2789bp:CGACTCAAATGTTCAACGAAGCCTATGATGCTTGCGAAAGGGC

AGGCATCAAACTCATTGAGACGCGGCAAGAAGAAATCGACTCGGTTGAAACGGTCAGTCGA

GCCTATAAGTACCATTATCCATCAATGTATCAAGACTTCTCAAAGGGACGGCCAACCGAAGT

TGATTATATTAATGGCTATATTGCTAAGATTGGGCGGGAACATGACTACGTTTGTCGCGTGCA

TGAGTTTGTGACTCACGAAGTCCATTTAGCAGAAATGATGCGGCAATACCGGCACCCGGAAA

TCCCCGTTGCTGAAAAATAAGATTATTAATGAAAAAGTTCTGAGTATTTTCAGAACTTTTTTT

GTGGGCGGAAAAACTTAATTTAGTAAGGTTGAAATCGGTTTCTTTAACATTGGTCTATACACT

TAGATGACCTTGTAAACTTGCAATGAAAACATGTTGTGATAGGATTTATAACGTGCTTAAAA

AGAGAATGAGGAGAAATTGGTATGAAGACATATTTTCAGAAAATCGGTCAGTCATTGATGCT

ACCGATTGCAACGTTGCCCGCAGCAGCGATTTTGGTTGGGATTGGGAATTACCTGCCAAAAC

AGTGGCTGTTTGCAAATTACTTGATTCAGGGTGGTAACGTTGTCCTGAATAACTTAGCCTTGC

TGTTTGCAGTCGGCTTAGCTATCGGGATGTCAGTCAATAAGGACGGCGCGGCAGCTATTGCT

GGTTTGATCGCGTTTGAAGTTCCGGTGATGGTTTTAAAACCAGCGACATTAGCGACGATGCT

GAATGTTAAAGTGAGTCAGATCAATCCAGCATTTAGCGCGTTAGATAACAATGTGCTGATTG

GAATCAGTGCTGGACTGATTGCGGCTGCGCTCTATAATCGGTTCCACGAAGTAAAATTACCA

ATGGCGCTCTCATTCTTTAGTGGTAAACGCCTAGTGCCAATTATGGCTGCTTTTGTGATGCTG

ATTGTGACGGCGGTATTGTACCTTGTTTGGCCGTTTGTCTACGATGCCATCGTCCTCTTTGCA

ACGGGAATTTCAAAATTAGGTTTCGTGGGGGCTGGTCTATATGGCTTTTTCAATCGTTTATTG

ATTCCGACTGGTTTGCACCACGCCTTAAATTCCGTATTCTGGTATAACGTTGCGGGGATCAAT

GATATTGGCAATTTCTGGGCCAGCCATGGTGTTAAAGGAATCACAGGGATGTATGAAGCTGG

ATTCTTCCCAATTATGATGTTTGGCTTGCCAGCTGGTGCGTATGCAATTTATCGTAATGCACG

ACCGGAACGAAAAAAAGAAGTGGGTTCATTGATGTTAGCTGGTGCGTTTGCGTCGTTCTTTA

CGGGGGTGACCGAACCGCTTGAATTTTCATTCATGTTCGTGGCATGGCCGCTCTACTTATTAC

ATGCCGTTTTTATGGGACTATCATTAGGATTCGCCGCGTTAATGCACTGGACCGCCAGCTTTT

CGTTTAGTGGCGGGTTAGTCGATTATTTATTGAGTTTTCGGATGCCGTTAGCCAACCAACCCT

ATATGTTACTGGTTCAAGGGCTGGTGATGGCCGTCATTTATTACTTTGGTTTTGACTTTGCAA

TCAAACGGTTTAATTTGAAGACGCCTGGTCGTGAAGTTGTGTCTGCAGATGTCGATGGGGTG

GGTGCACCTGCTAGTCCAGCGGTGGCCGTTGCAGCAACGGATGATAAGTATATGCGCCAAGC
AAAGCAAATTTATGCAGCTATTGGTGGTCATGACAATATTAGTGTGATCAACAATTGTACGA
CGCGGCTGCGGTTACAACTTAAGGATACGGAAAAAGTCGATCAGCCGGCCGTAATGGCTGCT
GGCGTGCCTGGTTTGAACGTACTGGATGTTCATAACATTCACATTGTGATCGGCACGGAAGT
CCAGTTCGTCGCGGAGGCTTTACAAAAATTATTTTCTGGTCAGGTGGCGACGACCCCGGCAT
CTGATGCTGAATCAAAGGCCCCGGTTGAACCGCAAACTGCTACTGTAACGGAAGCGCCGGTG
ACAACGATTTTGCGGGCACCGGCAACGGGACAATTAATGCCGATTAGTACGGTTGCGGATGA
GACGTTTGCTGGTAAACTCTTAGGTGATGGTTACGCTGTTGAACCCGAAGATGGTGAAGTCG
TTGCTCCGGTTAGTGGGACAGTGACAAGTGTCTTTCCGACCAAACACGCTATCGGTTTGAAG
ACGACGAGTGGTTTAGAAGTCTTATTACACATGGGGATCAATACCGTGGAAATGAATGGTAC
GCCGTTCAAGTTACACGTGGCAACGGGTGATGAAATTGCGGCCGGTACTGCGGTAGCGACGG
TTGACTTGGCTGCTATCAAGTCAGCCGGTAAAGCGACAACCATGATGGTCGTTATCACCAAT
ATGGATCACGTTAACAAACTAACCCTTAATCCAACTGGTCATGTGACTAGCGGTGATTTGATT
GGCGCAGCTGAATAAGTTACGCGCTGTCGACGTTGGGTTTATTGAAAAAGTTTCGTCTAAAT
CTTTTCGATTTAGGCGGAACTTTTTTTGTCGCATTGGTTATCGAGGCAACGGATGACAGCAAA
TGAACAGCTTGATGCGGGGACAGTGCGTTTGATTGCTTCCCGCTGACTCCCACGATATACTAT
TGGCATTATAGGGGGGCGGTAACATGATTGACGAAGTATTAGTATTGAACTCAACGCCAACG
GGGATTTTACAGAATTCAGTTTCAGGAAGCATGGTGATTGACGCGGCACTTGATGCACGACT
AGT

5.2.2 *Lactobacillus plantarum* NC8

>NC8_pts18_500_300_2789bp:CGACTCAAATGTTCAACGAAGCCTATGATGCTTGCGAAAGGGC
AGGCATCAAACTCATTGAGACGCGGCAAGAAGAAATCGACTCGGTTGAAACGGTCAGTCGA
GCCTATAAGTACCATTATCCATCAATGTATCAAGACTTCTCAAAGGGACGGCCAACCGAAGT
TGATTATATTAATGGCTATATTGCTAGGATTGGGCGGGAACATGACTACGTTTGTCGCGTGCA
TGAGTTTGTGACTCACGAAGTCCATTTAGCAGAAATGATGCGGCAATACCGGCACCCGGAAA
TCCCCGTTGCTGAAAAATAAGATTATTAATGAAAAAGTTCTGAGTATTTTCAGAACTTTTTTT
GTGGGCGGAAAAACTTAATTTAGTAAGGTTGAAATCGGTTTCTTTGACATTGGTCTATACACT
TAGATAACCTTGTAAACTTGCAATGAAAACATGTTGTGATAGGATTTATAACGTGCTTAAAA
AGAGAATGAGGAGAAATTGGTATGAAGACATATTTTCAGAAAATCGGTCAGTCATTGATGCT

GCCGATTGCAACGTTGCCCGCAGCAGCGATTTTGGTTGGGATTGGTAATTACCTGCCAAAAC

AGTGGCTGTTTGCAAATTACTTGATTCAGGGTGGTAACGTTGTCCTGAATAACTTAGCCTTGC

TGTTTGCAGTCGGCTTAGCTATCGGGATGTCAGTCAATAAGGACGGCGCGGCAGCTATTGCT

GGTTTGATCGCGTTTGAAGTTCCGGTGATGGTGTTAAAACCAGCGACATTAGCGACGATGCT

GAATGTTAAAGTGAGTCAGATCAATCCAGCATTTAGCGCGTTAGATAACAATGTGCTGATTG

GGATCAGTGCTGGACTGATTGCGGCTGCGCTCTATAATCGGTTCCACGAAGTAAAATTACCA

ATGGCGCTCTCATTCTTTAGTGGTAAACGCCTAGTGCCAATTATGGCTGCTTTTGTGATGCTG

ATTGTGACGGCGGTATTGTACCTTGTTTGGCCGTTTGTCTACGATGCCATTGTCCTCTTTGCAA

CGGGGATTTCAAAATTAGGTTTCGTGGGGGCTGGTCTATATGGCTTTTTCAATCGTTTATTGA

TTCCGACTGGTTTGCACCACGCCTTAAATTCCGTATTCTGGTATAACGTTGCTGGTATCAATG

ATATTGGCAATTTCTGGGCCAGCCATGGTGTTAAAGGAATCACAGGGATGTATGAAGCTGGA

TTCTTCCCAATTATGATGTTTGGCTTGCCAGCTGGTGCGTATGCAATTTATCGTAATGCACGA

CCGGAACGAAAAAAAGAAGTGGGTTCATTGATGTTAGCGGGTGCGTTTGCGTCGTTCTTTAC

GGGGGTGACCGAACCGCTTGAATTTTCATTCATGTTCGTGGCATGGCCGCTCTACTTATTACA

TGCCGTTTTTATGGGACTATCATTAGGATTCGCCGCGTTAATGCACTGGACCGCCAGCTTTTC

GTTTAGTGGCGGGTTAGTCGATTATTTATTGAGTTTTCGGATGCCGTTAGCCAACCAACCCTA

TATGTTACTGGTTCAAGGGCTGGTGATGGCCGTCATTTATTACTTTGGTTTTGACTTTGCAAT

CAAGCGGTTTAATTTGAAGACGCCTGGTCGTGAAGTTGTGTCTGCAGATGTCGATGGGGTGG

GTGCACCTGCTAGTCCAGCGGTGGCCGTTGCAGCAACGGATGATAAATATATGCGCCAAGCA

AAGCAAATTTATGCGGCTATTGGTGGTCATGACAATATTAGTGTGATCAACAATTGTACGAC

GCGGCTGCGGTTACAACTTAAGGATACGGAAAAAGTTGATCAGCCGGCCGTAATGGCTGCTG

GCGTGCCTGGTTTGAACGTACTGGATGTTCATAACATCCACATTGTGATCGGTACGGAAGTC

CAGTTCGTTGCGGAGGCTTTACAAAAATTATTTTCTGGTCAGGTGGCGACGACCCCGGCATC

TGATGCTGAATCAAAGGCCCCGCTTGAACCGCAAACTGCTACTGTAACGGAAGCGCCGGTGA

CAACGATTTTGCGGGCACCGGCAACGGGACAATTAATGCCGATTAGTGCGGTTGCGGATGAG

ACGTTTGCTGGTAAACTCTTAGGTGATGGTTACGCTGTTGAACCCGAAGATGGTGAAGTCGT

TGCTCCGGTTAGTGGGACAGTGACAAGTGTCTTTCCGACCAAACACGCTATCGGTTTGAAGA

CGACGAGTGGTTTAGAAGTCTTATTACACATGGGGATCAATACCGTGGAAATGAATGGTACG

CCGTTCAAGTTACACGTGGCAACGGGTGATGAAATTGCGGCCGGTACTGCGGTAGCGACGGT

TGACTTGGCTGCTATCAAGTCAGCCGGTAAAGCGACAACCATGATGGTCGTTATCACCAATA

TGGATCACGTTAACAAACTAACCCTTAATCCAACTGGTCATGTGACTAGCGGTGATTTGATTG
GCGCAGCTGAATAAGTTACGCGCTGTCGACGTTGGGTTTATTGAAAAAGTTTCGTCTAAATCT
TTTCGATTTAGGCGGAACTTTTTTTGTCGCATTGGTTATCGAGGCAACGGATGACAGCAAATG
AACAACTTGATGCGGGGACAGTGCGTTTGATTGCTTCCCACTGACTCCCACGATATACTATTG
GCATTATAGGGGGGCGGTAACATGATTGACGAAGTATTAGTATTGAACTCAACGCCAACGGG
GATTTTACAGAATTCAGTTTCAGGAAGCATGGTGATTGACGCGGCACTTGATGCACGACTAG
T

5.2.3 *Lactobacillus plantarum* subsp. plantarum (ATCC 14917)

>14917_pts18_500_300_2789bp:CGACTCAAATGTTCAACGAAGCCTATGATGCTTGCGAAAGGG
CAGGCATCAAACTCATTGAGACGCGGCAAGAAGAAATCGACTCGGTTGAAACGGTCAGTCG
AGCCTATAAGTACCATTATCCATCAATGTATCAAGACTTCTCAAAGGGACGGCCAACCGAAG
TTGATTATATTAATGGCTATATTGCTAGGATTGGGCGGGAACATGACTACGTTTGTCGCGTGC
ATGAGTTTGTGACTCACGAAGTCCATTTAGCAGAAATGATGCGGCAATACCGGCACCCGGAA
ATCCCCGTTGCTGAAAAATAAGATTATTAATGAAAAAGTTCTGAGTATTTTCAGAACTTTTTT
TGTGGGCGGAAAAACTTAATTTAGTAAGGTTGAAATCGGTTTCTTTGACATTGGTCTATACAC
TTAGATAACCTTGTAAACTTGCAATGAAAACATGTTGTGATAGGATTTATAACGTGCTTAAA
AAGAGAATGAGGAGAAATTGGTATGAAGACATATTTTCAGAAAATCGGTCAGTCATTGATGC
TGCCGATTGCAACGTTGCCCGCAGCAGCGATTTTGGTTGGGATTGGTAATTACCTGCCAAAA
CAGTGGCTGTTTGCAAATTACTTGATTCAGGGTGGTAACGTTGTCCTGAATAACTTAGCCTTG
CTGTTTGCAGTCGGCTTAGCTATCGGGATGTCAGTCAATAAGGACGGCGCGGCAGCTATTGC
TGGTTTGATCGCGTTTGAAGTTCCGGTGATGGTGTTAAAACCAGCGACATTAGCGACGATGC
TGAATGTTAAAGTGAGTCAGATCAATCCAGCATTTAGCGCGTTAGATAACAATGTGCTGATT
GGGATCAGTGCTGGACTGATTGCGGCTGCGCTCTATAATCGGTTCCACGAAGTAAAATTACC
AATGGCGCTCTCATTCTTTAGTGGTAAACGCCTAGTGCCAATTATGGCTGCTTTTGTGATGCT
GATTGTGACGGCGGTATTGTACCTTGTTTGGCCGTTTGTCTACGATGCCATTGTCCTCTTTGCA
ACGGGGATTTCAAAATTAGGTTTCGTGGGGGCTGGTCTATATGGCTTTTTCAATCGTTTATTG
ATTCCAACTGGTTTGCACCACGCCTTAAATTCCGTATTCTGGTATAACGTTGCTGGTATCAAT
GATATTGGCAATTTCTGGGCCAGCCATGGTGTTAAAGGAATCACAGGGATGTATGAAGCTGG
ATTCTTCCCAATTATGATGTTTGGCTTGCCAGCTGGTGCGTATGCAATTTATCGTAATGCACG

ACCGGAACGAAAAAAGAAAGTGGGTTCATTGATGTTAGCGGGTGCGTTTGCGTCGTTCTTTA
CGGGGGTGACCGAACCGCTTGAATTTTCATTCATGTTCGTGGCATGGCCGCTCTACTTATTAC
ATGCCGTTTTTATGGGACTATCATTAGGATTCGCCGCGTTAATGCACTGGACCGCCAGCTTTT
CGTTTAGTGGCGGGTTAGTCGATTATTTATTGAGTTTTCGGATGCCGTTAGCCAACCAACCCT
ATATGTTACTGGTTCAAGGGCTGGTGATGGCCGTCATTTATTACTTTGGTTTTGACTTTGCAA
TCAAGCGGTTTAATTTGAAGACGCCTGGTCGTGAAGTTGTGTCTGCAGATGTCGATGGGGTG
GGTGCACCTGCTAGTCCAGCGGTGGCCGTTGCAGCAACGGATGATAAATATATGCGCCAAGC
AAAGCAAATTTATGCGGCTATTGGTGGTCATGACAATATTAGTGTGATCAACAATTGTACGA
CGCGGCTGCGGTTACAACTTAAGGATACGGAAAAAGTTGATCAGCCGGCCGTAATGGCTGCT
GGCGTGCCTGGTTTGAACGTACTGGATGTTCATAACATCCACATTGTGATCGGTACGGAAGT
CCAGTTCGTTGCGGAGGCTTTACAAAAATTATTTTCTGGTCAGGTGGCGACGACCCCGGCAT
CTGATGCTGAATCAAAGGCCCCGCTTGAACCGCAAACTGCTACTGTAACGGAAGCGCCGGTG
ACAACGATTTTGCGGGCACCGGCAACGGGACAATTAATGCCGATTAGTGCGGTTGCGGATGA
GACGTTTGCTGGTAAACTCTTAGGTGATGGTTACGCTGTTGAACCCGAAGATGGTGAAGTCG
TTGCTCCGGTTAGTGGGACAGTGACAAGTGTCTTTCCGACCAAACACGCTATCGGTTTGAAG
ACGACGAGTGGTTTAGAAGTCTTATTACACATGGGGATCAATACCGTGGAAATGAATGGTAC
GCCGTTCAAGTTACACGTGGCAACGGGTGATGAAATTGCGGCCGGTACTGCGGTAGCGACGG
TTGACTTGGCTGCTATCAAGTCAGCCGGTAAAGCGACAACCATGATGGTCGTTATCACCAAT
ATGGATCACGTTAACAAACTAACCCTTAATCCAACTGGTCATGTGACTAGCGGTGATTTGATT
GGCGCAGCTGAATAAGTTACGCGCTGTCGACGTTGGGTTTATTGAAAAAGTTTCGTCTAAAT
CTTTTCGATTTAGGCGGAACTTTTTTTGTCGCATTGGTTATCGAGGCAACGGATGACAGCAAA
TGAACAACTTGATGCGGGGACAGTGCGTTTGATTGCTTCCCACTGACTCCCACGATATACTAT
TGGCATTATAGGGGGGCGGTAACATGATTGACGAAGTATTAGTATTGAACTCAACGCCAACG
GGGATTTTACAGAATTCAGTTTCAGGAAGCATGGTGATTGACGCGGCACTTGATGCACGACT
AGT

5.2.4 *pts18cba* sequence alignments

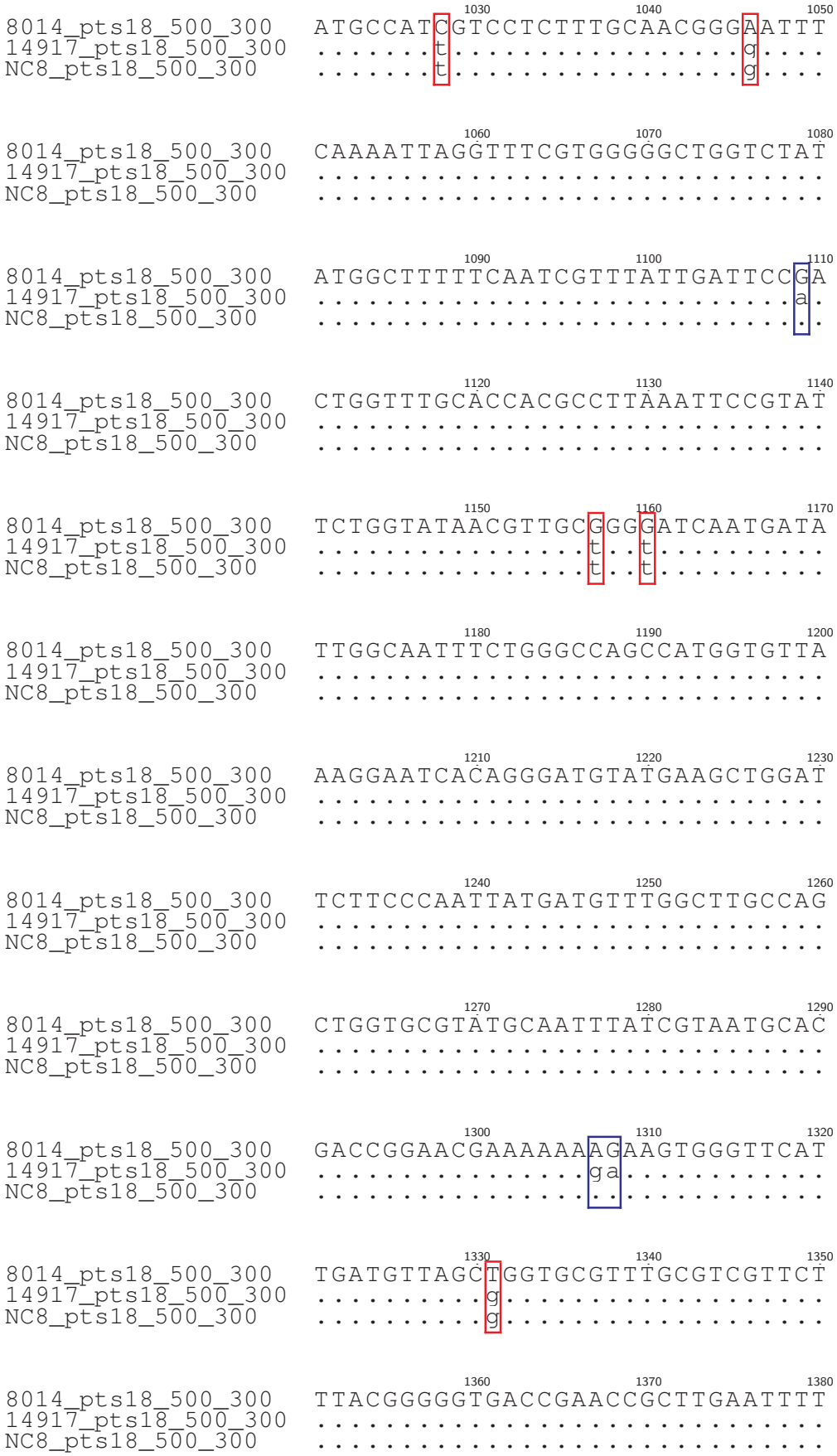
Gene sequence alignments of *pts18cba* from *Lactobacillus plantarum* 8014 compared to *Lactobacillus plantarum* NC8 and *Lactobacillus plantarum* subsp. plantarum (ATCC 14917). Consensus between the sequences is shown as full stops (.) while single-nucleotide

polymorphisms (SNPs) that appear in *Lactobacillus plantarum* 8014 that are not found in the other species are shown as the nucleotide letter surrounded by a red frame. SNPs found in the other genes that do not appear in *Lactobacillus plantarum* 8014 are indicated by the nucleotide letter surrounded by a blue frame.

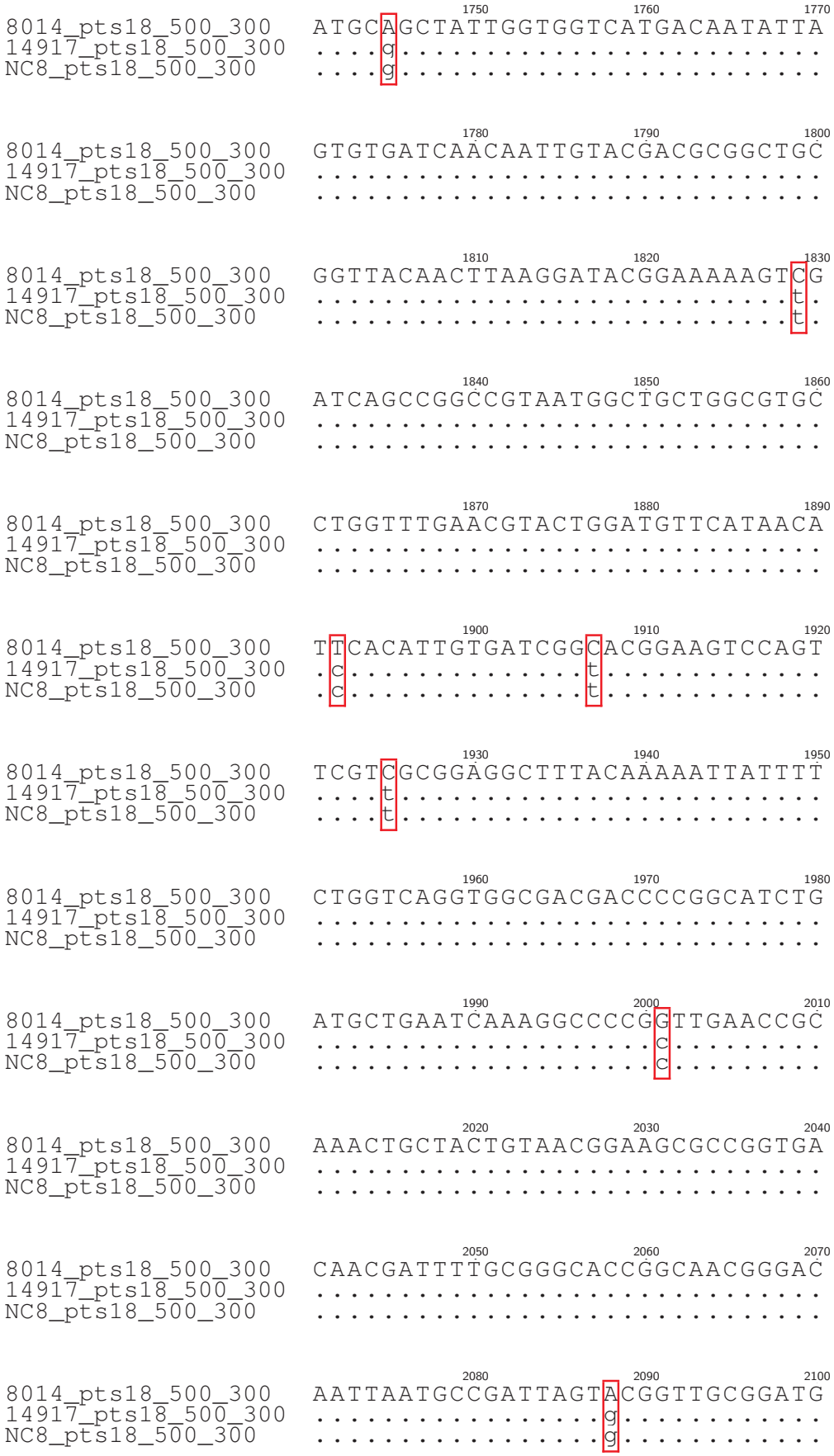
```
                                 10                  20                  30
8014_pts18_500_300   CGACTCAAATGTTCAACGAAGCCTATGATG
14917_pts18_500_300  ..............................
NC8_pts18_500_300    ..............................

                                 40                  50                  60
8014_pts18_500_300   CTTGCGAAAGGGCAGGCATCAAACTCATTG
14917_pts18_500_300  ..............................
NC8_pts18_500_300    ..............................

                                 70                  80                  90
8014_pts18_500_300   AGACGCGGCAAGAAGAAATCGACTCGGTTG
14917_pts18_500_300  ..............................
NC8_pts18_500_300    ..............................

                                100                 110                 120
8014_pts18_500_300   AAACGGTCAGTCGAGCCTATAAGTACCATT
14917_pts18_500_300  ..............................
NC8_pts18_500_300    ..............................

                                130                 140                 150
8014_pts18_500_300   ATCCATCAATGTATCAAGACTTCTCAAAGG
14917_pts18_500_300  ..............................
NC8_pts18_500_300    ..............................

                                160                 170                 180
8014_pts18_500_300   GACGGCCAACCGAAGTTGATTATATTAATG
14917_pts18_500_300  ..............................
NC8_pts18_500_300    ..............................

                                190                 200                 210
8014_pts18_500_300   GCTATATTGCTAAGATTGGGCGGGAACATG
14917_pts18_500_300  ............g.................
NC8_pts18_500_300    ............g.................

                                220                 230                 240
8014_pts18_500_300   ACTACGTTTGTCGCGTGCATGAGTTTGTGA
14917_pts18_500_300  ..............................
NC8_pts18_500_300    ..............................

                                250                 260                 270
8014_pts18_500_300   CTCACGAAGTCCATTTAGCAGAAATGATGC
14917_pts18_500_300  ..............................
NC8_pts18_500_300    ..............................

                                280                 290                 300
8014_pts18_500_300   GGCAATACCGGCACCCGGAAATCCCCGTTG
14917_pts18_500_300  ..............................
NC8_pts18_500_300    ..............................
```

```
                                  310                 320                 330
8014_pts18_500_300   CTGAAAAATAAGATTATTAATGAAAAAGTT
14917_pts18_500_300  ..............................
NC8_pts18_500_300    ..............................

                                  340                 350                 360
8014_pts18_500_300   CTGAGTATTTTCAGAACTTTTTTTGTGGGC
14917_pts18_500_300  ..............................
NC8_pts18_500_300    ..............................

                                  370                 380                 390
8014_pts18_500_300   GGAAAAACTTAATTTAGTAAGGTTGAAATC
14917_pts18_500_300  ..............................
NC8_pts18_500_300    ..............................

                                  400                 410                 420
8014_pts18_500_300   GGTTTCTTTAAACATTGGTCTATACACTTAG
14917_pts18_500_300  .........g....................
NC8_pts18_500_300    .........g....................

                                  430                 440                 450
8014_pts18_500_300   ATGACCTTGTAAACTTGCAATGAAAACATG
14917_pts18_500_300  ..a..........................
NC8_pts18_500_300    ..a..........................

                                  460                 470                 480
8014_pts18_500_300   TTGTGATAGGATTTATAACGTGCTTAAAAA
14917_pts18_500_300  ..............................
NC8_pts18_500_300    ..............................

                                  490                 500                 510
8014_pts18_500_300   GAGAATGAGGAGAAATTGGTATGAAGACAT
14917_pts18_500_300  ..............................
NC8_pts18_500_300    ..............................

                                  520                 530                 540
8014_pts18_500_300   ATTTTCAGAAAATCGGTCAGTCATTGATGC
14917_pts18_500_300  ..............................
NC8_pts18_500_300    ..............................

                                  550                 560                 570
8014_pts18_500_300   TACCGATTGCAACGTTGCCCGCAGCAGCGA
14917_pts18_500_300  .g............................
NC8_pts18_500_300    .g............................

                                  580                 590                 600
8014_pts18_500_300   TTTTGGTTGGGATTGGGAATTACCTGCCAA
14917_pts18_500_300  ................t.............
NC8_pts18_500_300    ................t.............

                                  610                 620                 630
8014_pts18_500_300   AACAGTGGCTGTTTGCAAATTACTTGATTC
14917_pts18_500_300  ..............................
NC8_pts18_500_300    ..............................

                                  640                 650                 660
8014_pts18_500_300   AGGGTGGTAACGTTGTCCTGAATAACTTAG
14917_pts18_500_300  ..............................
NC8_pts18_500_300    ..............................
```

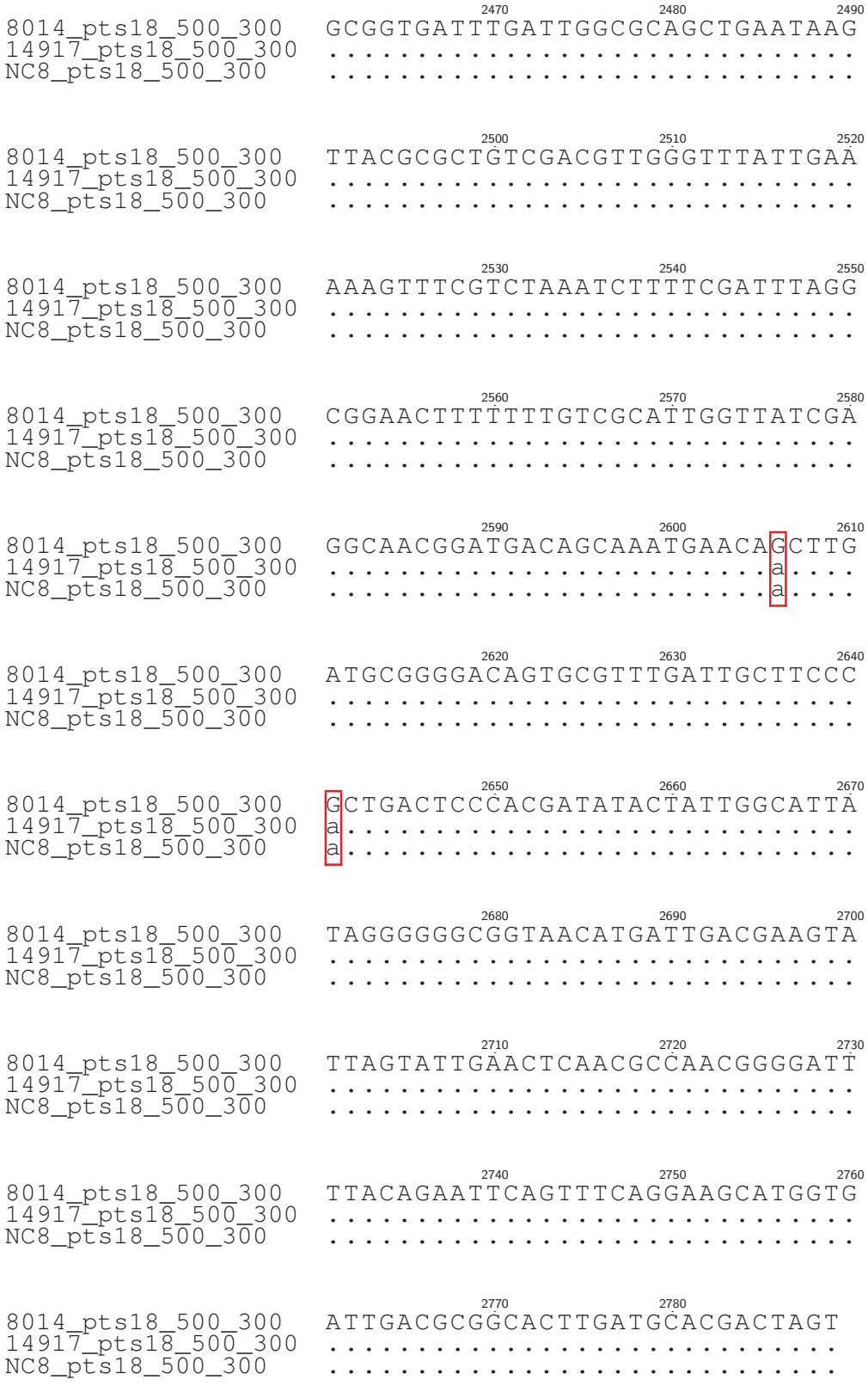
```
                                670                 680                 690
8014_pts18_500_300   CCTTGCTGTTTGCAGTCGGCTTAGCTATCG
14917_pts18_500_300  ..............................
NC8_pts18_500_300    ..............................

                                700                 710                 720
8014_pts18_500_300   GGATGTCAGTCAATAAGGACGGCGCGGCAG
14917_pts18_500_300  ..............................
NC8_pts18_500_300    ..............................

                                730                 740                 750
8014_pts18_500_300   CTATTGCTGGTTTGATCGCGTTTGAAGTTC
14917_pts18_500_300  ..............................
NC8_pts18_500_300    ..............................

                                760                 770                 780
8014_pts18_500_300   CGGTGATGGTTTTTAAAACCAGCGACATTAG
14917_pts18_500_300  ..........g.....................
NC8_pts18_500_300    ..........g.....................

                                790                 800                 810
8014_pts18_500_300   CGACGATGCTGAATGTTAAAGTGAGTCAGA
14917_pts18_500_300  ..............................
NC8_pts18_500_300    ..............................

                                820                 830                 840
8014_pts18_500_300   TCAATCCAGCATTTAGCGCGTTAGATAACA
14917_pts18_500_300  ..............................
NC8_pts18_500_300    ..............................

                                850                 860                 870
8014_pts18_500_300   ATGTGCTGATTGGAATCAGTGCTGGACTGA
14917_pts18_500_300  ............g.................
NC8_pts18_500_300    ............g.................

                                880                 890                 900
8014_pts18_500_300   TTGCGGCTGCGCTCTATAATCGGTTCCACG
14917_pts18_500_300  ..............................
NC8_pts18_500_300    ..............................

                                910                 920                 930
8014_pts18_500_300   AAGTAAAATTACCAATGGCGCTCTCATTCT
14917_pts18_500_300  ..............................
NC8_pts18_500_300    ..............................

                                940                 950                 960
8014_pts18_500_300   TTAGTGGTAAACGCCTAGTGCCAATTATGG
14917_pts18_500_300  ..............................
NC8_pts18_500_300    ..............................

                                970                 980                 990
8014_pts18_500_300   CTGCTTTTGTGATGCTGATTGTGACGGCGG
14917_pts18_500_300  ..............................
NC8_pts18_500_300    ..............................

                               1000                1010                1020
8014_pts18_500_300   TATTGTACCTTGTTTGGCCGTTTGTCTACG
14917_pts18_500_300  ..............................
NC8_pts18_500_300    ..............................
```


1030 1040 1050
8014_pts18_500_300 ATGCCATCGTCCTCTTTGCAACGGGAATTT
14917_pts18_500_300 ....t..................g....
NC8_pts18_500_300 ....t..................g....
1060 1070 1080
8014_pts18_500_300 CAAAATTAGGTTTCGTGGGGGCTGGTCTAT
14917_pts18_500_300 ....
NC8_pts18_500_300 ....
1090 1100 1110
8014_pts18_500_300 ATGGCTTTTTCAATCGTTTATTGATTCCGA
14917_pts18_500_300 ....a.
NC8_pts18_500_300 ....
1120 1130 1140
8014_pts18_500_300 CTGGTTTGCACCACGCCTTAAATTCCGTAT
14917_pts18_500_300 ....
NC8_pts18_500_300 ....
1150 1160 1170
8014_pts18_500_300 TCTGGTATAACGTTGCGGGGATCAATGATA
14917_pts18_500_300 ....t..t..........
NC8_pts18_500_300 ....t..t..........
1180 1190 1200
8014_pts18_500_300 TTGGCAATTTCTGGGCCAGCCATGGTGTTA
14917_pts18_500_300 ....
NC8_pts18_500_300 ....
1210 1220 1230
8014_pts18_500_300 AAGGAATCACAGGGATGTATGAAGCTGGAT
14917_pts18_500_300 ....
NC8_pts18_500_300 ....
1240 1250 1260
8014_pts18_500_300 TCTTCCCAATTATGATGTTTGGCTTGCCAG
14917_pts18_500_300 ....
NC8_pts18_500_300 ....
1270 1280 1290
8014_pts18_500_300 CTGGTGCGTATGCAATTTATCGTAATGCAC
14917_pts18_500_300 ....
NC8_pts18_500_300 ....
1300 1310 1320
8014_pts18_500_300 GACCGGAACGAAAAAAAGAAGTGGGTTCAT
14917_pts18_500_300 ....ga............
NC8_pts18_500_300 ....
1330 1340 1350
8014_pts18_500_300 TGATGTTAGCTGGTGCGTTTGCGTCGTTCT
14917_pts18_500_300 ....g...................
NC8_pts18_500_300 ....g...................
1360 1370 1380
8014_pts18_500_300 TTACGGGGGTGACCGAACCGCTTGAATTTT
14917_pts18_500_300 ....
NC8_pts18_500_300 ....

```
                                1390                1400                1410
8014_pts18_500_300   CATTCATGTTCGTGGCATGGCCGCTCTACT
14917_pts18_500_300  ..............................
NC8_pts18_500_300    ..............................

                                1420                1430                1440
8014_pts18_500_300   TATTACATGCCGTTTTTATGGGACTATCAT
14917_pts18_500_300  ..............................
NC8_pts18_500_300    ..............................

                                1450                1460                1470
8014_pts18_500_300   TAGGATTCGCCGCGTTAATGCACTGGACCG
14917_pts18_500_300  ..............................
NC8_pts18_500_300    ..............................

                                1480                1490                1500
8014_pts18_500_300   CCAGCTTTTCGTTTAGTGGCGGGTTAGTCG
14917_pts18_500_300  ..............................
NC8_pts18_500_300    ..............................

                                1510                1520                1530
8014_pts18_500_300   ATTATTTATTGAGTTTTCGGATGCCGTTAG
14917_pts18_500_300  ..............................
NC8_pts18_500_300    ..............................

                                1540                1550                1560
8014_pts18_500_300   CCAACCAACCCTATATGTTACTGGTTCAAG
14917_pts18_500_300  ..............................
NC8_pts18_500_300    ..............................

                                1570                1580                1590
8014_pts18_500_300   GGCTGGTGATGGCCGTCATTTATTACTTTG
14917_pts18_500_300  ..............................
NC8_pts18_500_300    ..............................

                                1600                1610                1620
8014_pts18_500_300   GTTTTGACTTTGCAATCAAACGGTTTAATT
14917_pts18_500_300  ..................g...........
NC8_pts18_500_300    ..................g...........

                                1630                1640                1650
8014_pts18_500_300   TGAAGACGCCTGGTCGTGAAGTTGTGTCTG
14917_pts18_500_300  ..............................
NC8_pts18_500_300    ..............................

                                1660                1670                1680
8014_pts18_500_300   CAGATGTCGATGGGGTGGGTGCACCTGCTA
14917_pts18_500_300  ..............................
NC8_pts18_500_300    ..............................

                                1690                1700                1710
8014_pts18_500_300   GTCCAGCGGTGGCCGTTGCAGCAACGGATG
14917_pts18_500_300  ..............................
NC8_pts18_500_300    ..............................

                                1720                1730                1740
8014_pts18_500_300   ATAAGTATATGCGCCAAGCAAAGCAAATTT
14917_pts18_500_300  ....a.........................
NC8_pts18_500_300    ....a.........................
```


1750 1760 1770
8014_pts18_500_300 ATGCAGCTATTGGTGGTCATGACAATATTA
14917_pts18_500_300 ....g.....
NC8_pts18_500_300 ....g.....
1780 1790 1800
8014_pts18_500_300 GTGTGATCAACAATTGTACGACGCGGCTGC
14917_pts18_500_300 .....
NC8_pts18_500_300 .....
1810 1820 1830
8014_pts18_500_300 GGTTACAACTTAAGGATACGGAAAAAGTCG
14917_pts18_500_300 .....t.
NC8_pts18_500_300 .....t.
1840 1850 1860
8014_pts18_500_300 ATCAGCCGGCCGTAATGGCTGCTGGCGTGC
14917_pts18_500_300 .....
NC8_pts18_500_300 .....
1870 1880 1890
8014_pts18_500_300 CTGGTTTGAACGTACTGGATGTTCATAACA
14917_pts18_500_300 .....
NC8_pts18_500_300 .....
1900 1910 1920
8014_pts18_500_300 TTCACATTGTGATCGGCACGGAAGTCCAGT
14917_pts18_500_300 .c.....t.....
NC8_pts18_500_300 .c.....t.....
1930 1940 1950
8014_pts18_500_300 TCGTCGCGGAGGCTTTACAAAAATTATTTT
14917_pts18_500_300 ....t.....
NC8_pts18_500_300 ....t.....
1960 1970 1980
8014_pts18_500_300 CTGGTCAGGTGGCGACGACCCCGGCATCTG
14917_pts18_500_300 .....
NC8_pts18_500_300 .....
1990 2000 2010
8014_pts18_500_300 ATGCTGAATCAAAGGCCCCGGGTTGAACCGC
14917_pts18_500_300 .....c.....
NC8_pts18_500_300 .....c.....
2020 2030 2040
8014_pts18_500_300 AAACTGCTACTGTAACGGAAGCGCCGGTGA
14917_pts18_500_300 .....
NC8_pts18_500_300 .....
2050 2060 2070
8014_pts18_500_300 CAACGATTTTGCGGGCACCGGCAACGGGAC
14917_pts18_500_300 .....
NC8_pts18_500_300 .....
2080 2090 2100
8014_pts18_500_300 AATTAATGCCGATTAGTACGGTTGCGGATG
14917_pts18_500_300 .....g.....
NC8_pts18_500_300 .....g.....

```
                                2110                2120                2130
8014_pts18_500_300   AGACGTTTGCTGGTAAACTCTTAGGTGATG
14917_pts18_500_300  ..............................
NC8_pts18_500_300    ..............................

                                2140                2150                2160
8014_pts18_500_300   GTTACGCTGTTGAACCCGAAGATGGTGAAG
14917_pts18_500_300  ..............................
NC8_pts18_500_300    ..............................

                                2170                2180                2190
8014_pts18_500_300   TCGTTGCTCCGGTTAGTGGGACAGTGACAA
14917_pts18_500_300  ..............................
NC8_pts18_500_300    ..............................

                                2200                2210                2220
8014_pts18_500_300   GTGTCTTTCCGACCAAACACGCTATCGGTT
14917_pts18_500_300  ..............................
NC8_pts18_500_300    ..............................

                                2230                2240                2250
8014_pts18_500_300   TGAAGACGACGAGTGGTTTAGAAGTCTTAT
14917_pts18_500_300  ..............................
NC8_pts18_500_300    ..............................

                                2260                2270                2280
8014_pts18_500_300   TACACATGGGGATCAATACCGTGGAAATGA
14917_pts18_500_300  ..............................
NC8_pts18_500_300    ..............................

                                2290                2300                2310
8014_pts18_500_300   ATGGTACGCCGTTCAAGTTACACGTGGCAA
14917_pts18_500_300  ..............................
NC8_pts18_500_300    ..............................

                                2320                2330                2340
8014_pts18_500_300   CGGGTGATGAAATTGCGGCCGGTACTGCGG
14917_pts18_500_300  ..............................
NC8_pts18_500_300    ..............................

                                2350                2360                2370
8014_pts18_500_300   TAGCGACGGTTGACTTGGCTGCTATCAAGT
14917_pts18_500_300  ..............................
NC8_pts18_500_300    ..............................

                                2380                2390                2400
8014_pts18_500_300   CAGCCGGTAAAGCGACAACCATGATGGTCG
14917_pts18_500_300  ..............................
NC8_pts18_500_300    ..............................

                                2410                2420                2430
8014_pts18_500_300   TTATCACCAATATGGATCACGTTAACAAAC
14917_pts18_500_300  ..............................
NC8_pts18_500_300    ..............................

                                2440                2450                2460
8014_pts18_500_300   TAACCCTTAATCCAACTGGTCATGTGACTA
14917_pts18_500_300  ..............................
NC8_pts18_500_300    ..............................
```


2470 2480 2490
8014_pts18_500_300 GCGGTGATTTGATTGGCGCAGCTGAATAAG
14917_pts18_500_300 .....
NC8_pts18_500_300 .....
2500 2510 2520
8014_pts18_500_300 TTACGCGCTGTCGACGTTGGGTTTATTGAA
14917_pts18_500_300 .....
NC8_pts18_500_300 .....
2530 2540 2550
8014_pts18_500_300 AAAGTTTCGTCTAAATCTTTTCGATTTAGG
14917_pts18_500_300 .....
NC8_pts18_500_300 .....
2560 2570 2580
8014_pts18_500_300 CGGAACTTTTTTTGTCGCATTGGTTATCGA
14917_pts18_500_300 .....
NC8_pts18_500_300 .....
2590 2600 2610
8014_pts18_500_300 GGCAACGGATGACAGCAAATGAACAGCTTG
14917_pts18_500_300 .....a....
NC8_pts18_500_300 .....a....
2620 2630 2640
8014_pts18_500_300 ATGCGGGGACAGTGCGTTTGATTGCTTCCC
14917_pts18_500_300 .....
NC8_pts18_500_300 .....
2650 2660 2670
8014_pts18_500_300 GCTGACTCCCACGATATACTATTGGCATTA
14917_pts18_500_300 a.....
NC8_pts18_500_300 a.....
2680 2690 2700
8014_pts18_500_300 TAGGGGGGCGGTAACATGATTGACGAAGTA
14917_pts18_500_300 .....
NC8_pts18_500_300 .....
2710 2720 2730
8014_pts18_500_300 TTAGTATTGAACTCAACGCCAACGGGGATT
14917_pts18_500_300 .....
NC8_pts18_500_300 .....
2740 2750 2760
8014_pts18_500_300 TTACAGAATTCAGTTTCAGGAAGCATGGTG
14917_pts18_500_300 .....
NC8_pts18_500_300 .....
2770 2780
8014_pts18_500_300 ATTGACGCGGCACTTGATGCACGACTAGT
14917_pts18_500_300 .....
NC8_pts18_500_300 .....

5.3 PTS18CBA protein properties and structure

5.3.1 Pts18cba Sequence translation

>8014_pts18:MKTYFQKIGQSLMLPIATLPAAAILVGIGNYLPKQWLFANYLIQGG
NVVLNNLALLFAVGLAIGMSVNKDGAAAIAGLIAFEVPVMVLKPATLATMLNV
KVSQINPAFSALDNNVLIGISAGLIAAALYNRFHEVKLPMALSFFSGKRLVPIMA
AFVMLIVTAVLYLVWPFVYDAIVLFATGISKLGFVGAGLYGFFNRLLIPTGLHH
ALNSVFWYNVAGINDIGNFWASHGVKGITGMYEAGFFPIMMFGLPAGAYAIYR
NARPERKKEVGSLMLAGAFASFFTGVTEPLEFSFMFVAWPLYLLHAVFMGLSL
GFAALMHWTASFSFSGGLVDYLLSFRMPLANQPYMLLVQGLVMAVIYYFGFD
FAIKRFNLKTPGREVVSADVDGVGAPASPAVAVAATDDKYMRQAKQIYAAIGG
HDNISVINNCTTRLRLQLKDTEKVDQPAVMAAGVPGLNVLDVHNIHIVIGTEVQ
FVAEALQKLFSGQVATTPASDAESKAPVEPQTATVTEAPVTTILRAPATGQLMPI
STVADETFAGKLLGDGYAVEPEDGEVVAPVSGTVTSVFPTKHAIGLKTTSGLEV
LLHMGINTVEMNGTPFKLHVATGDEIAAGTAVATVDLAAIKSAGKATTMMVVI
TNMDHVNKLTLNPTGHVTSGDLIGAAE

This amino acid sequence was specifically translated from *Lactobacillus plantarum* 8014 using a reading frame that totally encompasses the gene without any internal stop codons. The nucleotide sequence used was modeled to the same open reading frames found within the constructs developed during this study, and the constructs retrospectively checked using gene sequencing as confirmation.

5.3.2 ExPASy ProtParam

Table 17: Amino acid composition of PTS18CBA

Amino acid	Quantity	Percentage
Ala (A)	86	13.0%
Arg (R)	13	2.0%
Asn (N)	29	4.4%
Asp (D)	22	3.3%
Cys (C)	1	0.2%
Gln (Q)	16	2.4%
Glu (E)	22	3.3%
Gly (G)	60	9.1%
His (H)	14	2.1%
Ile (I)	39	5.9%
Leu (L)	71	10.7%
Lys (K)	27	4.1%
Met (M)	26	3.9%
Phe (F)	39	5.9%
Pro (P)	33	5.0%
Ser (S)	30	4.5%
Thr (T)	43	6.5%
Trp (W)	6	0.9%
Tyr (Y)	19	2.9%
Val (V)	66	10.0%
Pyl (O)	0	0.0%
Sec (U)	0	0.0%
(B)	0	0.0%
(Z)	0	0.0%
(X)	0	0.0%

5.3.3 Secondary structure of PTS18CBA by in silico analysis

5.3.3.1 TMpred helices identification

A)

TMPred: all possible inside to outside helices

From	To	Score	Center
12 (14)	33 (33)	1599	22
49 (49)	68 (68)	1956	58
53 (62)	87 (79)	985	70
71 (73)	94 (92)	767	84
110 (110)	129 (129)	1435	120
149 (151)	169 (169)	2657	159
169 (175)	198 (193)	1325	185
208 (210)	228 (228)	625	219
239 (241)	260 (260)	1377	250
272 (272)	296 (289)	1231	281
307 (307)	326 (326)	1931	317
346 (348)	365 (365)	1652	357
552 (554)	574 (574)	153	564
600 (602)	626 (619)	68	610

B)

TMPred: all possible outside to inside helices

From	To	Score	Center
14 (14)	33 (31)	1614	22
48 (48)	67 (65)	1581	57
71 (71)	88 (88)	1096	80
114 (114)	130 (130)	1341	122
149 (149)	167 (167)	2717	159
168 (175)	197 (194)	1589	185
200 (200)	219 (219)	439	210
242 (242)	259 (259)	1684	250
305 (305)	322 (322)	1911	313
350 (350)	369 (369)	1456	358
381 (381)	401 (401)	146	391
554 (554)	574 (574)	211	564
602 (604)	623 (620)	334	612

C)

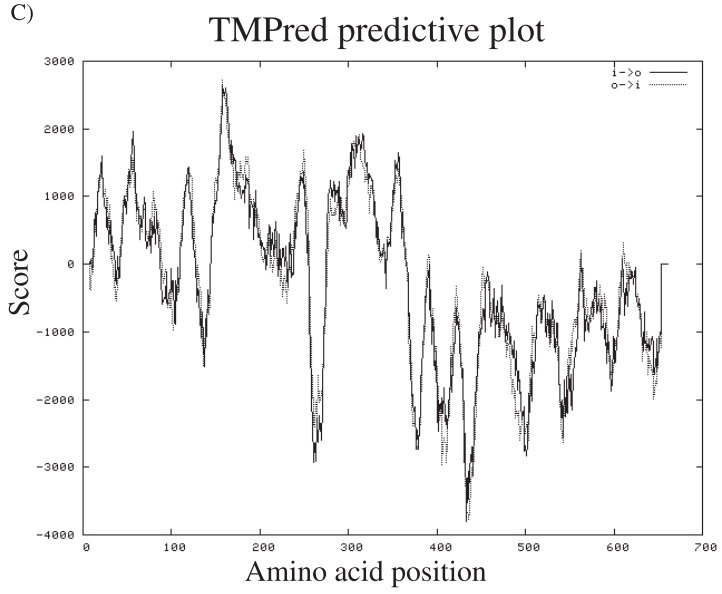


Figure 49: **TMpred: helices identifications**
(A) TMpred helical predictions for all possible helices that have originate from inside the cell and terminate outside the cell. **(B)** TMpred helical predictions for all possible helices that have originate from outside the cell and terminate inside the cell. **(C)** A TMpred hydrophobicity plot of the amino acid position against its relative hydrophobicity score used to determine the reliability of helical formations.

5.3.3.2 Topcons analysis & NCBI CDD cross referenced with Pfam

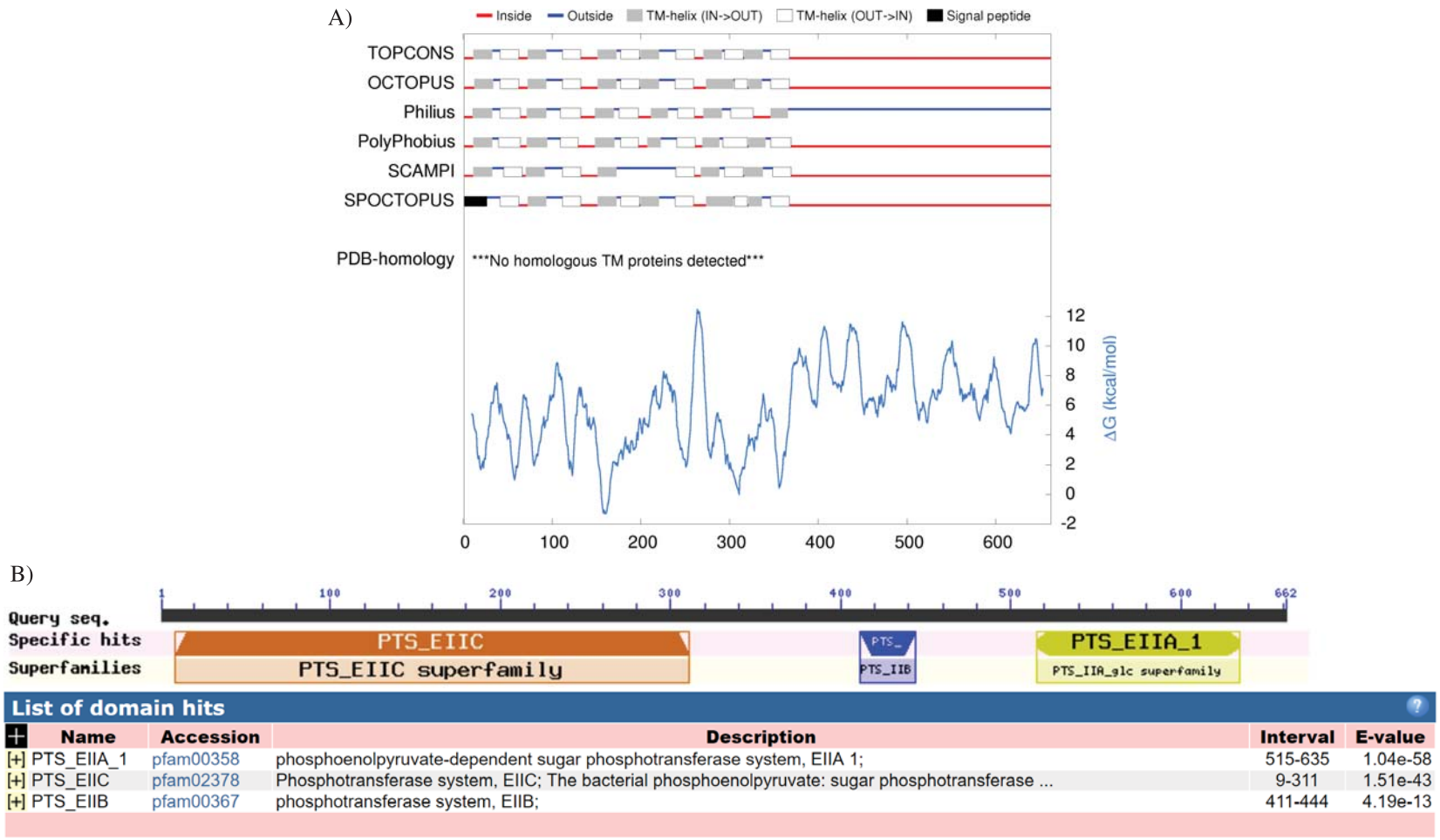


Figure 50: **TopCons and NCBI CDD predictions and domain identification**

(A) The graphical output from the TopCons helical prediction web server. TopCons makes use of the other web servers to predict possible helical formations and then compares them to find consensus between them that is represented as the TopCons output on the top line. **(B)** NCBI blast search cross referenced with pfam to identify any known structural regions. The search identified that regions of the submitted sequence are conserved to three different known superfamilies, the PEP-PTS EIIA (Green), EIIB (Blue) and EIIC (Brown).

5.3.3.3 Full MINNOU analysis with lipid accessibility

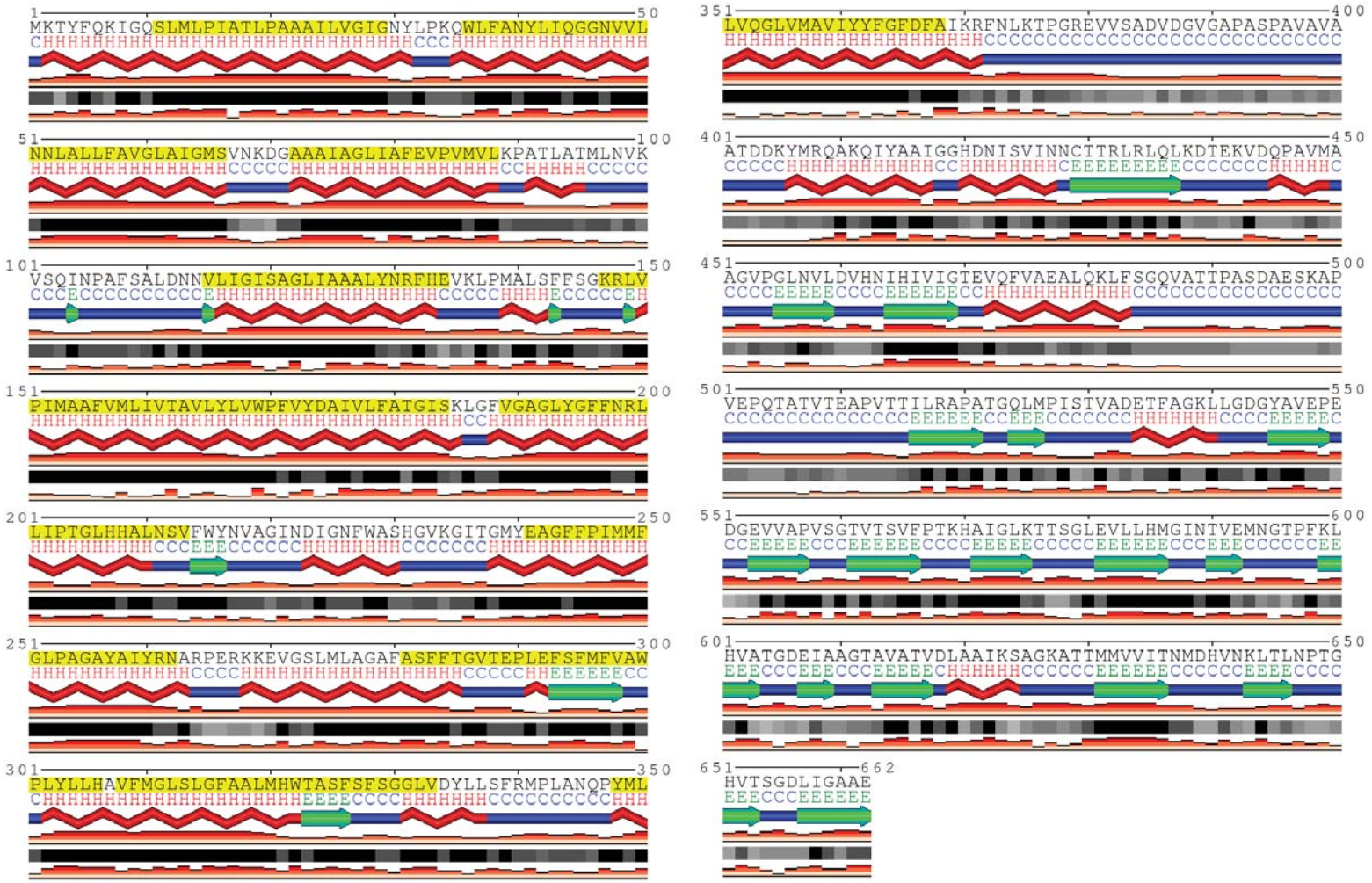


Figure 51: **MINNOU: Full primary & predicted secondary structure**
The graphical summary of the MINNOU secondary structure web server analysis. Amino acid number and identity shown by amino acid lettering and number bar (Grey). Hydrophobic amino acids predicted to interact with lipids are shown by yellow high lighting of the amino acids lettering. predicted secondary structure is shown as: helices = H and zig zags (Red), coils = C and bars (Blue), β-strands = E or directional arrows (Green). The confidence of each prediction is represented as a red probability graph and complementary coverage overlay (Black).

5.3.3.4 Full Jpred secondary structure & residue burial prediction

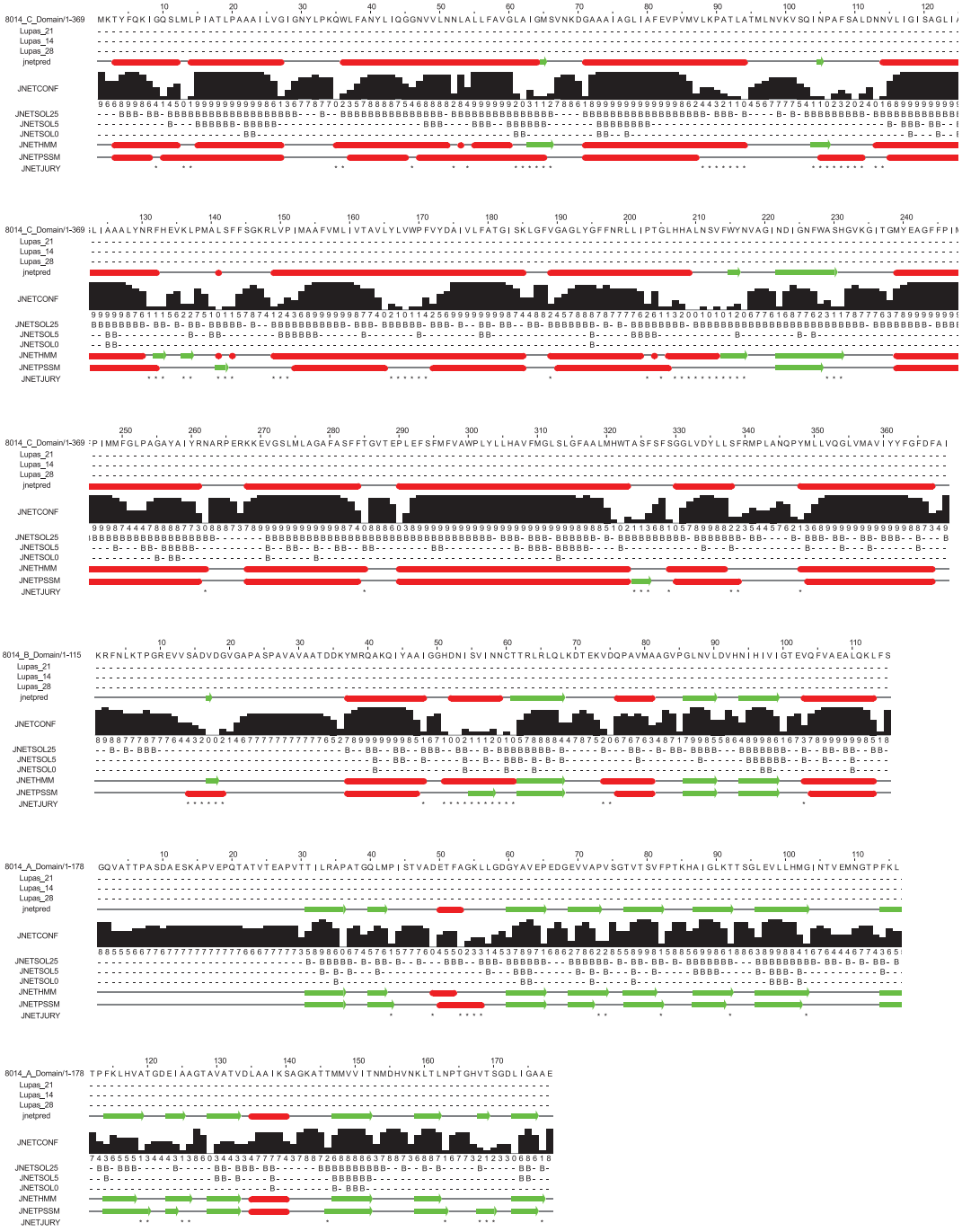


Figure 52: **Jpred: Full batch results separated into C, B and A domain**

The amino acid sequence and position are represented by the single amino acid nomenclature and measurement bar (first two lines). Helices are shown as red Bars, β-strands are indicated by green directional arrows and unstructured regions are left blank. The confidence of each amino acids predictions is shown by a score out of 10 and a complementary bar chart, the higher the score the more confident. Solvent accessible residues are represented by "-" while hydrophobic residues that either interact with lipids or fold into the hydrophobic core are represented with "B" for buried.

5.3.3.5 Phyre^2 Secondary structure and disorder predictions

Figure 53: **Phyre^2 graphical representation of secondary structure**

The amino acids sequence is represented by single letter nomenclature. Secondary structure is indicate by using spirals to indicate helical formations, directional arrows represent β-strands and blank spaces show coiled or unstructured regions. The surface of each structure predicted to interact with a lipid interface is shown by an alternative gold colouration of predicted secondary structural graphics. Secondary structure confidence is shown as a coloured bar that ranges from high confidence (Red) to low confidence (Blue). The disorder of the structure is indicated by the "?" symbol.

5.4 Full Construct sequences

5.4.1 pRV613-LacZ

ATAGCGGCGCATTAAGCGCGGGTGTGGTGGTTACGCGCAGCGTGACCGCTACACTTGCCAGC
GCCCTAGCGCCCGCTCCTTTCGCTTTCTTCCCTTCCTTTCTCGCCACGTTCGCCATCCATGGAA
GCTGTCAGTAGTATACCTAATAATTTATCTACATTCCCTTTAGTAACGTGTAACTTTCCAAAT
TTACAAAAGCGACTCATAGAATTATTTCCTCCCGTTAAATAATAGATAACTATTAAAAATAG
ACAATACTTGCTCATAAGTAACGGTACTTAAATTGTTTACTTTGGCGTGTTTCATTGCTTGAT
GAAACTGATTTTTAGTAAACAGTTGACGATATTCTCGATTGACCCATTTTGAAACAAAGTAC
GTATATAGCTTCCAATATTTATCTGGAACATCTGTGGTATGGCGGGTAAGTTTTATTAAGACA
CTGTTTACTTTTGGTTTAGGATGAAAGCATTCCGCTGGCAGCTTAAGCAATTGCTGAATCGAG
ACTTGAGTGTGCAAGAGCAACCCTAGTGTTCGGTGAATATCCAAGGTACGCTTGTAGAATCC
TTCTTCAACAATCAGATAGATGTCAGACGCATGGCTTTCAAAAACCACTTTTTTAATAATTTG
TGTGCTTAAATGGTAAGGAATACTCCCAACAATTTTATACCTCTGTTTGTTAGGGAATTGAAA
CTGTAGAATATCTTGGTGAATTAAAGTGACACGAGTATTCAGTTTTAATTTTTCTGACGATAA
GTTGAATAGATGACTGTCTAATTCAATAGACGTTACCTGTTTACTTATTTTAGCCAGTTTCGT
CGTTAAATGCCCTTTACCTGTTCCAATTTCGTAAACGGTATCGGTTTCTTTTAAATTCAATTGT
TTTATTATTTGGTTGAGTACTTTTTCACTCGTTAAAAAGTTTTGAGAATATTTTATATTTTTGT
TCATGTAATCACTCCTTCTTAATTACAAATTTTTAGCATCTAATTTAACTTCAATTCCTATTAT
ACAAAATTTTAAGATACTGCACTATCAACACACTCTTAAGTTTGCTTCTAAGTCTTATTTCCA
TAACTTCTTTTACGTTTCCGCCATTCTTTGCTGTTTCGATATGATATCAATTCATCAGCAAGTT
CTCTAATAGTTTTAGGCATGTTTTCCAAACCCCTTTCGGAGTTCAGCAAGCATTTCTTGTCTTG
TCTGATCTCTTATTTTTTTATCATGCTCAATCTGTCGTTCGGCTAAAATCTGTTTTTCAGTTGA
GCCTAGTGGTAGGCCTTTAACTTTATCGGTAGCGCGCCATTTTTCTTGTTCTGTTAATTCACCA
TTATTCTGGATATTAAAAAGTTTTTGACGCTCGTCTTGGAACTTACCTTGGGAAAAGTCGTTT
GCGTCTTTCTTTTCAGGTTTCCAAGTAAAGGTATAACCAATCACTGGTTTTCCACGACCTTTA
CCATATTTTTTCCTAACCGTTAACCCTCTAAATAAGGGAGTTAATTCTTCTTTAATGGGTTTTA
TAACAAATTTATCAACGTTAGATGGACTACTCCAGTAACTTTTAGGCATATCGAGTAATTCAA
AAAAATCTTCTTTAGAAAAATAAGCATATCCAGTGGTTCGGAACTGTTTTAGTAATCGAAAC
ATAGTTTTTGCGTAACTACTCTTCAAATCTCTGAACTCTGCTAGAGCATAACGAACCCAACTT
TCAAGCTTGTTTAAAAGGGGTAAAGCACGTTCATAAACTTTTACGTCTACATAAGGTTCCTTT
GCATCTCCATCGATTTTAAACTCTGTAAACAAGACAAAAAACTCTCGAGTTAATCCACTTTTA
CTTCGCCTACCAAAATGTAATCCCATCATTTTTTCATAAGTTCTCTGAATATCATCTTCAAAA
CGGTTATTTGCAGTTGGTTTATAAGCACTTAATTCTTTTAATTGATCAAAAGTAAAGCGAATG
GTCTGATTACTCTTATCACGCATACGAGAAATAATTGAAAAAAATAAATTCATTTCAACCGG
AGTAAATTTTCGAAGTGGAATAGTATTCAACTCTGGATCATATTTAATAATTTCGTTTCCCAT

GCTGCACCTCCTAATCATTACAAATTTAATATAAGTTAATTTTATCAAGTAGTCAACTTTATA
ATACTACTTGATAAACAGACAAAACCTACTCGATAAACAGACAAAACCTACTTGATAAACAG
ACAAAACCTACTTGATAAACAGACAAAACCTACTTGATAAAGTTCTGTAAGTCTTGGGGGAG
TAAGGCTCAAGAGGGGTCTAAAAGAGGTTTAAAAGAGGTTTATAAAAGAGGTATATAAGAG
GCACCACTGTACGAGATCAAAACGGGCCCATATCATGGCTTTCCCCGTCAAGCTCTAAATCG
GGGGCTCCCTTTAGGGTTCCGATTTAGTGCTTTACGGCACCTCGACCCCAAAAAACTTGATTA
GGGTGATGGTTCACGTAGTGGGCCATCGCCCTGATAGACGGTTTTTCGCCCTTTGACGTTGGA
GTCCACGTTCTTTAATAGTGGACTCTTGTTCCAAACTGGAACAACACTCAACCCTATCTCGGT
CTATTCTTTTGATTTATAAGGGATTTTGCCGATTTCGGCCTATTGGTTAAAAAATGAGCTGAT
TTAACAAAAATTTAACGCGAATTTTAACAAAATATTAACGCTTACAATTCGCGCGTAATACG
ACTCACTATAGGGCGAATTGGGTACCGAATTCCTGCAGCCCGGGTGTTAAGACAAAAGTAAT
TTTGTCTTAACACCCTTTTTTTGTACCCCAAAAGTATTAATGATTGACAGAAGTGCTTCTTGAT
TGTAAACTGAATATATAATTACAAATGTAAACGAAAGAGTGTGAGCAAGATGGCAGAAAAT
ACAAATGAAATTACGACTTCCGAATGGGAAGTGATGCGCATTGTCTGGTCACTTGGTCAAGT
CAACAGTCGTGACTTGATTGATTTACTTCAACCAAAGCGCGATTGGCAAGATTCAACCATCA
AAACATTGATTGGCCGCTTGGTTAAAAAAGGCTTCTTAAAAACAGAAAAGGAAGGGCGGCG
GTTTAATTACACAGCAACCGTTCCTGAGATTGAAGCGATGGAAAATGCCACACAAAGTCTTT
TTGAGCACTTGTGTGGCATGAAAAAAGGGCAGACATTAGCGGCTTTAATTGATCAAACGACT
TTGAGTCAGACGGATATTTTGCAGTTACAACAATTACTGACAGCCAAAGCAGCTACTGCACC
AGAAAAAGTGGCTTGCGACTGTTTGCCAACAACATACGAGCCGGAAGCATAAAGTGTAAAG
ATAAATGTGACTGTGAAAAGGAGGAATAACGATGAAACACGATGGATCCCCAGCTTGTTGA
TACACTAATGCTTTTATATAGGGAAAAGGTGGTGAACTACTGTGGAAGTTACTGACGTAAGA
TTACGGGTCGACCGGGAAAACCCTGGCGTTACCCAACTTAATCGCCTTGCAGCACATCCCCC
TTTCGCCAGCTGGCGTAATAGCGAAGAGGCCCGCACCGATCGCCCTTCCCAACAGTTGCGCA
GCCTGAATGGCGAATGGCGCTTTGCCTGGTTTCCGGCACCAGAAGCGGTGCCGGAAAGCTGG
CTGGAGTGCGATCTTCCTGAGGCCGATACTGTCGTCGTCCCCTCAAACTGGCAGATGCACGG
TTACGATGCGCCCATCTACACCAACGTGACCTATCCCATTACGGTCAATCCGCCGTTTGTTCC
CACGGAGAATCCGACGGGTTGTTACTCGCTCACATTTAATGTTGATGAAAGCTGGCTACAGG
AAGGCCAGACGCGAATTATTTTTGATGGCGTTAACTCGGCGTTTCATCTGTGGTGCAACGGG
CGCTGGGTCGGTTACGGCCAGGACAGTCGTTTGCCGTCTGAATTTGACCTGAGCGCATTTTTA
CGCGCCGGAGAAAACCGCCTCGCGGTGATGGTGCTGCGCTGGAGTGACGGCAGTTATCTGGA
AGATCAGGATATGTGGCGGATGAGCGGCATTTTCCGTGACGTCTCGTTGCTGCATAAACCGA
CTACACAAATCAGCGATTTCCATGTTGCCACTCGCTTTAATGATGATTTCAGCCGCGCTGTAC
TGGAGGCTGAAGTTCAGATGTGCGGCGAGTTGCGTGACTACCTACGGGTAACAGTTTCTTTA
TGGCAGGGTGAAACGCAGGTCGCCAGCGGCACCGCGCCTTTCGGCGGTGAAATTATCGATGA
GCGTGGTGGTTATGCCGATCGCGTCACACTACGTCTGAACGTCGAAAACCCGAAACTGTGGA
GCGCCGAAATCCCGAATCTCTATCGTGCGGTGGTTGAACTGCACACCGCCGACGGCACGCTG

ATTGAAGCAGAAGCCTGCGATGTCGGTTTCCGCGAGGTGCGGATTGAAAATGGTCTGCTGCT
GCTGAACGGCAAGCCGTTGCTGATTCGAGGCGTTAACCGTCACGAGCATCATCCTCTGCATG
GTCAGGTCATGGATGAGCAGACGATGGTGCAGGATATCCTGCTGATGAAGCAGAACAACTTT
AACGCCGTGCGCTGTTCGCATTATCCGAACCATCCGCTGTGGTACACGCTGTGCGACCGCTA
CGGCCTGTATGTGGTGGATGAAGCCAATATTGAAACCCACGGCATGGTGCCAATGAATCGTC
TGACCGATGATCCGCGCTGGCTACCGGCGATGAGCGAACGCGTAACGCGAATGGTGCAGCG
CGATCGTAATCACCCGAGTGTGATCATCTGGTCGCTGGGGAATGAATCAGGCCACGGCGCTA
ATCACGACGCGCTGTATCGCTGGATCAAATCTGTCGATCCTTCCCGCCCGGTGCAGTATGAA
GGCGGCGGAGCCGACACCACGGCCACCGATATTATTTGCCCGATGTACGCGCGCGTGGATGA
AGACCAGCCCTTCCCGGCTGTGCCGAAATGGTCCATCAAAAAATGGCTTTCGCTACCTGGAG
AGACGCGCCCGCTGATCCTTTGCGAATACGCCCACGCGATGGGTAACAGTCTTGGCGGTTTC
GCTAAATACTGGCAGGCGTTTCGTCAGTATCCCCGTTTACAGGGCGGCTTCGTCTGGGACTG
GGTGGATCAGTCGCTGATTAAATATGATGAAAACGGCAACCCGTGGTCGGCTTACGGCGGTG
ATTTTGGCGATACGCCGAACGATCGCCAGTTCTGTATGAACGGTCTGGTCTTTGCCGACCGCA
CGCCGCATCCAGCGCTGACGGAAGCAAAACACCAGCAGCAGTTTTTCCAGTTCCGTTTATCC
GGGCAAACCATCGAAGTGACCAGCGAATACCTGTTCCGTCATAGCGATAACGAGCTCCTGCA
CTGGATGGTGGCGCTGGATGGTAAGCCGCTGGCAAGCGGTGAAGTGCCTCTGGATGTCGCTC
CACAAGGTAAACAGTTGATTGAACTGCCTGAACTACCGCAGCCGGAGAGCGCCGGGCAACT
CTGGCTCACAGTACGCGTAGTGCAACCGAACGCGACCGCATGGTCAGAAGCCGGGCACATC
AGCGCCTGGCAGCAGTGGCGTCTGGCGGAAAACCTCAGTGTGACGCTCCCCGCCGCGTCCCA
CGCCATCCCGCATCTGACCACCAGCGAAATGGATTTTTGCATCGAGCTGGGTAATAAGCGTT
GGCAATTTAACCGCCAGTCAGGCTTTCTTTCACAGATGTGGATTGGCGATAAAAAACAACTG
CTGACGCCGCTGCGCGATCAGTTCACCCGTGCACCGCTGGATAACGACATTGGCGTAAGTGA
AGCGACCCGCATTGACCCTAACGCCTGGGTCGAACGCTGGAAGGCGGCGGGCCATTACCAG
GCCGAAGCAGCGTTGTTGCAGTGCACGGCAGATACACTTGCTGATGCGGTGCTGATTACGAC
CGCTCACGCGTGGCAGCATCAGGGGAAAACCTTATTTATCAGCCGGAAAACCTACCGGATTG
ATGGTAGTGGTCAAATGGCGATTACCGTTGATGTTGAAGTGGCGAGCGATACACCGCATCCG
GCGCGGATTGGCCTGAACTGCCAGCTGGCGCAGGTAGCAGAGCGGGTAAACTGGCTCGGAT
TAGGGCCGCAAGAAAACTATCCCGACCGCCTTACTGCCGCCTGTTTTGACCGCTGGGATCTG
CCATTGTCAGACATGTATACCCCGTACGTCTTCCCGAGCGAAAACGGTCTGCGCTGCGGGAC
GCGCGAATTGAATTATGGCCCACACCAGTGGCGCGGCGACTTCCAGTTCAACATCAGCCGCT
ACAGTCAACAGCAACTGATGGAAACCAGCCATCGCCATCTGCTGCACGCGGAAGAAGGCAC
ATGGCTGAATATCGACGGTTTCCATATGGGGATTGGTGGCGACGACTCCTGGAGCCCGTCAG
TATCGGCGGAATTACAGCTGAGCGCCGGTCGCTACCATTACCAGTTGGTCTGGTGTCAAAAT
AAATAATAACCGGGCAGGCCATGTCTGCCCGTATTTCGCGTAAGGAAATCCATTCAAACTCG
GGCAGCGTTGGGTCCTGGGGGATCCACTAGTTCTAGAGCGGCCGCCACCGCGGTGGAGCTCC
AGCTTTTGTTCCCTTTAGTGAGGGTTAATTGCGCGCTTGGCGTAATCATGGTCATAGCTGTTT

CCTGTGTGAAATTGTTATCCGCTCACAATTCCACACCTGGGGTGCCTAATGAGTGAGCTAACT
CACATTAATTGCGTTGCGCTCACTGCCCGCTTTCCAGTCGGGAAACCTGTCGTGCCAGCTGCA
TTAATGAATCGGCCAACGCGCGGGGAGAGGCGGTTTGCGTATTGGGCGCTCTTCCGCTTCCT
CGCTCACTGACTCGCTGCGCTCGGTCGTTCGGCTGCGGCGAGCGGTATCAGCTCACTCAAAG
GCGGTAATACGGTTATCCACAGAATCAGGGGATAACGCAGGAAAGAACATGTGAGCAAAAG
GCCAGCAAAAGGCCAGGAACCGTAAAAAGGCCGCGTTGCTGGCGTTTTTCCATAGGCTCCGC
CCCCCTGACGAGCATCACAAAAATCGACGCTCAAGTCAGAGGTGGCGAAACCCGACAGGAC
TATAAAGATACCAGGCGTTTCCCCCTGGAAGCTCCCTCGTGCGCTCTCCTGTTCCGACCCTGC
CGCTTACCGGATACCTGTCCGCCTTTCTCCCTTCGGGAAGCGTGGCGCTTTCTCATAGCTCAC
GCTGTAGGTATCTCAGTTCGGTGTAGGTCGTTCGCTCCAAGCTGGGCTGTGTGCACGAACCCC
CCGTTCAGCCCGACCGCTGCGCCTTATCCGGTAACTATCGTCTTGAGTCCAACCCGGTAAGAC
ACGACTTATCGCCACTGGCAGCAGCCACTGGTAACAGGATTAGCAGAGCGAGGTATGTAGG
CGGTGCTACAGAGTTCTTGAAGTGGTGGCCTAACTACGGCTACACTAGAAGGACAGTATTTG
GTATCTGCGCTCTGCTGAAGCCAGTTACCTTCGGAAAAAGAGTTGGTAGCTCTTGATCCGGC
AAACAAACCACCGCTGGTAGCGGTGGTTTTTTTGTTTGCAAGCAGCAGATTACGCGCAGAAA
AAAAGGATCTCAAGAAGATCCTTTGATCTTTTCTACGGGGTCTGACGCTCAGTGGAACGAAA
ACTCACGTTAAGGGATTTTGGTCATGAGATTATCAAAAAGGATCTTCACCTAGATCCTTTTAA
ATTAAAAATGAAGTTTTAAATCAATCTAAAGTATATATGAGTAAACTTGGTCTGACAGTTAC
CAATGCTTAATCAGTGAGGCACCTATCTCAGCGATCTGTCTATTTCGTTCATCCATAGTTGCC
TGACTCCCCGTCGTGTAGATAACTACGATACGGGAGGGCTTACCATCTGGCCCCAGTGCTGC
AATGATACCGCGAGACCCACGCTCACCGGCTCCAGATTTATCAGCAATAAACCAGCCAGCCG
GAAGGGCCGAGCGCAGAAGTGGTCCTGCAACTTTATCCGCCTCCATCCAGTCTATTAATTGTT
GCCGGGAAGCTAGAGTAAGTAGTTCGCCAGTTAATAGTTTGCGCAACGTTGTTGCCATTGCT
ACAGGCATCGTGGTGTCACGCTCGTCGTTTGGTATGGCTTCATTCAGCTCCGGTTCCCAACGA
TCAAGGCGAGTTACATGATCCCCCATGTTGTGCAAAAAAGCGGTTAGCTCCTTCGGTCCTCC
GATCGTTGTCAGAAGTAAGTTGGCCGCAGTGTTATCACTCATGGTTATGGCAGCACTGCATA
ATTCTCTTACTGTCATGCCATCCGTAAGATGCTTTTCTGTGACTGGTGAGTACTCAACCAAGT
CATTCTGAGAATAGTGTATGCGGCGACCGAGTTGCTCTTGCCCGGCGTCAATACGGGATAAT
ACCGCGCCACATAGCAGAACTTTAAAAGTGCTCATCATTGGAAAACGTTCTTCGGGGCGAAA
ACTCTCAAGGATCTTACCGCTGTTGAGATCCAGTTCGATGTAACCCACTCGTGCACCCAACTG
ATCTTCAGCATCTTTTACTTTCACCAGCGTTTCTGGGTGAGCAAAAACAGGAAGGCAAAATG
CCGCAAAAAAGGGAATAAGGGCGACACGGAAATGTTGAATACTCATACTCTTCCTTTTTCAA
TATTATTGAAGCATTTATCAGGGTTA

5.4.2 pRV613:pts18cba-FLAG

ATAGCGGCGCATTAAGCGCGGGTGTGGTGGTTACGCGCAGCGTGACCGCTACACTTGCCAGC
GCCCTAGCGCCCGCTCCTTTCGCTTTCTTCCCTTCCTTTCTCGCCACGTTCGCCATCCATGGAA
GCTGTCAGTAGTATACCTAATAATTTATCTACATTCCCTTTAGTAACGTGTAACTTTCCAAAT
TTACAAAAGCGACTCATAGAATTATTTCCTCCCGTTAAATAATAGATAACTATTAAAAATAG
ACAATACTTGCTCATAAGTAACGGTACTTAAATTGTTTACTTTGGCGTGTTTCATTGCTTGAT
GAAACTGATTTTTAGTAAACAGTTGACGATATTCTCGATTGACCCATTTTGAAACAAAGTAC
GTATATAGCTTCCAATATTTATCTGGAACATCTGTGGTATGGCGGGTAAGTTTTATTAAGACA
CTGTTTACTTTTGGTTTAGGATGAAAGCATTCCGCTGGCAGCTTAAGCAATTGCTGAATCGAG
ACTTGAGTGTGCAAGAGCAACCCTAGTGTTCGGTGAATATCCAAGGTACGCTTGTAGAATCC
TTCTTCAACAATCAGATAGATGTCAGACGCATGGCTTTCAAAAACCACTTTTTTAATAATTTG
TGTGCTTAAATGGTAAGGAATACTCCCAACAATTTTATACCTCTGTTTGTTAGGGAATTGAAA
CTGTAGAATATCTTGGTGAATTAAAGTGACACGAGTATTCAGTTTTAATTTTTCTGACGATAA
GTTGAATAGATGACTGTCTAATTCAATAGACGTTACCTGTTTACTTATTTTAGCCAGTTTCGT
CGTTAAATGCCCTTTACCTGTTCCAATTTCGTAAACGGTATCGGTTTCTTTTAAATTCAATTGT
TTTATTATTTGGTTGAGTACTTTTTCACTCGTTAAAAAGTTTTGAGAATATTTTATATTTTTGT
TCATGTAATCACTCCTTCTTAATTACAAATTTTTAGCATCTAATTTAACTTCAATTCCTATTAT
ACAAAATTTTAAGATACTGCACTATCAACACACTCTTAAGTTTGCTTCTAAGTCTTATTTCCA
TAACTTCTTTTACGTTTCCGCCATTCTTTGCTGTTTCGATATGATATCAATTCATCAGCAAGTT
CTCTAATAGTTTTAGGCATGTTTTCCAAACCCCTTTCGGAGTTCAGCAAGCATTTCTTGTCTTG
TCTGATCTCTTATTTTTTTATCATGCTCAATCTGTCGTTCGGCTAAAATCTGTTTTTCAGTTGA
GCCTAGTGGTAGGCCTTTAACTTTATCGGTAGCGCGCCATTTTTCTTGTTCTGTTAATTCACCA
TTATTCTGGATATTAAAAAGTTTTTGACGCTCGTCTTGGAACTTACCTTGGGAAAAGTCGTTT
GCGTCTTTCTTTTCAGGTTTCCAAGTAAAGGTATAACCAATCACTGGTTTTCCACGACCTTTA
CCATATTTTTTCCTAACCGTTAACCCTCTAAATAAGGGAGTTAATTCTTCTTTAATGGGTTTTA
TAACAAATTTATCAACGTTAGATGGACTACTCCAGTAACTTTTAGGCATATCGAGTAATTCAA
AAAAATCTTCTTTAGAAAAATAAGCATATCCAGTGGTTCGGAACTGTTTTAGTAATCGAAAC
ATAGTTTTTGCGTAACTACTCTTCAAATCTCTGAACTCTGCTAGAGCATAACGAACCCAACTT
TCAAGCTTGTTTAAAAGGGGTAAAGCACGTTCATAAACTTTTACGTCTACATAAGGTTCCTTT
GCATCTCCATCGATTTTAAACTCTGTAAACAAGACAAAAAACTCTCGAGTTAATCCACTTTTA
CTTCGCCTACCAAAATGTAATCCCATCATTTTTTCATAAGTTCTCTGAATATCATCTTCAAAA
CGGTTATTTGCAGTTGGTTTATAAGCACTTAATTCTTTTAATTGATCAAAAGTAAAGCGAATG
GTCTGATTACTCTTATCACGCATACGAGAAATAATTGAAAAAAATAAATTCATTTCAACCGG
AGTAAATTTTCGAAGTGGAATAGTATTCAACTCTGGATCATATTTAATAATTTCGTTTCCCAT
GCTGCACCTCCTAATCATTACAAATTTAATATAAGTTAATTTTATCAAGTAGTCAACTTTATA
ATACTACTTGATAAACAGACAAAACCTACTCGATAAACAGACAAAACCTACTTGATAAACAG

ACAAAACCTACTTGATAAACAGACAAAACCTACTTGATAAAGTTCTGTAAGTCTTGGGGGAG
TAAGGCTCAAGAGGGGTCTAAAAGAGGTTTAAAAGAGGTTTATAAAAGAGGTATATAAGAG
GCACCACTGTACGAGATCAAAACGGGCCCATATCATGGCTTTCCCCGTCAAGCTCTAAATCG
GGGGCTCCCTTTAGGGTTCCGATTTAGTGCTTTACGGCACCTCGACCCCAAAAAACTTGATTA
GGGTGATGGTTCACGTAGTGGGCCATCGCCCTGATAGACGGTTTTTCGCCCTTTGACGTTGGA
GTCCACGTTCTTTAATAGTGGACTCTTGTTCCAAACTGGAACAACACTCAACCCTATCTCGGT
CTATTCTTTTGATTTATAAGGGATTTTGCCGATTTCGGCCTATTGGTTAAAAAATGAGCTGAT
TTAACAAAAATTTAACGCGAATTTTAACAAAATATTAACGCTTACAATTCGCGCGTAATACG
ACTCACTATAGGGCGAATTGGGTACCGAATTCCTGCAGCCCGGGTGTTAAGACAAAAGTAAT
TTTGTCTTAACACCCTTTTTTTGTACCCCAAAAGTATTAATGATTGACAGAAGTGCTTCTTGAT
TGTAAACTGAATATATAATTACAAATGTAAACGAAAGAGTGTGAGCAAGATGGCAGAAAAT
ACAAATGAAATTACGACTTCCGAATGGGAAGTGATGCGCATTGTCTGGTCACTTGGTCAAGT
CAACAGTCGTGACTTGATTGATTTACTTCAACCAAAGCGCGATTGGCAAGATTCAACCATCA
AAACATTGATTGGCCGCTTGGTTAAAAAAGGCTTCTTAAAAACAGAAAAGGAAGGGCGGCG
GTTTAATTACACAGCAACCGTTCCTGAGATTGAAGCGATGGAAAATGCCACACAAAGTCTTT
TTGAGCACTTGTGTGGCATGAAAAAAGGGCAGACATTAGCGGCTTTAATTGATCAAACGACT
TTGAGTCAGACGGATATTTTGCAGTTACAACAATTACTGACAGCCAAAGCAGCTACTGCACC
AGAAAAAGTGGCTTGCGACTGTTTGCCAACAACATACGAGCCGGAAGCATAAAGTGTAAAG
ATAAATGTGACTGTGAAAAGGAGGAATAACGATGAAACACGATGGATCCGGACATATTTTC
AGAAAATCGGTCAGTCATTGATGCTACCGATTGCAACGTTGCCCGCAGCAGCGATTTTGGTT
GGGATTGGGAATTACCTGCCAAAACAGTGGCTGTTTGCAAATTACTTGATTCAGGGTGGTAA
CGTTGTCCTGAATAACTTAGCCTTGCTGTTTGCAGTCGGCTTAGCTATCGGGATGTCAGTCAA
TAAGGACGGCGCGGCAGCTATTGCTGGTTTGATCGCGTTTGAAGTTCCGGTGATGGTTTTAA
AACCAGCGACATTAGCGACGATGCTGAATGTTAAAGTGAGTCAGATCAATCCAGCATTTAGC
GCGTTAGATAACAATGTGCTGATTGGAATCAGTGCTGGACTGATTGCGGCTGCGCTCTATAA
TCGGTTCCACGAAGTAAAATTACCAATGGCGCTCTCATTCTTTAGTGGTAAACGCCTAGTGCC
AATTATGGCTGCTTTTGTGATGCTGATTGTGACGGCGGTATTGTACCTTGTTTGGCCGTTTGT
CTACGATGCCATCGTCCTCTTTGCAACGGGAATTTCAAAATTAGGTTTCGTGGGGGCTGGTCT
ATATGGCTTTTTCAATCGTTTATTGATTCCGACTGGTTTGCACCACGCCTTAAATTCCGTATTC
TGGTATAACGTTGCGGGGATCAATGATATTGGCAATTTCTGGGCCAGCCATGGTGTTAAAGG
AATCACAGGGATGTATGAAGCTGGATTCTTCCCAATTATGATGTTTGGCTTGCCAGCTGGTGC
GTATGCAATTTATCGTAATGCACGACCGGAACGAAAAAAAGAAGTGGGTTCATTGATGTTAG
CTGGTGCGTTTGCGTCGTTCTTTACGGGGGTGACCGAACCGCTTGAATTTTCATTCATGTTCG
TGGCATGGCCGCTCTACTTATTACATGCCGTTTTTATGGGACTATCATTAGGATTCGCCGCGT
TAATGCACTGGACCGCCAGCTTTTCGTTTAGTGGCGGGTTAGTCGATTATTTATTGAGTTTTC
GGATGCCGTTAGCCAACCAACCCTATATGTTACTGGTTCAAGGGCTGGTGATGGCCGTCATTT
ATTACTTTGGTTTTGACTTTGCAATCAAACGGTTTAATTTGAAGACGCCTGGTCGTGAAGTTG

TGTCTGCAGATGTCGATGGGGTGGGTGCACCTGCTAGTCCAGCGGTGGCCGTTGCAGCAACG
GATGATAAGTATATGCGCCAAGCAAAGCAAATTTATGCAGCTATTGGTGGTCATGACAATAT
TAGTGTGATCAACAATTGTACGACGCGGCTGCGGTTACAACTTAAGGATACGGAAAAAGTCG
ATCAGCCGGCCGTAATGGCTGCTGGCGTGCCTGGTTTGAACGTACTGGATGTTCATAACATTC
ACATTGTGATCGGCACGGAAGTCCAGTTCGTCGCGGAGGCTTTACAAAAATTATTTTCTGGTC
AGGTGGCGACGACCCCGGCATCTGATGCTGAATCAAAGGCCCCGGTTGAACCGCAAACTGCT
ACTGTAACGGAAGCGCCGGTGACAACGATTTTGCGGGCACCGGCAACGGGACAATTAATGC
CGATTAGTACGGTTGCGGATGAGACGTTTGCTGGTAAACTCTTAGGTGATGGTTACGCTGTTG
AACCCGAAGATGGTGAAGTCGTTGCTCCGGTTAGTGGGACAGTGACAAGTGTCTTTCCGACC
AAACACGCTATCGGTTTGAAGACGACGAGTGGTTTAGAAGTCTTATTACACATGGGGATCAA
TACCGTGGAAATGAATGGTACGCCGTTCAAGTTACACGTGGCAACGGGTGATGAAATTGCGG
CCGGTACTGCGGTAGCGACGGTTGACTTGGCTGCTATCAAGTCAGCCGGTAAAGCGACAACC
ATGATGGTCGTTATCACCAATATGGATCACGTTAACAAACTAACCCTTAATCCAACTGGTCA
TGTGACTAGCGGTGATTTGATTGGCGCAGCTGAAGCGGCCGCACTCGAGGACTACAAAGACG
ACGACGACAAATGAGGATCCACTAGTTCTAGAGCGGCCGCCACCGCGGTGGAGCTCCAGCTT
TTGTTCCCTTTAGTGAGGGTTAATTGCGCGCTTGGCGTAATCATGGTCATAGCTGTTTCCTGT
GTGAAATTGTTATCCGCTCACAATTCCACACCTGGGGTGCCTAATGAGTGAGCTAACTCACA
TTAATTGCGTTGCGCTCACTGCCCGCTTTCCAGTCGGGAAACCTGTCGTGCCAGCTGCATTAA
TGAATCGGCCAACGCGCGGGGAGAGGCGGTTTGCGTATTGGGCGCTCTTCCGCTTCCTCGCT
CACTGACTCGCTGCGCTCGGTCGTTCGGCTGCGGCGAGCGGTATCAGCTCACTCAAAGGCGG
TAATACGGTTATCCACAGAATCAGGGGATAACGCAGGAAAGAACATGTGAGCAAAAGGCCA
GCAAAAGGCCAGGAACCGTAAAAAGGCCGCGTTGCTGGCGTTTTTCCATAGGCTCCGCCCCC
CTGACGAGCATCACAAAAATCGACGCTCAAGTCAGAGGTGGCGAAACCCGACAGGACTATA
AAGATACCAGGCGTTTCCCCCTGGAAGCTCCCTCGTGCGCTCTCCTGTTCCGACCCTGCCGCT
TACCGGATACCTGTCCGCCTTTCTCCCTTCGGGAAGCGTGGCGCTTTCTCATAGCTCACGCTG
TAGGTATCTCAGTTCGGTGTAGGTCGTTCGCTCCAAGCTGGGCTGTGTGCACGAACCCCCCGT
TCAGCCCGACCGCTGCGCCTTATCCGGTAACTATCGTCTTGAGTCCAACCCGGTAAGACACG
ACTTATCGCCACTGGCAGCAGCCACTGGTAACAGGATTAGCAGAGCGAGGTATGTAGGCGGT
GCTACAGAGTTCTTGAAGTGGTGGCCTAACTACGGCTACACTAGAAGGACAGTATTTGGTAT
CTGCGCTCTGCTGAAGCCAGTTACCTTCGGAAAAAGAGTTGGTAGCTCTTGATCCGGCAAAC
AAACCACCGCTGGTAGCGGTGGTTTTTTTGTTTGCAAGCAGCAGATTACGCGCAGAAAAAAA
GGATCTCAAGAAGATCCTTTGATCTTTTCTACGGGGTCTGACGCTCAGTGGAACGAAAACTC
ACGTTAAGGGATTTTGGTCATGAGATTATCAAAAAGGATCTTCACCTAGATCCTTTTAAATTA
AAAATGAAGTTTTAAATCAATCTAAAGTATATATGAGTAAACTTGGTCTGACAGTTACCAAT
GCTTAATCAGTGAGGCACCTATCTCAGCGATCTGTCTATTTCGTTCATCCATAGTTGCCTGAC
TCCCCGTCGTGTAGATAACTACGATACGGGAGGGCTTACCATCTGGCCCCAGTGCTGCAATG
ATACCGCGAGACCCACGCTCACCGGCTCCAGATTTATCAGCAATAAACCAGCCAGCCGGAAG

GGCCGAGCGCAGAAGTGGTCCTGCAACTTTATCCGCCTCCATCCAGTCTATTAATTGTTGCCG
GGAAGCTAGAGTAAGTAGTTCGCCAGTTAATAGTTTGCGCAACGTTGTTGCCATTGCTACAG
GCATCGTGGTGTCACGCTCGTCGTTTGGTATGGCTTCATTCAGCTCCGGTTCCCAACGATCAA
GGCGAGTTACATGATCCCCCATGTTGTGCAAAAAAGCGGTTAGCTCCTTCGGTCCTCCGATC
GTTGTCAGAAGTAAGTTGGCCGCAGTGTTATCACTCATGGTTATGGCAGCACTGCATAATTCT
CTTACTGTCATGCCATCCGTAAGATGCTTTTCTGTGACTGGTGAGTACTCAACCAAGTCATTC
TGAGAATAGTGTATGCGGCGACCGAGTTGCTCTTGCCCGGCGTCAATACGGGATAATACCGC
GCCACATAGCAGAACTTTAAAAGTGCTCATCATTGGAAAACGTTCTTCGGGGCGAAAACTCT
CAAGGATCTTACCGCTGTTGAGATCCAGTTCGATGTAACCCACTCGTGCACCCAACTGATCTT
CAGCATCTTTTACTTTCACCAGCGTTTCTGGGTGAGCAAAAACAGGAAGGCAAAATGCCGCA
AAAAAGGGAATAAGGGCGACACGGAAATGTTGAATACTCATACTCTTCCTTTTTCAATATTA
TTGAAGCATTTATCAGGGTTA

5.4.3 pRV613:pts18c (AB deletion)

ATAGCGGCGCATTAAGCGCGGGTGTGGTGGTTACGCGCAGCGTGACCGCTACACTTGCCAGC
GCCCTAGCGCCCGCTCCTTTCGCTTTCTTCCCTTCCTTTCTCGCCACGTTCGCCATCCATGGAA
GCTGTCAGTAGTATACCTAATAATTTATCTACATTCCCTTTAGTAACGTGTAACTTTCCAAAT
TTACAAAAGCGACTCATAGAATTATTTCCTCCCGTTAAATAATAGATAACTATTAAAAATAG
ACAATACTTGCTCATAAGTAACGGTACTTAAATTGTTTACTTTGGCGTGTTTCATTGCTTGAT
GAAACTGATTTTTAGTAAACAGTTGACGATATTCTCGATTGACCCATTTTGAAACAAAGTAC
GTATATAGCTTCCAATATTTATCTGGAACATCTGTGGTATGGCGGGTAAGTTTTATTAAGACA
CTGTTTACTTTTGGTTTAGGATGAAAGCATTCCGCTGGCAGCTTAAGCAATTGCTGAATCGAG
ACTTGAGTGTGCAAGAGCAACCCTAGTGTTCGGTGAATATCCAAGGTACGCTTGTAGAATCC
TTCTTCAACAATCAGATAGATGTCAGACGCATGGCTTTCAAAAACCACTTTTTTAATAATTTG
TGTGCTTAAATGGTAAGGAATACTCCCAACAATTTTATACCTCTGTTTGTTAGGGAATTGAAA
CTGTAGAATATCTTGGTGAATTAAAGTGACACGAGTATTCAGTTTTAATTTTTCTGACGATAA
GTTGAATAGATGACTGTCTAATTCAATAGACGTTACCTGTTTACTTATTTTAGCCAGTTTCGT
CGTTAAATGCCCTTTACCTGTTCCAATTTCGTAAACGGTATCGGTTTCTTTTAAATTCAATTGT
TTTATTATTTGGTTGAGTACTTTTTCACTCGTTAAAAAGTTTTGAGAATATTTTATATTTTTGT
TCATGTAATCACTCCTTCTTAATTACAAATTTTTAGCATCTAATTTAACTTCAATTCCTATTAT
ACAAAATTTTAAGATACTGCACTATCAACACACTCTTAAGTTTGCTTCTAAGTCTTATTTCCA
TAACTTCTTTTACGTTTCCGCCATTCTTTGCTGTTTCGATATGATATCAATTCATCAGCAAGTT
CTCTAATAGTTTTAGGCATGTTTTCCAAACCCCTTTCGGAGTTCAGCAAGCATTTCTTGTCTTG
TCTGATCTCTTATTTTTTTATCATGCTCAATCTGTCGTTCGGCTAAAATCTGTTTTTCAGTTGA
GCCTAGTGGTAGGCCTTTAACTTTATCGGTAGCGCGCCATTTTTCTTGTTCTGTTAATTCACCA
TTATTCTGGATATTAAAAAGTTTTTGACGCTCGTCTTGGAACTTACCTTGGGAAAAGTCGTTT

GCGTCTTTCTTTTCAGGTTTCCAAGTAAAGGTATAACCAATCACTGGTTTTCCACGACCTTTA
CCATATTTTTTCCTAACCGTTAACCCTCTAAATAAGGGAGTTAATTCTTCTTTAATGGGTTTTA
TAACAAATTTATCAACGTTAGATGGACTACTCCAGTAACTTTTAGGCATATCGAGTAATTCAA
AAAAATCTTCTTTAGAAAAATAAGCATATCCAGTGGTTCGGAACTGTTTTAGTAATCGAAAC
ATAGTTTTTGCGTAACTACTCTTCAAATCTCTGAACTCTGCTAGAGCATAACGAACCCAACTT
TCAAGCTTGTTTAAAAGGGGTAAAGCACGTTCATAAACTTTTACGTCTACATAAGGTTCCTTT
GCATCTCCATCGATTTTAAACTCTGTAAACAAGACAAAAAACTCTCGAGTTAATCCACTTTTA
CTTCGCCTACCAAAATGTAATCCCATCATTTTTTCATAAGTTCTCTGAATATCATCTTCAAAA
CGGTTATTTGCAGTTGGTTTATAAGCACTTAATTCTTTTAATTGATCAAAAGTAAAGCGAATG
GTCTGATTACTCTTATCACGCATACGAGAAATAATTGAAAAAAATAAATTCATTTCAACCGG
AGTAAATTTTCGAAGTGGAATAGTATTCAACTCTGGATCATATTTAATAATTTCGTTTCCCAT
GCTGCACCTCCTAATCATTACAAATTTAATATAAGTTAATTTTATCAAGTAGTCAACTTTATA
ATACTACTTGATAAACAGACAAAACCTACTCGATAAACAGACAAAACCTACTTGATAAACAG
ACAAAACCTACTTGATAAACAGACAAAACCTACTTGATAAAGTTCTGTAAGTCTTGGGGGAG
TAAGGCTCAAGAGGGGTCTAAAAGAGGTTTAAAAGAGGTTTATAAAAGAGGTATATAAGAG
GCACCACTGTACGAGATCAAAACGGGCCCATATCATGGCTTTCCCCGTCAAGCTCTAAATCG
GGGGCTCCCTTTAGGGTTCCGATTTAGTGCTTTACGGCACCTCGACCCCAAAAAACTTGATTA
GGGTGATGGTTCACGTAGTGGGCCATCGCCCTGATAGACGGTTTTTCGCCCTTTGACGTTGGA
GTCCACGTTCTTTAATAGTGGACTCTTGTTCCAAACTGGAACAACACTCAACCCTATCTCGGT
CTATTCTTTTGATTTATAAGGGATTTTGCCGATTTCGGCCTATTGGTTAAAAAATGAGCTGAT
TTAACAAAAATTTAACGCGAATTTTAACAAAATATTAACGCTTACAATTCGCGCGTAATACG
ACTCACTATAGGGCGAATTGGGTACCGAATTCCTGCAGCCCGGGTGTTAAGACAAAAGTAAT
TTTGTCTTAACACCCTTTTTTTGTACCCCAAAAGTATTAATGATTGACAGAAGTGCTTCTTGAT
TGTAAACTGAATATATAATTACAAATGTAAACGAAAGAGTGTGAGCAAGATGGCAGAAAAT
ACAAATGAAATTACGACTTCCGAATGGGAAGTGATGCGCATTGTCTGGTCACTTGGTCAAGT
CAACAGTCGTGACTTGATTGATTTACTTCAACCAAAGCGCGATTGGCAAGATTCAACCATCA
AAACATTGATTGGCCGCTTGGTTAAAAAAGGCTTCTTAAAAACAGAAAAGGAAGGGCGGCG
GTTTAATTACACAGCAACCGTTCCTGAGATTGAAGCGATGGAAAATGCCACACAAAGTCTTT
TTGAGCACTTGTGTGGCATGAAAAAAGGGCAGACATTAGCGGCTTTAATTGATCAAACGACT
TTGAGTCAGACGGATATTTTGCAGTTACAACAATTACTGACAGCCAAAGCAGCTACTGCACC
AGAAAAAGTGGCTTGCGACTGTTTGCCAACAACATACGAGCCGGAAGCATAAAGTGTAAAG
ATAAATGTGACTGTGAAAAGGAGGAATAACGATGAAACACGATGGATCCGGACATATTTTC
AGAAAATCGGTCAGTCATTGATGCTACCGATTGCAACGTTGCCCGCAGCAGCGATTTTGGTT
GGGATTGGGAATTACCTGCCAAAACAGTGGCTGTTTGCAAATTACTTGATTCAGGGTGGTAA
CGTTGTCCTGAATAACTTAGCCTTGCTGTTTGCAGTCGGCTTAGCTATCGGGATGTCAGTCAA
TAAGGACGGCGCGGCAGCTATTGCTGGTTTGATCGCGTTTGAAGTTCCGGTGATGGTTTTAA
AACCAGCGACATTAGCGACGATGCTGAATGTTAAAGTGAGTCAGATCAATCCAGCATTTAGC

GCGTTAGATAACAATGTGCTGATTGGAATCAGTGCTGGACTGATTGCGGCTGCGCTCTATAA
TCGGTTCCACGAAGTAAAATTACCAATGGCGCTCTCATTCTTTAGTGGTAAACGCCTAGTGCC
AATTATGGCTGCTTTTGTGATGCTGATTGTGACGGCGGTATTGTACCTTGTTTGGCCGTTTGTC
TACGATGCCATCGTCCTCTTTGCAACGGGAATTTCAAAATTAGGTTTCGTGGGGGCTGGTCTA
TATGGCTTTTTCAATCGTTTATTGATTCCGACTGGTTTGCACCACGCCTTAAATTCCGTATTCT
GGTATAACGTTGCGGGGATCAATGATATTGGCAATTTCTGGGCCAGCCATGGTGTTAAAGGA
ATCACAGGGATGTATGAAGCTGGATTCTTCCCAATTATGATGTTTGGCTTGCCAGCTGGTGCG
TATGCAATTTATCGTAATGCACGACCGGAACGAAAAAAAGAAGTGGGTTCATTGATGTTAGC
TGGTGCGTTTGCGTCGTTCTTTACGGGGGTGACCGAACCGCTTGAATTTTCATTCATGTTCGT
GGCATGGCCGCTCTACTTATTACATGCCGTTTTTATGGGACTATCATTAGGATTCGCCGCGTT
AATGCACTGGACCGCCAGCTTTTCGTTTAGTGGCGGGTTAGTCGATTATTTATTGAGTTTTCG
GATGCCGTTAGCCAACCAACCCTATATGTTACTGGTTCAAGGGCTGGTGATGGCCGTCATTTA
TTACTTTGGTTTTGACTTTGCAATCAAACGGTTTAATTTGAAGACGCCTGGTCGTGAAGTTGT
GTCTGCAGATGTCGATGGGGTGGGTGCACCTGCTAGTCCAGCGGTGGCCGTTGCAGCAACGG
ATGATAAGTATATGCGCCAAGCAAAGCATAGGAGCATGCCATATGGACTACAAAGACGACG
ACGACAAATGAGGATCCACTAGTTCTAGAGCGGCCGCCACCGCGGTGGAGCTCCAGCTTTTG
TTCCCTTTAGTGAGGGTTAATTGCGCGCTTGGCGTAATCATGGTCATAGCTGTTTCCTGTGTG
AAATTGTTATCCGCTCACAATTCCACACCTGGGGTGCCTAATGAGTGAGCTAACTCACATTA
ATTGCGTTGCGCTCACTGCCCGCTTTCCAGTCGGGAAACCTGTCGTGCCAGCTGCATTAATGA
ATCGGCCAACGCGCGGGGAGAGGCGGTTTGCGTATTGGGCGCTCTTCCGCTTCCTCGCTCAC
TGACTCGCTGCGCTCGGTCGTTCGGCTGCGGCGAGCGGTATCAGCTCACTCAAAGGCGGTAA
TACGGTTATCCACAGAATCAGGGGATAACGCAGGAAAGAACATGTGAGCAAAAGGCCAGCA
AAAGGCCAGGAACCGTAAAAAGGCCGCGTTGCTGGCGTTTTTCCATAGGCTCCGCCCCCCTG
ACGAGCATCACAAAAATCGACGCTCAAGTCAGAGGTGGCGAAACCCGACAGGACTATAAAG
ATACCAGGCGTTTCCCCCTGGAAGCTCCCTCGTGCGCTCTCCTGTTCCGACCCTGCCGCTTAC
CGGATACCTGTCCGCCTTTCTCCCTTCGGGAAGCGTGGCGCTTTCTCATAGCTCACGCTGTAG
GTATCTCAGTTCGGTGTAGGTCGTTCGCTCCAAGCTGGGCTGTGTGCACGAACCCCCCGTTCA
GCCCGACCGCTGCGCCTTATCCGGTAACTATCGTCTTGAGTCCAACCCGGTAAGACACGACT
TATCGCCACTGGCAGCAGCCACTGGTAACAGGATTAGCAGAGCGAGGTATGTAGGCGGTGCT
ACAGAGTTCTTGAAGTGGTGGCCTAACTACGGCTACACTAGAAGGACAGTATTTGGTATCTG
CGCTCTGCTGAAGCCAGTTACCTTCGGAAAAAGAGTTGGTAGCTCTTGATCCGGCAAACAAA
CCACCGCTGGTAGCGGTGGTTTTTTTGTTTGCAAGCAGCAGATTACGCGCAGAAAAAAAGGA
TCTCAAGAAGATCCTTTGATCTTTTCTACGGGGTCTGACGCTCAGTGGAACGAAAACTCACG
TTAAGGGATTTTGGTCATGAGATTATCAAAAAGGATCTTCACCTAGATCCTTTTAAATTAAAA
ATGAAGTTTTAAATCAATCTAAAGTATATATGAGTAAACTTGGTCTGACAGTTACCAATGCTT
AATCAGTGAGGCACCTATCTCAGCGATCTGTCTATTTCGTTCATCCATAGTTGCCTGACTCCC
CGTCGTGTAGATAACTACGATACGGGAGGGCTTACCATCTGGCCCCAGTGCTGCAATGATAC

CGCGAGACCCACGCTCACCGGCTCCAGATTTATCAGCAATAAACCAGCCAGCCGGAAGGGC
CGAGCGCAGAAGTGGTCCTGCAACTTTATCCGCCTCCATCCAGTCTATTAATTGTTGCCGGGA
AGCTAGAGTAAGTAGTTCGCCAGTTAATAGTTTGCGCAACGTTGTTGCCATTGCTACAGGCA
TCGTGGTGTCACGCTCGTCGTTTGGTATGGCTTCATTCAGCTCCGGTTCCCAACGATCAAGGC
GAGTTACATGATCCCCCATGTTGTGCAAAAAAGCGGTTAGCTCCTTCGGTCCTCCGATCGTTG
TCAGAAGTAAGTTGGCCGCAGTGTTATCACTCATGGTTATGGCAGCACTGCATAATTCTCTTA
CTGTCATGCCATCCGTAAGATGCTTTTCTGTGACTGGTGAGTACTCAACCAAGTCATTCTGAG
AATAGTGTATGCGGCGACCGAGTTGCTCTTGCCCGGCGTCAATACGGGATAATACCGCGCCA
CATAGCAGAACTTTAAAAGTGCTCATCATTGGAAAACGTTCTTCGGGGCGAAAACTCTCAAG
GATCTTACCGCTGTTGAGATCCAGTTCGATGTAACCCACTCGTGCACCCAACTGATCTTCAGC
ATCTTTTACTTTCACCAGCGTTTCTGGGTGAGCAAAAACAGGAAGGCAAAATGCCGCAAAAA
AGGGAATAAGGGCGACACGGAAATGTTGAATACTCATACTCTTCCTTTTTCAATATTATTGA
AGCATTTATCAGGGTTA

5.4.4 Full pRV613:pts18cb-FLAG

ATAGCGGCGCATTAAGCGCGGGTGTGGTGGTTACGCGCAGCGTGACCGCTACACTTGCCAGC
GCCCTAGCGCCCGCTCCTTTCGCTTTCTTCCCTTCCTTTCTCGCCACGTTCGCCATCCATGGAA
GCTGTCAGTAGTATACCTAATAATTTATCTACATTCCCTTTAGTAACGTGTAACTTTCCAAAT
TTACAAAAGCGACTCATAGAATTATTTCCTCCCGTTAAATAATAGATAACTATTAAAAATAG
ACAATACTTGCTCATAAGTAACGGTACTTAAATTGTTTACTTTGGCGTGTTTCATTGCTTGAT
GAAACTGATTTTTAGTAAACAGTTGACGATATTCTCGATTGACCCATTTTGAAACAAAGTAC
GTATATAGCTTCCAATATTTATCTGGAACATCTGTGGTATGGCGGGTAAGTTTTATTAAGACA
CTGTTTACTTTTGGTTTAGGATGAAAGCATTCCGCTGGCAGCTTAAGCAATTGCTGAATCGAG
ACTTGAGTGTGCAAGAGCAACCCTAGTGTTCGGTGAATATCCAAGGTACGCTTGTAGAATCC
TTCTTCAACAATCAGATAGATGTCAGACGCATGGCTTTCAAAAACCACTTTTTTAATAATTTG
TGTGCTTAAATGGTAAGGAATACTCCCAACAATTTTATACCTCTGTTTGTTAGGGAATTGAAA
CTGTAGAATATCTTGGTGAATTAAAGTGACACGAGTATTCAGTTTTAATTTTTCTGACGATAA
GTTGAATAGATGACTGTCTAATTCAATAGACGTTACCTGTTTACTTATTTTAGCCAGTTTCGT
CGTTAAATGCCCTTTACCTGTTCCAATTTCGTAAACGGTATCGGTTTCTTTTAAATTCAATTGT
TTTATTATTTGGTTGAGTACTTTTTCACTCGTTAAAAAGTTTTGAGAATATTTTATATTTTTGT
TCATGTAATCACTCCTTCTTAATTACAAATTTTTAGCATCTAATTTAACTTCAATTCCTATTAT
ACAAAATTTTAAGATACTGCACTATCAACACACTCTTAAGTTTGCTTCTAAGTCTTATTTCCA
TAACTTCTTTTACGTTTCCGCCATTCTTTGCTGTTTCGATATGATATCAATTCATCAGCAAGTT
CTCTAATAGTTTTAGGCATGTTTTCCAAACCCCTTTCGGAGTTCAGCAAGCATTTCTTGTCTTG
TCTGATCTCTTATTTTTTTATCATGCTCAATCTGTCGTTCGGCTAAAATCTGTTTTTCAGTTGA
GCCTAGTGGTAGGCCTTTAACTTTATCGGTAGCGCGCCATTTTTCTTGTTCTGTTAATTCACCA

TTATTCTGGATATTAAAAAGTTTTTGACGCTCGTCTTGGAACTTACCTTGGGAAAAGTCGTTT
GCGTCTTTCTTTTCAGGTTTCCAAGTAAAGGTATAACCAATCACTGGTTTTCCACGACCTTTA
CCATATTTTTTCCTAACCGTTAACCCTCTAAATAAGGGAGTTAATTCTTCTTTAATGGGTTTTA
TAACAAATTTATCAACGTTAGATGGACTACTCCAGTAACTTTTAGGCATATCGAGTAATTCAA
AAAAATCTTCTTTAGAAAAATAAGCATATCCAGTGGTTCGGAACTGTTTTAGTAATCGAAAC
ATAGTTTTTGCGTAACTACTCTTCAAATCTCTGAACTCTGCTAGAGCATAACGAACCCAACTT
TCAAGCTTGTTTAAAAGGGGTAAAGCACGTTCATAAACTTTTACGTCTACATAAGGTTCCTTT
GCATCTCCATCGATTTTAAACTCTGTAAACAAGACAAAAAACTCTCGAGTTAATCCACTTTTA
CTTCGCCTACCAAAATGTAATCCCATCATTTTTTCATAAGTTCTCTGAATATCATCTTCAAAA
CGGTTATTTGCAGTTGGTTTATAAGCACTTAATTCTTTTAATTGATCAAAAGTAAAGCGAATG
GTCTGATTACTCTTATCACGCATACGAGAAATAATTGAAAAAAATAAATTCATTTCAACCGG
AGTAAATTTTCGAAGTGGAATAGTATTCAACTCTGGATCATATTTAATAATTTCGTTTCCCAT
GCTGCACCTCCTAATCATTACAAATTTAATATAAGTTAATTTTATCAAGTAGTCAACTTTATA
ATACTACTTGATAAACAGACAAAACCTACTCGATAAACAGACAAAACCTACTTGATAAACAG
ACAAAACCTACTTGATAAACAGACAAAACCTACTTGATAAAGTTCTGTAAGTCTTGGGGGAG
TAAGGCTCAAGAGGGGTCTAAAAGAGGTTTAAAAGAGGTTTATAAAAGAGGTATATAAGAG
GCACCACTGTACGAGATCAAAACGGGCCCATATCATGGCTTTCCCCGTCAAGCTCTAAATCG
GGGGCTCCCTTTAGGGTTCCGATTTAGTGCTTTACGGCACCTCGACCCCAAAAAACTTGATTA
GGGTGATGGTTCACGTAGTGGGCCATCGCCCTGATAGACGGTTTTTCGCCCTTTGACGTTGGA
GTCCACGTTCTTTAATAGTGGACTCTTGTTCCAAACTGGAACAACACTCAACCCTATCTCGGT
CTATTCTTTTGATTTATAAGGGATTTTGCCGATTTCGGCCTATTGGTTAAAAAATGAGCTGAT
TTAACAAAAATTTAACGCGAATTTTAACAAAATATTAACGCTTACAATTCGCGCGTAATACG
ACTCACTATAGGGCGAATTGGGTACCGAATTCCTGCAGCCCGGGTGTTAAGACAAAAGTAAT
TTTGTCTTAACACCCTTTTTTTGTACCCCAAAAGTATTAATGATTGACAGAAGTGCTTCTTGAT
TGTAAACTGAATATATAATTACAAATGTAAACGAAAGAGTGTGAGCAAGATGGCAGAAAAT
ACAAATGAAATTACGACTTCCGAATGGGAAGTGATGCGCATTGTCTGGTCACTTGGTCAAGT
CAACAGTCGTGACTTGATTGATTTACTTCAACCAAAGCGCGATTGGCAAGATTCAACCATCA
AAACATTGATTGGCCGCTTGGTTAAAAAAGGCTTCTTAAAAACAGAAAAGGAAGGGCGGCG
GTTTAATTACACAGCAACCGTTCCTGAGATTGAAGCGATGGAAAATGCCACACAAAGTCTTT
TTGAGCACTTGTGTGGCATGAAAAAAGGGCAGACATTAGCGGCTTTAATTGATCAAACGACT
TTGAGTCAGACGGATATTTTGCAGTTACAACAATTACTGACAGCCAAAGCAGCTACTGCACC
AGAAAAAGTGGCTTGCGACTGTTTGCCAACAACATACGAGCCGGAAGCATAAAGTGTAAAG
ATAAATGTGACTGTGAAAAGGAGGAATAACGATGAAACACGATGGATCCGGACATATTTTC
AGAAAATCGGTCAGTCATTGATGCTACCGATTGCAACGTTGCCCGCAGCAGCGATTTTGGTT
GGGATTGGGAATTACCTGCCAAAACAGTGGCTGTTTGCAAATTACTTGATTCAGGGTGGTAA
CGTTGTCCTGAATAACTTAGCCTTGCTGTTTGCAGTCGGCTTAGCTATCGGGATGTCAGTCAA
TAAGGACGGCGCGGCAGCTATTGCTGGTTTGATCGCGTTTGAAGTTCCGGTGATGGTTTTAA

AACCAGCGACATTAGCGACGATGCTGAATGTTAAAGTGAGTCAGATCAATCCAGCATTTAGC
GCGTTAGATAACAATGTGCTGATTGGAATCAGTGCTGGACTGATTGCGGCTGCGCTCTATAA
TCGGTTCCACGAAGTAAAATTACCAATGGCGCTCTCATTCTTTAGTGGTAAACGCCTAGTGCC
AATTATGGCTGCTTTTGTGATGCTGATTGTGACGGCGGTATTGTACCTTGTTTGGCCGTTTGT
CTACGATGCCATCGTCCTCTTTGCAACGGGAATTTCAAAATTAGGTTTCGTGGGGGCTGGTCT
ATATGGCTTTTTCAATCGTTTATTGATTCCGACTGGTTTGCACCACGCCTTAAATTCCGTATTC
TGGTATAACGTTGCGGGGATCAATGATATTGGCAATTTCTGGGCCAGCCATGGTGTTAAAGG
AATCACAGGGATGTATGAAGCTGGATTCTTCCCAATTATGATGTTTGGCTTGCCAGCTGGTGC
GTATGCAATTTATCGTAATGCACGACCGGAACGAAAAAAAGAAGTGGGTTCATTGATGTTAG
CTGGTGCGTTTGCGTCGTTCTTTACGGGGGTGACCGAACCGCTTGAATTTTCATTCATGTTCG
TGGCATGGCCGCTCTACTTATTACATGCCGTTTTTATGGGACTATCATTAGGATTCGCCGCGT
TAATGCACTGGACCGCCAGCTTTTCGTTTAGTGGCGGGTTAGTCGATTATTTATTGAGTTTTC
GGATGCCGTTAGCCAACCAACCCTATATGTTACTGGTTCAAGGGCTGGTGATGGCCGTCATTT
ATTACTTTGGTTTTGACTTTGCAATCAAACGGTTTAATTTGAAGACGCCTGGTCGTGAAGTTG
TGTCTGCAGATGTCGATGGGGTGGGTGCACCTGCTAGTCCAGCGGTGGCCGTTGCAGCAACG
GATGATAAGTATATGCGCCAAGCAAAGCAAATTTATGCAGCTATTGGTGGTCATGACAATAT
TAGTGTGATCAACAATTGTACGACGCGGCTGCGGTTACAACTTAAGGATACGGAAAAAGTCG
ATCAGCCGGCCGTAATGGCTGCTGGCGTGCCTGGTTTGAACGTACTGGATGTTCATAACATTC
ACATTGTGATCGGCACGGAAGTCCAGTTCGTCGCGGAGGCTTTACAAAAATTATTTTCTGGTC
AGGTGGCGACGACCCCGGCATCTGATGCTGAATCAAAGGCCCATAGGCATGCGGATATGGA
CTACAAAGACGACGACGACAAATGAGGATCCACTAGTTCTAGAGCGGCCGCCACCGCGGTG
GAGCTCCAGCTTTTGTTCCCTTTAGTGAGGGTTAATTGCGCGCTTGGCGTAATCATGGTCATA
GCTGTTTCCTGTGTGAAATTGTTATCCGCTCACAATTCCACACCTGGGGTGCCTAATGAGTGA
GCTAACTCACATTAATTGCGTTGCGCTCACTGCCCGCTTTCCAGTCGGGAAACCTGTCGTGCC
AGCTGCATTAATGAATCGGCCAACGCGCGGGGAGAGGCGGTTTGCGTATTGGGCGCTCTTCC
GCTTCCTCGCTCACTGACTCGCTGCGCTCGGTCGTTCGGCTGCGGCGAGCGGTATCAGCTCAC
TCAAAGGCGGTAATACGGTTATCCACAGAATCAGGGGATAACGCAGGAAAGAACATGTGAG
CAAAAGGCCAGCAAAAGGCCAGGAACCGTAAAAAGGCCGCGTTGCTGGCGTTTTTCCATAG
GCTCCGCCCCCCTGACGAGCATCACAAAAATCGACGCTCAAGTCAGAGGTGGCGAAACCCG
ACAGGACTATAAAGATACCAGGCGTTTCCCCCTGGAAGCTCCCTCGTGCGCTCTCCTGTTCCG
ACCCTGCCGCTTACCGGATACCTGTCCGCCTTTCTCCCTTCGGGAAGCGTGGCGCTTTCTCAT
AGCTCACGCTGTAGGTATCTCAGTTCGGTGTAGGTCGTTCGCTCCAAGCTGGGCTGTGTGCAC
GAACCCCCCGTTCAGCCCGACCGCTGCGCCTTATCCGGTAACTATCGTCTTGAGTCCAACCCG
GTAAGACACGACTTATCGCCACTGGCAGCAGCCACTGGTAACAGGATTAGCAGAGCGAGGT
ATGTAGGCGGTGCTACAGAGTTCTTGAAGTGGTGGCCTAACTACGGCTACACTAGAAGGACA
GTATTTGGTATCTGCGCTCTGCTGAAGCCAGTTACCTTCGGAAAAAGAGTTGGTAGCTCTTGA
TCCGGCAAACAAACCACCGCTGGTAGCGGTGGTTTTTTTGTTTGCAAGCAGCAGATTACGCG

CAGAAAAAAAGGATCTCAAGAAGATCCTTTGATCTTTTCTACGGGGTCTGACGCTCAGTGGA
ACGAAAACTCACGTTAAGGGATTTTGGTCATGAGATTATCAAAAAGGATCTTCACCTAGATC
CTTTTAAATTAAAAATGAAGTTTTAAATCAATCTAAAGTATATATGAGTAAACTTGGTCTGAC
AGTTACCAATGCTTAATCAGTGAGGCACCTATCTCAGCGATCTGTCTATTTCGTTCATCCATA
GTTGCCTGACTCCCCGTCGTGTAGATAACTACGATACGGGAGGGCTTACCATCTGGCCCCAG
TGCTGCAATGATACCGCGAGACCCACGCTCACCGGCTCCAGATTTATCAGCAATAAACCAGC
CAGCCGGAAGGGCCGAGCGCAGAAGTGGTCCTGCAACTTTATCCGCCTCCATCCAGTCTATT
AATTGTTGCCGGGAAGCTAGAGTAAGTAGTTCGCCAGTTAATAGTTTGCGCAACGTTGTTGC
CATTGCTACAGGCATCGTGGTGTCACGCTCGTCGTTTGGTATGGCTTCATTCAGCTCCGGTTC
CCAACGATCAAGGCGAGTTACATGATCCCCCATGTTGTGCAAAAAAGCGGTTAGCTCCTTCG
GTCCTCCGATCGTTGTCAGAAGTAAGTTGGCCGCAGTGTTATCACTCATGGTTATGGCAGCAC
TGCATAATTCTCTTACTGTCATGCCATCCGTAAGATGCTTTTCTGTGACTGGTGAGTACTCAA
CCAAGTCATTCTGAGAATAGTGTATGCGGCGACCGAGTTGCTCTTGCCCGGCGTCAATACGG
GATAATACCGCGCCACATAGCAGAACTTTAAAAGTGCTCATCATTGGAAAACGTTCTTCGGG
GCGAAAACTCTCAAGGATCTTACCGCTGTTGAGATCCAGTTCGATGTAACCCACTCGTGCAC
CCAACTGATCTTCAGCATCTTTTACTTTCACCAGCGTTTCTGGGTGAGCAAAAACAGGAAGG
CAAAATGCCGCAAAAAAGGGAATAAGGGCGACACGGAAATGTTGAATACTCATACTCTTCC
TTTTTCAATATTATTGAAGCATTTATCAGGGTTA

5.4.5 pRV613-pts18cba H584A

ATAGCGGCGCATTAAGCGCGGGTGTGGTGGTTACGCGCAGCGTGACCGCTACACTTGCCAGC
GCCCTAGCGCCCGCTCCTTTCGCTTTCTTCCCTTCCTTTCTCGCCACGTTCGCCATCCATGGAA
GCTGTCAGTAGTATACCTAATAATTTATCTACATTCCCTTTAGTAACGTGTAACTTTCCAAAT
TTACAAAAGCGACTCATAGAATTATTTCCTCCCGTTAAATAATAGATAACTATTAAAAATAG
ACAATACTTGCTCATAAGTAACGGTACTTAAATTGTTTACTTTGGCGTGTTTCATTGCTTGAT
GAAACTGATTTTTAGTAAACAGTTGACGATATTCTCGATTGACCCATTTTGAAACAAAGTAC
GTATATAGCTTCCAATATTTATCTGGAACATCTGTGGTATGGCGGGTAAGTTTTATTAAGACA
CTGTTTACTTTTGGTTTAGGATGAAAGCATTCCGCTGGCAGCTTAAGCAATTGCTGAATCGAG
ACTTGAGTGTGCAAGAGCAACCCTAGTGTTCGGTGAATATCCAAGGTACGCTTGTAGAATCC
TTCTTCAACAATCAGATAGATGTCAGACGCATGGCTTTCAAAAACCACTTTTTTAATAATTTG
TGTGCTTAAATGGTAAGGAATACTCCCAACAATTTTATACCTCTGTTTGTTAGGGAATTGAAA
CTGTAGAATATCTTGGTGAATTAAAGTGACACGAGTATTCAGTTTTAATTTTTCTGACGATAA
GTTGAATAGATGACTGTCTAATTCAATAGACGTTACCTGTTTACTTATTTTAGCCAGTTTCGT
CGTTAAATGCCCTTTACCTGTTCCAATTTCGTAAACGGTATCGGTTTCTTTTAAATTCAATTGT
TTTATTATTTGGTTGAGTACTTTTTCACTCGTTAAAAAGTTTTGAGAATATTTTATATTTTTGT
TCATGTAATCACTCCTTCTTAATTACAAATTTTTAGCATCTAATTTAACTTCAATTCCTATTAT

ACAAAATTTTAAGATACTGCACTATCAACACACTCTTAAGTTTGCTTCTAAGTCTTATTTCCA
TAACTTCTTTTACGTTTCCGCCATTCTTTGCTGTTTCGATATGATATCAATTCATCAGCAAGTT
CTCTAATAGTTTTAGGCATGTTTTCCAAACCCCTTTCGGAGTTCAGCAAGCATTTCTTGTCTTG
TCTGATCTCTTATTTTTTTATCATGCTCAATCTGTCGTTCGGCTAAAATCTGTTTTTCAGTTGA
GCCTAGTGGTAGGCCTTTAACTTTATCGGTAGCGCGCCATTTTTCTTGTTCTGTTAATTCACCA
TTATTCTGGATATTAAAAAGTTTTTGACGCTCGTCTTGGAACTTACCTTGGGAAAAGTCGTTT
GCGTCTTTCTTTTCAGGTTTCCAAGTAAAGGTATAACCAATCACTGGTTTTCCACGACCTTTA
CCATATTTTTTCCTAACCGTTAACCCTCTAAATAAGGGAGTTAATTCTTCTTTAATGGGTTTTA
TAACAAATTTATCAACGTTAGATGGACTACTCCAGTAACTTTTAGGCATATCGAGTAATTCAA
AAAAATCTTCTTTAGAAAAATAAGCATATCCAGTGGTTCGGAACTGTTTTAGTAATCGAAAC
ATAGTTTTTGCGTAACTACTCTTCAAATCTCTGAACTCTGCTAGAGCATAACGAACCCAACTT
TCAAGCTTGTTTAAAAGGGGTAAAGCACGTTCATAAACTTTTACGTCTACATAAGGTTCCTTT
GCATCTCCATCGATTTTAAACTCTGTAAACAAGACAAAAAACTCTCGAGTTAATCCACTTTTA
CTTCGCCTACCAAAATGTAATCCCATCATTTTTTCATAAGTTCTCTGAATATCATCTTCAAAA
CGGTTATTTGCAGTTGGTTTATAAGCACTTAATTCTTTTAATTGATCAAAAGTAAAGCGAATG
GTCTGATTACTCTTATCACGCATACGAGAAATAATTGAAAAAAATAAATTCATTTCAACCGG
AGTAAATTTTCGAAGTGGAATAGTATTCAACTCTGGATCATATTTAATAATTTCGTTTCCCAT
GCTGCACCTCCTAATCATTACAAATTTAATATAAGTTAATTTTATCAAGTAGTCAACTTTATA
ATACTACTTGATAAACAGACAAAACCTACTCGATAAACAGACAAAACCTACTTGATAAACAG
ACAAAACCTACTTGATAAACAGACAAAACCTACTTGATAAAGTTCTGTAAGTCTTGGGGGAG
TAAGGCTCAAGAGGGGTCTAAAAGAGGTTTAAAAGAGGTTTATAAAAGAGGTATATAAGAG
GCACCACTGTACGAGATCAAAACGGGCCCATATCATGGCTTTCCCCGTCAAGCTCTAAATCG
GGGGCTCCCTTTAGGGTTCCGATTTAGTGCTTTACGGCACCTCGACCCCAAAAAACTTGATTA
GGGTGATGGTTCACGTAGTGGGCCATCGCCCTGATAGACGGTTTTTCGCCCTTTGACGTTGGA
GTCCACGTTCTTTAATAGTGGACTCTTGTTCCAAACTGGAACAACACTCAACCCTATCTCGGT
CTATTCTTTTGATTTATAAGGGATTTTGCCGATTTCGGCCTATTGGTTAAAAAATGAGCTGAT
TTAACAAAAATTTAACGCGAATTTTAACAAAATATTAACGCTTACAATTCGCGCGTAATACG
ACTCACTATAGGGCGAATTGGGTACCGAATTCCTGCAGCCCGGGTGTTAAGACAAAAGTAAT
TTTGTCTTAACACCCTTTTTTTGTACCCCAAAAGTATTAATGATTGACAGAAGTGCTTCTTGAT
TGTAAACTGAATATATAATTACAAATGTAAACGAAAGAGTGTGAGCAAGATGGCAGAAAAT
ACAAATGAAATTACGACTTCCGAATGGGAAGTGATGCGCATTGTCTGGTCACTTGGTCAAGT
CAACAGTCGTGACTTGATTGATTTACTTCAACCAAAGCGCGATTGGCAAGATTCAACCATCA
AAACATTGATTGGCCGCTTGGTTAAAAAAGGCTTCTTAAAAACAGAAAAGGAAGGGCGGCG
GTTTAATTACACAGCAACCGTTCCTGAGATTGAAGCGATGGAAAATGCCACACAAAGTCTTT
TTGAGCACTTGTGTGGCATGAAAAAAGGGCAGACATTAGCGGCTTTAATTGATCAAACGACT
TTGAGTCAGACGGATATTTTGCAGTTACAACAATTACTGACAGCCAAAGCAGCTACTGCACC
AGAAAAAGTGGCTTGCGACTGTTTGCCAACAACATACGAGCCGGAAGCATAAAGTGTAAAG

ATAAATGTGACTGTGAAAAGGAGGAATAACGATGAAACACGATGGATCCGGACATATTTTC
AGAAAATCGGTCAGTCATTGATGCTACCGATTGCAACGTTGCCCGCAGCAGCGATTTTGGTT
GGGATTGGGAATTACCTGCCAAAACAGTGGCTGTTTGCAAATTACTTGATTCAGGGTGGTAA
CGTTGTCCTGAATAACTTAGCCTTGCTGTTTGCAGTCGGCTTAGCTATCGGGATGTCAGTCAA
TAAGGACGGCGCGGCAGCTATTGCTGGTTTGATCGCGTTTGAAGTTCCGGTGATGGTTTTAA
AACCAGCGACATTAGCGACGATGCTGAATGTTAAAGTGAGTCAGATCAATCCAGCATTTAGC
GCGTTAGATAACAATGTGCTGATTGGAATCAGTGCTGGACTGATTGCGGCTGCGCTCTATAA
TCGGTTCCACGAAGTAAAATTACCAATGGCGCTCTCATTCTTTAGTGGTAAACGCCTAGTGCC
AATTATGGCTGCTTTTGTGATGCTGATTGTGACGGCGGTATTGTACCTTGTTTGGCCGTTTGT
CTACGATGCCATCGTCCTCTTTGCAACGGGAATTTCAAAATTAGGTTTCGTGGGGGCTGGTCT
ATATGGCTTTTTCAATCGTTTATTGATTCCGACTGGTTTGCACCACGCCTTAAATTCCGTATTC
TGGTATAACGTTGCGGGGATCAATGATATTGGCAATTTCTGGGCCAGCCATGGTGTTAAAGG
AATCACAGGGATGTATGAAGCTGGATTCTTCCCAATTATGATGTTTGGCTTGCCAGCTGGTGC
GTATGCAATTTATCGTAATGCACGACCGGAACGAAAAAAAGAAGTGGGTTCATTGATGTTAG
CTGGTGCGTTTGCGTCGTTCTTTACGGGGGTGACCGAACCGCTTGAATTTTCATTCATGTTCG
TGGCATGGCCGCTCTACTTATTACATGCCGTTTTTATGGGACTATCATTAGGATTCGCCGCGT
TAATGCACTGGACCGCCAGCTTTTCGTTTAGTGGCGGGTTAGTCGATTATTTATTGAGTTTTC
GGATGCCGTTAGCCAACCAACCCTATATGTTACTGGTTCAAGGGCTGGTGATGGCCGTCATTT
ATTACTTTGGTTTTGACTTTGCAATCAAACGGTTTAATTTGAAGACGCCTGGTCGTGAAGTTG
TGTCTGCAGATGTCGATGGGGTGGGTGCACCTGCTAGTCCAGCGGTGGCCGTTGCAGCAACG
GATGATAAGTATATGCGCCAAGCAAAGCAAATTTATGCAGCTATTGGTGGTCATGACAATAT
TAGTGTGATCAACAATTGTACGACGCGGCTGCGGTTACAACTTAAGGATACGGAAAAAGTCG
ATCAGCCGGCCGTAATGGCTGCTGGCGTGCCTGGTTTGAACGTACTGGATGTTCATAACATTC
ACATTGTGATCGGCACGGAAGTCCAGTTCGTCGCGGAGGCTTTACAAAAATTATTTTCTGGTC
AGGTGGCGACGACCCCGGCATCTGATGCTGAATCAAAGGCCCCGGTTGAACCGCAAACTGCT
ACTGTAACGGAAGCGCCGGTGACAACGATTTTGCGGGCACCGGCAACGGGACAATTAATGC
CGATTAGTACGGTTGCGGATGAGACGTTTGCTGGTAAACTCTTAGGTGATGGTTACGCTGTTG
AACCCGAAGATGGTGAAGTCGTTGCTCCGGTTAGTGGGACAGTGACAAGTGTCTTTCCGACC
AAACACGCTATCGGTTTGAAGACGACGAGTGGTTTAGAAGTCTTATTACACATGGGGATCAA
TACCGTGGAAATGAATGGTACGCCGTTCAAGTTACACGTGGCAACGGGTGATGAAATTGCGG
CCGGTACTGCGGTAGCGACGGTTGACTTGGCTGCTATCAAGTCAGCCGGTAAAGCGACAACC
ATGATGGTCGTTATCACCAATATGGATCACGTTAACAAACTAACCCTTAATCCAACTGGTCA
TGTGACTAGCGGTGATTTGATTGGCGCAGCTGAAGCGGCCGCACTCGAGGACTACAAAGACG
ACGACGACAAATGAGGATCCACTAGTTCTAGAGCGGCCGCCACCGCGGTGGAGCTCCAGCTT
TTGTTCCCTTTAGTGAGGGTTAATTGCGCGCTTGGCGTAATCATGGTCATAGCTGTTTCCTGT
GTGAAATTGTTATCCGCTCACAATTCCACACCTGGGGTGCCTAATGAGTGAGCTAACTCACA
TTAATTGCGTTGCGCTCACTGCCCGCTTTCCAGTCGGGAAACCTGTCGTGCCAGCTGCATTAA

TGAATCGGCCAACGCGCGGGGAGAGGCGGTTTGCGTATTGGGCGCTCTTCCGCTTCCTCGCT
CACTGACTCGCTGCGCTCGGTCGTTCGGCTGCGGCGAGCGGTATCAGCTCACTCAAAGGCGG
TAATACGGTTATCCACAGAATCAGGGGATAACGCAGGAAAGAACATGTGAGCAAAAGGCCA
GCAAAAGGCCAGGAACCGTAAAAAGGCCGCGTTGCTGGCGTTTTTCCATAGGCTCCGCCCCC
CTGACGAGCATCACAAAAATCGACGCTCAAGTCAGAGGTGGCGAAACCCGACAGGACTATA
AAGATACCAGGCGTTTCCCCCTGGAAGCTCCCTCGTGCGCTCTCCTGTTCCGACCCTGCCGCT
TACCGGATACCTGTCCGCCTTTCTCCCTTCGGGAAGCGTGGCGCTTTCTCATAGCTCACGCTG
TAGGTATCTCAGTTCGGTGTAGGTCGTTCGCTCCAAGCTGGGCTGTGTGCACGAACCCCCCGT
TCAGCCCGACCGCTGCGCCTTATCCGGTAACTATCGTCTTGAGTCCAACCCGGTAAGACACG
ACTTATCGCCACTGGCAGCAGCCACTGGTAACAGGATTAGCAGAGCGAGGTATGTAGGCGGT
GCTACAGAGTTCTTGAAGTGGTGGCCTAACTACGGCTACACTAGAAGGACAGTATTTGGTAT
CTGCGCTCTGCTGAAGCCAGTTACCTTCGGAAAAAGAGTTGGTAGCTCTTGATCCGGCAAAC
AAACCACCGCTGGTAGCGGTGGTTTTTTTGTTTGCAAGCAGCAGATTACGCGCAGAAAAAAA
GGATCTCAAGAAGATCCTTTGATCTTTTCTACGGGGTCTGACGCTCAGTGGAACGAAAACTC
ACGTTAAGGGATTTTGGTCATGAGATTATCAAAAAGGATCTTCACCTAGATCCTTTTAAATTA
AAAATGAAGTTTTAAATCAATCTAAAGTATATATGAGTAAACTTGGTCTGACAGTTACCAAT
GCTTAATCAGTGAGGCACCTATCTCAGCGATCTGTCTATTTCGTTCATCCATAGTTGCCTGAC
TCCCCGTCGTGTAGATAACTACGATACGGGAGGGCTTACCATCTGGCCCCAGTGCTGCAATG
ATACCGCGAGACCCACGCTCACCGGCTCCAGATTTATCAGCAATAAACCAGCCAGCCGGAAG
GGCCGAGCGCAGAAGTGGTCCTGCAACTTTATCCGCCTCCATCCAGTCTATTAATTGTTGCCG
GGAAGCTAGAGTAAGTAGTTCGCCAGTTAATAGTTTGCGCAACGTTGTTGCCATTGCTACAG
GCATCGTGGTGTCACGCTCGTCGTTTGGTATGGCTTCATTCAGCTCCGGTTCCCAACGATCAA
GGCGAGTTACATGATCCCCCATGTTGTGCAAAAAAGCGGTTAGCTCCTTCGGTCCTCCGATC
GTTGTCAGAAGTAAGTTGGCCGCAGTGTTATCACTCATGGTTATGGCAGCACTGCATAATTCT
CTTACTGTCATGCCATCCGTAAGATGCTTTTCTGTGACTGGTGAGTACTCAACCAAGTCATTC
TGAGAATAGTGTATGCGGCGACCGAGTTGCTCTTGCCCGGCGTCAATACGGGATAATACCGC
GCCACATAGCAGAACTTTAAAAGTGCTCATCATTGGAAAACGTTCTTCGGGGCGAAAACTCT
CAAGGATCTTACCGCTGTTGAGATCCAGTTCGATGTAACCCACTCGTGCACCCAACTGATCTT
CAGCATCTTTTACTTTCACCAGCGTTTCTGGGTGAGCAAAAACAGGAAGGCAAAATGCCGCA
AAAAAGGGAATAAGGGCGACACGGAAATGTTGAATACTCATACTCTTCCTTTTTCAATATTA
TTGAAGCATTTATCAGGGTTA

5.4.6 Unpublished data

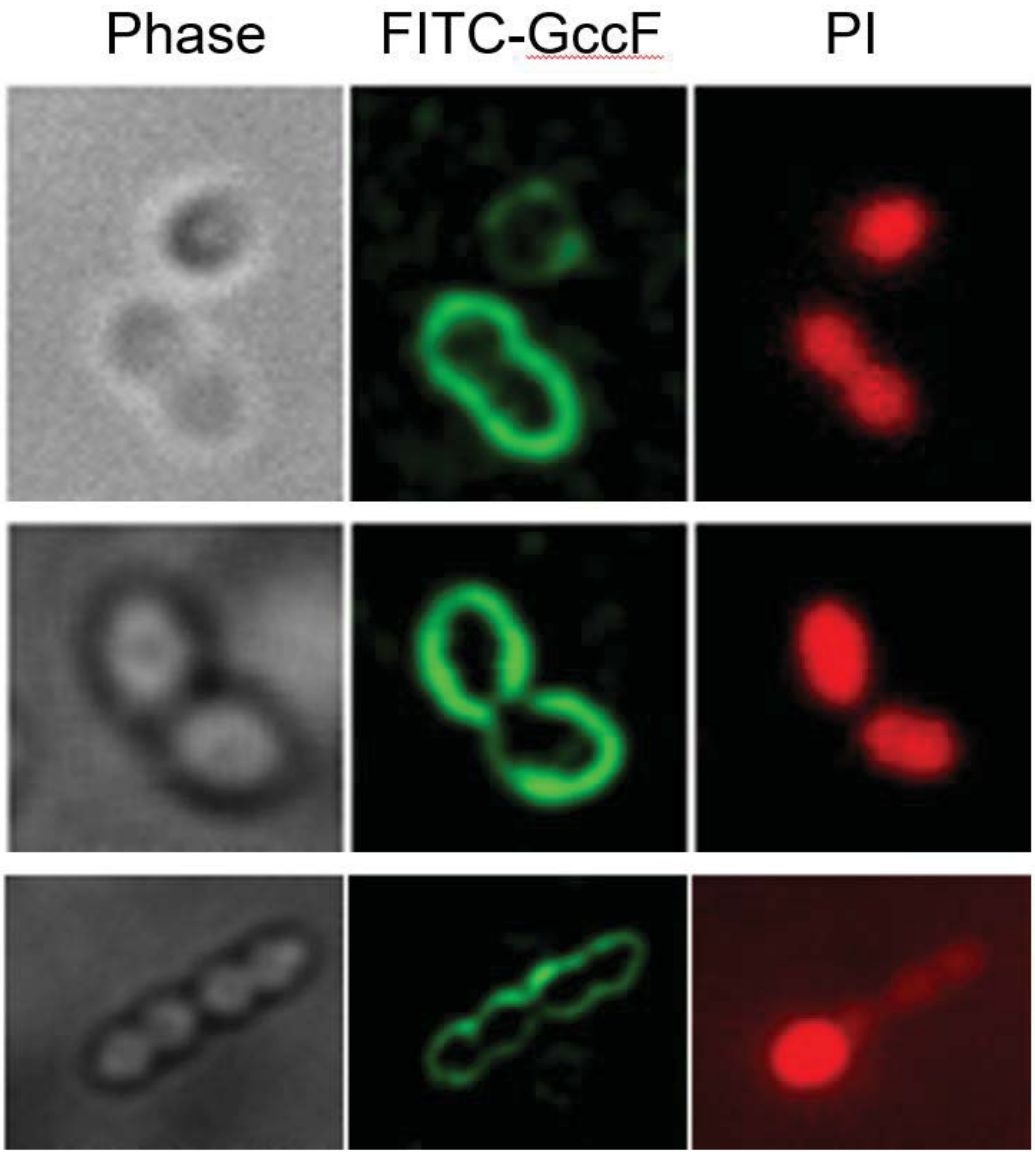


Figure 54: **GccF localises to the bacterial membrane**

L. plantarum 8014 cells treated with fluorescein isothiocyanate (FITC) labelled GccF (Green) and observed under Olympus IX17 inverted microscope in the presence of mounting media, Vectashield (Vector Laboratories). The cells DNA was counterstained with Propidium iodide (Red). Image provided by Associate professor Gill Norris, as personal communications.