The Bacteriostatic Diglycosylated 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

Kelvin Ross Drower

2014

Dedicated to Nana and Pop

Abstract

The increasing prevalence of antibiotic-resistance bacteria is threatening to end the antibiotic era established following Alexander Fleming's discovery of penicillin in 1928. Over-prescription and misuse of broad-spectrum antibiotics has hastened the development and spread of antibiotic resistance. This, combined with a lack of research and development (R&D) of new antibiotics by major pharmaceutical companies, may lead to a widespread recurrence of 'incurable' bacterial diseases. However while commercial R&D of antibiotics has waned, much research has been carried out to characterise bacteriocins, ribosomally-synthesised antimicrobial polypeptides thought to be produced by virtually all prokaryotes. Although hundreds of bacteriocins have been identified and characterised, only a handful of their cognate receptors on susceptible cells have been identified. Glycocin F is a bacteriostatic diglycosylated 43-amino acid bacteriocin produced by the Gram-positive bacterium *Lactobacillus plantarum* KW30 that inhibits the growth of a broad range of bacteria. The mechanism of action of glycocin F is unknown, however evidence suggested that glycocin F binds to cells via a N-acetylglucosamine (GlcNAc) specific phosphoenolpyruvate:carbohydrate-phosphotransferase system (PTS) transporter, as had been shown for lactococcin A, lactococcin B and microcin E492 that target a mannose specific PTS transporter. These other bacteriocins are, however, bactericidal suggesting that glycocin F uses a different mechanism of action to stop cell growth.

To test the hypothesis that one of the putative GlcNAc-specific PTS transporters identified in glycocin F-sensitive *L. plantarum* strains is the primary membrane receptor for glycocin F, a GlcNAc-specific PTS transporter gene knockout mutant was generated and analysed for glycocin F sensitivity. The GlcNAc-specific PTS transporter, *pts18CBA*, was successfully knocked out in *L. plantarum* NC8 which conferred the resulting *L. plantarum* NC8 Δ*pts18CBA* a degree of resistance to glycocin F confirming the GlcNAc-specific PTS transporter is a receptor of glycocin F. Additionally the genomes of wild-type (glycocin F sensitive) *L. plantarum* ATCC 8014, *L. plantarum* subsp. *plantarum* ATCC 14917, and multiple glycocin F-resistant mutants of these two strains were sequenced, assembled and comparatively analysed to identify changes consistent with increased resistance to glycocin F. Mutations, mapped to *pts18CBA* in all sequenced mutants, appeared to be deleterious to both the structure and function of PTS18CBA. A correlation of glycocin F resistance to the degree of mutation in the transmembrane domain of the *pts18CBA* gene was established confirming that glycocin F targets the EIIC transmembrane domain of PTS18CBA.

Acknowledgments

Thank you to Dr Gill Norris and Dr Mark Patchett for the patience, tolerance, support and guidance that you have shown me

Gratitude goes to Associate Professor Murray Cox for the initial direction provided for the genome assembly

Thank you Trevor Loo, Soyeon Ahn and all the members of X-lab and IFS

A thank you goes out to the team at the MMIC, specifically Jordan, for the use of the camera

Thank you Stuart, Fareeda, Nat and Jan for your help both in and out of the lab

To all members of my family, especially Mum, Dad, Jennifer, Eric and Brendon, thank you for your support

Thank you Mel for teaching me the lessons of life

List of Abbreviations

3D	Three-dimensional
Å	Ångström (0.1 nm)
ABC	ATP-binding cassette
ADP	Adenosine diphosphate
ATP	Adenosine triphosphate
bp	Base Pair
CCA	Carbon catabolite activation
CcpA	Carbon catabolite protein A
CCR	Carbon catabolite repression
CDS	Coding DNA sequence
cm	Centimetre
CRE	Catabolite responsive element
ChbC	N,N'-diacetylchitobiose-specific PTS from *B. cereus*
da	Dalton
DNA	Deoxyribonucleic acid
dNTP	Deoxyribonucleotide triphosphate
E-06	Micro
EDTA	Ethylenediaminetetraacetic acid
EII	Enzyme I
EII	Enzyme II
EIIA	Enzyme IIA
EIIB	Enzyme IIB
EIIC	Enzyme IIC
EIID	Enzyme IID
FBP	Fructose-1,6-bisphosphate
g	Gram
gDNA	Genomic DNA
GlcNAc	N-acetylglucosamine
GlpK	Glycerol kinase
His_6	Hexa-Histidine
Hpr	Histidine-phosphorylation protein
HPrK/P	HPr kinase/phosphatase
IPTG	Isopropyl β-D-1-thiogalactopyranoside
ITC	Isothermal titration calorimetry
kbp	Kilobasepair
kDa	Kilodalton
kPa	Kilopascal
L	Litre
LAB	Lactic acid bacteria
Lac	Lactose
LB	Luria-Bertani medium
LB agar	Luria-Bertani medium agar

M	Molar
MccE492	Microcin E492
MIC	Minimum inhibition concentration
MCS	Multiple cloning site
mg	Milligram
MGS	Massey genome service
ms	Millisecond
nL	Nanolitre
NCBI	National Center for Biotechnology Information
µL	Microlitre
µM	Micromolar
mL	Millilitre
mM	Millimolar
MOA	Mechanism of action
MLST	Multilocus sequence typing
MRS	De Man, Rogosa and Sharpe medium
MscL	Large-conductance mechanosensitive channel
MW	Molecular weight
NaCl	Sodium chloride
NGS	Next generation sequencing
°C	Degrees Celsius
OD_{600nm}	Optical density at 600 nm
PCR	Polymerase chain reaction
PDB	Protein data bank
PEG	Polyethylene glycol
PEP	Phosphoenolpyruvate
PMF	Proton motive force
PRD	PTS regulatory domain
PTM	Post-translational modification
PTS	Phosphoenolpyruvate:carbohydrate-phosphotransferase system
RBS	Ribosome binding site
RMSD	Root mean square deviation
SDS-PAGE	Sodium dodecyl sulfate polyacrylamide gel electrophoresis
TBE	Tris-Boric Acid-EDTA
TCBD	Transporter classification database
TEMED	*N,N,N′,N′*-tetramethylethylenediamine
Tm	Melting temperature
TMH	Transmembrane helices
UV	Ultra violet
V	Volts
v/v	Volume/volume
w/v	Weight/volume
WT	Wild-type
x g	Multiple of earth's gravitational force

Amino Acid and Nucleotide Abbreviations

Amino Acids

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

Nucleotides

Adenine	A
Thymine	T
Cytosine	C
Guanine	G
Uracil	U

Table of Contents

List of Figures

List of Tables

1. Introduction

1. Introduction

For thousands of millennia limited resources in closed habitats have driven the evolution of mechanisms that can enhance the survival of an organism. One mechanism that has evolved independently in all three taxonomic domains is the production and secretion of molecules that either kill or inhibit the growth of other organisms competing for the same resources (Cotter *et al.*, 2005, Takayuki and Van Der Hoorn, 2008, Frederick Harris *et al.*, 2009). Prokaryote species secrete ribosomally-synthesised antimicrobial polypeptides called bacteriocins. Bacteriocin producers are usually immune to the growth inhibiting effects of their own bacteriocins due to the production of bacteriocin-specific immunity proteins. The bacteriocin and immunity protein are usually part of a gene cluster which can be either chromosomal or located on a plasmid (Diep *et al.*, 2007, Kjos *et al.*, 2010b, Stepper *et al.*, 2011, van Belkum *et al.*, 2011). Much research has been undertaken to characterise them as they show great potential for inhibiting bacteria involved in pathogenesis and food spoilage (Cotter *et al.*, 2013).

1.1 Classification of bacteriocins

The structural, functional and phylogenetic diversity of the bacteriocins has been both a powerful driving force and a significant challenge for the development of a non-controversial classification scheme (Maisnier-patin *et al.*, 1996, Kalmokoff and Teather, 1997, Zamfir *et al.*, 1999, Bizani and Brandelli, 2002, Zouhir *et al.*, 2010). The extensive history of the bacteriocin classification schemes is beyond the scope of this introduction but, Cotter *et al.* (2005), J Nissen-Meyer *et al.* (2009), and Zouhir *et al.* (2010) all provide good reviews on the topic. An early classification scheme for LAB bacteriocins that was accepted by bacteriocin researchers was presented by Klaenhammer (1993). This divided bacteriocins into four major classes based on post-translational modifications, structural and chemical properties: (class I) post-translational modified (PTM) lantibiotics, (class II) unmodified bacteriocins with a molecular weight (MW) of 10 kilodaltons (kDa) or less, (class III) large heat-labile bacteriocins with a MW of 30 kDa or larger and (class IV) complex post-translationally modified bacteriocins containing lipid or carbohydrate moieties. More recently Cotter *et al.* (2013) proposed a classification scheme (Table 1.1) which has only two classes: class I containing modified bacteriocins and class II containing unmodified or cyclic bacteriocins. Each class is divided into subclasses based on 'distinctive features'. While this two-class classification scheme is problematic in that some bacteriocins fall into more than one subclass, and high molecular weight bacteriocins are excluded, it has the advantage of simplicity/clarity and will be used here for the classification of bacteriocins.

Table 1.1 The two class scheme presented by Cotter *et al.* (2013)

Group	Distinctive feature	Examples
Class I (modified)		
MccC7-C51-type bacteriocins	Is covalently attached to a carboxy-terminal aspartic acid	MccC7-C51
Lasso peptides	Have a lasso structure	MccJ25
Linear azole- or azoline-containing peptides	Possess heterocycles but not other modifications	MccB17
Lantibiotics	Possess lanthionine bridges	**Nisin**, planosporicin, mersacidin, actagardine, mutacin 1140
Linaridins	Have a linear structure and contain dehydrated amino acids	Cypemycin
Proteusins	Contain multiple hydroxylations, epimerizations and methylations	Polytheonamide A
Sactibiotics	Contain sulphur–α-carbon linkages	Subtilosin A, thuricin CD
Patellamide-like cyanobactins	Possess heterocycles and undergo macrocyclization	Patellamide A
Anacyclamide-like cyanobactins	Cyclic peptides consisting of proteinogenic amino acids with prenyl attachments	Anacyclamide A10
Thiopeptides	Contain a central pyridine, dihydropyridine or piperidine ring as well as heterocycles	Thiostrepton, nocathiacin I, GE2270 A, philipimycin
Bottromycins	Contain macrocyclic amidine, a decarboxylated carboxy-terminal thiazole and carbon-methylated amino acids	Bottromycin A2
Glycocins	Contain *S*-linked glycopeptides	**Sublancin 168**, g**lycocin F**, Thurandacin
MccC7-C51-type bacteriocins	Is covalently attached to a carboxy-terminal aspartic acid	MccC7-C51
Lasso peptides	Have a lasso structure	MccJ25
Class II (unmodified or cyclic)		
IIa peptides (pediocin PA-1-like bacteriocins)	Possess a conserved YGNGV motif (in which N represents any amino acid)	Pediocin PA-1, enterocin CRL35, carnobacteriocin BM1
IIb peptides	Two unmodified peptides are required for activity	ABP118, lactacin F
IIc peptides	Cyclic peptides	Enterocin AS-48, **garvicin ML**
IId peptides	Unmodified, linear, non-pediocin-like, single-peptide bacteriocins	MccV, MccS, epidermicin NI01, **lactococcin A**, **lactococcin B**
IIe peptides	Contain a serine-rich carboxy-terminal region with a non-ribosomal siderophore-type modification	**MccE492**, MccM

Mcc, microcin. Table reprinted, containing minor changes, with permission from Cotter *et al.* (2013). Bacteriocins in bold are those discussed in greater detail in later sections.

1.2 Bacteriocins mechanism of action

Bacteriocins generally affect susceptible cells in two ways: they can either kill the cell (bactericidal), or stop cell growth (bacteriostasis). Their antimicrobial spectrum of activity is generally phylogenetically specific and narrow, although occasionally the inhibitory action of a bacteriocin can affect a broad range of bacteria suggesting that cellular recognition mechanisms differ between bacteriocins (Muriana and Klaenhammer, 1991a, Muriana and Klaenhammer, 1991b, Stevens *et al.*, 1991, Jack *et al.*, 1995, Bizani and Brandelli, 2002). Hundreds of bacteriocins have been identified and characterised, yet the cognate receptors for susceptible cells have only been identified for a handful of cases. For some bacteriocins, such as the two-component class IIb peptide plantaricins, a specific receptor molecule is not necessarily involved in their mechanism of action (MOA). For example some of the plantaricins target the bacterial cell membrane and function by dissipating the membrane potential (Lopes *et al.*, 2013, Sharma and Srivastava, 2014). So far, receptors that have been identified are either specific proteins, explaining the defined phylogenetic inhibitory spectrum exhibited by some bacteriocins (Diep *et al.*, 2007, Kouwen *et al.*, 2009, Kjos *et al.*, 2011a, Gabrielsen *et al.*, 2012), or a ubiquitous non-proteinaceous molecule as is the case for the lantibiotic nisin (Hasper *et al.*, 2004, Wiedemann *et al.*, 2004). Specific proteins identified as bacteriocin receptors have either been transmembrane proteins, which is generally the case for the bacteriocins produced by Gram-positive bacteria (Diep *et al.*, 2007, Kouwen *et al.*, 2009, Gabrielsen *et al.*, 2012), or an intracellular enzyme, which is generally the case for bacteriocins produced by Gram-negative bacteria (Delgado *et al.*, 2001, Zamble *et al.*, 2001, Yuzenkova *et al.*, 2002, Metlitskaya *et al.*, 2006). As always, there are exceptions to these generalisations as shown by the Gram-negative bacteriocin microcin E492 which targets a mannose-specific PTS transporter (Biéler *et al.*, 2006, Biéler *et al.*, 2010), and is discussed in section 1.2.5.

The targeting of carbohydrate transmembrane transporters is becoming an increasingly common MOA for the Gram-positive bacteriocins. The permease subunit of the mannose-specific PTS transporter has been identified as the receptor for the Gram-positive bacteriocins lactococcin A and lactococcin B (Diep *et al.*, 2007), and is discussed at greater length in section 1.2.4. Another carbohydrate transporter, the maltose ATP-binding cassette (ABC) transporter, was reported by Gabrielsen *et al.* (2012) to be a receptor for the Gram-positive circular bacteriocin garvicin ML, and is discussed at greater length in section 1.2.3. Lactococcin A and lactococcin B kill target cells by forming pores in the cell membrane which leads to dissipation of the proton motive force (PMF) (van Belkum *et al.*, 1991, Venema *et al.*, 1993). The mechanism by which garvicin ML kills cells has yet to be shown. Another Gram-positive bacteriocin that has been shown

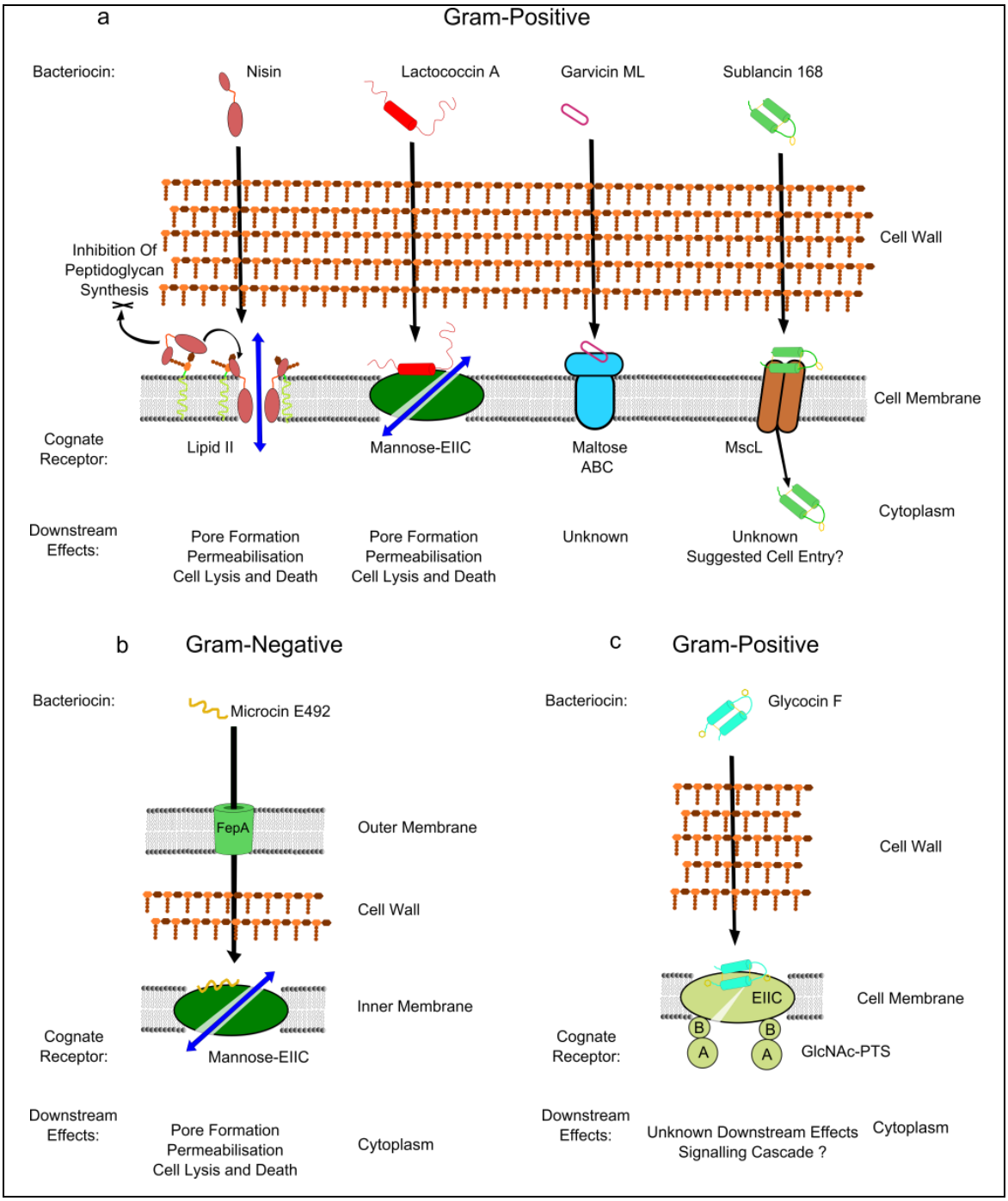


Figure 1.1 Mechanisms of action of selected bacteriocins

A schematic diagram showing selected bacteriocins, their cognate receptors and possible downstream consequences of the binding. Blue double ended arrows indicate membrane depolarisation. See text for details on mode of actions on (a) the Gram-positive bacteriocins nisin (section 1.2.1), lactococcin A (section 1.2.4), garvicin ML (section 1.2.3) and sublancin 168 (section 1.2.2). MscL is the large-conductance mechanosensitive channel. b) The Gram-negative bacteriocin microcin E492 (section 1.2.5). FepA is an outer-membrane catecholate siderophore receptor through which microcin E492 is translocated in a TonB dependent fashion (TonB not shown). c) The suggested mechanism of the Gram-positive bacteriocin glycocin F (section 1.4).

to bind to a transmembrane protein is sublancin 168, and is discussed at greater length in section 1.2.2. Like garvicin ML, however, the mechanism that sublancin 168 uses to kill target cells is unknown (Kouwen *et al.*, 2009). Sublancin 168 is one of three identified glycocins, peptides containing a cysteine linked sugar molecule (*S*-linked glycan) (Table 1.1), the others being glycocin F and thurandacin (Stepper *et al.*, 2011, Wang *et al.*, 2013). However glycocin F is different to sublancin 168, as it has two covalently linked sugars, one *S*-linked and the other a more conventionally *O*-linked, that are both essential for function (Stepper *et al.*, 2011). In contrast, the sugar linked to sublancin 168 is not essential for antimicrobial activity (Wang and van der Donk, 2011). As it was shown that free GlcNAc could reverse the action of glycocin F, and as PTS transporters are known to be involved in the action of microcin E492, lactococcin A and lactococcin B, it was predicted that a sugar transporter may be involved in the MOA of glycocin F (Stepper *et al.*, 2011).

1.2.1 Class I lantibiotics

Hundreds of putative lantibiotics have been identified but only a few have been characterised in detail. The 34 amino acid lantibiotic nisin was the first and so far the most widespread commercially used bacteriocin (Mattick and Hirsch, 1944, Chen and Hoover, 2003). Its antimicrobial activity was first shown by Rogers (1928) and it has become the most intensively studied bacteriocin to date with a detailed MOA being proposed by Wiedemann *et al.* (2004) who showed that the antimicrobial action of nisin is the result of two distinct mechanisms (Figure 1.1a). The first mechanism involves nisin binding to the cell wall peptidoglycan precursor lipid II (binding constant of 200 nM), which prevents the disaccharide moiety of lipid II from being incorporated into the peptidoglycan layer, thus halting cell wall synthesis (Brötz *et al.*, 1998). In the second mechanism, nisin-lipid II complexes induce cell death through pore formation in the membrane of the target cell. The resulting pores dissipate the proton motive force (PMF) and allow an efflux of cellular metabolites such as glutamate and other amino acids. This essentially halts the cellular biosynthetic pathways by removing biosynthetic substrate molecules and ATP (Elke Ruhr and Sahl, 1985, Sahl *et al.*, 1987, Brötz *et al.*, 1998, Wiedemann *et al.*, 2001). This second mechanism is dependent on lipid II, as in its absence nisin concentration required for pore formation increased from 15 nM to 1.5 µM (Christ *et al.*, 2007).

1.2.2 Class I glycocins

A distinctive feature of the glycocins is a cysteine-linked (*S*-linked) glycan (saccharide), which are rare in nature. Currently only three bacteriocins containing *S*-linked glycans have been

reported: glycocin F (Stepper *et al.*, 2011), sublancin 168 (Oman *et al.*, 2011) and thurandacin (Wang *et al.*, 2013). Of these, only the cognate receptor for sublancin 168 has been identified. Sublancin 168 was originally characterised by Paik *et al.* (1998) as a lantibiotic. It was later correctly identified by Oman *et al.* (2011) to be a glycocin and its 3D structure was recently solved by Garcia De Gonzalo *et al.* (2014) using NMR techniques (Figure 1.2). Its structure is similar to that of glycocin F (Figure 1.6), and many other toxins, and the structure is known as a Cs α/α fold (Möller *et al.*, 2005). These structures are characterised by a helix-loop-helix architecture stabilised by two nested disulfide bonds. The β-*S*-linked glucose is located in the loop region between two α-helices, and Wang and van der Donk (2011) have shown that the *S*-linked glucose of sublancin 168 is not required for antimicrobial activity. The antimicrobial activity of sublancin 168 was shown however to be attenuated by 5 % (w/v) NaCl which led to the identification of the large-conductance mechanosensitive channel (MscL) as the receptor for sublancin 168 (Figure 1.1a) by Kouwen *et al.* (2009) by constructing a knockout of the *mscL* gene.

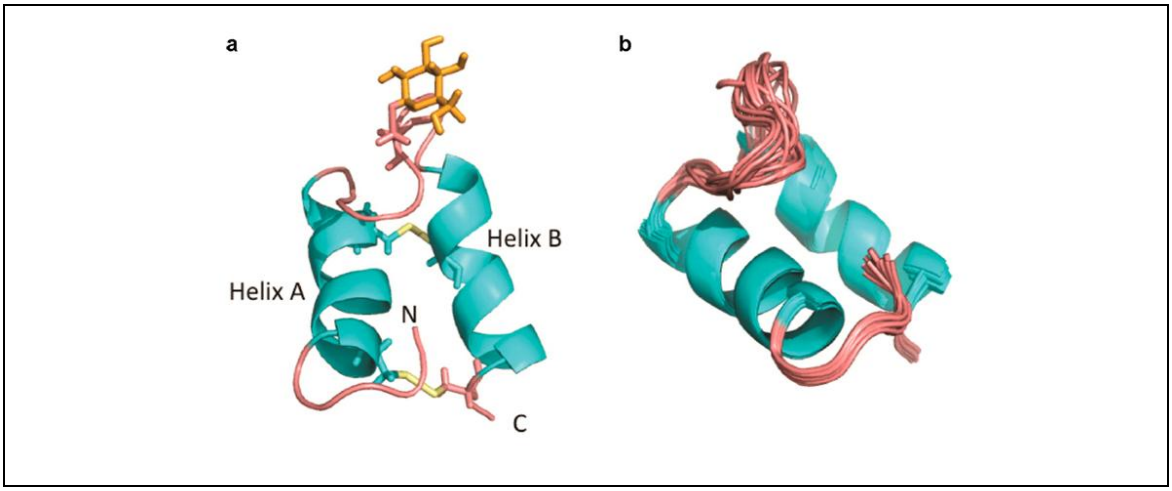


Figure 1.2 Structure of sublancin 168
a) The 3D structure of sublancin 168 and (b) the superposition of the 15 conformers with the lowest energy. The N- and C- terminus are labelled and the glucose moiety is coloured orange. The two helices are labelled 'Helix A' and 'Helix B' (cyan) with the nested disulfide bonds (yellow) located either side of the helices. A more detailed description of the structure is provided in text (section 1.2.2). Figure adapted with permission from Garcia De Gonzalo *et al.* (2014). Copyright (2014) American Chemical Society.

1.2.3 Class IIc peptides

Garvicin ML is a class IIc peptide, isolated from *Lactococcus garvieae* DCC43, that has broad-spectrum antimicrobial activity against Gram-positive bacteria (Sánchez *et al.*, 2007, Borrero *et*

al., 2011). Gabrielsen *et al.* (2012) isolated naturally-occurring garvicin ML-resistant *L. lactis* subsp. *lactis* Il1403 strains and identified, *via* multilocus sequence typing (MLST) and Polymerase Chain Reaction (PCR) analysis, that the resistance was due to the deletion of a 13.5 kbp region of the genome containing the maltose ABC-encoding genes *malE*, *malF* and *malG*. This region of DNA was flanked by transposable elements suggesting transposon-mediated DNA recombination was responsible for the deletion (de Visser *et al.*, 2004). Upon complementation of the maltose ABC-encoding genes the sensitivity of the garvicin ML-resistant strains was restored (Gabrielsen *et al.*, 2012) confirming the maltose ABC transporter as the receptor for garvicin ML (Figure 1.1a). The mechanism by which garvicin ML kills cell remains to be resolved.

1.2.4 Class IId peptides

Lactococcin A and lactococcin B are class IId thiol-activated bacteriocins produced by *Lactococcus lactis* subsp. *cremoris* LMG 2130 (Holo *et al.*, 1991) and *L lactis* subsp. *cremoris* 9B4 respectively (Venema *et al.*, 1993). They kill cells by permeabilising target membranes, which results in dissipation of the PMF, in a voltage-independent manner, allowing the efflux of small solutes such as alanine through the membrane. However both lactococcin A and lactococcin B do not dissipate the PMF of liposomes indicating that a membrane protein must be required for activity (van Belkum *et al.*, 1991, Venema *et al.*, 1993). Diep *et al.* (2007) identified this membrane protein receptor by characterising the complexes that lactococcin A immunity proteins make with membrane proteins. To do this they used *in vivo* pull down experiments and peptide mass fingerprinting to identify co-immunoprecipitated proteins. The mannose-PTS subunits EIIAB, EIIC and EIID were shown to immunoprecipitate with the lactococcin A immunity protein in a lactococcin A-dependent manner, which strongly suggested that the mannose-PTS transporter was a receptor for lactococcin A. The results indicated that the lactococcin A immunity protein formed a complex with the EIIC/D subunits of the mannose-PTS transporter but could not determine if the immunity protein interacted directly with the bacteriocin. Venema *et al.* (1993) had previously proposed three possible models for the mechanisms of immunity in lactococcin B: firstly the immunity proteins sterically or electrostatically hinder binding of the bacteriocin to its receptor; secondly the immunity protein physically interacts with the bacteriocin to inhibit bacteriocin-receptor interaction; thirdly the pores formed by the bacteriocin are closed by the immunity protein.

Diep *et al.* (2007) confirmed that the mannose-PTS transporter was the receptor for lactococcin A and lactococcin B in a series of experiments. A positive correlation between lactococcin A sensitivity to the abundance of mannose-PTS EIIC and EIID subunits in the cell was

established. Targeted deletion of the *ptn* operon, which contains the mannose-PTS EIIAB, EIIC and EIID genes, from the genome of the lactococcin A- and lactococcin B-susceptible strain *Lactococcus lactis* II1403 was correlated with lactococcin A and lactococcin B resistance. Complementation of the mannose-PTS genes into the *L. lactis* II1403 Δ*ptn* mutant resulted in re-sensitisation to both lactococcin A and lactococcin B. Diep *et al.* (2007) proposed that in the absence of extracellular lactococcin A the immunity protein interacts only transiently with the EIIC/D mannose-PTS subunits and that this interaction becomes more permanent when a bacteriocin binds the transmembrane EIIC permease domain. It was suggested that when a bacteriocin binds to the transmembrane EIIC domain, a conformational change is induced in the EIIC/D subunits resulting in the formation of a pore (Figure 1.1a). When they are in this pore conformation, the EIIC/D subunits have a high affinity for the immunity protein, which effectively blocks the pore, preventing cell death. Thus, the immunity protein is likely to bind to only one of the conformational states of EIIC/D, discussed at greater length in section 1.3.2, that the permease of the PTS transporter adapts (Figure 1.5b).

1.2.5 Class IIe peptides

Microcin E492 (MccE492) is a class IIe peptide produced by the Gram-negative *Klebsiella pneumoniae* RYC492 (de Lorenzo, 1984) that, dependent on the carbon and nitrogen concentration in the growth media, becomes post-translationally modified with a C-terminal siderophore-type modification (Thomas *et al.*, 2004). MccE492 primarily targets and kills species belonging to the *Enterobacteriaceae* family (de Lorenzo, 1984, de Lorenzo *et al.*, 1984, Destoumieux-Garzon *et al.*, 2003, Thomas *et al.*, 2004). The molecular properties of MccE492 have been characterised is great detail. The molecular 'machinery' for the attachment of the siderophore-type modification consists of four enzymes MceC, MceD, MceI and MceJ (Nolan *et al.*, 2007). The modification enhances the targeting of MccE492 to outer-membrane (Figure 1.1b) catecholate siderophore receptors (Patzer *et al.*, 2003, Thomas *et al.*, 2004, Strahsburger *et al.*, 2005) where it is then translocated into the periplasm in a TonB dependent manner (Pugsley *et al.*, 1986, Destoumieux-Garzon *et al.*, 2003). Studies by Biéler *et al.* (2006) and Biéler *et al.* (2010) suggest that once inside the periplasmic space, MccE492 interacts with the inner transmembrane mannose-PTS subunit ManYZ, creating a pore that results in depolarization, permeabilisation and cell lysis (de Lorenzo and Pugsley, 1985, Lagos *et al.*, 1993, Destoumieux-Garzon *et al.*, 2003). In the same study by Biéler *et al.* (2010), it was shown that co-expression of MceA, peptide backbone of microcin E492, and the microcin E492 immunity protein, MceB, in a microcin E492-susceptible cell, interferes with mannose metabolism. These results suggested that MceB does not act to dislodge the bacteriocin from ManYZ, but rather

protects against the toxicity of the bacteriocin-receptor complex, like the lactococcin A immunity protein (section 1.2.4).

1.3 Carbohydrate transporters

According to the Transporter Classification DataBase (TCDB) (Busch and Saier, 2002), transporters that translocate carbohydrates across the membranes fall into three broad classes. These are: (class 1) the channel-type facilitators, (class 2) the secondary carrier-type facilitators, (class 3) the primary active transporters and (class 4) the group translocators. Comparative genome studies by Lorca *et al.* (2007) and Cases *et al.* (2007) provided evidence that the majority of sugar transporters in *Firmicutes* genomes are class 4 phosphoenolpyruvate:carbohydrate-phosphotransferase systems (PTS). The PTS superfamily is divided into seven PTS families: (1) glucose-glucoside (Glc) family, (2) fructose-mannitol (Fru) family, (3) Lactose-N,N'-diacetylchitobiose-β-glucoside (Lac) family, (4) glucitol (Gut) family, (5) galactitol (Gat) family, (6) mannose-fructose-sorbose (Man) family and (7) L-ascorbate (L-Asc) family (Busch and Saier, 2002, Saier *et al.*, 2005, Nguyen *et al.*, 2006, Saier *et al.*, 2006, Saier *et al.*, 2009, Saier *et al.*, 2014).

1.3.1 Phosphoenolpyruvate:carbohydrate-phosphotransferase system

The first PTS transporter characterised, was purified from *E. coli* K235 by Kundig *et al.* (1964) and described as a system to implement the transfer of a phosphate group from phosphoenolpyruvate (PEP) to a hexose. At that time, the biological significance of their discovery was unclear but it would set in motion a new field of scientific research dedicated to bacterial PTSs (to date some 2,000 peer-reviewed scientific publications). Excellent literature reviews covering a greater breadth of material in more detail than can be covered here are recommended (Postma *et al.*, 1993, Saler and Reizer, 1994, Deutscher *et al.*, 2014, Pflüger-Grau and de Lorenzo, 2014).

PTSs were initially discovered and characterised as receptors, kinases and subsequently translocators of sugars (Kundig *et al.*, 1964, Simoni *et al.*, 1967, Simoni and Roseman, 1973, Simoni *et al.*, 1976). The literature now supports PTSs having additional functions involving regulation of carbon metabolism (Cases *et al.*, 1999, Schumacher *et al.*, 2004, Aranda-Olmedo *et al.*, 2005), regulation of nitrogen metabolism (Jin *et al.*, 1994, Powell *et al.*, 1995, Segura and Espin, 1998, King and O'Brian, 2001) and maintaining potassium homeostasis (Luttmann *et al.*, 2009, Untiet *et al.*, 2013). PTSs also play a role in bacterial host interactions. Evidence suggests PTSs are involved in the virulence of some pathogenic bacteria (Tan *et al.*, 1999, Higa and

Edelstein, 2001) and some *Pseudomonas*-rhizome colonisation interactions (Mavrodi *et al.*, 2006, Huang *et al.*, 2009). In addition PTSs seem to have been exploited by both bacteriophage and competing bacteria. It is well documented that the mannose-PTS is utilised by the bacteriophage λ for translocation of DNA across the inner membrane (Elliott and Arber, 1978, Williams *et al.*, 1986, Erni *et al.*, 1987, Esquinas-Rychen and Erni, 2001) and bactericidal bacteriocins target the mannose-PTS (section 1.2) (Biéler *et al.*, 2006, Kjos *et al.*, 2009, Biéler *et al.*, 2010, Kjos *et al.*, 2011b). Based on the different functions of PTSs they have been divided into two groups; the nitrogen-related PTSs that do not participate in sugar transport but are instead involved in the regulation of carbon and nitrogen metabolic processes and the canonical PTSs that function as sugar transporters and can be involved in the regulation of carbon and nitrogen metabolism.

1.3.2 Canonical PTSs

Canonical PTSs are generally composed of five protein domains: histidine-phosphorylation protein (HPr), Enzyme I (EI), Enzyme IIA (EIIA), Enzyme IIB (EIIB) and Enzyme IIC (EIIC) (Kundig *et al.*, 1964, Mitchell *et al.*, 1991) (Figure 1.3). An additional Enzyme IID (EIID) domain is found only in the mannose-PTS. The HPr and EI subunits of the PTS transporter are soluble cytoplasmic proteins that are not carbohydrate-specific and can be shared between the multiple PTS transporters that a bacterial cell will have. In contrast, the EII domains of the PTS are carbohydrate-specific and usually do not associate with the subunits from other PTS transporters. The EII domains can be fused so that a single polypeptide contains two or all three domains in a variety of arrangements (e.g. ABC, BCA, CBA, AB, AC, etc), or the domains can be expressed separately as the EIIC, EIIB and EIIA subunits (Reizer and Saier, 1997, Nguyen *et al.*, 2006). EIIC is a transmembrane spanning permease, EIIB is a hydrophilic membrane-associated protein and EIIA is soluble cytoplasmic protein. All these proteins participate in a phosphorylation cascade (Figure 1.3) that is initiated when EI is autophosphorylated by PEP. EI transfers it to histidine 15 (His-15) of HPr which transfers it to EIIA which in turn transfers it to EIIB which phosphorylates the carbohydrate that is concurrently translocated across the membrane by EIIC (Kundig *et al.*, 1964, Simoni and Roseman, 1973, Simoni *et al.*, 1976). Each step in this phosphorylation cascade is reversible, except for the sugar phosphorylation (Weigel *et al.*, 1982, Meadow and Roseman, 1996, Pelton *et al.*, 1996, Rohwer *et al.*, 2000).

Three-dimensional (3D) structures of all of the PTS domains, except EIID, have been solved from a variety a bacterial species (Cai *et al.*, 2003, Lei *et al.*, 2009, Cao *et al.*, 2011, Navdaeva *et al.*, 2011, Schumacher *et al.*, 2011). For the EIIC permeases only one 3D structure has been

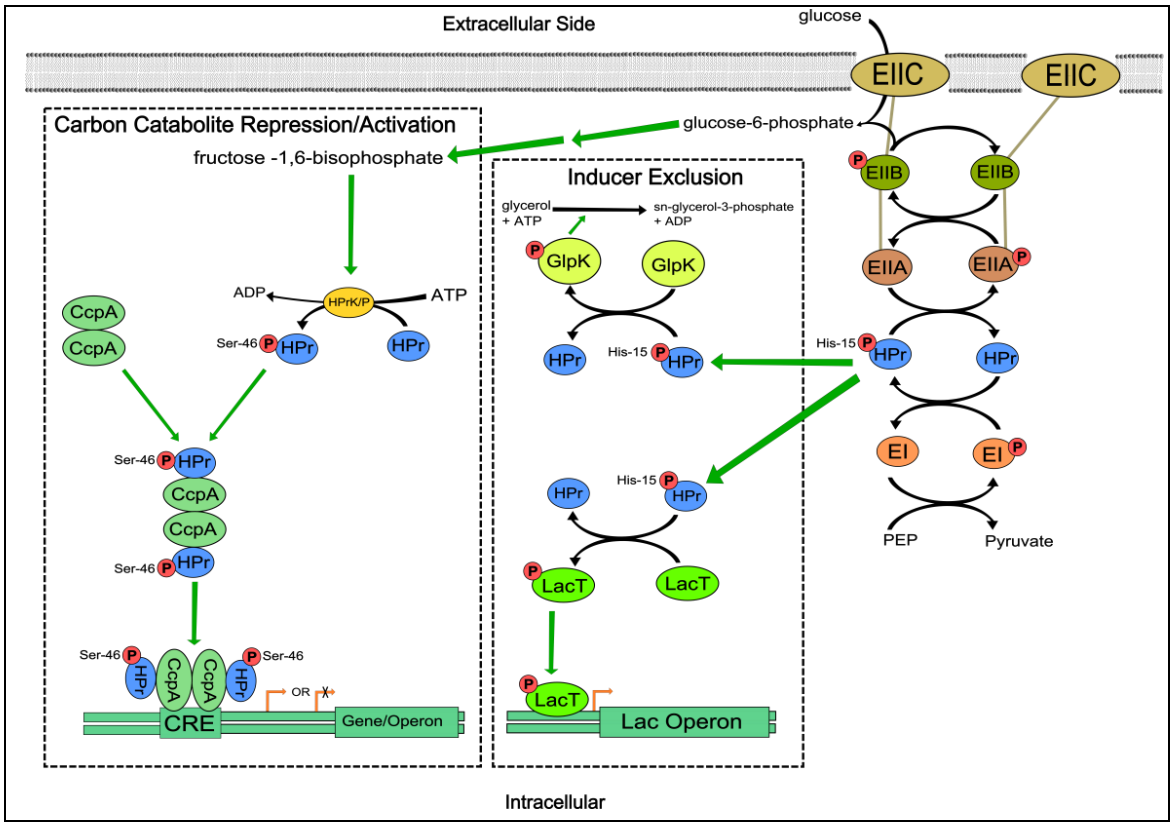


Figure 1.3 PTS Phosphorylation cascade and links to CCR, CCA and Inducer Exclusion

The phosphorylation cascade is initiated with Enzyme I (EI) accepting a phosphate from phosphoenolpyruvate (PEP), EI passes it on to histidine 15 (His-15) of histidine-phosphorylation protein (HPr) which passes it on to Enzyme IIA (EIIA) which passes it to Enzyme IIB (EIIB) which phosphorylates the sugar as it's translocated by Enzyme IIC (EIIC) across the membrane. Though it is not indicated in the diagram, the reaction from PEP to the phosphorylation of EIIB is reversible. Red spheres with 'P' indicate a phosphate group originating from either adenosine triphosphate (ATP) or PEP. HPr has a central role in the carbon catabolite repression (CCR), carbon catabolite activation (CCA) and inducer exclusion. In CCR/CCA the glycolytic intermediate, fructose-1,6-bisphosphate, allosterically activates HPrK kinase/phosphotase (HPrK/P). HPrK/P acts on serine 46 (Ser-46) of HPr, the phosphorylation of which promotes binding to carbon catabolite protein A (CcpA). This interaction recruits the complex to catabolite response elements (CREs) located in promoter and/or open reading frames of genes/operons. Depending on the location of the CRE this recruitment can then activate or repress transcription. The His-15 phosphorylated HPr can allosterically or covalently regulate non-PTS proteins. Shown are two examples of covalent modification of the glycerol kinase (GlpK) and the antiterminator LacT. GlpK is phosphorylated resulting in enhanced enzymatic activity and LacT is phosphorylated on a PTS regulatory domain (PRD) which recruits it to the lactose operon (Lac Operon) resulting in the prevention of transcriptional termination and increased transcription of the Lac Operon genes. The model is based on that proposed by Deutscher *et al.* (2014).

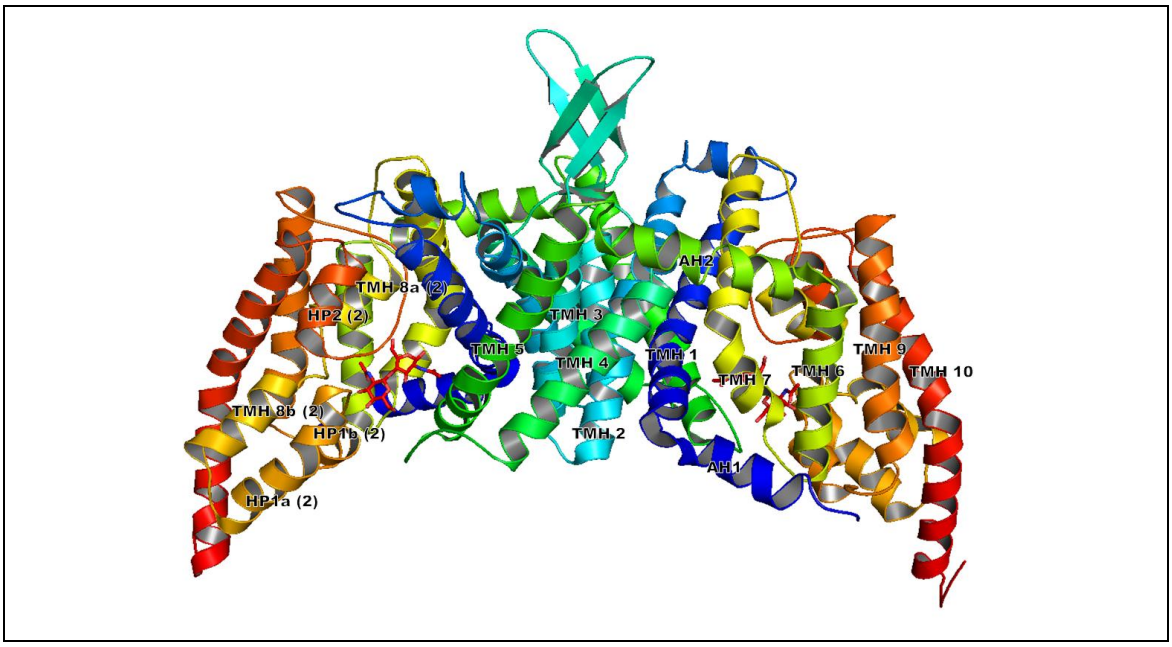


Figure 1.4 Structure of the N,N'-diacetylchitobiose-specific PTS from *B. cereus.*
The structure of the dimer of the EIIC domain of the N,N'-diacetylchitobiose-specific PTS from *B. cereus* was solved at a resolution of 3.3 Å with X-ray diffraction techniques by Cao *et al.* (2011). Transmembrane helices (TMH) and re-entrant loops (HP1a, HP1b, HP2, AH1 and AH2) are named as according to McCoy *et al.* (2014). Helices labelled with the suffix '(2)', are from the backside of the second subunit of the dimer that are not visible from the front view.

reported, that of a N,N'-diacetylchitobiose-specific PTS from *Bacillus cereus* (ChbC) (Cao *et al.*, 2011) (PDB# 3QNQ). The availability of this structure played an important role in the development of carbohydrate translocation mechanisms proposed by both Cao *et al.* (2011) and McCoy *et al.* (2014) (Figure 1.5). The structure (Figure 1.4) shows the N- and C-terminus are located in the cytoplasm of the cell with the interconnecting sequence containing 10 transmembrane helices (TMH) which together create two domains, an oligomerisation domain (TMH 1-5) and transport domain (TMH 6-10). The structure shows that the EIIC dimerises in the membrane, consistent with the conclusions of Lolkema and Robillard (1990) and Chen and Amster-Choder (1998). The structure of the EIIC dimer shows a possible ligand-binding site located on the intracellular side of an electronegative cleft buried by the loop region between TMH 4 and TMH 5 (Figure 1.4). The structure of the EIIC dimer does not reveal a clear passage for the sugar, suggesting conformational changes would have to occur to allow translocation of the sugar through the membrane. Cao *et al.* (2011) proposed that the EIIC dimer structure they solved is in an occluded state and that the EIIC dimer exists in two different states: (1) the outward-open state where the sugar binding interface is facing outward towards the periplasm and (2) the inward-open state where the sugar binding interface is facing the cytoplasm (Figure 1.5).

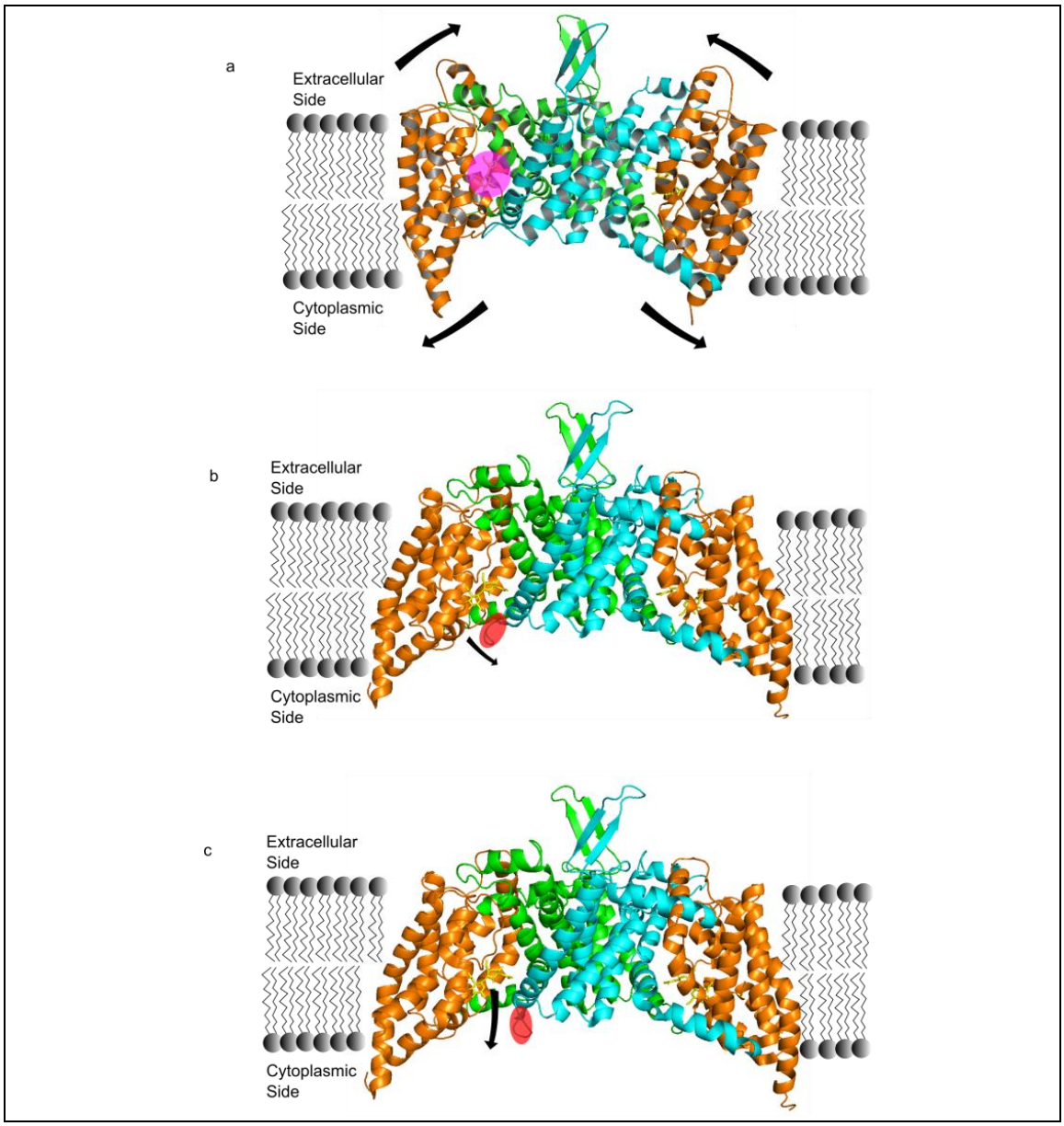


Figure 1.5 Rigid-body rotation and intracellular gate movement in ChbC

The 3D structure of N,N'-diacetylchitobiose-specific PTS from *Bacillus cereus* (ChbC) is predicted to be in an (b) occluded state between an (a) outward-open state and (c) inward-open state. It must be noted that (b) is the reported 3D structure and that (a) and (c) are predicted models based on the reports of Cao *et al.* (2011) and McCoy *et al.* (2014). The model of translocation, as proposed by Cao *et al.* (2011) and McCoy *et al.* (2014), involves carbohydrate binding at the ligand binding site (yellow molecule in all three structures is N,N'-diacetylchitobiose in the proposed ligand binding site, highlighted by opaque magenta circle in a) which induces a rigid body rotation, shown by the black arrows in (a), of the transport domains (orange) relative to the oligomerisation domains (green and cyan). After rotation of the transport domains, the ligand binding site now faces the cytoplasm, however its dissociation is blocked by an intracellular gate (highlighted by opaque red oval in b and c). This gate is predicted to open, as indicated by the black arrow in (b), allowing the dissociation of sugar as shown in (c).

The conformational changes that bridge these two states are thought to occur in the loop regions between TMH 4 and TMH 5 and the loop region between TMH7 and TMH 8a (Figure 1.4 and Figure 1.5) (Cao *et al.*, 2011, McCoy *et al.*, 2014). It is possible that in the outward-open state (Figure 1.5a) the ligand binds to the electronegative cleft in the transport domain, which induces changes in conformation of the loop region between TMH 7 and TMH 8a. This change in conformation acts like a hinge, allowing the transport domains of the dimer to rotate relative to the oligomerisation domains so that the ligand binding site can flip to face the cytoplasm. It then achieves an occluded state (Figure 1.5b) in which the ligand binding site is facing the cytoplasm, but the ligand is prevented from dissociating from the protein by an intracellular gate formed by the loop region between TMH 4 and TMH 5. A conformational change at the N-terminal of TMH 5 is thought to occur (Figure 1.5c), that results in the opening of the intracellular gate and the exposure of the ligand to the cytoplasm. During this translocation the sugar is concurrently phosphorylated by EIIB, which most likely aids the release of the sugar from the electronegative dense cleft and prevents re-association and subsequent efflux of the sugar out of the cell.

1.3.3 Regulatory functions of the canonical PTSs

The un-phosphorylated and the phosphorylated forms of HPr (HPr-His-15-P), EI (EI-P) and EIIA (EIIA-P) have been found to be essential components of the regulatory mechanisms termed 'inducer exclusion', 'carbon catabolite repression' (CCR) and 'carbon catabolite activation' (CCA) (Figure 1.3). These regulatory mechanisms affect a plethora of non-PTS proteins that include transcription factors, transcriptional antiterminators and catalytic enzymes. This action is mediated by three different processes: (1) phosphorylation of a regulatory domain fused to the non-PTS protein, (2) direct phosphorylation of the protein itself or (3) by binding to the protein (Charrier *et al.*, 1997, Greenberg *et al.*, 2002). Currently the literature suggests that the inducer exclusion and CCR mechanisms evolved in response to the need to increase the efficacy the carbon utilisation by ensuing that rapidly metabolised PTS-imported carbon sources are utilised before more complex non-PTS carbon sources (Gorke and Stulke, 2008). The two mechanisms are interconnected and act to convey signals about the current environmental and intracellular carbon situation to an array of metabolic pathways that are usually, though not necessarily, involved in carbon metabolism.

In models of prokaryote regulatory systems, the inducer exclusion mechanism has been shown to effect a variety of catalytic enzymes. These include glycerol kinase (GlpK) (Figure 1.3), lactose permease (LacY), melibiose permease (MelB) and the ATPase domain of the maltose-ABC transporter (MalK) (Osumi and Saier, 1982, Saier *et al.*, 1983, Deutscher, 1985, Deutscher and Sauerwald, 1986, Charrier *et al.*, 1997). However in *Firmicutes* the only catalytic enzyme

that is regulated by inducer exclusion and has a direct link to a PTS is GlpK (Deutscher, 1985, Deutscher and Sauerwald, 1986, Charrier *et al.*, 1997, Yeh *et al.*, 2009) (Figure 1.3). The inducer exclusion on GlpK is initiated by the PTS sugars which when being actively translocated, due to high extracellular concentrations, decrease the cytoplasmic pool of HPr-P and EI-P. These two phosphorylated proteins phosphorylate His232 of GplK increasing its enzymatic activity (Deutscher, 1985, Deutscher and Sauerwald, 1986, Charrier *et al.*, 1997, Yeh *et al.*, 2009). A decrease in the HPr-P:HPr and EI-P:EI ratios results in a decreased phosphorylation rate of GlpK and thus decreased enzymatic activity. This allows metabolic resources to be efficiently utilised for catabolism of the rapidly metabolisable PTS sugars instead of glycerol. Inducer exclusion on GlpK in *Proteobacteria* is identical, relating to the initiating signal being PTS sugars with the result being GlpK inactivation. However in *Proteobacteria*, phosphorylation of GlpK does not occur. Instead the unphosphorylated form of EIIA binds to GlpK with a 1:1 stoichiometry causing complete catalytic inactivation (Postma *et al.*, 1984, Hurley *et al.*, 1993, van der Vlag *et al.*, 1994). The convergent evolution of the PTS sugar-initiated inducer exclusion regulation on GlpK in the two bacterial phyla indicates the importance of efficient carbon utilisation.

Though GlpK is the only known non-PTS catalytic enzyme in *Firmicutes* that is regulated by inducer exclusion through a PTS, some transcription factors and transcriptional antiterminators are also regulated by inducer exclusion, usually through the 1st and 2nd processes above. Transcription factors and antiterminators under the regulatory influence of these PTS components include BgIG (beta-glucoside utilisation), LicT (beta-glucoside utilisation), LacT (lactose utilisation) (Figure 1.3), Mlc (carbon metabolism), ManR (mannose/glucose utilisation), MtlR (mannitol utilisation), SacY (sucrose utilisation) and YesS (pectin/ rhamnogalacturonan utilization) (Gosalbes *et al.*, 1999, Henstra *et al.*, 2000, Tortosa *et al.*, 2001, Poncet *et al.*, 2009, Joyet *et al.*, 2013, Rothe *et al.*, 2013, Wenzel and Altenbuchner, 2013, Heravi and Altenbuchner, 2014). Transcription factors and antiterminators under the regulatory influence of PTS components are usually fused to at least one of four domains: (a) PTS regulatory domain (PRD), (b) EI-like domain, (c) EIIA-like domain and (d) EIIB-like domain (Greenberg *et al.*, 2002). These domains contain conserved histidine or cysteine residues that are subjected to phosphorylation by the PTS subunits leading to either activation or inactivation of the protein and consequently either an increase or decrease in the expression of the regulated genes (Henstra *et al.*, 2000, Xue and Miller, 2007, Joyet *et al.*, 2010, Rothe *et al.*, 2012).

A specific example of inducer exclusion regulation is provided by the transcriptional antiterminator LacT (Figure 1.3). The *Lactobacillus casei* LacT is phosphoregulated by the

lactose-PTS subsequently regulating the transcription of the lactose operon (Alpert and Siebers, 1997, Gosalbes *et al.*, 1997, Gosalbes *et al.*, 1999). LacT contains two conserved PRD domains, PRD I and PRD II, of which the PRD I domain has been implicated in the lactose induced exclusion (Gosalbes *et al.*, 2002). The current theory is that increased lactose transport through the lactose-PTS results in decreased phosphorylation of the EII subunits, which in turn decreases the phosphorylation of the LacT PRD I domain to promote antitermination activity, resulting in increased transcription of the *lac* operon. Genes under the control of transcription factors and antiterminators that are regulated by these phosphorylation reactions are usually directly involved in carbon metabolism however there are cases being discovered where they are involved in other processes such as bacterial pathogenesis (Poolman *et al.*, 1989, Hung and Miller, 2009, Hondorp *et al.*, 2013, Venditti *et al.*, 2013).

A central component of the CCR is phosphorylated HPr, however in this context the phosphate group does not originate from EI and PEP, nor is it located on His15. Instead HPr is phosphorylated on serine 46 (Ser46) by HPr-kinase/phosphotase (HprK/P), using ATP as the phosphate donor (Deutscher and Saier, 1983, Reizer *et al.*, 1998), which is allosterically activated by fructose-1,6-bisphosphate, a glycolytic intermediate (Dossonnet *et al.*, 2000). Phosphorylation of Ser46 promotes HPr-protein interactions, the most prominent being to carbon catabolite protein A (CcpA); this interaction of HPr and CcpA promotes the binding of CcpA to a palindromic sequence termed the 'catabolite responsive element' (CRE) located in the promoters and open reading frames of genes regulated by this protein (Deutscher *et al.*, 1995, Fujita *et al.*, 1995, Kim *et al.*, 1995, Schumacher *et al.*, 2004). This binding results in the transcriptional activation (CCA) or transcriptional repression (CCR) of the regulated genes. It has been suggested that the transcriptional effect of HPr-CcpA-CRE binding is dependent on the location of the CRE relative to other elements in the promoter region such as the transcription start site, -10 and -35 boxes (Marciniak *et al.*, 2012). Genes regulated by CcpA were generally thought to be involved in carbon metabolism however it is being discovered that CcpA regulates diverse biological processes that are not always related to carbon metabolism, including arginine biosynthesis, biofilm formation and bacterial pathogenesis (Deutscher *et al.*, 2005, Sadykov *et al.*, 2011, Nuxoll *et al.*, 2012, Zheng *et al.*, 2012, Leyn *et al.*, 2013).

It is this HPr component of the PTS that interlinks CCA, CCR and inducer exclusion. It does this in response to the levels of PTS sugars and their metabolites to regulate the carbon metabolism at small scale, by regulating individual enzyme such as GlpK, at a larger scale, by changing transcription of a small subset of genes, and at a global scale though CcpA to ensure efficient utilisation of available carbon sources. Literature reviews solely dedicated to the

understanding of the regulatory functions of canonical PTS have been published by Deutscher *et al.* (2014), Joyet *et al.* (2013), Deutscher *et al.* (2006) and Stulke and Hillen (2000).

1.3.4 The PTSs of *Lactobacillus plantarum*

Lactobacillus plantarum is a lactic acid bacteria (LAB) belonging to the *Firmicutes* (Felis and Dellaglio, 2007). They are found in a wide range of habitats such as dairy products, meat products, vegetable products and the human intestinal tract (Ahrne *et al.*, 1998, Gardner *et al.*, 2001, Ercolini *et al.*, 2003, Aquilanti *et al.*, 2007). Exposure to a range of carbohydrates in these different habitats drove them to evolve a wide array of metabolic pathways for the utilisation of various sugars (Siezen and van Hylckama Vlieg, 2011, O'Donnell *et al.*, 2013). Carbohydrate uptake is the initial step in these pathways and within the *L. plantarum* genome a large number of complete PTSs have been identified that display a diverse specificity for their substrates (Kleerebezem *et al.*, 2003, Siezen and van Hylckama Vlieg, 2011).

GlcNAc utilisation in *L. gasseri* ATCC 33323 was shown to be PTS dependent. In the study by Francl *et al.* (2010) the EI knockout strain of *L. gasseri* ATCC 33323 could not utilise N-acetylglucosamine along with the 14 other sugars: "D-galactose, D-fructose, D-mannose," "amygdalin, arbutin, esculin ferric citrate, salicin, D-cellobiose, D-lactose, D-saccharose , D-trehalose, amidon , gentiobiose and D-tagatose" as carbon sources leaving only glucose and malstose metabolically utilisable. This provides evidence that in some *Lactobacillus* species, PTSs are solely responsible for GlcNAc import.

Within the genome of *L. plantarum* WCFSI (GenBank accession number NC_004567.2), two genes, *pts18CBA* (GenBank accession number lp_2531) and *pts22CBA* (GenBank accession number lp_2969), are predicted to be GlcNAc-specific PTS EIICBA subunits belonging to the PTS Glc family (TCDB: 4.A.1). Homologs of both *pts18CBA*, *nagE* (GenBank accession number HMPREF0531_11400), and *pts22CBA*, *nagE* (GenBank accession number HMPREF0531_12045), are present in the genome of another glycocin F-susceptible strain *L. plantarum* subsp. *plantarum* ATCC 14917 (GenBank accession number GCF_000143745.1). The naming convention of these two *nagE* genes, *pts18CBA* and *pts22CBA*, from *L. plantarum* WCFS1 will be used from here on in.

1.4 Glycocin F

Glycocin F is a diglycosylated 43-amino acid bacteriocin produced by the Gram-positive bacterium *Lactobacillus plantarum* KW30 isolated from fermenting corn (Kelly *et al.*, 1996, Stepper *et al.*, 2011). It has a bacteriostatic effect on a broad range of bacteria which include some *Enterococcus* species, *Streptococcus* species, *Bacillus* species and *Lactobacillus* species with *L. plantarum* strains suspected to be its natural target (Kelly *et al.*, 1996, Stepper *et al.*,

2011, Kerr, 2013). The structure of glycocin F was solved by Venugopal *et al.* (2011) using NMR techniques. It was shown to contain two nested disulfide bonds flanking two α-helices connected by a short loop (Figure 1.6), resulting in a hair-pin like conformation with the GlcNAc moieties located at the poles of the 3D structure. The loop region is constrained by the two α-helices while the C-terminal 'tail', from glycine 30 onwards is very flexible. The glycosylation pattern of glycocin F is unique, like sublancin 168 and thurandacin (see section 1.2.2), as one of its two GlcNAc moieties is linked through the Sγ atom of cysteine 43 with the other *O*-linked to serine 18 (Ser18) (Venugopal *et al.*, 2011). The sugar moieties have been shown by Stepper *et al.* (2011) to be necessary for bacteriostatic activity. Enzymatic removal of the GlcNAc linked to Ser18 rendered the bacteriocin completely inactive, while the chymotryptic removal of the His42:Cys43-GlcNAc increased the IC_{50} (concentration causing 50 % growth inhibition) from 2 nM to 130 nM leaving residual bacteriostatic activity. This showed that while both sugars are necessary for bacteriostatic activity, the GlcNAc linked to Ser18 is absolutely vital. The different properties of the *S*-glycosidic bond, compared to canonical *O*- and *N*-linked modifications, may play an important role in the MOA of glycocin F which has yet to be elucidated.

A vital step in determining the MOA of glycocin F is the identification of its cognate receptor. Kerr (2013) showed that the inhibitory growth effect of glycocin F on *Enterococcus faecalis*

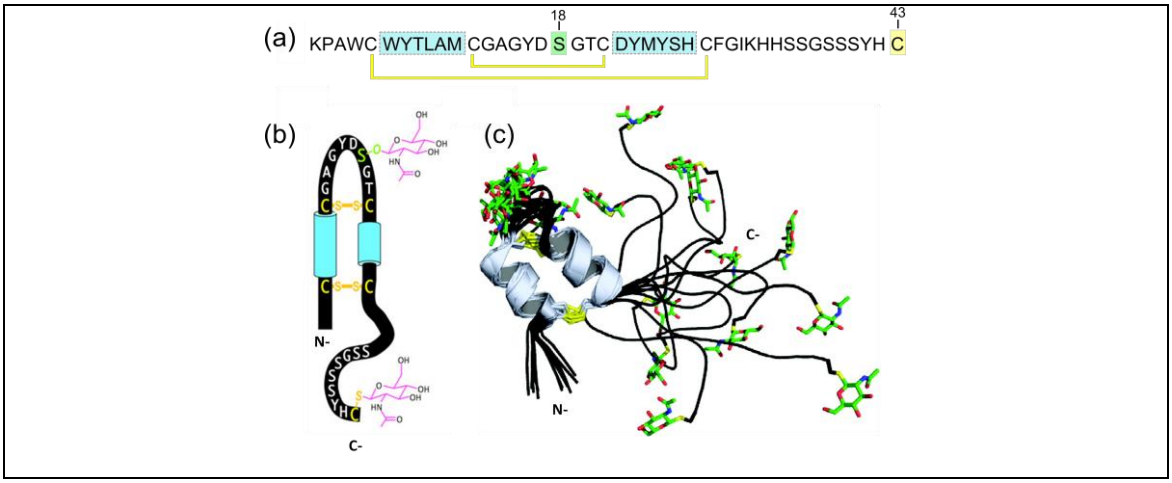


Figure 1.6 Structure of glycocin F

(a) Primary amino acid sequence and (b) secondary structure of glycocin F. Two nested disulfides (Cys5-Cys28, Cys12-Cys21) flank the two α-helices (cyan), and serine 18 (green) and cysteine 43 (yellow) side chains are both glycosylated with GlcNAc (pink). The N- and C-terminus are indicated. c) A superposition of 12 glycocin F conformers, that have the lowest energy, as determined by NMR. The N- and C-terminus are indicated and the 'tail' region can be seen to be very flexible. Helices (blue) are between the two nested disulfide bonds (yellow) with the two GlcNAc moieties (green) attached to serine 18 and cysteine 43. Adapted from Venugopal *et al.* (2011).

V583 could be reversed by treating the bacterial culture with GlcNAc, this protective effect of free GlcNAc had previously been shown for *L. plantarum* ATCC 8014 (Stepper *et al.*, 2011). Interestingly, sensitisation to glycocin F could also be accomplished by treating *E. faecalis* V583 cells with GlcNAc prior to treatment with the bacteriocin. Kerr (2013) suggested that this sensitisation could be a result of an up-regulation of the glycocin F receptor. The results by Stepper *et al.* (2011) and Kerr (2013) strongly suggested that a GlcNAc-specific transporter was targeted by glycocin F and that the sugar moieties of glycocin F are likely to be vital in its interaction with the transporter. As PTS transporters had been shown to be involved in the action of microcin E492 (1.2.5), lactococcin A and lactococcin B (1.2.4), it is not unreasonable that a GlcNAc-specific PTS transporter could be the receptor for glycocin F (Personal communication,Patchett and Norris, 2012). Evidence for this was provided by the identification of a mutation in the ribosome binding site (RBS) of a GlcNAc-specific PTS transporter, annotated as *pts18CBA*, from the genome sequencing of randomly generated glycocin F-resistant *L. plantarum* ATCC 8014 mutants (Personal communication, Patchett and Norris, 2012). The effect of this RBS mutation on the expression of *pts18CBA* was not carried out, though it was suggested that the mutation could result in a decreased translation of PTS18CBA. This was supported following glycocin F inhibitor assays of *E. coli* Δ*nagE* (*E. coli* Keio Knockout Collection; National Institute of Genetics, Japan) showed decreased sensitivity to glycocin F (Personal communication, Patchett and Norris, 2012). This identification of the *pts18CBA* gene product as a possible receptor coincides with both the enzymatic sugar removal studies by Stepper *et al.* (2011) and experiments by Kerr (2013). However the *E. coli* Δ*nagE* results were complicated, as pre-treatment with polymyxin was required to make the outer membrane permeable to glycocin F. Thus increased resistance to glycocin F in a *pts18CBA* deletion in the genome of a species naturally sensitive to glycocin F, e.g. *Lactobacillus plantarum*, would provide definitive evidence that glycocin F targets the GlcNAc-specific PTS transporter.

An efficient method to investigate if a specific gene product is a bacteriocin receptor is by targeted gene knockout as utilised by Diep *et al.* (2007). A Cre-Lox double crossover integration system was developed by Lambert *et al.* (2007) specifically for *Lactobacillus* species and has been shown to effectively knock out multiple genes in a single genetic background in *L. plantarum*. Using genetic knockout systems such as this coupled with complementation of knocked out genes using *Lactobacillus* protein expression vectors such as those developed by Sørvig *et al.* (2005), could provide evidence GlcNAc-PTS transporters are receptors for Glycocin F.

1.5 Aims of the study

The first aim of this project is to identity target cell proteins required for glycocin F-induced growth inhibition, namely its receptor, by using two approaches. The first approach involves comparative analysis of the genomes of glycocin F resistant *L. plantarum* ATCC 8014 and *L. plantarum* subsp. *plantarum* ATCC 14917 mutants. Mutations seen in glycocin F-resistant mutants should help to identify proteins that are required for sensitivity to glycocin F. This approach has the following objectives:

1. To obtain and isolate glycocin F-resistant *L. plantarum* subsp. *plantarum* ATCC 14917 and *L. plantarum* ATCC 8014 mutants (3.1).
2. To characterise the degree of resistance for the mutants and wild-type strains to glycocin F (3.2).
3. To sequence the genomes of the mutants and wild-type strains (3.3).
4. To assemble, annotate, and comparatively analyse the assembled genomes and identify genetic elements responsible for glycocin F resistance (3.3).

The second approach used to achieve the first aim is to test the hypothesis that the GlcNAc-PTS transporter, encoded by *pts18CBA*, is a receptor for glycocin F and that knocking out this gene will confer glycocin F resistance. A second GlcNAc PTS transporter in the *Lactobacillus* genome, *pts22CBA*, was not targeted due to time constraints. This approach had the following objectives:

1. To construct *pts18CBA* knockout plasmids (3.5 and 3.6).
2. To transform chemically competent *L. plantarum* strains with the knockout plasmids, isolate cell lines that are suspected to have undergone a homologous recombination event with the knockout plasmid and confirm the recombination event (3.7).
3. To characterise the degree of resistance for the *pts18CBA* knockout *L. plantarum* mutant to glycocin F (3.8).

The second aim of this project was to further characterise the relationship between glycocin F and *pts18CBA*, if it was confirmed to be a receptor for glycocin F in the first aim. This aim had the following objectives:

1. To construct a PTS18CBA production plasmid (3.9.3).
2. To produce recombinant PTS18CBA and confirm its identify by mass spectrometry (3.9.4).
3. To purify recombinant PTS18CBA to homogeny (3.9.4).
4. To perform biochemical assays to characterise interactions(s) between glycocin F and purified recombinant PTS18CBA (5).

2. Materials and Methods

2. Materials and Methods

2.1 General materials and methods

2.1.1 Water source

Sterile deionised water, obtained from a Barnstead™ Nanopure™ system (Thermo Scientific; Wilmington, DE, USA), was used to make up buffers and in experiments involving genetic manipulation. Tap water was used for bacterial growth media.

2.1.2 Bacterial growth media

2.1.2.1 Luria-Bertani medium

25 g of Luria broth base (Invitrogen™; Wilmington, DE, USA) was made up to 1 L with tap water and sterilised by autoclaving (121°C for 20 minutes).

2.1.2.2 Luria-Bertani medium agar plates

25 g of Luria broth base and 15 g 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 Luria agar was equilibrated to 50°C in a water bath and then poured into sterile Petri dishes.

2.1.2.3 de Man, Rogosa & Sharpe broth

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.

2.1.2.4 de Man, Rogosa & Sharpe agar plates

55 g of Lactobacilli MRS broth and 15 g of Agar Bacteriological was made up to 1 L with tap water and sterilised by autoclaving. The molten MRS agar was equilibrated to 50°C in a water bath and then poured into sterile Petri dishes.

2.1.3 MRS agar plates embedded with bacterial cells

1.375 g of Lactobacilli MRS broth and 0.225 g of Agar Bacteriological was made up to 24 mL with tap water in 50 mL tubes. These were sterilised by autoclaving and then equilibrated to 40°C in a water bath. 1 mL of a *L. plantarum* overnight culture was added to the 24 mL molten

MRS agar to obtain an OD_{600nm} of 0.01; the tubes were quickly capped and the cells evenly distributed through the solution by inverting several times. The molten MRS agar was then poured into a sterile Petri dish.

2.1.4 Sterilisation of media and buffers

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

2.1.5 Antibiotics

Antibiotics were added to media from stock solutions (Table 2.1) at concentrations that were dependent on the nature of the experiment. Unless otherwise indicated, the final concentrations of antibiotics used are listed in Table 2.2.

Table 2.1 Stock antibiotic concentrations

Antibiotic	Concentration (mg mL^{-1})	Solvent
Ampicillin	100	H_2O
Chloramphenicol	34	96 % (v/v) Ethanol
Erythromycin	20	96 % (v/v) Ethanol

Table 2.2 Typical final antibiotic concentrations used in media

Antibiotic	Antibiotic Concentration (µg mL^{-1})	
	Bacterial Species	
	L. plantarum	*E. coli*
Ampicillin	100	100
Chloramphenicol	15	15
Erythromycin	10-30	150

2.1.6 Storage and revival of bacteria

Bacterial stocks were made from overnight cell cultures that were made by adding sterile glycerol solutions to a final concentration of 10 % (v/v) storing at -80°C.

2.1.7 Growth conditions

2.1.7.1 *L. plantarum* overnight culture

A 10 mL MRS broth was inoculated with a scraping of bacterial cells from either a frozen *L. plantarum* glycerol stock or an MRS agar plate and incubated at 30°C with gentle shaking overnight or until stationary phase was reached.

2.1.7.2 *E. coli* overnight culture

A 10 mL LB culture was inoculated with a scraping of bacterial cells from either a frozen *E. coli* glycerol stock or an LB agar plate and incubated at 37°C with shaking overnight or until stationary phase was reached.

2.1.8 Optical density measurement of cell cultures

The optical density of bacterial cultures at 600 nm (OD_{600nm}) was routinely determined using a Bio-Rad Smartspec™ plus spectrophotometer (Bio-Rad Laboratories; USA) or a Cary 300 UV-Visible spectrophotometer (Agilent Technologies Inc; Santa Clara, CA, USA), after dilution in sterile growth media to an absorbance of less than 0.1-0.4 if required.

2.1.9 Plasmid purification

Plasmid-containing cells from *E. coli* overnight cultures (section 2.1.7) were harvested by centrifugation at 4,000 x g for 10 minutes at 4°C. Plasmids were purified using a Roche® High Pure Plasmid Purification Kit (Roche®; Basel, Switzerland) according to the manufacturer's instructions. For reasons described in section 3.7.3, the wash buffer was replaced with a solution of EMSURE® 80 % ethanol.

2.1.10 Drop dialysis

The osmolarity of plasmid solutions was reduced by drop dialysis. A piece of 0.025 μm 'V' series membrane (Millipore®; Billerica, USA) was floated shiny side up on 40 mL of Milli-Q water. The plasmid solution (0.05-0.1 mL) was carefully pipetted on to the top of the membrane and incubated for 3 hours at room temperature. After this time the plasmid solution was recovered and quantified using the NanoDrop® (section 2.1.14).

2.1.11 General PCR

PCRs for the purposes of colony screening and reactions not requiring high fidelity amplification were carried out using Roche® Taq DNA Polymerase in 12.5-50 µL reactions. Two master mix solutions separating the Taq polymerase and DNA components were made up (Table 2.3) and then combined in a 1:1 ratio prior to being placed in the thermocycler, which is an adaptation of the manufacturer's instructions. PCR reactions were run in 0.2 mL PCR Tubes (Axygen™; Pittsburgh, PA, USA) using a TGradient 96-Gradient Thermocycler (Biometra; Goettingen, Germany) with the temperature profile shown in Table 2.4. Primer annealing was done at ± 3°C of the melting temperature (Tm) of the longest primer as calculated using the nearest-neighbour two-state model utilised by OligoAnalyzer v3.1 (https://sg.idtdna.com /analyzer/Applications/OligoAnalyzer). The elongation temperature was either 72°C or 68°C for PCR products respectively shorter or longer than 1-kbp. DNA template concentration varied depending whether it was purified genomic DNA (gDNA) (2 ng μL^{-1}), purified PCR product (0.2 ng μL^{-1}) or purified plasmid (0.2 ng μL^{-1}). Colony screening PCR of whole cells involved taking a scraping of cells with a sterile loop and suspending them in 5 µL of Milli-Q H_2O, which was then substituted into the PCR reaction mixture. Extension times were calculated at 60 seconds per kbp of PCR product.

Table 2.3 General PCR component concentrations

Master Mix 1	
Component	Concentration
dNTPs	0.4 mM
Forward Primer	0.64 µM
Reverse Primer	0.64 µM
DNA template	2x of that specified above
Total Volume*	25 µL
Master Mix 2	
Component	Concentration
PCR reaction buffer	2 X
Taq Polymerase	0.05 U μL^{-1}
Total Volume*	25 µL
* Total volume of each master mix was made up to 25 µL with Milli-Q water	

Table 2.4 PCR temperature gradient profile

Step	Temperature (°C)	Time	Cycles
Initial Denaturation	94	2 minutes	1
Denaturation	94	30 seconds	
Annealing	Tm ± 3	30 seconds	27
Elongation*	68 or 72	60 seconds per kbp	
Final Elongation*	68 or 72	7 minutes	1
Storage	10	Indefinite	
* 68°C for PCR products >1 kbp, 72°C for PCR products ≤1 kbp			

2.1.12 High fidelity PCR

PCRs for the purposes of cloning or subsequent sequencing were performed using iProof™ High-Fidelity DNA Polymerase (Bio-Rad Laboratories; USA) in 12.5-50 µL reactions. A single master mix solution was made up as listed in Table 2.5. Reactions were done in 0.2 mL PCR Tubes in a TGradient 96-Gradient Thermocycler using the temperature profile shown in Table 2.6 adapted from the manufacturer's instructions. The annealing temperature was set at the highest Tm of the primers if they were shorter than 20 nucleotides or 3°C higher than the Tm of the longest primer if either of the primers were longer than 20 nucleotides. DNA template concentration varied depending on whether it was purified gDNA (5 ng μL^{-1}), purified PCR product (0.2 ng μL^{-1}) or purified plasmid (0.2 ng μL^{-1}). Extension times were calculated at 30 seconds per kbp of PCR product.

Table 2.5 High fidelity PCR component concentrations

Component	Final Concentration
5x iProof HF buffer	1x
dNTPs	0.2 mM
Forward Primer	0.32 µM
Reverse Primer	0.32 µM
DNA Template	Specified above
iProof DNA Polymerase	0.02 U μL^{-1}

Table 2.6 High fidelity PCR temperature gradient profile

Step	Temperature (°C)	Time (seconds)	Cycles
Initial Denaturation	98	30	1
Denaturation	98	10	
Annealing	Specified above	30	31
Elongation	72	30 seconds per kbp	
Final Elongation	72	420	1
Storage	10	Indefinite	

2.1.13 Agarose gel electrophoresis

2.1.13.1 Agarose gel preparation

HyAgarose™ (HydraGene Co, Ltd) agarose powder was dissolved in Tris-Base (10.8 g L^{-1}) Boric acid (5.5 g L^{-1}) ethylenediaminetetraacetic acid (EDTA) (2 mM, pH 8.0) buffer (TBE) at 0.6 % (w/v)-2.5 % (w/v), dependent on the length of the DNA as indicated in Table 2.7. This solution was heated in a microwave until the agarose had dissolved, allowed to cool slightly and then poured into a Bio-Rad gel tray in a gel caster (Bio-Rad Laboratories; USA) and left to set.

Table 2.7 Agarose % for DNA gel electrophoresis

DNA	Agarose % (w/v)
gDNA	0.6-1
5-kbp-10-kbp	0.7
1-kbp-5-kbp	1.5
0.05-kbp-1-kbp	2.5

2.1.13.2 DNA sample preparation

For analysis by electrophoresis, DNA samples were mixed with loading buffer (5 volumes of sample and 1 volume of 6x loading buffer). The 6x loading buffer consisted of 0.4 g mL^{-1} sucrose and an unspecified concentration of orange G dye in Milli-Q H_2O.

2.1.13.3 Electrophoresis

Agarose gels were equilibrated in TBE buffer at 4°C for at least 30 minutes in either a Mini-Sub® Cell GT Tank (Bio-Rad Laboratories; USA) or a Wide Mini-Sub® Cell GT Tank (Bio-Rad Laboratories; USA). DNA samples to be analyzed were pipetted into wells alongside 100 ng of

Invitrogen™ 1 kb Plus DNA Ladder containing bromophenol blue (0.05 % w/v) and an unspecified concentration of orange G. Loaded gels were electrophoresed at 1 V cm^{-1}-10 V cm^{-1} for 30 minutes-14 hours. Electrophoresis experiments that ran longer than 3 hours also had anode to cathode buffer recirculation.

2.1.13.4 Ethidium bromide staining

Agarose gels were soaked in water containing 2 µg mL^{-1} ethidium bromide (USB; Ohio, USA) for 30 minutes. Following this the gel was washed in clean water to remove unbound ethidium bromide.

2.1.13.5 Visualisation

Ethidium bromide-stained agarose gels were exposed to UV light in a Molecular Imager Gel Doc™ XR+ System (Bio-Rad Laboratories; USA) and the fluorescence imaged and recorded using Image Lab™ software (Bio-Rad Laboratories; USA).

2.1.13.6 DNA purification

Purification of DNA from agarose gels was done using the QIAquick® Gel Extraction Kit (Qiagen; Venlo, Limburg, Netherlands) following a protocol adapted from the manufacturer's instructions. DNA bands of interest were cut from the agarose gel with a sterile scalpel blade then placed in a clean 1.5 mL centrifuge tube. The agarose gel slices were weighed, suspended in QIAquick® Gel Extraction Kit Buffer QG at a 3 to 1 (v/w) ratio, and incubated at 50°C with vortexing every 2-3 minutes, for 10 minutes or until the gel had completely dissolved. To this, 1 volume of isopropanol (Scharlau; Sentmenat, Spain) was added and mixed by inversion. The solution was then passed through a QIAquick® Gel Extraction Kit column by centrifugation in a microcentrifuge at room temperature for 1 minute at 16,000 x g. The flow through was discarded and the column again centrifuged to remove residual extraction buffer. 80 % (v/v) EMSURE® ethanol (750 µL) was added to the column, which was then incubated at room temperature for 5 minutes before being centrifuged at room temperature at 16,000 x g in a microcentrifuge for 1 minute. This wash step was repeated 3 times followed by a 1 minute centrifugation at 16,000 x g to remove residual ethanol. 30-50 µL of Tris-HCl (20 mM, pH 8.0) was added straight onto the column membrane and incubated at room temperature for 1 minute. The column was again centrifuged at 16,000 x g in a microcentrifuge and the solution of eluted DNA stored at -80°C.

2.1.13.7 Quantitative analysis

DNA solutions were quantified by an agarose gel electrophoresis comparison of the sample with known amounts of Invitrogen™ 1 kb Plus DNA Ladder, 5 samples from 50 ng to 1 µg. Gels were visualized as before (section 2.1.13.5) and the mass of the ladder that had a fluorescence intensity most similar to that of the DNA sample was used to estimate sample DNA concentration.

2.1.14 Spectrophotometric quantification of DNA

The concentration of DNA in aqueous solution was estimated using a NanoDrop® ND-1000 Spectrophotometer v3.7 (Thermo Scientific; Wilmington, DE, USA) according to the manufacturer's instructions.

2.1.15 Restriction endonuclease digestion

Restriction endonuclease digestion was accomplished using either New England Biolabs® (New England Biolabs®; MA, USA), Roche™ (Roche™; Mannheim, Germany), Fermentas™ (Fermentas™; Wilmington, DE, USA) or Invitrogen™ (Invitrogen™; Wilmington, DE, USA) restriction endonucleases. The appropriate restriction endonuclease was used with the appropriate manufacturer-supplied buffer for the appropriate time at the recommended temperatures and concentrations to achieve cleavage of plasmid or PCR product. Products of restriction digestion reactions were either analysed by agarose gel electrophoresis and when necessary purified as described above (section 2.1.13.6) or purified using a GeneJET Gel Extraction and DNA Cleanup Micro Kit (Thermo Fisher Scientific; Wilmington, DE, USA) according to the manufacturer's instructions.

2.1.16 DNA ligation reactions

Ligation of purified restriction endonuclease digested plasmid and either purified PCR products and purified restriction endonuclease-digested PCR products was catalysed by Quick-Stick Ligase (Bioline™; London) or Roche®T4 DNA Ligase (Roche®; Basel, Switzerland) according to the manufacturer's instructions.

2.1.17 Sequencing of plasmid and PCR products

Plasmid DNA and PCR products were sequenced at the Palmerston North Massey Genome Service (MGS) using a 3730 DNA Analyzer (Applied Biosystems® Inc; Germany). Primer concentration, template DNA concentrations and total reaction volume were specified by MGS.

2.2 Bacterial manipulation techniques

2.2.1 Generation of chemically-competent *E. coli*

All chemically-competent *E. coli* were generated following a protocol established in the laboratory. An *E. coli* overnight culture was used to inoculate a 50 mL LB broth to an OD_{600nm} of 0.05. This was incubated with shaking at 37°C to an OD_{600nm} of 0.80. The culture was incubated on ice for 15 minutes, then centrifuged at 2,700 x g for 15 minutes at 4°C. The supernatant was discarded and the cell pellet suspended in 16 mL of RF1 buffer (pre-chilled to 4°C) and incubated on ice for 15 minutes (Table 2.8). After centrifugation at 2,700 x g for 15 minutes at 4°C, the supernatant was discarded and the cell pellet suspended in 1.5 mL of RF2 buffer (pre-chilled to 4°C) (Table 2.8) and incubated on ice for 15 minutes. Aliquots of 75 µL were then snap frozen in liquid nitrogen and stored at -80°C until required.

Table 2.8 RF1 and RF2 buffers for *E. coli* chemical competence

RF1	
Chemical component	Concentration
$RbCl_2$	100 mM
$MnCl_2.H_2O$	50 mM
Potassium Acetate	30 mM
$CaCl_2.2H_2O$	10 mM
Glycerol	15 % (v/v)
Acetic acid (glacial) was used to adjust pH to 5.8	
RF2	
Chemical component	Concentration
MOPS, pH 6.8	10 mM
$RbCl_2$	10 mM
$CaCl_2.2H_2O$	75 mM
Glycerol	15 % (v/v)

2.2.2 Transformation of chemically-competent *E. coli*

An aliquot of competent cells was thawed on ice for 15 minutes and 50 µL transferred to a sterile 1.5 mL centrifuge tube. Immediately after transfer, 20 ng-100 ng of plasmid was added and then incubated for 30 minutes on ice. The mixture was then heat shocked by incubation at 42°C for 40 seconds followed immediately by a 5 minute incubation on ice. To the heat-shocked cells 650 µL of room temperature LB was added and the cells incubated at 37°C with

shaking for 1 hour. After incubation 50 µL-100 µL of the heat-shocked cells were spread on to LB agar plates containing appropriate antibiotics and incubated at 37°C overnight.

2.2.3 Preparation of electrocompetent *L. plantarum*

To allow a comparison of the methods used for generation of chemically competent *L. plantarum* and the methods used for the electroporation of these cells, these methodologies are separated into the two sections 2.2.3 and 2.2.4. Each subsection in section 2.2.3 has a corresponding subsection in 2.2.4. This section describes eight methods, and variations, used to prepare electrocompetent *L. plantarum*.

2.2.3.1 Preparation of electrocompetent cells, adapted from Palomino *et al.* (2010)

100 mL of MRS medium supplemented with 0.9 M NaCl was inoculated (1 % (v/v)) from a *L. plantarum* subsp. *plantarum* ATCC 14917 overnight MRS culture. This was incubated at 37°C with gentle shaking to an OD_{600nm} of 2.0, incubated on ice for 10 minutes, then centrifuged at 2,590 x g for 15 minutes at 4°C. The cell pellet was washed three times with 50 mL of sterile Milli-Q water (4°C) then resuspended in 1 mL of sterile Milli-Q water and 50 µL aliquots were snap frozen in liquid nitrogen and stored at -80°C until required. Details of the electroporation are provided in section 2.2.4.1.

2.2.3.2 Preparation of electrocompetent cells, adapted from Lambert *et al.* (2007)

100 mL of MRS medium supplemented with 1 % (w/v) glycine was inoculated (1 % (v/v)) from a *L. plantarum* subsp. *plantarum* ATCC 14917 overnight MRS culture. This was incubated at 30°C with gentle shaking to an OD_{600nm} of 1.2, incubated on ice for 10 minutes, then centrifuged at 4,000 x g for 10 minutes at 4°C. The resulting cell pellet was resuspended in 100 mL of 30 % (v/v) polyethylene glycol 1500 (PEG-1500) (4°C) then incubated for 10 minutes on ice and centrifuged as before. The cell pellet was resuspended in 1 mL of 30 % (v/v) PEG-1500 (4°C) and then divided into 40 µL aliquots that, if not used immediately, were snap frozen in liquid nitrogen and stored at -80°C until required. Details of the electroporation step are given in section 2.2.4.2.

2.2.3.3 Preparation of electrocompetent cells, adapted from Alegre *et al.* (2004) and Berthier *et al.* (1996)

100 mL of MRS medium supplemented with 10 g L^{-1} glucose was inoculated (1 % (v/v)) from a *L. plantarum* subsp. *plantarum* ATCC 14917 overnight MRS culture. This was incubated at 30°C with gentle shaking to an OD_{600nm} of 1.6, incubated on ice for 10 minutes, then centrifuged at 4,000 x g for 10 minutes at 4°C. The resulting cell pellet was washed twice with 10 mL of 10 mM $MgCl_2$ (4°C) and once with 10 mL of 0.5 M sucrose in 10 % (v/v) glycerol (4°C). The final cell pellet was resuspended in 3 mL of the sucrose-glycerol solution and divided into 100 µL aliquots that, if not used immediately, were snap frozen in liquid nitrogen and stored at -80°C. Details of the electroporation step are provided in section 2.2.4.3.

2.2.3.4 Preparation of electrocompetent cells, adapted from Jin *et al.* (2012)

An overnight culture of *L. plantarum* subsp. *plantarum* ATCC 14917 was used to inoculate three 10 mL MRS broths to an OD_{600nm} of 0.05. These were incubated with gentle shaking at 30°C to an OD_{600nm} of 0.2 and then incubated on ice for 10 minutes. Ampicillin was added to a final concentration of either 0.01 µg mL^{-1}, 0.05 µg mL^{-1} or 0.1 µg mL^{-1}. The cultures were then incubated at 30°C with gentle shaking to an OD_{600nm} of 1.5, incubated on ice for 10 minutes then centrifuged at 4,000 x g for 10 minutes at 4°C. Cell pellets were suspended and washed twice in 2.5 mL of Milli-Q water (4°C). The final cell pellet was resuspended in 1 mL of suspension buffer (0.5 M sucrose, 0.5 mM K_2HPO_4, 0.5 mM KH_2PO_4, 1 mM $MgCl_2$, pH 7.3) (4°C)and split into two 500 µL aliquots. One was set aside as a control while the other was mixed with lysozyme (Sigma; Auckland, New Zealand, Grade 1 from chicken egg white, EC: 3.2.1.17, Lot: 39F8213; L-6876) (500 µL of a 1200 U mL^{-1} solution in suspension buffer) and incubated for 20 minutes at 37°C. Both the lysozyme-treated cells and control aliquots were centrifuged and the cell pellets washed with 1.5 mL of suspension buffer before being suspended in 150 µL of suspension buffer (4°C) (the cell pellet from the 0.1 µg mL^{-1} ampicillin and lysozyme treatment was suspended in 100 µL of suspension buffer (4°C)). The cells were stored at 4°C for a maximum of 2 hours prior to electroporation. Details of the electroporation step are provided in section 2.2.4.4.

2.2.3.5 Preparation of electrocompetent cells, adapted from Aukrust and Blom (1992)

Overnight cultures of *L. plantarum* subsp. *plantarum* ATCC 14917 and *L. plantarum* WCFS1 were used to inoculate 35 mL MRS broths to an OD_{600nm} of 0.1. These were incubated at 30°C with gentle shaking to an OD_{600nm} of 0.5, incubated on ice for 10 minutes, then centrifuged at 4,000 x g for 15 minutes at 4°C. The cell pellet was washed with 35 mL of 1 mM $MgCl_2$ (4°C) then with 35 mL of 30 % (w/v) PEG 1500 (4°C). The final cell pellet was resuspended in 30 % (w/v) PEG 1500 (4°C) in a volume up to 1/100 of the initial culture volume and divided into 100 µL aliquots, then snap frozen in liquid nitrogen and stored at -80°C. Details of the electroporation step are provided in section 2.2.4.5.

2.2.3.6 Preparation of electrocompetent cells, adapted from Aukrust and Blom (1992) and Lambert *et al.* (2007)

Overnight cultures of *L. plantarum* subsp. *plantarum* ATCC 14917 and *L. plantarum* WCFS1 were used to inoculate 100 mL MRS broths supplemented with 1.2 % (w/v) glycine to an OD_{600nm} of 0.1. These were incubated at 30°C with gentle shaking to an OD_{600nm} of 0.4, incubated on ice for 10 minutes, then centrifuged at 4,000 x g for 15 minutes at 4°C. The cell pellet was washed with 20 mL of 10 mM $MgCl_2$ (4°C) followed by a wash with 4 mL of 30 % (w/v) PEG 1500 (4°C). The cell pellet was resuspended in 1 mL of 30 % (w/v) PEG 1500 (4°C), then divided into 50 µL or 100 µL aliquots that were used immediately or snap frozen in liquid nitrogen and stored at -80°C. Electroporation details are provided in 2.2.4.6. This method was then adapted multiple times with variations to particular variables which are detailed in sections 2.2.3.6.a-2.2.3.6.d, if a variable is not mentioned then it remained unchanged compared to this protocol. Each variation has its own electroporation conditions, detailed in sections 2.2.4.6.a-2.2.4.6.d.

2.2.3.6.a.1st variation

The 1.2 % (w/v) glycine concentration in the MRS media was changed to 1.25 % (w/v) and 2.5 % (w/v). The concentration of $MgCl_2$ in the wash buffer was changed to 1 mM and 100 mM as shown in Table 2.13. Electroporation details are provided in section 2.2.4.6.a.

2.2.3.6.b.2nd variation

Electroporation parameters varied as detailed in Table 2.14 (section 2.2.4.6.b).

2.2.3.6.c.3rd variation

L. plantarum WCFS1 was not used. The initial OD_{600nm} inoculum was reduced to 0.05 and the OD_{600nm} at harvest was reduced to 0.1. Details of the electroporation are provided in Table 2.15 (section 2.2.4.6.c).

2.2.3.6.d.4th variation

L. plantarum ATCC 8014 was used instead of *L. plantarum* WCFS1. Initial inoculum OD_{600nm} for both *L. plantarum* ATCC 8014 and *L. plantarum* subsp. *plantarum* ATCC 14917 was 0.05. Details of the electroporation are provided in Table 2.16 (section 2.2.4.6.d).

2.2.3.7 Preparation of electrocompetent cells, adapted from Papagianni *et al.* (2007)

Overnight cultures of *L. plantarum* subsp. *plantarum* ATCC 14917 and *L. plantarum* ATCC 8014 were used to inoculate 55 mL MRS broths supplemented with 1 % (w/v) glycine to an OD_{600nm} of 0.05. These were incubated at 30°C with gentle shaking to an OD_{600nm} of 0.1, incubated on ice for 10 minutes and then centrifuged at 10,000 x g for 6 minutes at 4°C. The cell pellet was resuspended in 30 mL of resuspension buffer (100 mM LiAc, 10 mM dithiothreitol (DTT), 0.6 M sucrose, 10 mM Tris-HCl, pH 7.8) (4°C) and incubated at room temperature for 30 minutes. After centrifugation as above, cells were washed twice with 3 mL of 30 % (w/v) PEG 1500 (4°C), once with 3 mL of 50 mM EDTA (pH 7.8) (4°C), then twice with 3 mL of 30 % (w/v) PEG 1500 (4°C). The cell pellet was re-suspended in 500 µL of 30 % (w/v) PEG 1500 then divided into 100 µL aliquots that were immediately used for electroporation. Electroporation details are given in section 2.2.4.7.

2.2.3.8 Preparation of electrocompetent cells, adapted from Jin *et al.* (2012), Aukrust and Blom (1992) and Lambert *et al.* (2007)

MRS cultures of *L. plantarum* subsp. *plantarum* ATCC 14917, *L. plantarum* ATCC 8014 and *L. plantarum* NC8, grown with gentle shaking at 30°C to an OD_{600nm} of 1-3, were used to inoculate 50 mL MRS broths containing 0.05 µg mL^{-1} ampicillin to an OD_{600nm} of 0.01. These were incubated at 30°C with gentle shaking to an OD_{600nm} of 0.2, incubated on ice for 10 minutes, then centrifuged at 5,000 x g for 15 minutes at 4°C. The cell pellet was resuspended in 20 mL of 10 mM $MgCl_2$ (4°C), centrifuged as above, then washed with 4 mL of 30 % (w/v) PEG 1500 (4°C). The cell pellet was resuspended in 200 µL of 30 % (w/v) PEG 1500 (4°C), divided into 100 µL aliquots and kept cold (4°C) to be used immediately for electroporation. Electroporation

details are given in section 2.2.4.8. This method was then adapted once with variations to particular variables which are given in section 2.2.3.8.a, if a variable is not mentioned to have changed then it remained unchanged when compared to this protocol. The variation also had its own electroporation conditions given in section 2.2.4.8.a.

2.2.3.8.a.1st variation

MRS cultures were grown to an OD_{600nm} of 3-6 and used to inoculate 100 mL MRS broths to an OD_{600nm} of 0.05. For *L. plantarum* ATCC 8014 only, ampicillin was increased to 0.1 µg mL^{-1}. Cultures were centrifuged at 6,000 x g for 15 minutes at 4°C, resuspended in 40 mL of 100 mM $MgCl_2$ (4°C), then washed with 8 mL of 30 % (w/v) PEG 1500 (4°C), the cell pellet weighed and resuspended in 30 % (w/v) PEG 1500 (4°C) at 5 mL per gram of wet cell pellet. This was divided into 50 µL-100 µL aliquots and kept cold (4°C) to be used immediately for electroporation. Electroporation details are given in section 2.2.4.8.a.

2.2.4 Transformation of electrocompetent *L. plantarum*

This section (2.2.4) describes eight methods, and variations, used to prepare electrocompetent *L. plantarum*. Each subsection here has an equivalent subsection in 2.2.3.

2.2.4.1 Transformation protocol adapted from Palomino *et al.* (2010)

A Bio-Rad 0.2 cm electroporation cuvette (Bio-Rad Laboratories; USA) was chilled on ice. To the front inner surface of individual cuvettes plasmid solutions were added as detailed in Table 2.9. Immediately prior to electroporation, 50 µL of thawed chemically competent cells (section 2.2.3.1) were added to the cuvette so the liquid flowed over the surface on which the plasmid had been placed. The cuvette was then put inside the Bio-Rad shocking chamber, attached to a Bio-Rad Pulse controller and a Bio-Rad Gene Pulser, and electroporated with a single pulse at 2.5 kV, 200 Ω and 25 µF. 950 µL of room temperature MRS broth was immediately added to the cuvette followed by 10 minutes incubation on ice. The solution was transferred to a clean 1.5 mL centrifuge tube, incubated at 30°C for 3 hours, then diluted 5-fold with sterile MRS broth and 50-100 µL spread on MRS agar plates containing 10 µg mL^{-1} chloramphenicol. The plates were incubated at 30°C for 48 -72 hours in a sealed candle jar with O_2 levels depleted using wax candles.

Table 2.9 Details of electroporation experiment 2.2.4.1

Experiment	Cuvette	*L. plantarum* strain	Cell volume (µL)	Plasmid	Volume of plasmid solution (µL)	Plasmid mass (µg)
1	1	14917	50	pNZ5319_F1 _F2_14917	1	0.05
	1	14917	50	pTRk669	4.4	1
	2	14917	50	pTRK669	2.2	5
2	3	14917	50	-		
	4	14917	50	pNZ5319_F1 _F2_14917	5.9	0.5

2.2.4.2 Transformation protocol adapted from Lambert *et al.* (2007)

Bio-Rad 0.2 cm electroporation cuvettes were chilled on ice. To one cuvette a 2 µL volume of 50 ng μL^{-1} plasmid solution was applied to the front inner surface of the cuvette. A cuvette with no plasmid was used as a negative control and in later experiments plasmids pTRK669 and pRV613, all of which can replicate in *L. plantarum*, were used as positive controls. Immediately prior to electroporation a fresh 40 µL cell aliquot (section 2.2.3.2) was flowed over the front inner surface of the cuvette. The cuvette was then placed inside the Bio-Rad shocking chamber and electroporated with a single pulse at 1.5 kV, 200 Ω and 25 µF. The electroporated cells were transferred to a clean 1.5 mL centrifuge tube, incubated on ice for 10 minutes, diluted 10 fold with room temperature MRS broth then incubated at 30°C for 2 hours. 100 µL of these cell cultures were then spread on to MRS agar plates, containing 10 µg mL^{-1} chloramphenicol (raised to 15 µg mL^{-1} in latter experiments). The plates were incubated at 30°C for 72 hours in a sealed candle jar with O_2 levels depleted using wax candles.

2.2.4.3 Transformation protocol adapted from Alegre *et al.* (2004) and Berthier *et al.* (1996)

Bio-Rad 0.2 cm electroporation cuvettes were chilled on ice. To the front inner surface of individual cuvettes a plasmid solution was added as described in Table 2.10. A cuvette with no plasmid was used as a negative control and plasmid pTRK669 was used as a positive control. Immediately prior to electroporation a fresh 100 µL cell aliquot (section 2.2.3.3) was flowed over the front inner surface of the cuvette. The cuvette was then placed inside the Bio-Rad shocking chamber and electroporated with a single pulse at 1.3 kV, 200 Ω and 25 µF. The electroporated cells were transferred to a fresh 1.5 mL centrifuge tube, incubated on ice for 10

minutes, diluted 10-fold with room temperature MRS broth, and then incubated at 30°C for 2 hours. 100 µL of these cell cultures were spread on to MRS agar plates, containing 10 µg mL^{-1} chloramphenicol. The plates were incubated at 30°C for 48 hours in a sealed candle jar with O_2 levels depleted using wax candles.

Table 2.10 Details of electroporation experiment 2.2.4.3

Cuvette	*L. plantarum* strain	Cell volume (µL)	Plasmid	Volume of plasmid solution (µL)	Plasmid mass (µg)
1	14917	100	pTRK669	4.4	1
2	14917	100	pTRK669	2.2	0.5
3	14917	100	-	0	0
4	14917	100	pNZ5319_F1_F2_14917	5.9	0.5

2.2.4.4 Transformation protocol adapted from Jin *et al.* (2012)

Bio-Rad 0.2 cm electroporation cuvettes were chilled on ice. 1 µg (4.4 µL) of pTRK669 plasmid was pipetted onto the front inner surface of the cuvettes. Immediately prior to electroporation 70 µL of the cell solutions (section 2.2.3.4) were washed over the front inner surface of individual cuvettes. A smaller volume (50 µL) of the resuspended cell pellet from the 0.1 µg mL^{-1} ampicillin and lysozyme treatment was added to a cuvette. The cuvettes were incubated on ice for 5 minutes and then placed inside a Bio-Rad shocking chamber and electroporated with a single pulse at 2 kV, 400 Ω and 25 µF. After electroporation the cuvettes were incubated at 4°C for 5 minutes. A 900 µL volume of MRS containing 2 % (w/v) sucrose was added to the cuvettes. The solutions were transferred to clean 1.5 mL centrifuge tubes and incubated for 2 hours at 30°C. 100 µL of these cell solutions were spread on to MRS agarose plates containing 10 µg mL^{-1} chloramphenicol and incubated at 30°C for 30 hours.

2.2.4.5 Transformation protocol adapted from Aukrust and Blom (1992)

Bio-Rad 0.2 cm electroporation cuvettes were chilled on ice. To the front inner surface of individual cuvettes a plasmid solution was added as detailed in Table 2.11. Immediately prior to electroporation 50 µL-100 µL of thawed chemically competent cells (section 2.2.3.5) was flowed over the inner surface containing the plasmid. The cuvette was then placed inside the Bio-Rad shocking chamber and electroporated with a single pulse at 1.5 kV, 400 Ω and 25 µF.

The electroporated cells were then diluted 10 fold with room temperature MRS containing 0.4 M sucrose and 1 mM $MgCl_2$. The solution was then transferred to a clean 1.5 mL centrifuge tube and incubated at 30°C for 2 hours. The solutions were spread on to MRS agar plates, containing 7 μg mL^{-1}-10 μg mL^{-1} chloramphenicol, in 50 μL-100 μL volumes and incubated at 30°C for 72 hours.

Table 2.11 Details of electroporation experiment 2.2.4.5

Cuvette	*L. plantarum* strain	Cell volume (μL)	Plasmid	Volume of plasmid solution (μL)	Plasmid mass (μg)
1	WCFS1	100	pTRK669	15	1.4
2	14917	100	pTRK669	15	1.4
3	WCFS1	100	pNZ5319_F1_F2_14917	15	1.3
4	14917	100	pNZ5319_F1_F2_14917	15	1.3
5	14917	100	pNZ5319_F1_F2_14917	15	1.3
6	14917	100	pNZ5319_F1_F2_14917	15	1.3
7	WCFS1	50	pTRK669	5	0.47
8	WCFS1	50	pTRK669	5	0.47
9	WCFS1	50	pNZ5319_F1_F2_14917	5	0.5
10	WCFS1	50	pNZ5319_F1_F2_14917	5	0.5
11	14917	50	pTRK669	5	0.47
12	14917	50	pTRK669	5	0.47
13	14917	50	pNZ5319_F1_F2_14917	5	0.5
14	14917	50	pNZ5319_F1_F2_14917	5	0.5

2.2.4.6 Transformation protocol adapted from Aukrust and Blom (1992) and Lambert *et al.* (2007)

Bio-Rad 0.2 cm electroporation cuvettes were chilled on ice. To the front inner surface of individual cuvettes a plasmid solution was added as detailed in Table 2.12. Immediately prior to electroporation 50-100 μL of freshly prepared chemically competent cells (section 2.2.3.6) was pipetted over the inner surface on which the plasmid solution had been placed. The electroporation cuvette was then placed inside the Bio-Rad shocking chamber and electroporated with a single pulse at 1.5 kV, 400 Ω and 25 μF. The electroporated cells were then diluted 10-fold with room temperature MRS broth. Each cell solution was transferred to a clean 1.5 mL centrifuge tube and incubated at 30°C for 2 hours. The solutions were diluted 2-fold with MRS broth, 100 μL spread on to MRS agar plates containing 7 μg mL^{-1} chloramphenicol and then incubated at 30°C for 48 hours. This method was then used for the

electroporation of the variations of section 2.2.3.6 and unless a variable it stated to have changed in the following sections (2.2.4.6.a-2.2.4.6.d), it remains as described here.

Table 2.12 Details of electroporation experiment 2.2.4.6

Cuvette	*L. plantarum* Strain	Volume of cells (µL)	Plasmid	Volume of plasmid solution (µL)	Plasmid mass (µg)
1	WCFS1	50	pTRK669	5	0.47
2	WCFS1	100	pNZ5319_F1_F2_14917	10	1
3	14917	100	pNZ5319_F1_F2_14917	10	1
4	14917	100	pNZ5319_F1_F2_14917	10	1
5	14917	100	pNZ5319_F1_F2_14917	10	1
6	14917	100	pNZ5319_F1_F2_14917	10	1

2.2.4.6.a.1st variation

MRS agar plates were incubated for an extended time of 72 hours at 30°C. Electroporation details are given in Table 2.13.

Table 2.13 Details of electroporation experiment 2.2.4.6.a

Cuvette*	Glycine % (w/v)	[$MgCl_2$] (mM)	*L. plantarum* strain	Plasmid
1	2.5	10	WCSF1	pTRK669
2	2.5	10	14917	pTRK669
3	1.25	1	WCSF1	pTRK669
4	1.25	10	WCSF1	pTRK669
5	1.25	100	WCSF1	pTRK669
6	1.25	1	14917	pTRK669
7	1.25	10	14917	pTRK669
8	1.25	100	14917	pTRK669
9	2.5	10	WCSF1	pNZ5319_F1_F2_14917
10	2.5	10	14917	pNZ5319_F1_F2_14917
11	1.25	1	WCSF1	pNZ5319_F1_F2_14917
12	1.25	10	WCSF1	pNZ5319_F1_F2_14917
13	1.25	100	WCSF1	pNZ5319_F1_F2_14917
14	1.25	1	14917	pNZ5319_F1_F2_14917
15	1.25	10	14917	pNZ5319_F1_F2_14917
16	1.25	100	14917	pNZ5319_F1_F2_14917

* Cuvettes 1 to 8 contained 50 µL of cells and 5 µL (0.47 µg) of plasmid DNA. Cuvettes 9 to 16 contained 100 µL of cells and 10 µL (1 µg) of plasmid DNA.

2.2.4.6.b.2nd variation

Aliquots of frozen chemical-competent cells were thawed on ice. During electroporation the voltages of 0.5 kV, 1 kV, 1.5 kV and 2 kV were trialled. The MRS recovery media, which was added to the electroporated cells, was supplemented with 0.4 M sucrose, 0.1 M $MgCl_2$, 0.5 % (w/v) β-D-fructose and 0.5 % (w/v) GlcNAc. The recovery incubation was extended to 3 hours at 30°C. Electroporation details are given in Table 2.14.

Table 2.14 Details of electroporation experiment 2.2.4.6.b

Cuvette	*L. plantarum* strain	Volume of cells (µL)	Plasmid	Volume of plasmid solution (µL)	Plasmid mass (µg)
1	WCFS1	50	pNZ5319_F1_F2_14917	10	1
2	14917	50	pNZ5319_F1_F2_14917	10	1
3	WCFS1	50	pNZ5319_F1_F2_14917	10	1
4	14917	50	pNZ5319_F1_F2_14917	10	1
5	WCFS1	50	pNZ5319_F1_F2_14917	10	1
6	14917	50	pNZ5319_F1_F2_14917	10	1

2.2.4.6.c.3rd variation

MRS recovery media contained 0.45 M sucrose and 0.1 M $MgCl_2$. Electroporation details are given in Table 2.15.

Table 2.15 Details of electroporation experiment 2.2.4.6.c

Cuvette	*L. plantarum* strain	Volume of cells (µL)	Plasmid	Volume of plasmid solution (µL)	Plasmid mass (µg)
1	14917	50	pNZ5319_F1_F2_14917	10	1
2	14917	50	pNZ5319_F1_F2_14917	10	1
3	14917	50	pNZ5319_F1_F2_14917	10	1
4	14917	50	pNZ5319_F1_F2_14917	10	1
5	14917	50	pNZ5319_F1_F2_14917	10	1

2.2.4.6.d.4th variation

Electroporation conditions were changed to 1.5 kV, 200 Ω, and 25 µF. The recovery incubation was extended to 3 hours at 30°C. MRS agarose plates contained either 10 µg mL^{-1} chloramphenicol or 20 µg mL^{-1} erythromycin. Details of the electroporation are provided in Table 2.16.

Table 2.16 Details of electroporation experiment 2.2.4.6.d

Cuvette	*L. plantarum* strain	Volume of cells (µL)	Plasmid	Volume of plasmid solution (µL)	Plasmid mass (µg)
1	14917	100	pNZ5319_F1_F2_14917_Emr-	10	0.77
2	14917	100	pNZ5319_F1_F2_14917_Emr-	10	0.77
3	8014	100	pNZ5319_F1_F2_8014	10	1.26
4	8014	100	pNZ5319_F1_F2_8014	10	1.26
5	14917	50	pTRK669	5	1.77
6	8014	50	pTRK669	5	1.77
7	14917	50	pRV613	5	0.50
8	8014	50	pRV613	5	0.50

2.2.4.7 Transformation protocol adapted from Papagianni *et al.* (2007)

Bio-Rad 0.2 cm electroporation cuvettes were chilled on ice. To the front inner surface of individual cuvettes a plasmid solution was added as detailed in Table 2.17. Immediately prior to electroporation 100 µL of freshly prepared chemically competent cells (section 2.2.3.7) was flowed over the inner surface on which the plasmid was placed. Cuvettes were then placed inside a Bio-Rad shocking chamber and electroporated with a single pulse at 2.0 kV, 200 Ω and 25 µF. The cell solutions were then diluted 3.7-fold with room temperature MRS broth containing 0.5 M sucrose. These cell solutions were then transferred to a clean 1.5 mL centrifuge tube and incubated at 30°C for 2 hours. The solutions were diluted 2-fold with MRS broth and 100 µL spread on to MRS agar plates, containing 10 µg mL^{-1} chloramphenicol or 20 µg mL^{-1} erythromycin, and incubated at 30°C for 72 hours.

Table 2.17 Details of electroporation experiment 2.2.4.7

Cuvette	*L. plantarum* strain	Volume of cells (µL)	Plasmid	Volume of plasmid solution (µL)	Plasmid mass (µg)
1	14917	100	pRV613	10	1
2	8014	100	pRV613	10	1
3	8014	100	pNZ5319_F1_F2_8014	10	1.26
4	8014	100	pNZ5319_F1_F2_8014	10	1.26
5	14917	100	pNZ5319_F1_F2_14917	10	1
6	14917	100	pNZ5319_F1_F2_14917_Emr-	10	0.77
7	14917	100	pNZ5319_F1_F2_14917_Emr-	10	0.77

2.2.4.8 Transformation protocol adapted Jin *et al.* (2012), Aukrust and Blom (1992) and Lambert *et al.* (2007)

Bio-Rad 0.2 cm electroporation cuvettes were chilled on ice. To the front inner surface of individual cuvettes a plasmid solution was added as detailed in Table 2.18. Immediately prior to electroporation 100 µL of freshly prepared chemically competent cells (section 2.2.3.8) was flowed over the inner surface on which the plasmid was placed. Cuvette were then placed inside the Bio-Rad shocking chamber and electroporated with a single pulse at 2 kV, 400 Ω and 25 µF. The cells solutions were then diluted 10-fold with room temperature MRS broth containing 0.4 M sucrose, 0.1 M $MgCl_2$ and 0.5 % (w/v) GlcNAc. The cell solutions were then transferred to clean 1.5 mL centrifuge tubes and incubated at 30°C for 2 hours. 100 µL of the cell solutions were spread on to MRS agar plates supplemented with 0.5 % (w/v) GlcNAc, containing 5 µg mL^{-1} chloramphenicol or 5 µg mL^{-1} ampicillin, and incubated at 30°C for 168 hours. The plates were checked every 24 hours and any isolated colonies were spread onto MRS agar containing the appropriate antibiotic.

Table 2.18 Details of electroporation experiment 2.2.4.8

Cuvette	*L. plantarum* strain	Volume of cells (µL)	Plasmid	Volume of plasmid solution (µL)	Plasmid mass (µg)
1	NC8	100	pNZ5319_F1_F2_8014	10	1.26
2	NC8	100	pNZ5319_F1_F2_14917_Emr-	10	0.77
3	14917	100	pNZ5319_F1_F2_14917_Emr-	10	0.77
4	14917	100	pNZ5319_F1_F2_14917	10	1
5	8014	100	pNZ5319_F1_F2_8014	10	1.26
6	8014	100	pNZ5319_F1_F2_8014	10	1.26

2.2.4.8.a.1st variation

The recovery incubation period was extended to 3 hours at 30°C. Electroporation details are given in Table 2.19.

Table 2.19 Details of electroporation experiment 2.2.4.8.a

Cuvette	*L. plantarum* strain	Volume of cells (µL)	Plasmid	Volume of plasmid solution (µL)	Plasmid mass (µg)
1	NC8	50	pTRK669	5	0.53
2	NC8	50	pRV613	5	0.37
3	NC8	100	pNZ5319_F1_F2_8014	10	2
4	NC8	100	pNZ5319_F1_F2_14917	10	1
5	14917	50	pTRK669	5	0.53
6	14917	50	pRV613	5	0.36
7	14917	100	pNZ5319_F1_F2_14917	10	1
8	14917	100	pNZ5319_F1_F2_14917_Emr-	10	0.77
9	8014	50	pTRK669	5	0.53
10	8014	50	pRV613	5	0.37
11	8014	100	pNZ5319_F1_F2_8014	10	2
12	8014	100	pNZ5319_F1_F2_8014	10	2

2.2.5 Cellular density and viable cell counts

MRS cultures of *L. plantarum* subsp. *plantarum* ATCC 14917, *L. plantarum* ATCC 8014 and *L. plantarum* NC8 were cooled on ice and diluted in ice cold sterile MRS to OD_{600nm} of 0.0345, 0.0919 and 0.1052, respectively. Following the placement of glass coverslips on both the upper and lower chambers of a sterile Boeco Neubauer (Boeco; Hamburg, Germany), diluted cell cultures were pipetted under a coverslip, that had been placed on top of the Neubauer counting chambers, until both the upper and lower chambers were sufficiently filled according to the manufacturer's instructions. The Neubauer was placed in an optical microscope and the cells counted at 100x magnification. The number of cells in 7 of the 0.0025 mm^2 squares were counted in each chamber. Mean and standard errors of the cells mL^{-1} values were calculated for each strain and the results provided in Appendix 1.

Cell viability was determined by counting colonies formed from log phase cells diluted to an OD_{600nm} of $7.5x10^{-6}$ in sterile ice-cold MRS broth. 25 µL of this cell solution was embedded into 25 mL of molten MRS agar, as described in (section 2.1.3), and then poured into a sterile Petri dishes which were incubated at 30°C for 48 hours. Colony numbers were counted and the colony forming unit per mL of bacterial cells calculated (CFU mL^{-1}) and the results provided in Appendix 1.

2.2.6 Isolation of *L. plantarum* mutants with resistance to glycocin F

L. plantarum subsp. *plantarum* ATCC 14917 or *L. plantarum* ATCC 8014 cell-embedded MRS agar plates (section 2.1.3) containing 100 nM glycocin F were made. These plates were made based on bacteriocin resistance frequencies ranging from 10^{-5} to 10^{-7} so that addition of approximately 2.0 x 10^7 susceptible cells would yield 2-200 glycocin F-resistant mutants colonies (Gravesen *et al.*, 2002, Loessner *et al.*, 2003, Blake *et al.*, 2011, Kjos *et al.*, 2011b, Liu *et al.*, 2011). From log phase cultures of *L. plantarum* ATCC 8014 and *L. plantarum* subsp. *plantarum* ATCC 14917, aliquots containing 1.61 x 10^7 and 2.0 x 10^7 CFU respectively were diluted into 20 mL of molten (40°C) MRS agar containing 100 nM glycocin F and incubated for 48 hours at 30°C. Surface colonies were streaked on to MRS agar plates that contained 100 nM glycocin F. These were re-streaked once more onto MRS agar plates that contained 100 nM glycocin F and incubated overnight at 30°C. Final isolated colonies were used to inoculate 10 mL of MRS broth containing 100 nM glycocin F which were incubated at 30°C overnight. From these cultures 1 mL was made up to 10 % (v/v) glycerol from sterile 50 % (v/v) glycerol solution, snap frozen in liquid nitrogen and stored at -80°C for future use. Vital details

pertaining to the identification of these mutants and the wildtype strains are provided in Table 3.2.

2.2.7 Antimicrobial assays

2.2.7.1 Glycocin F minimum inhibition concentration (MIC) determination

The Minimum Inhibition Concentration (MIC) is defined as the lowest concentration of glycocin F at which either complete or near complete growth inhibition is observed. Room temperature MRS broth was inoculated with *L. plantarum* to OD_{600nm} of 0.05 from overnight MRS cultures. The cells were evenly distributed through the MRS broth and then 0.999 mL aliquots were added to 12 ethanol-sterilised 2 mL cuvettes (LP Italiana; Milano, Italy). The cuvettes were then placed in 6x6 multicell peltier, in a Cary 300 UV-Visible spectrophotometer, that had been equilibrated to 30°C and blanked using sterile MRS. 1 µL of a glycocin F solution in Milli-Q H_2O was added to 11 of these cuvettes (final glycocin F concentrations: 1 nM, 2.5 nM, 5 nM, 7.5 nM, 10 nM, 25 nM, 50 nM, 75 nM, 100 nM, 0.5 µM and 1 µM). 1 µL of Milli-Q H_2O was added to the last cuvette as a positive control for growth. To evenly distribute the glycocin F, the cell cultures were gently stirred using a sterile pipette tip. The OD_{600nm} was recorded every 30 seconds for 600 minutes; sample time set to 0.2 seconds with an averaging time of 0.2 seconds, using Cary Win UV Kinetics Application v4.10(461) software.

2.2.7.2 Liquid culture glycocin F sensitivity assay

The sensitivity of naturally-selected glycocin F-resistant *L. plantarum* mutants (section 2.2.6) to glycocin F was assessed. MRS broth cultures were inoculated with the five mutants and the wild-type of one of the *L. plantarum* strains and grown overnight at 30°C with gentle shaking, incubated on ice for 1 hour, then their OD_{600nm} recorded (section 2.1.8). Two aliquots from each one of the six cultures were diluted, up to 1 mL of sterile MRS broth, in two ethanol-sterilised cuvettes to an OD_{600nm} of 0.05. One set of these cuvettes was treated with glycocin F, either 1 µM or the MIC as detailed in Table 3.3, while set of cuvettes were not treated with glycocin F. Immediately after glycocin F was added to the cuvettes they were placed in a Varian 6X6 multicell peltier in a Cary 300 UV-Visible spectrophotometer that had been equilibrated to 30°C and blanked using sterile MRS broth. The OD_{600nm} was recorded every 30 seconds for 600 minutes. Sample time was set to 0.2 seconds with an averaging time of 0.2 seconds, using Cary Win UV Kinetics Application v4.10(461) software.

The sensitivity of *L. plantarum* NC8 Δ*pts18CBA* to glycocin F was assayed as described above with the exceptions that the high and low glycocin F concentration assays were performed together in a single experiment with identical untreated wild-type controls.

2.2.7.3 Agar plate glycocin F sensitivity assay

An MRS agar indicator plate assay was used to determine glycocin F sensitivity in *L. plantarum* wild-type and mutant strains. Glycocin F (100 µM) was diluted (2-fold dilution series) in Milli-Q water to the concentrations in the range of 97 nM-100 µM (Figure 2.1). A 2.5 µL aliquot of each glycocin F dilution was spotted onto *L. plantarum*-embedded MRS agar plates (section 2.1.3), according to the template shown in Figure 2.1, and the samples allowed to absorb into the agar. Both wild-type and mutant *L. plantarum* subsp. *plantarum* ATCC 14917 and *L. plantarum* ATCC 8014 plates were incubated for 24 hours at 30°C while both wild-type and mutant *L. plantarum* NC8 plates were incubated for 12 hours at 30°C.

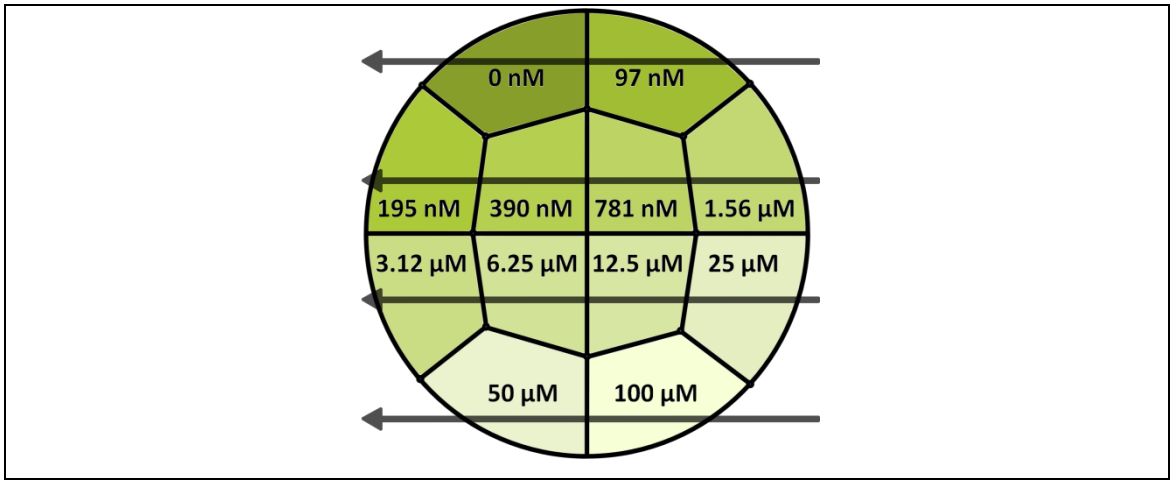


Figure 2.1 Layout of glycocin F spot plate assay

2.5 µL of glycocin F solution at the concentration indicated (2-fold dilution series) in each section of the plate was spotted onto MRS agar plates containing embedded wild-type and mutant *L. plantarum* strains. Colour shading and arrows indicate the direction the dilution series was applied.

2.2.8 Chemically defined minimal media agarose plates and carbon source utilisation assay

The chemical composition of all stock solutions are provided in Table 2.20. All stocks were filter-sterilised and all but the vitamin stock, which was frozen at -80°C prior to use, were frozen at -20°C.

Chemically defined media agarose plates containing a sole carbon source were made by mixing a series of stock solutions. To 31.9 mL of 0.1 g L^{-1} L-tyrosine the following solutions were added: 2.5 mL of 20x salt stock solution, 2.5 mL of 20 % (w/v) carbon source, 2.5 mL of 20x amino acid solution, 500 µL of a 100x trace element solution and 100 µL of a 500x vitamin stock solution. The solution containing all the individual components was equilibrated at 40°C before adding it to 10 mL of 5 % (w/v) HyAgarose™ agarose solution that had been equilibrated to 70°C following sterilisation by pressure cooking. The mixture was poured into two sterile Petri dishes and left at room temperature to solidify.

Chemically-defined media agar plates were partitioned into 8 equal-sized areas and to each area a loop of bacterial suspension was streaked. Bacterial suspensions were made as follows; 1 mL of an overnight culture was centrifuged in a microcentrifuge at 10,000 x g at room temperature for 1 minute. The cell pellet was washed once in 1 mL of sterile room temperature 0.1 g L^{-1} L-tyrosine solution and the final cell pellet suspended in 200 µL of 0.1 g L^{-1} L-tyrosine solution. The plates were left uncovered in a sterile laminar flow hood until all bacterial streaks had dried. The plates were incubated at 30°C for 24 hours then shifted and incubated at 25°C for 48 hours after which they were photographed at the Manawatu Microscopy and Imaging Centre. The results of these plates are provided in Appendix 2.

Table 2.20 Chemically-Defined Media Stock Solutions

Stock Solution	Chemical	Concentration		
		g L^{-1}	% (v/v)	% (w/v)
L-Tyrosine	L-Tyrosine	0.1		
Salt solution (20x)	KH_2PO_4 (anhydrous salt)	40		
pH, 6.4	K_2HPO_4 (anhydrous salt)	52		
	Ammonium citrate tribasic	40		
	Na acetate (anhydrous salt)	100		
Amino acid (20x)	L-tryptophan	2		
pH, ~4.5	L-alanine	2		
	L-aspartic acid (free acid)	2		
	glycine	2		
	L-isoleucine	2		
	L-leucine	2		
	L-methionine	2		
	L-threonine	2		
	L-valine	2		
	L-arginine	1.66		
	L-cysteine	5.5		

	L-glutamic acid (Na Salt)	4	
	L-histidine	1.5	
	L-lysine monohydrate (mono HCL)	1.8	
	L-phenylalanine	1	
	L-serine	3	
Trace element (100x)	$MgSO_4.7H_2O$	20	
	$CaCl_2.2H_2O$	0.2	
	$MnCl_2.4H_2O$	4	
	$FeSO_4.7H_2O$	2	
	$CoCl_2.6H_2O$	0.02	
	$ZnSO_4.7H_2O$	0.01	
	$NaMoO_4.2H_2O$	0.005	
	$NiCl_2.6H_2O$	0.0025	
	HCl (glacial)	1	
Vitamin (500x)	4-aminobenzoic acid (p-aminobenzoic acid)	0.05	
	biotin	0.025	
	vitamin B12	0.5	
	nicotinic acid	0.5	
	pantothenic acid	0.5	
	pyridoxal phosphate	0.65	
	pyridoxine.HCl	0.0125	
	riboflavin	0.5	
	thiamine	0.05	
Carbon source (20%)	D (+) fructose		20
	N-acetylglucosamine		20
	α-lactose		20
	Glucose		20

2.3 General DNA manipulation

2.3.1 Genomic DNA isolation

This genomic DNA (gDNA) isolation procedure was used for all *L. plantarum* strains. An overnight *L. plantarum* culture in MRS broth was diluted 100-fold into 35 mL of MRS, containing 1.2 % (w/v) glycine and incubated at 30°C. When the OD_{600nm} reached 1 the culture was chilled to 4°C and then centrifuged at 4,000 x g for 10 minutes at 4°C. The cell pellets were snap frozen in liquid nitrogen and stored at -80°C for later use.

A cell pellet from a 35 mL culture was defrosted on ice, resuspended in 2 mL of a 50 mM EDTA (pH 7.8) and 10 mM Tris-HCl (pH 8.0) solution (4°C) and then split into two equal volumes in 1.5 mL micro-centrifuge tubes. Lysozyme (Sigma; Auckland, New Zealand) was

added (400 µL of 20 mg mL^{-1} lysozyme in a solution containing 0.8 g L^{-1} $NaCl_2$, 0.02 g L^{-1} KCl, 0.024 g L^{-1} KH_2PO_4 and 0.268 g L^{-1} Na_2HPO_4) to each tube, then mixed by inversion and incubated at 37°C for 50 minutes with an inversion every 10 minutes. The mixtures were then centrifuged at 14,000 x g for 1 minute and the cell pellet resuspended in 800 µL of Promega® Wizard DNA purification kit Nuclei Lysis solution (Promega®; Madison, WI, USA). Cells were lysed by three cycles of freeze thawing; this involved submersing the tubes in liquid nitrogen for 1 minute then immediately transferring them to an 80°C waterbath for 2 minutes. The cells were cooled to 35°C and 5 µL of 5 mg mL^{-1} Promega® Wizard DNA purification kit RNase solution was added. This was incubated at 38°C for 45 minutes with inversions every 15 minutes, and then cooled to room temperature. 250 µL of Promega® Wizard DNA purification kit Protein Precipitation solution was added and the solution mixed by inversion then incubated on ice for 5 minutes. The solution was centrifuged at 14,000 x g for 3 minutes at room temperature and the supernatant collected and again centrifuged at 14,000 x g for 3 minutes. This step was repeated five times and each time the pellet was discarded. A 620 µL volume of isopropanol was carefully laid on top of the resulting supernatant and mixed by inversion. Precipitated gDNA was pelleted by centrifugation at 14,000 x g for 3 minutes, washed with 800 µL of 70 % (v/v) ethanol and centrifuged at 14,000 x g for 2 min and the supernatant discarded. The wash step was repeated once more before the final gDNA pellet was air dried and suspended in either 20 µL of Milli-Q water or 50 µL of 10 mM Tris-HCl (pH 8.0).

2.3.2 Genomic sequencing

Genome sequencing was performed at the Palmerston North MGS in collaboration with New Zealand Genomics Limited. *L. plantarum* subsp. *plantarum* ATCC 14917 and *L. plantarum* ATCC 8014 genomic samples were delivered to the MGS as 50 µL and 90 µL, respectively, in 10 mM Tris-HCl (pH 8.0), in accordance to sample preparation guidelines provided by the MGS. The MGS then performed the following quality control checks prior to library preparation: the Quant-iT™ dsDNA HS assay (Invitrogen™; Wilmington, DE, USA), Quant-iT™ RNA assay (Invitrogen™; Wilmington, DE, USA) and the Quant-iT™ Protein assay (Invitrogen™; Wilmington, DE, USA). Library preparation for sequencing was performed by the MGS using a TruSeq™ DNA Library preparation kit (Illumina®; San Diego, CA, USA) and subsequently sequenced using a MiSeq™ System (Illumina®; San Diego, CA, USA) to generate 250-bp paired-end reads. Prior to sequencing the library preparations were quality control checked using the DNA 1000 Labchip® (Agilent Technologies Inc; Santa Clara, CA, USA) coupled with the Agilent 2100 Bioanalyser (Agilent Technologies Inc; Santa Clara, CA, USA) and by performing the DNA, RNA and protein assays detailed above.

2.4 Genomic DNA assembly

2.4.1 Quality control

The raw sequenced reads were processed through an established quality control pipeline by the MGS: First the raw sequencing reads were mapped to the *Enterobacteria λ phiX174* sensu lato genome (GenBank accession number NC_001422) using the Burrows-Wheeler-Aligner software (BWA) (Li and Durbin, 2009), and any reads that mapped were removed from the BWA SAM output file. This SAM file was subsequently converted into a fastq file using the Picard suite 'SamToFastq.jar' tool (http://picard.sourceforge.net/index.shtml). The resultant fastq file was processed with the 'fastx-clipper' script (v1.1.2-318), from the ea-utils suite (https://code.google.com/p/ea-utils/) to remove contaminating TruSeq™ Universal Adapter (Illumina®; San Diego, CA, USA) sequences. The final 'forward and reverse fastq files' were subjected to further quality control before downstream *de novo* assembly.

From both the forward and reverse fastq files, reads were removed if they contained any ambiguities or were shorter than 150-bp using the 'AmbiguityFiltering' script from NGS QC Toolkit suite v2.3.2 (Patel and Jain, 2012) (http://59.163.192.90:8080/ngsqctoolkit/). The 'filtered forward and reverse fastq' files were used directly for *de novo* assembly via SPAdes (section 2.4.3.3). However, for *de novo* assembly using Velvet (section 2.4.3.2) they were further processed; the Fast Length Adjustment of SHort reads (FLASh) software v1.2.9 (Magoč and Salzberg, 2011) (http://ccb.jhu.edu/software/FLASH/) was used to combine the filtered forward and reverse fastq files. FLASh generates three output files; a fastq containing the combined reads and two fastq files containing forward and reverse paired-end reads that could not be combined. From these three fastq files any reads shorter than 203-bp in length were removed using the length filtering function from the 'AmbiguityFiltering' script from NGS QC toolkit. The resulting 'combined, forward and reverse fastq files' were used for *de novo* assembly via Velvet.

2.4.2 Genome size estimation

The size of the genomes was estimated using KmerGenie v1.6476 (Chikhi and Medvedev, 2014) (http://kmergenie.bx.psu.edu/) from the fastq files filtered for reads containing ambiguities or shorter than 150 nucleotides (section 2.4.1).

2.4.3 *De novo* assembly

2.4.3.1 Velvet optimiser

VelvetOptimiser v2.2.5 (Gladman and Seemann, 2008) (http://www.vicbioinformatics.com/) was used prior to Velvet (section 2.4.3.2) *de novo* assembly to determine the Velvet parameter values for: 'cov_cutoff' (coverage cut-off), 'hash_length' (kmer size), 'exp_cov' (expected coverage), 'ins_length' (average length of paired-end insert length) and ins_length_SD (standard deviation of the paired-end insert length). The optimisation functions in VelvetOptimiser were: kmer choice based on the N50, coverage cut-off based on the total number of bp in contigs larger than 1-kbp. Input files were the combined, forward and reverse fastq files generated during quality control (section 2.4.1) processing for *de novo* Velvet assembly. The parameter values output by VelvetOptimiser that resulted in the highest assembly score were then used for Velvet *de novo* assembly.

2.4.3.2 Velvet

De novo assembly of the sequenced *L. plantarum* wild-type strains and their glycocin F-resistant mutant strains were completed using the *de novo* assembler software Velvet v1.2.9 (Zerbino and Birney, 2008) (https://www.ebi.ac.uk/~zerbino/velvet/). Velvet consists of the two programs, velveth and velvetg. The combined, forward and reverse fastq files from the quality control processing (section 2.4.1) were inputs for velveth. The combined fastq file was input as a 'short' read category and the forward and reverse fastq files were input as the 'shortPaired' read category. All other parameters were left at default except for kmer size which for all assemblies was set at 201. The output files of velveth were then directly processed by velvetg. Default parameters for velvetg were used except for the following: cov_cutoff, exp_cov, ins_length, ins_length_SD, read_trkg (read tracking) and amos_file (AMOS file creation) with the latter two being set to true. Values for cov_cutoff, exp_cov, ins_length and ins_length_SD were determined using VelvetOptimiser (section 2.4.3.1) and are listed in Table 3.4 (page 76). Following successful assembly, Hawkeye software v2.0 (Schatz *et al.*, 2013), from the "A Modular, Open-Source whole genome assembler" (AMOS) software suite (http://sourceforge.net/apps/mediawiki/amos /index.php?title=AMOS), was used to generate a 'bank' folder from the AMOS '.afg' file. The consensus DNA sequence was then 'recalled' from the bank folder using AMOS 'recallConsensus' script. The recalled contigs were then extracted from the bank file using AMOS 'bank2fasta' script.

2.4.3.3 SPAdes

De novo assembly of the sequenced *L. plantarum* wild-type strains and their glycocin F mutant strains was completed using SPAdes software v3.0 (Anton Bankevich Sergey Nurk, 2012, Nurk *et al.*, 2013) (http://bioinf.spbau.ru/spades). The filtered forward and reverse fastq files generated during quality control processing (section 2.4.1) were used as input files. The SPAdes software was run with default settings except for the 'careful' parameter which was set 'true'. The SPAdes software automatically chose kmer sizes of 21, 33, 55, 77, 99 and 127 for all assemblies.

2.4.3.4 Edena

De novo assembly of the sequenced *L. plantarum* wild-type strains and their mutant strains was completed using the *de novo* assembler software Exact DE Novo Assembly v3.130110 (Edena) (Hernandez *et al.*, 2008) (http://www.genomic.ch/edena.php). For all the sequenced genomes, except for *L. plantarum* subsp. *plantarum* ATCC 14917 LP1_4**H**, an overlap graph was computed from the FLASh-generated combined fastq file only (section 2.4.1), which was input into Edena using the singleEnd function. Edena was run with default variables for both the overlap graph computation and second assembler stage.

An overlap graph for the sequenced genome of *L. plantarum* subsp. *plantarum* ATCC 14917 LP1_4**H** was computed from the files generated by the Sickle software v1.29 pairwise function (Joshi and Fass, 2011) (https://github.com/najoshi/sickle). Sickle was employed to filter out (i) reads shorter than 250-bp, (ii) those that contain ambiguities, (iii) those that fell below the phred quality threshold of 20 determined by the Sickle software (described by the software manual) and (iv) all singletons, i.e. reads lacking a corresponding paired read. Input files into Sickle were the forward and reverse fastq files (section 2.4.1) and the generated 'Sickle filtered forward and reverse fastq files' were used as input for Edena, using the paired function, and run with default settings for both the overlap graph computation and second assembler stage. This was performed due to the use of the FLASh-generated combined fastq file from this mutant in Edena resulted in a very poor assembly.

2.4.4 Contig integration using CISA

Contigs larger than 1000-bp from Velvet, SPADEs and Edena assemblies were integrated for each isolate using the Contig Integrator for Sequence Assembly software (CISA) v3.04 (Lin and Liao, 2013) (http://sb.nhri.org.tw/CISA/en/CISA). Minimum length was set to 1000 and Gap to

11 in the 'Merge.config' file. The genome size input required in the 'CISA.config' file was estimated using KmerGenie (2.4.2). The R2_Gap variable was left as a default of 0.95.

2.4.5 Annotation of contigs by Prokka

DNA contigs of the wild-type and mutant strains were annotated by Prokka v1.9, the 'Prokaryotic Genome Annotation System' developed by Seemann (2014) (http://www.vicbioinformatics.com/). Installed dependencies (required packages for optimal use of Prokka) were: Bioperl v1.6.901-3 (Stajich *et al.*, 2002) (http://www.bioperl.org/wiki/Main_Page), ARAGORN v1.2.36 (Laslett and Canback, 2004, Dean and Björn, 2008) (http://mbio-serv2.mbioekol.lu.se/ARAGORN/), RNAmmer v1.2 (Lagesen *et al.*, 2007) (http://www.cbs.dtu.dk/services/RNAmmer/), Prodigal v2.6 (Hyatt *et al.*, 2010) (http://prodigal.ornl.gov/), SignalP v4.1 (Petersen *et al.*, 2011) (http://www.cbs.dtu.dk/services/SignalP/), Blast+ v2.2.29 (Camacho *et al.*, 2009) (ftp://ftp.ncbi.nlm.nih.gov/blast/executables/blast+/LATEST/), HMMER v3.1b1 (Finn *et al.*, 2011) (http://hmmer.janelia.org/software) and Infernal v1.1 (Nawrocki and Eddy, 2013) (http://infernal.janelia.org/). Prokka was run using the parameters listed in Table 2.21 to comply with NCBI's GenBank submission requirements. The locustag prefix assigned to annotations was that provided by NCBI BioProject. Eight NCBI BioProjects were established for both wild-types and all mutants and the details listed in Table 3.2.

Table 2.21 Parameters used for Prokka annotations

Variable	Condition
compliant	True
center	MGS
locustag	Locus prefix tag assigned by NCBI GenBank
genus	Lactobacillus
species	plantarum
strain	'ATCC_8014' or 'subsp_plantarum_ATCC_14917'
gram	positive
usegenus	True
rfam	True
addgenes	True
mincontiglen	200

2.4.6 Assembly validation

Validation of the genome assemblies was performed by mapping the reads from the filtered forward and reverse fastq files back to the annotated contigs (section 2.4.5) using the 'Map To Reference' function in Geneious™ (Biomatters Ltd, Auckland, New Zealand) software v7.1.4 with low sensitivity settings.

2.4.7 Comparative analysis of glycocin F-resistant mutant genomes

To identify mutation in the genomes of the mutant strains, for each mutant strain the annotated contigs were aligned to the annotated contigs of the respective wild-type assembly. Alignments were performed using the 'Map To Reference' function in Geneious™ software v7.1.4 with high sensitivity settings and an extended maximum gap size of 10,000-bp.

2.5 General protein biochemical methods

2.5.1 Protein production of the recombinant GlcNAc-PTS transporter PTS18CBA

2.5.1.1 Small scale time course of expression

Expression trials were conducted in LB medium containing 100 µg mL^{-1} ampicillin with *E. coli* BL21 (DE3) cells transformed, as described in section 2.2.2, with plasmid pET-21b(+)_8014_His6_pts18CBA (section 3.9.3). 50 mL-100 mL of LB medium was inoculated to an OD_{600nm} of 0.1 from an *E. coli* BL21 (DE3) overnight culture (section 2.1.7.2) and incubated at 37°C with shaking (200 rpm) for 1.25 hours (OD_{600nm} of 0.68). Prior to induction with 1 mM isopropyl β-D-1-thiogalactopyranoside (IPTG) a 1 mL sample was taken and labelled as un-induced cells. Following induction the culture was incubated at 25°C with shaking for 12 hours, and 1 mL samples were taken at 30, 60, 90, 120, 180, 240 and 720 minutes. For each sample the OD_{600nm} was determined, followed by centrifugation for 1 minute at 16,000 x g in a microcentrifuge at room temperature. Cell pellets were resuspended in Milli-Q water to a total volume (µL) equal to 100x the 'OD_{600nm} x sample volume (mL)' to obtain an OD_{600nm} of 10 in all samples. The suspended cells were subjected to sonication (Misonix Sonicator® S-4000 Ultrasonic Processor, Misonix Inc, Farmingdale, NY, USA), 30 second bursts at 30 % amplitude with a 1 mm microtip, until the solutions were clear. Insoluble cellular debris and unlysed cells were pelleted by microcentrifugation as before. 4 volumes of each cell lysis supernatant was mixed with 1 volume of 5x SDS loading buffer (Table 2.24) and boiled for 10 minutes; the remaining supernatant was stored at -80°C until required. The boiled samples were

cooled on ice and centrifuged for 10 minutes at 16,000 x g in a microcentrifuge at room temperature, and the supernatants analysed by SDS PAGE (section 2.5.2).

2.5.1.2 Small scale protein solubility trial

1 mL of the chilled culture was taken as an induced cell sample and treated as described above (section 2.5.1.1) for SDS PAGE analysis. Cells were harvested by centrifugation at 5,000 x g for 15 minutes at 4°C. Cells were washed once with 25 mL of 150 mM NaCl in 20 mM Tris-HCl (pH 8.0), then resuspended in 12 mL of lysis buffer (Table 2.22). The suspension was passed four times through a French pressure cell (Aminco Instruments Co; USA) at 4 kPa using a Wabash hydraulic press (Wabash; USA). Deoxyribonucleate 5'- oligonucleotidohydrolase (DNase I) (Sigma; Auckland, New Zealand; #DN-25) was added to the cell lysate and incubated on ice for 15 minutes. Cellular debris and unlysed cells were pelleted by centrifugation at 15,000 x g for 30 minutes at 4°C. A sample of the cell lysis supernatant, containing soluble membrane fragments and cytoplasmic proteins, was diluted 2-fold in Milli-Q H_2O, then 4 volumes of this was mixed with 1 volume of 5x SDS loading buffer and heated as described in section 2.5.1.1 for SDS PAGE analysis. The cell lysis pellet, contains insoluble cellular debris, was resuspended in 12 mL of Milli-Q H_2O and a sample diluted 2-fold in Milli-Q H_2O was treated as described above for SDS PAGE analysis. The decanted cell lysis supernatant was subjected to ultra-centrifugation at 200,000 x g for 2 hours at 4°C. A sample of the supernatant, contains soluble cytoplasmic proteins, was taken and treated as described above for SDS PAGE analysis. The remaining supernatant was discarded and 1 mL of 20 mM Tris-HCl (pH 8.0), 100 mM NaCl and 20 % (v/v) glycerol, was laid on top the pellet, containing membrane fragments, and incubated with gentle agitation at 4°C for 12 hours. A sample of this resuspended membrane pellet was diluted 25-fold in Milli-Q H_2O, then treated as described above for SDS PAGE analysis.

Table 2.22 Lysis buffer

Component	Concentration
Tris-HCl, pH 8.0	20 mM
NaCl	300 mM
cOmplete® EDTA-free protease inhibitor*	1x
Phenylmethylsulfonyl fluoride (PMSF)	1 mM
*Roche®; Basel, Switzerland	

2.5.2 Polyacrylamide gel electrophoresis

Discontinuous SDS PAGE separates proteins in a cross-linked acrylamide matrix based on their molecular weight (Laemmli, 1970). Acrylamide gels were made with stacking and resolving gels that were 4 % (v/v) and 10 % (v/v) acrylamide respectively as described in Table 2.23. Prior to addition of *N,N,N′,N′*-tetramethylethylenediamine (TEMED) (Sigma; Auckland, New Zealand) the SDS PAGE gel solutions were 'degassed' for 30 minutes using a water vacuum pump. After addition of TEMED the SDS PAGE solutions were pipetted into a Bio-Rad PROTEAN II gel casting apparatus, first the resolving gel solution, which was allowed to polymerase, then the stacking gel solution. Gels were placed in a Bio-Rad PROTEAN II electrophoresis chamber containing SDS tank buffer (Table 2.25). Wells were loaded with protein samples and run alongside 6 µL of Benchmark™ Protein Ladder (Invitrogen™; Wilmington, DE, USA) at 180 V at room temperature for approximately 50 minutes or until the dye front was within 1 cm of the bottom of the gel. Gels were removed, rinsed with Milli-Q H_2O, suspended in Coomassie stain solution (0.1 % (w/v) Coomassie Brilliant Blue R-250, 40 % (v/v) methanol and 10 % (v/v) acetic acid) and heated in a microwave on high for 1 minute or until the stain solution started to boil. The gel was left to stain for at least 1 hour with gentle agitation at room temperature. The stain solution was decanted and the gel rinsed with Milli-Q H_2O. Excess Coomassie stain was removed by destaining the gel in Coomassie destain solution (40 % (v/v) methanol and 10 % (v/v) acetic acid). The gel was suspended in destain solution and heated in a microwave for 1 minute on high or until the destain solution started to boil. The gel was incubated in the destain solution until the background Coomassie stain was sufficiently decreased after which the gel was rinsed with Milli-Q H_2O and then suspended in Milli-Q H_2O overnight to remove residual Coomassie dye. The destained gel was imaged using the 'trans light source' in a Bio-Rad Molecular Imager Gel Doc™ XR+ System and recorded using Bio-Rad Image Lab™ software.

Table 2.23 SDS PAGE discontinuous gel mixtures

Component	Volume (mL)
Stacking Gel (5 mL)	
Milli-Q H_2O	3.2
Acrylamide:Bis 29.1:0.9, 40 % (v/v)	0.5
0.5 M Tris-HCl, pH 8.8	1.25
Ammonium persulfate, 10% (w/v)	0.025
SDS, 10 % (w/v)	0.05
TEMED	0.02
Resolving Gel (10 mL)	
Milli-Q H_2O	4.93
Acrylamide:Bis 29.1:0.9, 40 % (v/v)	2.5
1.5 M Tris-HCl, pH 6.8	2.5
Ammonium persulfate, 10% (w/v)	0.06
SDS, 10 % (w/v)	0.1
TEMED	0.013

Table 2.24 5x SDS loading buffer

Component	Concentration
Bromophenol Blue	0.25 % (w/v)
DTT	0.5 M
Glycerol	50 % (v/v)
SDS	10 % (w/v)
Tris-HCl	0.25 M
pH, 6.8	

Table 2.25 SDS tank buffer

Component	Concentration (g L^{-1})
Tris	3
Glycine	14.4
SDS	1

2.5.3 In gel tryptic digestion and mass spectrometry

The band on SDS PAGE gels suspected to be the rPTS18CBA protein was digested with trypsin and the resulting tryptic peptides analysed using mass spectrometry.

From a 10 % acrylamide SDS PAGE gel that had been thoroughly destained the band suspected to be rPTS18CBA protein was excised using a sterile scalpel blade and transferred to a MAXYMum Recovery™ 1.5 mL centrifuge tube (Axygen™; Pittsburgh, PA, USA). The excised band was cut into small pieces and further destained by washing in 300 µL of 20 % (v/v) methanol, 0.5 % (v/v) acetic acid, until the excised gel pieces were clear. The gel pieces were then dehydrated for 1 minute in 300 µL of 80 % (v/v) acetonitrile at room temperature. The gel pieces were centrifuged for 1 minute at top speed in a microcentrifuge at room temperature and the acetonitrile solution was carefully decanted; residual acetonitrile was removed using a centrifugal evaporator. The gel pieces were reduced in 30 µL of 10 mM DTT, 50 mM ammonium bicarbonate, for 1 hour at 37°C. The gel pieces were centrifuged as above and the reducing solution decanted. The gel pieces were suspended in 100 µL of 50 mM ammonium bicarbonate (pH >7.9) and incubated at room temperature for 5 minutes. The gel pieces were centrifuged as before and the solution decanted. The gel pieces were dehydrated in 100 µL of 80 % (v/v) acetonitrile for 1 minute. The gel pieces were centrifuged as before and the acetonitrile solution decanted; residual acetonitrile was removed by drying the gel pieces in a centrifugal evaporator. The gel pieces were then suspended in 30 µL of 20 mM iodoacetamide, 50 mM ammonium bicarbonate, and incubated for 20 minutes in the dark at room temperature. The gel pieces were centrifuged as before and the solution decanted. The gel pieces were suspended in 100 µL of 50 mM ammonium bicarbonate for 5 minutes, centrifuged as before and the solution decanted. The gel pieces were dehydrated in 100 µL of 80 % (v/v) acetonitrile for 1 minute, centrifuged as before and the acetonitrile decanted. This step was repeated and the residual acetonitrile solution removed by drying the gel pieces in a centrifugal evaporator. The gel pieces were rehydrated in 30 µL of trypsin digestion solution (20 ng uL^{-1} trypsin in 50 mM ammonium bicarbonate) and incubated on ice for 10 minutes. Excess solution was removed and 30 µL of 50 mM ammonium bicarbonate solution added. This was incubated overnight at 37°C.

Following the overnight digestion the sample was centrifuged and sonicated on high for 2 minutes in a waterbath sonicator (Soniclean Pty Ltd; Australia). The sample was centrifuged as before and the supernatant decanted and stored. The gel pieces were suspended in 60 µL of 5 % (v/v) formic acid, 50 % (v/v) acetonitrile, and sonicated as before. This was centrifuged as before and the supernatant decanted and stored. The gel pieces were then suspended in 60 µL

0.1 % (v/v) formic acid, 80 % (v/v) acetonitrile and sonicated as before. This was then centrifuged as before and the supernatant decanted and stored. This treatment was repeated to give a another supernatant fraction that was then stored. The stored fractions were pooled and reduced in volume to approximately 15 µL using a centrifugal evaporator. This was then transferred into a mass spectrometry (MS) glass sample vial (Thermo Fisher Scientific; Wilmington, DE, USA).

2 µL of the concentrated sample was subjected to electronspray ionisation (capillary voltage of 1,800 V) and MS/MS fragmentation (collision energy slope of 3.6 V/100 Da and offset of -4 C) in an Agilent 6520 Q-TOF mass spectrometer (Agilent Technologies, Germany). Total ion counts were processed by MassHunter Workstation Qualitative Analysis software vB.03.01 (Agilent Technologies, Santa Clara, CA, USA). Files exported from MassHunter were searched using the Walter and Eliza Hall Institute (Melbourne, Australia) based Mascot Server v2.4.1 (Matrix Science Ltd, London) with the parameters listed below in Table 2.26. Mascot results were filtered, based on individual ion scores, to identify peptides with identity or extensive homology ($p<0.05$); (Personal communication, Loo, 2014).

Table 2.26 Mascot search parameters

NCBI non-redundant protein database with a *Firmicutes* taxonomic restriction
Semi-tryptic digest with up to 2 miscleavages
Carbamidomethyl of cysteines as fixed modification
Peptide charge of +2
Decoy-enabled
Peptide and MS/MS tolerances of 25 and 75 parts per million

3. Results and Discussion

3. Results and Discussion

This chapter is divided into nine main sections. The first three sections describe the isolation (section 3.1) and comparative characterisation, both genomic (section 3.3) and phenotypic (section 3.2), of glycocin F- resistant mutants (Table 3.2) of *L. plantarum* ATCC 8014 and *L. plantarum* subsp. *plantarum* ATCC 14917. The fourth section (section 3.4) is an attempt to explain the major relevant mutations identified in section 3.3 by the use of *in silico* methods. The next four sections (sections 3.5-3.8) describe efforts to perform targeted gene knockout, a project that spanned 24 months and started 6 months prior the isolation of glycocin F-resistant mutants. The last section (section 3.9) describes attempts to produce recombinant PTS18CBA.

3.1 Natural selection and isolation of glycocin F-resistant mutants

3.1.1 Introduction

For antimicrobial compounds of biological origin it is possible to select for resistance using appropriate (culture) conditions. A simple yet powerful approach to identify genetic elements involved in glycocin F-induced bacteriostasis involves the isolation of *L. plantarum* strains resistant to glycocin F, followed by genome sequencing, assembly and comparative analysis to the genomes of wild-type parent *L. plantarum* strains that are susceptible to glycocin F. For this study *L. plantarum* ATCC 8014 and *L. plantarum* subsp. *plantarum* ATCC 14917 strains were used as they are the most sensitive to glycocin F (Personal communication Patchett and Norris, 2012).

3.1.2 Aim

- To obtain *L. plantarum* subsp. *plantarum* ATCC 14917 and *L. plantarum* ATCC 8014 mutants resistant to glycocin F.

3.1.3 Results and discussion

Apart from *L. plantarum* strains *L. plantarum* KW30 and *L. plantarum* A-1, that produce glycocin F-like bacteriocins (Hata *et al.*, 2010), naturally glycocin F-resistant *L. plantarum* strains may be relatively rare. The only strain identified so far is *L. plantarum* 965 (Personal communication, Patchett and Norris, 2012), isolated from cheddar cheese (Daeschel *et al.*, 1990). Interestingly, this strain is used as an indicator for the two component plantaricins PlnEF and PlnJK, presumably having lost its chromosomal copy of the corresponding immunity genes (Diep *et al.*, 1995). Plates were poured from molten MRS agar containing glycocin F and a

susceptible *L. plantarum* strain (section 2.2.6). The cells were embedded in the agar to achieve a more even dispersion to aid in colony counts and determination of the relative frequency. Because some glycocin F-resistant colonies developed embedded in agar, different colony morphologies were observed depending on the vertical position of the colony-forming cell (Figure 3.1). Cells that settled on the base of the plate grew wide, yet short and had a soft diffuse appearance, cells growing completely surrounded in MRS agar had an ellipsoid appearance and surface colonies were raised and circular in appearance. The colonies were counted and the frequency of resistance calculated (Table 3.1). Five surface colonies (Table 3.2) were picked and sub-cultured in the presence of 100 nM glycocin F to maintain selection (section 2.2.6).

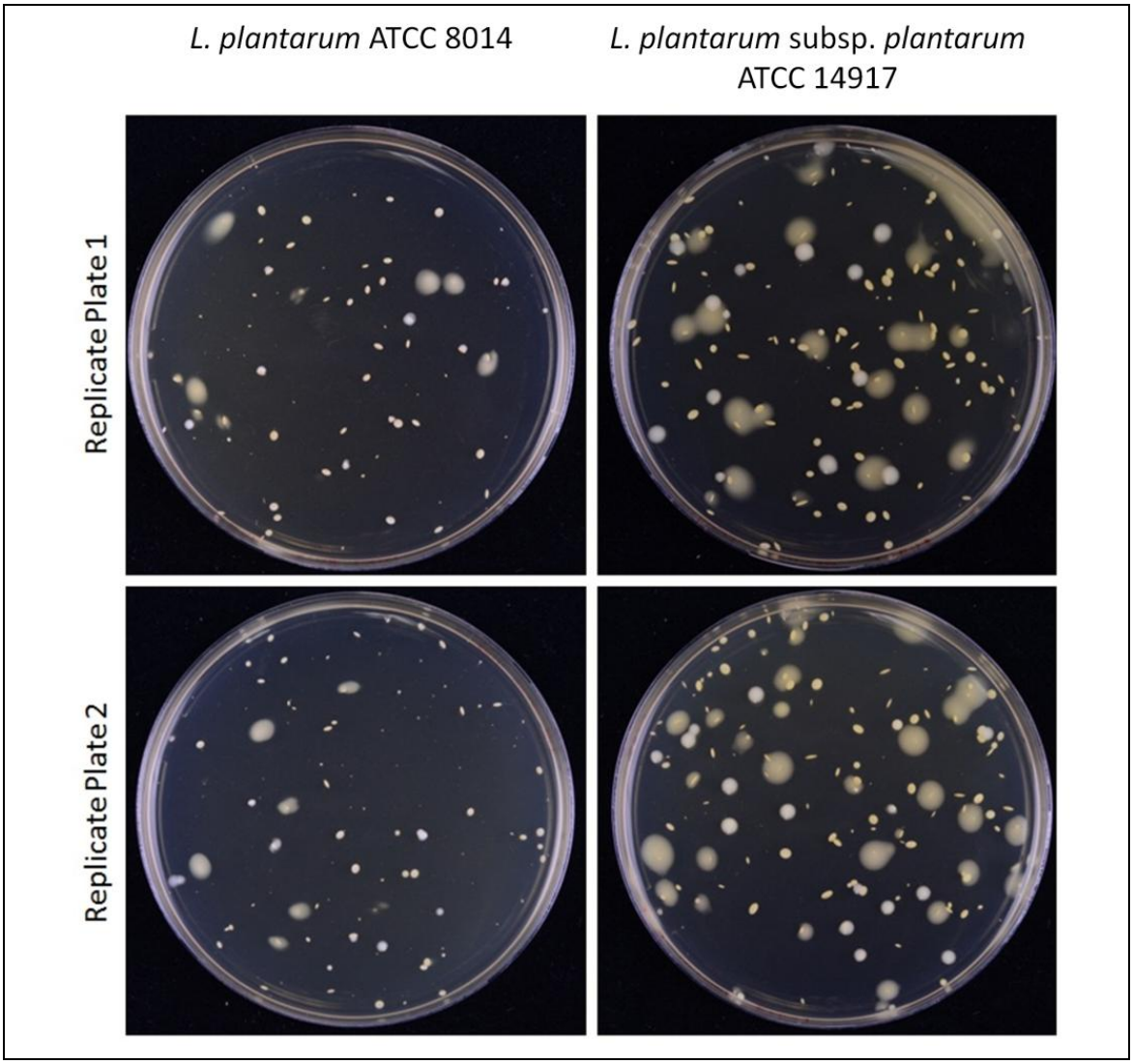


Figure 3.1 ***L. plantarum*** **colonies grown in the presence of 100 nM glycocin F**

Two replicate MRS agar plates containing 100 nM glycocin F and embedded cells of the *L. plantarum* strain indicated after incubation at 30°C for 48 hours.

The original aim (section 3.1.2) to obtain glycocin F-resistant mutants was achieved, as the cell lines isolated from the surface colonies were later shown to be highly or completely resistant to glycocin F (section 3.2). Based on the degree of glycocin F resistance, three mutants of each strain were subjected to genome sequencing (section 3.3). Interestingly, *L. plantarum* subsp. *plantarum* 14917, the more sensitive of the two wild-type strains (Figure 3.3a *cf.* Figure 3.3b), acquired resistance (average relative resistance frequency of $7.6x10^{-6}$) at almost double the frequency of *L. plantarum* ATCC 8014 (average relative resistance frequency of $4.0x10^{-6}$) (Figure 3.1). Later characterisation of the two strains (section 3.2) showed that the resistance acquired by the *L. plantarum* ATCC 8014 mutants was more complete compared to that of *L. plantarum* subsp. *plantarum* ATCC 14917 mutants. The implications of these observations are discussed in section 3.3.

Table 3.1 Glycocin F resistance colony counts and frequency

Lactobacillus plantarum	Replicate Plate	Colony Count	Relative Resistance Frequency
subsp. *plantarum* ATCC 14917	1	151	$7.4x10^{-6}$
	2	158	$7.8x10^{-6}$
ATCC 8014	1	66	$4.1x10^{-6}$
	2	63	$3.9x10^{-6}$

Table 3.2 Identification details of wildtype and mutants

Strain	**Original Colony**	**Mutant Identifier***	**Glycocin F Resistance Group**	**NCBI BioProject number**	**NCBI Locus Tag Prefix**
ATCC 8014	Wild-type	N/A	Indicator	247443	EV52
ATCC 8014	1	LP8_1**L**	Low	247442	EV51
ATCC 8014	2	LP8_2**H**	High	N/A	N/A
ATCC 8014	3	LP8_3**H**	High	N/A	N/A
ATCC 8014	4	LP8_4**H**	High	247441	EV50
ATCC 8014	5	LP8_5**H**	High	247440	EV49
ATCC 14917	Wild-type	N/A	Indicator	247439	EV48
ATCC 14917	1	LP1_1**H**	High	N/A	N/A
ATCC 14917	2	LP1_2**H**	High	247446	EV55
ATCC 14917	3	LP1_3**L**	Low	247445	EV54
ATCC 14917	4	LP1_4**H**	High	247444	EV53
ATCC 14917	5	LP1_5**L**	Low	N/A	N/A

N/A; not applicable, *Mutant identifier is explained in detail in section 3.2

3.2 Characterisation of glycocin F resistant mutants to glycocin F

3.2.1 Introduction

Prior to characterisation of the glycocin F-resistant *L. plantarum* mutants (Table 3.2) the glycocin F minimum inhibition concentration (MIC) (section 2.2.7.1) for wild-type *L. plantarum* subsp. *plantarum* ATCC 14917 and *L. plantarum* ATCC 8014 was determined. This was done so that the response of the glycocin F-resistant mutants to glycocin F could be assessed at the MIC. Characterisation of the glycocin F-resistant *L. plantarum* subsp. *plantarum* ATCC 14917 mutants and the glycocin F-resistant *L. plantarum* ATCC 8014 mutants to glycocin F was accomplished using liquid culture assays (section 2.2.7.2) and plate assays (section 2.2.7.3).

3.2.2 Aims

- To determine the glycocin F MIC for indicator strains *L. plantarum* subsp. *plantarum* ATCC 14917 and *L. plantarum* ATCC 8014.
- To characterise the effect of glycocin F on glycocin F-resistant *L. plantarum* subsp. *plantarum* ATCC 14917 and *L. plantarum* ATCC 8014 mutants.

3.2.3 Results and discussion

From triplicate liquid culture assays of glycocin F inhibition, the glycocin F MIC for both *L. plantarum* subsp. *plantarum* ATCC 14917 and *L. plantarum* ATCC 8014 was determined to be 25 nM (Figure 3.2). This MIC value was used as the lower concentration of glycocin F in the liquid culture assays, alongside the high 1 µM concentration (section 2.2.7.2).

Liquid culture glycocin F sensitivity assays of the glycocin F-resistant *L. plantarum* subsp. *plantarum* ATCC 14917 mutants showed that compared to the wild-type strain, the inhibition of growth by glycocin F, at both 25 nM and 1 µM, was reduced from an average of 95.7 % to an average of 50 % (Figure 3.4). At 25 nM glycocin F, resistance among the mutants was similar while at the higher concentration of 1 µM two resistance phenotypes were apparent; a high level resistance group (Group H) and a low level resistance group (Group L). High level resistance mutants included colonies 1 (LP1_1**H**), 2 (LP1_2**H**), and 4 (LP1_4**H**) and low level resistance mutants included colonies 3 (LP1_3**L**) and 5 (LP1_5**L**). The name in the parenthesis is the 'mutant identifier' and details the parent strain (*L. plantarum*, 14917 or 8014), original colony number (1-5) and relative resistance (**H**igh or **L**ow) and replaces the colony identification number as shown in Table 3.2.

Table 3.3 MIC of the three wild-type *L. plantarum*

L. plantarum	MIC (nM)
subsp. *plantarum* ATCC 14917	25
ATCC 8014	25
NC8	50

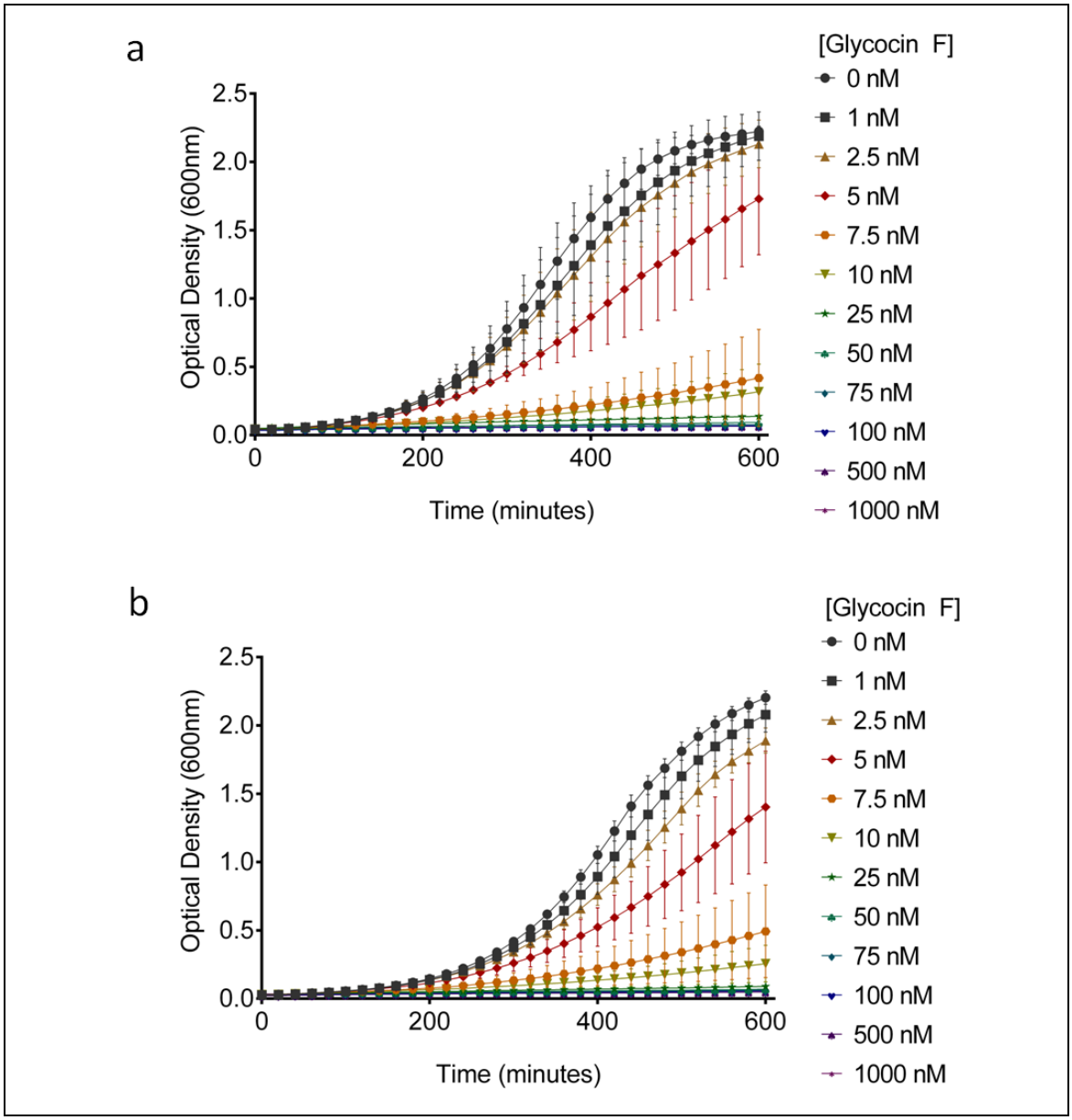


Figure 3.2 Liquid culture glycocin F assays of *L. plantarum* for MIC determination

Triplicate liquid culture assays of (a) *L. plantarum* subsp. *plantarum* ATCC 14917 and (b) *L. plantarum* ATCC 8014 inhibited by various concentrations of glycocin F. Vertical error bars represent one standard deviation from the mean. OD_{600nm} is plotted on the vertical axis and time in minutes is plotted on the horizontal axis. Legends to the right indicate glycocin F concentration added at time point 0 minutes.

After 400 minutes of growth (Figure 3.4) the differences in phenotype were not so clear but after 600 minutes of growth (Figure 3.5) the two groups became distinguishable when treated with 1 µM glycocin F. On average the OD_{600nm} at 600 minutes of those belonging to group H is 50 % higher than those belonging to group L (see inset scatter plot in Figure 3.5b). The glycocin F sensitivity plate assays (section 2.2.7.3) were consistent in showing that all mutants had acquired partial resistance to glycocin F (Figure 3.3a); the lowest glycocin F concentration that caused inhibition increased to 3.12 µM from 781 nM. The glycocin F sensitivity plate assays were consistent, showing mutant colonies 1 (LP1_1**H**), 2 (LP1_2**H**) and 4 (LP1_4**H**) have different resistance phenotypes to mutant colonies 3 (LP1_3**L**) and 5 (LP1_5**L**); clearing at 100 µM glycocin F was more complete for mutant colonies 3 (LP1_3**L**) and 5 (LP1_5**L**) than for mutant colonies 1 (LP1_1**H**), 2 (LP1_2**H**) and 4 (LP1_4**H**).

As with the glycocin F-resistant *L. plantarum* subsp. *plantarum* ATCC 14917 mutants, the glycocin F-resistant *L. plantarum* ATCC 8014 mutants displayed two resistance phenotypes. Glycocin F liquid culture sensitivity assays showed resistance across the mutants was fairly uniform at the MIC (Figure 3.6a and Figure 3.7a), but at 1 µM two phenotypes were obvious (Figure 3.6b and Figure 3.7b); mutant colonies 2 (LP8_2**H**), 3 (LP8_3**H**), 4 (LP8_4**H**) and 5 (LP8_5**H**) displayed a high level resistance phenotype (Group H) and mutant colony 1 (LP8_1**L**) displayed a low level resistance phenotype (Group L) (Inset scatter plot in Figure 3.7b). Remarkably the *L. plantarum* ATCC 8014 group H isolates seem to have acquired complete resistance to glycocin F. This was shown by the liquid culture assays where inhibition of the group H mutants at both 25 nM and 1 µM glycocin F, relative to the untreated controls, is not detectable (Figure 3.6 and Figure 3.7). It was again shown by the glycocin F sensitivity plate assays (Figure 3.3b), where inhibition of the wild-type strain begins at 1.56 µM while mutant colonies belonging to group H fail to exhibit bacterial clearing at all concentrations tested, the highest being 100 µM, providing strong evidence for complete resistance to glycocin F.

For *L. plantarum* subsp. *plantarum* ATCC 14917 it would appear that complete resistance was unattainable. This may be due to the manner in which complete resistance arises. The definitive reason for the two resistant phenotypes observed is unknown but can be potentially explained by the comparative analysis of the genome which is discussed in sections 3.3 and 3.4.

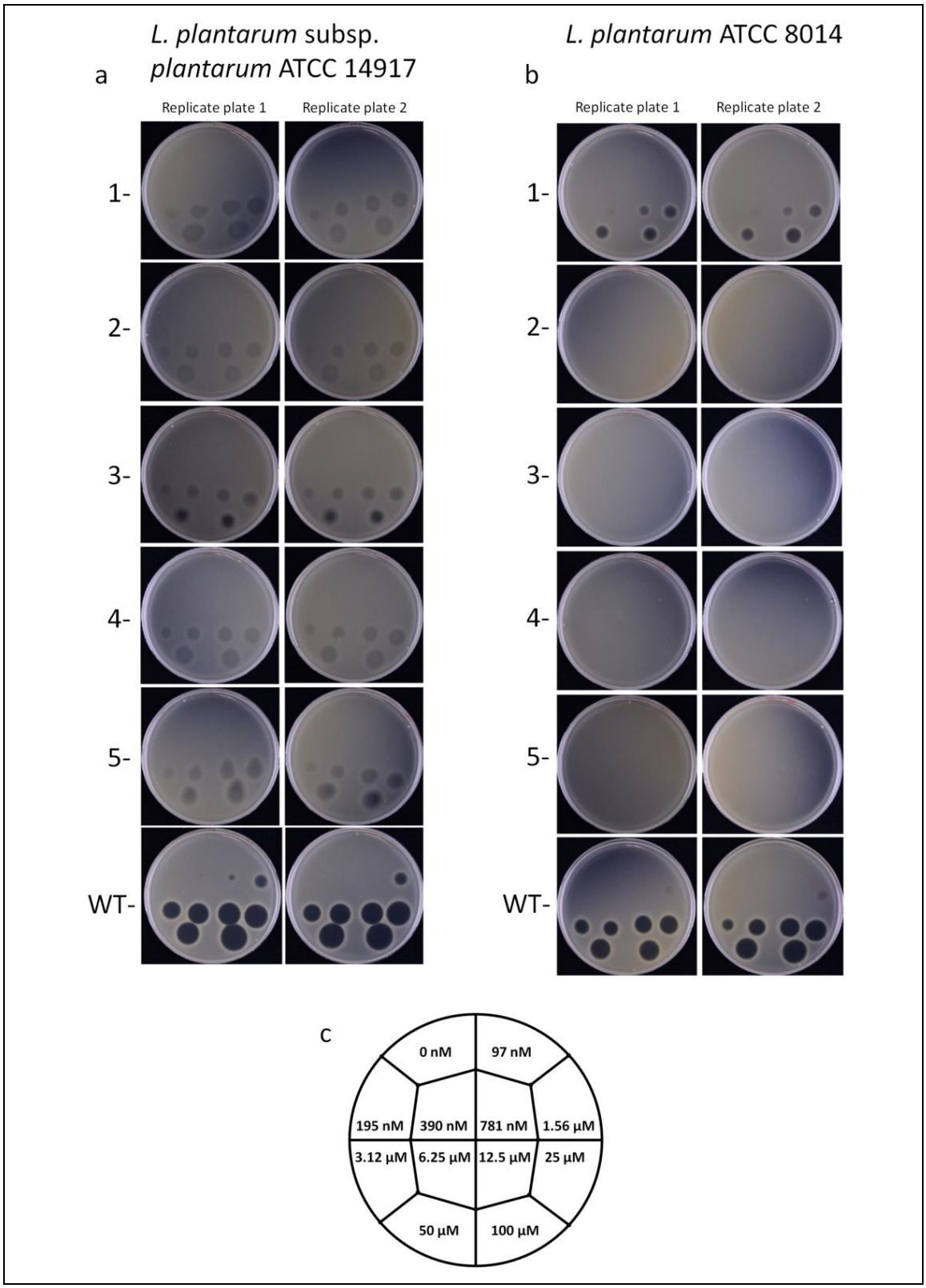


Figure 3.3 Glycocin F agar plate assays of glycocin F-resistant *L. plantarum* mutants

Wild-type (WT) and glycocin F-resistant *L. plantarum* mutants (1-5) were embedded in molten MRS agar, poured into two plates which were treated with 2.5 µL spots of glycocin F at the concentrations shown in (c). Label to the left of the plates indicate the mutant isolate.

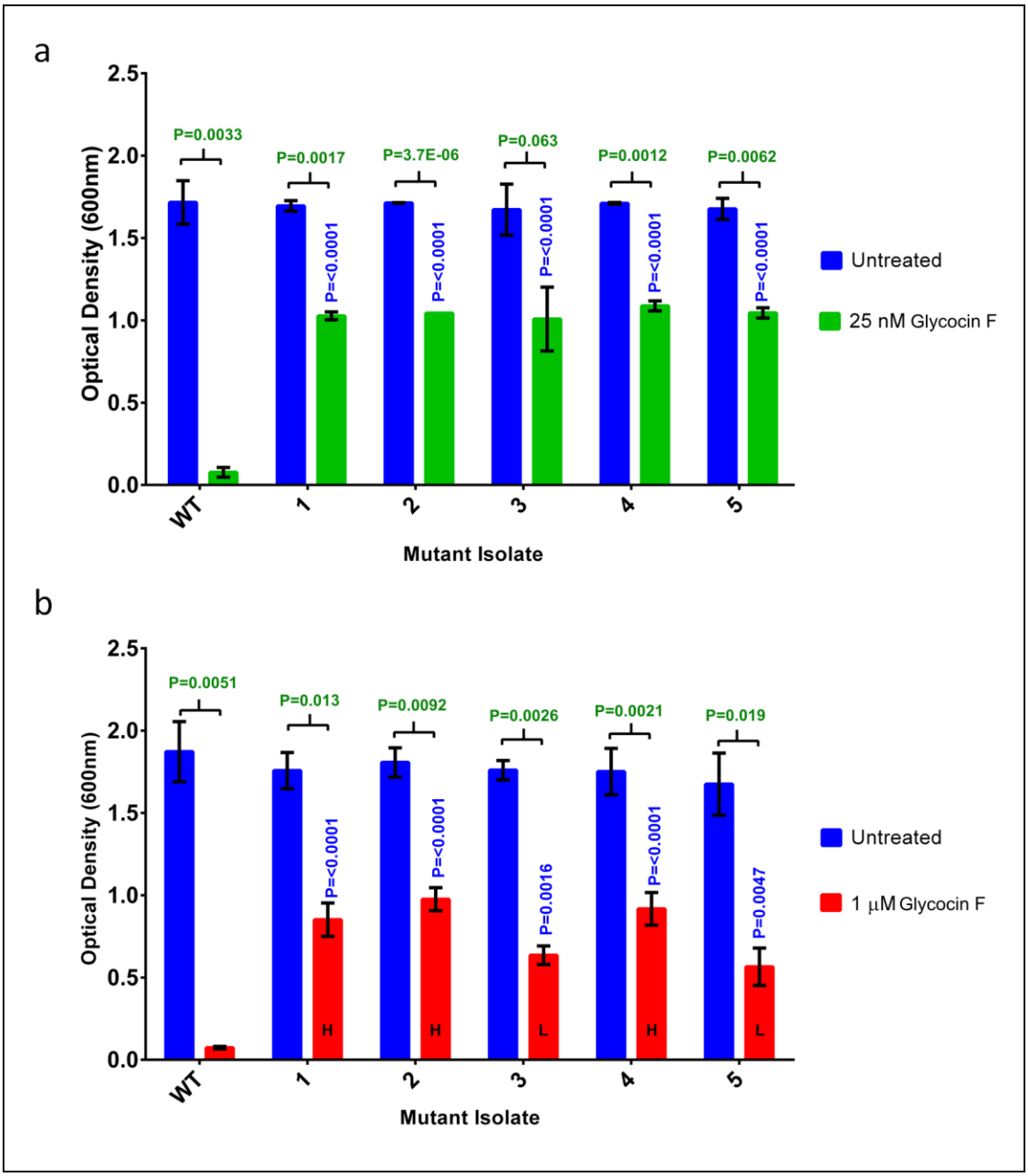


Figure 3.4 **Liquid culture growth inhibition of wild-type and glycocin F resistant *L. plantarum* subsp. *plantarum* ATCC 14917 mutants after 400 minutes**

Bar graphs represent the mean OD_{600nm} after 400 minutes (from growth curves in Figure 3.5) of duplicate experiments; error bars represent one standard deviation from the mean. Two cultures were inoculated at an OD_{600nm} of 0.05; one set was treated with either (a) 25 nM or (b) 1 μM glycocin F and the other left untreated. Statistical significance (P-value) between treated and untreated samples was determined using the Holm-Sidak test (values in green). Statistical significance (P-value) of the treated mutants (1-5) to the treated wild-type (WT) control was determined using the Holm multiple comparison test (values in blue). The letter H or L in the bar indicates which resistance group the mutant belongs to.

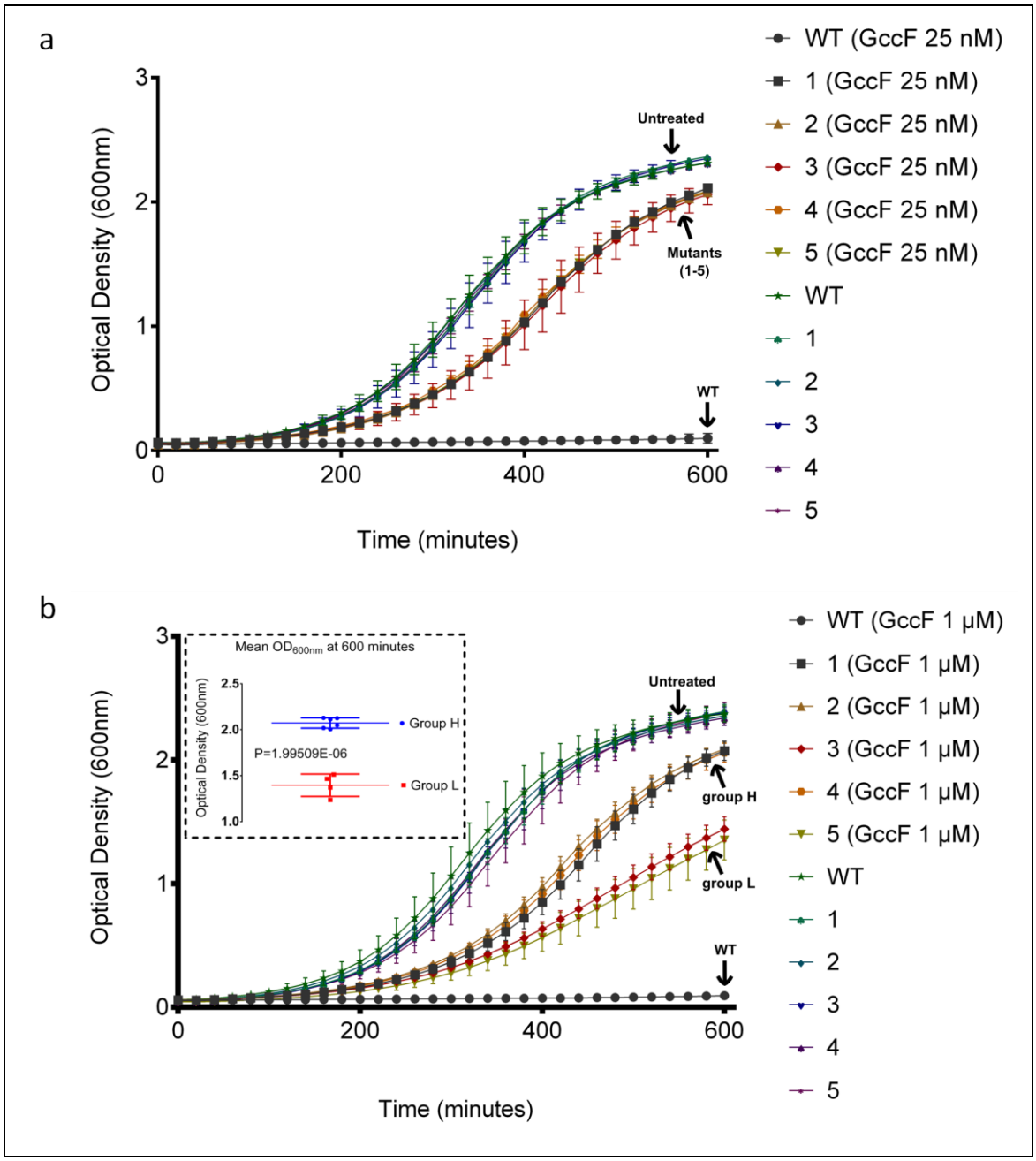


Figure 3.5 Growth curves of wild-type and mutant glycocin F-resistant *L. plantarum* subsp. *plantarum* ATCC 14917 mutants treated with glycocin F

Duplicate liquid culture growth inhibition assays, at (a) 25 nM glycocin F or (b) 1 µM glycocin F, of wildtype (WT) and glycocin F-resistant *L. plantarum* subsp. *plantarum* ATCC 14917 mutants (1-5). The mean OD_{600nm} at 20 minute intervals is plotted vs time (minutes). Legends on the right indicate the strain and final glycocin F (GccF) concentrations added at 0 minutes. Error bars indicate one standard deviation from the mean. The inset scatter plot in (b) is the mean OD_{600nm} after 600 minutes for group H and group L mutants. Error bars indicate one standard deviation from the mean. P-value signifies statistical significance between the group means.

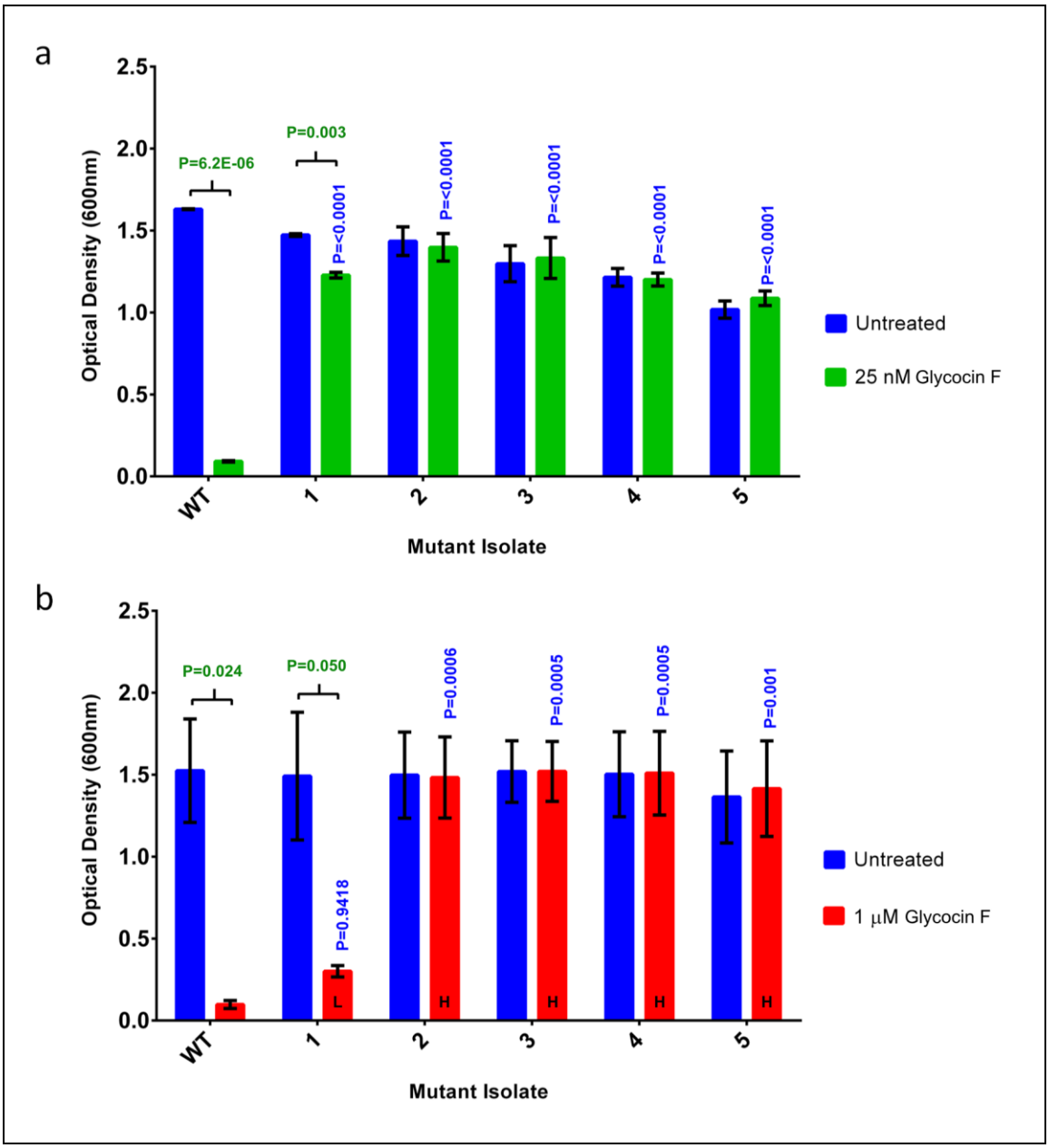


Figure 3.6 **Liquid culture growth inhibition of wild-type and glycocin F resistant *L. plantarum* ATCC 8014 mutants after 400 minutes**

Bar graphs represent the mean OD_{600nm} after 400 minutes (from growth curves in Figure 3.7) for duplicate experiments; error bars represent one standard deviation from the mean. Two cultures were inoculated at OD_{600nm} of 0.05 and set one was treated with either (a) 25 nM or (b) 1 μM glycocin F while the other was left untreated. Statistical significance (P-value) between treated and untreated samples was determined using the Holm-Sidak test (values in green). Statistical significance (P-value) of the treated mutant (1-5) to the treated wild-type (WT) control was determined using the Holm multiple comparison test (values in blue). There was no statistical significance between treated and untreated samples from mutants 2, 3, 4 and 5. The letter H or L in the bar indicates which resistance group the mutant belongs to.

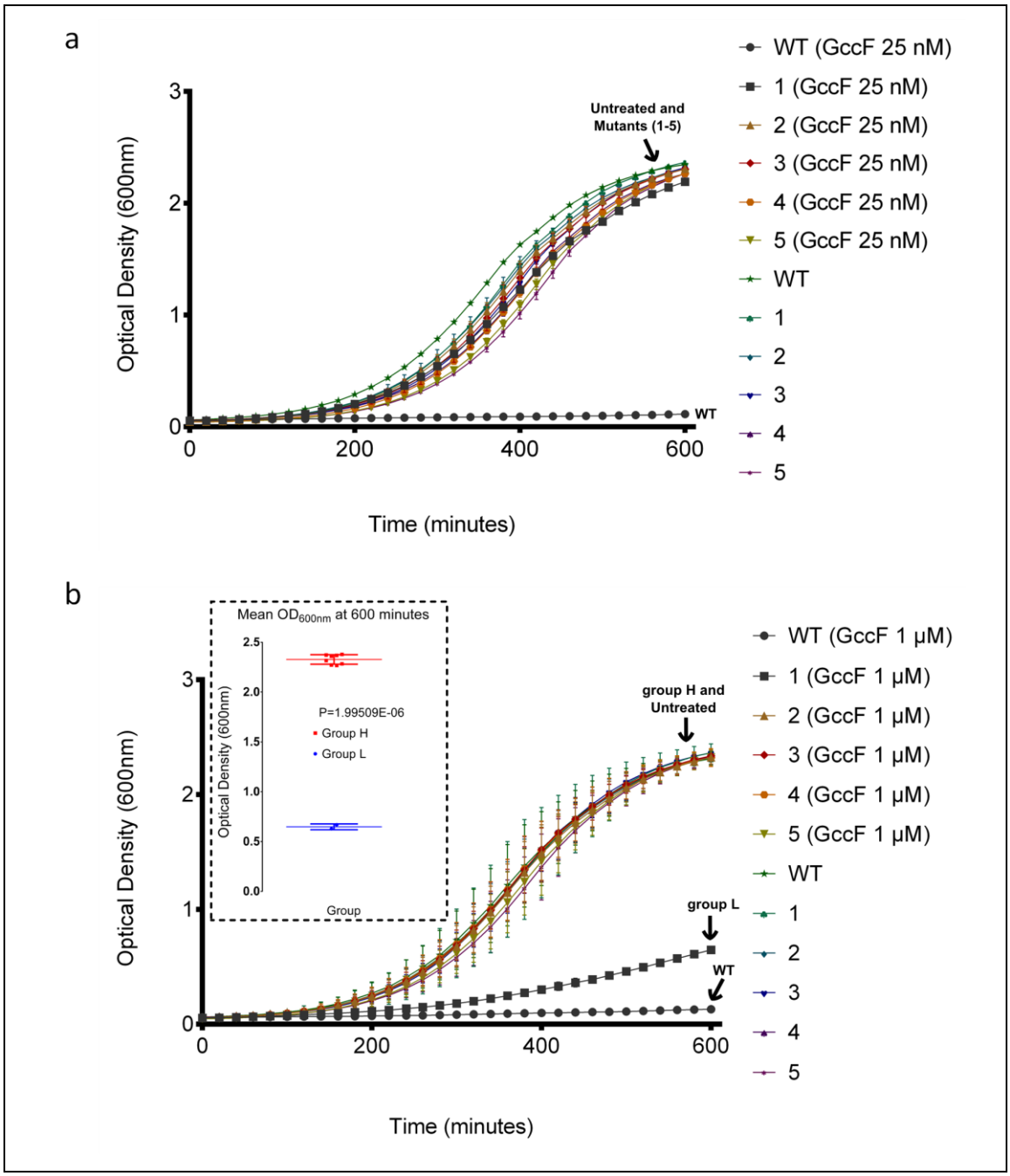


Figure 3.7 Glycocin F liquid culture assays of glycocin F resistant *L. plantarum* ATCC 8014 mutants

Duplicate liquid culture growth inhibition assays, at (a) 25 nM glycocin F or (b) 1 µM glycocin F, of wild-type (WT) and glycocin F-resistant *L. plantarum* ATCC 8014 mutants (1-5). The mean OD_{600nm} at 20 minute intervals is plotted vs time (minutes). Legends on the right indicate the strain and final glycocin F (GccF) concentrations added at 0 minutes. Error bars indicate one standard deviation from the mean. The inset scatter plot in (b) is the mean OD_{600nm} after 600 minutes for group H and group L mutants. Error bars indicate one standard deviation from the mean. P-value signifies statistical significance between the means of the groups.

3.3 Sequencing, assembly, annotation and comparative analysis of the genomes of glycocin F-resistant mutants

3.3.1 Introduction

gDNA from the glycocin F-resistant mutants (section 3.1) and their respective glycocin F-sensitive wild-type strains were sequenced using the Illumina® MiSeq™ system. The mutants were selected for genome sequencing (section 2.3.2) based on the findings of the glycocin F solid plate and liquid culture sensitivity assays (section 3.2); from each strain, two group H mutants and one group L mutant were selected: LP1_2**H**, LP1_3**L** and LP1_4**H** from *L. plantarum* subsp. *plantarum* ATCC 14917 mutants, and LP8_1**L**, LP8_4**H** and LP8_5**H** from *L. plantarum* ATCC 8014 mutants. Both group L and group H mutants were selected as it was assumed that mutations identified in the group L mutants would complement those identified in the group H mutants in relation to how mutations were conferring resistance to glycocin F. Three *de novo* assembly programs (two of these use a de Bruijn graph algorithm and one using an overlap-layout consensus algorithm), were used to perform *de novo* assembly of eight sequenced genomes, which were then annotated and comparatively analysed to identity proteins that might be required for sensitivity to glycocin F.

3.3.2 Aims

- To sequence the genomes of the selected glycocin F-resistant *L. plantarum* subsp. *plantarum* ATCC 14917 and *L. plantarum* ATCC 8014 mutants.
- To perform *de novo* assembly and annotate the sequenced genomes.
- To comparatively analyse the assembled genomes and identify mutation possibly responsible for glycocin F resistance.

3.3.3 Results and discussion of comparative genomic analysis

gDNA from the glycocin F-resistant mutants (section 3.1) and their respective glycocin F-sensitive wild-type strains was purified (section 2.3.1) and quantified using the NanoDrop® (section 2.1.14). Analysis of the purified gDNA using agarose gel electrophoresis (section 2.1.13) confirmed that the samples were not heavily contaminated with RNA (Appendix 3). Sequencing generated 250-bp paired-end sequencing reads with each direction having an average of 2,596,507 sequences which equates to an approximate 399-fold average coverage based on the estimated genome sizes (section 2.4.2). These raw reads were subjected to quality control processing (section 2.4.1) by MGS, and also as part of this work, prior to *de novo* assembly (section 2.4.3). Statistical details on the raw sequencing reads and quality control

processed reads are provided in Appendix 4. *De novo* assembly of combined, forward and reverse fastq files from the quality control processing (section 2.4.1) was performed using Velvet, SPAdes and Edena (section 2.4.3).The results of all assemblies are detailed in Appendix 5 and details pertaining to the *de novo* Velvet assemblies are detailed in Table 3.4, the reasoning for this is provided below. Contigs generated from all three *de novo* assemblies were processed by CISA (section 2.4.4) to extend the contigs from other representative contigs, identify misassembled contigs and split and trim these, and finally to merge contigs based on overlap and repetitive regions (Lin and Liao, 2013). Results of the integration are listed in Appendix 5. Validation of the CISA contigs was done by mapping the reads from the filtered forward and reverse fastq files (that had reads with ambiguities and those shorter than 150-bp removed) back to the CISA contigs (section 2.4.6).

Table 3.4 Velvet assembly parameters and results

	L. plantarum ATCC 8014			
Isolate	LP8_1**L**	LP8_4**H**	LP8_5**H**	WT
k-mer	201	201	201	201
exp_cov	66	62	73	66
cov_cutoff	25.62	31.31	28.34	10.65
ins_length	230	230	230	231
ins_length_SD	25	42	21	21
Nodes	76	89	76	71
Total nucleotides (bp)	3254122	3253155	3254022	3254764
Min length (bp)	401	68	401	401
Max length (bp)	358980	429109	358980	528679
Average length (bp)	42817.39	36552.3	42816.08	45841.75
N50 length (bp)	196709	91684	150664	229667
GC (%)	44.51	44.50	44.51	44.50
	L. plantarum subsp. *plantarum* ATCC 14917			
Isolate	LP1_2**H**	LP1_3**L**	LP1_4**H**	WT
k-mer	201	201	201	201
exp_cov	49	54	61	58
cov_cutoff	22.36	3.90	3.36	4.56
ins_length	232	229	231	230
ins_length_SD	19	36	40	22
Nodes	50	70	88	51
Total nucleotides (bp)	3232022	3238718	3244568	3234848
Min length (bp)	401	191	191	193
Max length (bp)	444383	545064	547839	507253
Average length (bp)	64640.44	46267.4	36870.09	63428.39
N50 length (bp)	244237	226553	191110	413585
GC (%)	44.48	44.47	44.48	44.48

In the assemblies of some mutant isolates, mis-assemblies were identified in the CISA contigs that were then traced back to the SPAdes *de novo* assemblies. These mis-assemblies were not identified in the contigs generated by *de novo* Velvet or Edena assemblies, thus due to time constraints the integrated contigs were discarded and comparative analysis (section 2.4.7) was carried out using the contigs generated by *de novo* Velvet assembly. The Velvet contigs from all strains were annotated using Prokka (section 2.4.5) (results are summarised in Table 3.5 and are provided in electronic Appendix 6). The locus tags prefixes from the NCBI BioProjects (http://www.ncbi.nlm.nih.gov/bioproject/) (Table 3.2), originally set up for CISA integrated contigs, were used in the annotation of the Velvet contigs. In due time these BioProjects will be adjusted to describe a *de novo* assembly by Velvet, and another eight BioProjects will be submitted to NCBI BioProjects representing contigs integrated using CISA. Submission of these genomes to the NCBI Genome (http://www.ncbi.nlm.nih.gov/genome) will be completed upon successful integration of the contigs using CISA.

Comparative analyses (section 2.4.7) were performed using the wild-type strains as reference sequences to which the three mutant isolate genomes were aligned. The following mutations were ignored: discrepancies within 250-bp of contig ends, silent mutations inside coding sequences and mutations in intergenic regions that were not likely to alter gene expression. Analysis of the *L. plantarum* ATCC 8014 genomes identified two areas in the mutant genomes that differed to the wild-type. These mutations are described at both the genetic level and

Table 3.5 Summary of Prokka annotation of velvet assemblies

	L. plantarum ATCC 8014			
	LP8_1**L**	LP8_4**H**	LP8_5**H**	WT
Gene	3162	3149	3162	3151
CDS	3032	3028	3031	3029
Signal peptide	155	156	156	154
Misc_RNA	66	67	66	67
tRNA	63	53	64	54
tmRNA	1	1	1	1
	L. plantarum subsp. *plantarum* ATCC 14917			
	LP1_2**H**	LP1_3**L**	LP1_4**H**	WT
Gene	3162	3157	3148	3160
CDS	3026	3022	3026	3026
Signal peptide	174	172	170	173
Misc_RNA	69	70	68	69
tRNA	66	64	53	64
tmRNA	1	1	1	1

protein translation level along with original colony identification and final mutant identifier in Table 3.6. A region of sequence in the wild-type genome (GenBank accession number EV52_contig000013) was found to contain a mutation in a sequence that is annotated as a *purR* CDS (GenBank accession number EV52_02355) due to its sequence similarity to a purine operon repressor (GenBank accession number YP_003061980.1) from *L. plantarum* JDM1. This mutation is only identified in the genome of isolate LP8_1**L** (GenBank accession number EV51_contig000001) and is an alanine to serine (A276S) missense mutation involving a cytosine to adenine substitution at nucleotide position 825 of the *purR* CDS (GenBank accession number: EV51_00058).

PurR has 70.7 % pairwise identity and 56.1 % sequence identity to two purine operon repressors that are deposited in the RCSB Protein Data Bank (PDB) (PDB# 1P4A and 1O57 respectively) (http://www.rcsb.org/). A276 is located three amino acids from the C-terminus in PurR and is not conserved in either of the PDB protein structures. The amino acids surrounding A276 also have low similarity. Both PurR structures show that the C-terminal tail is not involved in the integrity of the subunit structures or the dimerisation interfaces. Due to the conservative nature of this mutation and its location no major changes to the function of PurR are predicted to occur.

Table 3.6 Mutations identified in glycocin F resistant mutants

Strain	Resistance Group	Mutant Identifier	Genetic Mutation	Protein Mutation
ATCC 8014	Low	LP8_1**L**	*purR(C825A); pts18CBA(C610A)*	PurR(A276S);PTS18CBA(T204N)
ATCC 8014	High	LP8_2**H**	Unknown	Unknown
ATCC 8014	High	LP8_3**H**	Unknown	Unknown
ATCC 8014	High	LP8_4**H**	*pts18CBA(*Δ769-838)	PTS18CBA(Δ257-280)
ATCC 8014	High	LP8_5**H**	*pts18CBA(Am898)*	PTS18CBA(Δ300-662)
ATCC 14917	High	LP1_1**H**	Unknown	Unknown
ATCC 14917	High	LP1_2**H**	Δ*panE1*;*pts18CBA(Δ1-72)*	ΔPANE1;PTS18CBA(Δ1-24)
ATCC 14917	Low	LP1_3**L**	*pts18CBA(A630C)*	PTS18CBA(N211H)
ATCC 14917	High	LP1_4**H**	*pts18CBA(G727A)*	PTS18CBA(G243E)
ATCC 14917	Low	LP1_5**L**	Unknown	Unknown

The second difference was in a protein coding sequence in the wild-type genome (GenBank accession number EV21_contig000007) that contains a different mutation in each of the three sequenced glycocin F-resistant isolates (Appendix 7). This coding sequence is annotated as *pts18CBA_3* (GenBank accession number EV52_01314) due to its high similarity to the GlcNAc-specific PTS-EIICBA (GenBank accession number YP_003063622.1) from *L. plantarum* JDM1 (GenBank accession number NC_012984) (Figure 3.8). It should be noted that *pts18CBA_3* is the homolog of *pts18CBA* identified in *L. plantarum* WCFS1 (section 1.4) that was targeted for knockout (section 3.5) and that two other regions in the genome of *L. plantarum* ATCC 8014 were mis-annotated as being *pts18CBA* and *pts18CBA_2*. In the genome of mutant LP8_1**L** (GenBank accession number EV51_contig000029), a threonine to asparagine (T204N) substitution is the result of a cytosine to adenine missense mutation at nucleotide position 610. In the genome of mutant LP8_4**H** (EV50_000010) a 69-nucleotide deletion between positions 769 and 839 was identified in *pts18CBA_3*. This deletion, mapped in LP8_4**H**, results in the loss of 23 amino acids and does not cause a frame shift in the downstream sequence. In the genome of the isolate LP8_5**H** (GenBank accession number EV49_contig000029) there is an amber nonsense mutation resulting from a guanidine to adenine substitution at nucleotide position 898 in *pts18CBA_3*. This nonsense mutation removes a 363-amino acid segment from PTS18CBA, truncating the 662-amino acid protein by more than 50 %. Figure 3.9 shows the location of these mutations in the translated *pts18CBA_3*.

Comparative analyses of the *L. plantarum* subsp. *plantarum* ATCC 14917 genomes (Appendix 8) identified a single region in the wild-type genome (GenBank accession number EV48_contig000009) with varying degrees of change in the three mutant genomes. This region is homologous to the region in *L. plantarum* ATCC 8014 in which mutations were mapped to the *pts18CBA_3* CDS (GenBank accession number EV52_01314) (Figure 3.8). This region in *L. plantarum* subsp. *plantarum* ATCC 14917 was annotated by Prokka to contain a *pts18CBA* CDS (GenBank accession number EV48_02088), an obvious homolog of *pts18CBA_3* with 99.7 % amino acid pairwise identity. In mutant LP1_2**H** (GenBank accession number EV55_contig000009) a 1,285-bp deletion is observed at position 222,721 (GenBank accession number EV55_contig000009). This includes the entire *panE1* CDS (GenBank accession number EV48_02087), and 49-bp upstream and 261-bp downstream of *panE1*. The downstream region includes 72-bp of the *pts18CBA* CDS. The gene *panEI* encodes the enzyme 2-dehydropantoate 2-reductase (GenBank accession number YP_003063623.1) and is 189-bp upstream from the start of the *pst18CBA* CDS. In the genome of mutant LP1_3**L** (GenBank accession number EV54_contig000003) an asparagine to histidine (N211H) substitution is the result of an adenine to cytosine missense mutation at nucleotide position 630. In the genome of mutant LP1_4**H**

(GenBank accession number EV53_contig000052) a glycine to glutamic acid (G243E) substitution is the result of a guanine to adenine missense mutation at nucleotide position 727.

The genomes of two wild-type and six glycocin F-resistant mutants were partially assembled, then annotated, and comparative analysis revealed a different mutation in the *pts18CBA* gene in each of the six resistant mutants. These mutations are discussed both in greater detail and in context to their location in a predicted PTS18CBA 3D model (section 3.4).

Figure 3.8 DNA sequence Alignment of *L. plantarum* subsp. *plantarum* ATCC 14917 and *L. plantarum* ATCC 8014

DNA alignment of the *pts18CBA* coding region, and flanking DNA, from *L. plantarum* subsp. *plantarum* ATCC 14917 and *L. plantarum* ATCC 8014 was undertaken using Geneious™, a detailed view of the alignments are provided in Appendix 7 and Appendix 8 which were generated using CLUSTAL OMEGA (http://www.ebi.ac.uk/Tools/msa/clustalo/). The thin grey ribbons represent the two nucleotide sequences of *L. plantarum*; top *L. plantarum* subsp. *plantarum* ATCC 14917, bottom *L. plantarum* ATCC 8014. Vertical lines in the thin grey ribbons represent locations where the nucleotide sequence is different between the two strains: A red, T green, C blue and G yellow. Annotations above or below the sequence belong to the respective strain. Yellow bars represent protein coding sequences. Dark red bars are the region of sequence amplified by high fidelity PCR for knockout vector construction. Beige annotations are primer binding sites. Light red annotation is the T-box element annotated by Prokka. Lime annotations are promoter regions predicted using online promoter prediction software BPROM (Solovyev and Salamov, 2011) (http://linux1.softberry.com/all.htm).

Figure 3.8 DNA sequence Alignment of *L. plantarum* subsp. *plantarum* ATCC 14917 and *L. plantarum* ATCC 8014

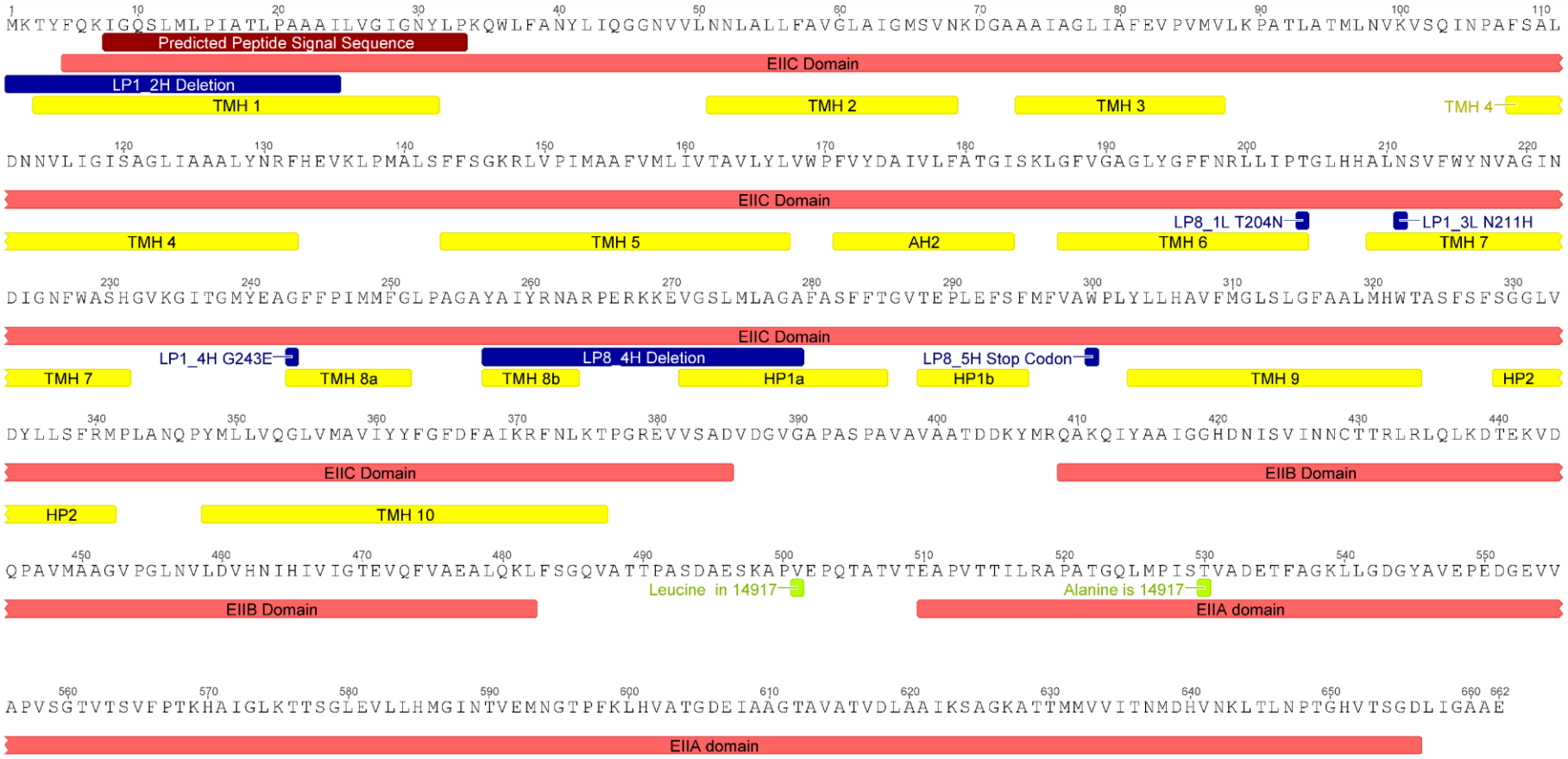


Figure 3.9 Amino acid sequence of PTS18CBA from *L. plantarum* ATCC 8014

The *pts18CBA* CDS from *L. plantarum* ATCC 8014 was translated using the bacterial translation table. The predicted EII domains are indicated by the red ribbons. The yellow bars are the positions of the transmembrane helices (TMH) and re-entrant loops from the threaded EIIC model. Their naming convention is based on 3D ChbC model. The two amino acid residues that differ in *L. plantarum* subsp. *plantarum* ATCC 14917 are indicated in lime. Mutations that are mapped in the glycocin F-resistant mutants are indicated by the blue annotations.

3.4 Explaining the PTS18CBA mutations using *in silico* methods

3.4.1 Introduction

PTS18CBA is the EII component of a carbohydrate PTS and is thought to be specific for GlcNAc. It is a multi-domain protein in which the consecutive EIIC, EIIB and EIIA domains are fused into a single polypeptide, as predicted by NCBIs Conserved Domain search (http://www.ncbi.nlm.nih.gov/Structure/cdd/wrpsb.cgi). Each of the six mutations mapped to *pts18CBA* in the six glycocin F-resistant isolates are located in or (in the case of LP1_2H) overlap the predicted CDS of the EIIC domain. The EIIC domain is the transmembrane component of the PTS that is responsible for carbohydrate recognition and membrane translocation of the sugar (Simoni *et al.*, 1976). How these mutations affect both the structure and function of the EIIC domain could be determined from experiments designed to assess particular properties of the protein, such as saccharide phosphorylation, carbohydrate-protein interaction, protein fold and domain interactions. Such investigations might use surface plasmon resonance (SPR), isothermal titration calorimetry (ITC), X-ray diffraction or NMR techniques. Due to time constraints it was not possible to experimentally investigate the effect of these mutations on PTS structure or function, but previous reports by Schnetz *et al.* (1990), Ruijter *et al.* (1992), Buhr *et al.* (1992), Weng *et al.* (1992), Cao *et al.* (2011) and McCoy *et al.* (2014), combined with *in silico* modelling methods, similar to what was performed by Kjos *et al.* (2010a), provided the bases for informed speculation on the structural and functional consequences of these mutations.

3.4.2 Results and Discussion

PTS18CBA amino acids 1-410, containing the PTS18CBA EIIC domain (PTS18C) were subjected to homology-based tertiary structure prediction using the Phyre^2 web server (Kelley and Sternberg, 2009) (http://www.sbg.bio.ic.ac.uk/phyre2). The top ranked model (EIIC-PHY) predicted by Phyre^2 (Figure 3.10a), based on sequence alignments, secondary structure predictions and statistical probabilities, was PTS18C threaded to the structure of a N,N'-diacetylchitobiose-specific PTS-EIIC permease (PDB# 3QNQ) (Figure 3.10b) from *Bacillus cereus* (ChbC) with which it shares only 14 % amino acid sequence identify. This threading resulted in 93 % of the amino acids being modelled with >90 % confidence (Appendix 9 and Appendix 10). The SuperPose v1.0 web server was utilised, with default settings, to compare the threaded EIIC-PHY model to the structure of ChbC (Maiti *et al.*, 2004) (http://wishart.biology.ualberta.ca/SuperPose/). Not surprisingly, the alignment showed the root

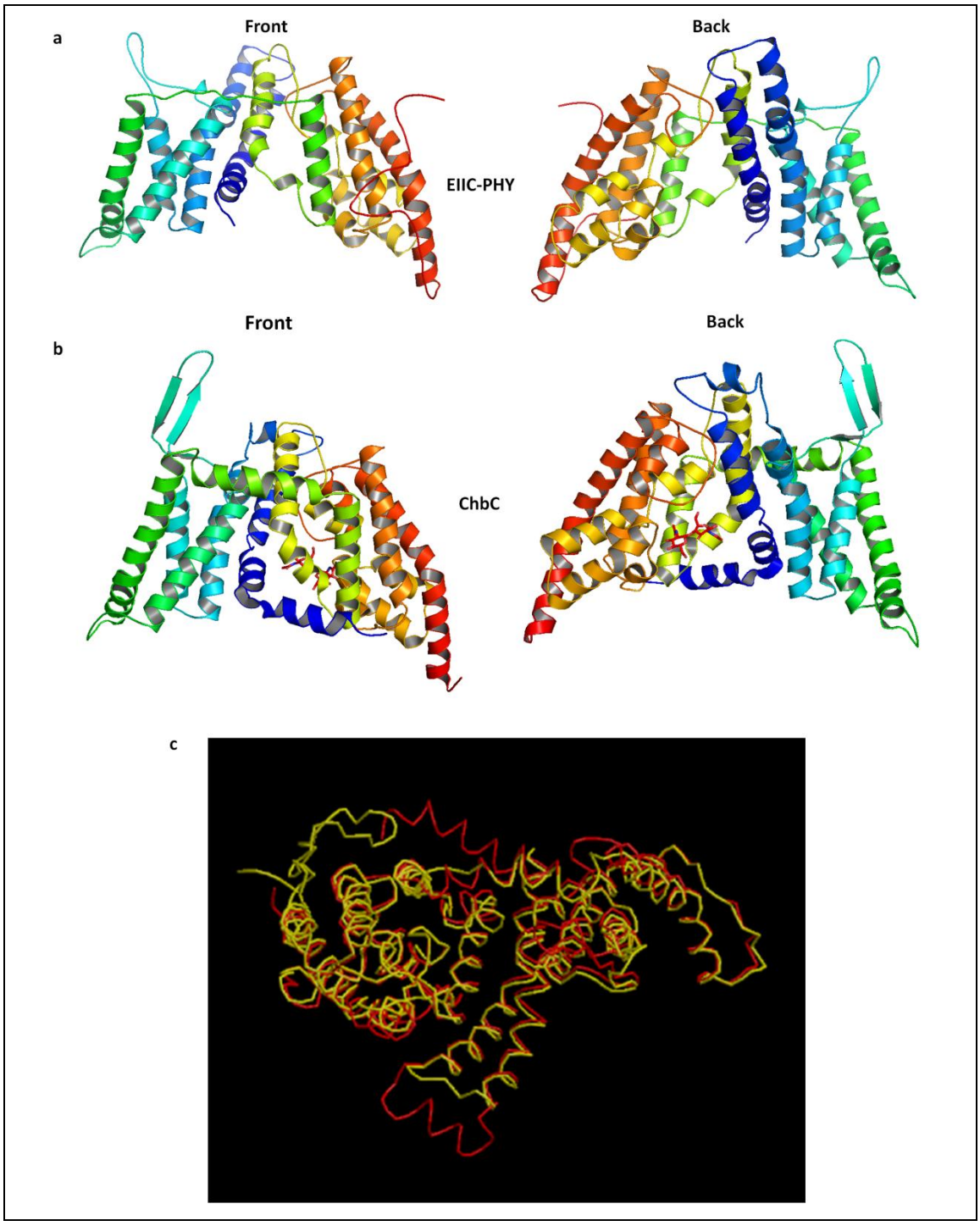


Figure 3.10 Predicted model of the transmembrane EIIC domain from *L. plantarum* ATCC 8014 PTS18CBA

PTS18CBA amino acids 1-410 (includes the entire EIIC domain) were submitted to the Phyre[2] web server to produce a model of the 3D structure (EIIC-PHY), shown in (a). The structure used to model EIIC-PHY was ChbC, shown in (b). Both the front and back view of the EIIC-PHY model and solved 3D structure of ChbC are shown. The N-terminal is blue and the C-terminal is red. c) The superposition of the backbone atoms of ChbC (red) and EIIC-PHY (yellow).

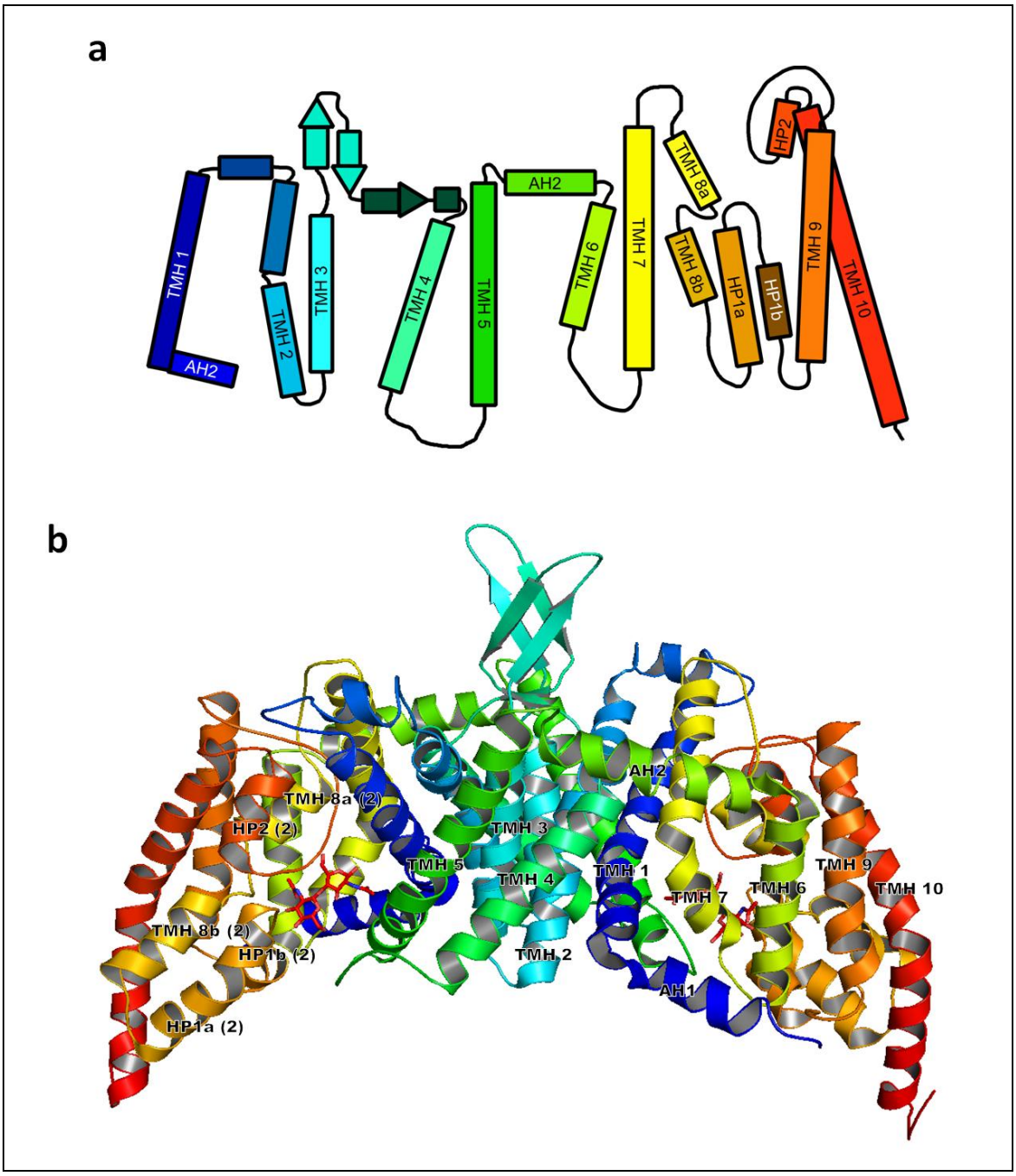


Figure 3.11 Topology of the TMHs of ChbC

a) Figure adapted with permission from McCoy *et al.* (2014) showing the topology of the 10 TMHs in ChbC, colour-coded to match the TMHs in (b). b) The 3D structure of the dimer of the EIIC domain of the N,N'-diacetylchitobiose-specific PTS from *B. cereus* was solved to 3.3 Å using X-ray diffraction by Cao *et al.* (2011). TMHs and re-entrant loops are numbered according to McCoy *et al.* (2014). The five helices labelled with a '(2)' are from the back side view of the second subunit of the dimer that are not visible from the front view.

mean square deviation (RMSD) between the backbone atoms of the two structures to be 4.64 Å, showing they have the same fold and 3D structure (Figure 3.10c). To provide some weight to the threaded EIIC-PHY model, PTS18C was also subjected to homology based tertiary structure prediction using the web server RaptorX (Kallberg *et al.*, 2012) (http://raptorx.uchicago.edu/). Again PTS18C was threaded into the structure of ChbC and was reported as the top ranked model with a P value of $4.2e^{-7}$ indicating the threaded model by RaptorX is of relatively good quality. The predicted model and relevant results generated by RaptorX are provided (Appendix 11). Superpose was used to compare the RaptorX model and Phyre^2 model, a RMSD between the backbone of the atoms of the two structures was calculated to be 5.11 Å showing them to have the same fold and 3D structure (Appendix 11).

The secondary structure and transmembrane prediction by Psi-pred 2.5 (Jones, 1999) and Memsat SVM (Jones, 2007) respectively, implemented in Phyre^2, predicted the presence of 12 TMHs. These 12 TMHs correlate to the helices observed in the threaded EIIC-PHY model, which suggest that the EIIC domain of PTS18CBA has very similar membrane topology to ChbC. However two of the TMH predicted by Psi-pres 2.5 and Memsat SVM, in both the EIIC-PHY model structure and the ChbC structure, are re-entrant loops, not TMHs. The threaded EIIC-PHY model is also missing the N-terminal helix, AH1 present in the ChbC structure. Due to the high degree of similarity in the topology of the TMHs between the threaded EIIC-PHY and ChbC structures, the naming convention for all helices in the threaded EIIC-PHY model is based on ChbC topology (Figure 3.9 and Figure 3.11). It is perhaps fortuitous for these modelling studies that the only PTS-EIIC domain for which an X-ray crystal structure is available (ChbC), transports a GlcNAc disaccharide.

The mutations in the genomes of all six resistant strains were mapped to PTS18CBA and fell into two categories. Either a relatively large mutation, as occurs in LP1_2**H**, LP8_4**H** and LP8_5**H**, or a single missense mutation, as occurs in LP8_1**L**, LP1_3**L** and LP1_4**H**. The largest of the mutations, mapped in mutant LP1_2**H**, is a 1,285-bp deletion that results in the of: (a) loss of the entire *panE1* CDS, (b) the -35 and -10 promoter elements of *pts18CBA* and (c) the -10 element of *panE1*, predicted by online promoter prediction software BPROM (Solovyev and Salamov, 2011) (http://linux1.softberry.com/all.htm), and (d) the first 24 codons of the *pts18CBA* CDS (Figure 3.8 and Figure 3.9). This deletion does not introduce a frame shift into the remaining 1,917-bp of *pts18CBA* and the first codon from the remaining coding sequence is a TTG codon which can initiate translation, albeit relatively infrequently (Villegas and Kropinski, 2008). However due to the deletion of the promoter region, the remainder of the *pts18CBA* gene would not be expressed.

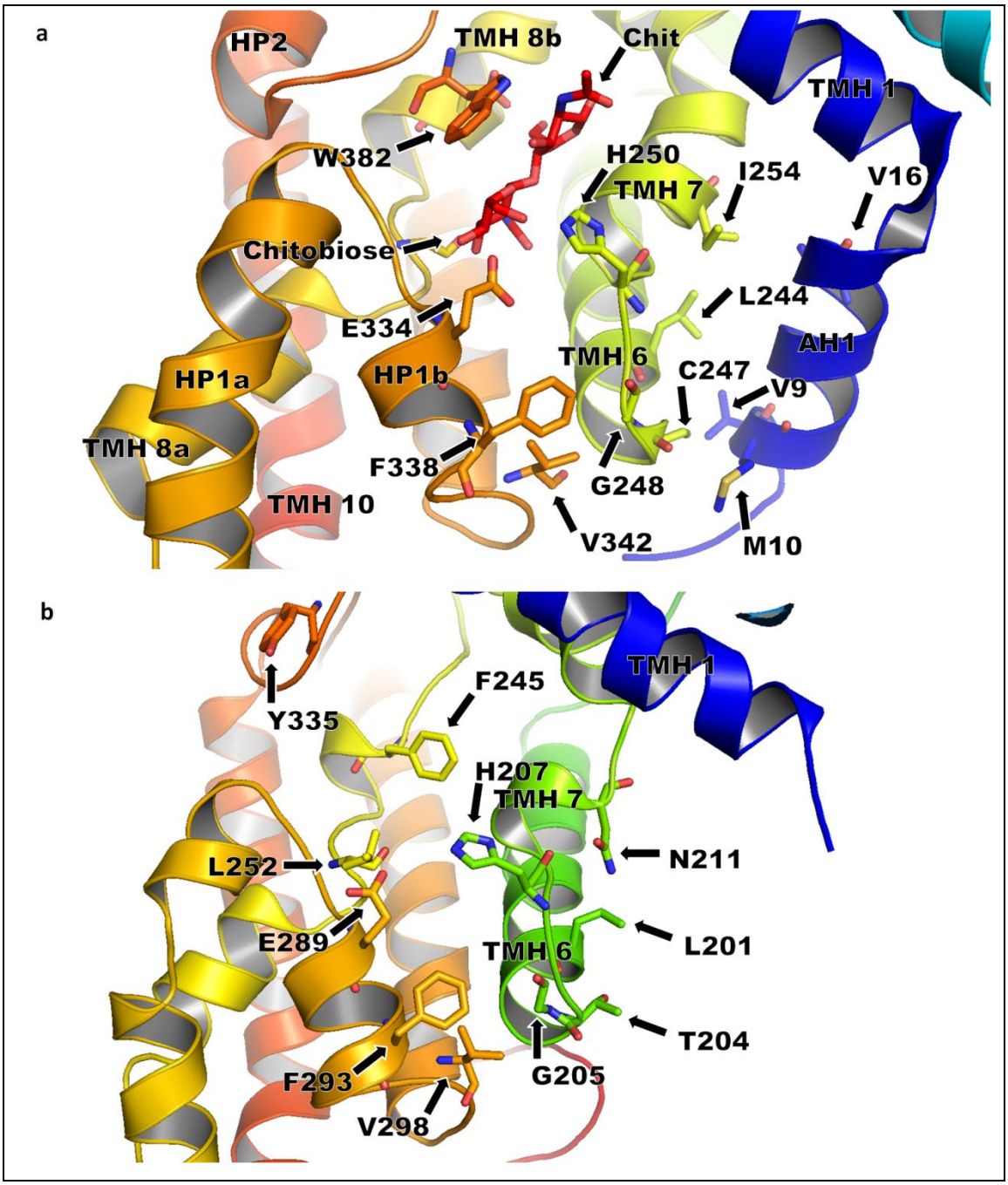


Figure 3.12 Substrate sugar binding site of ChbC and EIIC-PHY

a) Substrate sugar binding site located in the structure of ChbC. The 26 amino acid conserved region comprised of TMHs 6-7 and can be observed forming one side of the binding site for N,N'-diacetylchitobiose (Chit, red), along with the essential histidine 250. Glycine 248 is present at the C-terminus of TMH 6 packing between F338 and V342. The same region from the threaded EIIC-PHY model is provided in (b) for comparison, showing the location of the mutated residues T204 and N211.

The second largest mutation is the nonsense mutation mapped in mutant LP8_5**H**. This mutation results in the loss of the EIIB and EIIA domains and 80 residues from the C-terminal region of the EIIC domain (Figure 3.9). The *L. plantarum* ATCC 8014 genome contains only two putative GlcNAc-specific PTS transporters, *pts18CBA 3* (GenBank accession number: EV52_01314) and *pts22CBA* (GenBank accession number: EV52_01147) and in both cases the EIIB and EIIA domains are fused to the EIIC domain so loss of the EIIB and/or EIIA domain from *pts18CBA* will almost certainly render this protein functionally inactive due to an incomplete phosphorylation cascade. EII domains from other PTSs would not be able to complement the loss as discussed in section 1.3.2, and it would therefore be predicted to abolish transport and regulatory functions. The model, EIIC-PHY, showed that the 80 EIIC residues to make up TMH 9, HP2 and TMH 10 (Figure 3.9). In ChbC (Figure 3.11) these TMHs are part of the transport domain, interface with the membrane, and also contain residues predicted by Cao *et al.* (2011) and McCoy *et al.* (2014) to be involved in stabilising N,N-diacetylchitobiose through stacking interactions (Figure 3.12). Their loss is predicted to cause major structural changes to the transport domain of EIIC, resulting in conformational instability most likely leading to proteolytic degradation of the mutated PTS transporter (Striebel *et al.*, 2014).

The last of the large mutations is a 23-amino acid deletion in the genome of mutant LP8_4**H**. In the model EIIC-PHY these are residues 257-279 which correspond to residues 302-324 in ChbC (Appendix 9). These 23 residues make up the C-terminal part of TMH 8b and the N-terminal part of HP1a (Figure 3.9 and Figure 3.11). In the EII-PHY model and the ChbC structure, TMH 8b and HP1a are part of the transport domain and interface with the membrane while the loop region between them is exposed to the cytosol (Figure 3.11). The loss of the 23 amino acids from this region of the protein would most likely disrupt the structure and render it non-functional. In a similar fashion to the mutation observed in LP8_5**H,** the 23-amino acid loss is predicted to result in protein insolubility and possible proteolytic degradation.

To help predict the effects of the single missense mutations the conserved nature of these residues was determined using NCBI's Position-Specific Iterated Basic Local Alignment Search Tool (PSI-BLAST) (http://blast.ncbi.nlm.nih.gov/Blast.cgi) and NCBI's Constraint-Based multiple Alignment Tool (COBALT) (http://www.st-va.ncbi.nlm.nih.gov/tools/cobalt/). Residues 169-274 from *L. plantarum* ATCC 8014 PTS18CBA, corresponding to the C-terminus of TMH 5 to the N-terminus of TMH 9, were aligned to proteins in NCBI Protein Reference Sequence database using PSI-BLAST with the gap cost set to the non-default values: 'existence: 9, extension: 1'. This alignment was performed twice, once open to proteins containing an EIIC domain from all prokaryotes and again with a taxonomic restriction that included EIIC containing proteins from *Firmicutes* only. The sequences identified by PSI-BLAST were then

subjected to COBALT analysis and the conservation was identified using a conservation bit setting of 2.

In all prokaryotes a single region was identified as being conserved. This region is made up of 26 amino acids between glycine 190 and asparagine 217. In *Firmicutes* EIIC domains this region was still the only area of conservation detected, although it is expanded to include glycine 190. In both the threaded EIIC-PHY model and the ChbC structure, this conserved region consists of TMH 6, the loop region between THM 6 and TMH 7, and the N-terminal half of TMH 7 (Figure 3.9 and Figure 3.11). Within this conserved region there are some amino acids that are found to be particularly highly conserved. In *Firmicutes* the amino acids glycine 205 (ChbC: glycine 248) and leucine 206 (ChbC: isoleucine 249) are completely conserved while histidine 207 (ChbC: histidine 250) and histidine 208 (ChbC: glycine 251) have a high degree of conservation. In all prokaryotes this glycine 205, from EIIC-PHY, is completely conserved. Its equivalent glycine in ChbC, glycine 248, is located at the C-terminus of THM 6 (Figure 3.12a). Its peptide nitrogen makes a hydrogen bond with the backbone carbonyl oxygen of leucine 244 (EIIC-PHY: leucine 201), which is located at the C-terminus of THM 6, and its α-carbon packs tightly between the hydrophobic side chains of valine 342 (EIIC-PHY: valine 298) and phenylalanine 338 (EIIC-PHY: phenylalanine 293) from HP1b, which may explain glycine's conservation, as no other residue would fit in this small space (Figure 3.12a).

These two THMs, 6 and 7, are part of the transport domain and contain residues vital to the PTS transporters function. In ChbC, histidine 250 (H250) (EIIC-PHY: histidine 207) makes a hydrogen bond with the substrate sugar (Figure 3.12), and site-directed-mutagenesis of the equivalent histidine in *E. coli* mannitol-EIIC (Weng *et al.*, 1992, Weng and Jacobson, 1993) showed that this residue is involved in both the binding and phosphorylation of the mannitol. Weng *et al.* (1992) cloned and produced the mannitol-EIIC protein with the following mutations: H195N, H195R and H195A. They assayed the effect of these mutations on the phosphorylation of mannitol *in vitro* and showed that although H195N and H195R had activity comparable to the wild-type protein, H195A rendered the protein almost completely inactive. In a later study, Weng and Jacobson (1993) showed that the same three mutations increased the dissociation constant of mannitol and EIIC by an average factor of 216. An interesting observation by Weng *et al.* (1992) was that all three of these mutations rendered the recombinant protein more susceptible to endogenous proteolytic degradation suggesting that they may have disrupted protein conformation and as a result exposed some regions of the protein to proteases. These studies show the importance of the residues located in TMH6 to TMH 7.

The two amino acids T204 and N211, mutated in LP8_1**L** and LP1_3**L** respectively, are both located in this highly conserved region (Figure 3.9). Comparisons of these residues in the COBALT analyses shows that residue 204 is usually hydrophobic, and can be valine, phenyalanine, methionine, leucine, alanine or isoleucine. Residue 211 on the other hand is more varied, and amino acids at this position include tyrosine, histidine, tryptophan, serine, leucine, valine, threonine, cysteine, glycine, alanine and proline. The T204N mutation is conservative, but introduces a side chain containing a carboxamide group, that is capable of being involved in 3 hydrogen bonds-in contrast to the single hydrogen bond of the threonine side chain. The N211H mutation introduces a positive charge, at pH <6.5, that could potentially form an electrostatic bond or contribute to electrostatic repulsion.

G243, mutated in LP1_4**H**, is not part of the TMH 6-7 region or part of another large conserved region. Analysis of G243 shows that it is fairly conserved throughout the Bacterial domain and an analysis of the *Firmicutes* shows that the small number of variations that occur at this position are very conservative. It is usually replaced with small polar residues such as serine or threonine, and occasionally with glutamate. The mutation to E introduces a large negatively-charged side chain which could clash both sterically and electrostatically with neighbouring residues which make up the conserved region mentioned above. This could result in a disruption to the ligand-binding site or the transport domain. It is suggested that the mutation G243E disrupts the protein's conformation, possibly leading to proteolytic degradation while the T204N and N211H mutations may compromise GlcNAc binding and/or transporter function. Compromised GlcNAc binding would also be expected to decrease the binding affinity of the PTS18CBA EIIC GlcNAc-binding site for the GlcNAc moieties of glycocin F.

All the large mutations of LP1_2**H**, LP8_4**H** and LP8_5**H** result in relatively high level or complete resistance to glycocin F, while the missense mutations of LP8_1**L** and LP1_3**L** result in relatively low level resistance. Interestingly the G243E missense mutation in LP1_4**H** also results in a relatively high level of glycocin F resistance. It is possible that the large deletions are so deleterious that they result in PTS18CBA either being not transcribed, or being incorrectly folded leading to pupylation and proteolytic degradation (Striebel *et al.*, 2014). The effect of the single missense mutations correlated with low level resistance could result in a decreased substrate affinity or a non-functional transport domain, while the G243E mutation may cause a change in the stability of the transporter, leading to its degradation. Nevertheless, mutations which may impair the correct folding of PTS18CBA, or destabilise the folded state, result in a relatively high level of glycocin F resistance. What then is the basis of the residual glycocin F bacteriostatic activity against the glycocin F resistant *L. plantarum* subsp. *plantarum* ATCC 14917? One explanation is the presence of another membrane receptor (discussed in

section 3.8.3), which would also account for increased sensitivity of this strain to glycocin F compared to *L. plantarum* ATCC 8014.

3.5 Construction of *pts18CBA* knockout plasmids

3.5.1 Introduction

Plasmid pNZ5319, constructed by Lambert *et al.* (2007), was a kind gift from Richard Bongers at the Wageningen Centre for Food Science, Health and Safety Department, P.O. Box 20, 6710 BA Ede, The Netherlands. Plasmid pNZ5319 (Appendix 12) is a medium copy number plasmid used to construct knockouts of specific genes in *L. plantarum* WCFS1 (Lambert *et al.*, 2007). It uses a suicide mutagenesis strategy in which the plasmid that transforms competent *L. plantarum* is destroyed in the process of homologous recombination. pNZ5319 contains two selectable marker genes encoding chloramphenicol and erythromycin resistance, the former is incorporated into the genome upon homologous recombination, providing a simple method for screening clones that have undergone homologous recombination. Using Cre recombinase the selectable marker can be removed resulting in a system that allows multiple gene deletions in a single *L. plantarum* genetic background.

3.5.2 Aims

- To clone two approximately 1-kbp segments of DNA flanking *pts18CBA* from the genome of *L. plantarum* subsp. *plantarum* ATCC 14917 and *L. plantarum* ATCC 8014 into the multiple cloning sites of plasmid pNZ5319.

3.5.3 Primer design

Primers were designed according to the recommendations of Lambert *et al.* (2007) (Appendix 13). The primer pair 80_term and 81 were designed to generate '14917_F1' from the genome of *L. plantarum* subsp. *plantarum* ATCC 14917 so that 45 nucleotides at the 3′ end of 14917_F1 would be amplified from within the coding region of *pts18CBA* (Figure 3.8). The primer pair ‘82a and 83b’ were designed to generate '14917_F2' and '8014_F2 from the genomes of *L. plantarum* subsp. *plantarum* ATCC 14917 and *L. plantarum* ATCC 8014, so that 330 nucleotides at the 5′ end of 14917_F2 and 8014_F2 would be amplified from the coding region of *pts18CBA*. Primer pair 'F8014_1 and R8014_1_STP' were designed to generate '8014_F1' from the genome of *L. plantarum* ATCC 8014 so that 36 nucleotides at the 3′ end of 8014_F1 would be amplified from within the coding region of *pts18CBA*. The sequence TTA was present at the 5′ end of primer 80_term and F8014_1_STP to introduce a stop codon at amino acid

residue 16 of *pts18CBA* thus preventing the potential expression of a possible transcript from the *pts18CBA* after homologous recombination that might be toxic. No restriction enzyme recognition sites were included in the primers as all cloning was to be blunt end so as to negate restriction digestion of PCR products and dephosphorylation of linearised plasmid.

3.5.4 PCR amplification of flanking regions from *L. plantarum* gDNA

14917_F1, 14917_F2, 8014_F1 and 8014_F2 were amplified from *L. plantarum* subsp. *plantarum* ATCC 14917 and *L. plantarum* ATCC 8014 gDNA by high fidelity PCR (section 2.1.12) to generate blunt ended PCR products. The PCR reactions were analysed using agarose gel electrophoresis (section 2.1.13) and bands corresponding to 14917_F1, 14917_F2, 8014_F1 and 8014_F2, predicted to have lengths of 920-bp, 868-bp, 977-bp and 868-bp, respectively were identified as shown in Figure 3.13. The products were purified from the gel (2.1.13.6), their concentrations determined (section 2.1.14) and the products cloned into pNZ5319 (sections 3.5.6 and 3.5.9).

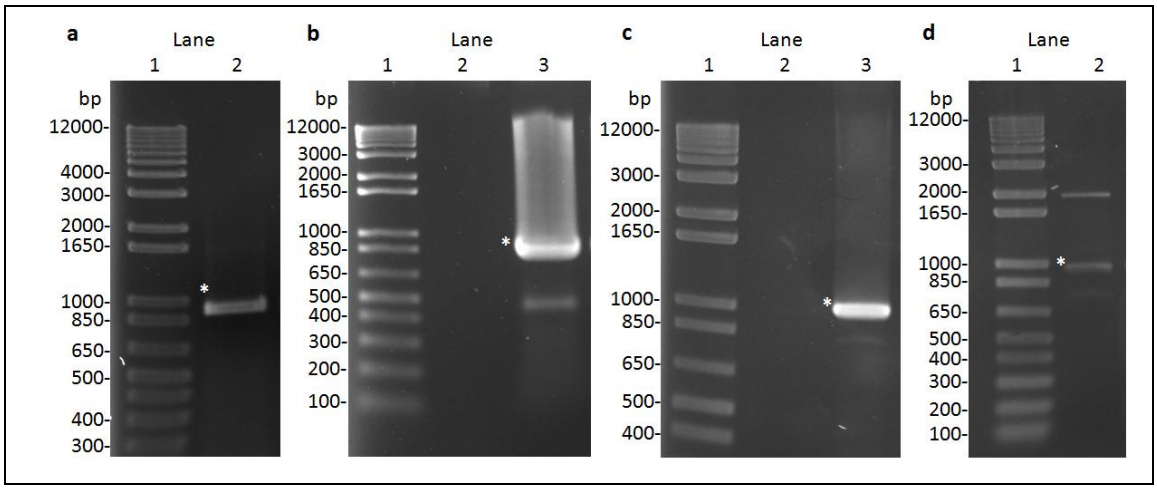


Figure 3.13 Agarose gel of high fidelity PCR products from *L. plantarum* gDNA

a) PCR product 14917_F1 indicated by an asterisk (lane 2). b) PCR product 14917_F2 indicated by the asterisk (lane 3). A smaller secondary PCR product is present which is likely the result of non-specific primer binding (lane 3). Lane 2 is empty. c) PCR product 8014_F1 indicated by the asterisk (lane 3). A smaller secondary PCR product is present which is likely the result of non-specific primer binding (lane 3). Lane 2 is empty. d) PCR product 8014_F2 indicated by the asterisk (lane 2). Multiple other PCR products were generated which are likely the result of non-specific primer binding (lane 2). Invitrogen™ 1 kb Plus DNA Ladder (Invitrogen™; Wilmington, DE, USA) is present in lane 1 of each gel.

3.5.5 Restriction enzyme linearization of pNZ5319

The multiple cloning site one (MCS1) of pNZ5319 has two unique restriction enzyme recognition sites, *Pme*I and *Xho*I, that are approximately 1.1-kbp upstream from a unique *Bgl*II recognition site present at multiple cloning site two (MCS2) (Appendix 12). To determine the best restriction enzyme to linearise pNZ5319 at MCS1, the insert site for 14917_F1 and 8014_F1, pNZ5319 purified from *E. coli* XL1-Blue (section 2.1.9) was subjected to restriction digestion (section 2.1.15) using *Pme*I, *Xho*I and *Bgl*II (New England Biolabs®; MA, USA) for 3 hours at 37°C (Figure 3.14). A control reaction to detect possible star activity of both restriction enzymes *Pme*I and *Xho*I was set up by incubating pNZ5319 with either enzyme overnight at 37°C. Agarose gel analysis (section 2.1.13) of pNZ5319 sequentially digested by *Xho*I and *Bgl*II showed that there was still a proportion of pNZ5319 that had not been fully digested in the 3 hour incubation (Figure 3.14a, Lane 5). This was not the case for the *Pme*I-*Bgl*II combination (Figure 3.14a, Lane 4) suggesting that pNZ5319 was cut by *Bgl*II but not *Xho*I. Star activity from either *Pme*I or *Xho*I was not detected in the overnight incubation (Figure 3.14a, Lane 2 and 3). Initial restriction enzyme digestion of pNZ5319 with *Pme*I and *Xho*I showed that compared to *Pme*I, XhoI did not efficiently linearise pNZ5319. Thus pNZ5319 was linearised for cloning by incubation with *Pme*I for 3 hours at 37^{o}C in NEBuffer 4, followed by 20 minutes at 65°C to inactivate the restriction enzyme. The reaction products were analysed by agarose gel electrophoresis (section 2.1.13) and the band corresponding to *Pme*I-linearised pNZ5319 was extracted and purified (section 2.1.13.6) then used for cloning (section 3.5.6) (Figure 3.14b).

3.5.6 Cloning 14917_F1 and 8014_F1 into *Pme*I linearised pNZ5319

The two F1 fragments (section 3.5.4) were each added to *Pme*I linearised pNZ5319 (section 3.5.5) in a 5:1 molar ratio and subjected to ligation (section 2.1.16). Reactions were incubated for 5-10 minutes as was a control reaction containing only *Pme*I linearised pNZ5319, which served to determine the efficiency of the ligation reaction. To improve cloning efficiency the ligation reaction was treated with 10 U of *Pme*I (New England Biolabs®; MA, USA) for 30 minutes at 37°C. Successful ligation of the insert into MCS1 resulted in the loss of the *Pme*I recognition site, whereas, in any re-circularised plasmid, the recognition site will be re-established. The final solution was used to transform *E. coli* XL1 cells (section 2.2.2). Cells were spread on LB agar plates (section 2.1.2.2) containing 17 µg mL^{-1} chloramphenicol and incubated at 37°C for 24 hours.

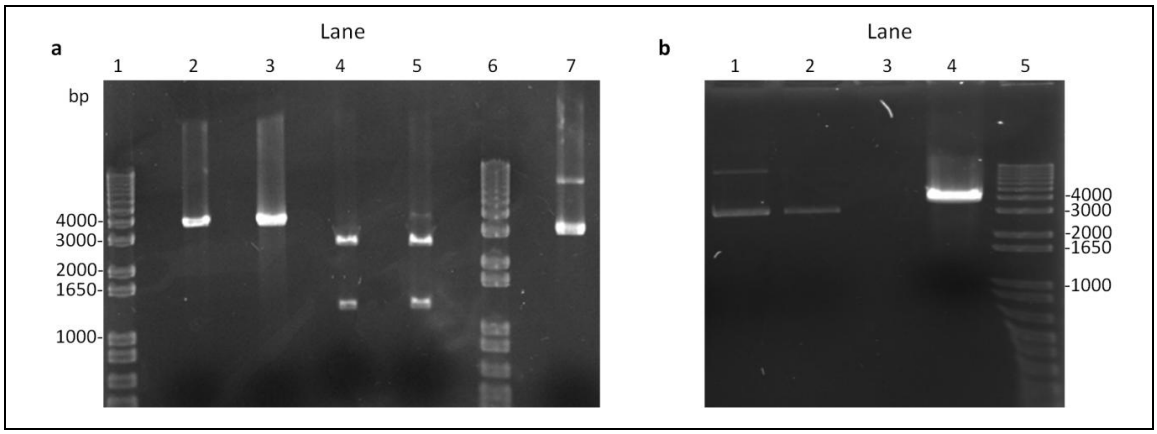


Figure 3.14 Restriction digestion of pNZ5319 for 8014_F1 and 14917_F1 cloning

a) Agarose gel of the products from overnight restriction enzyme digestion at MCS1 in pNZ5319 by *Xho*I (Lane 2) and *Pme*I (Lane 3). Products from a 3 hour digestion of pNZ5319 with both *Pme*I and *Bgl*II (Lane 4) and both *Xho*I and *Bgl*II (Lane 5). A control of undigested pNZ5319 is included (Lane 7). In lanes 1 and 6 is the Invitrogen™ 1 kb Plus DNA Ladder (Invitrogen™; Wilmington, DE, USA). b) Agarose gel with the control of pNZ5319 stock kept at -20°C (Lane 1) and a control of pNZ5319 treated the same as the digested plasmid but lacking any restriction enzyme (Lane 2). Products from a reaction of pNZ5319 digested with *Pme*I that was extracted and purified for cloning (Lane 4). In lane 5 is the Invitrogen™ 1 kb Plus DNA Ladder. Lane 3 is empty.

3.5.7 14917_F1 and 8014_F1 colony PCR screening

F1 orientation-dependent PCR (section 2.1.11) was carried out using a reverse primer that annealed to the plasmid backbone and a forward primer that annealed to the 5' end of the insert. 14917_F1 and 8014_F1 orientation-dependent PCR was carried out using primer pairs '85 and 81' and '85 and F8014_1' (Appendix 13), respectively. From the 14917_F1 cloning experiment, 18 randomly selected colonies were selected. Three colonies were identified: 12, 15 and 18, where 14917_F1 was in the correct orientation (Figure 3.15a). These colonies were streaked and the plasmid purified (section 2.1.9). Sequencing (section 2.1.17) across MCS1 using primers F1_seq and 85 was carried out using the purified plasmids in order to confirm that no nucleotides had mutated during the cloning procedure. This showed that plasmids isolated from both colonies 12 and 15 contained deletions in the introduced stop codon while the plasmid isolated from colony 15 had an additional deletion in the p32 promoter. Plasmid isolated from colony 18 had no nucleotide insertions or deletions and was thus designated as plasmid pNZ5319_14917_F1 (Appendix 14). *E. coli* EC100 cells (Appendix 15) were transformed with pNZ5319_14917_F1, incubated overnight (section 2.1.7.2) and the plasmid purified (section 2.1.9) before being used in restriction digestion experiments (section 3.5.8). From the 8014_F1 cloning experiment, 6 colonies were selected. Of these colony 2 was shown to contain 8014_F1

in the correct orientation (Figure 3.15b). Cells from this colony were streaked onto LB agar plates containing chloramphenicol (section 2.1.5), grown in culture overnight (section 2.1.7.2) and the plasmid purified (section 2.1.9) and then sequenced (section 2.1.17). Sequencing with primers 85 and F1_seq across the MCS1 of the purified plasmid confirmed that 8014_F1 had been incorporated correctly into MCS1 with the introduced stop codon present at amino acid position 12 as designed. This was designated as plasmid pNZ5319_8014_F1 (Appendix 14) and used for restriction digestion experiments (section 3.5.8).

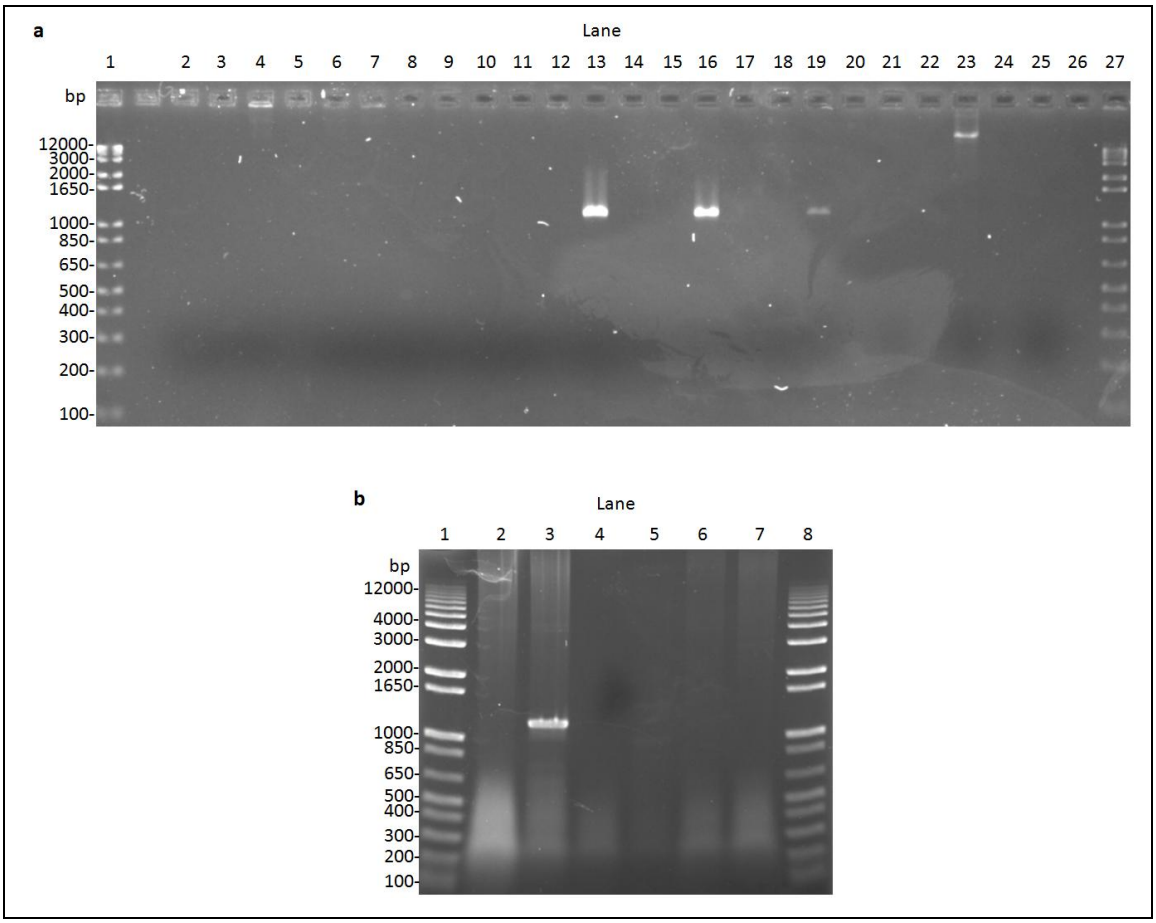


Figure 3.15 F1 Orientation dependent colony PCR

a) Agarose gel showing the products of 14917_F1 orientation-dependent PCR screening, primers 85 and 81,of 18 colonies transformed the plasmid DNA from the *Pme*I digested ligation reaction (Lanes 2-19, colony number is offset by 1). Negative controls are *E. coli* XL1 gDNA as template DNA (Lane 21), pNZ5319 as DNA template (Lane 23) and a reaction where the DNA template was substituted with ddH_2O (Lane 25). Lanes 20, 22, 24 and 26 are empty. b) Agarose gel containing the products of 8014_F1 orientation-specific PCR screening, primers 85 and F8014_1, of 6 colonies transformed with the plasmid DNA from the *Pme*I digested ligation reaction (Lanes 2-7, colony number is offset by 1). Negative controls as performed for 14917_F1 PCR were not performed. Invitrogen™ 1 kb Plus DNA Ladder (Invitrogen™; Wilmington, DE, USA) is present in the outermost lanes of each gel.

3.5.8 Restriction enzyme linearization of pNZ5319 containing F1

The MCS2 of pNZ5319 contains unique recognition sites for the blunt end restriction enzymes *Sma*I and *Ecl*136II (Appendix 12). MCS2 is approximately 1.1-kbp downstream from a unique *Pme*I recognition site, 1.3-kbp up stream of a unique *Eco*RV recognition site and 2.9-kbp downstream of a unique *Nde*I site introduced by 8014_F1. *Ecl*136II could not be used for cloning 8014_F2 as 8014_F1 contains an *Ecl*136II recognition site. No star activity was observed upon pNZ5319 digestion (section 2.1.15) by *Ecl*136II (Figure 3.16a). No star activity was detected when pNZ5319_F1_8014 was digested with *Sma*I but was detected when digested with *Eco*RV (Figure 3.16b). *Ecl*136II was selected for 14917_F1 cloning and *Sma*I for 8014_F2 after initial trial digestions showed that both enzymes cut efficiently. pNZ5319_14917_F1 and pNZ5319_8014_F1 were each subjected to *Ecl*136II or *Sma*I restriction digestion, then the products of digestion were analysed using agarose gel electrophoresis (section 2.1.13) before being purified (section 2.1.13.6) for cloning (section 3.5.9) (Figure 3.16c and Figure 3.16d).

3.5.9 Cloning 8014_F2 and 14917_F2

The F2 fragments (section 3.5.4) were added to *Sma*I or *Ecl*136II linearised plasmid respectively (section 3.5.8) in a 5:1 molar ratio and subjected to ligation (section 2.1.16). The reactions were incubated for 5-10 minutes at room temperature along with a control reaction containing digested plasmid with no insert to serve as an efficacy control. Both insert and insert with vector ligation reactions for 14917_F2 cloning were treated with 10 U of *Ecl*136II to prevent re-circularisation of the digested plasmid. *E. coli* XL1 cells were then transformed (section 2.2.2) with the products of the 14917_F2 ligation reactions and the transformed cells spread on LB agar plates containing 17 µg mL^{-1} chloramphenicol and incubated at 37°C for 24 hours. The 8014_F2 ligation reaction was not treated with *Sma*I as it was an initial trial ligation. These were not treated with restriction enzyme; if initial trial ligations failed to generate the correct product they were then later treated with restriction enzyme. *E. coli* EC100 cells were transformed with the products of the 8014_F2 cloning reactions and eventually spread on LB agar plates containing 17 µg mL^{-1} chloramphenicol and incubated at 37°C for 24 hours.

3.5.10 8014_F2 and 8014_F2 colony PCR screening

F2 orientation-dependent PCR was carried out using a forward primer annealing to the plasmid backbone and a reverse primer annealing to the 3' end of the insert. 14917_F2 (Figure 3.17a) and 8014_F2 (Figure 3.17b) orientation-dependent PCR was carried out using primer pairs '87

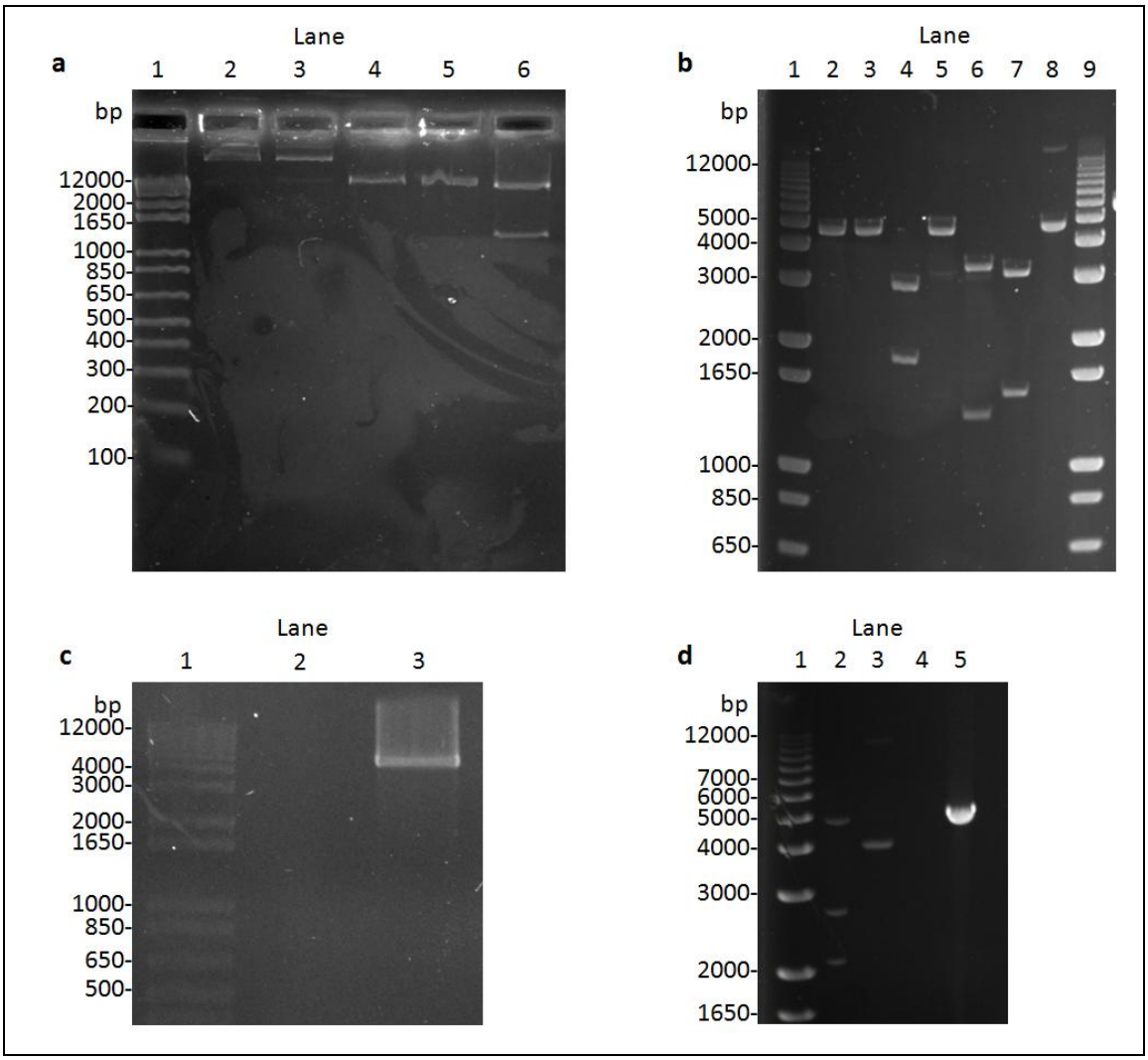


Figure 3.16 Restriction digestion for 8014_F2 and 14917_F2 cloning

a) Trial digestions of pNZ5319: lane 1, Invitrogen™ 1 kb Plus DNA Ladder; lane 2, pNZ5319 - 20°C control; lane 3, pNZ5319 treated using the same conditions as used to digest the plasmid but lacking restriction enzyme; lane 4, pNZ5319 treated with *Pme*I; lane 5, pNZ5319 treated with *Ecl*136II; lane 6, pNZ5319 treated with *Pme*I and *Ecl*136II. b) pNZ5319_8014_F1 trial digestions: lane 1, Invitrogen™ 1 kb Plus DNA Ladder; lane 2, pNZ5319_8014_F1 treated with *Sma*I; lane 3, pNZ5319_8014_F1 treated with *Nde*I; lane 4, pNZ5319_8014_F1 treated with *Sma*I and *Nde*I; lane 5, pNZ5319_8014_F1 treated with *Eco*RV; lane 6, pNZ5319_8014_F1 treated with *Sma*I and *Eco*RV; lane 7, pNZ5319_8014_F1 treated with *Nde*I and *Eco*RV; lane 8, pNZ5319_8014_F1 untreated control; lane 9, Invitrogen™ 1 kb Plus DNA Ladder. c) pNZ5319_14917_F1 digestion purified for cloning: lane 1, Invitrogen™ 1 kb Plus DNA Ladder; lane 2, empty; lane 3, pNZ5319_14917_F1 treated with *Ecl*136II; lane 3, Invitrogen™ 1 kb Plus DNA Ladder. d) pNZ5319_8014_F1 digestion purified for cloning: lane 1, . Invitrogen™ 1 kb Plus DNA Ladder; lane 2, pNZ5319_8014_F1 digested with both *Sma*I and *Sac*II showing poor efficiency of *Sac*II, a site introduced by 8014_F1; lane 3, untreated pNZ5319_8014_F1; lane 4, empty; lane 5, pNZ5319_8014_F1 treated with *Sma*I.

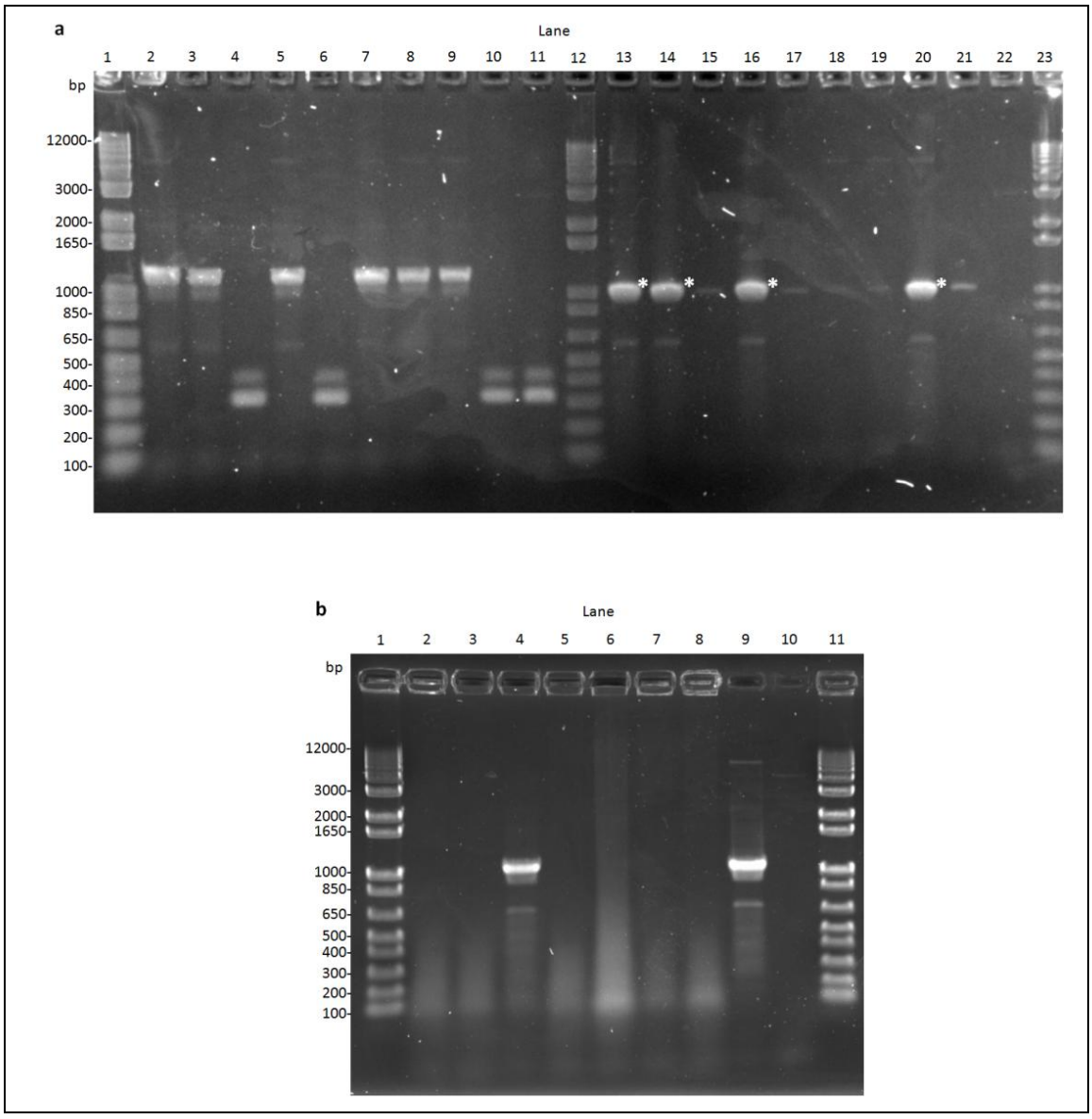


Figure 3.17 Colony PCR screening of F2 cloning

a) Agarose gel showing the products of 14917_F2 orientation-independent PCR screening of 9 colonies (lanes 2-10, colony number offset by 1). Orientation dependent PCR screening of 9 colonies (lanes 13-21, colony number offset by 12). pNZ5319 is a PCR DNA template negative control (lanes 11 and 22). b) Agarose gel showing the products of 8014_F2 orientation dependent PCR screening of 7 colonies (lanes 2-8, colony number offset by 1). pNZ5319_14917_F1_F2 was used as the DNA template for a positive control (lane 9) and pNZ5319 was used as the DNA template for a negative control (lane 10). Invitrogen™ 1 kb Plus DNA Ladder is present in the outermost lanes of each gel and lane 12 of (a).

and 83b' (Appendix 13). Orientation-independent PCR screening on 14917_F2 was carried out using primer pairs 82a and 83b (Figure 3.17a). Phusion High Fidelity Polymerase (Thermo Scientific; Wilmington, DE, USA) was used due to the lab stock of Roche® Taq DNA Polymerase being compromised. From the 14917_F2 reactions twelve colonies were analysed. Four colonies were shown to contain 14917_F2 in the correct orientation as indicated by a 1-kbp PCR product: colonies 1, 2, 4 and 8 (Figure 3.17a, Lanes 13-22). These colonies were streaked, cultured overnight (section 2.1.7.2), the plasmid purified (section 2.1.9) and then sequenced (section 2.1.17) across MCS2 using primers F2_seq and 87 to confirm no nucleotides had been mutated during the cloning procedure. Plasmid purified from colonies 1, 2 and 8 contained no nucleotide insertions or deletions and the plasmid isolated from colony 8 was designated as plasmid pNZ5319_14917_F1_F2 (Appendix 14). *E. coli* XL1 cells were transformed (section 2.2.2) with pNZ5319_14917_F1_F2, cultured, then the plasmid was purified (section 2.1.9) for *pts18CBA* knockout (section 3.7) and further restriction digestion experiments (section 3.6). From the 8014_F2 clones, seven colonies were analysed. One colony, colony 3, was shown to contain 8014_F2 in the correct orientation (Figure 3.17b). This colony was picked and streaked and the subsequent culture used for plasmid purification. Sequencing across MCS2 using primers F2_Seq and 87 was done to confirm the insert was correctly orientated before it was designated as pNZ5319_8014_F1_F2. *E. coli* EC100 cells were transformed with pNZ5319_8014_F1_F2 and the plasmid purified for *pts18CBA* knockout experiments. Thus, knockout plasmids for *pts18CBA* from both *L. plantarum* ATCC 8014 and *L. plantarum* subsp. *plantarum* ATCC 14917 were constructed from pNZ5319, which were later used with the aim to knockout *pts18CBA* from *L. plantarum* strains. Initial attempts to knockout *pts18CBA* with these plasmids failed, discussed in section 3.7, which prompted the construction of a knockout plasmid that was smaller in size (section 3.6).

3.6 Construction of a size reduced *pts18CBA* knockout plasmid

3.6.1 Introduction

In a series of experiments, chemically competent *L. plantarum* WCFS1 and *L. plantarum* subsp. *plantarum* ATCC 14917 were successfully transformed with pTRK669 while transformation with pNZ5319_14917_F1_F2 failed (section 3.7). pTRK669 is a small plasmid, 2.437-kbp, constructed by Russell and Klaenhammer (2001) for chromosomal site specific homologous recombination. The ease at which transformation was achieved with pTRK669 suggested that the difficulties experienced with transforming various *L. plantarum* cells with pNZ5319_14917_F1_F2 may be due to its relatively large size (5.459-kbp). A survey of the

literature also showed a negative correlation of plasmid size with transformation rate; transformation of several *Lactobacillus sake* strains by Berthier *et al.* (1996) showed a strong correlation between transformation rate and plasmid size, with lower transformation rates predominating when larger plasmids were used. This was not a clear pattern however, as some strains had an equal or even increased efficacy of transformation with larger plasmid. A more compelling example was published by Ohse *et al.* (1995) where *Bacillus subtilis* ISW1214 was transformed with a range of plasmids ranging from 2.9-kbp to 12.6-kbp. Although the differences in transformation efficacy between the 'smaller' and 'larger' plasmids were minor in comparison to the transformation efficacy of the individual plasmids they did show that the smaller the plasmid, the higher the transformation efficiency. A protocol was therefore devised to decrease the size of pNZ5319_14917_F1_F2 by removing a non-essential erythromycin resistance cassette (Appendix 12).

3.6.2 Aim

- To remove the erythromycin resistance coding sequence from pNZ5319_14917_F1_F2 to acquire a smaller pts18CBA knockout plasmid.

3.6.3 Restriction enzyme linearization of pNZ5319_14917_F1_F2

The erythromycin resistance cassette is flanked by two *Pst*I recognition sites (Appendix 12). Restriction enzyme digestion of pNZ5319_14917_F1_F2 with *Pst*I should thus generate two products: a 4.039-kbp fragment and a 1.099-kbp fragment containing the excised erythromycin resistance cassette. pNZ5319_14917_F1_F2 purified from *E. coli* XLI-Blue (section 2.1.9) was subjected to restriction digestion (section 2.1.15) using Roche® *Pst*I. The products of the digestion were analysed using agarose gel electrophoresis (section 2.1.13) and the band corresponding to digested pNZ5319_14917_F1_F2 (4.039-kbp) was excised and purified (section 2.1.13.6) (Figure 3.18).

3.6.4 Cloning of *Pst*I digested pNZ5319_14917_F1_F2

Purified digested plasmid (section 3.6.3) was re-ligated (section 2.1.16) and *E. coli* XLI were transformed with the ligation reaction (section 2.2.2). These cells were spread onto chloramphenicol LB agar plates and incubated for 24 hours at 37°C. 32 bacterial colonies were spotted onto LB agar plates containing either chloramphenicol or erythromycin (section 2.1.5) and incubated for 48 hours at 37°C. All colonies grew in the presence of chloramphenicol but only 2 grew in the presence of erythromycin; 10 erythromycin-sensitive colonies were picked from the corresponding chloramphenicol culture and streaked onto fresh chloramphenicol LB

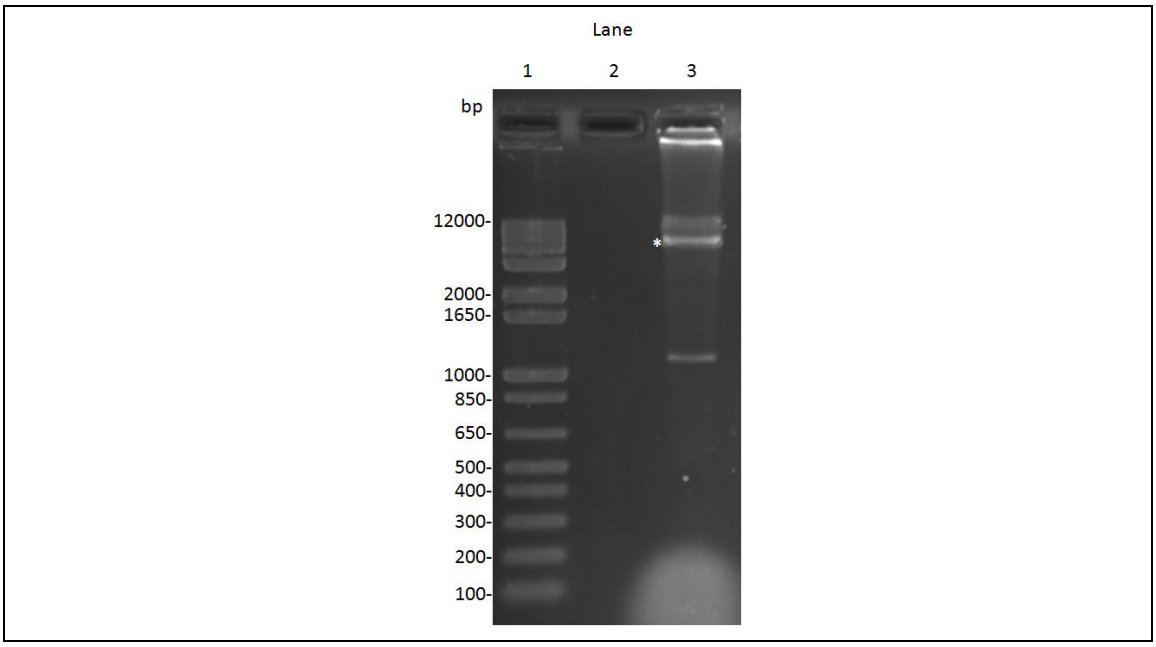


Figure 3.18 Restriction enzyme digestion of pNZ5319_14917_F1_F2

Lane 1, Invitrogen™ 1 kb Plus DNA Ladder; lane 2, empty; lane 3, 5 µg of pNZ5319_14917_F1_F2 treated with 5 U of Roche® *Pst*I for 1 hour at 37°C. The band that corresponds to *Pst*I digested pNZ5319_14917_F1_F2 is highlighted by the asterisk.

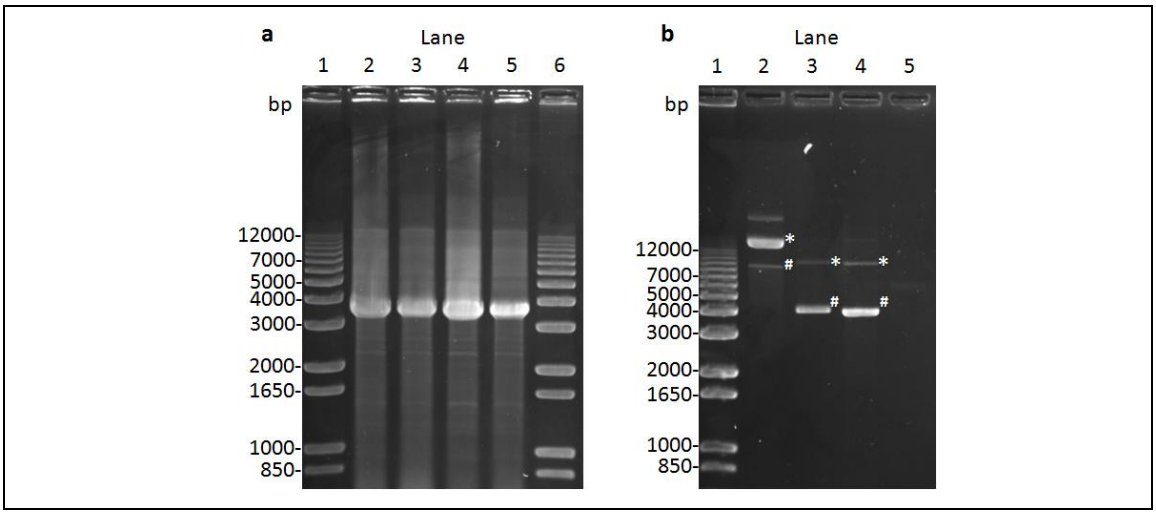


Figure 3.19 Agarose gel electrophoresis of purified plasmid and colony PCR screening

a) Agarose gel of PCR products of colony screening of the 4 colonies (lane 2-5, colony number offset by 1). Invitrogen™ 1 kb Plus DNA Ladder (Invitrogen™; Wilmington, DE, USA) is loaded in lanes 1 and 6. b) Agarose gel of purified plasmid of the 4 colonies (lanes2-5, colony number offset by 1). Plasmid in the open circular form is indicated with an asterisk and plasmid in the supercoiled form is indicated by the hashtag. Invitrogen™ 1 kb Plus DNA Ladder is loaded in lane 1.

agar plates, then incubated for 24 hours at 37°C. Isolated colonies were again streaked on to chloramphenicol and erythromycin LB agar and incubated as previously described. Plasmids from 4 colonies were subjected to PCR (section 2.1.11) using primers 85 and 87, which amplify over the erythromycin cassette (Appendix 12), and agarose gel electrophoresis (section 2.1.13) to confirm the absence/presence of the erythromycin cassette. The presence of a 3.5-kbp rather than a 4.8-kbp PCR product confirmed the absence of the erythromycin cassette in all 4 colonies (Figure 3.19a).

These colonies were streaked, and the plasmids purified from resulting cultures were analysed using agarose gel electrophoresis, which showed that plasmids from colonies 1 and 3 were not the expected size (Figure 3.19b). The predicted size of pNZ5319_14917_F1_F2_Emr- was 4.4-kbp. The results shown in Figure 3.19b suggest that the ligation reaction may have generated concatemerised forms of the plasmid, as the plasmid isolated from colony 1 was almost double the expected size. Figure 3.19b showed that plasmid purification from colony 4 failed and that plasmid from colony 2 was approximately 4.4-kbp. This plasmid was therefore designated pNZ5319_14917_F1_F2_Emr-. It was hoped that this smaller knockout plasmid would have a higher efficacy in the transformation experiments (section 3.7). Unfortunately, this was not the case as only a single *pts18CBA* knockout cell line was generated using the larger knockout plasmid pNZ5319_14917_F1_F2 (section 3.7.3).

3.7 Transformation of *L. plantarum* with *pts18CBA* knockout plasmids

3.7.1 Introduction

The multitude of methods that swarm the literature designed with the intent to transform Gram-positive bacteria with foreign DNA make use of the phenomena where an electric field can bring about the formation of nano-sized pores in lipid membranes (Weaver and Chizmadzhev, 1996, Tieleman, 2004). Known as electroporation, this technique is used in the majority of methods that are optimised for *Lactobacillus* species. The following publications describe methods to achieve chemical competence and transformation: Palomino *et al.* (2010), Lambert *et al.* (2007), Alegre *et al.* (2004), Berthier *et al.* (1996), Jin *et al.* (2012), Aukrust and Blom (1992) and Papagianni *et al.* (2007). The methods from these publications were appropriately adapted (sections 2.2.1 and 2.2.2) to accomplish the goal of transforming *L. plantarum* subsp. *plantarum* ATCC 14917, *L. plantarum* WCFS1, *L. plantarum* ATCC 8014 and *L. plantarum* NC8 with knockout plasmids (sections 3.5 and 3.6) along with control plasmids pTRK669 and pRV613 (Appendix 14) which contain origin of replications that replicate in *Lactobacillus* species (Russell and Klaenhammer, 2001, Crutz-Le Coq and Zagorec, 2008).

3.7.2 Aim

- To generate chemically-competent *L. plantarum* subsp. *plantarum* ATCC 14917, *L. plantarum* WCFS1, *L. plantarum* ATCC 8014 and *L. plantarum* NC8.
- To transform chemically-competent *L. plantarum* with *pts18CBA* knockout plasmids using electroporation.
- To screen for transformants and identify *L. plantarum* that have undergone homologous recombination with a knockout plasmid to replace *pts18CBA* coding sequence with the chloramphenicol resistance selectable marker.

3.7.3 Results and discussion

During the initial experiments following the method adapted by Palomino *et al.* (2010) (sections 2.2.3.1/2.2.4.1) and Lambert *et al.* (2007) (sections 2.2.3.2/2.2.4.2) a significant amount of electrical discharge inside the cuvette was observed during electroporation. This electrical discharge was only present in cuvettes that had had plasmid present suggesting that the plasmid solution contained a high concentration of salts which was increasing the conductivity of the solution and leading to the electrical discharge. As a solution to this problem, the plasmid was de-salted using drop dialysis (section 2.1.10). After this, electrical discharge did not occur during electroporation using identical conditions as were used in earlier experiments. After this success, plasmid were purified prior to electroporation using a methods adapted from the plasmid purification procedure (section 2.1.9). This involved washing the plasmid with EMSURE® 80 % ethanol which resulted in a significant reduction of the osmolarity of the plasmid solution so that drop dialysis was no longer required on future plasmid solutions. The methods of Palomino *et al.* (2010) (sections 2.2.3.1/2.2.4.1) and Lambert *et al.* (2007) (sections 2.2.3.2/2.2.4.2) did not yield any successful transformations. The following protocols also failed to generate successful transformations: Alegre *et al.* (2004) and Berthier *et al.* (1996) (sections 2.2.3.3/2.2.4.3), Jin *et al.* (2012) (sections 2.2.3.4/2.2.4.4), Aukrust and Blom (1992) (sections 2.2.3.5/2.2.4.5) and Papagianni *et al.* (2007) (sections 2.2.3.7/2.2.4.7).

On the MRS agar plates from iteration 1 of the protocol adapted from Aukrust and Blom (1992) and Lambert *et al.* (2007) (sections 2.2.3.6.a/2.2.4.6.a) (Table 2.13) bacterial colonies resembling *L. plantarum* were identified and counted (Table 3.7). From one of the three plates originating from cuvettes 1, 3, 4, 5 and 8, a single colony was streaked onto chloramphenicol (7 µg mL^{-1}) MRS agar plates which were incubated overnight at 37°C. PCR screening (section 2.1.11) with primer pair TRKCmSeqR3 and TRKCmSeqF2 (Appendix 13) that amplifies a 1.5-kbp product from within the chloramphenicol acetyltransferase gene of pTRK669, was used to

determine if the isolated *L. plantarum* WCFSI and *L. plantarum* subsp. *plantarum* 14917 colonies were transformed with pTRK669. Agarose gel electrophoresis (section 2.1.13) of the PCR products showed that all the isolated bacterial colonies had been successfully transformed (Figure 3.20a). The outcome of this experiment indicates that high concentrations of $MgCl_2$, in the cell wash buffer, hinder transformation efficiency for *L. plantarum* WCFS1. However because the only *L. plantarum* subsp. *plantarum* ATCC 14917 transformant obtained was from cuvette 8, which had a $MgCl_2$ concentration of 100 mM in the cell wash buffer, this concentration of $MgCl_2$ was used for later iterations following this protocol, all of which failed to generate knockouts. The 100 mM $MgCl_2$ concentration in the cell wash buffer however was included into the subsequent protocols adapted from Jin *et al.* (2012), Aukrust and Blom (1992) and Lambert *et al.* (2007) (sections 2.2.3.8/2.2.4.8) which did yield knockouts, as is discussed next.

Table 3.7 Transformation rates from Aukrust and Blom (1992) and Lambert *et al.* (2007) (2.2.3.6.a/2.2.4.6.a/Table 2.13)

		Replicate MRS agar plate			
Cuvette	Time constant (ms)	1	2	3	CFU μg^{-1}
1	9.1	>1000	>1000	>1000	>6000
2	9.2	0	0	0	0
3	9.1	>1000	>1000	>1000	>6000
4	8.5	90	75	74	1614.21
5	8.9	10	11	2	155.34
6	9.2	0	0	0	0
7	9.2	0	0	0	0
8	9.1	1	0	0	6.75
9	8.6	0	0	0	0
10	8.8	0	0	0	0
11	8.5	0	0	0	0
12	8.3	0	0	0	0
13	8.2	0	0	0	0
14	8.8	0	0	0	0
15	8.8	0	0	0	0
16	7.9	0	0	0	0

On the transformation plates from the protocol adapted from Jin *et al.* (2012) and Aukrust and Blom (1992) and Lambert *et al.* (2007) (sections 2.2.3.8/2.2.4.8), detailed in Table 2.18, bacterial colonies resembling *L. plantarum* were identified and counted (Table 3.8). Four colonies were selected from the MRS agar plates spread with bacterial cells from cuvette one and three colonies were selected from the MRS agar plates spread with bacterial cells from cuvette two. These colonies were individually streaked onto chloramphenicol MRS agar plates

and incubated for 48 hours at 30°C. From these new plates, a single colony was streaked onto chloramphenicol MRS agar plates. PCR screening of these colonies was undertaken using the primer pair CAT_F and CAT_R (Appendix 13) which bind inside the chloramphenicol acetyltransferase gene amplifying a 511-bp product. A single colony, colony 6, originating from an agar plate spread with bacteria from cuvette two tentatively contained the knockout plasmid as indicated by a 511-bp PCR product (Figure 3.20b; lane 7).

Table 3.8 Transformation rates from Jin *et al.* (2012), Aukrust and Blom (1992) and Lambert *et al.* (2007) (2.2.3.8/2.2.4.8/Table 2.18)

Cuvette	Time Constant (ms)	Replicate MRS agar plate			CFU μg^{-1}
		1	2	3	
1	8.9	5	2	3	26.72
2	9	6	3	12	91.46
3	9	0	0	0	0
4	9	0	0	0	0
5	9	0	0	0	0
6	9	0	0	0	0

Colony 6 was *L. plantarum* NC8 that was transformed with pNZ5319_14917_F1_F2_Emr-. Subsequent colony PCR analysis on colony 6 failed to replicate these positive results of Figure 3.20b. It was suspected that the subsequent PCR failures were due to difficulty obtaining DNA template from the bacterial colony. To solve this, an overnight culture was inoculated with colony 6, glycerol stocks made, and the gDNA was purified (section 2.3.1). The presence of the *cat* gene in colony 6 was confirmed by the amplification of a 500-bp product from the purified gDNA using primer pair CAT_F and CAT_R (Figure 3.20d). Colony 6 gDNA was used as a DNA template for PCR with primer pair P1R and PTS18CBA_START; primers P1R and PTS18CBA_START respectively bind to the 3' and 5' end of the *pts18CBA* coding region. PCR amplification with these primers across the chloramphenicol gene originating from the knockout plasmid is predicted to produce a 1.5-kbp band while PCR amplification across the native *pt18CBA* coding sequence is predicted to produce a 1.9-kbp band. Confusingly, both of these products were amplified from the gDNA of colony 6 (Figure 3.20e) while control *L. plantarum* NC8 gDNA only produced the 1.9-kbp band (Figure 3.20f, Lane 3). These results suggest homologous recombination with *pts18CBA* (Locus tag: nc8_2163) had not occurred and that either homologous recombination had occurred at another location in the genome or that the plasmid had undergone a recombination event giving rise to an ability to replicate in *L. plantarum*. PCR colony screening was attempted multiple times with the remaining colonies but failed to produce results suggesting homologous recombination had not occurred.

On the MRS agar plates from the iteration of the protocol (sections 2.2.3.8.a/2.2.4.8.a), detailed in Table 2.19, adapted from Jin *et al.* (2012) and Aukrust and Blom (1992) and Lambert *et al.* (2007) that had been used before (sections 2.2.3.8/2.2.4.8), bacterial colonies resembling *L. plantarum* were identified and counted (Table 3.9). A number of colonies from these plates, except for plates originating from cuvette 10 as no primers were currently available for PCR screening of that particular plasmid, were streaked out onto selective MRS agar plates and incubated for 48 hours at 37°C. PCR screening of these colonies was undertaken with primer pair CAT_F and CAT_R which in addition to being able to amplify a product from pNZ5319 can also amplify a 511 bp product from pTRK669. A single colony, colony 4, originating from an agar plate spread with bacteria from cuvette four was shown to possibly contain the knockout plasmid (Figure 3.20c). Two colonies, 11 and 13, were shown to have been successfully transformed with pTRK669 as indicated by the 511 bp PCR product (Figure 3.20c).

Figure 3.20 PCR screening of *L. plantarum* for transformation

a) Agarose gel containing PCR products of primer pair TRKCmSeqR3 and TRKCmSeqF2 from screening of *L. plantarum* colonies from experiments 2.2.3.6.a/2.2.4.6.a (Table 2.13 and Table 3.7). PCR products from colonies 1-7 (lanes 2-8). Invitrogen™ 1 kb Plus DNA Ladder is present in lane 1. b) Agarose gel containing the PCR products of primer pair CAT_F and CAT_R from colony screening of colonies isolated from experiment 2.2.3.8/2.2.4.8 (Table 2.18 and Table 3.8) (Lanes 2-8). PCR products of colonies 1-4 (lane 2-5) from cuvette 1 and colonies 5-7 (lanes 6-8) from cuvette 2. Invitrogen™ 1 kb Plus DNA Ladder is present in lane 1. c) Agarose gel containing PCR products of primer pair CAT_F and CAT_R from colony screening of colonies isolated from iteration 1 of experiment 2.2.3.8.a/2.2.4.8.a (Table 2.19 and Table 3.9) (Lanes 1-20 and 22-31). PCR products of colonies: 1-3 (Lanes 1-3) from cuvette 3, 4-10 (Lanes 4-10) from cuvette 4, 11-13 (Lanes 11-13) from cuvette 5, 14 & 15 (Lane 14 & 15) from cuvette 7, 16-18 (Lanes 16-18) from cuvette 9, 19-24 (Lane 19-20 & 23-26) from cuvette 11 and 25-29 (Lanes 27-31) from cuvette 12. Invitrogen™ 1 kb Plus DNA Ladder is present in lane 1. d) Agarose gel containing PCR product of primer pair CAT_F and CAT_R from colony 6 gDNA (Lane 2). Invitrogen™ 1 kb Plus DNA Ladder is present in lane 1. e) Agarose gel containing PCR product of primer pair P1R and PTS18CBA_START from colony 6 gDNA (Lane 2). Invitrogen™ 1 kb Plus DNA Ladder is present in lane 1. f) Agarose gel containing PCR products of primer pairs CAT_F and CAT_R (Lanes 2 and 4) and P1R and PTS18CBA_START (Lanes 3 and 5) from *L. plantarum* NC8 gDNA (Lanes 2-3) and colony 4 gDNA (Lanes 4-5). Invitrogen™ 1 kb Plus DNA Ladder is present in lanes 1 and 6. g) Agarose gel containing PCR product of primer pair 81 and 83b from *L. plantarum* NC8 gDNA (Lane 2) and colony 4 gDNA (Lane 3). Invitrogen™ 1 kb Plus DNA Ladder is present in lanes 1 and 4.

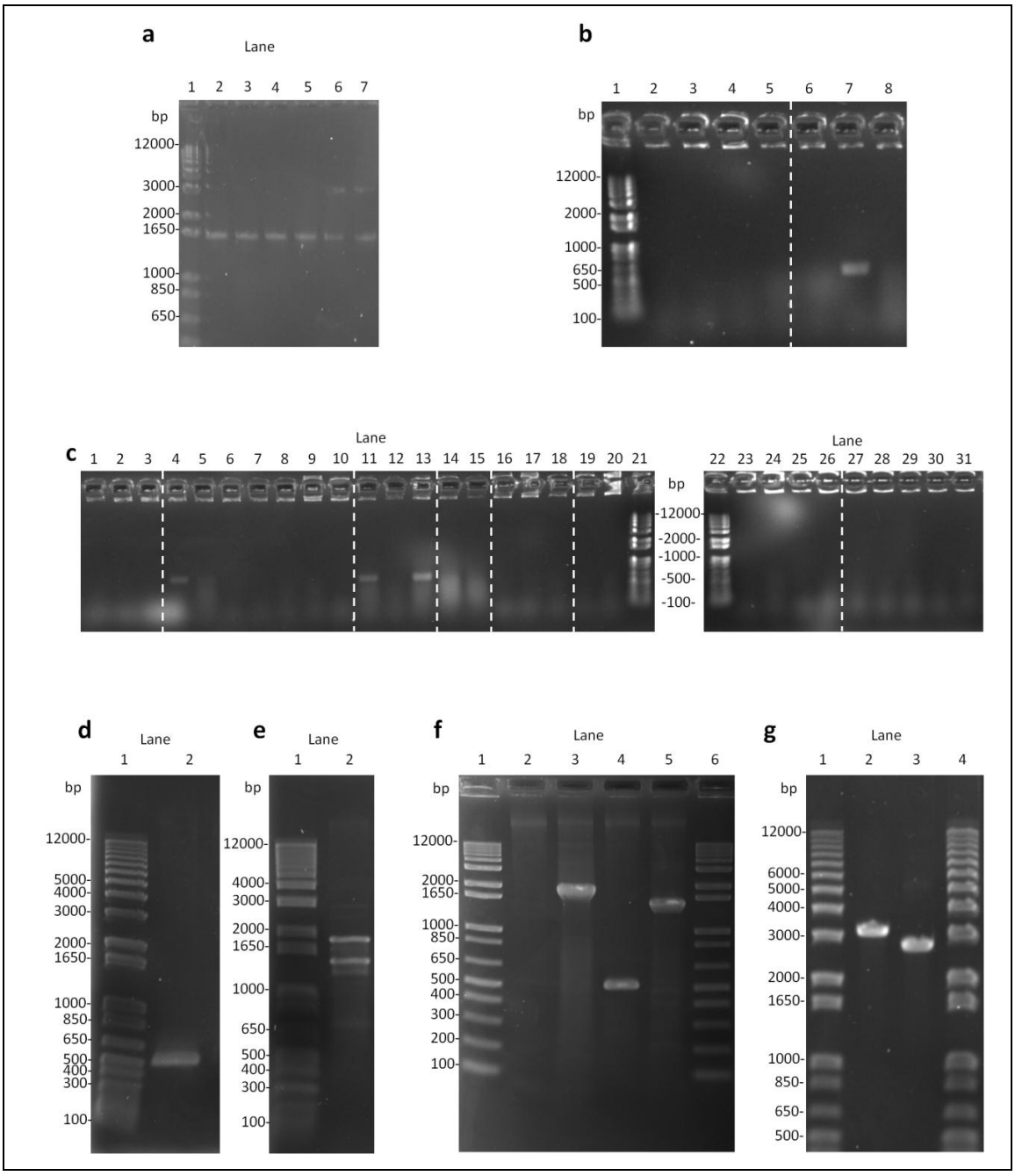


Figure 3.20 PCR screening of *L. plantarum* for transformation

Table 3.9 Transformation rates from Jin *et al.* (2012), Aukrust and Blom (1992) and Lambert *et al.* (2007) (2.2.3.8.a/2.2.4.8.a/Table 2.19)

		MRS agar plate			
Cuvette	Time Constant (ms)	1	2	3	CFU μg^{-1}
1	8.6	215	215	258	2379.87
2	8.6	12	15	3	150.69
3	8.0	1	1	1	5.50
4	7.9	1	1	5	25.67
5	8.7	22	40	58	415.09
6	8.7	64	29	10	517.36
7	8.3	1	1	0	7.33
8	8.2	0	0	0	0
9	8.9	16	5	43	221.38
10	9.0	1	1	0	10.046
11	8.2	6	9	3	33
12	8.3	3	10	0	23.83

Colony 4 was *L. plantarum* NC8 that had been transformed with pNZ5319_14917_F1_F2. Subsequent PCR analysis on cell isolated from colony 4 failed to replicate the results shown in Figure 3.20c so an overnight culture was inoculated with colony 4, glycerol stocks made, and the gDNA was purified. This was used as a template for PCR using primer pair CAT_F and CAT_R and pair P1R and PTS18CBA_START, which was analysed by agarose gel electrophoresis (Figure 3.20f). Amplification using P1R and PTS18CBA_START from *L. plantarum* NC8 gDNA produced the expected 1.9-kbp product, colony 4 gDNA produced a 1.5-kbp product indicating homologous recombination had occurred (Figure 3.20f). The region of the gDNA encompassing the *pts18CBA* coding sequence targeted for homologous recombination was amplified by high fidelity PCR (section 2.1.12) from *L. plantarum* NC8 gDNA and colony 4 gDNA using primer pair 81 and 83b. The PCR products were analysed using agarose gel electrophoresis and the bands purified (Figure 3.20g). The PCR products amplified from *L. plantarum* NC8 and colony 4 were respectively 3.1-kbp and 2.9-kbp which coincides with that observed in Figure 3.20f. They were both sequenced (section 2.1.17) with primers listed in Table 3.10, then comparatively analysed using Geneious™ software (Figure 3.21). Sequencing confirmed that homologous recombination had occurred between the genome of *L. plantarum* NC8 and the knockout plasmid pNZ5319_14917_F1_F2 to generate *L. plantarum* NC8 Δ*pts18CBA*.

Table 3.10 ***L. plantarum*** **NC8** ***pts18CBA*** **sequencing primers**

82a	GCGGATGAGACGTTTGCTGG
83b	CAGTCCCGTTGCTTCTGGATAAATC
87	GCCGACTGTACTTTCGGATCCT
P1R	GCTAGTCACATGACCAGTTG
CAT_F	TACCGAAACATAAAACAAGAAGGA
CAT_R	CTGACAATTCCTGAATAGAGTTCA
PTS18CBA_START	ATGAAGACATATTTTCAGAAAATC
85	GTTTTTTTCTAGTCCAAGCTCACA
80_term	TTACCGCAGCATCAATGACTGACC
81	AGGACCAGTACGTTTTCGCCGC

Sequencing showed there was a 43-bp insertion in the middle of the P_{32} promoter which NCBI BLASTn identified as "*Streptococcus cremoris* promoter 32 DNA fragment" (E value = $2x10^{-14}$). pNZ5319 is a plasmid that has been developed from many constructions. It was constructed from the plasmid pGIZ850, which had obtained its P_{32} element from the plasmid pMG36e (Goffin *et al.*, 2004) which obtained its P_{32} element from pGKV432 (van de Guchte *et al.*, 1989), which is a derivative of pGKV232 (van der Vossen *et al.*, 1987), which was the result of a promoter fragment P_{32} isolated from *Streptococcus cremoris* Wg2 chromosomal DNA being cloned into the promoter selection vector pGKV210 (van der Vossen *et al.*, 1987). Most likely pNZ5319 contains this 43-bps in its P_{32} element and the pNZ5319 sequence (GenBank AI: DQ104847.1) submitted to NCBI GenBank by Lambert *et al.* (2007) is incorrect. Sequencing also identified a di-nucleotide deletion in the middle of the primer 87 binding site, 24 nucleotides downstream of the *cat* stop codon. Due to its location this deletion is predicted to have little or no effect on surrounding elements.

This series of experiments (section 3.7.3) demonstrated the difficulty of generating chemically competent *L. plantarum*, transformation of the chemically competent *L. plantarum* with a plasmid and eventually the generation of a *pts18CBA* knockout cell line. Both the aims of achieving chemically competent *L. plantarum* and transformation of these with knockout plasmid (section 3.7.2) was achieved using protocols detailed in sections 2.2.3.6.a/2.2.4.6.a (page 35/41), sections 2.2.3.8/2.2.4.8 (page 36/44) and sections 2.2.3.8.a/2.2.4.8.a (page 37/45). The final aim, to identify a *pts18CBA* knockout *L. plantarum*, was achieved with *L. plantarum* NC8 using the protocol detailed in sections 2.2.3.8.a/2.2.4.8.a. In the early transformation experiments *L. plantarum* NC8 was not originally targeted for *pts18CBA* knockout. The original target strains were *L. plantarum* ATCC 8014 and *L. plantarum* subsp. *plantarum* ATCC 14917 as these were the established glycocin F indicator strains in the laboratory. However the inability to transform these two strains with a plasmid prompted the trial with both *L. plantarum*

WCFS1 and *L. plantarum* NC8 alongside *L. plantarum* ATCC 8014 and *L. plantarum* subsp. *plantarum* ATCC 14917. These trials were feasible because of a high degree of conservation in the flanking regions of the *pts18CBA* gene that were cloned into the knockout plasmids. The low transformation efficacy of all four *L. plantarum* strains was thought to be due to the thick cell wall of these bacterial species. The relatively high transformation rates achieved using protocol detailed in sections 2.2.3.8.a/2.2.4.8.a (Table 3.9) was probably due to the addition of ampicillin (0.1 µg mL^{-1}) in the growth media, which weakened the cell wall. Another factor that was believed to have had a deleterious effect was the presence of natural plasmids in *L. plantarum* ATCC 8014, *L. plantarum* subsp. *plantarum* ATCC 14917 and *L. plantarum* WCFS1. *L. plantarum* NC8 harbours no plasmids (Axelsson *et al.*, 2012) while the other three stains each harbour multiple plasmids (Leer *et al.*, 1992, van Kranenburg *et al.*, 2005) (Appendix 3 shows *L. plantarum* subsp. *plantarum* ATCC 14917 harbouring multiple plasmids).

Figure 3.21 DNA sequence alignment of *expected L.* plantarum NC8 Δpts18CBA to sequenced colony 4 gDNA.

DNA alignment of the expected *L. plantarum* NC8 Δ*pts18CBA* to sequenced colony 4 gDNA using Geneious™ software. A detailed version of this alignment is provided in Appendix 16. The large shaded yellow box is the expected *L. plantarum* NC8 Δ*pts18CBA* sequence with nucleotide identities indicated by colour (A:red, T:green, C:blue, G:yellow) and its respective annotations. gDNA and the dark grey annotation is sequence from pNZ5319_14917_F1_F2 after the expected recombination event. Rose pink annotations are primer binding sites with direction of amplification indicated by an arrow. Magenta annotations indicate flanking regions cloned into pNZ5319. Yellow annotations represent gene coding sequences. Dark blue annotations are the background *L. plantarum* NC8 Below this are the eleven colony 4 sequencing reads. The blue shading represents the relative quality of base call with dark blue indicating low quality and light blue indicating high quality. Those that were of poor quality, not allowing base to be called, were trimmed and are not shown. On the left are the names of the primers that were used to generate the corresponding sequence read, the direction of the sequence read in relation to the expected knockout sequence (FWD: forward and REV: reverse) and a numbering system for identification (2-12). The mutation in the colony 4 gDNA, downstream of the *cat* gene, can be identified at position 1,130 in sequence reads 5 and 6. The mutation in the colony 4 gDNA p32 promoter can be identified at position 1,900 in sequence reads 8 and 9.

Figure 3.21 DNA sequence alignment of expected *L. plantarum* NC8 Δ*pts18CBA* to sequenced colony 4 gDNA.

3.8 *L. plantarum* NC8 Δ*pts18CBA* glycocin F characterisation

3.8.1 Introduction

Biochemical characterisation of *L. plantarum* NC8 Δ*pts18CBA* was accomplished using both a plate assay (section 2.2.7.3) and liquid culture assay (section 2.2.7.2). Prior to characterisation of *L. plantarum* NC8 Δ*pts18CBA*, the glycocin F minimum inhibition concentration (MIC) (section 2.2.7.1) was determined. This was done so that the response to glycocin F at either end of the inhibitory scale could be measured for *L. plantarum* NC8 Δ*pts18CBA*.

3.8.2 Aims

- To determine a glycocin F MIC for *L. plantarum* NC8.
- To characterise the effect glycocin F has on *L. plantarum* NC8 Δ*pts18CBA*.

3.8.3 Results and discussion

From triplicate glycocin F assays carried out in liquid culture (section 2.2.7.1), the MIC for *L. plantarum* NC8 was determined to be 50 nM (Table 3.3 and Figure 3.22). This concentration was used in liquid culture glycocin F sensitivity assays (sections 2.2.7.2 and 2.2.7.3). Plate assays comparing *L. plantarum* NC8 Δ*pts18CBA* and wild-type NC8 as indicator strain showed that at similar concentrations of glycocin F, the amount of clearing was significantly less for the knockout strain (Figure 3.23). However due to the heavy bacterial inoculum and lengthy incubation time, the results from the plate assays were not as convincing as those seen for the naturally selected glycocin F resistant *L. plantarum* mutants (Figure 3.3). Time constraints did not allow troubleshooting. Liquid culture glycocin F assays using *L. plantarum* NC8 Δ*pts18CBA* as the targeted strain showed that at the MIC, inhibition of growth decreased from 93 % to 50 %, and from 95 % to 66 % at 1 µM glycocin F convincingly demonstrating that knocking out *pts18CBA* confers a degree of glycocin F resistance to *L. plantarum* NC8 (Figure 3.24).

These glycocin F sensitivity assay results are consistent with the results obtained for the Group H glycocin F resistant mutants (Table 3.6), characterised in section 3.2. Mutations shown to disrupt PTS18CBA conferred complete resistance to glycocin F for *L. plantarum* ATCC 8014 and a very high level of glycocin F resistance for *L. plantarum* subsp. *plantarum* ATCC 14917 strongly indicating that PTS18CBA is the physiological receptor of glycocin F. As seen for glycocin F resistant *L. plantarum* subsp. *plantarum* ATCC 14917 mutants (Figure 3.4, Figure 3.5 and Figure 3.3a), complete resistance was not obtained upon knocking out PTS18CBA,

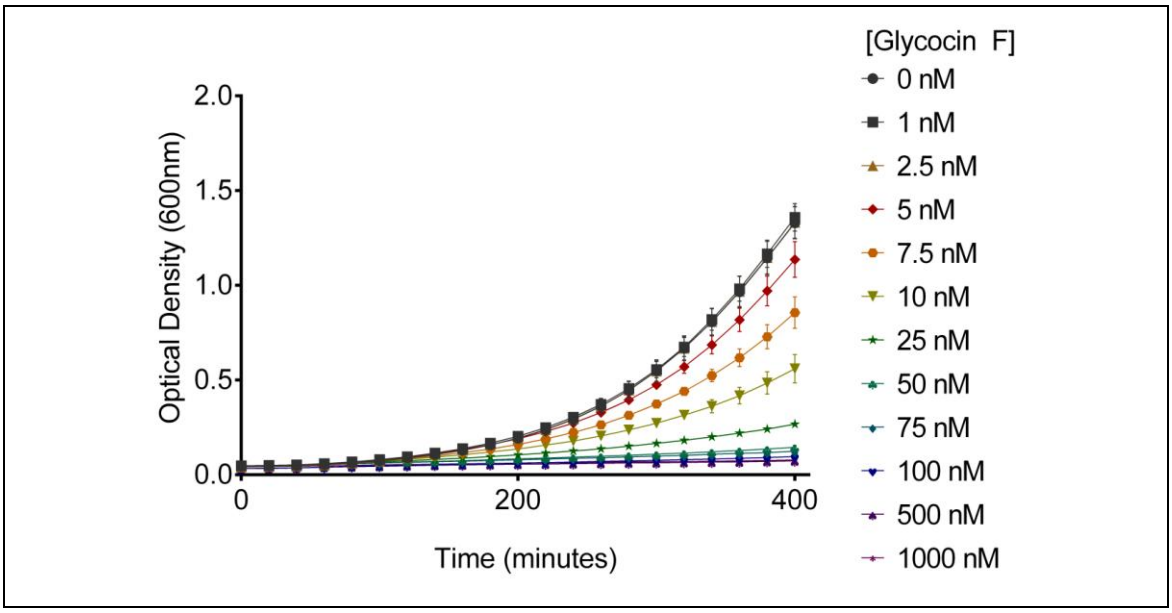


Figure 3.22 Liquid glycocin F assay of *L. plantarum* NC8 for MIC determination

The mean OD_{600nm} from triplicate assays. Vertical error bars represent one standard deviation from the mean. Legend to the right indicates the concentration of glycocin F added at time point 0.

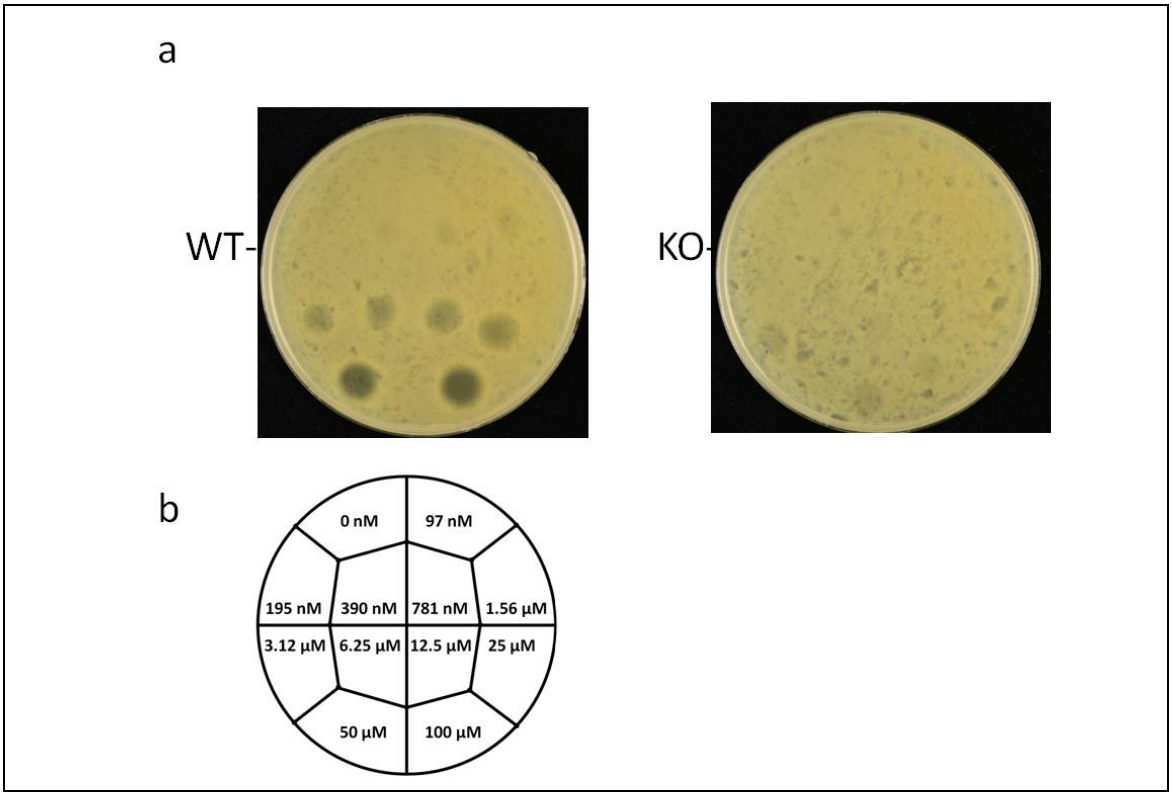


Figure 3.23 *L. plantarum* solid glycocin F assays

a) MRS agar plates from solid glycocin F assay of *L. plantarum* NC8 (WT) and *L. plantarum* NC8 Δ*pts18CBA*. b) Solid glycocin F plate layout.

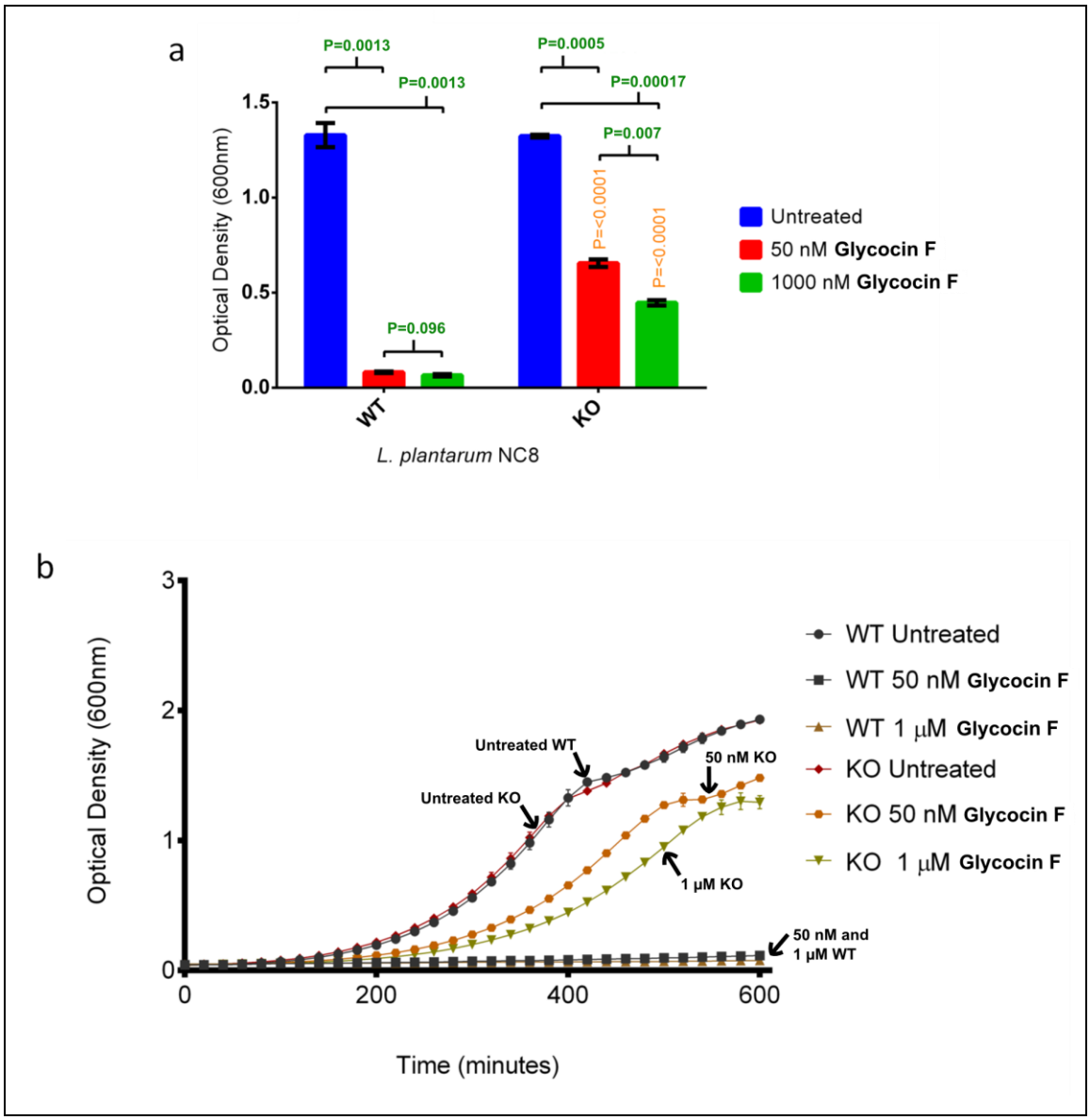


Figure 3.24 Liquid glycocin F assay of *L. plantarum* NC8 Δ*pts18CBA*

From duplicate liquid glycocin F assays of *L. plantarum* NC8 (WT) and *L. plantarum* NC8 Δ*pts18CBA* (KO). a) Bar graphs show the mean OD_{600nm} after 400 minutes from the growth curves shown in (b). Error bars represent the standard deviation. Statistical significance (P-value) between untreated and treated and between 50 nM treated and 1 µM treated samples was determined using the Holm-Sidak test (values in green). Statistical significance (P-value) of treated mutant isolates to their respective counterpart treated WT control was determined using the Holm multiple comparison test (values in blue). b) Three cultures were inoculated at an OD_{600nm} of 0.05, two were treated with either 50 nM or 1 µM glycocin F and the other left untreated. The OD_{600nm} was measured every 30 seconds using a Cary 300 UV-Visible spectrophotometer and every 20 minutes the mean OD_{600nm} of duplicate cuvettes was plotted. Vertical error bars represent the standard deviation from the mean.

suggesting that at least in these two strains PTS18CBA may not be the only receptor. The other possible receptor identified (section 1.4, page 18) was the gene product of *pts22CBA*. Alignments of PTS22CBA, from *L. plantarum* NC8, *L. plantarum* ATCC 8014 and *L. plantarum* subsp. *plantarum* ATCC 14917, showed no amino acid differences in the EIIC domain (Appendix 17). Extending this search to the promoter region of *pts22CBA* showed no differences in sequences between *L. plantarum* NC8 and *L. plantarum* subsp. *plantarum* ATCC 14917. However in *L. plantarum* ATCC 8014 three nucleotides were found to differ compared to the other two strains (Figure 3.25). The -35, -10 and RBS were identified, based on consensus sequence similarity, though the differences were not directly located within these elements. How the differences in this region affect the expression of *pts22CBA* needs to be experimentally verified and may eventually explain the increased sensitivity to glycocin F observed by *L. plantarum* subsp. *plantarum* ATCC 14917. It may be that these changes may prevent expression of *pts22CBA* so that PTS18CBA is left as the only receptor for glycocin F in the membrane. The question remains, just how does glycocin F induce its bacteriostatic effect through this GlcNAc-specific PTS transporter. This is discussed in chapter 4.

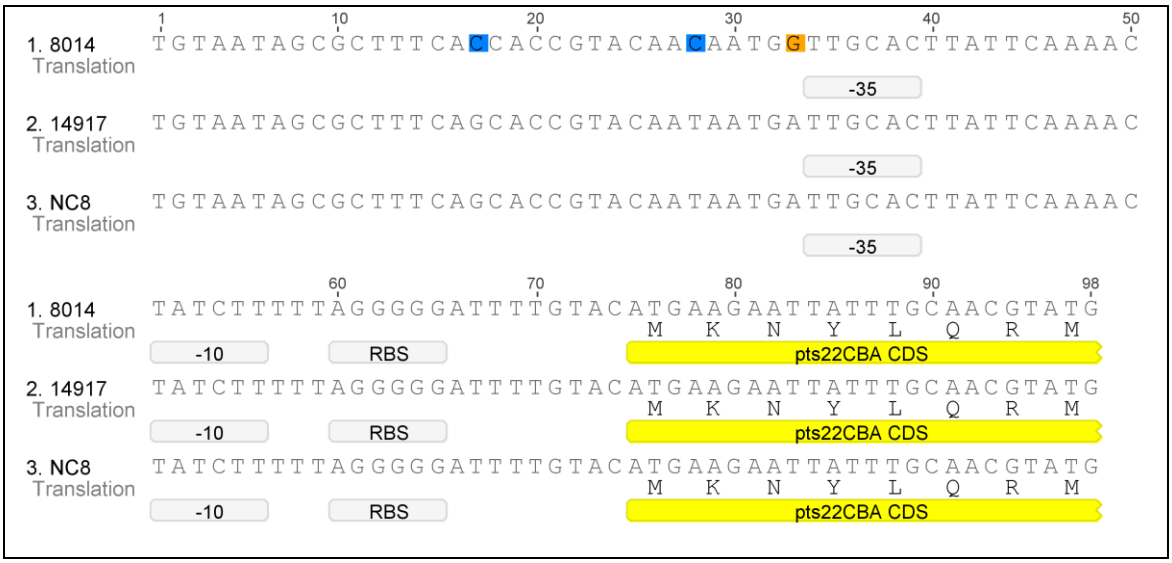


Figure 3.25 *pts22CBA* nucleotide alignment
The nucleotide region upstream from *pts22CBA* from (2) *L. plantarum* ATCC subsp. *plantarum* ATCC 14917 (14917), (1) *L. plantarum* ATCC 8014 (8014) and (3) *L. plantarum* NC8 (NC8), were aligned. The coloured nucleotides in (1) are the differences identified in *L. plantarum* ATCC 8014. The yellow bar is the *pts22CBA* coding sequence with the predicted translation above. Promoter elements are highlighted by silver bars.

3.9 Production of PTS18CBA protein

3.9.1 Introduction

Results of the *L. plantarum* NC8 Δ*pts18CBA* (section 3.8.3) and glycocin F resistant *L. plantarum* mutants (section 3.4.2) confirmed the presence of an interaction between PTS18CBA and glycocin F. To further study this interaction requires purified protein, i.e. an active PTS18CBA to allow the investigation of its functionality using methods such as isothermal titration calorimetry or surface plasmon resonance to measure the interaction between the transporter and glycocin F. Ultimately, a three dimensional structure would be desirable. The general membrane protein purification protocol developed by Newby *et al.* (2009) was initially used with the aim of producing C-terminal hexa-histidine tagged PTS18CBA (rPTS18CBA). The pET-21b(+) plasmid (Appendix 18) was chosen as the vector because its multiple cloning site (MCS) allows the use of *Nde*I and *Not*I restriction endonucleases to produce a hexa-histidine tag at the C-terminus of PTS18CBA to aid in protein purification. Cloning with the *Nde*I and *Not*I will also result in the loss of the T7 tag immediately downstream of the start codon.

Gram-positive bacteria target proteins to the membrane using one of multiple pathways. Using the numerous signal peptide prediction algorithms, no significant signal sequences for either of these three pathways could be deduced in PTS18CBA. However at the N-terminus there is a string of 26 hydrophobic residues flanked by positively charged residues, two at the N-terminal and one at the C-terminal (Figure 3.9). This is a pattern recognised by the Sec translocase mediated pathway and the T7 tag sequence could potentially pose a problem for membrane insertion. Due to time constraints, codon optimisation was not investigated.

3.9.2 Aims

- To clone the open reading frame of *pts18CBA* from *L. plantarum* ATCC 8014 and *L. plantarum* subsp. *plantarum* ATCC 14917 into the T7 expression plasmid pET-21b(+).
- To purify rPTS18CBA from pET-21b(+)_8014_His6_pts18CBA expression plasmid for use in subsequent experiments aimed to characterise the interaction between glycocin F and rPTS18CBA.

3.9.3 Constructing PTS18CBA production plasmids

pET-21b(+) purified from *E. coli* EC100 (section 2.1.9) was treated with New England Biolabs® *Nde*I, *Not*I and *Pvu*I (section 2.1.15: 5 U of *Not*I, 10 U of *Nde*I and 5 U of *Pvu*I were used in the

reactions. All reactions contained 1x concentration of 100x New England Biolabs® Bovine serum albumin and 1x concentration of New England Biolabs® Buffer 3. 150 ng of pET-21b(+) purified from *E. coli* EC100 cell was present in each reactions which were incubated at 37°C for 4 hours) and analysed using agarose gel electrophoresis (section 2.1.13). Both *Nde*I and *Not*I digested pET-21b(+) with no observable star activity whereas *Pvu*I was inefficient (Figure 3.26). pET-21b(+) was digested, as described above, with New England Biolabs® *Nde*I and *Not*I and the digested plasmid purified (section 2.1.13.6).

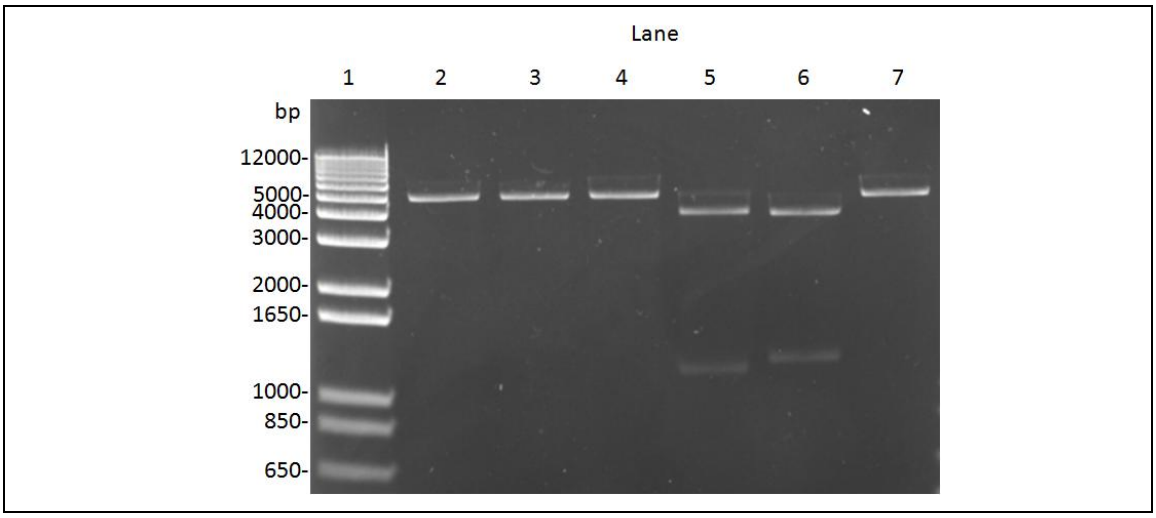


Figure 3.26 pET-21b(+) trial digestions

Agarose gel containing pet-21b(+) treated with: lane 2, New England Biolabs® *Not*I; lane 3, New England Biolabs® *Nde*I; lane 4, New England Biolabs® *Pvu*I; lane 5, *Not*I and *Pvu*I; lane 6, *Nde*I and *Pvu*I; lane 7, *Nde*I and *Not*I. Lane 1, Invitrogen™ 1 kb Plus DNA Ladder.

The *pts18CBA* coding region was amplified out of both *L. plantarum* subsp. *plantarum* ATCC 14917 and *L. plantarum* ATCC 8014 gDNA using high fidelity PCR (section 2.1.12) and the primer pairs pts18CBA_F_NdeI and pts18CBA_R_NotI (Appendix 13). PCR products were analysed using agarose gel electrophoresis (Figure 3.27a) and those of the appropriate size were purified (section 2.1.13.6). Primers pts18CBA_F_NdeI and pts18CBA_R_NotI, contained *Nde*I and *Not*I restriction enzyme recognition sites respectively. The purified PCR products were treated with both New England Biolabs® *Nde*I and *Not*I (section 2.1.15: 500 ng of the PCR products were treated with 2.5 U of New England Biolabs® *Not*I and 5 U of New England Biolabs® *Nde*I for 30 minutes at 37°C followed by a 20 minute incubation at 65°C; New England Biolabs® Bovine serum albumin and New England Biolabs® Buffer 3 were included in the reactions at 1x concentration), analysed by agarose gel electrophoresis (Figure 3.27b) and the bands of the appropriate size excised and purified as described above. The purified PCR and plasmid digestion products were subjected to ligation (section 2.1.16) at a molar ratio of 3:1

respectively, and *E. coli* EC100 cells transformed (section 2.2.2) with the ligation mixtures. The cells were plated onto ampicillin (section 2.1.5) LB agar plates and incubated for 48 hours at 37°C. Colonies were selected and streaked out onto ampicillin LB agar plates then incubated for 24 hours at 37°C. Two of the streaked colonies originating from the 8014 ligation and five streaked colonies originating from the 14917 ligation grew. These were screened for *pts18CBA* insertion by PCR (section 2.1.11) using primer pair pet_32_T7_term and lac_oprn that amplifies across the MCS; a PCR product of 2.3-kbp is expected if the gene is present, while a product of 250-bp will be present if the gene has not been inserted. Four of the 14917 and both of the 8014 colonies were shown to contain the gene (Figure 3.28). Both the 8014 colonies and colonies 2 and 3 from the 14917 reaction were cultured and the plasmid purified (section 2.1.9). The purified plasmids were subjected to sequence analysis (section 2.1.17) which confirmed them to contain the gene. They were named pET-21b(+)_8014_His6_pts18CBA and pET-21b(+)_14917_His6_pts18CBA (Appendix 14).

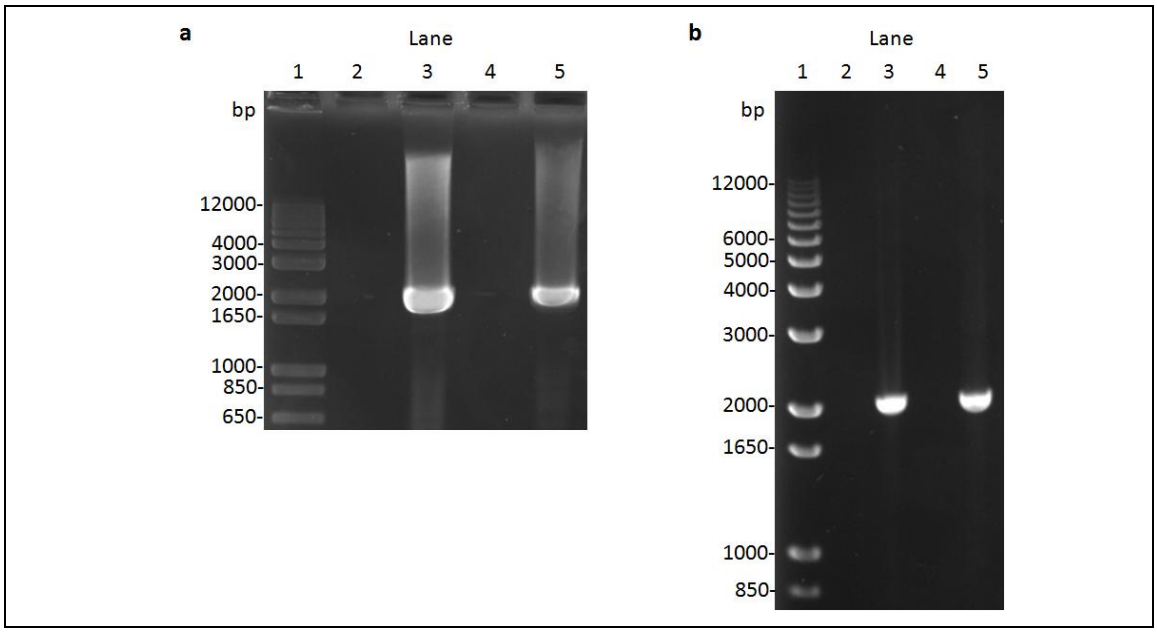


Figure 3.27 Agarose gels of *pts18CBA* inserts for pET-21b(+) expression construction

a): Lane 1, Invitrogen™ 1 kb Plus DNA Ladder; lane 2, empty; lane 3, high fidelity PCR product amplified from *L. plantarum* subsp. *plantarum* ATCC 14917 gDNA using primers pts18CBA_F_NdeI and pts18CBA_R_NotI; lane 4, empty; lane 5, high fidelity PCR product amplified from *L. plantarum* ATCC 8014 gDNA using primers pts18CBA_F_NdeI and pts18CBA_R_NotI. b) Lane 1, Invitrogen™ 1 kb Plus DNA Ladder; lane 2, empty; lane 3, restriction enzyme digestions of 8014 *pts18CBA* PCR product; lane 4, empty; lane 5, restriction enzyme digestions of 14917 *pts18CBA* PCR product.

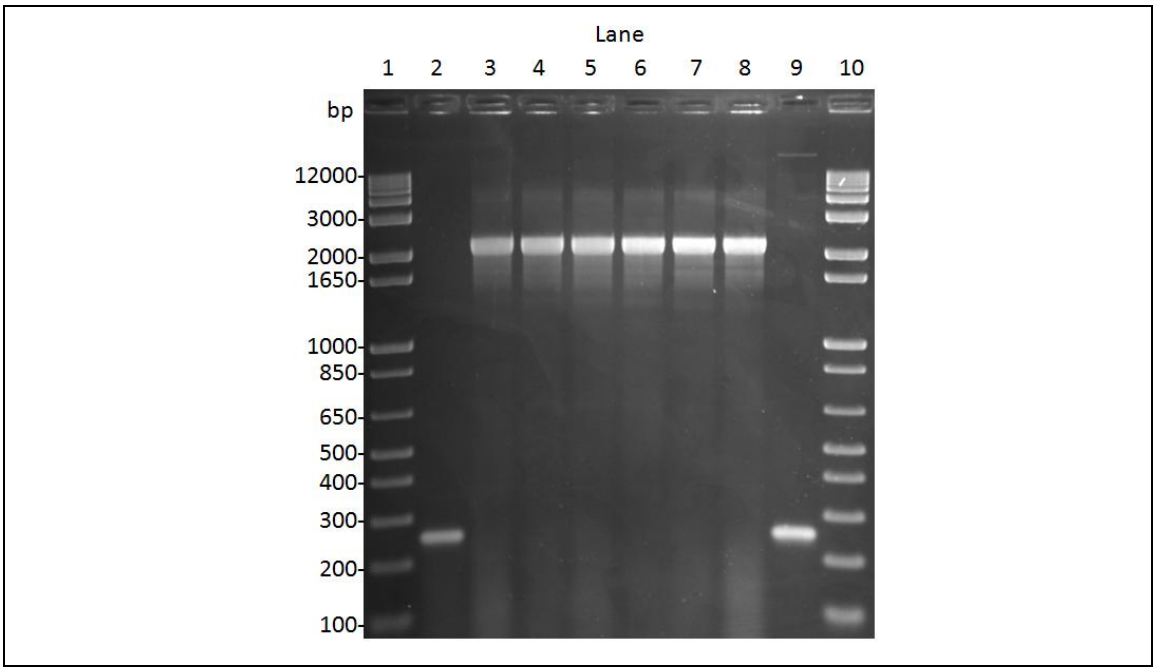


Figure 3.28 Colony PCR screening of pET-21b(+) *pts18CBA* 8014 and 14917 cloning

Agarose gel showing colony PCR products from primer pair pet_32_T7_term and lac_oprn products of 14917 (lanes 2-5, colony number offset by 1) and 8014 (lane 7 and 8, colony number offset by 6) cloning experiments. pET-21b(+) was used as DNA template to serve as a negative control (lane 9). A successful clone is indicated by a 2.3-kbp band and negative by a 250-bp band. Invitrogen™ 1 kb Plus DNA Ladder is present in lanes 1 and 10.

3.9.4 Production and purification of rPTS18CBA

SDS-PAGE analysis (section 2.5.2) of the time course trial (section 2.5.1.1) showed a band that corresponds to the predicted molecular weight of rPTS18CBA, 71.60 kDa, to be increasing in presence with time (Figure 3.29a). The longest induction time point trialled was 12 hours. The SDS-polyacrylamide gel was however, overloaded with this sample, and thus the 240 minute is the greatest time point with distinguishable bands. 240 minutes was therefore chosen as the induction time for subsequent solubility trials. Small scale solubility trials determined rPTS18CBA to be approximately 40 % soluble, and it appears to be targeted to the cell membrane (Figure 3.29b). Two aliquots of the sample loaded in lane 7 of Figure 3.29b were subjected to SDS-PAGE, identical conditions as above, and the band suspected to be rPTS18CBA was excised, digested with trypsin and analysed by tandem mass spectrometry (section 2.5.3). The results of the Mascot NCBI non-redundant database search (section 2.5.3), data generated by electron spray ionisation MS/MS, from one of the replicates suggested that it contains a GlcNAc/glucose-specific PTS EIICB from *L. plantarum* (peptide score of 47, peptide

scores > 46 indicate identity or extensive homology with a p<0.05) (electronic Appendix 19). The second replicate contained no peptides with a score above the threshold (peptide scores > 47 indicate identity or extensive homology with a p<0.05).

Although the first replicate does identify the protein extracted from the band as belonging to the GlcNAc/glucose-specific PTSs, the results are not as reliable as they should be due to the low peptide scores from both replicates. The low peptide scores are probably due to the low concentration of the protein in the sample shown in lane 7 of Figure 3.29b. This could be resolved by concentrating the sample so that more protein can be loaded on the gel. It is also possible that there are two proteins in this band, the rPTS18CBA and some other protein. The presence of peptides from two sources would complicate the spectra, and have the effect of lowering the score. Limited time did not allow for either further purification attempts of rPTS18CBA or further attempts to confirm the identity of the extracted protein, which is required before using purified rPTS18CBA in subsequent experiments.

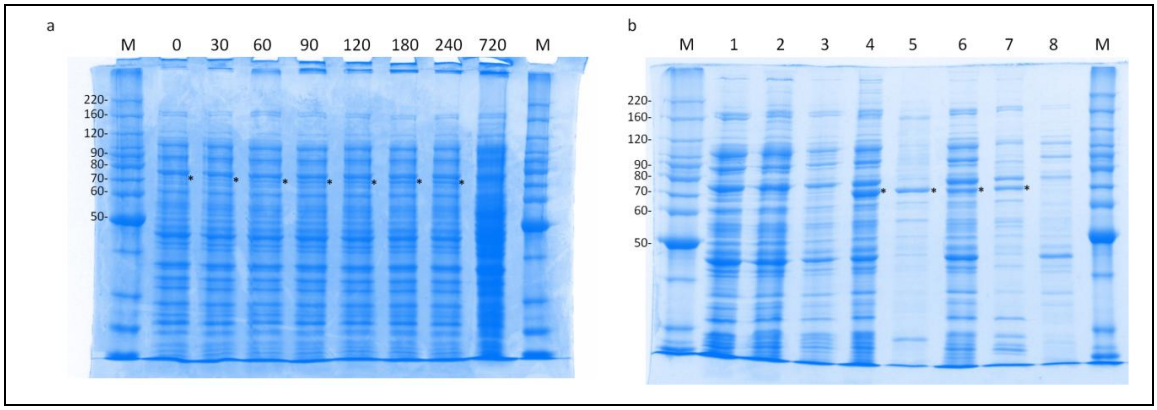


Figure 3.29 SDS-PAGE gels of 8014 rPTS18CBA production and purification

a) SDS-polyacrylamide gel containing protein from whole cells sampled at the time points (minutes) indicated above the lanes. An asterisk marks the putative rPTS18CBA. b) SDS-polyacrylamide gel of small scale protein solubility trial. An asterisk marks the putative rPTS18CBA. Lane 1 is the total protein from un-induced *E. coli* BL21 transformed with empty pet-21b(+) vector. Lane 2 is the total protein from induced *E. coli* BL21 transformed with empty pet-21b(+) vector. Lane 3 is total protein from un-induced *E. coli* BL21 transformed with pET-21b(+)_8014_pts18CBA expression plasmid. Lane 4 is total protein from induced *E. coli* BL21 transformed with pET-21b(+)_8014_pts18CBA expression plasmid. Lane 5 is insoluble fraction following cell lysis. Lane 6 is the soluble fraction following cell lysis. Lane 7 is the membrane fraction following ultracentrifugation. Lane 8 is the soluble fraction following ultracentrifugation. Bands corresponding to rPTS18CBA are indicated by asterisks. The protein marker (M) BenchMark™ Protein Ladder was loaded in the outermost lanes in both gels.

4.General Discussion and Conclusion

4. General Discussion and Conclusions

A major aim of this study was to identify proteins in *L. plantarum* strains that are required for sensitivity to glycocin F. One approach was to use comparative genome analysis of *L. plantarum* ATCC 8014 and *L. plantarum* subsp. *plantarum* ATCC 14917 mutants selected for resistance to glycocin F-mediated bacteriostasis. This provided evidence that PTS18CBA is probably the main receptor for glycocin F. As shown in section 3.2, all six glycocin F-resistant isolates had mutations in or overlapping the *pts18CBA* gene, those predicted to cause complete loss of transporter function being correlated with either complete resistance (in ATCC 8014) or very high level of resistance (in ATCC 14917) to glycocin F. On the other hand mutations predicted to decrease substrate affinity or result in a non-functional transport domain seemed to be associated with a relatively low level of resistance. This approach was supported by the successful targeted gene disruption (knock out) of *pts18CBA* in *L. plantarum* NC8 (section 3.8), which also identified PTS18CBA as the probable cognate receptor for glycocin F. The *L. plantarum* NC8 Δ*pts18CBA* strain was generated only after repeated attempts to use targeted gene disruption in *L. plantarum* ATCC 8014 and *L. plantarum* subsp. *plantarum* ATCC 14917 had failed. A second aim of this study was to characterise the interaction of glycocin F with PTS18CBA, but unfortunately time did not permit this.

4.1 Glycocin F binding to PTS18CBA

This study has provided definitive evidence that the putative GlcNAc-specific PTS transporter, PTS18CBA, is required for sensitivity to glycocin F. Given that both GlcNAc moieties of glycocin F contribute to its inhibitory activity, and indeed the Ser18-*O*-GlcNAc is essential for activity, it seems plausible that glycocin F-mediated inhibition requires one or both of the GlcNAc moieties to be bound in the GlcNAc-binding sites in the EIIC domain of PTS18CBA. One hypothesis is that the serine-linked GlcNAc initiates the interaction by occupying one GlcNAc binding site, and then the flexible 'tail' of glycocin F 'reaches' above, around and down the oligomerisation domain to the other ligand binding site. Glycocin F was modelled into the predicted outward-open state of ChbC that was suggested by Cao *et al.* (2011) and McCoy *et al.* (2014), using PyMOL (Schrodinger, 2010). This model, shown in Figure 4.1, shows that glycocin F could physically dock with the protein in such a way that the two GlcNAc moieties could each occupy a sugar-binding site.

The model shows that this interaction would be favoured by electrostatic bonds between the oligomerisation domain of EIIC and the positively charged residues in the 'tail' of glycocin F

(residues H27, K32, H33, H34 and H42). Hydrogen bonds and stacking interactions between the GlcNAc moieties of glycocin F and amino acids located in the ligand binding site of EIIC would also favour this interaction. These hypothetical interactions are consistent with the effects of the single missense mutations identified in LP8_1**L** and LP1_3**L**, which are thought to result in a decreased affinity for GlcNAc. The results of Stepper *et al.* (2011) and Kerr (2013), that pre- and post-exposure to GlcNAc can increase or alleviate the bacteriostatic effect of glycocin F respectively, depending on the target species, also support this model.

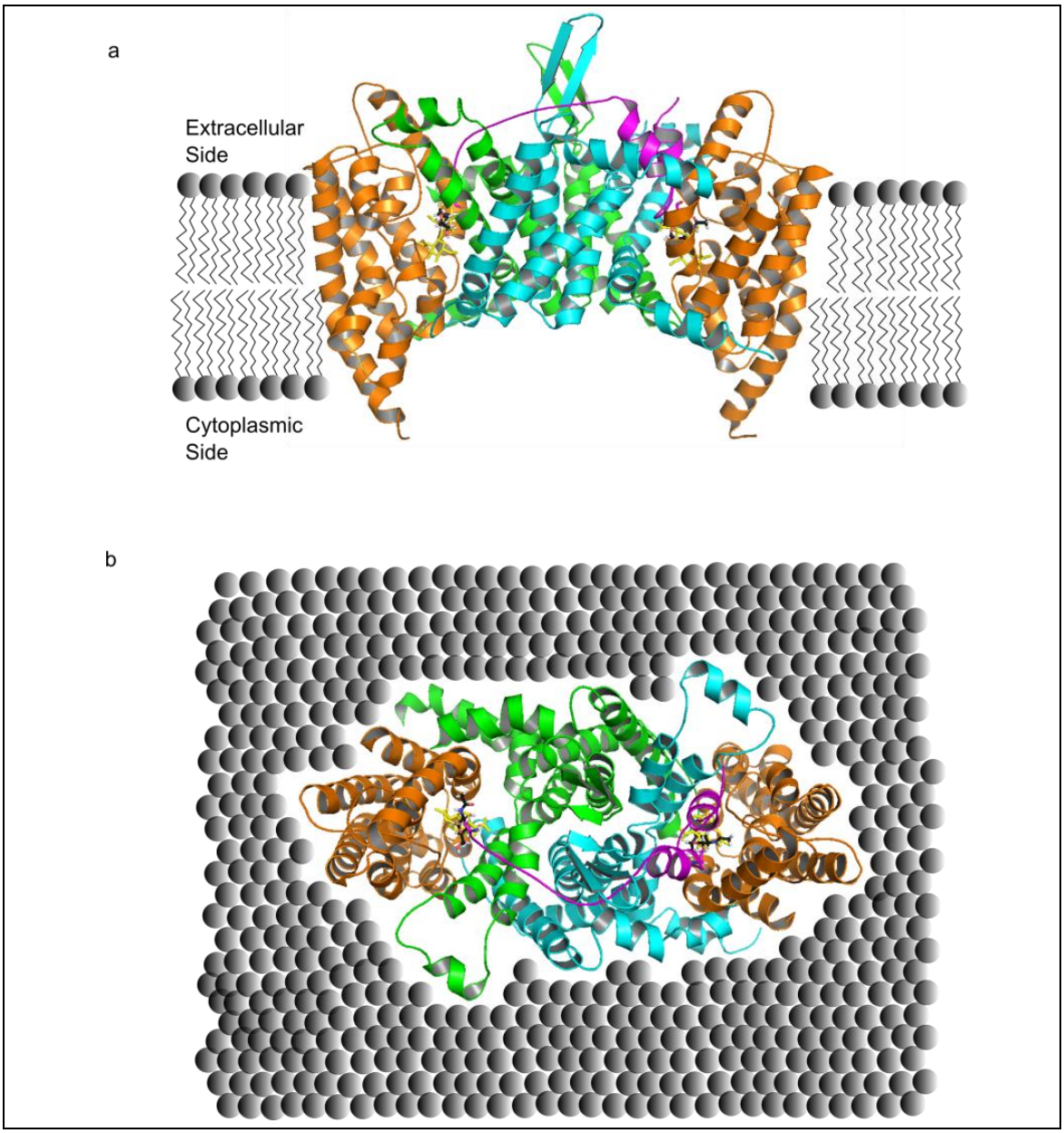


Figure 4.1 Model of the outward-open state of ChbC with glycocin F
a) Side on and (b) top down views of ChbC in the outward-open state with glycocin F (pink) bound in the two N,N'-diacetylchitobiose (yellow)-binding sites. The transport domains are coloured orange, and the oligomerisation domains are coloured cyan and green.

4.2 How might formation of a glycocin F:PTS18CBA complex cause bacteriostasis?

The function of the GlcNAc residues on glycocin F could simply be to recruit the bacteriocin to the cell membrane through binding to the PTS. Once in position, glycocin F could then interact with other non-PTS molecules to exert its inhibitory effects. This is, however, unlikely because very high concentrations of glycocin F should still be inhibiting in the absence of the PTS, this was shown not to be the case for *L. plantarum* ATCC 8014 group H mutants (section 3.2, Figure 3.3b). Instead, it is possible that glycocin F traps EIIC in the outward-open state (Figure 1.5) leading to a build-up in the concentration of phosphorylated EIIB and EIIA at the membrane (see phosphorylation cascade, section 1.3.2, page 11). Many PTS studies suggest that accumulation of the phosphorylated forms of these two PTS subunits/domains is enough to induce an increase in the concentration of the phosphorylated forms of cytoplasmic HPr and EI (Weigel *et al.*, 1982, Meadow and Roseman, 1996, Pelton *et al.*, 1996, Rohwer *et al.*, 2000, Francke *et al.*, 2002, Francke *et al.*, 2003), which in turn could increase the phosphorylation of certain non-PTS proteins (see section 1.3.3, page 15).

In *Firmicutes* species, glycerol kinase is the only non-PTS enzyme known to be regulated by an inducer exclusion mechanism (section 1.3.3, page 15), though bacteriostasis due to the effects of a hyper-phosphorylated glycerol kinase seems unlikely. A number of transcription factors and antiterminators have been shown to be regulated by the inducer exclusion mechanisms, though none of these have been shown to have a role in the regulation of cell division. It may be that there are unidentified enzymes, that have a role in cell division, or transcriptional regulators, that regulate proteins that have a role in cell division, that are phosphoregulated by components of the PTS. Such a mechanism could explain how glycocin F can induce bacteriostasis in minutes, at nanomolar concentrations, as the chemotaxis study in *E. coli* by Lux *et al.* (1999) showed that in a matter of seconds nanomolar concentrations of glucose were capable of causing a significant change in the phosphorylation of EI leading to the dephosphorylation of the non-PTS protein CheY.

Alternatively, the build-up of phosphorylated PTS subunits/domains may affect other regulatory pathways. The bacteriostatic effect induced by glycocin F is reminiscent of the cell cycle arrest characteristic of the 'stringent response', reviewed by Katarzyna and Michael (2008) and Srivatsan and Wang (2008), which occurs in most bacteria in response to nutrient stress. In *E. coli*, the two proteins central to this response are RelA and SpoT, which synthesise the 'alarmone' nucleotides guanosine tetraphosphate/pentaphosphate ((p)ppGpp), although SpoT also catalyses the reverse reaction (Xiao *et al.*, 1991, Gropp *et al.*, 2001). (p)ppGpp interacts

with the transcription machinery (Ross *et al.*, 2013) and changes the expression of hundreds of genes (Schreiber *et al.*, 1995, Barker *et al.*, 2001, Durfee *et al.*, 2008). In both *E. coli* and *B. subtilis*, the cell cycle arrest is due to the inhibition of DNA replication (Levine *et al.*, 1991, Ferullo and Lovett, 2008). It was shown recently by Karstens *et al.* (2014), that in the Gram-negative bacterium *Ralstonia eutropha* H16, the unphosphorylated nitrogen related-PTS EIIA subunit interacts with SpoT. They were, however, unable to replicate this interaction in *E. coli*. The physiological role of the interaction between SpoT and the unphosphorylated EIIA is unknown. However, a related interaction could be the missing link between the formation of a glycocin F:PTS18CBA complex and bacteriostasis. For example, the interaction of unphosphorylated EIIA with SpoT might suppress synthesis of (p)ppGpp, whereas phosphorylation of EIIA could disrupt this interaction leading to (p)ppGpp synthesis, and subsequent cell cycle arrest. This is very speculative however, and is based on experimental observations in organisms only distantly related to *Lactobacillus plantarum* species, glycocin F's natural targets. This could be experimentally investigated using a metabolomics approach to determine if changes in the levels of (p)ppGpp occur upon addition of glycocin F, or by using a bacterial two-hybrid methodology (Karstens *et al.*, 2014) to identify PTS18CBA binding partners.

Both phosphorylation of non-PTS proteins, and triggering of the stringent response could also explain how the glycocin F putative immunity protein might act in the glycocin F producer strain, *L. plantarum* KW30. It has been suggested that the immunity proteins of bacteriocins that use PTS transporters as their cognate receptors might function similarly. The work of Diep *et al.* (2007), discussed in section 1.2.4 (page 8), showed the immunity protein of lactococcin A interacts with the EIIAB, EIIC and EIID subunits of the mannose-specific PTS transporter in a lactococcin A dependent manner. As lactococcin A has bactericidal activity and works by dissipating the proton motive force, Diep *et al.* (2007) proposed that the immunity protein might simply plug a bacteriocin-induced pore formed in the transmembrane EIIC/D domains. The identity of the glycocin F immunity protein is unknown, although likely candidates are the gene products of *gccI* and *gccH*, both of which are located in the *gcc* gene cluster in close proximity to the glycocin F coding sequence (Stepper *et al.*, 2011). Regardless, the glycocin F immunity protein, may be binding phosphorylated EIIB or EIIA domains in a fashion similar to the lactococcin A immunity protein, effectively concealing these phosphorylated domains and so preventing molecular interactions that would otherwise lead to bacteriostasis. However this is speculative and experimental validation is necessary. This could be accomplished by *in vivo* pull down experiments much like those carried out by Diep *et al.* (2007) or by using the bacterial two-hybrid method.

A possible inhibitory mechanism that the results of this study do not exclude is that glycocin F could be inducing pore formation through PTS18CBA, similar to what was suggested for lactococcin A (section 1.2.4, page 8) and microcin E492 (section 1.2.5, page 9). The results of the ATP efflux experiments by Kerr (2013), that showed a small amount of ATP efflux upon treatment with glycocin F, are consistent with this theory. However, the known pore-forming bacteriocins are all bactericidal, while glycocin F is clearly bacteriostatic (Kerr, 2013) suggesting this is not how it inhibits growth.

Another possible mechanism for glycocin F involves entering the cell through the PTS18CBA, similar to some of the class IIe peptides (Heddle *et al.*, 2001). Once inside the cell it may interact with proteins that have a vital role in cell division, similar to microcin B17 which acts to inhibit DNA gyrase (Pierrat and Maxwell, 2003). This, however is unlikely as there are no documented precedents of PTSs moving bacteriocins into cells, even though the mannose PTS is known to be the receptor for many bacteriocins.

4.3 A second receptor or mechanism of action for glycocin F?

Interestingly isolates of *L. plantarum* subsp. *plantarum* ATCC 14917 and *L. plantarum* NC8 in which PTS18CBA was compromised by partial or complete deletion were not completely resistant to 1 µM glycocin F. This suggests that in these strains, high concentrations of glycocin F can either inhibit by a mechanism unrelated to PTS transport, or more likely that another receptor is present in the cell membrane. It is possible that this second receptor could be another putative GlcNAc-specific transporter, PTS22CBA. An amino acid alignment of PTS22CBA (Appendix 17) from the three *L. plantarum* strains showed no differences in the transmembrane EIIC domain. However, in contrast to *L. plantarum* subsp. *plantarum* ATCC 14917 and *L. plantarum* NC8, differences were identified upstream in the 5′-untranslated region of *L. plantarum* ATCC 8014 *pts22CBA* (Figure 3.25). These differences may decrease the constitutive expression of *pts22CBA*, leaving PTS18CBA as the sole glycocin F receptor. At first this may seem contradictory to the observation that all mutants, including *L. plantarum* NC8 Δ*pts18CBA*, can still utilise GlcNAc as a sole carbon source (Appendix 2). However it is possible that more abundant PTS transporters responsible for glucose and mannose transport are sufficiently non-specific to allow them to also transport GlcNAc. Less likely, but still possible is that in *L. plantarum* ATCC 8014 only, the putative glycocin F:PTS22CBA interaction does not inhibit growth. This is based on a theory that within the three strains, the GlcNAc-specific PTS transporter homologues have different regulatory functions. Evidence for this may have already been shown by the carbon utilisation assays (Appendix 2). As discussed in section 1.3.3, the PTSs of *Firmicutes* regulate the *lac* operon; the increased transport of lactose through the

lactose-specific PTS transporter is thought to result in the up-regulation of the *lac* operon. The glycocin F-resistant *L. plantarum* subsp. *plantarum* ATCC 14917 mutants gained the ability to utilise lactose, and potentially fructose. The same effect was not observed in either glycocin F-resistant *L. plantarum* 8014 mutants or *L. plantarum* NC8 Δ*pts18CBA*, indicating that PTS18CBA down-regulates lactose and fructose metabolic pathways in *L. plantarum* subsp. *plantarum* 14917, but not in *L. plantarum* NC8 or *L. plantarum* ATCC 8014. Caution is necessary however in coming to this conclusion as the gained ability to utilise lactose could be the effect of mutations not within the boundaries of the assembled contigs that remain to be identified. Either explanation (differential expression of *pts22CBA* or lack of downstream inhibitory signalling) could account for the complete resistance observed for the *L. plantarum* ATCC 8014 mutants while the other two *L. plantarum* strains retained residual sensitivity to glycocin F. However, if PTS22CBA is another receptor for glycocin F, the question remains why mutations in its coding sequence were not identified in the glycocin F-resistant *L. plantarum* subsp. *plantarum* ATCC 14917 mutants.

4.4 Possible mechanism of actions for glycocin F

A model for the mechanism of action (MOA) of glycocin F is proposed: glycocin F manipulates either the PTS18CBA-mediated GlcNAc inducer exclusion regulation pathway or the stringent response, tricking the cell into 'thinking' it is in a carbon-deficient environment and that its best chance of surviving is to go into a hibernation-like state until the suitable carbon sources are introduced into the environment. The postulated mechanism involves glycocin F binding to the outward-open state of the transmembrane EIIC dimer locking it in the open state, leading to a build up of phosphorylated EIIB. This leads in turn to a build up of phosphorylated EIIA (and potentially HPr and EI), that alter the activity of various transcriptional regulators, some of which control the production of proteins involved in cell division. This may also feed into the stringent response, through SpoT, which then induces cell cycle arrest. The downstream elements linked to this phosphorylation cascade that are involved in the regulation of cell division, have yet to be identified in *Firmicutes* species. Thus the next major step in elucidating the MOA of glycocin F is to determine if glycocin F is manipulating inducer exclusion regulatory pathways, or inducing the stringent response, using approaches summarised in section 5.

5. Future Directions

5. Future Directions

The binding of glycocin F to PTS18CBA should be characterised *in vitro* by producing and purifying rPTS18CBA. This is not trivial since about half of PTS18CBA is an integral membrane protein domain (EIIC), and these present specific challenges in protein purification and crystallisation. Nevertheless, significant progress towards this goal is reported in this thesis.

Such characterisation could involve analysis of the binding between glycocin F and PTS18CBA using techniques such as isothermal titration calorimetry, circular dichroism, small angle X-ray scattering and protein crystallisation and X-ray diffraction. Crystallisation of a dimer of the multi-domain recombinant EIICBA would likely be extremely challenging, but the chance of success might improve if the isolated EIIC domain was the crystallisation target. Glycocin F could perhaps be co-crystallised with rPTS18CBA, or even soaked into the crystal.

The hypothesis that glycocin F exerts its antimicrobial activity through binding to PTS18CBA and induces bacteriostasis through a build up of phosphorylated PTS proteins requires experimental validation. This can be achieved by either studying the expression of genes that are regulated by transcription factors or antiterminators that are themselves known to be regulated by PTS-mediated inducer exclusion. Based on the differential lactose utilisation displayed by the glycocin F resistant mutants, the immediate target for such investigations would be LacT. Also studying glycerol kinase activity in response to glycocin F could be informative. This can be done using transcriptomic, proteomic and metabolomic methods. Proteomic methods include comparative 1- dimensional and/or 2-dimensional SDS-PAGE, with differential gel electrophoresis, of cell lysates with and without glycocin F treatment coupled to in-gel tryptic digestion and MS/MS to identify proteins with changed abundance. The expression of these genes could then be investigated using real-time PCR. A metabolomic approach using mass-spectrometry and/or NMR would allow the intracellular activity of glycerol kinase to be evaluated in response to glycocin F by looking at the relative levels of glycerol and glycerol-3-phosphate (Brockmöller *et al.*, 2011). These investigations would be carried out to link glycocin F to inducer exclusion regulation, not necessarily to the cell cycle arrest induced by the stringent response. Instead, to investigate if glycocin F uses the stringent response to induce bacteriostasis, metabolomics could be used to analyse the levels of (p)ppGpp (Kriel *et al.*, 2012).

The results of the *pts18CBA* knockout in *L. plantarum* NC8 need further validation through complementation of the knocked out *pts18CBA* to show recovery of sensitivity to glycocin F. The suspicion that there is a second receptor, possibly PTS22CBA, that is responsible for the

residual sensitivity to glycocin F seen in some glycocin F-selected mutants, and in NC8 KO, needs further examination. An experiment that has potential, is to repeat the natural selection experiments, generating glycocin F-resistant *L. plantarum* at concentrations of glycocin F much higher than was used originally. This could also be completed using the glycocin F-resistant mutants that are only partially resistant in an attempt to force them to become completely resistant. This may allow the identification of proteins 'downstream' of glycocin F binding PTS18CBA or a second receptor.

Additionally, the genome assembly, integration and annotation of other strains with varying susceptibility to glycocin F should be continued as such comparisons have the potential to yield valuable information about the MOA of glycocin F. Other *de novo* assembly algorithms exist which could be used for the generation of contigs for CISA integration. Using the latest release of Spades may avoid the misassembles reported here and allow these contigs to be included in CISA integration. Also, the genomes of *L. plantarum* subsp. *plantarum* ATCC 14917 and *L. plantarum* ATCC 8014 could be completed by closing the gaps between the contigs, which if performed for the mutants isolates may reveal other mutations that could be responsible for the glycocin F resistance. Additionally, the targeted sequencing of *pts18CBA* of the four glycocin F resistant mutants that were not analysed by Illumina sequencing could identify additional mutations in PTS18CBA.

Further biochemical characterisation of *L. plantarum* NC8 Δ*pts18CBA* and the glycocin F-resistant mutants would be informative. Such characterisation could include investigating growth rates in the presence of PTS sugars that are imported up by PTS transporters, with and without glycocin F. An experiment that could provide more details on the effects on the mutations on PTS18CBA function would be characterising their sensitivity to the antimicrobial GlcNAc analogue streptozotocin which would likely use PTS18CBA to enter the cell. This could reveal if some of the partial glycocin F-resistant mutants have an impaired ability to translocate GlcNAc, as if they retained wildtype sensitivity to streptozotocin then this could suggest there are unidentified mutations that confer glycocin F-resistance.

Evidence is provided that PTS18CBA is a receptor for the diglycosylated bacteriostatic glycocin F. Future investigations aimed at elucidating the mechanism of action of glycocin F show potential that may lead to discoveries involving the regulatory pathway that glycocin F hijacks. These studies may also lead to the development of a synthetic antibiotic resembling glycocin F, two GlcNAc moieties separated by a positively charged linker.

Bibliography

AHRNE, S., NOBAEK, S., JEPPSSON, B., ADLERBERTH, I., WOLD, A. E. & MOLIN, G. 1998. The normal *Lactobacillus* flora of healthy human rectal and oral mucosa. *Journal Applied Microbiology,* 85**,** 88-94.

ALEGRE, T., CARMEN RODRíGUEZ, M. & MESAS, J. M. 2004. Transformation of *Lactobacillus plantarum* by electroporation with in vitro modified plasmid DNA. *FEMS Microbiology Letters,* 241**,** 73-77.

ALPERT, C. A. & SIEBERS, U. 1997. The *lac* operon of *Lactobacillus casei* contains *lacT*, a gene coding for a protein of the Bg1G family of transcriptional antiterminators. *Journal of Bacteriology,* 179**,** 1555-1562.

ANTON BANKEVICH SERGEY NURK, D. A., ALEXEY A. GUREVICH, MIKHAIL DVORKIN, ALEXANDER S. KULIKOV, VALERY M. LESIN, SERGEY I. NIKOLENKO, SON PHAM, ANDREY D. PRJIBELSKI, ALEXEY V. PYSHKIN, ALEXANDER V. SIROTKIN, NIKOLAY VYAHHI, GLENN TESLER, MAX A. ALEKSEYEV, AND PAVEL A. PEVZNER 2012. SPAdes: A New Genome Assembly Algorithm and Its Applications to Single-Cell Sequencing. *Journal of Computational Biology,* 19**,** 455-477.

AQUILANTI, L., SANTARELLI, S., SILVESTRI, G., OSIMANI, A., PETRUZZELLI, A. & CLEMENTI, F. 2007. The microbial ecology of a typical Italian salami during its natural fermentation. *International Journal of Food Microbiology,* 120**,** 136-145.

ARANDA-OLMEDO, I., RAMOS, J. L. & MARQUÉS, S. 2005. Integration of Signals through Crc and PtsN in Catabolite Repression of *Pseudomonas putida* TOL Plasmid pWW0. *Applied and Environmental Microbiology,* 71**,** 4191-4198.

AUKRUST, T. & BLOM, H. 1992. Transformation of *Lactobacillus* strains used in meat and vegetable fermentations. *Food Research International,* 25**,** 253-261.

AXELSSON, L., RUD, I., NATERSTAD, K., BLOM, H., RENCKENS, B., BOEKHORST, J., KLEEREBEZEM, M., VAN HIJUM, S. & SIEZEN, R. J. 2012. Genome Sequence of the Naturally Plasmid-Free *Lactobacillus plantarum* Strain NC8 (CCUG 61730). *Journal of Bacteriology,* 194**,** 2391-2392.

BARKER, M. M., GAAL, T., JOSAITIS, C. A. & GOURSE, R. L. 2001. Mechanism of regulation of transcription initiation by ppGpp. I. Effects of ppGpp on transcription initiation in vivo and in vitro. *Journal of Molecular Biology,* 305**,** 673-688.

BERTHIER, F., ZAGOREC, M., CHAMPOMIER-VERGÈS, M., EHRLICH, S. D. & MOREL-DEVILLE, F. 1996. Efficient transformation of *Lactobacillus sake* by electroporation. *Microbiology,* 142**,** 1273-1279.

BIÉLER, S., SILVA, F. & BELIN, D. 2010. The polypeptide core of Microcin E492 stably associates with the mannose permease and interferes with mannose metabolism. *Research in Microbiology,* 161**,** 706-710.

BIÉLER, S., SILVA, F., SOTO, C. & BELIN, D. 2006. Bactericidal activity of both secreted and nonsecreted microcin E492 requires the mannose permease. *Journal of Bacteriology,* 188**,** 7049-7061.

BIZANI, D. & BRANDELLI, A. 2002. Characterization of a bacteriocin produced by a newly isolated *Bacillus* sp. Strain 8 A. *Journal of Applied Microbiology,* 93**,** 512-519.

BLAKE, K. L., RANDALL, C. P. & O'NEILL, A. J. 2011. In Vitro Studies Indicate a High Resistance Potential for the Lantibiotic Nisin in *Staphylococcus* aureus and Define a Genetic Basis for Nisin Resistance. *Antimicrobial Agents and Chemotherapy,* 55**,** 2362-2368.

BORRERO, J., BREDE, D. A., SKAUGEN, M., DIEP, D. B., HERRANZ, C., NES, I. F., CINTAS, L. M. & HERNÁNDEZ, P. E. 2011. Characterization of garvicin ML, a novel circular bacteriocin produced by *Lactococcus garvieae* DCC43, isolated from Mallard Ducks (*Anas platyrhynchos*). *Applied and Environmental Microbiology,* 77**,** 369-373.

BROCKMÖLLER, S. F., BUCHER, E., MÜLLER, B. M., BUDCZIES, J., HILVO, M., GRIFFIN, J. L., OREŠIČ, M., KALLIONIEMI, O., ILJIN, K., LOIBL, S., DARB-ESFAHANI, S., SINN, B. V., KLAUSCHEN, F., PRINZLER, J., BANGEMANN, N., ISMAEEL, F., FIEHN, O., DIETEL, M. & DENKERT, C. 2011. Integration of Metabolomics and Expression of Glycerol-3-phosphate Acyltransferase (GPAM) in Breast Cancer—Link to Patient Survival, Hormone Receptor Status, and Metabolic Profiling. *Journal of Proteome Research,* 11**,** 850-860.

BRÖTZ, H., JOSTEN, M., WIEDEMANN, I., SCHNEIDER, U., GÖTZ, F., BIERBAUM, G. & SAHL, H.-G. 1998. Role of lipid-bound peptidoglycan precursors in the formation of pores by nisin, epidermin and other lantibiotics. *Molecular Microbiology,* 30**,** 317-327.

BUHR, A., DANIELS, G. A. & ERNI, B. 1992. The glucose transporter of *Escherichia coli*. Mutants with impaired translocation activity that retain phosphorylation activity. *Journal of Biological Chemistry,* 267**,** 3847-3851.

BUSCH, W. & SAIER, M. H., JR. 2002. The transporter classification (TC) system, 2002. *Critical Reviews in Biochemistry and Molecular Biology,* 37**,** 287-337.

CAI, M., WILLIAMS, D. C., JR., WANG, G., LEE, B. R., PETERKOFSKY, A. & CLORE, G. M. 2003. 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.

CAMACHO, C., COULOURIS, G., AVAGYAN, V., MA, N., PAPADOPOULOS, J., BEALER, K. & MADDEN, T. 2009. BLAST+: architecture and applications. *BMC Bioinformatics,* 10**,** 1-9.

CAO, Y., JIN, X., LEVIN, E. J., HUANG, H., ZONG, Y., QUICK, M., WENG, J., PAN, Y., LOVE, J., PUNTA, M., ROST, B., HENDRICKSON, W. A., JAVITCH, J. A., RAJASHANKAR, K. R. & ZHOU, M. 2011. Crystal structure of a phosphorylation-coupled saccharide transporter. *Nature,* 473**,** 50-54.

CASES, I., PEREZ-MARTIN, J. & DE LORENZO, V. 1999. The IIANtr (PtsN) protein of *Pseudomonas putida* mediates the C source inhibition of the sigma54-dependent Pu promoter of the TOL plasmid. *Journal of Biological Chemistry,* 274**,** 15562-15568.

CASES, I., VELÁZQUEZ, F. & DE LORENZO, V. 2007. The ancestral role of the phosphoenolpyruvate–carbohydrate phosphotransferase system (PTS) as exposed by comparative genomics. *Research in Microbiology,* 158**,** 666-670.

CHARRIER, V., BUCKLEY, E., PARSONAGE, D., GALINIER, A., DARBON, E., JAQUINOD, M., FOREST, E., DEUTSCHER, J. & CLAIBORNE, A. 1997. Cloning and sequencing of two enterococcal *glpK* genes and regulation of the encoded glycerol kinases by

phosphoenolpyruvate-dependent, phosphotransferase system-catalyzed phosphorylation of a single histidyl residue. *Journal of Biological Chemistry,* 272**,** 14166-14174.

CHEN, H. & HOOVER, D. G. 2003. Bacteriocins and their food applications. *Comprehensive Reviews in Food Science and Food Safety,* 2**,** 82-100.

CHEN, Q. & AMSTER-CHODER, O. 1998. BglF, the sensor of the bgl system and the beta-glucosides permease of *Escherichia coli*: evidence for dimerization and intersubunit phosphotransfer. *Biochemistry,* 37**,** 8714-8723.

CHIKHI, R. & MEDVEDEV, P. 2014. Informed and automated k-mer size selection for genome assembly. *Bioinformatics,* 30**,** 31-37.

CHRIST, K., WIEDEMANN, I., BAKOWSKY, U., SAHL, H.-G. & BENDAS, G. 2007. The role of lipid II in membrane binding of and pore formation by nisin analyzed by two combined biosensor techniques. *Biochimica et Biophysica Acta (BBA) - Biomembranes,* 1768**,** 694-704.

COTTER, P. D., HILL, C. & ROSS, R. P. 2005. Bacteriocins: developing innate immunity for food. *Nature Reviews Microbiology,* 3**,** 777-788.

COTTER, P. D., ROSS, R. P. & HILL, C. 2013. Bacteriocins - a viable alternative to antibiotics? *Nature Reviews Microbiology,* 11**,** 95-105.

CRUTZ-LE COQ, A.-M. & ZAGOREC, M. 2008. Vectors for *lactobacilli* and other Gram-positive bacteria based on the minimal replicon of pRV500 from *Lactobacillus sakei*. *Plasmid,* 60**,** 212-220.

DAESCHEL, A., MCKENNEY, M. C. & MCDONALD, L. C. 1990. Bacteriocidal activity of *Lactobacillus plantarum* C-11. *Food Microbiology,* 7**,** 91-98.

DE LORENZO, V. 1984. Isolation and characterization of microcin E492 from *Klebsiella pneumoniae*. *Archives of Microbiology,* 139**,** 72-75.

DE LORENZO, V., MARTINEZ, J. L. & ASENSIO, C. 1984. Microcin-mediated interactions between *Klebsiella pneumoniae* and *Escherichia coli* strains. *Journal of General Microbiology,* 130**,** 391-400.

DE LORENZO, V. & PUGSLEY, A. P. 1985. Microcin E492, a low-molecular-weight peptide antibiotic which causes depolarization of the *Escherichia coli* cytoplasmic membrane. *Antimicrobial Agents and Chemotherapy,* 27**,** 666-669.

DE VISSER, J., AKKERMANS, A. D. L., HOEKSTRA, R. F. & DE VOS, W. M. 2004. Insertion-sequence-mediated mutations isolated during adaptation to growth and starvation in *Lactococcus lactis*. *Genetics,* 168**,** 1145-1157.

DEAN, L. & BJÖRN, C. 2008. ARWEN: a program to detect tRNA genes in metazoan mitochondrial nucleotide sequences. *Bioinformatics,* 24**,** 172-175.

DELGADO, M. A., RINTOUL, M. R., FARIAS, R. N. & SALOMON, R. A. 2001. *Escherichia coli* RNA polymerase is the target of the cyclopeptide antibiotic microcin J25. *Journal of Bacteriology,* 183**,** 4543-4550.

DESTOUMIEUX-GARZON, D., THOMAS, X., SANTAMARIA, M., GOULARD, C., BARTHELEMY, M., BOSCHER, B., BESSIN, Y., MOLLE, G., PONS, A. M., LETELLIER, L., PEDUZZI, J. & REBUFFAT, S. 2003. Microcin E492 antibacterial activity: evidence for a TonB-dependent inner membrane permeabilization on *Escherichia coli*. *Molecular Microbiology,* 49**,** 1031-1041.

DEUTSCHER, J. 1985. Phosphoenolpyruvate-dependent phosphorylation of a 55-kDa protein of *Streptococcus faecalis* catalyzed by the phosphotransferase system. *FEMS Microbiology Letters,* 29**,** 237-243.

DEUTSCHER, J., AKÉ, F. M. D., DERKAOUI, M., ZÉBRÉ, A. C., CAO, T. N., BOURAOUI, H., KENTACHE, T., MOKHTARI, A., MILOHANIC, E. & JOYET, P. 2014. The Bacterial Phosphoenolpyruvate:Carbohydrate Phosphotransferase System: Regulation by Protein Phosphorylation and Phosphorylation-Dependent Protein-Protein Interactions. *Microbiology and Molecular Biology Reviews,* 78**,** 231-256.

DEUTSCHER, J., FRANCKE, C. & POSTMA, P. W. 2006. How Phosphotransferase System-Related Protein Phosphorylation Regulates Carbohydrate Metabolism in Bacteria. *Microbiology and Molecular Biology Reviews,* 70**,** 939-1031.

DEUTSCHER, J., HERRO, R., BOURAND, A., MIJAKOVIC, I. & PONCET, S. 2005. P-Ser-HPr—a link between carbon metabolism and the virulence of some pathogenic bacteria. *Biochimica et Biophysica Acta (BBA) - Proteins and Proteomics,* 1754**,** 118-125.

DEUTSCHER, J., KUSTER, E., BERGSTEDT, U., CHARRIER, V. & HILLEN, W. 1995. Protein kinase-dependent HPr/CcpA interaction links glycolytic activity to carbon catabolite repression in Gram-positive bacteria. *Molecular Microbiology,* 15**,** 1049-1053.

DEUTSCHER, J. & SAIER, M. H., JR. 1983. ATP-dependent protein kinase-catalyzed phosphorylation of a seryl residue in HPr, a phosphate carrier protein of the phosphotransferase system in *Streptococcus pyogenes*. *Proceedings of the National Academy of Sciences of the United States of America,* 80**,** 6790-6794.

DEUTSCHER, J. & SAUERWALD, H. 1986. Stimulation of dihydroxyacetone and glycerol kinase activity in *Streptococcus faecalis* by phosphoenolpyruvate-dependent phosphorylation catalyzed by enzyme I and HPr of the phosphotransferase system. *Journal of Bacteriology,* 166**,** 829-836.

DIEP, D. B., H»VARSTEIN, L. S. & NES, I. F. 1995. A bacteriocin-like peptide induces bacteriocin synthesis in *Lactobacillus plantarum* C11. *Molecular Microbiology,* 18**,** 631-639.

DIEP, D. B., SKAUGEN, M., SALEHIAN, Z., HOLO, H. & NES, I. F. 2007. Common mechanisms of target cell recognition and immunity for class II bacteriocins. *Proceedings of the National Academy of Sciences of the United States of America,* 104**,** 2384-2389.

DOSSONNET, V., MONEDERO, V., ZAGOREC, M., GALINIER, A., PÉREZ-MARTÍNEZ, G. & DEUTSCHER, J. 2000. Phosphorylation of HPr by the Bifunctional HPr Kinase/P-Ser-HPr Phosphatase from *Lactobacillus casei* Controls Catabolite Repression and Inducer Exclusion but Not Inducer Expulsion. *Journal of Bacteriology,* 182**,** 2582-2590.

DURFEE, T., HANSEN, A.-M., ZHI, H., BLATTNER, F. R. & JIN, D. J. 2008. Transcription Profiling of the Stringent Response in *Escherichia coli*. *Journal of Bacteriology,* 190**,** 1084-1096.

ELKE RUHR & SAHL, H.-G. 1985. 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.

ELLIOTT, J. & ARBER, W. 1978. *E. coli* K-12 pel mutants, which block phage λ DNA injection, coincide with ptsM, which determines a component of a sugar transport system. *Molecular and General Genetics MGG,* 161**,** 1-8.

ERCOLINI, D., HILL, P. J. & DODD, C. E. 2003. Bacterial community structure and location in Stilton cheese. *Applied and Environmental Microbiology,* 69**,** 3540-3548.

ERNI, B., ZANOLARI, B. & KOCHER, H. P. 1987. The mannose permease of Escherichia coli consists of three different proteins. Amino acid sequence and function in sugar transport, sugar phosphorylation, and penetration of phage lambda DNA. *Journal of Biological Chemistry,* 262**,** 5238-5247.

ESQUINAS-RYCHEN, M. & ERNI, B. 2001. Facilitation of bacteriophage lambda DNA injection by inner membrane proteins of the bacterial phosphoenolpyruvate : carbohydrate phosphotransferase system (PTS). *Journal of Molecular Microbiology and Biotechnology,* 3**,** 361-370.

FELIS, G. E. & DELLAGLIO, F. 2007. Taxonomy of *Lactobacilli* and *Bifidobacteria*. *Current Issues in Intestinal Microbiology,* 8**,** 44-61.

FERULLO, D. J. & LOVETT, S. T. 2008. The Stringent Response and Cell Cycle Arrest in *Escherichia coli*. *PLoS Genetics,* 4**,** e1000300.

FINN, R. D., CLEMENTS, J. & EDDY, S. R. 2011. HMMER web server: interactive sequence similarity searching. *Nucleic Acids Research,* 39**,** W29-W37.

FRANCKE, C., POSTMA, P. W., WESTERHOFF, H. V., BLOM, J. G. & PELETIER, M. A. 2003. Why the Phosphotransferase System of *Escherichia coli* Escapes Diffusion Limitation. *Biophysical Journal,* 85**,** 612-622.

FRANCKE, C., WESTERHOFF, H. V., BLOM, J. G. & PELETIER, M. A. 2002. Flux control of the bacterial phosphoenolpyruvate:glucose phosphotransferase system and the effect of diffusion. *Molecular Biology Reports,* 29**,** 21-26.

FRANCL, A. L., THONGARAM, T. & MILLER, M. J. 2010. The PTS transporters of *Lactobacillus gasseri* ATCC 33323. *BioMed Central Microbiology,* 10**,** 1-13.

FREDERICK HARRIS, SARAH R. DENNISON & PHOENIX, D. A. 2009. Anionic antimicrobial peptides from eukaryotic organisms. *Current Protein and Peptide Science,* 10**,** 585-606.

FUJITA, Y., MIWA, Y., GALINIER, A. & DEUTSCHER, J. 1995. Specific recognition of the *Bacillus subtilis* gnt cis-acting catabolite-responsive element by a protein complex formed between CcpA and seryl-phosphorylated HPr. *Molecular Microbiology,* 17**,** 953-960.

GABRIELSEN, C., DAG A BREDE, PABLO E HERNÁNDEZ, INGOLF F NES & DIEP, D. B. 2012. The Maltose Abc Transporter In *Lactococcus Lactis* Facilitates High-Level Sensitivity To The Circular Bacteriocin Garvicin Ml. *Antimicrobial Agents and Chemotherapy,* 56**,** 2908-2915.

GARCIA DE GONZALO, C. V., ZHU, L., OMAN, T. J. & VAN DER DONK, W. A. 2014. NMR Structure of the *S*-Linked Glycopeptide Sublancin 168. *ACS Chemical Biology,* 9**,** 796-801.

GARDNER, N. J., SAVARD, T., OBERMEIER, P., CALDWELL, G. & CHAMPAGNE, C. P. 2001. Selection and characterization of mixed starter cultures for lactic acid fermentation of carrot, cabbage, beet and onion vegetable mixtures. *International Journal of Food Microbiology,* 64**,** 261-275.

GLADMAN, S. & SEEMANN, T. 2008. VelvetOptimiser. 2.2.5 ed. http://www.vicbioinformatics.com/software.velvetoptimiser.shtml.

GOFFIN, P., LORQUET, F., KLEEREBEZEM, M. & HOLS, P. 2004. Major Role of NAD-Dependent Lactate Dehydrogenases in Aerobic Lactate Utilization in *Lactobacillus plantarum* during Early Stationary Phase. *Journal of Bacteriology,* 186**,** 6661-6666.

GORKE, B. & STULKE, J. 2008. Carbon catabolite repression in bacteria: many ways to make the most out of nutrients. *Nature Reviews Microbiology,* 6**,** 613-624.

GOSALBES, M. A. J., ESTEBAN, C. D. & PÉREZ-MARTíNEZ, G. 2002. In vivo effect of mutations in the antiterminator LacT in *Lactobacillus casei*. *Microbiology,* 148**,** 695-702.

GOSALBES, M. J., MONEDERO, V., ALPERT, C.-A. & PÉREZ-MARTÍNEZ, G. 1997. Establishing a model to study the regulation of the lactose operon in *Lactobacillus casei*. *FEMS Microbiology Letters,* 148**,** 83-89.

GOSALBES, M. J., MONEDERO, V. & PÉREZ-MARTÍNEZ, G. 1999. Elements Involved in Catabolite Repression and Substrate Induction of the Lactose Operon in *Lactobacillus casei*. *Journal of Bacteriology,* 181**,** 3928-3934.

GRAVESEN, A., JYDEGAARD AXELSEN, A.-M., MENDES DA SILVA, J., HANSEN, T. B. & KNØCHEL, S. 2002. Frequency of Bacteriocin Resistance Development and Associated Fitness Costs in *Listeria monocytogenes*. *Applied and Environmental Microbiology,* 68**,** 756-764.

GREENBERG, D. B., STULKE, J. & SAIER, M. H., JR. 2002. Domain analysis of transcriptional regulators bearing PTS regulatory domains. *Research in Microbiology,* 153**,** 519-526.

GROPP, M., STRAUSZ, Y., GROSS, M. & GLASER, G. 2001. Regulation of *Escherichia coli* RelA Requires Oligomerization of the C-Terminal Domain. *Journal of Bacteriology,* 183**,** 570-579.

HASPER, H. E., DE KRUIJFF, B. & BREUKINK, E. 2004. Assembly and stability of nisin−lipid II pores. *Biochemistry,* 43**,** 11567-11575.

HATA, T., TANAKA, R. & OHMOMO, S. 2010. Isolation and characterization of plantaricin ASM1: A new bacteriocin produced by *Lactobacillus plantarum* A-1. *International Journal of Food Microbiology,* 137**,** 94-99.

HEDDLE, J. G., BLANCE, S. J., ZAMBLE, D. B., HOLLFELDER, F., MILLER, D. A., WENTZELL, L. M., WALSH, C. T. & MAXWELL, A. 2001. The antibiotic microcin B17 is a DNA gyrase poison: characterisation of the mode of inhibition. *Journal of Molecular Biology,* 307**,** 1223-1234.

HENSTRA, S. A., DUURKENS, R. H. & ROBILLARD, G. T. 2000. Multiple phosphorylation events regulate the activity of the mannitol transcriptional regulator MtlR of the *Bacillus stearothermophilus* phosphoenolpyruvate-dependent mannitol phosphotransferase system. *Journal of Biological Chemistry,* 275**,** 7037-7044.

HERAVI, K. M. & ALTENBUCHNER, J. 2014. Regulation of the *Bacillus subtilis* mannitol utilization genes: promoter structure and transcriptional activation by the wild-type regulator (MtlR) and its mutants. *Microbiology,* 160**,** 91-101.

HERNANDEZ, D., FRANÇOIS, P., FARINELLI, L., ØSTERÅS, M. & SCHRENZEL, J. 2008. *De novo* bacterial genome sequencing: Millions of very short reads assembled on a desktop computer. *Genome Research,* 18**,** 802-809.

HIGA, F. & EDELSTEIN, P. H. 2001. Potential Virulence Role of the *Legionella pneumophila ptsP* Ortholog. *Infection and Immunity,* 69**,** 4782-4789.

HOLO, H., NILSSEN, O. & NES, I. F. 1991. Lactococcin A, a new bacteriocin from *Lactococcus lactis* subsp. *cremoris*: isolation and characterization of the protein and its gene. *Journal of Bacteriology,* 173**,** 3879-3887.

HONDORP, E. R., HOU, S. C., HAUSE, L. L., GERA, K., LEE, C.-E. & MCIVER, K. S. 2013. PTS phosphorylation of Mga modulates regulon expression and virulence in the group A *streptococcus*. *Molecular Microbiology,* 88**,** 1176-1193.

HUANG, J., XU, Y., ZHANG, H., LI, Y., HUANG, X., REN, B. & ZHANG, X. 2009. Temperature-Dependent Expression of *phzM* and Its Regulatory Genes *lasI* and *ptsP* in Rhizosphere Isolate *Pseudomonas* sp. Strain M18. *Applied and Environmental Microbiology,* 75**,** 6568-6580.

HUNG, V. K. & MILLER, K. W. 2009. Regulation of mannose phosphotransferase system permease and virulence gene expression in *Listeria monocytogenes* by the EIIt Man transporter. *Applied and Environmental Microbiology,* 75**,** 6671-6678.

HURLEY, J. H., FABER, H. R., WORTHYLAKE, D., MEADOW, N. D., ROSEMAN, S., PETTIGREW, D. W. & REMINGTON, S. J. 1993. Structure of the regulatory complex of *Escherichia coli* IIIGlc with glycerol kinase. *Science,* 259**,** 673-677.

HYATT, D., CHEN, G. L., LOCASCIO, P. F., LAND, M. L., LARIMER, F. W. & HAUSER, L. J. 2010. Prodigal: prokaryotic gene recognition and translation initiation site identification. *BMC Bioinformatics,* 11**,** 1-11.

J NISSEN-MEYER, P ROGNE, C OPPEGÂRD, H HAUGEN & KRISTIANSEN, P. 2009. Structure-function relationships of the non-lanthionine-containing peptide (class II) bacteriocins produced by Gram-positive bacteria. *Current Pharmaceutical Biotechnology,* 10**,** 19-37.

JACK, R. W., TAGG, J. R. & RAY, B. 1995. Bacteriocins of Gram-positive bacteria. *Microbiological Reviews,* 59**,** 171-200.

JIN, Q., EOM, H. J., JUNG, J. Y., MOON, J. S., KIM, J. H. & HAN, N. S. 2012. Optimization of electrotransformation conditions for *Leuconostoc mesenteroides* subsp. *mesenteroides* ATCC8293. *Letters in Applied Microbiology,* 55**,** 314-321.

JIN, S., ISHIMOTO, K. & LORY, S. 1994. Nucleotide sequence of the *rpoN* gene and characterization of two downstream open reading frames in *Pseudomonas aeruginosa*. *Journal of Bacteriology,* 176**,** 1316-1322.

JONES, D. T. 1999. Protein secondary structure prediction based on position-specific scoring matrices. *Journal of Molecular Biology,* 292**,** 195-202.

JONES, D. T. 2007. Improving the accuracy of transmembrane protein topology prediction using evolutionary information. *Bioinformatics,* 23**,** 538-544.

JOSHI, N. A. & FASS, J. N. 2011. Sickle: A sliding-window, adaptive, quality-based trimming tool for FastQ files (Version 1.29). Available at https://github.com/najoshi/sickle.

JOYET, P., BOURAOUI, H., AKÉ, F. M. D., DERKAOUI, M., ZÉBRÉ, A. C., CAO, T. N., VENTROUX, M., NESSLER, S., NOIROT-GROS, M.-F., DEUTSCHER, J. & MILOHANIC, E. 2013. 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.

JOYET, P., DERKAOUI, M., PONCET, S. & DEUTSCHER, J. 2010. Control of *Bacillus subtilis* mtl operon expression by complex phosphorylation-dependent regulation of the transcriptional activator MtlR. *Molecular Microbiology,* 76**,** 1279-1294.

KALLBERG, M., WANG, H., WANG, S., PENG, J., WANG, Z., LU, H. & XU, J. 2012. Template-based protein structure modeling using the RaptorX web server. *Nature Protocols,* 7**,** 1511-1522.

KALMOKOFF, M. L. & TEATHER, R. M. 1997. Isolation and characterization of a bacteriocin (Butyrivibriocin AR10) from the ruminal anaerobe *Butyrivibrio fibrisolvens* AR10: evidence in support of the widespread occurrence of bacteriocin-like activity among ruminal isolates of *B. fibrisolvens*. *Applied and Environmental Microbiology,* 63**,** 394-402.

KARSTENS, K., ZSCHIEDRICH, C. P., BOWIEN, B., STÜLKE, J. & GÖRKE, B. 2014. Phosphotransferase protein EIIANtr interacts with SpoT, a key enzyme of the stringent response, in *Ralstonia eutropha* H16. *Microbiology,* 160**,** 711-722.

KATARZYNA, P. & MICHAEL, C. 2008. (p)ppGpp: Still Magical? *Annual Review of Microbiology,* 62**,** 35-51.

KELLEY, L. A. & STERNBERG, M. J. 2009. Protein Structure Prediction on the Web: A Case Study Using the Phyre Server. *Nature Protocols,* 4**,** 363-371.

KELLY, W. J., ASMUNDSON, R. V. & HUANG, C. M. 1996. Characterization of plantaricin KW30, a bacteriocin produced by *Lactobacillus plantarum*. *Journal of Applied Microbiology,* 81**,** 657-662.

KERR, A. 2013. *The Bacteriostatic Spectrum and Inhibitory Mechansim of Glycocin F, A Bacteriocin from Lactobacillus plantarum KW30.* Massey University, IMBS.

KIM, J. H., GUVENER, Z. T., CHO, J. Y., CHUNG, K. C. & CHAMBLISS, G. H. 1995. Specificity of DNA binding activity of the *Bacillus subtilis* catabolite control protein CcpA. *Journal of Bacteriology,* 177**,** 5129-5134.

KING, N. D. & O'BRIAN, M. R. 2001. Evidence for direct interaction between enzyme I(Ntr) and aspartokinase to regulate bacterial oligopeptide transport. *Journal of Biological Chemistry,* 276**,** 21311-21316.

KJOS, M., BORRERO, J., OPSATA, M., BIRRI, D. J., HOLO, H., CINTAS, L. M., SNIPEN, L., HERNÁNDEZ, P. E., NES, I. F. & DIEP, D. B. 2011a. Target recognition, resistance, immunity and genome mining of class II bacteriocins from Gram-positive bacteria. *Microbiology,* 157**,** 3256-3267.

KJOS, M., NES, I. F. & DIEP, D. B. 2009. Class II one-peptide bacteriocins target a phylogenetically defined subgroup of mannose phosphotransferase systems on sensitive cells. *Microbiology,* 155**,** 2949-2961.

KJOS, M., NES, I. F. & DIEP, D. B. 2011b. Mechanisms of Resistance to Bacteriocins Targeting the Mannose Phosphotransferase System. *Applied and Environmental Microbiology,* 77**,** 3335-3342.

KJOS, M., SALEHIAN, Z., NES, I. F. & DIEP, D. B. 2010a. An Extracellular Loop of the Mannose Phosphotransferase System Component IIC Is Responsible for Specific Targeting by Class IIa Bacteriocins. *Journal of Bacteriology,* 192**,** 5906-5913.

KJOS, M., SNIPEN, L., SALEHIAN, Z., NES, I. F. & DIEP, D. B. 2010b. The Abi proteins and their involvement in bacteriocin self-immunity. *Journal of Bacteriology,* 192**,** 2068-2076.

KLAENHAMMER, T. R. 1993. Genetics of bacteriocins produced by lactic acid bacteria. *FEMS Microbiology Reviews,* 12**,** 39-85.

KLEEREBEZEM, M., BOEKHORST, J., VAN KRANENBURG, R., MOLENAAR, D., KUIPERS, O. P., LEER, R., TARCHINI, R., PETERS, S. A., SANDBRINK, H. M., FIERS, M. W., STIEKEMA, W., LANKHORST, R. M., BRON, P. A., HOFFER, S. M., GROOT, M. N., KERKHOVEN, R., DE VRIES, M., URSING, B., DE VOS, W. M. & SIEZEN, R. J. 2003. Complete genome sequence of *Lactobacillus plantarum* WCFS1. *Proceedings of the National Academy of Sciences of the United States of America,* 100**,** 1990-1995.

KOUWEN, T. R. H. M., TRIP, E. N., DENHAM, E. L., SIBBALD, M. J. J. B., DUBOIS, J.-Y. F. & VAN DIJL, J. M. 2009. The large mechanosensitive channel MscL determines bacterial susceptibility to the bacteriocin sublancin 168. *Antimicrobial Agents and Chemotherapy,* 53**,** 4702-4711.

KRIEL, A., BITTNER, A. N., KIM, S. H., LIU, K., TEHRANCHI, A. K., ZOU, W. Y., RENDON, S., CHEN, R., TU, B. P. & WANG, J. D. 2012. Direct regulation of GTP homeostasis by (p)ppGpp: a critical component of viability and stress resistance. *Molecular Cell,* 48**,** 231-241.

KUNDIG, W., GHOSH, S. & ROSEMAN, S. 1964. Phosphate bound to histidine in a protein as an intermediate in a novel phospho-transferase system. *Proceedings of the National Academy of Sciences of the United States of America,* 52**,** 1067-1074.

LAEMMLI, U. K. 1970. Cleavage of structural proteins during the assembly of the head of bacteriophage T4. *Nature,* 227**,** 680-685.

LAGESEN, K., HALLIN, P., RØDLAND, E. A., STÆRFELDT, H.-H., ROGNES, T. & USSERY, D. W. 2007. RNAmmer: consistent and rapid annotation of ribosomal RNA genes. *Nucleic Acids Research,* 35**,** 3100-3108.

LAGOS, R., WILKENS, M., VERGARA, C., CECCHI, X. & MONASTERIO, O. 1993. Microcin E492 forms ion channels in phospholipid bilayer membrane. *FEBS Letters,* 321**,** 145-148.

LAMBERT, J. M., BONGERS, R. S. & KLEEREBEZEM, M. 2007. Cre-Lox-based system for multiple gene deletions and selectable-marker removal in *Lactobacillus plantarum*. *Applied and Environmental Microbiology,* 73**,** 1126-1135.

LASLETT, D. & CANBACK, B. 2004. ARAGORN, a program to detect tRNA genes and tmRNA genes in nucleotide sequences. *Nucleic Acids Research,* 32**,** 11-16.

LEER, R. J., VAN LUIJK, N., POSNO, M. & POUWELS, P. H. 1992. Structural and functional analysis of two cryptic plasmids from *Lactobacillus pentosus* MD353 and *Lactobacillus plantarum* ATCC 8014. *Molecular and General Genetics: MGG,* 234**,** 265-274.

LEI, J., LI, L. F. & SU, X. D. 2009. Crystal structures of phosphotransferase system enzymes PtxB (IIB(Asc)) and PtxA (IIA(Asc)) from *Streptococcus mutans*. *Journal of Molecular Biology,* 386**,** 465-475.

LEVINE, A., VANNIER, F., DEHBI, M., HENCKES, G. & SÉROR, S. J. 1991. The stringent response blocks DNA replication outside the ori region in *Bacillus subtilis* and at the origin in *Escherichia coli*. *Journal of Molecular Biology,* 219**,** 605-613.

LEYN, S. A., KAZANOV, M. D., SERNOVA, N. V., ERMAKOVA, E. O., NOVICHKOV, P. S. & RODIONOV, D. A. 2013. Genomic Reconstruction of the Transcriptional Regulatory Network in *Bacillus subtilis*. *Journal of Bacteriology,* 195**,** 2463-2473.

LI, H. & DURBIN, R. 2009. Fast and accurate short read alignment with Burrows–Wheeler transform. *Bioinformatics,* 25**,** 1754-1760.

LIN, S.-H. & LIAO, Y.-C. 2013. CISA: Contig Integrator for Sequence Assembly of bacterial genomes. *PLoS ONE,* 8**,** e60843.

LIU, B.-S., LI, G.-G., YU, Z.-Q., HAN, B. & ZHANG, R.-J. 2011. Evaluation of bacteriocin resistance in *Staphylococcus aureus* against the bacteriocin complex secreted by *Bacillus subtilis* LFB112. *Journal of Animal and Veterinary Advances,* 10**,** 1743-1749.

LOESSNER, M., GUENTHER, S., STEFFAN, S. & SCHERER, S. 2003. A Pediocin-Producing *Lactobacillus plantarum* Strain Inhibits *Listeria monocytogenes* in a Multispecies Cheese Surface Microbial Ripening Consortium. *Applied and Environmental Microbiology,* 69**,** 1854-1857.

LOLKEMA, J. S. & ROBILLARD, G. T. 1990. Subunit structure and activity of the mannitol-specific enzyme II of the *Escherichia coli* phosphoenolpyruvate-dependent phosphotransferase system solubilized in detergent. *Biochemistry,* 29**,** 10120-10125.

LOO, T. 2014. *RE: Personal Communication: Details of Mass Spectrometery Analysis.*

LOPES, J. L. S., GÓMARA, M. J., HARO, I., TONARELLI, G. & BELTRAMINI, L. M. 2013. Contribution of the Tyr-1 in plantaricin149a to disrupt phospholipid model membranes. *International Journal of Molecular Sciences,* 14**,** 12313-12328.

LORCA, G. L., BARABOTE, R. D., ZLOTOPOLSKI, V., TRAN, C., WINNEN, B., HVORUP, R. N., STONESTROM, A. J., NGUYEN, E., HUANG, L.-W., KIM, D. S. & SAIER JR, M. H. 2007.

Transport capabilities of eleven gram-positive bacteria: Comparative genomic analyses. *Biochimica et Biophysica Acta (BBA) - Biomembranes,* 1768**,** 1342-1366.

LUTTMANN, D., HEERMANN, R., ZIMMER, B., HILLMANN, A., RAMPP, I. S., JUNG, K. & GORKE, B. 2009. Stimulation of the potassium sensor KdpD kinase activity by interaction with the phosphotransferase protein IIA(Ntr) in *Escherichia coli*. *Molecular Microbiology,* 72**,** 978-994.

LUX, R., MUNASINGHE, V. R., CASTELLANO, F., LENGELER, J. W., CORRIE, J. E. & KHAN, S. 1999. Elucidation of a PTS-carbohydrate chemotactic signal pathway in *Escherichia coli* using a time-resolved behavioral assay. *Molecular Biology of the Cell,* 10**,** 1133-1346.

MAGOČ, T. & SALZBERG, S. L. 2011. FLASH: fast length adjustment of short reads to improve genome assemblies. *Bioinformatics,* 27**,** 2957-2963.

MAISNIER-PATIN, S., FORNI, E. & RICHARD, J. 1996. Purification, partial characterisation and mode of action of enterococcin EFS2, an antilisterial bacteriocin produced by a strain of *Enterococcus faecalis* isolated from a cheese. *International Journal of Food Microbiology,* 30**,** 255-270.

MAITI, R., VAN DOMSELAAR, G. H., ZHANG, H. & WISHART, D. S. 2004. SuperPose: a simple server for sophisticated structural superposition. *Nucleic Acids Research,* 32**,** W590-W594.

MARCINIAK, B. C., PABIJANIAK, M., DE JONG, A., DUHRING, R., SEIDEL, G., HILLEN, W. & KUIPERS, O. P. 2012. High- and low-affinity cre boxes for CcpA binding in *Bacillus subtilis* revealed by genome-wide analysis. *BMC Genomics,* 13**,** 1-15.

MATTICK, A. T. R. & HIRSCH, A. 1944. A powerful inhibitory substance produced by group N *Streptococci*. *Nature,* 154**,** 551-551.

MAVRODI, O. V., MAVRODI, D. V., WELLER, D. M. & THOMASHOW, L. S. 2006. Role of *ptsP*, orfT, and *sss* Recombinase Genes in Root Colonization by *Pseudomonas fluorescens* Q8r1-96. *Applied and Environmental Microbiology,* 72**,** 7111-7122.

MCCOY, J. G., LEVIN, E. J. & ZHOU, M. 2014. Structural insight into the PTS sugar transporter EIIC. *Biochimica et Biophysica Acta (BBA) - General Subjects* [Online]. Available: http://www.sciencedirect.com/science/article/pii/S0304416514001172.

MEADOW, N. D. & ROSEMAN, S. 1996. Rate and Equilibrium Constants for Phosphoryltransfer between Active Site Histidines of Escherichia coli HPr and the Signal Transducing Protein IIIGlc. *Journal of Biological Chemistry,* 271**,** 33440-33445.

METLITSKAYA, A., KAZAKOV, T., KOMMER, A., PAVLOVA, O., PRAETORIUS-IBBA, M., IBBA, M., KRASHENINNIKOV, I., KOLB, V., KHMEL, I. & SEVERINOV, K. 2006. Aspartyl-tRNA synthetase is the target of peptide nucleotide antibiotic Microcin C. *Journal of Biological Chemistry,* 281**,** 18033-18042.

MITCHELL, W. J., SHAW, J. E. & ANDREWS, L. 1991. Properties of the glucose phosphotransferase system of *Clostridium acetobutylicum* NCIB 8052. *Applied and Environmental Microbiology,* 57**,** 2534-2539.

MÖLLER, C., RAHMANKHAH, S., LAUER-FIELDS, J., BUBIS, J., FIELDS, G. B. & MARÍ, F. 2005. A Novel Conotoxin Framework with a Helix−Loop−Helix (Cs α/α) Fold†. *Biochemistry,* 44**,** 15986-15996.

MURIANA, P. M. & KLAENHAMMER, T. R. 1991a. Cloning, phenotypic expression, and DNA sequence of the gene for lactacin F, an antimicrobial peptide produced by *Lactobacillus* spp. *Journal of Bacteriology,* 173**,** 1779-1788.

MURIANA, P. M. & KLAENHAMMER, T. R. 1991b. Purification and partial characterization of lactacin F, a bacteriocin produced by *Lactobacillus acidophilus* 11088. *Applied and Environmental Microbiology,* 57**,** 114-121.

NAVDAEVA, V., ZURBRIGGEN, A., WALTERSPERGER, S., SCHNEIDER, P., OBERHOLZER, A. E., BAHLER, P., BACHLER, C., GRIEDER, A., BAUMANN, U. & ERNI, B. 2011. Phosphoenolpyruvate: sugar phosphotransferase system from the hyperthermophilic *Thermoanaerobacter tengcongensis*. *Biochemistry,* 50**,** 1184-1193.

NAWROCKI, E. P. & EDDY, S. R. 2013. Infernal 1.1: 100-fold faster RNA homology searches. *Bioinformatics,* 29**,** 2933-2935.

NEWBY, Z. E. R., O'CONNELL, J. D., GRUSWITZ, F., HAYS, F. A., HARRIES, W. E. C., HARWOOD, I. M., HO, J. D., LEE, J. K., SAVAGE, D. F., MIERCKE, L. J. W. & STROUD, R. M. 2009. A general protocol for the crystallization of membrane proteins for X-ray structural investigation. *Nature Protocols,* 4**,** 619-637.

NGUYEN, T. X., YEN, M. R., BARABOTE, R. D. & SAIER, M. H., JR. 2006. Topological predictions for integral membrane permeases of the phosphoenolpyruvate:sugar phosphotransferase system. *Journal of Molecular Microbiology and Biotechnology,* 11**,** 345-360.

NOLAN, E. M., FISCHBACH, M. A., KOGLIN, A. & WALSH, C. T. 2007. Biosynthetic tailoring of microcin E492m: post-translational modification affords an antibacterial siderophore-peptide conjugate. *Journal of the American Chemical Society,* 129**,** 14336-14347.

NURK, S., BANKEVICH, A., ANTIPOV, D., GUREVICH, A., KOROBEYNIKOV, A., LAPIDUS, A., PRJIBELSKY, A., PYSHKIN, A., SIROTKIN, A., SIROTKIN, Y., STEPANAUSKAS, R., MCLEAN, J., LASKEN, R., CLINGENPEEL, S., WOYKE, T., TESLER, G., ALEKSEYEV, M. & PEVZNER, P. 2013. Assembling Genomes and Mini-metagenomes from Highly Chimeric Reads. *In:* DENG, M., JIANG, R., SUN, F. & ZHANG, X. (eds.) *Research in Computational Molecular Biology.* Springer Berlin Heidelberg.

NUXOLL, A. S., HALOUSKA, S. M., SADYKOV, M. R., HANKE, M. L., BAYLES, K. W., KIELIAN, T., POWERS, R. & FEY, P. D. 2012. CcpA regulates arginine biosynthesis in *Staphylococcus aureus* through repression of proline catabolism. *PLoS Pathogens,* 8**,** e1003033.

O'DONNELL, M. M., O'TOOLE, P. W. & ROSS, R. P. 2013. Catabolic flexibility of mammalian-associated lactobacilli. *Microbial Cell Factories,* 12**,** 1-11.

OHSE, M., TAKAHASHI, K., KADOWAKI, Y. & KUSAOKE, H. 1995. Effects of Plasmid DNA Sizes and Several Other Factors on Transformation of *Bacillus subtilis* ISW1214 with Plasmid DNA by Electroporation. *Bioscience, Biotechnology, and Biochemistry,* 59**,** 1433-1437.

OMAN, T. J., BOETTCHER, J. M., WANG, H., OKALIBE, X. N. & VAN DER DONK, W. A. 2011. Sublancin is not a lantibiotic but an s-linked glycopeptide. *Nature Chemical Biology,* 7**,** 78-80.

OSUMI, T. & SAIER, M. H., JR. 1982. Regulation of lactose permease activity by the phosphoenolpyruvate:sugar phosphotransferase system: evidence for direct binding of the glucose-specific enzyme III to the lactose permease. *Proceedings of the National Academy of Sciences of the United States of America,* 79**,** 1457-1461.

PAIK, S. H., CHAKICHERLA, A. & HANSEN, J. N. 1998. Identification and characterization of the structural and transporter genes for, and the chemical and biological properties of, sublancin 168, a novel lantibiotic produced by *Bacillus subtilis* 168. *Journal of Biological Chemistry,* 273**,** 23134-23142.

PALOMINO, M. M., ALLIEVI, M. C., PRADO-ACOSTA, M., SANCHEZ-RIVAS, C. & RUZAL, S. M. 2010. New method for electroporation of *Lactobacillus* species grown in high salt. *Journal of Microbiological Methods,* 83**,** 164-167.

PAPAGIANNI, M., AVRAMIDIS, N. & FILIOUSSIS, G. 2007. High efficiency electrotransformation of *Lactococcus lactis* spp. *lactis* cells pretreated with lithium acetate and dithiothreitol. *BMC Biotechnology,* 7**,** 1-6.

PATCHETT, M. L. & NORRIS, G. E. 2012. *RE: Personal Communication.*

PATEL, R. K. & JAIN, M. 2012. NGS QC Toolkit: A Toolkit for Quality Control of Next Generation Sequencing Data. *PLoS ONE,* 7**,** e30619.

PATZER, S. I., BAQUERO, M. R., BRAVO, D., MORENO, F. & HANTKE, K. 2003. The colicin G, H and X determinants encode microcins M and H47, which might utilize the catecholate siderophore receptors FepA, Cir, Fiu and IroN. *Microbiology,* 149**,** 2557-2570.

PELTON, J. G., TORCHIA, D. A., REMINGTON, S. J., MURPHY, K. P., MEADOW, N. D. & ROSEMAN, S. 1996. Structures of Active Site Histidine Mutants of IIIGlc, a Major Signal-transducing Protein in Escherichia coli: Effects On The Mechanisms Of Regulation And Phosphoryl Transfer. *Journal of Biological Chemistry,* 271**,** 33446-33456.

PETERSEN, T. N., BRUNAK, S., VON HEIJNE, G. & NIELSEN, H. 2011. SignalP 4.0: discriminating signal peptides from transmembrane regions. *Nature Methods,* 8**,** 785-786.

PFLÜGER-GRAU, K. & DE LORENZO, V. 2014. From the phosphoenolpyruvate phosphotransferase system to selfish metabolism: a story retraced in *Pseudomonas putida*. *FEMS Microbiology Letters,* 356**,** 144-153.

PIERRAT, O. A. & MAXWELL, A. 2003. The action of the bacterial toxin microcin B17. Insight into the cleavage-religation reaction of DNA gyrase. *Journal of Biological Chemistry,* 278**,** 35016-35023.

PONCET, S., SORET, M., MERVELET, P., DEUTSCHER, J. & NOIROT, P. 2009. Transcriptional activator YesS is stimulated by histidine-phosphorylated HPr of the *Bacillus subtilis* phosphotransferase system. *Journal of Biological Chemistry,* 284**,** 28188-28197.

POOLMAN, B., ROYER, T. J., MAINZER, S. E. & SCHMIDT, B. F. 1989. Lactose transport system of *Streptococcus thermophilus*: a hybrid protein with homology to the melibiose carrier

and enzyme III of phosphoenolpyruvate-dependent phosphotransferase systems. *Journal of Bacteriology,* 171**,** 244-253.

POSTMA, P. W., EPSTEIN, W., SCHUITEMA, A. R. & NELSON, S. O. 1984. Interaction between IIIGlc of the phosphoenolpyruvate:sugar phosphotransferase system and glycerol kinase of *Salmonella typhimurium*. *Journal of Bacteriology,* 158**,** 351-353.

POSTMA, P. W., LENGELER, J. W. & JACOBSON, G. R. 1993. Phosphoenolpyruvate:carbohydrate phosphotransferase systems of bacteria. *Microbiological Reviews,* 57**,** 543-594.

POWELL, B. S., COURT, D. L., INADA, T., NAKAMURA, Y., MICHOTEY, V., CUI, X., REIZER, A., SAIER, M. H., JR. & REIZER, J. 1995. Novel proteins of the phosphotransferase system encoded within the *rpoN* operon of Escherichia coli. Enzyme IIANtr affects growth on organic nitrogen and the conditional lethality of an erats mutant. *Journal of Biological Chemistry,* 270**,** 4822-4839.

PUGSLEY, A. P., MORENO, F. & DE LORENZO, V. 1986. Microcin-E492-insensitive mutants of *Escherichia coli* K12. *Journal of General Microbiology,* 132**,** 3253-3259.

REIZER, J., HOISCHEN, C., TITGEMEYER, F., RIVOLTA, C., RABUS, R., STÜLKE, J., KARAMATA, D., SAIER JR, M. H. & HILLEN, W. 1998. A novel protein kinase that controls carbon catabolite repression in bacteria. *Molecular Microbiology,* 27**,** 1157-1169.

REIZER, J. & SAIER, M. H., JR. 1997. Modular multidomain phosphoryl transfer proteins of bacteria. *Current Opinion in Structural Biology,* 7**,** 407-415.

ROGERS, L. A. 1928. The inhibiting effect of *Streptococcus lactis* on *Lactobacillus bulgaricus*. *Journal of Bacteriology,* 16**,** 321-325.

ROHWER, J. M., MEADOW, N. D., ROSEMAN, S., WESTERHOFF, H. V. & POSTMA, P. W. 2000. Understanding Glucose Transport by the Bacterial Phosphoenolpyruvate:Glycose Phosphotransferase System on the Basis of Kinetic Measurements in Vitro. *Journal of Biological Chemistry,* 275**,** 34909-34921.

ROSS, W., VRENTAS, CATHERINE E., SANCHEZ-VAZQUEZ, P., GAAL, T. & GOURSE, RICHARD L. 2013. The Magic Spot: A ppGpp Binding Site on *E. coli* RNA Polymerase Responsible for Regulation of Transcription Initiation. *Molecular Cell,* 50**,** 420-429.

ROTHE, F. M., BAHR, T., STULKE, J., RAK, B. & GORKE, B. 2012. Activation of *Escherichia coli* antiterminator BglG requires its phosphorylation. *Proceedings of the National Academy of Sciences of the United States of America,* 109**,** 15906-15911.

ROTHE, F. M., WREDE, C., LEHNIK-HABRINK, M., GÖRKE, B. & STÜLKE, J. 2013. Dynamic Localization of a Transcription Factor in *Bacillus subtilis*: the LicT Antiterminator Relocalizes in Response to Inducer Availability. *Journal of Bacteriology,* 195**,** 2146-2154.

RUIJTER, G. J., VAN MEURS, G., VERWEY, M. A., POSTMA, P. W. & VAN DAM, K. 1992. Analysis of mutations that uncouple transport from phosphorylation in enzyme IIGlc of the *Escherichia coli* phosphoenolpyruvate-dependent phosphotransferase system. *Journal of Bacteriology,* 174**,** 2843-2850.

RUSSELL, W. M. & KLAENHAMMER, T. R. 2001. Efficient system for directed integration into the *Lactobacillus acidophilus* and *Lactobacillus gasseri* chromosomes via homologous recombination. *Applied and Environmental Microbiology,* 67**,** 4361-4364.

SADYKOV, M. R., HARTMANN, T., MATTES, T. A., HIATT, M., JANN, N. J., ZHU, Y., LEDALA, N., LANDMANN, R., HERRMANN, M., ROHDE, H., BISCHOFF, M. & SOMERVILLE, G. A. 2011. CcpA coordinates central metabolism and biofilm formation in *Staphylococcus epidermidis*. *Microbiology,* 157**,** 3458-3468.

SAHL, H. G., KORDEL, M. & BENZ, R. 1987. Voltage-dependent depolarization of bacterial membranes and artificial lipid bilayers by the peptide antibiotic nisin. *Archives of Microbiology,* 149**,** 120-124.

SAIER, M. H., HVORUP, R. N. & BARABOTE, R. D. 2005. Evolution of the bacterial phosphotransferase system: from carriers and enzymes to group translocators. *Biochemical Society Transactions,* 33**,** 220-224.

SAIER, M. H., JR., NOVOTNY, M. J., COMEAU-FUHRMAN, D., OSUMI, T. & DESAI, J. D. 1983. Cooperative binding of the sugar substrates and allosteric regulatory protein (enzyme IIIGlc of the phosphotransferase system) to the lactose and melibiose permeases in *Escherichia coli* and *Salmonella typhimurium*. *Journal of Bacteriology,* 155**,** 1351-1357.

SAIER, M. H., REDDY, V. S., TAMANG, D. G. & VÄSTERMARK, Å. 2014. The Transporter Classification Database. *Nucleic Acids Research,* 42**,** D251-D258.

SAIER, M. H., TRAN, C. V. & BARABOTE, R. D. 2006. TCDB: the Transporter Classification Database for membrane transport protein analyses and information. *Nucleic Acids Research,* 34**,** D181-D186.

SAIER, M. H., YEN, M. R., NOTO, K., TAMANG, D. G. & ELKAN, C. 2009. The Transporter Classification Database: recent advances. *Nucleic Acids Research,* 37**,** D274-D278.

SALER, M. H. & REIZER, J. 1994. The bacterial phosphotransferase system: new frontiers 30 years later. *Molecular Microbiology,* 13**,** 755-764.

SÁNCHEZ, J., BASANTA, A., GÓMEZ-SALA, B., HERRANZ, C., CINTAS, L. M. & HERNÁNDEZ, P. E. 2007. Antimicrobial and safety aspects, and biotechnological potential of bacteriocinogenic *enterococci* isolated from Mallard ducks (*Anas platyrhynchos*). *International Journal of Food Microbiology,* 117**,** 295-305.

SCHATZ, M. C., PHILLIPPY, A. M., SOMMER, D. D., DELCHER, A. L., PUIU, D., NARZISI, G., SALZBERG, S. L. & POP, M. 2013. Hawkeye and AMOS: visualizing and assessing the quality of genome assemblies. *Briefings in Bioinformatics,* 14**,** 213-224.

SCHNETZ, K., SUTRINA, S. L., SAIER, M. H., JR. & RAK, B. 1990. Identification of catalytic residues in the beta-glucoside permease of *Escherichia coli* by site-specific mutagenesis and demonstration of interdomain cross-reactivity between the beta-glucoside and glucose systems. *Journal of Biological Chemistry,* 265**,** 13464-13471.

SCHREIBER, G., RON, E. & GLASER, G. 1995. ppGpp-mediated regulation of DNA replication and cell division in *Escherichia coli*. *Current Microbiology,* 30**,** 27-32.

SCHRODINGER 2010. The PyMOL Molecular Graphics System, Version 1.3.

SCHUMACHER, M. A., ALLEN, G. S., DIEL, M., SEIDEL, G., HILLEN, W. & BRENNAN, R. G. 2004. Structural Basis for Allosteric Control of the Transcription Regulator CcpA by the Phosphoprotein HPr-Ser46-P. *Cell,* 118**,** 731-741.

SCHUMACHER, M. A., SPREHE, M., BARTHOLOMAE, M., HILLEN, W. & BRENNAN, R. G. 2011. Structures of carbon catabolite protein A–(HPr-Ser46-P) bound to diverse catabolite response element sites reveal the basis for high-affinity binding to degenerate DNA operators. *Nucleic Acids Research,* 39**,** 2931-2942.

SEEMANN, T. 2014. Prokka: rapid prokaryotic genome annotation. *Bioinformatics,* 30**,** 2068-2069.

SEGURA, D. & ESPIN, G. 1998. Mutational inactivation of a gene homologous to *Escherichia coli ptsP* affects poly-beta-hydroxybutyrate accumulation and nitrogen fixation in *Azotobacter vinelandii*. *Journal of Bacteriology,* 180**,** 4790-4798.

SHARMA, A. & SRIVASTAVA, S. 2014. Anti-Candida activity of two-peptide bacteriocins, plantaricins (Pln E/F and J/K) and their mode of action. *Fungal Biology,* 118**,** 264-275.

SIEZEN, R. & VAN HYLCKAMA VLIEG, J. 2011. Genomic diversity and versatility of *Lactobacillus plantarum*, a natural metabolic engineer. *Microbial Cell Factories,* 10**,** 1-13.

SIMONI, R. D., LEVINTHAL, M., KUNDIG, F. D., KUNDIG, W., ANDERSON, B., HARTMAN, P. E. & ROSEMAN, S. 1967. Genetic evidence for the role of a bacterial phosphotransferase system in sugar transport. *Proceedings of the National Academy of Sciences of the United States of America,* 58**,** 1963-1970.

SIMONI, R. D. & ROSEMAN, S. 1973. Sugar transport. *Journal of Biological Chemistry,* 248**,** 966-973.

SIMONI, R. D., ROSEMAN, S. & SAIER, M. H. 1976. Sugar transport. Properties of mutant bacteria defective in proteins of the phosphoenolpyruvate: sugar phosphotransferase system. *Journal of Biological Chemistry,* 251**,** 6584-6597.

SOLOVYEV, V. & SALAMOV, A. 2011. Automatic Annotation of Microbial Genomes and Metagenomic Sequences. *In Metagenomics and its Applications in Agriculture, Biomedicine and Environmental Studies.* Nova Science Publishers.

SØRVIG, E., MATHIESEN, G., NATERSTAD, K., EIJSINK, V. G. H. & AXELSSON, L. 2005. High-level, inducible gene expression in *Lactobacillus sakei* and *Lactobacillus plantarum* using versatile expression vectors. *Microbiology,* 151**,** 2439-2449.

SRIVATSAN, A. & WANG, J. D. 2008. Control of bacterial transcription, translation and replication by (p)ppGpp. *Current Opinion in Microbiology,* 11**,** 100-105.

STAJICH, J. E., BLOCK, D., BOULEZ, K., BRENNER, S. E., CHERVITZ, S. A., DAGDIGIAN, C., FUELLEN, G., GILBERT, J. G. R., KORF, I., LAPP, H., LEHVÄSLAIHO, H., MATSALLA, C., MUNGALL, C. J., OSBORNE, B. I., POCOCK, M. R., SCHATTNER, P., SENGER, M., STEIN, L. D., STUPKA, E., WILKINSON, M. D. & BIRNEY, E. 2002. The Bioperl Toolkit: Perl Modules for the Life Sciences. *Genome Research,* 12**,** 1611-1618.

STEPPER, J., SHASTRI, S., LOO, T. S., PRESTON, J. C., NOVAK, P., MAN, P., MOORE, C. H., HAVLÍČEK, V., PATCHETT, M. L. & NORRIS, G. E. 2011. Cysteine s-glycosylation, a new

post-translational modification found in glycopeptide bacteriocins. *FEBS letters,* 585**,** 645-650.

STEVENS, K. A., SHELDON, B. W., KLAPES, N. A. & KLAENHAMMER, T. R. 1991. Nisin treatment for inactivation of *Salmonella* species and other gram-negative bacteria. *Applied and Environmental Microbiology,* 57**,** 3613-3615.

STRAHSBURGER, E., BAEZA, M., MONASTERIO, O. & LAGOS, R. 2005. Cooperative uptake of microcin E492 by receptors FepA, Fiu, and Cir and inhibition by the siderophore enterochelin and its dimeric and trimeric hydrolysis products. *Antimicrobial Agents and Chemotherapy,* 49**,** 3083-3086.

STRIEBEL, F., IMKAMP, F., ÖZCELIK, D. & WEBER-BAN, E. 2014. Pupylation as a signal for proteasomal degradation in bacteria. *Biochimica et Biophysica Acta (BBA) - Molecular Cell Research,* 1843**,** 103-113.

STULKE, J. & HILLEN, W. 2000. Regulation of carbon catabolism in *Bacillus* species. *Annual Review of Microbiology,* 54**,** 849-880.

TAKAYUKI, S. & VAN DER HOORN, R. A. L. 2008. Papain-like cysteine proteases: key players at molecular battlefields employed by both plants and their invaders. *Molecular Plant Pathology,* 9**,** 119-125.

TAN, M. W., RAHME, L. G., STERNBERG, J. A., TOMPKINS, R. G. & AUSUBEL, F. M. 1999. *Pseudomonas aeruginosa* killing of *Caenorhabditis elegans* used to identify *P. aeruginosa* virulence factors. *Proceedings of the National Academy of Sciences of the United States of America,* 96**,** 2408-2413.

THOMAS, X., DESTOUMIEUX-GARZON, D., PEDUZZI, J., AFONSO, C., BLOND, A., BIRLIRAKIS, N., GOULARD, C., DUBOST, L., THAI, R., TABET, J. C. & REBUFFAT, S. 2004. Siderophore peptide, a new type of post-translationally modified antibacterial peptide with potent activity. *Journal of Biological Chemistry,* 279**,** 28233-28242.

TIELEMAN, D. P. 2004. The molecular basis of electroporation. *BMC Biochemistry,* 5**,** 1-12.

TORTOSA, P., DECLERCK, N., DUTARTRE, H., LINDNER, C., DEUTSCHER, J. & LE COQ, D. 2001. Sites of positive and negative regulation in the *Bacillus subtilis* antiterminators LicT and SacY. *Molecular Microbiology,* 41**,** 1381-1393.

UNTIET, V., KARUNAKARAN, R., KRAMER, M., POOLE, P., PRIEFER, U. & PRELL, J. 2013. ABC transport is inactivated by the PTS(Ntr) under potassium limitation in *Rhizobium leguminosarum* 3841. *PLoS ONE,* 8**,** e64682.

VAN BELKUM, M. J., KOK, J., VENEMA, G., HOLO, H., NES, I. F., KONINGS, W. N. & ABEE, T. 1991. The bacteriocin lactococcin A specifically increases permeability of *lactococcal* cytoplasmic membranes in a voltage-independent, protein-mediated manner. *Journal of Bacteriology,* 173**,** 7934-7941.

VAN BELKUM, M. J., MARTIN-VISSCHER, L. A. & VEDERAS, J. C. 2011. Structure and genetics of circular bacteriocins. *Trends in Microbiology,* 19**,** 411-418.

VAN DE GUCHTE, M., VAN DER VOSSEN, J. M., KOK, J. & VENEMA, G. 1989. Construction of a lactococcal expression vector: expression of hen egg white lysozyme in *Lactococcus lactis* subsp. *lactis*. *Applied and Environmental Microbiology,* 55**,** 224-228.

VAN DER VLAG, J., VAN DAM, K. & POSTMA, P. W. 1994. Quantification of the regulation of glycerol and maltose metabolism by IIAGlc of the phosphoenolpyruvate-dependent glucose phosphotransferase system in *Salmonella typhimurium*. *Journal of Bacteriology,* 176**,** 3518-3526.

VAN DER VOSSEN, J. M., VAN DER LELIE, D. & VENEMA, G. 1987. Isolation and characterization of *Streptococcus cremoris* Wg2-specific promoters. *Applied and Environmental Microbiology,* 53**,** 2452-2457.

VAN KRANENBURG, R., GOLIC, N., BONGERS, R., LEER, R. J., DE VOS, W. M., SIEZEN, R. J. & KLEEREBEZEM, M. 2005. Functional Analysis of Three Plasmids from *Lactobacillus plantarum*. *Applied and Environmental Microbiology,* 71**,** 1223-1230.

VENDITTI, V., GHIRLANDO, R. & CLORE, G. M. 2013. Structural basis for enzyme i inhibition by α-ketoglutarate. *ACS Chemical Biology,* 8**,** 1232-1240.

VENEMA, K., ABEE, T., HAANDRIKMAN, A. J., LEENHOUTS, K. J., KOK, J., KONINGS, W. N. & VENEMA, G. 1993. Mode of action of lactococcin B, a thiol-activated bacteriocin from *Lactococcus lactis*. *Applied and Environmental Microbiology,* 59**,** 1041-1048.

VENUGOPAL, H., EDWARDS, P. J. B., SCHWALBE, M., CLARIDGE, J. K., LIBICH, D. S., STEPPER, J., LOO, T., PATCHETT, M. L., NORRIS, G. E. & PASCAL, S. M. 2011. Structural, dynamic, and chemical characterization of a novel S-glycosylated bacteriocin. *Biochemistry,* 50**,** 2748-2755.

VILLEGAS, A. & KROPINSKI, A. M. 2008. An analysis of initiation codon utilization in the Domain Bacteria – concerns about the quality of bacterial genome annotation. *Microbiology,* 154**,** 2559-2661.

WANG, H., OMAN, T. J., ZHANG, R., GARCIA DE GONZALO, C. V., ZHANG, Q. & VAN DER DONK, W. A. 2013. The Glycosyltransferase Involved in Thurandacin Biosynthesis Catalyzes Both *O*- and *S*-Glycosylation. *Journal of the American Chemical Society,* 136**,** 84-87.

WANG, H. & VAN DER DONK, W. A. 2011. Substrate Selectivity of the Sublancin S-Glycosyltransferase. *Journal of the American Chemical Society,* 133**,** 16394-16397.

WEAVER, J. C. & CHIZMADZHEV, Y. A. 1996. Theory of electroporation: A review. *Bioelectrochemistry and Bioenergetics,* 41**,** 135-160.

WEIGEL, N., KUKURUZINSKA, M. A., NAKAZAWA, A., WAYGOOD, E. B. & ROSEMAN, S. 1982. Sugar transport by the bacterial phosphotransferase system. Phosphoryl transfer reactions catalyzed by enzyme I of *Salmonella typhimurium*. *Journal of Biological Chemistry,* 257**,** 14477-14491.

WENG, Q. P., ELDER, J. & JACOBSON, G. R. 1992. Site-specific mutagenesis of residues in the *Escherichia coli* mannitol permease that have been suggested to be important for its phosphorylation and chemoreception functions. *Journal of Biological Chemistry,* 267**,** 19529-19535.

WENG, Q. P. & JACOBSON, G. R. 1993. Role of a conserved histidine residue, His-195, in the activities of the *Escherichia coli* mannitol permease. *Biochemistry,* 32**,** 11211-11216.

WENZEL, M. & ALTENBUCHNER, J. 2013. The *Bacillus subtilis* mannose regulator, ManR, a DNA-binding protein regulated by HPr and its cognate PTS transporter ManP. *Molecular Microbiology,* 88**,** 562-576.

WIEDEMANN, I., BENZ, R. & SAHL, H.-G. 2004. Lipid II-mediated pore formation by the peptide antibiotic nisin: a black lipid membrane study. *Journal of Bacteriology,* 186**,** 3259-3261.

WIEDEMANN, I., BREUKINK, E., VAN KRAAIJ, C., KUIPERS, O. P., BIERBAUM, G., DE KRUIJFF, B. & SAHL, H.-G. 2001. Specific binding of nisin to the peptidoglycan precursor lipid II combines pore formation and inhibition of cell wall biosynthesis for potent antibiotic activity. *Journal of Biological Chemistry,* 276**,** 1772-1779.

WILLIAMS, N., FOX, D. K., SHEA, C. & ROSEMAN, S. 1986. Pel, the protein that permits lambda DNA penetration of *Escherichia coli*, is encoded by a gene in ptsM and is required for mannose utilization by the phosphotransferase system. *Proceedings of the National Academy of Sciences of the United States of America,* 83**,** 8934-8938.

XIAO, H., KALMAN, M., IKEHARA, K., ZEMEL, S., GLASER, G. & CASHEL, M. 1991. Residual guanosine 3',5'-bispyrophosphate synthetic activity of *relA* null mutants can be eliminated by *spoT* null mutations. *Journal of Biological Chemistry,* 266**,** 5980-5990.

XUE, J. & MILLER, K. W. 2007. Regulation of the mpt Operon in *Listeria* innocua by the ManR Protein. *Applied and Environmental Microbiology,* 73**,** 5648-5652.

YEH, J. I., KETTERING, R., SAXL, R., BOURAND, A., DARBON, E., JOLY, N., BRIOZZO, P. & DEUTSCHER, J. 2009. Structural Characterizations of Glycerol Kinase: Unraveling Phosphorylation-Induced Long-Range Activation. *Biochemistry,* 48**,** 346-356.

YUZENKOVA, J., DELGADO, M., NECHAEV, S., SAVALIA, D., EPSHTEIN, V., ARTSIMOVITCH, I., MOONEY, R. A., LANDICK, R., FARIAS, R. N., SALOMON, R. & SEVERINOV, K. 2002. Mutations of bacterial RNA polymerase leading to resistance to microcin j25. *Journal of Biological Chemistry,* 277**,** 50867-50875.

ZAMBLE, D. B., MILLER, D. A., HEDDLE, J. G., MAXWELL, A., WALSH, C. T. & HOLLFELDER, F. 2001. In vitro characterization of DNA gyrase inhibition by microcin B17 analogs with altered bisheterocyclic sites. *Proceedings of the National Academy of Sciences of the United States of America,* 98**,** 7712-7717.

ZAMFIR, M., CALLEWAERT, R., CORNEA, P. C., SAVU, L., VATAFU, I. & DE VUYST, L. 1999. Purification and characterization of a bacteriocin produced by *Lactobacillus acidophilus* IBB 801. *Journal of Applied Microbiology,* 87**,** 923-931.

ZERBINO, D. R. & BIRNEY, E. 2008. Velvet: Algorithms for *de novo* short read assembly using de Bruijn graphs. *Genome Research,* 18**,** 821-829.

ZHENG, L., CHEN, Z., ITZEK, A., HERZBERG, M. C. & KRETH, J. 2012. CcpA regulates biofilm formation and competence in *Streptococcus gordonii*. *Molecular Oral Microbiology,* 27**,** 83-94.

ZOUHIR, A., HAMMAMI, R., FLISS, I. & HAMIDA, J. 2010. A new structure-based classification of Gram-positive bacteriocins. *The Protein Journal,* 29**,** 432-439.

Appendices

Appendix 1 Cellular Density and Viable Cell Count Results

Aim

- To determine the relationships between OD_{600nm} with cell density and then cell viability for *Lactobacillus plantarum* strains *L. plantarum* ATCC 8014, *L. plantarum* NC8 and *L. plantarum* subsp. *plantarum* ATCC 14917.

Results

The results of experiments designed to determine the cellular density and cell viability (2.2.5) of *L. plantarum* ATCC 8014, *L. plantarum* NC8 and *L. plantarum* subsp. *plantarum* ATCC 14917 are listed below. These cell density and viable cell counts were used in subsequent experiments that required the a determination of cell density and or cell viability from OD_{600nm}.

Cell density counts of *L. plantarum*

	ATCC 8014		NC8		subsp. *plantarum* ATCC 14917	
	Chamber		Chamber		Chamber	
	1	2	1	2	1	2
Cells count in the 0.025 mm^2 square	106	134	94	105	49	43
	110	116	102	135	50	53
	158	141	129	111	52	64
	145	137	100	101	65	68
	110	130	141	103	59	66
	144	138	125	107	60	88
	132	151	124	117	68	55
Mean (cells mL^{-1})	$5.03x10^9$		$4.96x10^9$		$6.96x10^9$	
Standard Deviation (cells mL^{-1})	$6.28x10^8$		$6.53x10^8$		$1.32x10^9$	

Viable cell count results for *L. plantarum* strains

Lactobacillus plantarum strain	Colony Count (Replicate plate 1)	Colony Count (Replicate plate 2)	Colony Forming Units per mL^{-1} at 1 OD_{600nm}
NC8	94	71	$4.40x10^8$
subsp. *plantarum* ATCC 14917	89	68	$4.19x10^8$
ATCC 8014	66	58	$3.31x10^8$

Appendix 2 Carbon source utilisation assay

Introduction

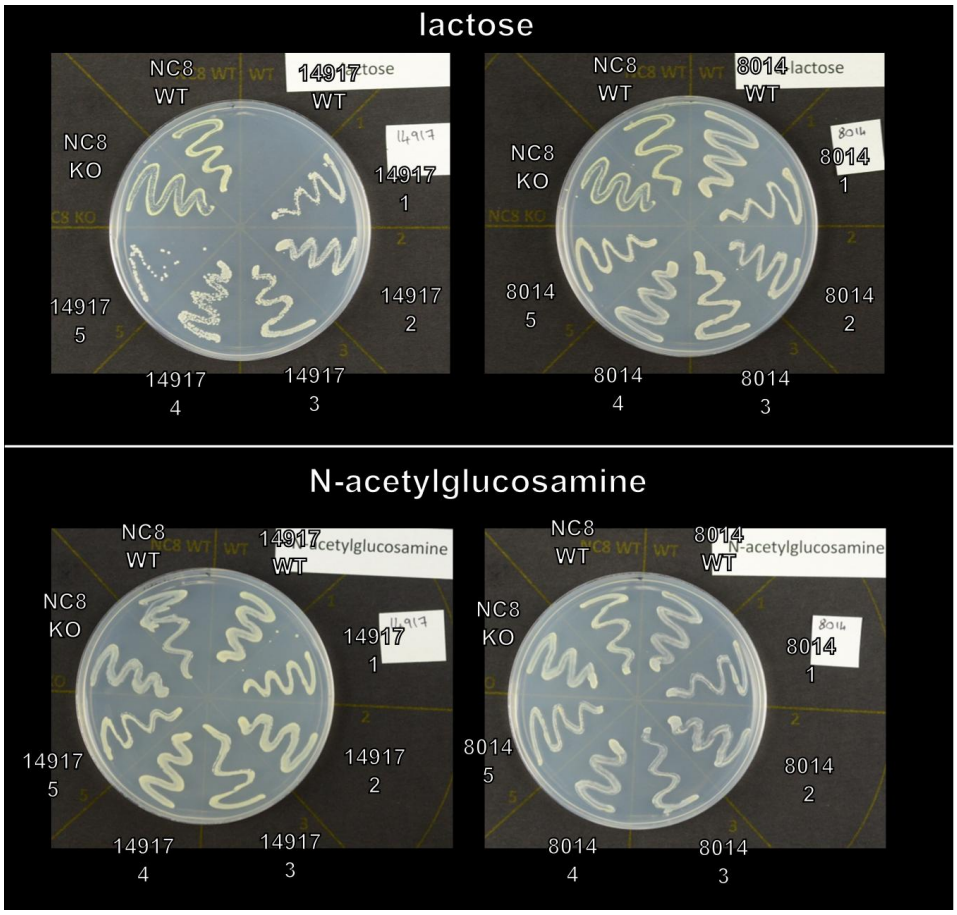
Metabolic utilisation of carbon sources was investigated for all the glycocin F-resistant mutants and *L. plantarum* NC8 Δ*pts18CBA* using a chemically defined media carbon source assay (2.2.8)

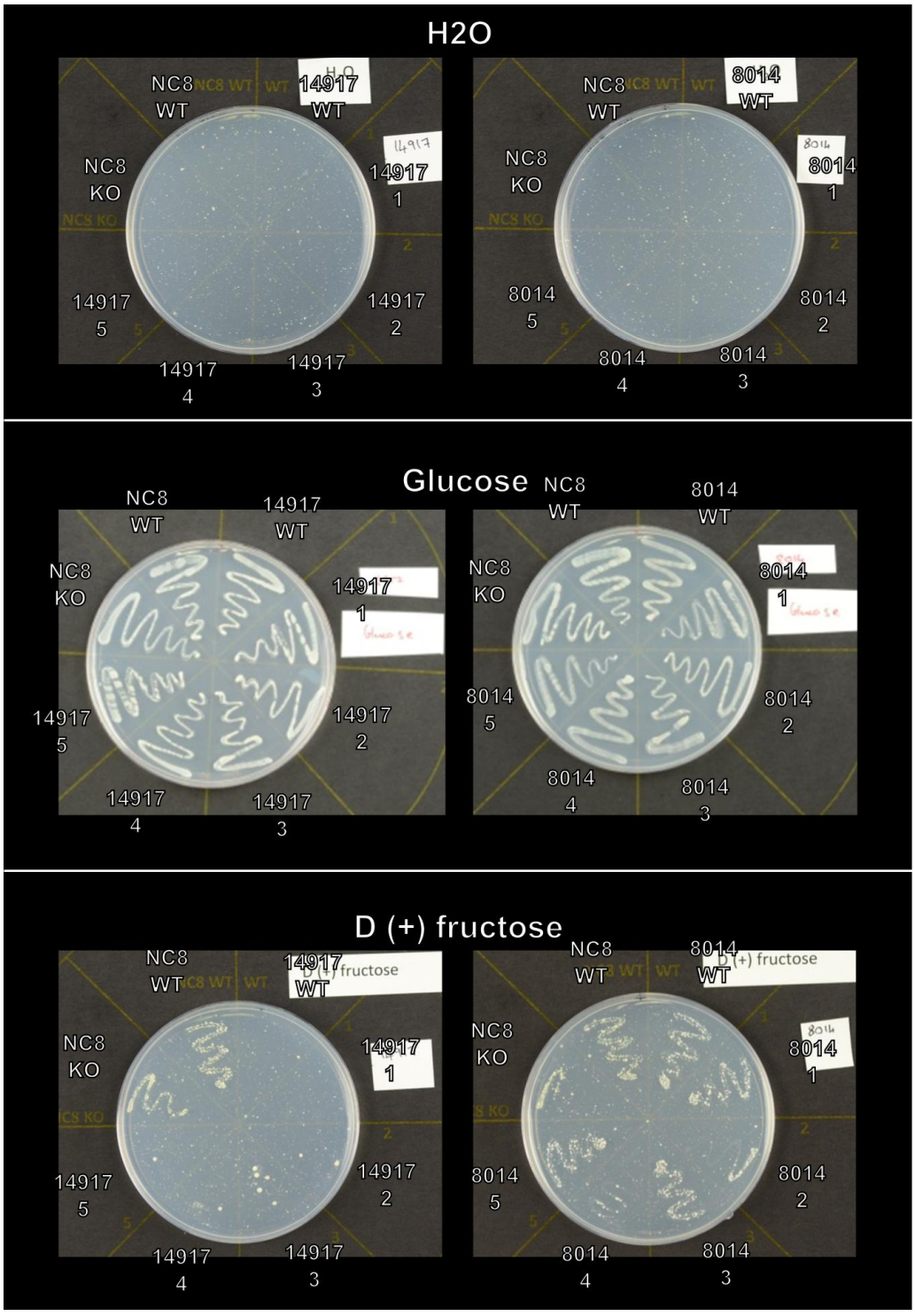
Aim

To identify changes in carbon source utilisation ability in *L. plantarum* NC8 Δ*pts18CBA* and the natural Glycocin F resistant *L. plantarum* subsp. *plantarum* ATCC 14917 and *L. plantarum* ATCC 8014 isolates

Results

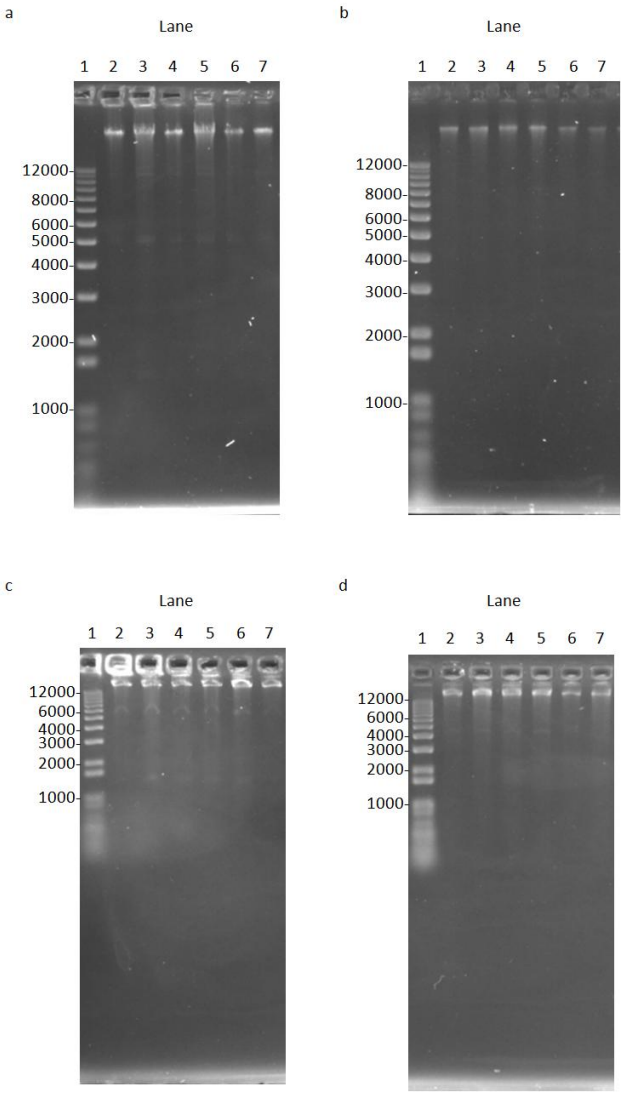
Six segments were streaked with one of the strains (8014 or 14917) and its mutants (1-5), the two other segments were streaked with WT NC8 and the *pts18CBA* knockout strain. Growth is indicated by a streak or small colonies. Light white speckles are clumps of insoluble amino acids. KO, knockout; WT, wildtype.




H2O
NC8 WT
14917 WT
NC8 KO
14917 1
14917 5
14917 2
14917 4
14917 3
NC8 WT
8014 WT
NC8 KO
8014 1
8014 5
8014 2
8014 4
8014 3
Glucose
NC8 WT
14917 WT
NC8 KO
14917 1
14917 5
14917 2
14917 4
14917 3
NC8 WT
8014 WT
NC8 KO
8014 1
8014 5
8014 2
8014 4
8014 3
D (+) fructose
NC8 WT
14917 WT
NC8 KO
14917 1
14917 5
14917 2
14917 4
14917 3
NC8 WT
8014 WT
NC8 KO
8014 1
8014 5
8014 2
8014 4
8014 3

Appendix 3 Agarose gel electrophoresis of genomic DNA purified for genomic sequencing

a) Agarose gel (0.6 %) electrophoresis of gDNA from *L. plantarum* ATCC 8014 isolates 1 (Lane 2), 2 (Lane 3), 3 (Lane 4), 4 (Lane 5), 5 (Lane 6) and wt (Lane 7). Agarose gel (0.6 %) electrophoresis of gDNA from *L. plantarum* subsp. *plantarum* ATCC 14917 isolates 1 (Lane 2), 2 (Lane 3), 3 (Lane 4), 4 (Lane 5), 5 (Lane 6) and wt (Lane 7). c) Agarose gel (1.0 %) electrophoresis of gDNA from *L. plantarum* ATCC 8014 isolates 1 (Lane 2), 2 (Lane 3), 3 (Lane 4), 4 (Lane 5), 5 (Lane 6) and wt (Lane 7). Agarose gel (1.0 %) electrophoresis of gDNA from *L. plantarum* subsp. *plantarum* ATCC 14917 isolates 1 (Lane 2), 2 (Lane 3), 3 (Lane 4), 4 (Lane 5), 5 (Lane 6) and wt (Lane 7).



Appendix 4 Statistical details on the raw sequencing reads and quality control processed reads

Relevant details pertaining to the five different stages of quality control.

L. plantarum ATCC 8014 WT

Processing State	Min Read Length	Max Read Length	Mean Read Length	Sequences	Total BP	Estimated Genome Size	Coverage
Forward Raw Sequencing reads	251	251	251	2785931	699268681	3263224	428.5753482
Reverse Raw Sequencing reads	251	251	251	2785931	699268681		
Forward Post MGS processing	15	251	248.9	2776720	691125608		424.0950814
Reverse Post MGS processing	15	251	249.5	2776720	692791640		
Forward No ambiguities 150 bp	150	251	249.9	2736780	683921322		419.420695
Reverse No ambiguities 150 bp	150	251	250.2	2736780	684742356		
Combined Flashed reads	170	492	338.5	2648545	896532482.5		287.1059944
Forward Flashed Reads	150	251	228	88235	20117580		
Reverse Flash Reads	150	251	229.4	88235	20241109		
Combined Flashed Trimmed 205	205	492	338.5	2648542	896531467		284.8224577
Forward Flashed Trimmed 205	205	251	241.6	68048	16440396.8		
Reverse Flashed Trimmed 205	205	251	242	68048	16467616		

L. plantarum ATCC 8014 WT

Processing State	≥ Q20 (%)	≥ Q30 (%)	A (%)	C (%)	G (%)	T (%)	N (%)	GC (%)
Forward Raw Sequencing reads	95.80	93.90	27.82	22.25	22.12	27.72	0.09	44.37
Reverse Raw Sequencing reads	90.70	86.80	27.79	22.16	22.38	27.62	0.05	44.54
Forward Post MGS processing	96.10	94.30	27.81	22.20	22.12	27.76	0.09	44.33
Reverse Post MGS processing	90.90	87.10	27.74	22.20	22.32	27.68	0.04	44.52
Forward No ambiguities 150 bp	96.20	94.40	27.83	22.23	22.15	27.79	0.00	44.37
Reverse No ambiguities 150 bp	91.10	87.30	27.75	22.21	22.33	27.70	0.00	44.54
Combined Flashed reads	97.90	96.70	27.81	22.22	22.13	27.84	0.00	44.34
Forward Flashed Reads	90.90	87.50	27.52	23.04	22.27	27.18	0.00	45.31
Reverse Flash Reads	70.00	62.40	26.90	22.81	23.30	27.01	0.00	46.11
Combined Flashed Trimmed 205	97.90	96.70	27.81	22.22	22.13	27.84	0.00	44.34
Forward Flashed Trimmed 205	89.60	85.60	27.42	23.13	22.30	27.14	0.00	45.43
Reverse Flashed Trimmed 205	65.80	57.20	26.66	22.94	23.41	26.98	0.00	46.35

L. plantarum ATCC 8014 LP8_1L

Processing State	Min Read Length	Max Read Length	Mean Read Length	Sequences	Total BP	Estimated Genome Size	Coverage
Forward Raw Sequencing reads	251	251	251	2604139	653638889	3270494	399.7187514
Reverse Raw Sequencing reads	251	251	251	2604139	653638889		
Forward Post MGS processing	15	251	248.4	2594916	644577134.4		394.7326337
Reverse Post MGS processing	15	251	249.1	2594916	646393575.6		
Forward No ambiguities 150 bp	150	251	249.7	2554298	637808210.6		390.2721456
Reverse No ambiguities 150 bp	150	251	250	2554298	638574500		
Combined Flashed reads	184	492	350.4	2468524	864970809.6		276.2659777
Forward Flashed Reads	150	251	224	85774	19213376		
Reverse Flash Reads	150	251	225.5	85774	19342037		
Combined Flashed Trimmed 205	205	492	350.4	2468515	864967656		273.4892153
Forward Flashed Trimmed 205	205	251	241.2	61042	14723330.4		
Reverse Flashed Trimmed 205	205	251	241.7	61042	14753851.4		

L. plantarum ATCC 8014 LP8_1L

Processing State	≥ Q20 (%)	≥ Q30 (%)	A (%)	C (%)	G (%)	T (%)	N (%)	GC (%)
Forward Raw Sequencing reads	95.80	93.90	27.82	22.29	22.15	27.65	0.09	44.44
Reverse Raw Sequencing reads	90.50	86.40	27.76	22.19	22.42	27.59	0.05	44.61
Forward Post MGS processing	96.10	94.30	27.78	22.27	22.18	27.70	0.09	44.44
Reverse Post MGS processing	90.70	86.70	27.71	22.24	22.36	27.66	0.04	44.60
Forward No ambiguities 150 bp	96.30	94.50	27.79	22.28	22.19	27.72	0.00	44.47
Reverse No ambiguities 150 bp	90.90	87.00	27.71	22.25	22.36	27.68	0.00	44.62
Combined Flashed reads	97.70	96.40	27.78	22.26	22.17	27.78	0.00	44.43
Forward Flashed Reads	90.40	86.90	27.60	22.97	22.34	27.10	0.00	45.31
Reverse Flash Reads	72.60	65.70	27.09	22.61	23.37	26.91	0.00	45.98
Combined Flashed Trimmed 205	97.70	96.40	27.78	22.26	22.17	27.78	0.00	44.43
Forward Flashed Trimmed 205	88.40	84.20	27.44	23.15	22.42	26.98	0.00	45.57
Reverse Flashed Trimmed 205	67.50	59.10	26.81	22.77	23.60	26.84	0.00	46.38

L. plantarum ATCC 8014 LP8_4H

Processing State	Min Read Length	Max Read Length	Mean Read Length	Sequences	Total BP	Estimated Genome Size	Coverage
Forward Raw Sequencing reads	251	251	251	2412570	605555070	3262343	371.2393639
Reverse Raw Sequencing reads	251	251	251	2412570	605555070		
Forward Post MGS processing	15	251	249.1	2406782	599529396.2		
Reverse Post MGS processing	15	251	249.6	2406782	600732787.2		367.9141597
Forward No ambiguities 150 bp	150	251	250	2373494	593373500		
Reverse No ambiguities 150 bp	150	251	250.3	2373494	594085548.2		363.9896382
Combined Flashed reads	194	492	357.2	2308497	824595128.4		
Forward Flashed Reads	150	251	228.8	64997	14871313.6		
Reverse Flash Reads	150	251	230.4	64997	14975308.8		261.9104585
Combined Flashed Trimmed 205	212	492	357.2	2308495	824594414		
Forward Flashed Trimmed 205	205	251	243.9	49336	12033050.4		
Reverse Flashed Trimmed 205	205	251	244.4	49336	12057718.4		260.1459083

L. plantarum ATCC 8014 LP8_4H

Processing State	≥ Q20 (%)	≥ Q30 (%)	A (%)	C (%)	G (%)	T (%)	N (%)	GC (%)
Forward Raw Sequencing reads	95.90	94.10	27.83	22.25	22.10	27.72	0.09	44.35
Reverse Raw Sequencing reads	90.70	86.70	27.78	22.16	22.37	27.65	0.05	44.53
Forward Post MGS processing	96.20	94.40	27.81	22.23	22.11	27.74	0.09	44.34
Reverse Post MGS processing	90.90	87.00	27.74	22.19	22.33	27.70	0.04	44.52
Forward No ambiguities 150 bp	96.30	94.50	27.85	22.26	22.13	27.78	0.00	44.39
Reverse No ambiguities 150 bp	91.00	87.20	27.75	22.20	22.34	27.72	0.00	44.54
Combined Flashed reads	97.50	96.20	27.82	22.23	22.12	27.83	0.00	44.35
Forward Flashed Reads	88.80	84.70	27.41	23.23	22.29	27.07	0.00	45.52
Reverse Flash Reads	67.00	58.70	26.79	22.75	23.53	26.91	0.00	46.29
Combined Flashed Trimmed 205	97.50	96.20	27.83	22.23	22.12	27.83	0.00	44.35
Forward Flashed Trimmed 205	86.90	82.10	27.26	23.39	22.35	26.98	0.00	45.74
Reverse Flashed Trimmed 205	62.00	52.30	26.50	22.92	23.73	26.85	0.00	46.65

L. plantarum ATCC 8014 LP8_5H

Processing State	Min Read Length	Max Read Length	Mean Read Length	Sequences	Total BP	Estimated Genome Size	Coverage
Forward Raw Sequencing reads	251	251	251	3054791	766752541	3263804	469.8520751
Reverse Raw Sequencing reads	251	251	251	3054791	766752541		
Forward Post MGS processing	15	251	247.9	3044940	754840626		463.1124344
Reverse Post MGS processing	15	251	248.5	3044940	756667590		
Forward No ambiguities 150 bp	150	251	249.3	2991288	745728098.4		457.2436284
Reverse No ambiguities 150 bp	150	251	249.6	2991288	746625484.8		
Combined Flashed reads	180	492	338.6	2866777	970690692.2		314.2574584
Forward Flashed Reads	150	251	220.5	124511	27454675.5		
Reverse Flash Reads	150	251	221.1	124511	27529382.1		
Combined Flashed Trimmed 205	206	492	338.6	2866771	970688660.6		309.7280381
Forward Flashed Trimmed 205	205	251	239.3	83931	20084688.3		
Reverse Flashed Trimmed 205	205	251	239.7	83931	20118260.7		

L. plantarum ATCC 8014 LP8_5H

Processing State	≥ Q20 (%)	≥ Q30 (%)	A (%)	C (%)	G (%)	T (%)	N (%)	GC (%)
Forward Raw Sequencing reads	95.90	94.00	27.88	22.24	22.12	27.67	0.09	44.36
Reverse Raw Sequencing reads	89.40	85.00	27.80	22.15	22.40	27.61	0.05	44.54
Forward Post MGS processing	96.20	94.50	27.82	22.21	22.13	27.73	0.09	44.34
Reverse Post MGS processing	89.70	85.40	27.73	22.21	22.30	27.70	0.04	44.51
Forward No ambiguities 150 bp	96.40	94.60	27.85	22.23	22.16	27.76	0.00	44.39
Reverse No ambiguities 150 bp	89.90	85.60	27.75	22.23	22.31	27.73	0.00	44.54
Combined Flashed reads	97.50	96.20	27.83	22.21	22.14	27.82	0.00	44.35
Forward Flashed Reads	91.60	88.40	27.67	22.88	22.32	27.14	0.00	45.19
Reverse Flash Reads	76.50	70.30	27.12	22.65	23.14	27.11	0.00	45.79
Combined Flashed Trimmed 205	97.50	96.20	27.83	22.21	22.14	27.82	0.00	44.35
Forward Flashed Trimmed 205	89.60	85.70	27.51	23.07	22.36	27.04	0.00	45.44
Reverse Flashed Trimmed 205	70.70	63.00	26.86	22.80	23.33	27.00	0.00	46.14

L. plantarum subsp. *plantarum* ATCC 14917 WT

Processing State	Min Read Length	Max Read Length	Mean Read Length	Sequences	Total BP	Estimated Genome Size	Coverage
Forward Raw Sequencing reads	251	251	251	2601578	652996078	3278114	398.3974
Reverse Raw Sequencing reads	251	251	251	2601578	652996078		
Forward Post MGS processing	15	251	248.7	2594036	645136753.2		393.9981
Reverse Post MGS processing	15	251	249.2	2594036	646433771.2		
Forward No ambiguities 150 bp	150	251	249.8	2556788	638685642.4		389.8225
Reverse No ambiguities 150 bp	150	251	250	2556788	639197000		
Combined Flashed reads	189	492	336.4	2466846	829846994.4		265.6041
Forward Flashed Reads	150	251	226.3	89942	20353874.6		
Reverse Flash Reads	150	251	227.7	89942	20479793.4		
Combined Flashed Trimmed 205	213	492	336.4	2466843	829845985.2		263.0063
Forward Flashed Trimmed 205	205	251	241.8	66774	16145953.2		
Reverse Flashed Trimmed 205	205	251	242.2	66774	16172662.8		

L. plantarum subsp. *plantarum* ATCC 14917 WT

Processing State	≥ Q20 (%)	≥ Q30 (%)	A (%)	C (%)	G (%)	T (%)	N (%)	GC (%)
Forward Raw Sequencing reads	95.30	93.20	27.63	22.38	22.34	27.56	0.09	44.72
Reverse Raw Sequencing reads	89.70	85.30	27.65	22.35	22.52	27.44	0.04	44.87
Forward Post MGS processing	95.60	93.50	27.61	22.36	22.35	27.60	0.09	44.71
Reverse Post MGS processing	89.90	85.70	27.61	22.40	22.46	27.50	0.04	44.85
Forward No ambiguities 150 bp	95.70	93.70	27.63	22.38	22.36	27.62	0.00	44.74
Reverse No ambiguities 150 bp	90.10	85.90	27.62	22.41	22.47	27.52	0.00	44.87
Combined Flashed reads	97.60	96.20	27.63	22.35	22.33	27.69	0.00	44.68
Forward Flashed Reads	89.80	86.00	27.32	23.18	22.63	26.89	0.00	45.81
Reverse Flash Reads	69.90	62.30	26.95	22.83	23.57	26.67	0.00	46.40
Combined Flashed Trimmed 205	97.60	96.20	27.63	22.35	22.33	27.69	0.00	44.68
Forward Flashed Trimmed 205	88.00	83.50	27.20	23.32	22.68	26.79	0.00	46.00
Reverse Flashed Trimmed 205	64.80	55.80	26.75	22.93	23.74	26.58	0.00	46.67

L. plantarum subsp. *plantarum* ATCC 14917 LP1_2H

Processing State	Min Read Length	Max Read Length	Mean Read Length	Sequences	Total BP	Estimated Genome Size	Coverage
Forward Raw Sequencing reads	251	251	251	2639160	662429160	3284742	403.3371
Reverse Raw Sequencing reads	251	251	251	2639160	662429160		
Forward Post MGS processing	15	251	249	2633252	655679748		399.7085
Reverse Post MGS processing	15	251	249.6	2633252	657259699.2		
Forward No ambiguities 150 bp	150	251	250	2597449	649362250		395.5391
Reverse No ambiguities 150 bp	150	251	250.2	2597449	649881739.8		
Combined Flashed reads	168	492	319.4	2500708	798726135.2		256.887
Forward Flashed Reads	150	251	232.3	96741	22472934.3		
Reverse Flash Reads	150	251	233.7	96741	22608371.7		
Combined Flashed Trimmed 205	205	492	319.4	2500702	798724218.8		254.801
Forward Flashed Trimmed 205	205	251	244.3	78183	19100106.9		
Reverse Flashed Trimmed 205	205	251	244.7	78183	19131380.1		

L. plantarum subsp. *plantarum* ATCC 14917 LP1_2H

Processing State	≥ Q20 (%)	≥ Q30 (%)	A (%)	C (%)	G (%)	T (%)	N (%)	GC (%)
Forward Raw Sequencing reads	95.90	94.00	27.66	22.35	22.27	27.63	0.09	44.62
Reverse Raw Sequencing reads	90.80	87.00	27.70	22.29	22.46	27.50	0.05	44.75
Forward Post MGS processing	96.10	94.30	27.65	22.32	22.28	27.67	0.09	44.60
Reverse Post MGS processing	91.00	87.30	27.66	22.33	22.41	27.55	0.04	44.73
Forward No ambiguities 150 bp	96.20	94.40	27.67	22.33	22.30	27.69	0.00	44.63
Reverse No ambiguities 150 bp	91.20	87.50	27.68	22.34	22.42	27.58	0.00	44.77
Combined Flashed reads	98.30	97.10	27.66	22.32	22.28	27.74	0.00	44.59
Forward Flashed Reads	92.00	88.90	27.43	23.02	22.45	27.11	0.00	45.47
Reverse Flash Reads	58.30	49.10	27.15	22.55	23.13	27.17	0.00	45.69
Combined Flashed Trimmed 205	98.30	97.10	27.66	22.32	22.28	27.74	0.00	44.59
Forward Flashed Trimmed 205	91.20	87.80	27.36	23.08	22.47	27.09	0.00	45.55
Reverse Flashed Trimmed 205	52.90	42.50	27.05	22.60	23.16	27.21	0.00	45.76

L. plantarum subsp. *plantarum* ATCC 14917 LP1_3L

Processing State	Min Read Length	Max Read Length	Mean Read Length	Sequences	Total BP	Estimated Genome Size	Coverage
Forward Raw Sequencing reads	251	251	251	2178224	546734224	3244121	337.0615
Reverse Raw Sequencing reads	251	251	251	2178224	546734224		
Forward Post MGS processing	15	251	248.7	2172824	540381328.8		333.5469
Reverse Post MGS processing	15	251	249.3	2172824	541685023.2		
Forward No ambiguities 150 bp	150	251	249.8	2140803	534772589.4		329.8852
Reverse No ambiguities 150 bp	150	251	250.1	2140803	535414830.3		
Combined Flashed reads	163	492	359.1	2069463	743144163.3		239.0952
Forward Flashed Reads	150	251	227.3	71340	16215582		
Reverse Flash Reads	150	251	228.4	71340	16294056		
Combined Flashed Trimmed 205	207	492	359.1	2069457	743142008.7		237.03
Forward Flashed Trimmed 205	205	251	243.4	52980	12895332		
Reverse Flashed Trimmed 205	205	251	243.8	52980	12916524		

L. plantarum subsp. *plantarum* ATCC 14917 LP1_3L

Processing State	≥ Q20 (%)	≥ Q30 (%)	A (%)	C (%)	G (%)	T (%)	N (%)	GC (%)
Forward Raw Sequencing reads	95.80	93.80	27.77	22.36	22.17	27.61	0.09	44.53
Reverse Raw Sequencing reads	89.80	85.60	27.72	22.19	22.52	27.53	0.05	44.71
Forward Post MGS processing	96.00	94.20	27.74	22.36	22.17	27.65	0.09	44.53
Reverse Post MGS processing	90.00	85.80	27.68	22.24	22.46	27.58	0.04	44.70
Forward No ambiguities 150 bp	96.20	94.30	27.76	22.37	22.19	27.68	0.00	44.56
Reverse No ambiguities 150 bp	90.20	86.00	27.69	22.25	22.47	27.61	0.00	44.72
Combined Flashed reads	97.30	95.90	27.74	22.35	22.16	27.76	0.00	44.51
Forward Flashed Reads	88.60	84.50	27.47	23.22	22.64	26.68	0.00	45.87
Reverse Flash Reads	68.00	60.00	26.85	22.75	23.90	26.53	0.00	46.64
Combined Flashed Trimmed 205	97.30	95.90	27.74	22.35	22.16	27.76	0.00	44.51
Forward Flashed Trimmed 205	86.50	81.70	27.36	23.36	22.72	26.55	0.00	46.07
Reverse Flashed Trimmed 205	62.40	52.80	26.61	22.87	24.13	26.40	0.00	46.99

L. plantarum subsp. *plantarum* ATCC 14917 LP1_4H

Processing State	Min Read Length	Max Read Length	Mean Read Length	Sequences	Total BP	Estimated Genome Size	Coverage
Forward Raw Sequencing reads	251	251	251	2495665	626411915	3262640	383.9908
Reverse Raw Sequencing reads	251	251	251	2495665	626411915		
Forward Post MGS processing	15	251	249.3	2488981	620502963.3		380.8264
Reverse Post MGS processing	15	251	249.9	2488981	621996351.9		
Forward No ambiguities 150 bp	150	251	250.2	2456828	614698365.6		377.0363
Reverse No ambiguities 150 bp	150	251	250.5	2456828	615435414		
Combined Flashed reads	195	492	357.1	2386127	852085951.7		271.3754
Forward Flashed Reads	150	251	234.6	70701	16586454.6		
Reverse Flash Reads	150	251	236.6	70701	16727856.6		
Combined Flashed Trimmed 205	210	492	357.1	2386126	852085594.6		269.9769
Forward Flashed Trimmed 205	205	251	245.9	58403	14361297.7		
Reverse Flashed Trimmed 205	205	251	246.4	58403	14390499.2		

L. plantarum subsp. *plantarum* ATCC 14917 LP1_4H

Processing State	≥ Q20 (%)	≥ Q30 (%)	A (%)	C (%)	G (%)	T (%)	N (%)	GC (%)
Forward Raw Sequencing reads	95.80	93.80	27.69	22.40	22.23	27.60	0.09	44.63
Reverse Raw Sequencing reads	90.20	86.00	27.64	22.28	22.51	27.52	0.05	44.79
Forward Post MGS processing	96.00	94.10	27.67	22.38	22.25	27.62	0.09	44.63
Reverse Post MGS processing	90.30	86.30	27.62	22.31	22.48	27.56	0.04	44.79
Forward No ambiguities 150 bp	96.10	94.20	27.69	22.40	22.26	27.64	0.00	44.66
Reverse No ambiguities 150 bp	90.50	86.50	27.62	22.31	22.49	27.57	0.00	44.80
Combined Flashed reads	97.50	96.20	27.68	22.37	22.24	27.71	0.00	44.61
Forward Flashed Reads	88.10	83.70	27.19	23.46	22.58	26.75	0.00	46.04
Reverse Flash Reads	58.90	49.00	26.58	22.84	23.88	26.69	0.00	46.72
Combined Flashed Trimmed 205	97.50	96.20	27.68	22.37	22.24	27.71	0.00	44.61
Forward Flashed Trimmed 205	86.90	82.00	27.10	23.57	22.64	26.68	0.00	46.21
Reverse Flashed Trimmed 205	54.90	43.90	26.38	22.93	24.03	26.68	0.00	46.96

Appendix 5 *De novo* Genome Assembly and CISA integration Results

De-novo Genome Assembly and CISA Integration Results

	L. plantarum ATCC 8014 LP8_1L				*L. plantarum* ATCC 8014 LP8_4H			
	SPAdes	Edena	Velvet	CISA	SPAdes	Edena	Velvet	CISA
Nodes	197	197	76	23	237	443	89	21
Total Nucleotides (bp)	3319654	3227994	3254122	3253729	3333537	3114710	3253155	3242548
Min Length (bp)	128	541	401	1041	295	564	68	2389
Max Length (bp)	528541	125873	358980	528679	528541	109178	429109	528679
Average Length (bp)	16851.04	16385.76	42817.39	141466.48	14065.56	7030.95	36552.3	154407.05
Median Length (bp)	404	10181	1155.5	45311	408	3914	4589	46355
N25 length (bp)	465662	53629	307340	465800	465663	28655	245296	465801
N50 length (bp)	418934	29983	196709	419125	402672	12227	91684	402740
N75 length (bp)	247482	18197	80885	253249	247221	6399	75905	247363
N90 length (bp)	80752	9979	36376	80896	45778	3213	30713	80885
N95 length (bp)	29989	6924	27743	45311	28440	2160	24184	45311
As	27.82%	27.73%	27.77%	27.79%	27.80%	27.79%	27.71%	27.85%
Ts	27.66%	27.80%	27.72%	27.73%	27.61%	27.81%	27.79%	27.66%
Gs	22.33%	21.92%	22.50%	22.68%	22.41%	22.24%	22.51%	22.03%
Cs	22.19%	22.56%	22.01%	21.79%	22.18%	22.16%	22.00%	22.47%
(A + T)s	55.48%	55.52%	55.49%	55.53%	55.41%	55.60%	55.50%	55.51%
(G + C)s	44.52%	44.48%	44.51%	44.47%	44.59%	44.40%	44.50%	44.49%
Ns	0.00%	0.00%	0.00%	0.00%	0.00%	0.00%	0.00%	0.00%

De-novo Genome Assembly and CISA Integration Results

	L. plantarum ATCC 8014 LP8_5H				*L. plantarum* ATCC 8014 WT			
	SPAdes	Edena	Velvet	CISA	SPAdes	Edena	Velvet	CISA
Nodes	173	154	76	24	132	78	71	20
Total Nucleotides (bp)	3306701	3241933	3254022	3243663	3289999	3260438	3254764	3245610
Min Length (bp)	238	522	401	234	295	478	401	2659
Max Length (bp)	528541	161131	358980	528679	528541	359168	528679	528838
Average Length (bp)	19113.88	21051.51	42816.08	135152.62	24924.23	41800.49	45841.75	162280.5
Median Length (bp)	406	9254.5	782.5	40178.5	420	9076.5	647	64769.5
N25 length (bp)	465663	90391	267904	465801	465662	309917	358980	465904
N50 length (bp)	418984	47877	150664	419125	418999	179480	229667	419187
N75 length (bp)	247482	25792	92992	253395	247221	88229	104013	253405
N90 length (bp)	80752	12804	36376	93017	80749	27126	38958	93168
N95 length (bp)	29989	6901	24378	44937	35204	19599	28573	45400
As	27.80%	27.78%	27.76%	27.84%	27.82%	27.75%	27.76%	27.74%
Ts	27.65%	27.74%	27.74%	27.66%	27.67%	27.74%	27.73%	27.77%
Gs	22.36%	22.56%	22.94%	22.26%	22.70%	22.41%	23.21%	23.39%
Cs	22.19%	21.92%	21.56%	22.23%	21.81%	22.09%	21.29%	21.10%
(A + T)s	55.44%	55.51%	55.49%	55.50%	55.49%	55.50%	55.50%	55.51%
(G + C)s	44.56%	44.49%	44.51%	44.50%	44.51%	44.50%	44.50%	44.49%
Ns	0.00%	0.00%	0.00%	0.00%	0.00%	0.00%	0.00%	0.00%

De-novo Genome Assembly and CISA Integration Results

	L. plantarum subsp. *plantarum* ATCC 14917 LP1_2H				*L. plantarum* subsp. *plantarum* ATCC 14917 LP1_3L			
	SPAdes	Edena	Velvet	CISA	SPAdes	Edena	Velvet	CISA
Nodes	229	65	50	24	307	55	70	24
Total Nucleotides (bp)	3349652	3234828	3232022	3245832	3365897	3237070	3238718	3243858
Min Length (bp)	156	461	401	1207	229	474	191	1063
Max Length (bp)	691448	415083	444383	691721	691597	444738	545064	691733
Average Length (bp)	14627.3	49766.58	64640.44	135243	10963.83	58855.82	46267.4	135160.75
Median Length (bp)	498	5123	1827.5	4036.5	412	2935	1350.5	6230
N25 length (bp)	633662	266089	415100	633662	633327	408626	384376	633407
N50 length (bp)	482026	195651	244237	482068	481968	233563	226553	482022
N75 length (bp)	433801	100846	165275	433983	418824	126779	96644	418972
N90 length (bp)	125314	39838	68085	240351	73950	56565	59729	125895
N95 length (bp)	22222	34292	39984	74239	12847	38829	33982	74266
As	27.80%	27.83%	27.83%	27.81%	27.71%	27.78%	27.71%	27.84%
Ts	27.84%	27.70%	27.69%	27.74%	27.72%	27.74%	27.81%	27.70%
Gs	21.51%	21.88%	22.65%	20.72%	22.56%	22.39%	21.87%	21.60%
Cs	22.85%	22.59%	21.83%	23.73%	22.02%	22.08%	22.60%	22.86%
(A + T)s	55.64%	55.52%	55.52%	55.54%	55.43%	55.52%	55.53%	55.54%
(G + C)s	44.36%	44.48%	44.48%	44.46%	44.57%	44.48%	44.47%	44.46%
Ns	0.00%	0.00%	0.00%	0.00%	0.00%	0.00%	0.00%	0.00%

De-novo Genome Assembly and CISA Integration Results

	L. plantarum subsp. *plantarum* ATCC 14917 LP1_4H				*L. plantarum* subsp. *plantarum* ATCC 14917 WT			
	SPAdes	Edena	Velvet	CISA	SPAdes	Edena	Velvet	CISA
Nodes	320	136	88	40	190	56	51	19
Total Nucleotides (bp)	3396531	3228331	3244568	3267966	3312006	3234041	3234848	3237548
Min Length (bp)	222	393	191	1054	229	461	193	1184
Max Length (bp)	631600	164606	547839	691680	633401	481782	507253	691821
Average Length (bp)	10614.16	23737.73	36870.09	81699.15	17431.61	57750.73	63428.39	170397.26
Median Length (bp)	415	13091	445	1842.5	432.5	14062	598	22551
N25 length (bp)	482040	80816	306063	633404	482225	419669	458113	633401
N50 length (bp)	444250	54159	191110	482562	444248	191858	413585	482582
N75 length (bp)	245167	30498	91652	435176	244091	146164	226553	418957
N90 length (bp)	73950	17627	40734	125460	125314	39853	64912	125460
N95 length (bp)	2792	9411	27881	74114	59583	25006	39980	74234
As	27.76%	27.82%	27.83%	27.80%	27.77%	27.79%	27.77%	27.70%
Ts	27.73%	27.72%	27.69%	27.78%	27.83%	27.73%	27.76%	27.83%
Gs	21.95%	21.71%	23.01%	22.52%	21.54%	21.90%	23.40%	21.71%
Cs	22.56%	22.75%	21.47%	21.90%	22.86%	22.58%	21.07%	22.76%
(A + T)s	55.49%	55.54%	55.52%	55.57%	55.60%	55.52%	55.52%	55.53%
(G + C)s	44.51%	44.46%	44.48%	44.43%	44.40%	44.48%	44.48%	44.47%
Ns	0.00%	0.00%	0.00%	0.00%	0.00%	0.00%	0.00%	0.00%

Appendix 6 *De novo* velvet assembled contigs annotated by Prokka

Contigs provided as electronic material on CD.

Appendix 7 Sequence alignment of *pts18CBA* region of *L. plantarum* ATCC 8014

CLUSTAL OMEGA (http://www.ebi.ac.uk/Tools/msa/clustalo/) multiple sequence alignment performed with default settings from the region of the genomes containing the *pts18BCA* CDS from *L. plantarum* ATCC 8014. The start and stop codons or *pts18CBA* are coloured red and green respectively.

```
EV50_contig000010       ATTATTTAGGTCGCGCATCGCGGTTAATCATCGTAAACACATAGTTGTTCGACATCACGT    60
EV52_contig000007       ATTATTTAGGTCGCGCATCGCGGTTAATCATCGTAAACACATAGTTGTTCGACATCACGT    60
EV49_contig000024       ATTATTTAGGTCGCGCATCGCGGTTAATCATCGTAAACACATAGTTGTTCGACATCACGT    60
EV51_contig000029       ATTATTTAGGTCGCGCATCGCGGTTAATCATCGTAAACACATAGTTGTTCGACATCACGT    60
                        ************************************************************

EV50_contig000010       GACAGCACATAGTAGATCGTAAGCACGGGCCCGTTTATTGGAAAAGCACTGAAGCGAGCG    120
EV52_contig000007       GACAGCACATAGTAGATCGTAAGCACGGGCCCGTTTATTGGAAAAGCACTGAAGCGAGCG    120
EV49_contig000024       GACAGCACATAGTAGATCGTAAGCACGGGCCCGTTTATTGGAAAAGCACTGAAGCGAGCG    120
EV51_contig000029       GACAGCACATAGTAGATCGTAAGCACGGGCCCGTTTATTGGAAAAGCACTGAAGCGAGCG    120
                        ************************************************************

EV50_contig000010       GGGAACATATACACATAGTGGGCAAGACGAAAGGATGGTTTTACTTATGAAATATGGCAT    180
EV52_contig000007       GGGAACATATACACATAGTGGGCAAGACGAAAGGATGGTTTTACTTATGAAATATGGCAT    180
EV49_contig000024       GGGAACATATACACATAGTGGGCAAGACGAAAGGATGGTTTTACTTATGAAATATGGCAT    180
EV51_contig000029       GGGAACATATACACATAGTGGGCAAGACGAAAGGATGGTTTTACTTATGAAATATGGCAT    180
                        ************************************************************

EV50_contig000010       TATTGGCGCAGGCGCAATGGGATATCGGTATGGTGTAATGTTACAAGAAAACGCCGGTGT    240
EV52_contig000007       TATTGGCGCAGGCGCAATGGGATATCGGTATGGTGTAATGTTACAAGAAAACGCCGGTGT    240
EV49_contig000024       TATTGGCGCAGGCGCAATGGGATATCGGTATGGTGTAATGTTACAAGAAAACGCCGGTGT    240
EV51_contig000029       TATTGGCGCAGGCGCAATGGGATATCGGTATGGTGTAATGTTACAAGAAAACGCCGGTGT    240
                        ************************************************************

EV50_contig000010       TGACGTTGATTTTATCGACACGTGGGAACCAAACGTCGCAAAGGTCCGTGAGCAGGGCAG    300
EV52_contig000007       TGACGTTGATTTTATCGACACGTGGGAACCAAACGTCGCAAAGGTCCGTGAGCAGGGCAG    300
EV49_contig000024       TGACGTTGATTTTATCGACACGTGGGAACCAAACGTCGCAAAGGTCCGTGAGCAGGGCAG    300
EV51_contig000029       TGACGTTGATTTTATCGACACGTGGGAACCAAACGTCGCAAAGGTCCGTGAGCAGGGCAG    300
                        ************************************************************

EV50_contig000010       CGTTGATGTGGCGCGTGATCATCAAAATCACCATGTCGTTCCAATCAATATTTATTATCC    360
EV52_contig000007       CGTTGATGTGGCGCGTGATCATCAAAATCACCATGTCGTTCCAATCAATATTTATTATCC    360
EV49_contig000024       CGTTGATGTGGCGCGTGATCATCAAAATCACCATGTCGTTCCAATCAATATTTATTATCC    360
EV51_contig000029       CGTTGATGTGGCGCGTGATCATCAAAATCACCATGTCGTTCCAATCAATATTTATTATCC    360
                        ************************************************************

EV50_contig000010       TGAAGAATATCAGGGTCATCCAGATGTTTGGATTGTTTTTAAGAAGCAAATGCAGTTAGC    420
EV52_contig000007       TGAAGAATATCAGGGTCATCCAGATGTTTGGATTGTTTTTAAGAAGCAAATGCAGTTAGC    420
EV49_contig000024       TGAAGAATATCAGGGTCATCCAGATGTTTGGATTGTTTTTAAGAAGCAAATGCAGTTAGC    420
EV51_contig000029       TGAAGAATATCAGGGTCATCCAGATGTTTGGATTGTTTTTAAGAAGCAAATGCAGTTAGC    420
                        ************************************************************

EV50_contig000010       CGATGAGCTCAAGCGCGATGCGTCGTTATTTCATGAGGACCAGTACGTTTTCGCCGCAAT    480
EV52_contig000007       CGATGAGCTCAAGCGCGATGCGTCGTTATTTCATGAGGACCAGTACGTTTTCGCCGCAAT    480
EV49_contig000024       CGATGAGCTCAAGCGCGATGCGTCGTTATTTCATGAGGACCAGTACGTTTTCGCCGCAAT    480
EV51_contig000029       CGATGAGCTCAAGCGCGATGCGTCGTTATTTCATGAGGACCAGTACGTTTTCGCCGCAAT    480
                        ************************************************************

EV50_contig000010       GAACGGGATGGGGCACTTTGAAAAGATTGCACAATATTTCCCAGAAAATCACATTATTGG    540
EV52_contig000007       GAACGGGATGGGGCACTTTGAAAAGATTGCACAATATTTCCCAGAAAATCACATTATTGG    540
EV49_contig000024       GAACGGGATGGGGCACTTTGAAAAGATTGCACAATATTTCCCAGAAAATCACATTATTGG    540
EV51_contig000029       GAACGGGATGGGGCACTTTGAAAAGATTGCACAATATTTCCCAGAAAATCACATTATTGG    540
                        ************************************************************

EV50_contig000010       TGGTACGGCGATGATTGCAACGGTCTTGAATGGTCCGGGTACGGTTGATTTTATGGGGCC    600
EV52_contig000007       TGGTACGGCGATGATTGCAACGGTCTTGAATGGTCCGGGTACGGTTGATTTTATGGGGCC    600
EV49_contig000024       TGGTACGGCGATGATTGCAACGGTCTTGAATGGTCCGGGTACGGTTGATTTTATGGGGCC    600
EV51_contig000029       TGGTACGGCGATGATTGCAACGGTCTTGAATGGTCCGGGTACGGTTGATTTTATGGGGCC    600
                        ************************************************************

EV50_contig000010       TAAGGGTAGTGAGGCGATGCATATGAGTAAGTATGCAGGTCATATTGATACGACTACCAA    660
EV52_contig000007       TAAGGGTAGTGAGGCGATGCATATGAGTAAGTATGCAGGTCATATTGATACGACTACCAA    660
EV49_contig000024       TAAGGGTAGTGAGGCGATGCATATGAGTAAGTATGCAGGTCATATTGATACGACTACCAA    660
EV51_contig000029       TAAGGGTAGTGAGGCGATGCATATGAGTAAGTATGCAGGTCATATTGATACGACTACCAA    660
                        ************************************************************
```

```
EV50_contig000010      GCAAGTGATGGCCGATTTTAAGGCGGCTGATTTAAATCCGATTTGGTCTGATAATTTTAT    720
EV52_contig000007      GCAAGTGATGGCCGATTTTAAGGCGGCTGATTTAAATCCGATTTGGTCTGATAATTTTAT    720
EV49_contig000024      GCAAGTGATGGCCGATTTTAAGGCGGCTGATTTAAATCCGATTTGGTCTGATAATTTTAT    720
EV51_contig000029      GCAAGTGATGGCCGATTTTAAGGCGGCTGATTTAAATCCGATTTGGTCTGATAATTTTAT    720
                       ************************************************************

EV50_contig000010      GGGCATGTGCATGTCCAAAGTTGTGTTTAACGCGGTCACCAATAGTTTGTGCACGATGTT    780
EV52_contig000007      GGGCATGTGCATGTCCAAAGTTGTGTTTAACGCGGTCACCAATAGTTTGTGCACGATGTT    780
EV49_contig000024      GGGCATGTGCATGTCCAAAGTTGTGTTTAACGCGGTCACCAATAGTTTGTGCACGATGTT    780
EV51_contig000029      GGGCATGTGCATGTCCAAAGTTGTGTTTAACGCGGTCACCAATAGTTTGTGCACGATGTT    780
                       ************************************************************

EV50_contig000010      CGAGATTCAGATGGGCCAATTCATCGAATATCCGGGCGTCAAGGATATGGCGACTCAAAT    840
EV52_contig000007      CGAGATTCAGATGGGCCAATTCATCGAATATCCGGGCGTCAAGGATATGGCGACTCAAAT    840
EV49_contig000024      CGAGATTCAGATGGGCCAATTCATCGAATATCCGGGCGTCAAGGATATGGCGACTCAAAT    840
EV51_contig000029      CGAGATTCAGATGGGCCAATTCATCGAATATCCGGGCGTCAAGGATATGGCGACTCAAAT    840
                       ************************************************************

EV50_contig000010      GTTCAACGAAGCCTATGATGCTTGCGAAAGGGCAGGCATCAAACTCATTGAGACGCGGCA    900
EV52_contig000007      GTTCAACGAAGCCTATGATGCTTGCGAAAGGGCAGGCATCAAACTCATTGAGACGCGGCA    900
EV49_contig000024      GTTCAACGAAGCCTATGATGCTTGCGAAAGGGCAGGCATCAAACTCATTGAGACGCGGCA    900
EV51_contig000029      GTTCAACGAAGCCTATGATGCTTGCGAAAGGGCAGGCATCAAACTCATTGAGACGCGGCA    900
                       ************************************************************

EV50_contig000010      AGAAGAAATCGACTCGGTTGAAACGGTCAGTCGAGCCTATAAGTACCATTATCCATCAAT    960
EV52_contig000007      AGAAGAAATCGACTCGGTTGAAACGGTCAGTCGAGCCTATAAGTACCATTATCCATCAAT    960
EV49_contig000024      AGAAGAAATCGACTCGGTTGAAACGGTCAGTCGAGCCTATAAGTACCATTATCCATCAAT    960
EV51_contig000029      AGAAGAAATCGACTCGGTTGAAACGGTCAGTCGAGCCTATAAGTACCATTATCCATCAAT    960
                       ************************************************************

EV50_contig000010      GTATCAAGACTTCTCAAAGGGACGGCCAACCGAAGTTGATTATATTAATGGCTATATTGC   1020
EV52_contig000007      GTATCAAGACTTCTCAAAGGGACGGCCAACCGAAGTTGATTATATTAATGGCTATATTGC   1020
EV49_contig000024      GTATCAAGACTTCTCAAAGGGACGGCCAACCGAAGTTGATTATATTAATGGCTATATTGC   1020
EV51_contig000029      GTATCAAGACTTCTCAAAGGGACGGCCAACCGAAGTTGATTATATTAATGGCTATATTGC   1020
                       ************************************************************

EV50_contig000010      TAAGATTGGGCGGGAACATGACTACGTTTGTCGCGTGCATGAGTTTGTGACTCACGAAGT   1080
EV52_contig000007      TAAGATTGGGCGGGAACATGACTACGTTTGTCGCGTGCATGAGTTTGTGACTCACGAAGT   1080
EV49_contig000024      TAAGATTGGGCGGGAACATGACTACGTTTGTCGCGTGCATGAGTTTGTGACTCACGAAGT   1080
EV51_contig000029      TAAGATTGGGCGGGAACATGACTACGTTTGTCGCGTGCATGAGTTTGTGACTCACGAAGT   1080
                       ************************************************************

EV50_contig000010      CCATTTAGCAGAAATGATGCGGCAATACCGGCACCCGGAAATCCCCGTTGCTGAAAAATA   1140
EV52_contig000007      CCATTTAGCAGAAATGATGCGGCAATACCGGCACCCGGAAATCCCCGTTGCTGAAAAATA   1140
EV49_contig000024      CCATTTAGCAGAAATGATGCGGCAATACCGGCACCCGGAAATCCCCGTTGCTGAAAAATA   1140
EV51_contig000029      CCATTTAGCAGAAATGATGCGGCAATACCGGCACCCGGAAATCCCCGTTGCTGAAAAATA   1140
                       ************************************************************

EV50_contig000010      AGATTATTAATGAAAAAGTTCTGAGTATTTTCAGAACTTTTTTTGTGGGCGGAAAAACTT   1200
EV52_contig000007      AGATTATTAATGAAAAAGTTCTGAGTATTTTCAGAACTTTTTTTGTGGGCGGAAAAACTT   1200
EV49_contig000024      AGATTATTAATGAAAAAGTTCTGAGTATTTTCAGAACTTTTTTTGTGGGCGGAAAAACTT   1200
EV51_contig000029      AGATTATTAATGAAAAAGTTCTGAGTATTTTCAGAACTTTTTTTGTGGGCGGAAAAACTT   1200
                       ************************************************************

EV50_contig000010      AATTTAGTAAGGTTGAAATCGGTTTCTTTAACATTGGTCTATACACTTAGATGACCTTGT   1260
EV52_contig000007      AATTTAGTAAGGTTGAAATCGGTTTCTTTAACATTGGTCTATACACTTAGATGACCTTGT   1260
EV49_contig000024      AATTTAGTAAGGTTGAAATCGGTTTCTTTAACATTGGTCTATACACTTAGATGACCTTGT   1260
EV51_contig000029      AATTTAGTAAGGTTGAAATCGGTTTCTTTAACATTGGTCTATACACTTAGATGACCTTGT   1260
                       ************************************************************

EV50_contig000010      AAACTTGCAATGAAAACATGTTGTGATAGGATTTATAACGTGCTTAAAAAGAGAATGAGG   1320
EV52_contig000007      AAACTTGCAATGAAAACATGTTGTGATAGGATTTATAACGTGCTTAAAAAGAGAATGAGG   1320
EV49_contig000024      AAACTTGCAATGAAAACATGTTGTGATAGGATTTATAACGTGCTTAAAAAGAGAATGAGG   1320
EV51_contig000029      AAACTTGCAATGAAAACATGTTGTGATAGGATTTATAACGTGCTTAAAAAGAGAATGAGG   1320
                       ************************************************************

EV50_contig000010      AGAAATTGGTATGAAGACATATTTTCAGAAAATCGGTCAGTCATTGATGCTACCGATTGC   1380
EV52_contig000007      AGAAATTGGTATGAAGACATATTTTCAGAAAATCGGTCAGTCATTGATGCTACCGATTGC   1380
EV49_contig000024      AGAAATTGGTATGAAGACATATTTTCAGAAAATCGGTCAGTCATTGATGCTACCGATTGC   1380
EV51_contig000029      AGAAATTGGTATGAAGACATATTTTCAGAAAATCGGTCAGTCATTGATGCTACCGATTGC   1380
                       ************************************************************

EV50_contig000010      AACGTTGCCCGCAGCAGCGATTTTGGTTGGGATTGGGAATTACCTGCCAAAACAGTGGCT   1440
EV52_contig000007      AACGTTGCCCGCAGCAGCGATTTTGGTTGGGATTGGGAATTACCTGCCAAAACAGTGGCT   1440
EV49_contig000024      AACGTTGCCCGCAGCAGCGATTTTGGTTGGGATTGGGAATTACCTGCCAAAACAGTGGCT   1440
EV51_contig000029      AACGTTGCCCGCAGCAGCGATTTTGGTTGGGATTGGGAATTACCTGCCAAAACAGTGGCT   1440
                       ************************************************************

EV50_contig000010      GTTTGCAAATTACTTGATTCAGGGTGGTAACGTTGTCCTGAATAACTTAGCCTTGCTGTT   1500
EV52_contig000007      GTTTGCAAATTACTTGATTCAGGGTGGTAACGTTGTCCTGAATAACTTAGCCTTGCTGTT   1500
EV49_contig000024      GTTTGCAAATTACTTGATTCAGGGTGGTAACGTTGTCCTGAATAACTTAGCCTTGCTGTT   1500
EV51_contig000029      GTTTGCAAATTACTTGATTCAGGGTGGTAACGTTGTCCTGAATAACTTAGCCTTGCTGTT   1500
                       ************************************************************
```

```
EV50_contig000010      TGCAGTCGGCTTAGCTATCGGGATGTCAGTCAATAAGGACGGCGCGGCAGCTATTGCTGG    1560
EV52_contig000007      TGCAGTCGGCTTAGCTATCGGGATGTCAGTCAATAAGGACGGCGCGGCAGCTATTGCTGG    1560
EV49_contig000024      TGCAGTCGGCTTAGCTATCGGGATGTCAGTCAATAAGGACGGCGCGGCAGCTATTGCTGG    1560
EV51_contig000029      TGCAGTCGGCTTAGCTATCGGGATGTCAGTCAATAAGGACGGCGCGGCAGCTATTGCTGG    1560
                       ************************************************************

EV50_contig000010      TTTGATCGCGTTTGAAGTTCCGGTGATGGTTTTAAAACCAGCGACATTAGCGACGATGCT    1620
EV52_contig000007      TTTGATCGCGTTTGAAGTTCCGGTGATGGTTTTAAAACCAGCGACATTAGCGACGATGCT    1620
EV49_contig000024      TTTGATCGCGTTTGAAGTTCCGGTGATGGTTTTAAAACCAGCGACATTAGCGACGATGCT    1620
EV51_contig000029      TTTGATCGCGTTTGAAGTTCCGGTGATGGTTTTAAAACCAGCGACATTAGCGACGATGCT    1620
                       ************************************************************

EV50_contig000010      GAATGTTAAAGTGAGTCAGATCAATCCAGCATTTAGCGCGTTAGATAACAATGTGCTGAT    1680
EV52_contig000007      GAATGTTAAAGTGAGTCAGATCAATCCAGCATTTAGCGCGTTAGATAACAATGTGCTGAT    1680
EV49_contig000024      GAATGTTAAAGTGAGTCAGATCAATCCAGCATTTAGCGCGTTAGATAACAATGTGCTGAT    1680
EV51_contig000029      GAATGTTAAAGTGAGTCAGATCAATCCAGCATTTAGCGCGTTAGATAACAATGTGCTGAT    1680
                       ************************************************************

EV50_contig000010      TGGAATCAGTGCTGGACTGATTGCGGCTGCGCTCTATAATCGGTTCCACGAAGTAAAATT    1740
EV52_contig000007      TGGAATCAGTGCTGGACTGATTGCGGCTGCGCTCTATAATCGGTTCCACGAAGTAAAATT    1740
EV49_contig000024      TGGAATCAGTGCTGGACTGATTGCGGCTGCGCTCTATAATCGGTTCCACGAAGTAAAATT    1740
EV51_contig000029      TGGAATCAGTGCTGGACTGATTGCGGCTGCGCTCTATAATCGGTTCCACGAAGTAAAATT    1740
                       ************************************************************

EV50_contig000010      ACCAATGGCGCTCTCATTCTTTAGTGGTAAACGCCTAGTGCCAATTATGGCTGCTTTTGT    1800
EV52_contig000007      ACCAATGGCGCTCTCATTCTTTAGTGGTAAACGCCTAGTGCCAATTATGGCTGCTTTTGT    1800
EV49_contig000024      ACCAATGGCGCTCTCATTCTTTAGTGGTAAACGCCTAGTGCCAATTATGGCTGCTTTTGT    1800
EV51_contig000029      ACCAATGGCGCTCTCATTCTTTAGTGGTAAACGCCTAGTGCCAATTATGGCTGCTTTTGT    1800
                       ************************************************************

EV50_contig000010      GATGCTGATTGTGACGGCGGTATTGTACCTTGTTTGGCCGTTTGTCTACGATGCCATCGT    1860
EV52_contig000007      GATGCTGATTGTGACGGCGGTATTGTACCTTGTTTGGCCGTTTGTCTACGATGCCATCGT    1860
EV49_contig000024      GATGCTGATTGTGACGGCGGTATTGTACCTTGTTTGGCCGTTTGTCTACGATGCCATCGT    1860
EV51_contig000029      GATGCTGATTGTGACGGCGGTATTGTACCTTGTTTGGCCGTTTGTCTACGATGCCATCGT    1860
                       ************************************************************

EV50_contig000010      CCTCTTTGCAACGGGAATTTCAAAATTAGGTTTCGTGGGGGCTGGTCTATATGGCTTTTT    1920
EV52_contig000007      CCTCTTTGCAACGGGAATTTCAAAATTAGGTTTCGTGGGGGCTGGTCTATATGGCTTTTT    1920
EV49_contig000024      CCTCTTTGCAACGGGAATTTCAAAATTAGGTTTCGTGGGGGCTGGTCTATATGGCTTTTT    1920
EV51_contig000029      CCTCTTTGCAACGGGAATTTCAAAATTAGGTTTCGTGGGGGCTGGTCTATATGGCTTTTT    1920
                       ************************************************************

EV50_contig000010      CAATCGTTTATTGATTCCGACTGGTTTGCACCACGCCTTAAATTCCGTATTCTGGTATAA    1980
EV52_contig000007      CAATCGTTTATTGATTCCGACTGGTTTGCACCACGCCTTAAATTCCGTATTCTGGTATAA    1980
EV49_contig000024      CAATCGTTTATTGATTCCGACTGGTTTGCACCACGCCTTAAATTCCGTATTCTGGTATAA    1980
EV51_contig000029      CAATCGTTTATTGATTCCGAATGGTTTGCACCACGCCTTAAATTCCGTATTCTGGTATAA    1980
                       ********************.***************************************

EV50_contig000010      CGTTGCGGGGATCAATGATATTGGCAATTTCTGGGCCAGCCATGGTGTTAAAGGAATCAC    2040
EV52_contig000007      CGTTGCGGGGATCAATGATATTGGCAATTTCTGGGCCAGCCATGGTGTTAAAGGAATCAC    2040
EV49_contig000024      CGTTGCGGGGATCAATGATATTGGCAATTTCTGGGCCAGCCATGGTGTTAAAGGAATCAC    2040
EV51_contig000029      CGTTGCGGGGATCAATGATATTGGCAATTTCTGGGCCAGCCATGGTGTTAAAGGAATCAC    2040
                       ************************************************************

EV50_contig000010      AGGGATGTATGAAGCTGGATTCTTCCCAATTATGATGTTTGGCTTGCCAGCTGGTG----    2096
EV52_contig000007      AGGGATGTATGAAGCTGGATTCTTCCCAATTATGATGTTTGGCTTGCCAGCTGGTGCGTA    2100
EV49_contig000024      AGGGATGTATGAAGCTGGATTCTTCCCAATTATGATGTTTGGCTTGCCAGCTGGTGCGTA    2100
EV51_contig000029      AGGGATGTATGAAGCTGGATTCTTCCCAATTATGATGTTTGGCTTGCCAGCTGGTGCGTA    2100
                       ************************************************************

EV50_contig000010      ------------------------------------------------------------    2096
EV52_contig000007      TGCAATTTATCGTAATGCACGACCGGAACGAAAAAAAGAAGTGGGTTCATTGATGTTAGC    2160
EV49_contig000024      TGCAATTTATCGTAATGCACGACCGGAACGAAAAAAAGAAGTGGGTTCATTGATGTTAGC    2160
EV51_contig000029      TGCAATTTATCGTAATGCACGACCGGAACGAAAAAAAGAAGTGGGTTCATTGATGTTAGC    2160


EV50_contig000010      -----CGTTTGCGTCGTTCTTTACGGGGGTGACCGAACCGCTTGAATTTTCATTCATGTT    2151
EV52_contig000007      TGGTGCGTTTGCGTCGTTCTTTACGGGGGTGACCGAACCGCTTGAATTTTCATTCATGTT    2220
EV49_contig000024      TGGTGCGTTTGCGTCGTTCTTTACGGGGGTGACCGAACCGCTTGAATTTTCATTCATGTT    2220
EV51_contig000029      TGGTGCGTTTGCGTCGTTCTTTACGGGGGTGACCGAACCGCTTGAATTTTCATTCATGTT    2220
                            ****************************************************

EV50_contig000010      CGTGGCATGGCCGCTCTACTTATTACATGCCGTTTTTATGGGACTATCATTAGGATTCGC    2211
EV52_contig000007      CGTGGCATGGCCGCTCTACTTATTACATGCCGTTTTTATGGGACTATCATTAGGATTCGC    2280
EV49_contig000024      CGTGGCATAGCCGCTCTACTTATTACATGCCGTTTTTATGGGACTATCATTAGGATTCGC    2280
EV51_contig000029      CGTGGCATGGCCGCTCTACTTATTACATGCCGTTTTTATGGGACTATCATTAGGATTCGC    2280
                       ********.***************************************************

EV50_contig000010      CGCGTTAATGCACTGGACCGCCAGCTTTTCGTTTAGTGGCGGGTTAGTCGATTATTTATT    2271
EV52_contig000007      CGCGTTAATGCACTGGACCGCCAGCTTTTCGTTTAGTGGCGGGTTAGTCGATTATTTATT    2340
EV49_contig000024      CGCGTTAATGCACTGGACCGCCAGCTTTTCGTTTAGTGGCGGGTTAGTCGATTATTTATT    2340
EV51_contig000029      CGCGTTAATGCACTGGACCGCCAGCTTTTCGTTTAGTGGCGGGTTAGTCGATTATTTATT    2340
                       ************************************************************
```

```
EV50_contig000010      GAGTTTTCGGATGCCGTTAGCCAACCAACCCTATATGTTACTGGTTCAAGGGCTGGTGAT    2331
EV52_contig000007      GAGTTTTCGGATGCCGTTAGCCAACCAACCCTATATGTTACTGGTTCAAGGGCTGGTGAT    2400
EV49_contig000024      GAGTTTTCGGATGCCGTTAGCCAACCAACCCTATATGTTACTGGTTCAAGGGCTGGTGAT    2400
EV51_contig000029      GAGTTTTCGGATGCCGTTAGCCAACCAACCCTATATGTTACTGGTTCAAGGGCTGGTGAT    2400
                       ************************************************************

EV50_contig000010      GGCCGTCATTTATTACTTTGGTTTTGACTTTGCAATCAAACGGTTTAATTTGAAGACGCC    2391
EV52_contig000007      GGCCGTCATTTATTACTTTGGTTTTGACTTTGCAATCAAACGGTTTAATTTGAAGACGCC    2460
EV49_contig000024      GGCCGTCATTTATTACTTTGGTTTTGACTTTGCAATCAAACGGTTTAATTTGAAGACGCC    2460
EV51_contig000029      GGCCGTCATTTATTACTTTGGTTTTGACTTTGCAATCAAACGGTTTAATTTGAAGACGCC    2460
                       ************************************************************

EV50_contig000010      TGGTCGTGAAGTTGTGTCTGCAGATGTCGATGGGGTGGGTGCACCTGCTAGTCCAGCGGT    2451
EV52_contig000007      TGGTCGTGAAGTTGTGTCTGCAGATGTCGATGGGGTGGGTGCACCTGCTAGTCCAGCGGT    2520
EV49_contig000024      TGGTCGTGAAGTTGTGTCTGCAGATGTCGATGGGGTGGGTGCACCTGCTAGTCCAGCGGT    2520
EV51_contig000029      TGGTCGTGAAGTTGTGTCTGCAGATGTCGATGGGGTGGGTGCACCTGCTAGTCCAGCGGT    2520
                       ************************************************************

EV50_contig000010      GGCCGTTGCAGCAACGGATGATAAGTATATGCGCCAAGCAAAGCAAATTTATGCAGCTAT    2511
EV52_contig000007      GGCCGTTGCAGCAACGGATGATAAGTATATGCGCCAAGCAAAGCAAATTTATGCAGCTAT    2580
EV49_contig000024      GGCCGTTGCAGCAACGGATGATAAGTATATGCGCCAAGCAAAGCAAATTTATGCAGCTAT    2580
EV51_contig000029      GGCCGTTGCAGCAACGGATGATAAGTATATGCGCCAAGCAAAGCAAATTTATGCAGCTAT    2580
                       ************************************************************

EV50_contig000010      TGGTGGTCATGACAATATTAGTGTGATCAACAATTGTACGACGCGGCTGCGGTTACAACT    2571
EV52_contig000007      TGGTGGTCATGACAATATTAGTGTGATCAACAATTGTACGACGCGGCTGCGGTTACAACT    2640
EV49_contig000024      TGGTGGTCATGACAATATTAGTGTGATCAACAATTGTACGACGCGGCTGCGGTTACAACT    2640
EV51_contig000029      TGGTGGTCATGACAATATTAGTGTGATCAACAATTGTACGACGCGGCTGCGGTTACAACT    2640
                       ************************************************************

EV50_contig000010      TAAGGATACGGAAAAAGTCGATCAGCCGGCCGTAATGGCTGCTGGCGTGCCTGGTTTGAA    2631
EV52_contig000007      TAAGGATACGGAAAAAGTCGATCAGCCGGCCGTAATGGCTGCTGGCGTGCCTGGTTTGAA    2700
EV49_contig000024      TAAGGATACGGAAAAAGTCGATCAGCCGGCCGTAATGGCTGCTGGCGTGCCTGGTTTGAA    2700
EV51_contig000029      TAAGGATACGGAAAAAGTCGATCAGCCGGCCGTAATGGCTGCTGGCGTGCCTGGTTTGAA    2700
                       ************************************************************

EV50_contig000010      CGTACTGGATGTTCATAACATTCACATTGTGATCGGCACGGAAGTCCAGTTCGTCGCGGA    2691
EV52_contig000007      CGTACTGGATGTTCATAACATTCACATTGTGATCGGCACGGAAGTCCAGTTCGTCGCGGA    2760
EV49_contig000024      CGTACTGGATGTTCATAACATTCACATTGTGATCGGCACGGAAGTCCAGTTCGTCGCGGA    2760
EV51_contig000029      CGTACTGGATGTTCATAACATTCACATTGTGATCGGCACGGAAGTCCAGTTCGTCGCGGA    2760
                       ************************************************************

EV50_contig000010      GGCTTTACAAAAATTATTTTCTGGTCAGGTGGCGACGACCCCGGCATCTGATGCTGAATC    2751
EV52_contig000007      GGCTTTACAAAAATTATTTTCTGGTCAGGTGGCGACGACCCCGGCATCTGATGCTGAATC    2820
EV49_contig000024      GGCTTTACAAAAATTATTTTCTGGTCAGGTGGCGACGACCCCGGCATCTGATGCTGAATC    2820
EV51_contig000029      GGCTTTACAAAAATTATTTTCTGGTCAGGTGGCGACGACCCCGGCATCTGATGCTGAATC    2820
                       ************************************************************

EV50_contig000010      AAAGGCCCCGGTTGAACCGCAAACTGCTACTGTAACGGAAGCGCCGGTGACAACGATTTT    2811
EV52_contig000007      AAAGGCCCCGGTTGAACCGCAAACTGCTACTGTAACGGAAGCGCCGGTGACAACGATTTT    2880
EV49_contig000024      AAAGGCCCCGGTTGAACCGCAAACTGCTACTGTAACGGAAGCGCCGGTGACAACGATTTT    2880
EV51_contig000029      AAAGGCCCCGGTTGAACCGCAAACTGCTACTGTAACGGAAGCGCCGGTGACAACGATTTT    2880
                       ************************************************************

EV50_contig000010      GCGGGCACCGGCAACGGGACAATTAATGCCGATTAGTACGGTTGCGGATGAGACGTTTGC    2871
EV52_contig000007      GCGGGCACCGGCAACGGGACAATTAATGCCGATTAGTACGGTTGCGGATGAGACGTTTGC    2940
EV49_contig000024      GCGGGCACCGGCAACGGGACAATTAATGCCGATTAGTACGGTTGCGGATGAGACGTTTGC    2940
EV51_contig000029      GCGGGCACCGGCAACGGGACAATTAATGCCGATTAGTACGGTTGCGGATGAGACGTTTGC    2940
                       ************************************************************

EV50_contig000010      TGGTAAACTCTTAGGTGATGGTTACGCTGTTGAACCCGAAGATGGTGAAGTCGTTGCTCC    2931
EV52_contig000007      TGGTAAACTCTTAGGTGATGGTTACGCTGTTGAACCCGAAGATGGTGAAGTCGTTGCTCC    3000
EV49_contig000024      TGGTAAACTCTTAGGTGATGGTTACGCTGTTGAACCCGAAGATGGTGAAGTCGTTGCTCC    3000
EV51_contig000029      TGGTAAACTCTTAGGTGATGGTTACGCTGTTGAACCCGAAGATGGTGAAGTCGTTGCTCC    3000
                       ************************************************************

EV50_contig000010      GGTTAGTGGGACAGTGACAAGTGTCTTTCCGACCAAACACGCTATCGGTTTGAAGACGAC    2991
EV52_contig000007      GGTTAGTGGGACAGTGACAAGTGTCTTTCCGACCAAACACGCTATCGGTTTGAAGACGAC    3060
EV49_contig000024      GGTTAGTGGGACAGTGACAAGTGTCTTTCCGACCAAACACGCTATCGGTTTGAAGACGAC    3060
EV51_contig000029      GGTTAGTGGGACAGTGACAAGTGTCTTTCCGACCAAACACGCTATCGGTTTGAAGACGAC    3060
                       ************************************************************

EV50_contig000010      GAGTGGTTTAGAAGTCTTATTACACATGGGGATCAATACCGTGGAAATGAATGGTACGCC    3051
EV52_contig000007      GAGTGGTTTAGAAGTCTTATTACACATGGGGATCAATACCGTGGAAATGAATGGTACGCC    3120
EV49_contig000024      GAGTGGTTTAGAAGTCTTATTACACATGGGGATCAATACCGTGGAAATGAATGGTACGCC    3120
EV51_contig000029      GAGTGGTTTAGAAGTCTTATTACACATGGGGATCAATACCGTGGAAATGAATGGTACGCC    3120
                       ************************************************************

EV50_contig000010      GTTCAAGTTACACGTGGCAACGGGTGATGAAATTGCGGCCGGTACTGCGGTAGCGACGGT    3111
EV52_contig000007      GTTCAAGTTACACGTGGCAACGGGTGATGAAATTGCGGCCGGTACTGCGGTAGCGACGGT    3180
EV49_contig000024      GTTCAAGTTACACGTGGCAACGGGTGATGAAATTGCGGCCGGTACTGCGGTAGCGACGGT    3180
EV51_contig000029      GTTCAAGTTACACGTGGCAACGGGTGATGAAATTGCGGCCGGTACTGCGGTAGCGACGGT    3180
                       ************************************************************
```

```
EV50_contig000010      TGACTTGGCTGCTATCAAGTCAGCCGGTAAAGCGACAACCATGATGGTCGTTATCACCAA    3171
EV52_contig000007      TGACTTGGCTGCTATCAAGTCAGCCGGTAAAGCGACAACCATGATGGTCGTTATCACCAA    3240
EV49_contig000024      TGACTTGGCTGCTATCAAGTCAGCCGGTAAAGCGACAACCATGATGGTCGTTATCACCAA    3240
EV51_contig000029      TGACTTGGCTGCTATCAAGTCAGCCGGTAAAGCGACAACCATGATGGTCGTTATCACCAA    3240
                       ************************************************************

EV50_contig000010      TATGGATCACGTTAACAAACTAACCCTTAATCCAACTGGTCATGTGACTAGCGGTGATTT    3231
EV52_contig000007      TATGGATCACGTTAACAAACTAACCCTTAATCCAACTGGTCATGTGACTAGCGGTGATTT    3300
EV49_contig000024      TATGGATCACGTTAACAAACTAACCCTTAATCCAACTGGTCATGTGACTAGCGGTGATTT    3300
EV51_contig000029      TATGGATCACGTTAACAAACTAACCCTTAATCCAACTGGTCATGTGACTAGCGGTGATTT    3300
                       ************************************************************

EV50_contig000010      GATTGGCGCAGCTGAATAAGTTACGCGCTGTCGACGTTGGGTTTATTGAAAAAGTTTCGT    3291
EV52_contig000007      GATTGGCGCAGCTGAATAAGTTACGCGCTGTCGACGTTGGGTTTATTGAAAAAGTTTCGT    3360
EV49_contig000024      GATTGGCGCAGCTGAATAAGTTACGCGCTGTCGACGTTGGGTTTATTGAAAAAGTTTCGT    3360
EV51_contig000029      GATTGGCGCAGCTGAATAAGTTACGCGCTGTCGACGTTGGGTTTATTGAAAAAGTTTCGT    3360
                       ************************************************************

EV50_contig000010      CTAAATCTTTTCGATTTAGGCGGAACTTTTTTTGTCGCATTGGTTATCGAGGCAACGGAT    3351
EV52_contig000007      CTAAATCTTTTCGATTTAGGCGGAACTTTTTTTGTCGCATTGGTTATCGAGGCAACGGAT    3420
EV49_contig000024      CTAAATCTTTTCGATTTAGGCGGAACTTTTTTTGTCGCATTGGTTATCGAGGCAACGGAT    3420
EV51_contig000029      CTAAATCTTTTCGATTTAGGCGGAACTTTTTTTGTCGCATTGGTTATCGAGGCAACGGAT    3420
                       ************************************************************

EV50_contig000010      GACAGCAAATGAACAGCTTGATGCGGGGACAGTGCGTTTGATTGCTTCCCGCTGACTCCC    3411
EV52_contig000007      GACAGCAAATGAACAGCTTGATGCGGGGACAGTGCGTTTGATTGCTTCCCGCTGACTCCC    3480
EV49_contig000024      GACAGCAAATGAACAGCTTGATGCGGGGACAGTGCGTTTGATTGCTTCCCGCTGACTCCC    3480
EV51_contig000029      GACAGCAAATGAACAGCTTGATGCGGGGACAGTGCGTTTGATTGCTTCCCGCTGACTCCC    3480
                       ************************************************************

EV50_contig000010      ACGATATACTATTGGCATTATAGGGGGGCGGTAACATGATTGACGAAGTATTAGTATTGA    3471
EV52_contig000007      ACGATATACTATTGGCATTATAGGGGGGCGGTAACATGATTGACGAAGTATTAGTATTGA    3540
EV49_contig000024      ACGATATACTATTGGCATTATAGGGGGGCGGTAACATGATTGACGAAGTATTAGTATTGA    3540
EV51_contig000029      ACGATATACTATTGGCATTATAGGGGGGCGGTAACATGATTGACGAAGTATTAGTATTGA    3540
                       ************************************************************

EV50_contig000010      ACTCAACGCCAACGGGGATTTTACAGAATTCAGTTTCAGGAAGCATGGTGATTGACGCGG    3531
EV52_contig000007      ACTCAACGCCAACGGGGATTTTACAGAATTCAGTTTCAGGAAGCATGGTGATTGACGCGG    3600
EV49_contig000024      ACTCAACGCCAACGGGGATTTTACAGAATTCAGTTTCAGGAAGCATGGTGATTGACGCGG    3600
EV51_contig000029      ACTCAACGCCAACGGGGATTTTACAGAATTCAGTTTCAGGAAGCATGGTGATTGACGCGG    3600
                       ************************************************************

EV50_contig000010      CACTTGATGCACGACTAGTCAAGCGACTAATGCAGGCGCTAACTCAATCGATGCTTAAGG    3591
EV52_contig000007      CACTTGATGCACGACTAGTCAAGCGACTAATGCAGGCGCTAACTCAATCGATGCTTAAGG    3660
EV49_contig000024      CACTTGATGCACGACTAGTCAAGCGACTAATGCAGGCGCTAACTCAATCGATGCTTAAGG    3660
EV51_contig000029      CACTTGATGCACGACTAGTCAAGCGACTAATGCAGGCGCTAACTCAATCGATGCTTAAGG    3660
                       ************************************************************

EV50_contig000010      TCAGGGTTGCAACCGCGGACGACTTACCGGTGACCGACCACCAATTGCTGGTCGTCAGTC    3651
EV52_contig000007      TCAGGGTTGCAACCGCGGACGACTTACCGGTGACCGACCACCAATTGCTGGTCGTCAGTC    3720
EV49_contig000024      TCAGGGTTGCAACCGCGGACGACTTACCGGTGACCGACCACCAATTGCTGGTCGTCAGTC    3720
EV51_contig000029      TCAGGGTTGCAACCGCGGACGACTTACCGGTGACCGACCACCAATTGCTGGTCGTCAGTC    3720
                       ************************************************************

EV50_contig000010      GACCAGTGGCAGCCGTGTTGGACTTACAAGCTAGTGCGCGAGTTCTGATTTATCCAGAAG    3711
EV52_contig000007      GACCAGTGGCAGCCGTGTTGGACTTACAAGCTAGTGCGCGAGTTCTGATTTATCCAGAAG    3780
EV49_contig000024      GACCAGTGGCAGCCGTGTTGGACTTACAAGCTAGTGCGCGAGTTCTGATTTATCCAGAAG    3780
EV51_contig000029      GACCAGTGGCAGCCGTGTTGGACTTACAAGCTAGTGCGCGAGTTCTGATTTATCCAGAAG    3780
                       ************************************************************

EV50_contig000010      CAACGGGACTGACAGCCGCTGGAATGACACAGCTAGCAGTGCAGATTGATCAACGATTGA    3771
EV52_contig000007      CAACGGGACTGACAGCCGCTGGAATGACACAGCTAGCAGTGCAGATTGATCAACGATTGA    3840
EV49_contig000024      CAACGGGACTGACAGCCGCTGGAATGACACAGCTAGCAGTGCAGATTGATCAACGATTGA    3840
EV51_contig000029      CAACGGGACTGACAGCCGCTGGAATGACACAGCTAGCAGTGCAGATTGATCAACGATTGA    3840
                       ************************************************************

EV50_contig000010      CTAAAATATCCAGATAACGATCAAAAAACGCCCAGGGCAGTTGTTCGCAAGTTACTGAAC    3831
EV52_contig000007      CTAAAATATCCAGATAACGATCAAAAAACGCCCAGGGCAGTTGTTCGCAAGTTACTGAAC    3900
EV49_contig000024      CTAAAATATCCAGATAACGATCAAAAAACGCCCAGGGCAGTTGTTCGCAAGTTACTGAAC    3900
EV51_contig000029      CTAAAATATCCAGATAACGATCAAAAAACGCCCAGGGCAGTTGTTCGCAAGTTACTGAAC    3900
                       ************************************************************

EV50_contig000010      AGTTGCCCTGGGCGTTTTTAATGATTAGTCTTTGTGGTGTTTTTTAATGATAGTTACTAA    3891
EV52_contig000007      AGTTGCCCTGGGCGTTTTTAATGATTAGTCTTTGTGGTGTTTTTTAATGATAGTTACTAA    3960
EV49_contig000024      AGTTGCCCTGGGCGTTTTTAATGATTAGTCTTTGTGGTGTTTTTTAATGATAGTTACTAA    3960
EV51_contig000029      AGTTGCCCTGGGCGTTTTTAATGATTAGTCTTTGTGGTGTTTTTTAATGATAGTTACTAA    3960
                       ************************************************************

EV50_contig000010      TTTCAATCAAGTTGTTATCAGGATCACGAACATACAGTGACGTCAACTTTCCGTGGGCAC    3951
EV52_contig000007      TTTCAATCAAGTTGTTATCAGGATCACGAACATACAGTGACGTCAACTTTCCGTGGGCAC    4020
EV49_contig000024      TTTCAATCAAGTTGTTATCAGGATCACGAACATACAGTGACGTCAACTTTCCGTGGGCAC    4020
EV51_contig000029      TTTCAATCAAGTTGTTATCAGGATCACGAACATACAGTGACGTCAACTTTCCGTGGGCAC    4020
                       ************************************************************
```

```
EV50_contig000010      CAGTTCTTTCAACCGGCCCGGCAATCACATCCACGAAATAACTCTTTAAGTGATGCTGAA    4011
EV52_contig000007      CAGTTCTTTCAACCGGCCCGGCAATCACATCCACGAAATAACTCTTTAAGTGATGCTGAA    4080
EV49_contig000024      CAGTTCTTTCAACCGGCCCGGCAATCACATCCACGAAATAACTCTTTAAGTGATGCTGAA    4080
EV51_contig000029      CAGTTCTTTCAACCGGCCCGGCAATCACATCCACGAAATAACTCTTTAAGTGATGCTGAA    4080
                       ************************************************************

EV50_contig000010      TATCATCAATATTATCCTTCGCAATCAAGCAAAGATCTGCACTGCCGGGTGTCGGCTTAG    4071
EV52_contig000007      TATCATCAATATTATCCTTCGCAATCAAGCAAAGATCTGCACTGCCGGGTGTCGGCTTAG    4140
EV49_contig000024      TATCATCAATATTATCCTTCGCAATCAAGCAAAGATCTGCACTGCCGGGTGTCGGCTTAG    4140
EV51_contig000029      TATCATCAATATTATCCTTCGCAATCAAGCAAAGATCTGCACTGCCGGGTGTCGGCTTAG    4140
                       ************************************************************

EV50_contig000010      CCGCAATGGGTTGGTGTGGCTGGTTAGTGGTCTGAAAGTTAATTTTTTGTTTACCCACTA    4131
EV52_contig000007      CCGCAATGGGTTGGTGTGGCTGGTTAGTGGTCTGAAAGTTAATTTTTTGTTTACCCACTA    4200
EV49_contig000024      CCGCAATGGGTTGGTGTGGCTGGTTAGTGGTCTGAAAGTTAATTTTTTGTTTACCCACTA    4200
EV51_contig000029      CCGCAATGGGTTGGTGTGGCTGGTTAGTGGTCTGAAAGTTAATTTTTTGTTTACCCACTA    4200
                       ************************************************************

EV50_contig000010      AGACAGCTTGACGGTCACCATCGAAGGTCACAATGGGAAGGTCAAATACTTCATGGTAAA    4191
EV52_contig000007      AGACAGCTTGACGGTCACCATCGAAGGTCACAATGGGAAGGTCAAATACTTCATGGTAAA    4260
EV49_contig000024      AGACAGCTTGACGGTCACCATCGAAGGTCACAATGGGAAGGTCAAATACTTCATGGTAAA    4260
EV51_contig000029      AGACAGCTTGACGGTCACCATCGAAGGTCACAATGGGAAGGTCAAATACTTCATGGTAAA    4260
                       ************************************************************

EV50_contig000010      ACCGTAATGACCGTGCAATGTCAGTAACGGTTAGGGTCAAATGATCGATATCGCGAATAT    4251
EV52_contig000007      ACCGTAATGACCGTGCAATGTCAGTAACGGTTAGGGTCAAATGATCGATATCGCGAATAT    4320
EV49_contig000024      ACCGTAATGACCGTGCAATGTCAGTAACGGTTAGGGTCAAATGATCGATATCGCGAATAT    4320
EV51_contig000029      ACCGTAATGACCGTGCAATGTCAGTAACGGTTAGGGTCAAATGATCGATATCGCGAATAT    4320
                       ************************************************************

EV50_contig000010      TCACAAGCTCAAATCCCTTCAAAAATAATTCATTATTAAACGTGTCGTACTCATCAGATT    4311
EV52_contig000007      TCACAAGCTCAAATCCCTTCAAAAATAATTCATTATTAAACGTGTCGTACTCATCAGATT    4380
EV49_contig000024      TCACAAGCTCAAATCCCTTCAAAAATAATTCATTATTAAACGTGTCGTACTCATCAGATT    4380
EV51_contig000029      TCACAAGCTCAAATCCCTTCAAAAATAATTCATTATTAAACGTGTCGTACTCATCAGATT    4380
                       ************************************************************

EV50_contig000010      TGGCATGAACCATGATTAAAATGTACCAGTTAGATTGCTGTTCGTCAGCCGTTGCGCGTC    4371
EV52_contig000007      TGGCATGAACCATGATTAAAATGTACCAGTTAGATTGCTGTTCGTCAGCCGTTGCGCGTC    4440
EV49_contig000024      TGGCATGAACCATGATTAAAATGTACCAGTTAGATTGCTGTTCGTCAGCCGTTGCGCGTC    4440
EV51_contig000029      TGGCATGAACCATGATTAAAATGTACCAGTTAGATTGCTGTTCGTCAGCCGTTGCGCGTC    4440
                       ************************************************************

EV50_contig000010      CTATGCCGACCTCCGGGGGCTGGGTGCCAATTGCTGGAACGCAGCCGACATCGATTTGAGC    4431
EV52_contig000007      CTATGCCGACCTCCGGGGGCTGGGTGCCAATTGCTGGAACGCAGCCGACATCGATTTGAGC    4500
EV49_contig000024      CTATGCCGACCTCCGGGGGCTGGGTGCCAATTGCTGGAACGCAGCCGACATCGATTTGAGC    4500
EV51_contig000029      CTATGCCGACCTCCGGGGGCTGGGTGCCAATTGCTGGAACGCAGCCGACATCGATTTGAGC    4500
                       ************************************************************

EV50_contig000010      TAACGCATAACCCGCGTCATCTCAAATACGAGTCTTATTCTAAGCCGGAAGCAGACC        4488
EV52_contig000007      TAACGCATAACCCGCGTCATCTCAAATACGAGTCTTATTCTAAGCCGGAAGCAGACC        4557
EV49_contig000024      TAACGCATAACCCGCGTCATCTCAAATACGAGTCTTATTCTAAGCCGGAAGCAGACC        4557
EV51_contig000029      TAACGCATAACCCGCGTCATCTCAAATACGAGTCTTATTCTAAGCCGGAAGCAGACC        4557
                       ************************************************************
```

Appendix 8 Sequence alignment of *pts18CBA* region of *L. plantarum* subsp. *plantarum* 14917

CLUSTAL OMEGA (http://www.ebi.ac.uk/Tools/msa/clustalo/) multiple sequence alignment performed with default settings from the region of the genomes containing the *pts18BCA* CDS from *L. plantarum* subsp. plantarum ATCC 14917 . The start and stop codons or *pts18CBA* are coloured red and green respectively.

```
EV48_contig000009      GAGGGTGCGACAATGACAACGACAATTGAAGTGGGCATGTTAGGGTTAGGAACTGTGGGT    60
EV55_contig000014      GAGGGTGCGACAATGACAACGACAATTGAAGTGGGCATGTTAGGGTTAGGAACTGTGGGT    60
EV54_contig000003      GAGGGTGCGACAATGACAACGACAATTGAAGTGGGCATGTTAGGGTTAGGAACTGTGGGT    60
EV53_contig000052      GAGGGTGCGACAATGACAACGACAATTGAAGTGGGCATGTTAGGGTTAGGAACTGTGGGT    60
                       ************************************************************
```

```
EV48_contig000009      AGTGGGGTGGTTGAACGGTTGACCCGCTCGGCTGCCAAAATTGAACAAACGCAGGGTATT    120
EV55_contig000014      AGTGGGGTGGTTGAACGGTTGACCCGCTCGGCTGCCAAAATTGAACAAACGCAGGGTATT    120
EV54_contig000003      AGTGGGGTGGTTGAACGGTTGACCCGCTCGGCTGCCAAAATTGAACAAACGCAGGGTATT    120
EV53_contig000052      AGTGGGGTGGTTGAACGGTTGACCCGCTCGGCTGCCAAAATTGAACAAACGCAGGGTATT    120
                       ************************************************************

EV48_contig000009      CGGTTGCACCTAGCAGCAGTTGCCGTTAATCATTTGAACGCGCCTCGGACTGTTCAGTTA    180
EV55_contig000014      CGGTTGCACCTAGCAGCAGTTGCCGTTAATCATTTGAACGCGCCTCGGACTGTTCAGTTA    180
EV54_contig000003      CGGTTGCACCTAGCAGCAGTTGCCGTTAATCATTTGAACGCGCCTCGGACTGTTCAGTTA    180
EV53_contig000052      CGGTTGCACCTAGCAGCAGTTGCCGTTAATCATTTGAACGCGCCTCGGACTGTTCAGTTA    180
                       ************************************************************

EV48_contig000009      CCGATTGGGACGCGCTTGACCGATTCATTAACGAACGTTGTTTGTGACCGAAAAATCCAA    240
EV55_contig000014      CCGATTGGGACGCGCTTGACCGATTCATTAACGAACGTTGTTTGTGACCGAAAAATCCAA    240
EV54_contig000003      CCGATTGGGACGCGCTTGACCGATTCATTAACGAACGTTGTTTGTGACCGAAAAATCCAA    240
EV53_contig000052      CCGATTGGGACGCGCTTGACCGATTCATTAACGAACGTTGTTTGTGACCGAAAAATCCAA    240
                       ************************************************************

EV48_contig000009      CTAGTCATTGAAGTCATGGGGACGGTCGCAATTGCTAAAAAGGCCATTGTGGCAGCACTT    300
EV55_contig000014      CTAGTCATTGAAGTCATGGGGACGGTCGCAATTGCTAAAAAGGCCATTGTGGCAGCACTT    300
EV54_contig000003      CTAGTCATTGAAGTCATGGGGACGGTCGCAATTGCTAAAAAGGCCATTGTGGCAGCACTT    300
EV53_contig000052      CTAGTCATTGAAGTCATGGGGACGGTCGCAATTGCTAAAAAGGCCATTGTGGCAGCACTT    300
                       ************************************************************

EV48_contig000009      AATCAGGGTAAGGCTGTCGTGACCGCTAATAAAGACTTAATTGCAACGGCGGGACCTGAA    360
EV55_contig000014      AATCAGGGTAAGGCTGTCGTGACCGCTAATAAAGACTTAATTGCAACGGCGGGACCTGAA    360
EV54_contig000003      AATCAGGGTAAGGCTGTCGTGACCGCTAATAAAGACTTAATTGCAACGGCGGGACCTGAA    360
EV53_contig000052      AATCAGGGTAAGGCTGTCGTGACCGCTAATAAAGACTTAATTGCAACGGCGGGACCTGAA    360
                       ************************************************************

EV48_contig000009      TTGGCTGCGTTAGCCAAAAAACAGGGCTGTGATTTATTTTACGAGGCTAGTGTTGCGGGG    420
EV55_contig000014      TTGGCTGCGTTAGCCAAAAAACAGGGCTGTGATTTATTTTACGAGGCTAGTGTTGCGGGG    420
EV54_contig000003      TTGGCTGCGTTAGCCAAAAAACAGGGCTGTGATTTATTTTACGAGGCTAGTGTTGCGGGG    420
EV53_contig000052      TTGGCTGCGTTAGCCAAAAAACAGGGCTGTGATTTATTTTACGAGGCTAGTGTTGCGGGG    420
                       ************************************************************

EV48_contig000009      GGGATTCCGATCTTGCGGACATTAACAGATAGTTACGTGACGGACAATGTGCAGGCGGTC    480
EV55_contig000014      GGGATTCCGATCTTGCGGACATTAACAGATAGTTACGTGACGGACAATGTGCAGGCGGTC    480
EV54_contig000003      GGGATTCCGATCTTGCGGACATTAACAGATAGTTACGTGACGGACAATGTGCAGGCGGTC    480
EV53_contig000052      GGGATTCCGATCTTGCGGACATTAACAGATAGTTACGTGACGGACAATGTGCAGGCGGTC    480
                       ************************************************************
EV48_contig000009      AGCGGAATCATCAACGGGACGGCTAATTATATGCTAAGTGCAATGGCAACCGGGCAATCG    540
EV55_contig000014      AGCGGAATCATCAACGGGACGGCTAATTATATGCTAAGTGCAATGGCAACCGGGCAATCG    540
EV54_contig000003      AGCGGAATCATCAACGGGACGGCTAATTATATGCTAAGTGCAATGGCAACCGGGCAATCG    540
EV53_contig000052      AGCGGAATCATCAACGGGACGGCTAATTATATGCTAAGTGCAATGGCAACCGGGCAATCG    540
                       ************************************************************

EV48_contig000009      TACGAACAAGCTTTAGCTACTGCACAAGCAGCCGGGTACGCCGAAGCTGATCCGACTAAT    600
EV55_contig000014      TACGAACAAGCTTTAGCTACTGCACAAGCAGCCGGGTACGCCGAAGCTGATCCGACTAAT    600
EV54_contig000003      TACGAACAAGCTTTAGCTACTGCACAAGCAGCCGGGTACGCCGAAGCTGATCCGACTAAT    600
EV53_contig000052      TACGAACAAGCTTTAGCTACTGCACAAGCAGCCGGGTACGCCGAAGCTGATCCGACTAAT    600
                       ************************************************************

EV48_contig000009      GATGTTGCGGGCATTGATGCGGCATACAAATTGATGATCCTCAGTCGCTTTGCTTTTGGG    660
EV55_contig000014      GATGTTGCGGGCATTGATGCGGCATACAAATTGATGATCCTCAGTCGCTTTGCTTTTGGG    660
EV54_contig000003      GATGTTGCGGGCATTGATGCGGCATACAAATTGATGATCCTCAGTCGCTTTGCTTTTGGG    660
EV53_contig000052      GATGTTGCGGGCATTGATGCGGCATACAAATTGATGATCCTCAGTCGCTTTGCTTTTGGG    660
                       ************************************************************

EV48_contig000009      CAAGAGTTGAGCCTGCCGCAGATTGCACCCACGGGAATCACGCATTTAAGCGCTTCGGTT    720
EV55_contig000014      CAAGAGTTGAGCCTGCCGCAGATTGCACCCACGGGAATCACGCATTTAAGCGCTTCGGTT    720
EV54_contig000003      CAAGAGTTGAGCCTGCCGCAGATTGCACCCACGGGAATCACGCATTTAAGCGCTTCGGTT    720
EV53_contig000052      CAAGAGTTGAGCCTGCCGCAGATTGCACCCACGGGAATCACGCATTTAAGCGCTTCGGTT    720
                       ************************************************************

EV48_contig000009      TGTCGATTAGCAGCTGAAAATGGCTGGCAGATTAAATTGCTCGCGCAGATTCAGCGACAC    780
EV55_contig000014      TGTCGATTAGCAGCTGAAAATGGCTGGCAGATTAAATTGCTCGCGCAGATTCAGCGACAC    780
EV54_contig000003      TGTCGATTAGCAGCTGAAAATGGCTGGCAGATTAAATTGCTCGCGCAGATTCAGCGACAC    780
EV53_contig000052      TGTCGATTAGCAGCTGAAAATGGCTGGCAGATTAAATTGCTCGCGCAGATTCAGCGACAC    780
                       ************************************************************

EV48_contig000009      GGATCTGGTTTGTATTGTCGTGTGGCGCCAGTGGCCGTGCCGGTTGATCAACCCCTGAGT    840
EV55_contig000014      GGATCTGGTTTGTATTGTCGTGTGGCGCCAGTGGCCGTGCCGGTTGATCAACCCCTGAGT    840
EV54_contig000003      GGATCTGGTTTGTATTGTCGTGTGGCGCCAGTGGCCGTGCCGGTTGATCAACCCCTGAGT    840
EV53_contig000052      GGATCTGGTTTGTATTGTCGTGTGGCGCCAGTGGCCGTGCCGGTTGATCAACCCCTGAGT    840
                       ************************************************************

EV48_contig000009      CAGATCAATGGGGTCCAAAACGCTGTAGCTGTTCAAAGTGAAGCCATTGGTACCAGCTTG    900
EV55_contig000014      CAGATCAATGGGGTCCAAAACGCTGTAGCTGTTCAAAGTGAAGCCATTGGTACCAGCTTG    900
EV54_contig000003      CAGATCAATGGGGTCCAAAACGCTGTAGCTGTTCAAAGTGAAGCCATTGGTACCAGCTTG    900
EV53_contig000052      CAGATCAATGGGGTCCAAAACGCTGTAGCTGTTCAAAGTGAAGCCATTGGTACCAGCTTG    900
                       ************************************************************
```

```
EV48_contig000009      TATACCGGACCGGGTGCGGGAAGTACGGCGACGGCTAATAGTGTTCTAAATGACGTCTTA    960
EV55_contig000014      TATACCGGACCGGGTGCGGGAAGTACGGCGACGGCTAATAGTGTTCTAAATGACGTCTTA    960
EV54_contig000003      TATACCGGACCGGGTGCGGGAAGTACGGCGACGGCTAATAGTGTTCTAAATGACGTCTTA    960
EV53_contig000052      TATACCGGACCGGGTGCGGGAAGTACGGCGACGGCTAATAGTGTTCTAAATGACGTCTTA    960
                       ************************************************************

EV48_contig000009      GTAGCGGCCAAGCACCTTATCAATGGTTATCGAATGATTCGTAACCAGAAAAGAAGTGTC    1020
EV55_contig000014      GTAGCGGCCAAGCACCTTATCAATGGTTATCGAATGATTCGTAACCAGAAAAGAAGTGTC    1020
EV54_contig000003      GTAGCGGCCAAGCACCTTATCAATGGTTATCGAATGATTCGTAACCAGAAAAGAAGTGTC    1020
EV53_contig000052      GTAGCGGCCAAGCACCTTATCAATGGTTATCGAATGATTCGTAACCAGAAAAGAAGTGTC    1020
                       ************************************************************

EV48_contig000009      TCGTTAACGGTCACTCGATTTGACCAGTTACCACAAACATACTTAGGCATTGGTCCGGTC    1080
EV55_contig000014      TCGTTAACGGTCACTCGATTTGACCAGTTACCACAAACATACTTAGGCATTGGTCCGGTC    1080
EV54_contig000003      TCGTTAACGGTCACTCGATTTGACCAGTTACCACAAACATACTTAGGCATTGGTCCGGTC    1080
EV53_contig000052      TCGTTAACGGTCACTCGATTTGACCAGTTACCACAAACATACTTAGGCATTGGTCCGGTC    1080
                       ************************************************************

EV48_contig000009      GCACCAGAAGCCGTGAAATATTACGCTAATGAAGCTGATTTCTCAGTAAAAACGATTGCC    1140
EV55_contig000014      GCACCAGAAGCCGTGAAATATTACGCTAATGAAGCTGATTTCTCAGTAAAAACGATTGCC    1140
EV54_contig000003      GCACCAGAAGCCGTGAAATATTACGCTAATGAAGCTGATTTCTCAGTAAAAACGATTGCC    1140
EV53_contig000052      GCACCAGAAGCCGTGAAATATTACGCTAATGAAGCTGATTTCTCAGTAAAAACGATTGCC    1140
                       ************************************************************

EV48_contig000009      GATGATTGTTATCAGATTACTGGATTAAGTGCGGTGAAACGCAAAAAGCTGCGTGCGAAC    1200
EV55_contig000014      GATGATTGTTATCAGATTACTGGATTAAGTGCGGTGAAACGCAAAAAGCTGCGTGCGAAC    1200
EV54_contig000003      GATGATTGTTATCAGATTACTGGATTAAGTGCGGTGAAACGCAAAAAGCTGCGTGCGAAC    1200
EV53_contig000052      GATGATTGTTATCAGATTACTGGATTAAGTGCGGTGAAACGCAAAAAGCTGCGTGCGAAC    1200
                       ************************************************************

EV48_contig000009      TTGCCACTGACTTTGATACCAATTGCCGGTACAACGGACTGGAAGCAAGTGACCGAAAAT    1260
EV55_contig000014      TTGCCACTGACTTTGATACCAATTGCCGGTACAACGGACTGGAAGCAAGTGACCGAAAAT    1260
EV54_contig000003      TTGCCACTGACTTTGATACCAATTGCCGGTACAACGGACTGGAAGCAAGTGACCGAAAAT    1260
EV53_contig000052      TTGCCACTGACTTTGATACCAATTGCCGGTACAACGGACTGGAAGCAAGTGACCGAAAAT    1260
                       ************************************************************

EV48_contig000009      TCAATAACATCTCCCATAAATGTTGCAATTTGAAAAAGAATGCTCAGGTACAGTATACTG    1320
EV55_contig000014      TCAATAACATCTCCCATAAATGTTGCAATTTGAAAAAGAATGCTCAGGTACAGTATACTG    1320
EV54_contig000003      TCAATAACATCTCCCATAAATGTTGCAATTTGAAAAAGAATGCTCAGGTACAGTATACTG    1320
EV53_contig000052      TCAATAACATCTCCCATAAATGTTGCAATTTGAAAAAGAATGCTCAGGTACAGTATACTG    1320
                       ************************************************************

EV48_contig000009      TATCTGAGCGTCTTTTTTTTGATGCTAAGTTCCCATGGGATATCCATGGCTAAGTGGATGA    1380
EV55_contig000014      TATCTGAGCGTCTTTTTTTTGATGCTAAGTTCCCATGGGATATCCATGGCTAAGTGGATGA    1380
EV54_contig000003      TATCTGAGCGTCTTTTTTTTGATGCTAAGTTCCCATGGGATATCCATGGCTAAGTGGATGA    1380
EV53_contig000052      TATCTGAGCGTCTTTTTTTTGATGCTAAGTTCCCATGGGATATCCATGGCTAAGTGGATGA    1380
                       ************************************************************

EV48_contig000009      GCTAATTGAAATGAGTTAGCGAGACGTGGCAGTTGGTTGACTGACACGATGAGGAGTGAA    1440
EV55_contig000014      GCTAATTGAAATGAGTTAGCGAGACGTGGCAGTTGGTTGACTGACACGATGAGGAGTGAA    1440
EV54_contig000003      GCTAATTGAAATGAGTTAGCGAGACGTGGCAGTTGGTTGACTGACACGATGAGGAGTGAA    1440
EV53_contig000052      GCTAATTGAAATGAGTTAGCGAGACGTGGCAGTTGGTTGACTGACACGATGAGGAGTGAA    1440
                       ************************************************************

EV48_contig000009      ATTATGAAAATTGCGGTGGTGACCGAAAGTTCTGCTAATTTAAAAGCGCAGGCGGTGAAA    1500
EV55_contig000014      ATTATGAAAATTGCGGTGGTGACCGAAAGTTCTGCTAATTTAAAAGCGCAGGCGGTGAAA    1500
EV54_contig000003      ATTATGAAAATTGCGGTGGTGACCGAAAGTTCTGCTAATTTAAAAGCGCAGGCGGTGAAA    1500
EV53_contig000052      ATTATGAAAATTGCGGTGGTGACCGAAAGTTCTGCTAATTTAAAAGCGCAGGCGGTGAAA    1500
                       ************************************************************

EV48_contig000009      GACTACCAAATTACGGTGGTGAACGACCCGATTATGTTCGGTAATCATGTTTACCACGAG    1560
EV55_contig000014      GACTACCAAATTACGGTGGTGAACGACCCGATTATGTTCGGTAATCATGTTTACCACGAG    1560
EV54_contig000003      GACTACCAAATTACGGTGGTGAACGACCCGATTATGTTCGGTAATCATGTTTACCACGAG    1560
EV53_contig000052      GACTACCAAATTACGGTGGTGAACGACCCGATTATGTTCGGTAATCATGTTTACCACGAG    1560
                       ************************************************************

EV48_contig000009      AACGTAGATATCACGACGGATCAATTTTACCGATTACTAAAAGTTGAAAAAGAAATTCCA    1620
EV55_contig000014      AACGTAGATATCACGACGGATCAATTTTACCGATTACTAAAAGTTGAAAAAGAAATTCCA    1620
EV54_contig000003      AACGTAGATATCACGACGGATCAATTTTACCGATTACTAAAAGTTGAAAAAGAAATTCCA    1620
EV53_contig000052      AACGTAGATATCACGACGGATCAATTTTACCGATTACTAAAAGTTGAAAAAGAAATTCCA    1620
                       ************************************************************

EV48_contig000009      ACGATTTCTCAAATTTCAATGCCAGAGATGCAAGTCGTTTTTGATCAGCTGCAAGCGGCG    1680
EV55_contig000014      ACGATTTCTCAAATTTCAATGCCAGAGATGCAAGTCGTTTTTGATCAGCTGCAAGCGGCG    1680
EV54_contig000003      ACGATTTCTCAAATTTCAATGCCAGAGATGCAAGTCGTTTTTGATCAGCTGCAAGCGGCG    1680
EV53_contig000052      ACGATTTCTCAAATTTCAATGCCAGAGATGCAAGTCGTTTTTGATCAGCTGCAAGCGGCG    1680
                       ************************************************************

EV48_contig000009      GGCTATGACCAAATTCTAGTTATTGGGTTGAGTAGTGGTATCAGTGGCTGGATCAATAAT    1740
EV55_contig000014      GGCTATGACCAAATTCTAGTTATTGGGTTGAGTAGTGGTATCAGTGGCTGGATCAATAAT    1740
EV54_contig000003      GGCTATGACCAAATTCTAGTTATTGGGTTGAGTAGTGGTATCAGTGGCTGGATCAATAAT    1740
EV53_contig000052      GGCTATGACCAAATTCTAGTTATTGGGTTGAGTAGTGGTATCAGTGGCTGGATCAATAAT    1740
                       ************************************************************
```

```
EV48_contig000009      TTGAAGACCTATGCGCCGTCAGTTGACGGTATTGACGTGCGGGTGTTTGATTCACGAACT    1800
EV55_contig000014      TTGAAGACCTATGCGCCGTCAGTTGACGGTATTGACGTGCGGGTGTTTGATTCACGAACT    1800
EV54_contig000003      TTGAAGACCTATGCGCCGTCAGTTGACGGTATTGACGTGCGGGTGTTTGATTCACGAACT    1800
EV53_contig000052      TTGAAGACCTATGCGCCGTCAGTTGACGGTATTGACGTGCGGGTGTTTGATTCACGAACT    1800
                       ************************************************************

EV48_contig000009      GCTTGCGCGGGGACGGCGAACATGGTGAAGCTGGCTGCGGCCATGGCATTGGCGGACTAT    1860
EV55_contig000014      GCTTGCGCGGGGACGGCGAACATGGTGAAGCTGGCTGCGGCCATGGCATTGGCGGACTAT    1860
EV54_contig000003      GCTTGCGCGGGGACGGCGAACATGGTGAAGCTGGCTGCGGCCATGGCATTGGCGGACTAT    1860
EV53_contig000052      GCTTGCGCGGGGACGGCGAACATGGTGAAGCTGGCTGCGGCCATGGCATTGGCGGACTAT    1860
                       ************************************************************

EV48_contig000009      ACCTTAGATGAAATTATTGCGCAACTGACGCGCTTGCGGGCGGCTACGACGACGGTCTTA    1920
EV55_contig000014      ACCTTAGATGAAATTATTGCGCAACTGACGCGCTTGCGGGCGGCTACGACGACGGTCTTA    1920
EV54_contig000003      ACCTTAGATGAAATTATTGCGCAACTGACGCGCTTGCGGGCGGCTACGACGACGGTCTTA    1920
EV53_contig000052      ACCTTAGATGAAATTATTGCGCAACTGACGCGCTTGCGGGCGGCTACGACGACGGTCTTA    1920
                       ************************************************************

EV48_contig000009      ATCGTTAACTCGATGCGCCATCTAGTTAAAAATGGCCGGATTTATAACAATAGTAGTCTG    1980
EV55_contig000014      ATCGTTAACTCGATGCGCCATCTAGTTAAAAATGGCCGGATTTATAACAATAGTAGTCTG    1980
EV54_contig000003      ATCGTTAACTCGATGCGCCATCTAGTTAAAAATGGCCGGATTTATAACAATAGTAGTCTG    1980
EV53_contig000052      ATCGTTAACTCGATGCGCCATCTAGTTAAAAATGGCCGGATTTATAACAATAGTAGTCTG    1980
                       ************************************************************

EV48_contig000009      GCCGGAACGATGGTGTTGCGCAATAAGCCGATTATTACGTTCAATGGGCATGGGAAGATT    2040
EV55_contig000014      GCCGGAACGATGGTGTTGCGCAATAAGCCGATTATTACGTTCAATGGGCATGGGAAGATT    2040
EV54_contig000003      GCCGGAACGATGGTGTTGCGCAATAAGCCGATTATTACGTTCAATGGGCATGGGAAGATT    2040
EV53_contig000052      GCCGGAACGATGGTGTTGCGCAATAAGCCGATTATTACGTTCAATGGGCATGGGAAGATT    2040
                       ************************************************************

EV48_contig000009      CAACTTTTAGGTAAGGAACGGACGCTTAAAAATGCACAACAGGCTGTTCAAGATGAATTT    2100
EV55_contig000014      CAACTTTTAGGTAAGGAACGGACGCTTAAAAATGCACAACAGGCTGTTCAAGATGAATTT    2100
EV54_contig000003      CAACTTTTAGGTAAGGAACGGACGCTTAAAAATGCACAACAGGCTGTTCAAGATGAATTT    2100
EV53_contig000052      CAACTTTTAGGTAAGGAACGGACGCTTAAAAATGCACAACAGGCTGTTCAAGATGAATTT    2100
                       ************************************************************

EV48_contig000009      AGTCAGTTTGTGATGAACAGTCAATTACCAGTTCGGCTGGACGTCATCAGCGCTAATGAC    2160
EV55_contig000014      AGTCAGTTTGTGATGAACAGTCAATTACCAGTTCGGCTGGACGTCATCAGCGCTAATGAC    2160
EV54_contig000003      AGTCAGTTTGTGATGAACAGTCAATTACCAGTTCGGCTGGACGTCATCAGCGCTAATGAC    2160
EV53_contig000052      AGTCAGTTTGTGATGAACAGTCAATTACCAGTTCGGCTGGACGTCATCAGCGCTAATGAC    2160
                       ************************************************************

EV48_contig000009      GCGCAGATTGCCGAAGATTGGGCGAGCGAACTACGTTATCAGTTTCCAGCAGTCCGGGTC    2220
EV55_contig000014      GCGCAGATTGCCGAAGATTGGGCGAGCGAACTACGTTATCAGTTTCCAGCAGTCCGGGTC    2220
EV54_contig000003      GCGCAGATTGCCGAAGATTGGGCGAGCGAACTACGTTATCAGTTTCCAGCAGTCCGGGTC    2220
EV53_contig000052      GCGCAGATTGCCGAAGATTGGGCGAGCGAACTACGTTATCAGTTTCCAGCAGTCCGGGTC    2220
                       ************************************************************

EV48_contig000009      AAGACTGGGGTTATCGGTCCAGTTTTAGGCGTCTTGACTGGAGACCATGCGCTCGGAATG    2280
EV55_contig000014      AAGACTGGGGTTATCGGTCCAGTTTTAGGCGTCTTGACTGGAGACCATGCGCTCGGAATG    2280
EV54_contig000003      AAGACTGGGGTTATCGGTCCAGTTTTAGGCGTCTTGACTGGAGACCATGCGCTCGGAATG    2280
EV53_contig000052      AAGACTGGGGTTATCGGTCCAGTTTTAGGCGTCTTGACTGGAGACCATGCGCTCGGAATG    2280
                       ************************************************************

EV48_contig000009      ATCTGGTCATTAGATTGGAAGCCATTGTTGACGACCCTCAATCAGGAGCAAAAACAATAG    2340
EV55_contig000014      ATCTGGTCATTAGATTGGAAGCCATTGTTGACGACCCTCAATCAGGAGCAAAAACAATAG    2340
EV54_contig000003      ATCTGGTCATTAGATTGGAAGCCATTGTTGACGACCCTCAATCAGGAGCAAAAACAATAG    2340
EV53_contig000052      ATCTGGTCATTAGATTGGAAGCCATTGTTGACGACCCTCAATCAGGAGCAAAAACAATAG    2340
                       ************************************************************

EV48_contig000009      GCCGTTGGGGTTGCCCGTGTGCGGTTTGCTGATGATACTGTCGTATACAGGCAGCTGGGT    2400
EV55_contig000014      GCCGTTGGGGTTGCCCGTGTGCGGTTTGCTGATGATACTGTCGTATACAGGCAGCTGGGT    2400
EV54_contig000003      GCCGTTGGGGTTGCCCGTGTGCGGTTTGCTGATGATACTGTCGTATACAGGCAGCTGGGT    2400
EV53_contig000052      GCCGTTGGGGTTGCCCGTGTGCGGTTTGCTGATGATACTGTCGTATACAGGCAGCTGGGT    2400
                       ************************************************************

EV48_contig000009      TTTAAAAAAGTTTTGCATCGAAGGTTGACATTTTACTAAAATGATATTATGGTTATAACA    2460
EV55_contig000014      TTTAAAAAAGTTTTGCATCGAAGGTTGACATTTTACTAAAATGATATTATGGTTATAACA    2460
EV54_contig000003      TTTAAAAAAGTTTTGCATCGAAGGTTGACATTTTACTAAAATGATATTATGGTTATAACA    2460
EV53_contig000052      TTTAAAAAAGTTTTGCATCGAAGGTTGACATTTTACTAAAATGATATTATGGTTATAACA    2460
                       ************************************************************

EV48_contig000009      TTAAAACTTGATAAAAAGCTATAATGCTTTAAGGAGTAACTGACTTGTGATGCTCAGAAA    2520
EV55_contig000014      TTAAAACTTGATAAAAAGCTATAATGCTTTAAGGAGTAACTGACTTGTGATGCTCAGAAA    2520
EV54_contig000003      TTAAAACTTGATAAAAAGCTATAATGCTTTAAGGAGTAACTGACTTGTGATGCTCAGAAA    2520
EV53_contig000052      TTAAAACTTGATAAAAAGCTATAATGCTTTAAGGAGTAACTGACTTGTGATGCTCAGAAA    2520
                       ************************************************************

EV48_contig000009      GCGCCGTTGCTGTGAAGGTGTGATGACAAGTGAGTAAAAGTTAAAGGATTGGCGTGAATT    2580
EV55_contig000014      GCGCCGTTGCTGTGAAGGTGTGATGACAAGTGAGTAAAAGTTAAAGGATTGGCGTGAATT    2580
EV54_contig000003      GCGCCGTTGCTGTGAAGGTGTGATGACAAGTGAGTAAAAGTTAAAGGATTGGCGTGAATT    2580
EV53_contig000052      GCGCCGTTGCTGTGAAGGTGTGATGACAAGTGAGTAAAAGTTAAAGGATTGGCGTGAATT    2580
                       ************************************************************
```

```
EV48_contig000009      TTACGAGATTCACTGGTAGCACCCATTATTGTGCACGCGTTTAGTTGTTGTGAACGCGAT    2640
EV55_contig000014      TTACGAGATTCACTGGTAGCACCCATTATTGTGCACGCGTTTAGTTGTTGTGAACGCGAT    2640
EV54_contig000003      TTACGAGATTCACTGGTAGCACCCATTATTGTGCACGCGTTTAGTTGTTGTGAACGCGAT    2640
EV53_contig000052      TTACGAGATTCACTGGTAGCACCCATTATTGTGCACGCGTTTAGTTGTTGTGAACGCGAT    2640
                       ************************************************************

EV48_contig000009      GAGTAGCCGGATGTGAGTTCGGTTAGCAATAAGGTGGTACCGCGCGTAAGCGCCCTTTGG    2700
EV55_contig000014      GAGTAGCCGGATGTGAGTTCGGTTAGCAATAAGGTGGTACCGCGCGTAAGCGCCCTTTGG    2700
EV54_contig000003      GAGTAGCCGGATGTGAGTTCGGTTAGCAATAAGGTGGTACCGCGCGTAAGCGCCCTTTGG    2700
EV53_contig000052      GAGTAGCCGGATGTGAGTTCGGTTAGCAATAAGGTGGTACCGCGCGTAAGCGCCCTTTGG    2700
                       ************************************************************

EV48_contig000009      ATAATTTGATCGATTATTCAAAGGGCGCTTTTTCGTCAAGTGACACATAGCATTATTTTT    2760
EV55_contig000014      ATAATTTGATCGATTATTCAAAGGGCGCTTTTTCGTCAAGTGACACATAGCATTATTTTT    2760
EV54_contig000003      ATAATTTGATCGATTATTCAAAGGGCGCTTTTTCGTCAAGTGACACATAGCATTATTTTT    2760
EV53_contig000052      ATAATTTGATCGATTATTCAAAGGGCGCTTTTTCGTCAAGTGACACATAGCATTATTTTT    2760
                       ************************************************************

EV48_contig000009      TAGTTAACACATAGTAGTCATTATTTAGGTCGCGCATCGCGGTTAATCATCGTAAACACA    2820
EV55_contig000014      TAGTTAACACATAGTAGTCATTATTTAGGTCGCGCATCGCGGTTAATCATCGTAAACACA    2820
EV54_contig000003      TAGTTAACACATAGTAGTCATTATTTAGGTCGCGCATCGCGGTTAATCATCGTAAACACA    2820
EV53_contig000052      TAGTTAACACATAGTAGTCATTATTTAGGTCGCGCATCGCGGTTAATCATCGTAAACACA    2820
                       ************************************************************

EV48_contig000009      TAGTTGTTCGACATCACGTGACAGCACATAGTAGATCGTAAGCACGGGCCCGTTTATTGG    2880
EV55_contig000014      TAGTTGTTCGACATCACGTGACAGCACATAGTAGATCGTAAGCACGGGCCCGTTTATTGG    2880
EV54_contig000003      TAGTTGTTCGACATCACGTGACAGCACATAGTAGATCGTAAGCACGGGCCCGTTTATTGG    2880
EV53_contig000052      TAGTTGTTCGACATCACGTGACAGCACATAGTAGATCGTAAGCACGGGCCCGTTTATTGG    2880
                       ************************************************************

EV48_contig000009      AAAAGCACTGAAGCGAGCGGGGAACATATACACATAGTGGGCAAGACGAAAGGATGGTTT    2940
EV55_contig000014      AAAAGCACTGAAGCGA----------------------------------------------    2896
EV54_contig000003      AAAAGCACTGAAGCGAGCGGGGAACATATACACATAGTGGGCAAGACGAAAGGATGGTTT    2940
EV53_contig000052      AAAAGCACTGAAGCGAGCGGGGAACATATACACATAGTGGGCAAGACGAAAGGATGGTTT    2940
                       ****************

EV48_contig000009      TACTTATGAAATATGGCATTATTGGCGCAGGCGCAATGGGATATCGGTATGGTGTAATGT    3000
EV55_contig000014      ------------------------------------------------------------    2896
EV54_contig000003      TACTTATGAAATATGGCATTATTGGCGCAGGCGCAATGGGATATCGGTATGGTGTAATGT    3000
EV53_contig000052      TACTTATGAAATATGGCATTATTGGCGCAGGCGCAATGGGATATCGGTATGGTGTAATGT    3000

EV48_contig000009      TACAAGAAAACGCCGGTGTTGACGTTGATTTTATCGACACGTGGGAACCAAACGTCGCAA    3060
EV55_contig000014      ------------------------------------------------------------    2896
EV54_contig000003      TACAAGAAAACGCCGGTGTTGACGTTGATTTTATCGACACGTGGGAACCAAACGTCGCAA    3060
EV53_contig000052      TACAAGAAAACGCCGGTGTTGACGTTGATTTTATCGACACGTGGGAACCAAACGTCGCAA    3060

EV48_contig000009      AGGTCCGTGAGCAGGGCGGCGTTGATGTGGCGCGTGATCATCAAAATCACCATGTCGTTC    3120
EV55_contig000014      ------------------------------------------------------------    2896
EV54_contig000003      AGGTCCGTGAGCAGGGCGGCGTTGATGTGGCGCGTGATCATCAAAATCACCATGTCGTTC    3120
EV53_contig000052      AGGTCCGTGAGCAGGGCGGCGTTGATGTGGCGCGTGATCATCAAAATCACCATGTCGTTC    3120

EV48_contig000009      CAATTAATATTTATTATCCTGAAGAATATCAGGGTCATCCAGATGTTTGGATTGTTTTTA    3180
EV55_contig000014      ------------------------------------------------------------    2896
EV54_contig000003      CAATTAATATTTATTATCCTGAAGAATATCAGGGTCATCCAGATGTTTGGATTGTTTTTA    3180
EV53_contig000052      CAATTAATATTTATTATCCTGAAGAATATCAGGGTCATCCAGATGTTTGGATTGTTTTTA    3180

EV48_contig000009      AGAAGCAAATGCAGTTAGCCGATGAGCTCAAGCGCGATGCGTCGTTATTTCATGAGGACC    3240
EV55_contig000014      ------------------------------------------------------------    2896
EV54_contig000003      AGAAGCAAATGCAGTTAGCCGATGAGCTCAAGCGCGATGCGTCGTTATTTCATGAGGACC    3240
EV53_contig000052      AGAAGCAAATGCAGTTAGCCGATGAGCTCAAGCGCGATGCGTCGTTATTTCATGAGGACC    3240

EV48_contig000009      AGTACGTTTTCGCCGCAATGAACGGGATGGGGCACTTTGAAAAGATTGCACAATATTTCC    3300
EV55_contig000014      ------------------------------------------------------------    2896
EV54_contig000003      AGTACGTTTTCGCCGCAATGAACGGGATGGGGCACTTTGAAAAGATTGCACAATATTTCC    3300
EV53_contig000052      AGTACGTTTTCGCCGCAATGAACGGGATGGGGCACTTTGAAAAGATTGCACAATATTTCC    3300

EV48_contig000009      CGGAAAATCACATTATTGGTGGTACGGCGATGATTGCAACGGTCTTGAATGGTCCGGGTG    3360
EV55_contig000014      ------------------------------------------------------------    2896
EV54_contig000003      CGGAAAATCACATTATTGGTGGTACGGCGATGATTGCAACGGTCTTGAATGGTCCGGGTG    3360
EV53_contig000052      CGGAAAATCACATTATTGGTGGTACGGCGATGATTGCAACGGTCTTGAATGGTCCGGGTG    3360

EV48_contig000009      CGGTTGATTTTATGGGGCCTAAGGGTAGTGAGGCGATGCATATGAGTAAGTATGCAGGTC    3420
EV55_contig000014      ------------------------------------------------------------    2896
EV54_contig000003      CGGTTGATTTTATGGGGCCTAAGGGTAGTGAGGCGATGCATATGAGTAAGTATGCAGGTC    3420
EV53_contig000052      CGGTTGATTTTATGGGGCCTAAGGGTAGTGAGGCGATGCATATGAGTAAGTATGCAGGTC    3420
```

```
EV48_contig000009      CCATTGATACGACTACTAAGCAAGTGATGGCCGATTTTAAGGCGGCTGATTTAAATCCGA    3480
EV55_contig000014      ------------------------------------------------------------    2896
EV54_contig000003      CCATTGATACGACTACTAAGCAAGTGATGGCCGATTTTAAGGCGGCTGATTTAAATCCGA    3480
EV53_contig000052      CCATTGATACGACTACTAAGCAAGTGATGGCCGATTTTAAGGCGGCTGATTTAAATCCGA    3480


EV48_contig000009      TTTGGTCTGATAATTTTATGGGCATGTGCATTTCCAAAGTTGTGTTTAACGCGGTCACCA    3540
EV55_contig000014      ------------------------------------------------------------    2896
EV54_contig000003      TTTGGTCTGATAATTTTATGGGCATGTGCATTTCCAAAGTTGTGTTTAACGCGGTCACCA    3540
EV53_contig000052      TTTGGTCTGATAATTTTATGGGCATGTGCATTTCCAAAGTTGTGTTTAACGCGGTCACCA    3540


EV48_contig000009      ATAGTTTGTGCACGATGTTCGAGATTCAGATGGGCCAATTCATCGAATATCCGGGCGTCA    3600
EV55_contig000014      ------------------------------------------------------------    2896
EV54_contig000003      ATAGTTTGTGCACGATGTTCGAGATTCAGATGGGCCAATTCATCGAATATCCGGGCGTCA    3600
EV53_contig000052      ATAGTTTGTGCACGATGTTCGAGATTCAGATGGGCCAATTCATCGAATATCCGGGCGTCA    3600


EV48_contig000009      AGGATATGGCGACTCAAATGTTCAACGAAGCCTATGATGCTTGCGAAAGGGCAGGCATCA    3660
EV55_contig000014      ------------------------------------------------------------    2896
EV54_contig000003      AGGATATGGCGACTCAAATGTTCAACGAAGCCTATGATGCTTGCGAAAGGGCAGGCATCA    3660
EV53_contig000052      AGGATATGGCGACTCAAATGTTCAACGAAGCCTATGATGCTTGCGAAAGGGCAGGCATCA    3660


EV48_contig000009      AACTCATTGAGACGCGGCAAGAAGAAATCGACTCGGTTGAAACGGTCAGTCGAGCCTATA    3720
EV55_contig000014      ------------------------------------------------------------    2896
EV54_contig000003      AACTCATTGAGACGCGGCAAGAAGAAATCGACTCGGTTGAAACGGTCAGTCGAGCCTATA    3720
EV53_contig000052      AACTCATTGAGACGCGGCAAGAAGAAATCGACTCGGTTGAAACGGTCAGTCGAGCCTATA    3720


EV48_contig000009      AGTACCATTATCCATCAATGTATCAAGACTTCTCAAAGGGACGGCCAACCGAAGTTGATT    3780
EV55_contig000014      ------------------------------------------------------------    2896
EV54_contig000003      AGTACCATTATCCATCAATGTATCAAGACTTCTCAAAGGGACGGCCAACCGAAGTTGATT    3780
EV53_contig000052      AGTACCATTATCCATCAATGTATCAAGACTTCTCAAAGGGACGGCCAACCGAAGTTGATT    3780


EV48_contig000009      ATATTAATGGCTATATTGCTAGGATTGGGCGGGAACATGACTACGTTTGTCGCGTGCATG    3840
EV55_contig000014      ------------------------------------------------------------    2896
EV54_contig000003      ATATTAATGGCTATATTGCTAGGATTGGGCGGGAACATGACTACGTTTGTCGCGTGCATG    3840
EV53_contig000052      ATATTAATGGCTATATTGCTAGGATTGGGCGGGAACATGACTACGTTTGTCGCGTGCATG    3840


EV48_contig000009      AGTTTGTGACTCACGAAGTCCATTTAGCAGAAATGATGCGGCAATACCGGCACCCGGAAA    3900
EV55_contig000014      ------------------------------------------------------------    2896
EV54_contig000003      AGTTTGTGACTCACGAAGTCCATTTAGCAGAAATGATGCGGCAATACCGGCACCCGGAAA    3900
EV53_contig000052      AGTTTGTGACTCACGAAGTCCATTTAGCAGAAATGATGCGGCAATACCGGCACCCGGAAA    3900


EV48_contig000009      TCCCCGTTGCTGAAAAATAAGATTATTAATGAAAAAGTTCTGAGTATTTTCAGAACTTTT    3960
EV55_contig000014      ------------------------------------------------------------    2896
EV54_contig000003      TCCCCGTTGCTGAAAAATAAGATTATTAATGAAAAAGTTCTGAGTATTTTCAGAACTTTT    3960
EV53_contig000052      TCCCCGTTGCTGAAAAATAAGATTATTAATGAAAAAGTTCTGAGTATTTTCAGAACTTTT    3960


EV48_contig000009      TTTGTGGGCGGAAAAACTTAATTTAGTAAGGTTGAAATCGGTTTCTTTGACATTGGTCTA    4020
EV55_contig000014      ------------------------------------------------------------    2896
EV54_contig000003      TTTGTGGGCGGAAAAACTTAATTTAGTAAGGTTGAAATCGGTTTCTTTGACATTGGTCTA    4020
EV53_contig000052      TTTGTGGGCGGAAAAACTTAATTTAGTAAGGTTGAAATCGGTTTCTTTGACATTGGTCTA    4020


EV48_contig000009      TACACTTAGATAACCTTGTAAACTTGCAATGAAAACATGTTGTGATAGGATTTATAACGT    4080
EV55_contig000014      ------------------------------------------------------------    2896
EV54_contig000003      TACACTTAGATAACCTTGTAAACTTGCAATGAAAACATGTTGTGATAGGATTTATAACGT    4080
EV53_contig000052      TACACTTAGATAACCTTGTAAACTTGCAATGAAAACATGTTGTGATAGGATTTATAACGT    4080


EV48_contig000009      GCTTAAAAAGAGAATGAGGAGAAATTGGTATGAAGACATATTTTCAGAAAATCGGTCAGT    4140
EV55_contig000014      ------------------------------------------------------------    2896
EV54_contig000003      GCTTAAAAAGAGAATGAGGAGAAATTGGTATGAAGACATATTTTCAGAAAATCGGTCAGT    4140
EV53_contig000052      GCTTAAAAAGAGAATGAGGAGAAATTGGTATGAAGACATATTTTCAGAAAATCGGTCAGT    4140


EV48_contig000009      CATTGATGCTGCCGATTGCAACGTTGCCCGCAGCAGCGATTTTGGTTGGGATTGGTAATT    4200
EV55_contig000014      ----------------------------------------TTGGTTGGGATTGGTAATT    2915
EV54_contig000003      CATTGATGCTGCCGATTGCAACGTTGCCCGCAGCAGCGATTTTGGTTGGGATTGGTAATT    4200
EV53_contig000052      CATTGATGCTGCCGATTGCAACGTTGCCCGCAGCAGCGATTTTGGTTGGGATTGGTAATT    4200
                                                       ******************


EV48_contig000009      ACCTGCCAAAACAGTGGCTGTTTGCAAATTACTTGATTCAGGGTGGTAACGTTGTCCTGA    4260
EV55_contig000014      ACCTGCCAAAACAGTGGCTGTTTGCAAATTACTTGATTCAGGGTGGTAACGTTGTCCTGA    2975
EV54_contig000003      ACCTGCCAAAACAGTGGCTGTTTGCAAATTACTTGATTCAGGGTGGTAACGTTGTCCTGA    4260
EV53_contig000052      ACCTGCCAAAACAGTGGCTGTTTGCAAATTACTTGATTCAGGGTGGTAACGTTGTCCTGA    4260
                       ************************************************************
```

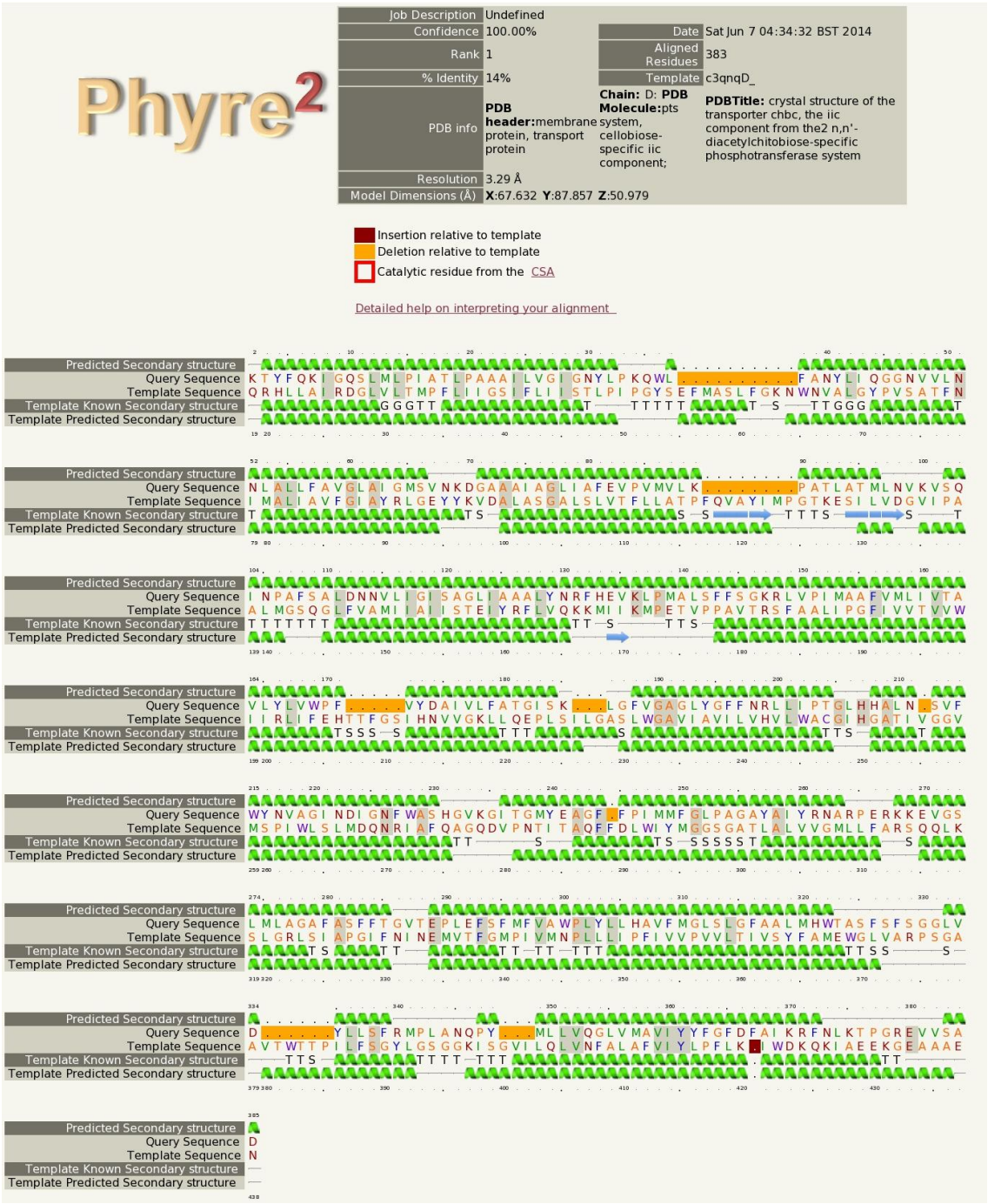
```
EV48_contig000009      ATAACTTAGCCTTGCTGTTTGCAGTCGGCTTAGCTATCGGGATGTCAGTCAATAAGGACG    4320
EV55_contig000014      ATAACTTAGCCTTGCTGTTTGCAGTCGGCTTAGCTATCGGGATGTCAGTCAATAAGGACG    3035
EV54_contig000003      ATAACTTAGCCTTGCTGTTTGCAGTCGGCTTAGCTATCGGGATGTCAGTCAATAAGGACG    4320
EV53_contig000052      ATAACTTAGCCTTGCTGTTTGCAGTCGGCTTAGCTATCGGGATGTCAGTCAATAAGGACG    4320
                       ************************************************************

EV48_contig000009      GCGCGGCAGCTATTGCTGGTTTGATCGCGTTTGAAGTTCCGGTGATGGTGTTAAAACCAG    4380
EV55_contig000014      GCGCGGCAGCTATTGCTGGTTTGATCGCGTTTGAAGTTCCGGTGATGGTGTTAAAACCAG    3095
EV54_contig000003      GCGCGGCAGCTATTGCTGGTTTGATCGCGTTTGAAGTTCCGGTGATGGTGTTAAAACCAG    4380
EV53_contig000052      GCGCGGCAGCTATTGCTGGTTTGATCGCGTTTGAAGTTCCGGTGATGGTGTTAAAACCAG    4380
                       ************************************************************

EV48_contig000009      CGACATTAGCGACGATGCTGAATGTTAAAGTGAGTCAGATCAATCCAGCATTTAGCGCGT    4440
EV55_contig000014      CGACATTAGCGACGATGCTGAATGTTAAAGTGAGTCAGATCAATCCAGCATTTAGCGCGT    3155
EV54_contig000003      CGACATTAGCGACGATGCTGAATGTTAAAGTGAGTCAGATCAATCCAGCATTTAGCGCGT    4440
EV53_contig000052      CGACATTAGCGACGATGCTGAATGTTAAAGTGAGTCAGATCAATCCAGCATTTAGCGCGT    4440
                       ************************************************************

EV48_contig000009      TAGATAACAATGTGCTGATTGGGATCAGTGCTGGACTGATTGCGGCTGCGCTCTATAATC    4500
EV55_contig000014      TAGATAACAATGTGCTGATTGGGATCAGTGCTGGACTGATTGCGGCTGCGCTCTATAATC    3215
EV54_contig000003      TAGATAACAATGTGCTGATTGGGATCAGTGCTGGACTGATTGCGGCTGCGCTCTATAATC    4500
EV53_contig000052      TAGATAACAATGTGCTGATTGGGATCAGTGCTGGACTGATTGCGGCTGCGCTCTATAATC    4500
                       ************************************************************

EV48_contig000009      GGTTCCACGAAGTAAAATTACCAATGGCGCTCTCATTCTTTAGTGGTAAACGCCTAGTGC    4560
EV55_contig000014      GGTTCCACGAAGTAAAATTACCAATGGCGCTCTCATTCTTTAGTGGTAAACGCCTAGTGC    3275
EV54_contig000003      GGTTCCACGAAGTAAAATTACCAATGGCGCTCTCATTCTTTAGTGGTAAACGCCTAGTGC    4560
EV53_contig000052      GGTTCCACGAAGTAAAATTACCAATGGCGCTCTCATTCTTTAGTGGTAAACGCCTAGTGC    4560
                       ************************************************************

EV48_contig000009      CAATTATGGCTGCTTTTGTGATGCTGATTGTGACGGCGGTATTGTACCTTGTTTGGCCGT    4620
EV55_contig000014      CAATTATGGCTGCTTTTGTGATGCTGATTGTGACGGCGGTATTGTACCTTGTTTGGCCGT    3335
EV54_contig000003      CAATTATGGCTGCTTTTGTGATGCTGATTGTGACGGCGGTATTGTACCTTGTTTGGCCGT    4620
EV53_contig000052      CAATTATGGCTGCTTTTGTGATGCTGATTGTGACGGCGGTATTGTACCTTGTTTGGCCGT    4620
                       ************************************************************

EV48_contig000009      TTGTCTACGATGCCATTGTCCTCTTTGCAACGGGGATTTCAAAATTAGGTTTCGTGGGGG    4680
EV55_contig000014      TTGTCTACGATGCCATTGTCCTCTTTGCAACGGGGATTTCAAAATTAGGTTTCGTGGGGG    3395
EV54_contig000003      TTGTCTACGATGCCATTGTCCTCTTTGCAACGGGGATTTCAAAATTAGGTTTCGTGGGGG    4680
EV53_contig000052      TTGTCTACGATGCCATTGTCCTCTTTGCAACGGGGATTTCAAAATTAGGTTTCGTGGGGG    4680
                       ************************************************************

EV48_contig000009      CTGGTCTATATGGCTTTTTCAATCGTTTATTGATTCCAACTGGTTTGCACCACGCCTTAA    4740
EV55_contig000014      CTGGTCTATATGGCTTTTTCAATCGTTTATTGATTCCAACTGGTTTGCACCACGCCTTAA    3455
EV54_contig000003      CTGGTCTATATGGCTTTTTCAATCGTTTATTGATTCCAACTGGTTTGCACCACGCCTTAC    4740
EV53_contig000052      CTGGTCTATATGGCTTTTTCAATCGTTTATTGATTCCAACTGGTTTGCACCACGCCTTAA    4740
                       ************************************************************

EV48_contig000009      ATTCCGTATTCTGGTATAACGTTGCTGGTATCAATGATATTGGCAATTTCTGGGCCAGCC    4800
EV55_contig000014      ATTCCGTATTCTGGTATAACGTTGCTGGTATCAATGATATTGGCAATTTCTGGGCCAGCC    3515
EV54_contig000003      ATTCCGTATTCTGGTATAACGTTGCTGGTATCAATGATATTGGCAATTTCTGGGCCAGCC    4800
EV53_contig000052      ATTCCGTATTCTGGTATAACGTTGCTGGTATCAATGATATTGGCAATTTCTGGGCCAGCC    4800
                       ************************************************************

EV48_contig000009      ATGGTGTTAAAGGAATCACAGGGATGTATGAAGCTGGATTCTTCCCAATTATGATGTTTG    4860
EV55_contig000014      ATGGTGTTAAAGGAATCACAGGGATGTATGAAGCTGGATTCTTCCCAATTATGATGTTTG    3575
EV54_contig000003      ATGGTGTTAAAGGAATCACAGGGATGTATGAAGCTGGATTCTTCCCAATTATGATGTTTG    4860
EV53_contig000052      ATGGTGTTAAAGGAATCACAGGGATGTATGAAGCTGAATTCTTCCCAATTATGATGTTTG    4860
                       ********************************** ***********************

EV48_contig000009      GCTTGCCAGCTGGTGCGTATGCAATTTATCGTAATGCACGACCGGAACGAAAAAAAGAAG    4920
EV55_contig000014      GCTTGCCAGCTGGTGCGTATGCAATTTATCGTAATGCACGACCGGAACGAAAAAAAGAAG    3635
EV54_contig000003      GCTTGCCAGCTGGTGCGTATGCAATTTATCGTAATGCACGACCGGAACGAAAAAAAGAAG    4920
EV53_contig000052      GCTTGCCAGCTGGTGCGTATGCAATTTATCGTAATGCACGACCGGAACGAAAAAAAGAAG    4920
                       ************************************************************

EV48_contig000009      TGGGTTCATTGATGTTAGCGGGTGCGTTTGCGTCGTTCTTTACGGGGGTGACCGAACCGC    4980
EV55_contig000014      TGGGTTCATTGATGTTAGCGGGTGCGTTTGCGTCGTTCTTTACGGGGGTGACCGAACCGC    3695
EV54_contig000003      TGGGTTCATTGATGTTAGCGGGTGCGTTTGCGTCGTTCTTTACGGGGGTGACCGAACCGC    4980
EV53_contig000052      TGGGTTCATTGATGTTAGCGGGTGCGTTTGCGTCGTTCTTTACGGGGGTGACCGAACCGC    4980
                       ************************************************************

EV48_contig000009      TTGAATTTTCATTCATGTTCGTGGCATGGCCGCTCTACTTATTACATGCCGTTTTTATGG    5040
EV55_contig000014      TTGAATTTTCATTCATGTTCGTGGCATGGCCGCTCTACTTATTACATGCCGTTTTTATGG    3755
EV54_contig000003      TTGAATTTTCATTCATGTTCGTGGCATGGCCGCTCTACTTATTACATGCCGTTTTTATGG    5040
EV53_contig000052      TTGAATTTTCATTCATGTTCGTGGCATGGCCGCTCTACTTATTACATGCCGTTTTTATGG    5040
                       ************************************************************

EV48_contig000009      GACTATCATTAGGATTCGCCGCGTTAATGCACTGGACCGCCAGCTTTTCGTTTAGTGGCG    5100
EV55_contig000014      GACTATCATTAGGATTCGCCGCGTTAATGCACTGGACCGCCAGCTTTTCGTTTAGTGGCG    3815
EV54_contig000003      GACTATCATTAGGATTCGCCGCGTTAATGCACTGGACCGCCAGCTTTTCGTTTAGTGGCG    5100
EV53_contig000052      GACTATCATTAGGATTCGCCGCGTTAATGCACTGGACCGCCAGCTTTTCGTTTAGTGGCG    5100
                       ************************************************************
```

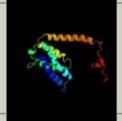
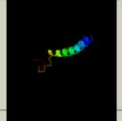
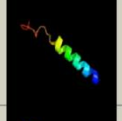
```
EV48_contig000009      GGTTAGTCGATTATTTATTGAGTTTTCGGATGCCGTTAGCCAACCAACCCTATATGTTAC    5160
EV55_contig000014      GGTTAGTCGATTATTTATTGAGTTTTCGGATGCCGTTAGCCAACCAACCCTATATGTTAC    3875
EV54_contig000003      GGTTAGTCGATTATTTATTGAGTTTTCGGATGCCGTTAGCCAACCAACCCTATATGTTAC    5160
EV53_contig000052      GGTTAGTCGATTATTTATTGAGTTTTCGGATGCCGTTAGCCAACCAACCCTATATGTTAC    5160
                       ************************************************************

EV48_contig000009      TGGTTCAAGGGCTGGTGATGGCCGTCATTTATTACTTTGGTTTTGACTTTGCAATCAAGC    5220
EV55_contig000014      TGGTTCAAGGGCTGGTGATGGCCGTCATTTATTACTTTGGTTTTGACTTTGCAATCAAGC    3935
EV54_contig000003      TGGTTCAAGGGCTGGTGATGGCCGTCATTTATTACTTTGGTTTTGACTTTGCAATCAAGC    5220
EV53_contig000052      TGGTTCAAGGGCTGGTGATGGCCGTCATTTATTACTTTGGTTTTGACTTTGCAATCAAGC    5220
                       ************************************************************

EV48_contig000009      GGTTTAATTTGAAGACGCCTGGTCGTGAAGTTGTGTCTGCAGATGTCGATGGGGTGGGTG    5280
EV55_contig000014      GGTTTAATTTGAAGACGCCTGGTCGTGAAGTTGTGTCTGCAGATGTCGATGGGGTGGGTG    3995
EV54_contig000003      GGTTTAATTTGAAGACGCCTGGTCGTGAAGTTGTGTCTGCAGATGTCGATGGGGTGGGTG    5280
EV53_contig000052      GGTTTAATTTGAAGACGCCTGGTCGTGAAGTTGTGTCTGCAGATGTCGATGGGGTGGGTG    5280
                       ************************************************************

EV48_contig000009      CACCTGCTAGTCCAGCGGTGGCCGTTGCAGCAACGGATGATAAATATATGCGCCAAGCAA    5340
EV55_contig000014      CACCTGCTAGTCCAGCGGTGGCCGTTGCAGCAACGGATGATAAATATATGCGCCAAGCAA    4055
EV54_contig000003      CACCTGCTAGTCCAGCGGTGGCCGTTGCAGCAACGGATGATAAATATATGCGCCAAGCAA    5340
EV53_contig000052      CACCTGCTAGTCCAGCGGTGGCCGTTGCAGCAACGGATGATAAATATATGCGCCAAGCAA    5340
                       ************************************************************

EV48_contig000009      AGCAAATTTATGCGGCTATTGGTGGTCATGACAATATTAGTGTGATCAACAATTGTACGA    5400
EV55_contig000014      AGCAAATTTATGCGGCTATTGGTGGTCATGACAATATTAGTGTGATCAACAATTGTACGA    4115
EV54_contig000003      AGCAAATTTATGCGGCTATTGGTGGTCATGACAATATTAGTGTGATCAACAATTGTACGA    5400
EV53_contig000052      AGCAAATTTATGCGGCTATTGGTGGTCATGACAATATTAGTGTGATCAACAATTGTACGA    5400
                       ************************************************************

EV48_contig000009      CGCGGCTGCGGTTACAACTTAAGGATACGGAAAAAGTTGATCAGCCGGCCGTAATGGCTG    5460
EV55_contig000014      CGCGGCTGCGGTTACAACTTAAGGATACGGAAAAAGTTGATCAGCCGGCCGTAATGGCTG    4175
EV54_contig000003      CGCGGCTGCGGTTACAACTTAAGGATACGGAAAAAGTTGATCAGCCGGCCGTAATGGCTG    5460
EV53_contig000052      CGCGGCTGCGGTTACAACTTAAGGATACGGAAAAAGTTGATCAGCCGGCCGTAATGGCTG    5460
                       ************************************************************

EV48_contig000009      CTGGCGTGCCTGGTTTGAACGTACTGGATGTTCATAACATCCACATTGTGATCGGTACGG    5520
EV55_contig000014      CTGGCGTGCCTGGTTTGAACGTACTGGATGTTCATAACATCCACATTGTGATCGGTACGG    4235
EV54_contig000003      CTGGCGTGCCTGGTTTGAACGTACTGGATGTTCATAACATCCACATTGTGATCGGTACGG    5520
EV53_contig000052      CTGGCGTGCCTGGTTTGAACGTACTGGATGTTCATAACATCCACATTGTGATCGGTACGG    5520
                       ************************************************************

EV48_contig000009      AAGTCCAGTTCGTTGCGGAGGCTTTACAAAAATTATTTTCTGGTCAGGTGGCGACGACCC    5580
EV55_contig000014      AAGTCCAGTTCGTTGCGGAGGCTTTACAAAAATTATTTTCTGGTCAGGTGGCGACGACCC    4295
EV54_contig000003      AAGTCCAGTTCGTTGCGGAGGCTTTACAAAAATTATTTTCTGGTCAGGTGGCGACGACCC    5580
EV53_contig000052      AAGTCCAGTTCGTTGCGGAGGCTTTACAAAAATTATTTTCTGGTCAGGTGGCGACGACCC    5580
                       ************************************************************

EV48_contig000009      CGGCATCTGATGCTGAATCAAAGGCCCCGCTTGAACCGCAAACTGCTACTGTAACGGAAG    5640
EV55_contig000014      CGGCATCTGATGCTGAATCAAAGGCCCCGCTTGAACCGCAAACTGCTACTGTAACGGAAG    4355
EV54_contig000003      CGGCATCTGATGCTGAATCAAAGGCCCCGCTTGAACCGCAAACTGCTACTGTAACGGAAG    5640
EV53_contig000052      CGGCATCTGATGCTGAATCAAAGGCCCCGCTTGAACCGCAAACTGCTACTGTAACGGAAG    5640
                       ************************************************************

EV48_contig000009      CGCCGGTGACAACGATTTTGCGGGCACCGGCAACGGGACAATTAATGCCGATTAGTGCGG    5700
EV55_contig000014      CGCCGGTGACAACGATTTTGCGGGCACCGGCAACGGGACAATTAATGCCGATTAGTGCGG    4415
EV54_contig000003      CGCCGGTGACAACGATTTTGCGGGCACCGGCAACGGGACAATTAATGCCGATTAGTGCGG    5700
EV53_contig000052      CGCCGGTGACAACGATTTTGCGGGCACCGGCAACGGGACAATTAATGCCGATTAGTGCGG    5700
                       ************************************************************

EV48_contig000009      TTGCGGATGAGACGTTTGCTGGTAAACTCTTAGGTGATGGTTACGCTGTTGAACCCGAAG    5760
EV55_contig000014      TTGCGGATGAGACGTTTGCTGGTAAACTCTTAGGTGATGGTTACGCTGTTGAACCCGAAG    4475
EV54_contig000003      TTGCGGATGAGACGTTTGCTGGTAAACTCTTAGGTGATGGTTACGCTGTTGAACCCGAAG    5760
EV53_contig000052      TTGCGGATGAGACGTTTGCTGGTAAACTCTTAGGTGATGGTTACGCTGTTGAACCCGAAG    5760
                       ************************************************************

EV48_contig000009      ATGGTGAAGTCGTTGCTCCGGTTAGTGGGACAGTGACAAGTGTCTTTCCGACCAAACACG    5820
EV55_contig000014      ATGGTGAAGTCGTTGCTCCGGTTAGTGGGACAGTGACAAGTGTCTTTCCGACCAAACACG    4535
EV54_contig000003      ATGGTGAAGTCGTTGCTCCGGTTAGTGGGACAGTGACAAGTGTCTTTCCGACCAAACACG    5820
EV53_contig000052      ATGGTGAAGTCGTTGCTCCGGTTAGTGGGACAGTGACAAGTGTCTTTCCGACCAAACACG    5820
                       ************************************************************

EV48_contig000009      CTATCGGTTTGAAGACGACGAGTGGTTTAGAAGTCTTATTACACATGGGGATCAATACCG    5880
EV55_contig000014      CTATCGGTTTGAAGACGACGAGTGGTTTAGAAGTCTTATTACACATGGGGATCAATACCG    4595
EV54_contig000003      CTATCGGTTTGAAGACGACGAGTGGTTTAGAAGTCTTATTACACATGGGGATCAATACCG    5880
EV53_contig000052      CTATCGGTTTGAAGACGACGAGTGGTTTAGAAGTCTTATTACACATGGGGATCAATACCG    5880
                       ************************************************************

EV48_contig000009      TGGAAATGAATGGTACGCCGTTCAAGTTACACGTGGCAACGGGTGATGAAATTGCGGCCG    5940
EV55_contig000014      TGGAAATGAATGGTACGCCGTTCAAGTTACACGTGGCAACGGGTGATGAAATTGCGGCCG    4655
EV54_contig000003      TGGAAATGAATGGTACGCCGTTCAAGTTACACGTGGCAACGGGTGATGAAATTGCGGCCG    5940
EV53_contig000052      TGGAAATGAATGGTACGCCGTTCAAGTTACACGTGGCAACGGGTGATGAAATTGCGGCCG    5940
                       ************************************************************
```

```
EV48_contig000009      GTACTGCGGTAGCGACGGTTGACTTGGCTGCTATCAAGTCAGCCGGTAAAGCGACAACCA    6000
EV55_contig000014      GTACTGCGGTAGCGACGGTTGACTTGGCTGCTATCAAGTCAGCCGGTAAAGCGACAACCA    4715
EV54_contig000003      GTACTGCGGTAGCGACGGTTGACTTGGCTGCTATCAAGTCAGCCGGTAAAGCGACAACCA    6000
EV53_contig000052      GTACTGCGGTAGCGACGGTTGACTTGGCTGCTATCAAGTCAGCCGGTAAAGCGACAACCA    6000
                       ************************************************************

EV48_contig000009      TGATGGTCGTTATCACCAATATGGATCACGTTAACAAACTAACCCTTAATCCAACTGGTC    6060
EV55_contig000014      TGATGGTCGTTATCACCAATATGGATCACGTTAACAAACTAACCCTTAATCCAACTGGTC    4775
EV54_contig000003      TGATGGTCGTTATCACCAATATGGATCACGTTAACAAACTAACCCTTAATCCAACTGGTC    6060
EV53_contig000052      TGATGGTCGTTATCACCAATATGGATCACGTTAACAAACTAACCCTTAATCCAACTGGTC    6060
                       ************************************************************

EV48_contig000009      ATGTGACTAGCGGTGATTTGATTGGCGCAGCTGAATAAGTTACGCGCTGTCGACGTTGGG    6120
EV55_contig000014      ATGTGACTAGCGGTGATTTGATTGGCGCAGCTGAATAAGTTACGCGCTGTCGACGTTGGG    4835
EV54_contig000003      ATGTGACTAGCGGTGATTTGATTGGCGCAGCTGAATAAGTTACGCGCTGTCGACGTTGGG    6120
EV53_contig000052      ATGTGACTAGCGGTGATTTGATTGGCGCAGCTGAATAAGTTACGCGCTGTCGACGTTGGG    6120
                       ************************************************************

EV48_contig000009      TTTATTGAAAAAGTTTCGTCTAAATCTTTTCGATTTAGGCGGAACTTTTTTTGTCGCATT    6180
EV55_contig000014      TTTATTGAAAAAGTTTCGTCTAAATCTTTTCGATTTAGGCGGAACTTTTTTTGTCGCATT    4895
EV54_contig000003      TTTATTGAAAAAGTTTCGTCTAAATCTTTTCGATTTAGGCGGAACTTTTTTTGTCGCATT    6180
EV53_contig000052      TTTATTGAAAAAGTTTCGTCTAAATCTTTTCGATTTAGGCGGAACTTTTTTTGTCGCATT    6180
                       ************************************************************

EV48_contig000009      GGTTATCGAGGCAACGGATGACAGCAAATGAACAACTTGATGCGGGGACAGTGCGTTTGA    6240
EV55_contig000014      GGTTATCGAGGCAACGGATGACAGCAAATGAACAACTTGATGCGGGGACAGTGCGTTTGA    4955
EV54_contig000003      GGTTATCGAGGCAACGGATGACAGCAAATGAACAACTTGATGCGGGGACAGTGCGTTTGA    6240
EV53_contig000052      GGTTATCGAGGCAACGGATGACAGCAAATGAACAACTTGATGCGGGGACAGTGCGTTTGA    6240
                       ************************************************************

EV48_contig000009      TTGCTTCCCACTGACTCCCACGATATACTATTGGCATTATAGGGGGGCGGTAACATGATT    6300
EV55_contig000014      TTGCTTCCCACTGACTCCCACGATATACTATTGGCATTATAGGGGGGCGGTAACATGATT    5015
EV54_contig000003      TTGCTTCCCACTGACTCCCACGATATACTATTGGCATTATAGGGGGGCGGTAACATGATT    6300
EV53_contig000052      TTGCTTCCCACTGACTCCCACGATATACTATTGGCATTATAGGGGGGCGGTAACATGATT    6300
                       ************************************************************

EV48_contig000009      GACGAAGTATTAGTATTGAACTCAACGCCAACGGGGATTTTACAGAATTCAGTTTCAGGA    6360
EV55_contig000014      GACGAAGTATTAGTATTGAACTCAACGCCAACGGGGATTTTACAGAATTCAGTTTCAGGA    5075
EV54_contig000003      GACGAAGTATTAGTATTGAACTCAACGCCAACGGGGATTTTACAGAATTCAGTTTCAGGA    6360
EV53_contig000052      GACGAAGTATTAGTATTGAACTCAACGCCAACGGGGATTTTACAGAATTCAGTTTCAGGA    6360
                       ************************************************************

EV48_contig000009      AGCATGGTGATTGACGCGGCACTTGATGCACGACTAGTCAAGCGACTAATGCAGGCGCTA    6420
EV55_contig000014      AGCATGGTGATTGACGCGGCACTTGATGCACGACTAGTCAAGCGACTAATGCAGGCGCTA    5135
EV54_contig000003      AGCATGGTGATTGACGCGGCACTTGATGCACGACTAGTCAAGCGACTAATGCAGGCGCTA    6420
EV53_contig000052      AGCATGGTGATTGACGCGGCACTTGATGCACGACTAGTCAAGCGACTAATGCAGGCGCTA    6420
                       ************************************************************

EV48_contig000009      ACTCAATCGATGCTTAAGGTCAGGGTTGCAACCGCGGACGACTTACCGGTGACCGACCAC    6480
EV55_contig000014      ACTCAATCGATGCTTAAGGTCAGGGTTGCAACCGCGGACGACTTACCGGTGACCGACCAC    5195
EV54_contig000003      ACTCAATCGATGCTTAAGGTCAGGGTTGCAACCGCGGACGACTTACCGGTGACCGACCAC    6480
EV53_contig000052      ACTCAATCGATGCTTAAGGTCAGGGTTGCAACCGCGGACGACTTACCGGTGACCGACCAC    6480
                       ************************************************************

EV48_contig000009      CAATTGCTGGTCGTCAGTCGACCAGTGGCAGCCGTGTTGGACTTACAAGCTAGTGCGCGA    6540
EV55_contig000014      CAATTGCTGGTCGTCAGTCGACCAGTGGCAGCCGTGTTGGACTTACAAGCTAGTGCGCGA    5255
EV54_contig000003      CAATTGCTGGTCGTCAGTCGACCAGTGGCAGCCGTGTTGGACTTACAAGCTAGTGCGCGA    6540
EV53_contig000052      CAATTGCTGGTCGTCAGTCGACCAGTGGCAGCCGTGTTGGACTTACAAGCTAGTGCGCGA    6540
                       ************************************************************

EV48_contig000009      GTTCTGATTTATCCAGAAGCAACGGGACTGACAGCCGCTGGAATGACACAGCTAGCAGTG    6600
EV55_contig000014      GTTCTGATTTATCCAGAAGCAACGGGACTGACAGCCGCTGGAATGACACAGCTAGCAGTG    5315
EV54_contig000003      GTTCTGATTTATCCAGAAGCAACGGGACTGACAGCCGCTGGAATGACACAGCTAGCAGTG    6600
EV53_contig000052      GTTCTGATTTATCCAGAAGCAACGGGACTGACAGCCGCTGGAATGACACAGCTAGCAGTG    6600
                       ************************************************************

EV48_contig000009      CAGATTGATCAACGATTGACTAAAATATCCAGATAACGATCAAAAAACGCCCAGGGCAGT    6660
EV55_contig000014      CAGATTGATCAACGATTGACTAAAATATCCAGATAACGATCAAAAAACGCCCAGGGCAGT    5375
EV54_contig000003      CAGATTGATCAACGATTGACTAAAATATCCAGATAACGATCAAAAAACGCCCAGGGCAGT    6660
EV53_contig000052      CAGATTGATCAACGATTGACTAAAATATCCAGATAACGATCAAAAAACGCCCAGGGCAGT    6660
                       ************************************************************

EV48_contig000009      TGTTCGCAAGTTACTGAACAGTTGCCCTGGGCGTTTTTAATGATTAGTCTTTGTGGTGTT    6720
EV55_contig000014      TGTTCGCAAGTTACTGAACAGTTGCCCTGGGCGTTTTTAATGATTAGTCTTTGTGGTGTT    5435
EV54_contig000003      TGTTCGCAAGTTACTGAACAGTTGCCCTGGGCGTTTTTAATGATTAGTCTTTGTGGTGTT    6720
EV53_contig000052      TGTTCGCAAGTTACTGAACAGTTGCCCTGGGCGTTTTTAATGATTAGTCTTTGTGGTGTT    6720
                       ************************************************************

EV48_contig000009      TTTTAATGATAGTTACTAATTTCAATCAAGTTGTTATCAGGATCACGAACATACAGTGAC    6780
EV55_contig000014      TTTTAATGATAGTTACTAATTTCAATCAAGTTGTTATCAGGATCACGAACATACAGTGAC    5495
EV54_contig000003      TTTTAATGATAGTTACTAATTTCAATCAAGTTGTTATCAGGATCACGAACATACAGTGAC    6780
EV53_contig000052      TTTTAATGATAGTTACTAATTTCAATCAAGTTGTTATCAGGATCACGAACATACAGTGAC    6780
                       ************************************************************
```

```
EV48_contig000009      GTCAACTTTCCGTGGGCACCAGTTCTTTCAACCGGCCCGGCAATCACATCCACGAAATAA    6840
EV55_contig000014      GTCAACTTTCCGTGGGCACCAGTTCTTTCAACCGGCCCGGCAATCACATCCACGAAATAA    5555
EV54_contig000003      GTCAACTTTCCGTGGGCACCAGTTCTTTCAACCGGCCCGGCAATCACATCCACGAAATAA    6840
EV53_contig000052      GTCAACTTTCCGTGGGCACCAGTTCTTTCAACCGGCCCGGCAATCACATCCACGAAATAA    6840
                       ************************************************************

EV48_contig000009      CTCTTTAAGTGATGCTGAATATCATCAATATTATCCTTCGCAATCAAGCAAAGATCTGCA    6900
EV55_contig000014      CTCTTTAAGTGATGCTGAATATCATCAATATTATCCTTCGCAATCAAGCAAAGATCTGCA    5615
EV54_contig000003      CTCTTTAAGTGATGCTGAATATCATCAATATTATCCTTCGCAATCAAGCAAAGATCTGCA    6900
EV53_contig000052      CTCTTTAAGTGATGCTGAATATCATCAATATTATCCTTCGCAATCAAGCAAAGATCTGCA    6900
                       ************************************************************

EV48_contig000009      CTGCCGGGTGTCGGCTTAGCCGCAATGGGTTGGTGTGGCTGGTTAGTGGTCTGAAAGTTA    6960
EV55_contig000014      CTGCCGGGTGTCGGCTTAGCCGCAATGGGTTGGTGTGGCTGGTTAGTGGTCTGAAAGTTA    5675
EV54_contig000003      CTGCCGGGTGTCGGCTTAGCCGCAATGGGTTGGTGTGGCTGGTTAGTGGTCTGAAAGTTA    6960
EV53_contig000052      CTGCCGGGTGTCGGCTTAGCCGCAATGGGTTGGTGTGGCTGGTTAGTGGTCTGAAAGTTA    6960
                       ************************************************************

EV48_contig000009      ATTTTTTGTTTACCCACTAAGACAGCTTGACGGTCACCATCGAAGGTCACAATGGGAAGG    7020
EV55_contig000014      ATTTTTTGTTTACCCACTAAGACAGCTTGACGGTCACCATCGAAGGTCACAATGGGAAGG    5735
EV54_contig000003      ATTTTTTGTTTACCCACTAAGACAGCTTGACGGTCACCATCGAAGGTCACAATGGGAAGG    7020
EV53_contig000052      ATTTTTTGTTTACCCACTAAGACAGCTTGACGGTCACCATCGAAGGTCACAATGGGAAGG    7020
                       ************************************************************

EV48_contig000009      TCAAATACTTCATGGTAAAACCGTAATGACCGTGCAATGTCAGTAACGGTTAGGGTCAAA    7080
EV55_contig000014      TCAAATACTTCATGGTAAAACCGTAATGACCGTGCAATGTCAGTAACGGTTAGGGTCAAA    5795
EV54_contig000003      TCAAATACTTCATGGTAAAACCGTAATGACCGTGCAATGTCAGTAACGGTTAGGGTCAAA    7080
EV53_contig000052      TCAAATACTTCATGGTAAAACCGTAATGACCGTGCAATGTCAGTAACGGTTAGGGTCAAA    7080
                       ************************************************************

EV48_contig000009      TGATCGATATCGCGAATATTCACAAGCTCAAATCCCTTCAAAAATAATTCATTATTAAAC    7140
EV55_contig000014      TGATCGATATCGCGAATATTCACAAGCTCAAATCCCTTCAAAAATAATTCATTATTAAAC    5855
EV54_contig000003      TGATCGATATCGCGAATATTCACAAGCTCAAATCCCTTCAAAAATAATTCATTATTAAAC    7140
EV53_contig000052      TGATCGATATCGCGAATATTCACAAGCTCAAATCCCTTCAAAAATAATTCATTATTAAAC    7140
                       ************************************************************

EV48_contig000009      GTGTCGTACTCATCAGATTTGGCATGAACCATGATTAAAATGTACCAGTTAGATTGCTGT    7200
EV55_contig000014      GTGTCGTACTCATCAGATTTGGCATGAACCATGATTAAAATGTACCAGTTAGATTGCTGT    5915
EV54_contig000003      GTGTCGTACTCATCAGATTTGGCATGAACCATGATTAAAATGTACCAGTTAGATTGCTGT    7200
EV53_contig000052      GTGTCGTACTCATCAGATTTGGCATGAACCATGATTAAAATGTACCAGTTAGATTGCTGT    7200
                       ************************************************************

EV48_contig000009      TCGTCAGCCGTTGCGCGTCCTATGCCGACCTCCGGGGCTGGGTGCCAATTGCTGGAACGC    7260
EV55_contig000014      TCGTCAGCCGTTGCGCGTCCTATGCCGACCTCCGGGGCTGGGTGCCAATTGCTGGAACGC    5975
EV54_contig000003      TCGTCAGCCGTTGCGCGTCCTATGCCGACCTCCGGGGCTGGGTGCCAATTGCTGGAACGC    7260
EV53_contig000052      TCGTCAGCCGTTGCGCGTCCTATGCCGACCTCCGGGGCTGGGTGCCAATTGCTGGAACGC    7260
                       ************************************************************

EV48_contig000009      AGCCGACATCGATTTGAGCTAACGCATAACCCGCGTCATCTCAAATACGAGTCTTATTCT    7320
EV55_contig000014      AGCCGACATCGATTTGAGCTAACGCATAACCCGCGTCATCTCAAATACGAGTCTTATTCT    6035
EV54_contig000003      AGCCGACATCGATTTGAGCTAACGCATAACCCGCGTCATCTCAAATACGAGTCTTATTCT    7320
EV53_contig000052      AGCCGACATCGATTTGAGCTAACGCATAACCCGCGTCATCTCAAATACGAGTCTTATTCT    7320
                       ************************************************************

EV48_contig000009      AAGCCGGAAGCAGACCACTTCCAGCTTAGAATAATTTGGCTACTGAGCATTGTCACCCAG    7380
EV55_contig000014      AAGCCGGAAGCAGACCACTTCCAGCTTAGAATAATTTGGCTACTGAGCATTGTCACCCAG    6095
EV54_contig000003      AAGCCGGAAGCAGACCACTTCCAGCTTAGAATAATTTGGCTACTGAGCATTGTCACCCAG    7380
EV53_contig000052      AAGCCGGAAGCAGACCACTTCCAGCTTAGAATAATTTGGCTACTGAGCATTGTCACCCAG    7380
                       ************************************************************

EV48_contig000009      CCCCTCCAGTCTCAAGCCTGTGAGAGGCCACGGGTTGAAGCCGAGCATTGTCTAGCCAAC    7440
EV55_contig000014      CCCCTCCAGTCTCAAGCCTGTGAGAGGCCACGGGTTGAAGCCGAGCATTGTCTAGCCAAC    6155
EV54_contig000003      CCCCTCCAGTCTCAAGCCTGTGAGAGGCCACGGGTTGAAGCCGAGCATTGTCTAGCCAAC    7440
EV53_contig000052      CCCCTCCAGTCTCAAGCCTGTGAGAGGCCACGGGTTGAAGCCGAGCATTGTCTAGCCAAC    7440
                       ************************************************************

EV48_contig000009      GGATTAATGCGGTTTGGGCATTAG        7464
EV55_contig000014      GGATTAATGCGGTTTGGGCATTAG        6179
EV54_contig000003      GGATTAATGCGGTTTGGGCATTAG        7464
EV53_contig000052      GGATTAATGCGGTTTGGGCATTAG        7464
                       ************************
```

Appendix 9 Phyre^2 Threading and Results



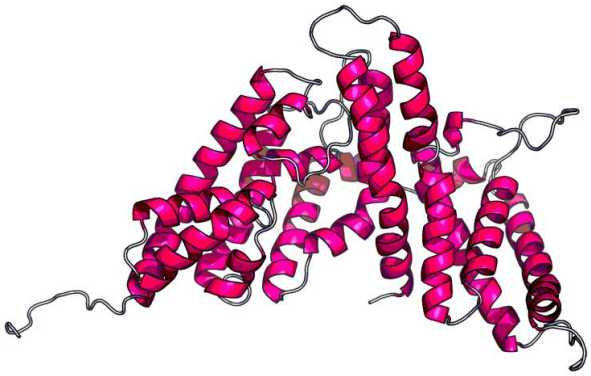
#	Template	Alignment Coverage	3D Model	Confidence	% i.d.	Template Information
1	c3qnqD_	Alignment		100.0	14	**PDB header:**membrane protein, transport protein **Chain:** D: **PDB Molecule:**pts system, cellobiose-specific iic component; **PDBTitle:** crystal structure of the transporter chbc, the iic component from the2 n,n'-diacetylchitobiose-specific phosphotransferase system
2	c3b9yA_	Alignment		52.8	12	**PDB header:**transport protein **Chain:** A: **PDB Molecule:**ammonium transporter family rh-like protein; **PDBTitle:** crystal structure of the nitrosomonas europaea rh protein
3	c2b6pA_	Alignment		44.8	12	**PDB header:**membrane protein **Chain:** A: **PDB Molecule:**lens fiber major intrinsic protein; **PDBTitle:** x-ray structure of lens aquaporin-0 (aqp0) (lens mip) in an open pore2 state
4	c3hd6A_	Alignment		29.1	13	**PDB header:**membrane protein, transport protein **Chain:** A: **PDB Molecule:**ammonium transporter rh type c; **PDBTitle:** crystal structure of the human rhesus glycoprotein rhcg
5	c4djiA_	Alignment		25.3	10	**PDB header:**transport protein **Chain:** A: **PDB Molecule:**probable glutamate/gamma-aminobutyrate antiporter; **PDBTitle:** structure of glutamate-gaba antiporter gadc
6	c2kncA_	Alignment		22.5	17	**PDB header:**cell adhesion **Chain:** A: **PDB Molecule:**integrin alpha-iib; **PDBTitle:** platelet integrin alfaiib-beta3 transmembrane-cytoplasmic2 heterocomplex
7	c2b2hA_	Alignment		15.6	9	**PDB header:**transport protein **Chain:** A: **PDB Molecule:**ammonium transporter; **PDBTitle:** ammonium transporter amt-1 from a. fulgidus (as)
8	c2hnhA_	Alignment		13.2	36	**PDB header:**transferase **Chain:** A: **PDB Molecule:**dna polymerase iii alpha subunit; **PDBTitle:** crystal structure of the catalytic alpha subunit of e. coli2 replicative dna polymerase iii
9	c3a0hi_	Alignment		11.0	20	**PDB header:**electron transport **Chain:** I: **PDB Molecule:**photosystem ii reaction center protein i; **PDBTitle:** crystal structure of i-substituted photosystem ii complex
10	d2axti1	Alignment		11.0	20	**Fold:**Single transmembrane helix **Superfamily:**Photosystem II reaction center protein I, PsbI **Family:**PsbI-like
11	c2rddB_	Alignment		10.0	32	**PDB header:**membrane protein/transport protein **Chain:** B: **PDB Molecule:**upf0092 membrane protein yajc; **PDBTitle:** x-ray crystal structure of acrb in complex with a novel2 transmembrane helix.
12	c4j05A_	Alignment		8.8	8	**PDB header:**transport protein **Chain:** A: **PDB Molecule:**phosphate transporter; **PDBTitle:** crystal structure of a eukaryotic phosphate transporter
13	c4apsB_	Alignment		8.1	11	**PDB header:**transport protein **Chain:** B: **PDB Molecule:**di-or tripeptide h+ symporter; **PDBTitle:** crystal structure of a pot family peptide transporter in an inward2 open conformation.
14	d2nr9a1	Alignment		6.7	21	**Fold:**Rhomboid-like **Superfamily:**Rhomboid-like **Family:**Rhomboid-like

Appendix 10 Phyre2 homology based tertiary structure prediction of *L. plantarum* ATCC 8014 PTS18CBA EIIC domain

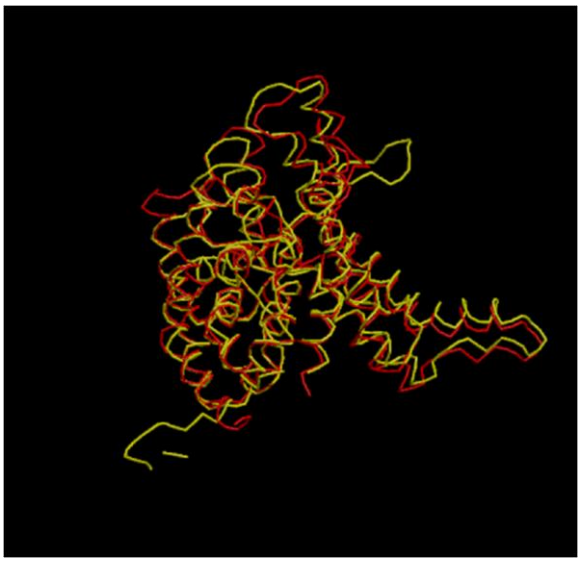
Raw results of the threading is provided as electronic material on CD.

Appendix 11 Superpose of the RaptorX and Phyre^2 model of PTS18C threaded into the structure of ChbC

Rank	P-value	Score	uGDT/GDT	uSeqID/SeqID	ModelName	Template
1	4.2e-07	269	239/65	47/13	3qnqA-66985_1	3qnqA

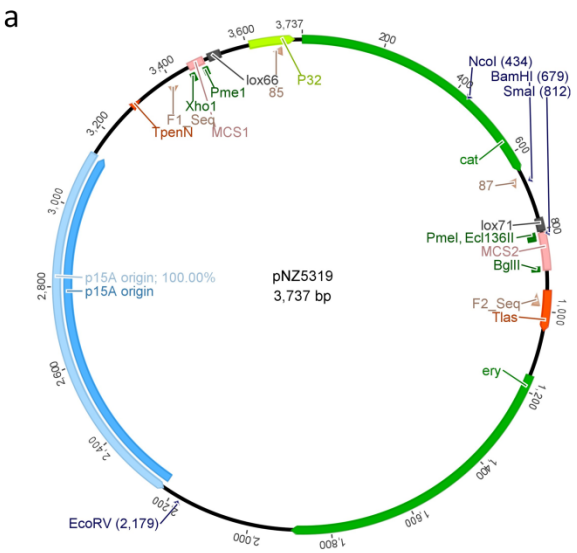


The predicted model by RaptorX (red) and Phyre^2 (yellow) were superposed with default settings.

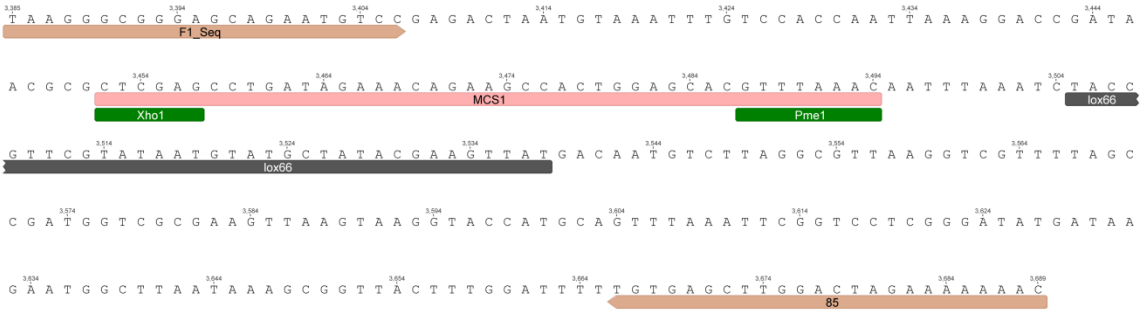


Appendix 12 Plasmid map of pNZ5319

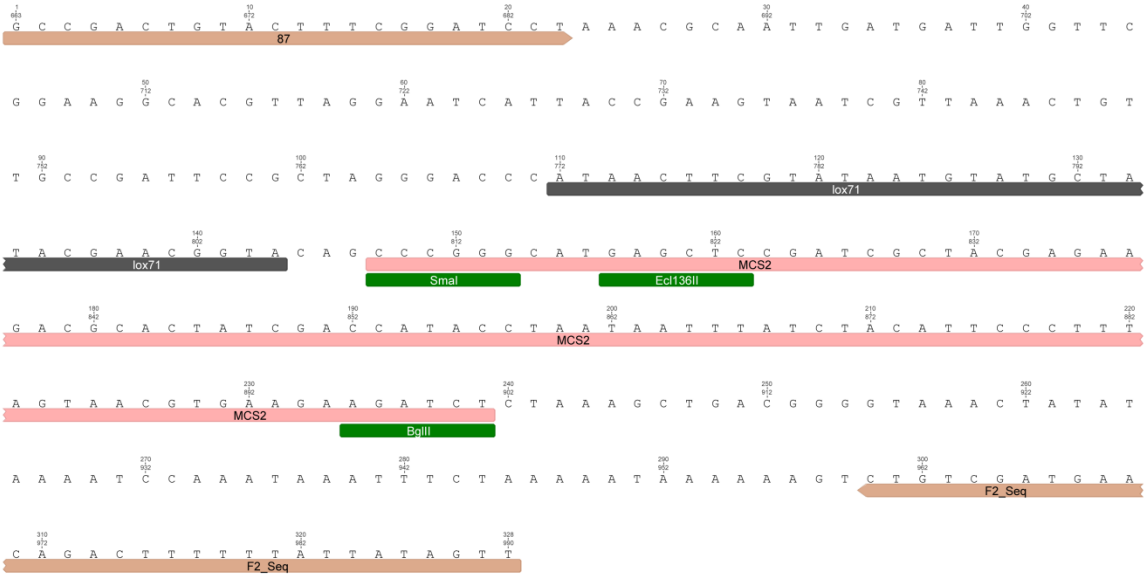
a



b



c



Plasmid Map of pNZ5319

a) Schematic diagram of mutagenic plasmid pNZ5319 constructed by and obtained from Lambert, J. M., et al. (2007). Cloning was performed at the Multiple cloning sites 1 (b) and 2 (c); annotated as a pink ribbon. Restriction enzyme recognition sites used in this study are shown (dark blue). Two selectable markers chloramphenicol resistance (Cm^R) and erythromycin resistance (Em^R) are encoded by the *cat* gene (dark green) and *ery* gene (dark green) which are under the control of the P_{32} promoter (light green). Plasmid replication is controlled by the pACYC184-derived origin of replication, Ori (blue). T_{penN} and T_{las} (orange) are Lactococcal transcription terminators. Cre recombinase recognition sites, lox66 and lox71 (grey), flank *cat*. Primers (brown) located flanking the MCSs were used during screening experiments.

Appendix 13 Oligonucleotides used in this study

Name	Sequence
81	AGGACCAGTACGTTTTCGCCGC
85	GTTTTTTTCTAGTCCAAGCTCACA
87	GCCGACTGTACTTTCGGATCCT
80_term	TTACCGCAGCATCAATGACTGACC
82a	GCGGATGAGACGTTTGCTGG
83b	CAGTCCCGTTGCTTCTGGATAAATC
CAT_F	TACCGAAACATAAAACAAGAAGGA
CAT_R	CTGACAATTCCTGAATAGAGTTCA
F1_seq	TAAGGGCGGGAGCAGAATGTCC
F2_seq	AACTATAATAAAAAAGTCTGTTCATCGACAG
F8014_1	GGATTGTTTTTAAGAAGCAAATGC
lac_oprn	NOT DISCLOSED
P1R	GCTAGTCACATGACCAGTTG
pet_32_T7_term	NOT DISCLOSED
pts18CBA_F_NdeI	GCTACGCATATGAAGACATATTTTCAGAAAATCG
pts18CBA_R_NotI	CTATAGCGGCCGCTTCAGCTGCGCCAAT
PTS18CBA_START	ATGAAGACATATTTTCAGAAAATC
R8014_1_STP	TTATGACTGACCGATTTTCTGAAAATA
TRKCmSeqF2	NOT DISCLOSED
TRKCmSeqR3	NOT DISCLOSED

Appendix 14 Plasmid used in study

Plasmid Name	Relevant Details	Size (kbp)	Source/Reference
pNZ5319	Em^R,Cm^R,	3.5	NZIO
pNZ5319_F1_14917	Em^R,Cm^R	4.5	This Study
pNZ5319_F1_F2_14917	Em^R,Cm^R	5.5	This Study
pNZ5319_F1_F2_14917_Emr-$^-$	Cm^R	4.427	This Study
pNZ5319_F1_8014	Em^R,Cm^R	4.5	This Study
pNZ5319_F2_8014	Em^R,Cm^R	4.5	This Study
pNZ5319_F1_F2_8014	Em^R,Cm^R	5.5	This Study
pTRK669	CM^R, RepA1, ori (pWV01)	2.437	(Russell and Klaenhammer, 2001)
pRV613	Amp^R, Em^R	8.69	(Crutz-Le Coq and Zagorec, 2008)
pET-21b(+)	Amp^R, T7 promoter, ori (pBR322)	5.44	Novagen
pET-21b(+)_8014_His6_pts18CBA	Amp^R, T7 promoter, ori (pBR322), His_6 8014 *pts18CBA*,	7.362	This Study
pET-21b(+)_14917_His6_ pts18CBA	Amp^R, T7 promoter, ori (pBR322),His_6 14917 *pts18CBA*	7.362	This Study

Amp^R is ampicillin resistance, Cm^R is chloramphenicol resistance, His_6 is a hex-histidine tag, Em_R is erythromycin resistance. NZIO is Wageningen Centre for Food Science, Health and Safety Department, P.O. Box 20, 6710 BA Ede, The Netherlands.

Appendix 15 Organisms and strains used in this study

Organism/Strain	Genotype	Reference/Source
Lactobacillus plantarum		
NC8	Wild-type	(Aukrust and Blom, 1992)
NC8 Δ*pts18CA*	Ω*pts18CBA*::Cm^R;glycocin-F^R;	This study
ATCC 8014	Wild-type	ATCC
ATCC 8014 LP8_1**L**	*purR(C825A); pts18CBA(C610A);* glycocin-F^R	This study
ATCC 8014 LP8_2**H**	glycocin-F^R	This study
ATCC 8014 LP8_3**H**	glycocin-F^R	This study
ATCC 8014 LP8_4**H**	*pts18CBA(Δ769-838)*; glycocin-F^R	This study
ATCC 8014 LP8_5**H**	*pts18CBA(Am898);* glycocin-F^R	This study
WCFS1	Wild-type	
Lactobacillus plantarum* subsp. *plantarum		
ATCC 14917	Wild-type	ATCC
ATCC 14917 LP1_1**H**	glycocin-F^R	This study
ATCC 14917 LP1_2**H**	Δ*panE1*;*pts18CBA(Δ1-72);*glycocin-F^R	This study
ATCC 14917 LP1_3**L**	*pts18CBA(A630C);*glycocin-F^R	This study
ATCC 14917 LP1_4**H**	*pts18CBA(G727A);*glycocin-F^R	This study
ATCC 14917 LP1_5**L**	glycocin-F^R	This study
Escherichia coli		
XLI Blue	Endonuclease deficient	Stratagene: CA, USA
EC100	RepA1 MC1000	Lucigen; WI, USA
Bl21 (DE3)	Contains IPTG-inducible T7 polymerase,	Stratagene: CA, USA
glycocin F^R is resistance to glycocin F, Cm^R is chloramphenicol resistance		

Appendix 16 Sequence alignment of the expected *L. plantarum* NC8 Δ*pts18CBA* to sequenced colony 4 gDNA

CLUSTAL OMEGA (http://www.ebi.ac.uk/Tools/msa/clustalo/) multiple sequence alignment performed with default settings.

```
PTS18CBA_START_1(reversed)      ------------------------------------------------------------      0
PTS18CBA_START_2(reversed)      ------------------------------------------------------------      0
CAT_F_(reversed)                ------------------------------------------------------------      0
P1R                             ------------------------------------------------------------      0
81_(reversed)                   ------------------------------------------------------------      0
80_term                         ------------------------------------------------------------      0
85                              ------------------------------------------------------------      0
CAT_R                           ------------------------------------------------------------      0
87_(reversed)                   ------------------------------------------------------------      0
expected_knockout               AGGGCAACTGTTCAGTAACTTGCGAACAACTCGCCGTGGCGTTTTTTGATCGTTATCTGG      60
82a_(reversed)                  ------------------------------------------------------------      0
83b                             ------------------------------------------------------------      0
```

```
PTS18CBA_START_1(reversed)      ------------------------------------------------------------          0
PTS18CBA_START_2(reversed)      ------------------------------------------------------------          0
CAT_F_(reversed)                ------------------------------------------------------------          0
P1R                             ------------------------------------------------------------          0
81_(reversed)                   ------------------------------------------------------------          0
80_term                         ------------------------------------------------------------          0
85                              ------------------------------------------------------------          0
CAT_R                           ------------------------------------------------------------          0
87_(reversed)                   ------------------------------------------------------------          0
expected_knockout               ATATTTTAGTCAATCGTTGATCAATCTGCACTGCTAGCTGTGTCATTCCAGCGGCTGTCA          120
82a_(reversed)                  ------------------------------------------------------------          0
83b                             ------------------------------------------------------------          0


PTS18CBA_START_1(reversed)      ------------------------------------------------------------          0
PTS18CBA_START_2(reversed)      ------------------------------------------------------------          0
CAT_F_(reversed)                ------------------------------------------------------------          0
P1R                             ------------------------------------------------------------          0
81_(reversed)                   ------------------------------------------------------------          0
80_term                         ------------------------------------------------------------          0
85                              ------------------------------------------------------------          0
CAT_R                           ------------------------------------------------------------          0
87_(reversed)                   ------------------------------------------------------------          0
expected_knockout               GTCCCGTTGCTTCTGGATAAATCAGAACTCGCGCACTAGCTTGTAAGTCCAACACGGCTG          180
82a_(reversed)                  ------------------------------------------------------------          0
83b                             ------------------------------------------------------------          0


PTS18CBA_START_1(reversed)      ------------------------------------------------------------          0
PTS18CBA_START_2(reversed)      ------------------------------------------------------------          0
CAT_F_(reversed)                ------------------------------------------------------------          0
P1R                             ------------------------------------------------------------          0
81_(reversed)                   ------------------------------------------------------------          0
80_term                         ------------------------------------------------------------          0
85                              ------------------------------------------------------------          0
CAT_R                           ------------------------------------------------------------          0
87_(reversed)                   ------------------------------------------------------------          0
expected_knockout               CCACTGGTCGACTGACGACCAGCAATTGGTGGTCGGTCACCGGTAAGTCGTCCGCGGTTG          240
82a_(reversed)                  -------------------------------------------------AGTCGTCCGCGGTTG          15
83b                             -----------------------------------------------------CGCGGTTG          8


PTS18CBA_START_1(reversed)      ------------------------------------------------------------          0
PTS18CBA_START_2(reversed)      ------------------------------------------------------------          0
CAT_F_(reversed)                ------------------------------------------------------------          0
P1R                             ------------------------------------------------------------          0
81_(reversed)                   ------------------------------------------------------------          0
80_term                         ------------------------------------------------------------          0
85                              ------------------------------------------------------------          0
CAT_R                           ------------------------------------------------------------          0
87_(reversed)                   ------------------------------------------------------------          0
expected_knockout               CAACCCTGACCTTAAGCATCGATTGAGTTAGCGCCTGCATTAGTCGCTTGACTAGTCGTG          300
82a_(reversed)                  CAACCCTGACCTTAAGCATCGATTGAGTTAGCGCCTGCATTAGTCGCTTGACTAGTCGTG          75
83b                             CAACCCTGACCTTAAGCATCGATTGAGTTAGCGCCTGCATTAGTCGCTTGACTAGTCGTG          68


PTS18CBA_START_1(reversed)      ------------------------------------------------------------          0
PTS18CBA_START_2(reversed)      ------------------------------------------------------------          0
CAT_F_(reversed)                ------------------------------------------------------------          0
P1R                             ------------------------------------------------------------          0
81_(reversed)                   ------------------------------------------------------------          0
80_term                         ------------------------------------------------------------          0
85                              ------------------------------------------------------------          0
CAT_R                           ------------------------------------------------------------          0
87_(reversed)                   ------------------------------------------------------------          0
expected_knockout               CATCAAGTGCCGCGTCAATCACCATGCTTCCTGAAACTGAATTCTGTAAAATCCCCGTTG          360
82a_(reversed)                  CATCAAGTGCCGCGTCAATCACCATGCTTCCTGAAACTGAATTCTGTAAAATCCCCGTTG          135
83b                             CATCAAGTGCCGCGTCAATCACCATGCTTCCTGAAACTGAATTCTGTAAAATCCCCGTTG          128


PTS18CBA_START_1(reversed)      ------------------------------------------------------------          0
PTS18CBA_START_2(reversed)      ------------------------------------------------------------          0
CAT_F_(reversed)                ------------------------------------------------------------          0
P1R                             ------------------------------------------------------------          0
81_(reversed)                   ------------------------------------------------------------          0
80_term                         ------------------------------------------------------------          0
85                              ------------------------------------------------------------          0
CAT_R                           ------------------------------------------------------------          0
87_(reversed)                   ----------------------------------------------------------CAATA          5
expected_knockout               GCGTTGAGTTCAATACTAATACTTCGTCAATCATGTTACCGCCCCCCCTATAATGCCAATA          420
82a_(reversed)                  GCGTTGAGTTCAATACTAATACTTCGTCAATCATGTTACCGCCCCCCCTATAATGCCAATA          195
83b                             GCGTTGAGTTCAATACTAATACTTCGTCAATCATGTTACCGCCCCCCCTATAATGCCAATA          188


PTS18CBA_START_1(reversed)      ------------------------------------------------------------          0
PTS18CBA_START_2(reversed)      ------------------------------------------------------------          0
CAT_F_(reversed)                ------------------------------------------------------------          0
P1R                             ------------------------------------------------------------          0
81_(reversed)                   ------------------------------------------------------------          0
80_term                         ------------------------------------------------------------          0
85                              ------------------------------------------------------------          0
CAT_R                           ------------------------------------------------------------          0
87_(reversed)                   GTATATCGTGGGAGTCAGTGGGAAGCAATCAAACGCACTGTCCCCGCATCAAGTTGTTCA          65
expected_knockout               GTATATCGTGGGAGTCAGTGGGAAGCAATCAAACGCACTGTCCCCGCATCAAGTTGTTCA          480
82a_(reversed)                  GTATATCGTGGGAGTCAGTGGGAAGCAATCAAACGCACTGTCCCCGCATCAAGTTGTTCA          255
83b                             GTATATCGTGGGAGTCAGTGGGAAGCAATCAAACGCACTGTCCCCGCATCAAGTTGTTCA          248
```

```
PTS18CBA_START_1(reversed)      ------------------------------------------------------------      0
PTS18CBA_START_2(reversed)      ------------------------------------------------------------      0
CAT_F_(reversed)                ------------------------------------------------------------      0
P1R                             ------------------------------------------------------------      0
81_(reversed)                   ------------------------------------------------------------      0
80_term                         ------------------------------------------------------------      0
85                              ------------------------------------------------------------      0
CAT_R                           ------------------------------------------------------------      0
87_(reversed)                   TTTGCTGTCATCCGTTGCCTCGATAACCAATGCGACAAAAAAAGTTCCGCCTAAATCGAA      125
expected_knockout               TTTGCTGTCATCCGTTGCCTCGATAACCAATGCGACAAAAAAAGTTCCGCCTAAATCGAA      540
82a_(reversed)                  TTTGCTGTCATCCGTTGCCTCGATAACCAATGCGACAAAAAAAGTTCCGCCTAAATCGAA      315
83b                             TTTGCTGTCATCCGTTGCCTCGATAACCAATGCGACAAAAAAAGTTCCGCCTAAATCGAA      308


PTS18CBA_START_1(reversed)      ------------------------------------------------------------      0
PTS18CBA_START_2(reversed)      ------------------------------------------------------------      0
CAT_F_(reversed)                ------------------------------------------------------------      0
P1R                             ------------------------------------------------------------      0
81_(reversed)                   ------------------------------------------------------------      0
80_term                         ------------------------------------------------------------      0
85                              ------------------------------------------------------------      0
CAT_R                           ------------------------------------------------------------      0
87_(reversed)                   AAGATTTAGACGAAACTTTTTCAATAAACCCAACGTCGACAGCGCGTAACTTATTCAGCT      185
expected_knockout               AAGATTTAGACGAAACTTTTTCAATAAACCCAACGTCGACAGCGCGTAACTTATTCAGCT      600
82a_(reversed)                  AAGATTTAGACGAAACTTTTTCAATAAACCCAACGTCGACAGCGCGTAACTTATTCAGCT      375
83b                             AAGATTTAGACGAAACTTTTTCAATAAACCCAACGTCGACAGCGCGTAACTTATTCAGCT      368


PTS18CBA_START_1(reversed)      ------------------------------------------------------------      0
PTS18CBA_START_2(reversed)      ------------------------------------------------------------      0
CAT_F_(reversed)                ------------------------------------------------------------      0
P1R                             ------------------------------------------------------------      0
81_(reversed)                   ------------------------------------------------------------      0
80_term                         ------------------------------------------------------------      0
85                              ------------------------------------------------------------      0
CAT_R                           ------------------------------------------------------------      0
87_(reversed)                   GCGCCAATCAAATCACCGCTAGTCACATGACCAGTTGGATTAAGGGTTAGTTTGTTAACG      245
expected_knockout               GCGCCAATCAAATCACCGCTAGTCACATGACCAGTTGGATTAAGGGTTAGTTTGTTAACG      660
82a_(reversed)                  GCGCCAATCAAATCACCGCTAGTCACATGACCAGTTGGATTAAGGGTTAGTTTGTTAACG      435
83b                             GCGCCAATCAAATCACCGCTAGTCACATGACCAGTTGGATTAAGGGTTAGTTTGTTAACG      428


PTS18CBA_START_1(reversed)      ------------------------------------------------------------      0
PTS18CBA_START_2(reversed)      ------------------------------------------------------------      0
CAT_F_(reversed)                ------------------------------------------------------------      0
P1R                             ------------------------------GGTTGTCGCTTTACCGGCTGACTTGATAGCA       31
81_(reversed)                   ------------------------------------------------------------      0
80_term                         ------------------------------------------------------------      0
85                              ------------------------------------------------------------      0
CAT_R                           ------------------------------------------------------------      0
87_(reversed)                   TGATCCATATTGGTGATAACGACCATCATGGTTGTCGCTTTACCGGCTGACTTGATAGCA      305
expected_knockout               TGATCCATATTGGTGATAACGACCATCATGGTTGTCGCTTTACCGGCTGACTTGATAGCA      720
82a_(reversed)                  TGATCCATATTGGTGATAACGACCATCATGGTTGTCGCTTTACCGGCTGACTTGATAGCA      495
83b                             TGATCCATATTGGTGATAACGACCATCATGGTTGTCGCTTTACCGGCTGACTTGATAGCA      488


PTS18CBA_START_1(reversed)      ------------------------------------------------------------      0
PTS18CBA_START_2(reversed)      ------------------------------------------------------------      0
CAT_F_(reversed)                ------------------------------------------------------------      0
P1R                             GCCAAGTCAACCGTCGCTACCGCAGTACCGGCCGCAATTTCATCACCCGTTGCCACGTGT       91
81_(reversed)                   ------------------------------------------------------------      0
80_term                         ------------------------------------------------------------      0
85                              ------------------------------------------------------------      0
CAT_R                           ------------------------------------------------------------      0
87_(reversed)                   GCCAAGTCAACCGTCGCTACCGCAGTACCGGCCGCAATTTCATCACCCGTTGCCACGTGT      365
expected_knockout               GCCAAGTCAACCGTCGCTACCGCAGTACCGGCCGCAATTTCATCACCCGTTGCCACGTGT      780
82a_(reversed)                  GCCAAGTCAACCGTCGCTACCGCAGTACCGGCCGCAATTTCATCACCCGTTGCCACGTGT      555
83b                             GCCAAGTCAACCGTCGCTACCGCAGTACCGGCCGCAATTTCATCACCCGTTGCCACGTGT      548


PTS18CBA_START_1(reversed)      ------------------------------------------------------------      0
PTS18CBA_START_2(reversed)      ------------------------------------------------------------      0
CAT_F_(reversed)                --------------------------CACGGTATTGATCCCCATGTGTAATAAGACTTCT       34
P1R                             AACTTGAACGGCGTACCATTCATTTCCACGGTATTGATCCCCATGTGTAATAAGACTTCT      151
81_(reversed)                   ------------------------------------------------------------      0
80_term                         ------------------------------------------------------------      0
85                              ------------------------------------------------------------      0
CAT_R                           ------------------------------------------------------------      0
87_(reversed)                   AACTTGAACGGCGTACCATTCATTTCCACGGTATTGATCCCCATGTGTAATAAGACTTCT      425
expected_knockout               AACTTGAACGGCGTACCATTCATTTCCACGGTATTGATCCCCATGTGTAATAAGACTTCT      840
82a_(reversed)                  AACTTGAACGGCGTACCATTCATTTCCACGGTATTGATCCCCATGTGTAATAAGACTTCT      615
83b                             AACTTGAACGGCGTACCATTCATTTCCACGGTATTGATCCCCATGTGTAATAAGACTTCT      608


PTS18CBA_START_1(reversed)      ------------------------------------------------------------      0
PTS18CBA_START_2(reversed)      ------------------------------------------------------------      0
CAT_F_(reversed)                AAACCACTCGTCGTCTTCAAACCGATAGCGTGTTTGGTCGGAAAGACACTTGTCACTGTC       94
P1R                             AAACCACTCGTCGTCTTCAAACCGATAGCGTGTTTGGTCGGAAAGACACTTGTCACTGTC      211
81_(reversed)                   ------------------------------------------------------------      0
80_term                         ------------------------------------------------------------      0
85                              ------------------------------------------------------------      0
CAT_R                           ------------------------------------------------------------      0
87_(reversed)                   AAACCACTCGTCGTCTTCAAACCGATAGCGTGTTTGGTCGGAAAGACACTTGTCACTGTC      485
expected_knockout               AAACCACTCGTCGTCTTCAAACCGATAGCGTGTTTGGTCGGAAAGACACTTGTCACTGTC      900
82a_(reversed)                  AAACCACTCGTCGTCTTCAAACCGATAGCGTGTTTGGTCGGAAAGACACTTGTCACTGTC      675
83b                             AAACCACTCGTCGTCTTCAAACCGATAGCGTGT----------------------------      641
```

```
PTS18CBA_START_1(reversed)      ------------------------------------------------------------          0
PTS18CBA_START_2(reversed)      ------------------------------------------------------------          0
CAT_F_(reversed)                CCACTAACCGGAGCAACGACTTCACCATCTTCGGGTTCAACAGCGTAACCATCACCTAAG          154
P1R                             CCACTAACCGGAGCAACGACTTCACCATCTTCGGGTTCAACAGCGTAACCATCACCTAAG          271
81_(reversed)                   ------------------------------------------------------------          0
80_term                         ------------------------------------------------------------          0
85                              ------------------------------------------------------------          0
CAT_R                           ------------------------------------------------------------          0
87_(reversed)                   CCACTAACCGGAGCAACGACTTCACCATCTTCGGGTTCAACAGCGTAACCATCACCTAAG          545
expected_knockout               CCACTAACCGGAGCAACGACTTCACCATCTTCGGGTTCAACAGCGTAACCATCACCTAAG          960
82a_(reversed)                  CC----------------------------------------------------------          677
83b                             ------------------------------------------------------------          641


PTS18CBA_START_1(reversed)      ------------------------------------------------------------          0
PTS18CBA_START_2(reversed)      ------------------------------------------------------------          0
CAT_F_(reversed)                AGTTTACCAGCAAACGTCTCATCCGCCTCATGCCCGGGCTGTACCGTTCGTATAGCATAC          214
P1R                             AGTTTACCAGCAAACGTCTCATCCGCCTCATGCCCGGGCTGTACCGTTCGTATAGCATAC          331
81_(reversed)                   ------------------------------------------------------------          0
80_term                         ------------------------------------------------------------          0
85                              ------------------------------------------------------------          0
CAT_R                           ------------------------------------------------------------          0
87_(reversed)                   AGTTTACCAGCAAACGTCTCATCCGCCTCATGCCCGGGCTGTACCGTTCGTATAGCATAC          605
expected_knockout               AGTTTACCAGCAAACGTCTCATCCGCCTCATGCCCGGGCTGTACCGTTCGTATAGCATAC          1020
82a_(reversed)                  ------------------------------------------------------------          677
83b                             ------------------------------------------------------------          641


PTS18CBA_START_1(reversed)      ------------------------------------------------------------          0
PTS18CBA_START_2(reversed)      ------------------------------------------------------------          0
CAT_F_(reversed)                ATTATACGAAGTTATGGGTCCCTAGCGGAATCGGCAACAGTTTAACGATTACTTCGGTAA          274
P1R                             ATTATACGAAGTTATGGGTCCCTAGCGGAATCGGCAACAGTTTAACGATTACTTCGGTAA          391
81_(reversed)                   ------------------------------------------------------------          0
80_term                         ------------------------------------------------------------          0
85                              ------------------------------------------------------------          0
CAT_R                           ------------------------------------------------------------          0
87_(reversed)                   ATTATACGAAGTTATGGGTCCCTAGCGGAATCGGCA----------------------------          641
expected_knockout               ATTATACGAAGTTATGGGTCCCTAGCGGAATCGGCAACAGTTTAACGATTACTTCGGTAA          1080
82a_(reversed)                  ------------------------------------------------------------          677
83b                             ------------------------------------------------------------          641


PTS18CBA_START_1(reversed)      ------------------------------------------------------------          0
PTS18CBA_START_2(reversed)      ------------------------------------------------------------          0
CAT_F_(reversed)                TGATTCCTAACGTGCCTTCCGAACCAATCATCAATTGCGTTTAGGATCCA--AGTACAGT          332
P1R                             TGATTCCTAACGTGCCTTCCGAACCAATCATCAATTGCGTTTAGGATCCA--AGTACAGT          449
81_(reversed)                   ------------------------------------------------------------          0
80_term                         ------------------------------------------------------------          0
85                              ------------------------------------------------------------          0
CAT_R                           ------------------------------------------------------------          0
87_(reversed)                   ------------------------------------------------------------          641
expected_knockout               TGATTCCTAACGTGCCTTCCGAACCAATCATCAATTGCGTTTAGGATCCGAAAGTACAGT          1140
82a_(reversed)                  ------------------------------------------------------------          677
83b                             ------------------------------------------------------------          641


PTS18CBA_START_1(reversed)      ------------------------------------------------------------          0
PTS18CBA_START_2(reversed)      ------------------------------------------------------------          0
CAT_F_(reversed)                CGGCATTATCTCATATTATAAAAGCCAGTCATTAGGCCTATCTGACAATTCCTGAATAGA          392
P1R                             CGGCATTATCTCATATTATAAAAGCCAGTCATTAGGCCTATCTGACAATTCCTGAATAGA          509
81_(reversed)                   ------------------------------------------------------------          0
80_term                         ------------------------------------------------------------          0
85                              ------------------------------------------------------------          0
CAT_R                           ------------------------------------------------------------          0
87_(reversed)                   ------------------------------------------------------------          641
expected_knockout               CGGCATTATCTCATATTATAAAAGCCAGTCATTAGGCCTATCTGACAATTCCTGAATAGA          1200
82a_(reversed)                  ------------------------------------------------------------          677
83b                             ------------------------------------------------------------          641


PTS18CBA_START_1(reversed)      ------------------------------------------------------------          0
PTS18CBA_START_2(reversed)      ------------------------------------------------------------          0
CAT_F_(reversed)                GTTCATAAACAATCCTGCATGATAACCATCACAAACAGAATGATGTACCTGTAAAGATAG          452
P1R                             GTTCATAAACAATCCTGCATGATAACCATCACAAACAGAATGATGTACCTGTAAAGATAG          569
81_(reversed)                   ------------------------------------------------------------          0
80_term                         ------------------------------------------------------------          0
85                              ------------------------------------------------------------          0
CAT_R                           ------------------------------------------------------------          0
87_(reversed)                   ------------------------------------------------------------          641
expected_knockout               GTTCATAAACAATCCTGCATGATAACCATCACAAACAGAATGATGTACCTGTAAAGATAG          1260
82a_(reversed)                  ------------------------------------------------------------          677
83b                             ------------------------------------------------------------          641


PTS18CBA_START_1(reversed)      ------------------------------------------------------------          0
PTS18CBA_START_2(reversed)      ------------------------------------------------------------          0
CAT_F_(reversed)                CGGTAAATATATTGAATTACCTTTATTAATGAATTTTCCTGCTGTAATAATGGGTAGAAG          512
P1R                             CGGTAAATATATTGAATTACCTTTATTAATGAATTTTCCTGCTGTAATAATGGGTAGAAG          629
81_(reversed)                   ------------------------------------------------------------          0
80_term                         ------------------------------------------------------------          0
85                              ------------------------------------------------------------          0
CAT_R                           ----------------------------------------------------ATGGGTAGAAG          11
87_(reversed)                   ------------------------------------------------------------          641
expected_knockout               CGGTAAATATATTGAATTACCTTTATTAATGAATTTTCCTGCTGTAATAATGGGTAGAAG          1320
82a_(reversed)                  ------------------------------------------------------------          677
83b                             ------------------------------------------------------------          641
```

```
PTS18CBA_START_1(reversed)      ---------------ATTGATATTTAAGTTAAACCCAGTAAATGAAGTCCATGGAATAAT          45
PTS18CBA_START_2(reversed)      ------------------------------------------------------------          0
CAT_F_(reversed)                GTAATTACTATTATTATTGATATTTAAGTTAAACCCAGTAAATGAAGTCCATGGAATAAT          572
P1R                             GTAATTACTATTATTATTGATATTTAAGTTAAACCCAGTAAATGAAGTCCATGGAATAAT          689
81_(reversed)                   ------------------------------------------------------------          0
80_term                         ------------------------------------------------------------          0
85                              ------------------------------------------------------------          0
CAT_R                           GTAATTACTATTATTATTGATATTTAAGTTAAACCCAGTAAATGAAGTCCATGGAATAAT          71
87_(reversed)                   ------------------------------------------------------------          641
expected_knockout               GTAATTACTATTATTATTGATATTTAAGTTAAACCCAGTAAATGAAGTCCATGGAATAAT          1380
82a_(reversed)                  ------------------------------------------------------------          677
83b                             ------------------------------------------------------------          641


PTS18CBA_START_1(reversed)      AGAAAGAGAAAAAGCATTTTCAGGTATAGGTGTTTTGGGAAACAATTTCCCCGAACCATT          105
PTS18CBA_START_2(reversed)      ------------------------------------------------------------          0
CAT_F_(reversed)                AGAAAGAGAAAAAGCATTTTCAGGTATAGGTGTTTTGGGAAACAATTTCCCCGAACCATT          632
P1R                             AGAAAGAGAAAAAGCATTTTCAGGTATAGGTGTTTTGGGAAACAATTTCCCCGAACCATT          749
81_(reversed)                   ------------------------------------------------------------          0
80_term                         ------------------------------------------------------------          0
85                              ------------------------------------------------------------          0
CAT_R                           AGAAAGAGAAAAAGCATTTTCAGGTATAGGTGTTTTGGGAAACAATTTCCCCGAACCATT          131
87_(reversed)                   ------------------------------------------------------------          641
expected_knockout               AGAAAGAGAAAAAGCATTTTCAGGTATAGGTGTTTTGGGAAACAATTTCCCCGAACCATT          1440
82a_(reversed)                  ------------------------------------------------------------          677
83b                             ------------------------------------------------------------          641


PTS18CBA_START_1(reversed)      ATATTTCTCTACATCAGAAAGGTATAAATCATAAAACTCTTTGAAGTCATTCTTTACAGG          165
PTS18CBA_START_2(reversed)      ----------------------------TCATAAAACTCTTTGAAGTCATTCTTTACAGG          32
CAT_F_(reversed)                ATATTTCTCTACATCAGAAAGGTATAAATCATAAAACTCTTTGAAGTCATTCTTTACAGG          692
P1R                             ATATTTCTCTACATCAGAAAGGTATAAATCATAAAACTCTTTGAAGTCATTCTTTACAGG          809
81_(reversed)                   ------------------------------------------------------------          0
80_term                         ------------------------------------------------------------          0
85                              ------------------------------------------------------------          0
CAT_R                           ATATTTCTCTACATCAGAAAGGTATAAATCATAAAACTCTTTGAAGTCATTCTTTACAGG          191
87_(reversed)                   ------------------------------------------------------------          641
expected_knockout               ATATTTCTCTACATCAGAAAGGTATAAATCATAAAACTCTTTGAAGTCATTCTTTACAGG          1500
82a_(reversed)                  ------------------------------------------------------------          677
83b                             ------------------------------------------------------------          641


PTS18CBA_START_1(reversed)      AGTCCAAATACCAGAGAATGTTTTAGATACACCATCAAAAATTGTATAAAGTGGCTCTAA          225
PTS18CBA_START_2(reversed)      AGTCCAAATACCAGAGAATGTTTTAGATACACCATCAAAAATTGTATAAAGTGGCTCTAA          92
CAT_F_(reversed)                AGTCCAAATACCAGAGAATGTTTTAGATACACCATCAAAAATTGTATAAAGTGGCTCTAA          752
P1R                             AGTCCA------------------------------------------------------          815
81_(reversed)                   ------------------------------------------------------------          0
80_term                         ------------------------------------------------------------          0
85                              ------------------------------------------------------------          0
CAT_R                           AGTCCAAATACCAGAGAATGTTTTAGATACACCATCAAAAATTGTATAAAGTGGCTCTAA          251
87_(reversed)                   ------------------------------------------------------------          641
expected_knockout               AGTCCAAATACCAGAGAATGTTTTAGATACACCATCAAAAATTGTATAAAGTGGCTCTAA          1560
82a_(reversed)                  ------------------------------------------------------------          677
83b                             ------------------------------------------------------------          641


PTS18CBA_START_1(reversed)      CTTATCCCAATAACCTAACTCTCCGTCGCTATTGTAACCAGTTCTAAAAGCTGTATTTGA          285
PTS18CBA_START_2(reversed)      CTTATCCCAATAACCTAACTCTCCGTCGCTATTGTAACCAGTTCTAAAAGCTGTATTTGA          152
CAT_F_(reversed)                CTTATCCCAATAACCTAACTCTCCGTCGCTATTGTAACCAGTTCTAAAAGCTGTATT---          809
P1R                             ------------------------------------------------------------          815
81_(reversed)                   ------------------------------------------------------------          0
80_term                         ------------------------------------------------------------          0
85                              ------------------------------------------------------------          0
CAT_R                           CTTATCCCAATAACCTAACTCTCCGTCGCTATTGTAACCAGTTCTAAAAGCTGTATTTGA          311
87_(reversed)                   ------------------------------------------------------------          641
expected_knockout               CTTATCCCAATAACCTAACTCTCCGTCGCTATTGTAACCAGTTCTAAAAGCTGTATTTGA          1620
82a_(reversed)                  ------------------------------------------------------------          677
83b                             ------------------------------------------------------------          641


PTS18CBA_START_1(reversed)      GTTTATCACCCTTGTCACTAAGAAAATAAATGCAGGGTAAAATTTATATCCTTCTTGTTT          345
PTS18CBA_START_2(reversed)      GTTTATCACCCTTGTCACTAAGAAAATAAATGCAGGGTAAAATTTATATCCTTCTTGTTT          212
CAT_F_(reversed)                ------------------------------------------------------------          809
P1R                             ------------------------------------------------------------          815
81_(reversed)                   ------------------------------------------------------------          0
80_term                         ------------------------------------------------------------          0
85                              ------------------------------------------------------------          0
CAT_R                           GTTTATCACCCTTGTCACTAAGAAAATAAATGCAGGGTAAAATTTATATCCTTCTTGTTT          371
87_(reversed)                   ------------------------------------------------------------          641
expected_knockout               GTTTATCACCCTTGTCACTAAGAAAATAAATGCAGGGTAAAATTTATATCCTTCTTGTTT          1680
82a_(reversed)                  ------------------------------------------------------------          677
83b                             ------------------------------------------------------------          641


PTS18CBA_START_1(reversed)      TATGTTTCGGTATAAAACACTAATATCAATTTCTGTGGTTATACTAAAAGTCGTTTGTTG          405
PTS18CBA_START_2(reversed)      TATGTTTCGGTATAAAACACTAATATCAATTTCTGTGGTTATACTAAAAGTCGTTTGTTG          272
CAT_F_(reversed)                ------------------------------------------------------------          809
P1R                             ------------------------------------------------------------          815
81_(reversed)                   ------------------------------------------------------------          0
80_term                         ------------------------------------------------------------          0
85                              ------------------------------------------------------------          0
CAT_R                           TATGTTTCGGTATAAAACACTAATATCAATTTCTGTGGTTATACTAAAAGTCGTTTGTTG          431
87_(reversed)                   ------------------------------------------------------------          641
expected_knockout               TATGTTTCGGTATAAAACACTAATATCAATTTCTGTGGTTATACTAAAAGTCGTTTGTTG          1740
82a_(reversed)                  ------------------------------------------------------------          677
83b                             ------------------------------------------------------------          641
```

```
PTS18CBA_START_1(reversed)      GTTCAAATAATGATTAAATATCTCTTTTCTCTTCCAATTGTCTAAATCAATTTTATTAAA        465
PTS18CBA_START_2(reversed)      GTTCAAATAATGATTAAATATCTCTTTTCTCTTCCAATTGTCTAAATCAATTTTATTAAA        332
CAT_F_(reversed)                ------------------------------------------------------------        809
P1R                             ------------------------------------------------------------        815
81_(reversed)                   ------------------------------------------------------------        0
80_term                         ------------------------------------------------------------        0
85                              ------------------------------------------------------------        0
CAT_R                           GTTCAAATAATGATTAAATATCTCTTTTCTCTTCCAATTGTCTAAATCAATTTTATTAAA        491
87_(reversed)                   ------------------------------------------------------------        641
expected_knockout               GTTCAAATAATGATTAAATATCTCTTTTCTCTTCCAATTGTCTAAATCAATTTTATTAAA        1800
82a_(reversed)                  ------------------------------------------------------------        677
83b                             ------------------------------------------------------------        641


PTS18CBA_START_1(reversed)      GTTCATTTGATATGCCTCCTAAATTTTTATCTACCTAGTATAGCATTTTGTGAAGTTTTT        525
PTS18CBA_START_2(reversed)      GTTCATTTGATATGCCTCCTAAATTTTTATCTACCTAGTATAGCATTTTGTGAAGTTTTT        392
CAT_F_(reversed)                ------------------------------------------------------------        809
P1R                             ------------------------------------------------------------        815
81_(reversed)                   ------------------------------------------------------------        0
80_term                         ------------------------------------------------------------        0
85                              ------------------------------------------------------------        0
CAT_R                           GTTCATTTGATATGCCTCCTAAATTTTTATCTACCTAGTATAGCATTTTGTGAAGTTTTT        551
87_(reversed)                   ------------------------------------------------------------        641
expected_knockout               GTTCATTTGATATGCCTCCTAAATTTTTATCTACCTAGTATAGCATTTTGTGAAGTTTTT        1860
82a_(reversed)                  ------------------------------------------------------------        677
83b                             ------------------------------------------------------------        641


PTS18CBA_START_1(reversed)      TTCTAGTCCAAGCTCACAAAAATCCAAAGTAACCGCTTTATTAAGCCATTCTTAAATAAA        585
PTS18CBA_START_2(reversed)      TTCTAGTCCAAGCTCACAAAAATCCAAAGTAACCGCTTTATTAAGCCATTCTTAAATAAA        452
CAT_F_(reversed)                ------------------------------------------------------------        809
P1R                             ------------------------------------------------------------        815
81_(reversed)                   ------------------------------------------------------------        0
80_term                         ------------------------------------------------------------        0
85                              ------------------------------------------------------------        0
CAT_R                           TTCTAGTCCAAGCTCACAAAAATCCAAAGTAACCGCTT-----------------------        589
87_(reversed)                   ------------------------------------------------------------        641
expected_knockout               TTCTAGTCCAAGCTCACAAAAATCCAAAGTAACCGCTTTATTAAGCCATTC---------        1911
82a_(reversed)                  ------------------------------------------------------------        677
83b                             ------------------------------------------------------------        641


PTS18CBA_START_1(reversed)      AATAAAAAAAGATTAATAGCTAAAACTATTAATCTTATCATATCCCGAGGACCGAATTTA        645
PTS18CBA_START_2(reversed)      AATAAAAAAAGATTAATAGCTAAAACTATTAATCTTATCATATCCCGAGGACCGAATTTA        512
CAT_F_(reversed)                ------------------------------------------------------------        809
P1R                             ------------------------------------------------------------        815
81_(reversed)                   ------------------------------------------------------------        0
80_term                         ------------------------------------------------------------        0
85                              ----------------------------------------------------AGGACCGAATTTA        13
CAT_R                           ------------------------------------------------------------        589
87_(reversed)                   ------------------------------------------------------------        641
expected_knockout               ------------------------------------TTATCATATCCCGAGGACCGAATTTA        1937
82a_(reversed)                  ------------------------------------------------------------        677
83b                             ------------------------------------------------------------        641


PTS18CBA_START_1(reversed)      AACTGCATGGTACCTTACTTAACTTCGCGACCATCGGCTAAAACGACCTTAACGCCTAAG        705
PTS18CBA_START_2(reversed)      AACTGCATGGTACCTTACTTAACTTCGCGACCATCGGCTAAAACGACCT-----------        561
CAT_F_(reversed)                ------------------------------------------------------------        809
P1R                             ------------------------------------------------------------        815
81_(reversed)                   ------------------------------------------------------------        0
80_term                         ------------------------------------------------------------        0
85                              AACTGCATGGTACCTTACTTAACTTCGCGACCATCGGCTAAAACGACCTTAACGCCTAAG        73
CAT_R                           ------------------------------------------------------------        589
87_(reversed)                   ------------------------------------------------------------        641
expected_knockout               AACTGCATGGTACCTTACTTAACTTCGCGACCATCGGCTAAAACGACCTTAACGCCTAAG        1997
82a_(reversed)                  ------------------------------------------------------------        677
83b                             ------------------------------------------------------------        641


PTS18CBA_START_1(reversed)      ACATTGTCATAACTTCGTATAGCATACATTATACGAACGG----------------------        745
PTS18CBA_START_2(reversed)      ------------------------------------------------------------        561
CAT_F_(reversed)                ------------------------------------------------------------        809
P1R                             ------------------------------------------------------------        815
81_(reversed)                   ------------------------------------------------------------        0
80_term                         ------------------------------------------------------------        0
85                              ACATTGTCATAACTTCGTATAGCATACATTATACGAACGGTAGATTTAAATTGTTTACGG        133
CAT_R                           ------------------------------------------------------------        589
87_(reversed)                   ------------------------------------------------------------        641
expected_knockout               ACATTGTCATAACTTCGTATAGCATACATTATACGAACGGTAGATTTAAATTGTTTACGG        2057
82a_(reversed)                  ------------------------------------------------------------        677
83b                             ------------------------------------------------------------        641


PTS18CBA_START_1(reversed)      ------------------------------------------------------------        745
PTS18CBA_START_2(reversed)      ------------------------------------------------------------        561
CAT_F_(reversed)                ------------------------------------------------------------        809
P1R                             ------------------------------------------------------------        815
81_(reversed)                   ------------------------------------------------------------        0
80_term                         ------------------------------------------------------------        0
85                              CAGCATCAATGACTGACCGATTTTCTGAAAATATGTCTTCATACCAATTTCTCCTCATTC        193
CAT_R                           ------------------------------------------------------------        589
87_(reversed)                   ------------------------------------------------------------        641
expected_knockout               CAGCATCAATGACTGACCGATTTTCTGAAAATATGTCTTCATACCAATTTCTCCTCATTC        2117
82a_(reversed)                  ------------------------------------------------------------        677
83b                             ------------------------------------------------------------        641
```

```
PTS18CBA_START_1(reversed)      ------------------------------------------------------------  745
PTS18CBA_START_2(reversed)      ------------------------------------------------------------  561
CAT_F_(reversed)                ------------------------------------------------------------  809
P1R                             ------------------------------------------------------------  815
81_(reversed)                   ------------------------------------------------------------  0
80_term                         ----------------------------------------CATTGCAAGTTTACAAGGTT  20
85                              TCTTTTTAAGCACGTTATAAATCCTATCACAACATGTTTTCATTGCAAGTTTACAAGGTT  253
CAT_R                           ------------------------------------------------------------  589
87_(reversed)                   ------------------------------------------------------------  641
expected_knockout               TCTTTTTAAGCACGTTATAAATCCTATCACAACATGTTTTCATTGCAAGTTTACAAGGTT  2177
82a_(reversed)                  ------------------------------------------------------------  677
83b                             ------------------------------------------------------------  641


PTS18CBA_START_1(reversed)      ------------------------------------------------------------  745
PTS18CBA_START_2(reversed)      ------------------------------------------------------------  561
CAT_F_(reversed)                ------------------------------------------------------------  809
P1R                             ------------------------------------------------------------  815
81_(reversed)                   ----------------AATGTCAAAGAAACCGATTTCAACCTTACTAAATTAAGTTTTT  43
80_term                         ATCTAAGTGTATAGACCAATGTCAAAGAAACCGATTTCAACCTTACTAAATTAAGTTTTT  80
85                              ATCTAAGTGTATAGACCAATGTCAAAGAAACCGATTTCAACCTTACTAAATTAAGTTTTT  313
CAT_R                           ------------------------------------------------------------  589
87_(reversed)                   ------------------------------------------------------------  641
expected_knockout               ATCTAAGTGTATAGACCAATGTCAAAGAAACCGATTTCAACCTTACTAAATTAAGTTTTT  2237
82a_(reversed)                  ------------------------------------------------------------  677
83b                             ------------------------------------------------------------  641


PTS18CBA_START_1(reversed)      ------------------------------------------------------------  745
PTS18CBA_START_2(reversed)      ------------------------------------------------------------  561
CAT_F_(reversed)                ------------------------------------------------------------  809
P1R                             ------------------------------------------------------------  815
81_(reversed)                   CCGCCCACAAAAAAAGTTCTGAAAATACTCAGAACTTTTTCATTAATAATCTTATTTTTC  103
80_term                         CCGCCCACAAAAAAAGTTCTGAAAATACTCAGAACTTTTTCATTAATAATCTTATTTTTC  140
85                              CCGCCCACAAAAAAAGTTCTGAAAATACTCAGAACTTTTTCATTAATAATCTTATTTTTC  373
CAT_R                           ------------------------------------------------------------  589
87_(reversed)                   ------------------------------------------------------------  641
expected_knockout               CCGCCCACAAAAAAAGTTCTGAAAATACTCAGAACTTTTTCATTAATAATCTTATTTTTC  2297
82a_(reversed)                  ------------------------------------------------------------  677
83b                             ------------------------------------------------------------  641


PTS18CBA_START_1(reversed)      ------------------------------------------------------------  745
PTS18CBA_START_2(reversed)      ------------------------------------------------------------  561
CAT_F_(reversed)                ------------------------------------------------------------  809
P1R                             ------------------------------------------------------------  815
81_(reversed)                   AGCAACGGGGATTTCCGGGTGCCGGTATTGCCGCATCATTTCTGCTAAATGGACTTCGTG  163
80_term                         AGCAACGGGGATTTCCGGGTGCCGGTATTGCCGCATCATTTCTGCTAAATGGACTTCGTG  200
85                              AGCAACGGGGATTTCCGGGTGCCGGTATTGCCGCATCATTTCTGCTAAATGGACTTCGTG  433
CAT_R                           ------------------------------------------------------------  589
87_(reversed)                   ------------------------------------------------------------  641
expected_knockout               AGCAACGGGGATTTCCGGGTGCCGGTATTGCCGCATCATTTCTGCTAAATGGACTTCGTG  2357
82a_(reversed)                  ------------------------------------------------------------  677
83b                             ------------------------------------------------------------  641


PTS18CBA_START_1(reversed)      ------------------------------------------------------------  745
PTS18CBA_START_2(reversed)      ------------------------------------------------------------  561
CAT_F_(reversed)                ------------------------------------------------------------  809
P1R                             ------------------------------------------------------------  815
81_(reversed)                   AGTCACAAACTCATGCACGCGACAAACGTAGTCATGTTCCCGCCCAATCCTAGCAATATA  223
80_term                         AGTCACAAACTCATGCACGCGACAAACGTAGTCATGTTCCCGCCCAATCCTAGCAATATA  260
85                              AGTCACAAACTCATGCACGCGACAAACGTAGTCATGTTCCCGCCCAATCCTAGCAATATA  493
CAT_R                           ------------------------------------------------------------  589
87_(reversed)                   ------------------------------------------------------------  641
expected_knockout               AGTCACAAACTCATGCACGCGACAAACGTAGTCATGTTCCCGCCCAATCCTAGCAATATA  2417
82a_(reversed)                  ------------------------------------------------------------  677
83b                             ------------------------------------------------------------  641


PTS18CBA_START_1(reversed)      ------------------------------------------------------------  745
PTS18CBA_START_2(reversed)      ------------------------------------------------------------  561
CAT_F_(reversed)                ------------------------------------------------------------  809
P1R                             ------------------------------------------------------------  815
81_(reversed)                   GCCATTAATATAATCAACTTCGGTTGGCCGTCCCTTTGAGAAGTCTTGATACATTGATGG  283
80_term                         GCCATTAATATAATCAACTTCGGTTGGCCGTCCCTTTGAGAAGTCTTGATACATTGATGG  320
85                              GCCATTAATATAATCAACTTCGGTTGGCCGTCCCTTTGAGAAGTCTTGATACATTGATGG  553
CAT_R                           ------------------------------------------------------------  589
87_(reversed)                   ------------------------------------------------------------  641
expected_knockout               GCCATTAATATAATCAACTTCGGTTGGCCGTCCCTTTGAGAAGTCTTGATACATTGATGG  2477
82a_(reversed)                  ------------------------------------------------------------  677
83b                             ------------------------------------------------------------  641


PTS18CBA_START_1(reversed)      ------------------------------------------------------------  745
PTS18CBA_START_2(reversed)      ------------------------------------------------------------  561
CAT_F_(reversed)                ------------------------------------------------------------  809
P1R                             ------------------------------------------------------------  815
81_(reversed)                   ATAATGGTACTTATAGGCTCGACTGACCGTTTCAACCGAGTCGATTTCTTCTTGCCGCGT  343
80_term                         ATAATGGTACTTATAGGCTCGACTGACCGTTTCAACCGAGTCGATTTCTTCTTGCCGCGT  380
85                              ATAATGGTACTTATAGGCTCGACTGACCGTTTCAACCGAGTCGATTTCTTCTTGCCGCGT  613
CAT_R                           ------------------------------------------------------------  589
87_(reversed)                   ------------------------------------------------------------  641
expected_knockout               ATAATGGTACTTATAGGCTCGACTGACCGTTTCAACCGAGTCGATTTCTTCTTGCCGCGT  2537
82a_(reversed)                  ------------------------------------------------------------  677
83b                             ------------------------------------------------------------  641
```

```
PTS18CBA_START_1(reversed)      ------------------------------------------------------------  745
PTS18CBA_START_2(reversed)      ------------------------------------------------------------  561
CAT_F_(reversed)                ------------------------------------------------------------  809
P1R                             ------------------------------------------------------------  815
81_(reversed)                   CTCAATGAGTTTGATGCCTGCCCTTTCGCAAGCATCATAGGCTTCGTTGAACATTTGAGT  403
80_term                         CTCAATGAGTTTGATGCCTGCCCTTTCGCAAGCATCATAGGCTTCGTTGAACATTTGAGT  440
85                              CTCAATGAGTTTGATGCCTGCCCTTTCGCAAGCATCATAGGCTTCGTTGAACATTTGA--  671
CAT_R                           ------------------------------------------------------------  589
87_(reversed)                   ------------------------------------------------------------  641
expected_knockout               CTCAATGAGTTTGATGCCTGCCCTTTCGCAAGCATCATAGGCTTCGTTGAACATTTGAGT  2597
82a_(reversed)                  ------------------------------------------------------------  677
83b                             ------------------------------------------------------------  641


PTS18CBA_START_1(reversed)      ------------------------------------------------------------  745
PTS18CBA_START_2(reversed)      ------------------------------------------------------------  561
CAT_F_(reversed)                ------------------------------------------------------------  809
P1R                             ------------------------------------------------------------  815
81_(reversed)                   CGCCATATCCTTGACGCCCGGATATTCGATGAATTGGCCCATCTGAATCTCGAACATCGT  463
80_term                         CGCCATATCCTTGACGCCCGGATATTCGATGAATTGGCCCATCTGAATCTCGAACATCGT  500
85                              ------------------------------------------------------------  671
CAT_R                           ------------------------------------------------------------  589
87_(reversed)                   ------------------------------------------------------------  641
expected_knockout               CGCCATATCCTTGACGCCCGGATATTCGATGAATTGGCCCATCTGAATCTCGAACATCGT  2657
82a_(reversed)                  ------------------------------------------------------------  677
83b                             ------------------------------------------------------------  641


PTS18CBA_START_1(reversed)      ------------------------------------------------------------  745
PTS18CBA_START_2(reversed)      ------------------------------------------------------------  561
CAT_F_(reversed)                ------------------------------------------------------------  809
P1R                             ------------------------------------------------------------  815
81_(reversed)                   GCACAAACTATTGGTGACCGCGTTAAACACAACTTTGGAAATGCACATGCCCATAAAATT  523
80_term                         GCACAAACTATTGGTGACCGCGTTAAACACAACTTTGGAAATGCACATGCCCATAAAATT  560
85                              ------------------------------------------------------------  671
CAT_R                           ------------------------------------------------------------  589
87_(reversed)                   ------------------------------------------------------------  641
expected_knockout               GCACAAACTATTGGTGACCGCGTTAAACACAACTTTGGAAATGCACATGCCCATAAAATT  2717
82a_(reversed)                  ------------------------------------------------------------  677
83b                             ------------------------------------------------------------  641


PTS18CBA_START_1(reversed)      ------------------------------------------------------------  745
PTS18CBA_START_2(reversed)      ------------------------------------------------------------  561
CAT_F_(reversed)                ------------------------------------------------------------  809
P1R                             ------------------------------------------------------------  815
81_(reversed)                   ATCAGACCAAATCGGATTTAAATCAGCCGCCTTAAAATCGGCCATCACTTGCTTAGTAGT  583
80_term                         ATCAGACCAAATCGGATTTAAATCAGCCGCCTTAAAATCGGCCATCACTTGCTTAGTAGT  620
85                              ------------------------------------------------------------  671
CAT_R                           ------------------------------------------------------------  589
87_(reversed)                   ------------------------------------------------------------  641
expected_knockout               ATCAGACCAAATCGGATTTAAATCAGCCGCCTTAAAATCGGCCATCACTTGCTTAGTAGT  2777
82a_(reversed)                  ------------------------------------------------------------  677
83b                             ------------------------------------------------------------  641


PTS18CBA_START_1(reversed)      ------------------------------------------------------------  745
PTS18CBA_START_2(reversed)      ------------------------------------------------------------  561
CAT_F_(reversed)                ------------------------------------------------------------  809
P1R                             ------------------------------------------------------------  815
81_(reversed)                   CGTATCAATGGGACCTGCATACTTACTCATATGCATCGCCTCACTACCCTTAGGCCCCAT  643
80_term                         CGTATCAATGGGACCTGCATACTTACTCATATGCATCGCCTCACTACCCTTAGGCCCCAT  680
85                              ------------------------------------------------------------  671
CAT_R                           ------------------------------------------------------------  589
87_(reversed)                   ------------------------------------------------------------  641
expected_knockout               CGTATCAATGGGACCTGCATACTTACTCATATGCATCGCCTCACTACCCTTAGGCCCCAT  2837
82a_(reversed)                  ------------------------------------------------------------  677
83b                             ------------------------------------------------------------  641


PTS18CBA_START_1(reversed)      ------------------------------------------------------------  745
PTS18CBA_START_2(reversed)      ------------------------------------------------------------  561
CAT_F_(reversed)                ------------------------------------------------------------  809
P1R                             ------------------------------------------------------------  815
81_(reversed)                   AAAATCAACCGCACCCGGACCATTCAAGACCGTTGCAATCATCGCCG-------------  690
80_term                         AAAATCA----------------------------------------------------  687
85                              ------------------------------------------------------------  671
CAT_R                           ------------------------------------------------------------  589
87_(reversed)                   ------------------------------------------------------------  641
expected_knockout               AAAATCAACCGCACCCGGACCATTCAAGACCGTTGCAATCATCGCCGTACCACCAATAAT  2897
82a_(reversed)                  ------------------------------------------------------------  677
83b                             ------------------------------------------------------------  641


PTS18CBA_START_1(reversed)      ------------------------------------------------------------  745
PTS18CBA_START_2(reversed)      ------------------------------------------------------------  561
CAT_F_(reversed)                ------------------------------------------------------------  809
P1R                             ------------------------------------------------------------  815
81_(reversed)                   ------------------------------------------------------------  690
80_term                         ------------------------------------------------------------  687
85                              ------------------------------------------------------------  671
CAT_R                           ------------------------------------------------------------  589
87_(reversed)                   ------------------------------------------------------------  641
expected_knockout               GTGATTTTCCGGGAAATATTGTGCAATCTTTTCAAAGTGCCCCATCCCGTTCATTGCGGC  2957
82a_(reversed)                  ------------------------------------------------------------  677
83b                             ------------------------------------------------------------  641
```

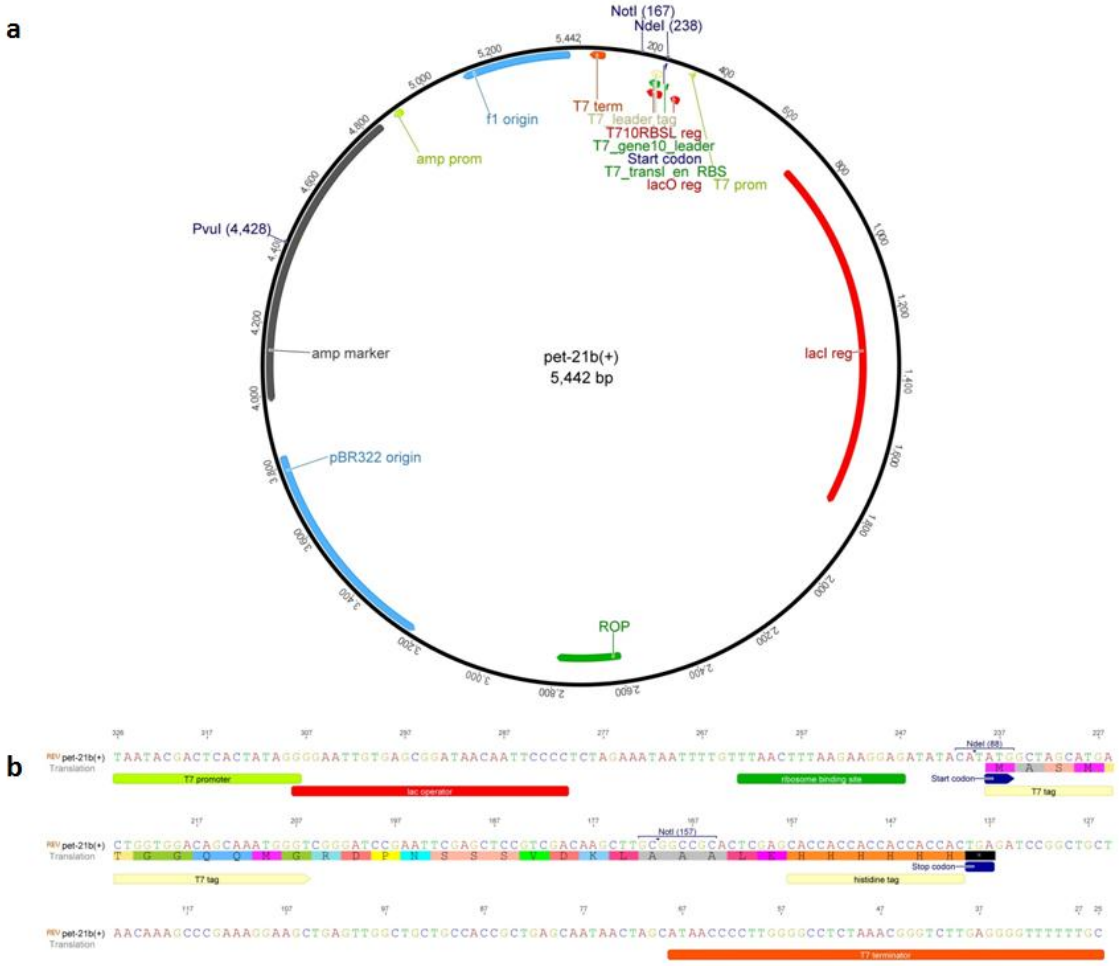
```
PTS18CBA_START_1(reversed)      ------------------------------------------------------------          745
PTS18CBA_START_2(reversed)      ------------------------------------------------------------          561
CAT_F_(reversed)                ------------------------------------------------------------          809
P1R                             ------------------------------------------------------------          815
81_(reversed)                   ------------------------------------------------------------          690
80_term                         ------------------------------------------------------------          687
85                              ------------------------------------------------------------          671
CAT_R                           ------------------------------------------------------------          589
87_(reversed)                   ------------------------------------------------------------          641
expected_knockout               GAAAACGTACTGGTCCTCATGAAATAACGACGCATCGCGCTTGAGCTCATCGGCTAACTG          3017
82a_(reversed)                  ------------------------------------------------------------          677
83b                             ------------------------------------------------------------          641


PTS18CBA_START_1(reversed)      ------------------------------------------------------------          745
PTS18CBA_START_2(reversed)      ------------------------------------------------------------          561
CAT_F_(reversed)                ------------------------------------------------------------          809
P1R                             ------------------------------------------------------------          815
81_(reversed)                   ------------------------------------------------------------          690
80_term                         ------------------------------------------------------------          687
85                              ------------------------------------------------------------          671
CAT_R                           ------------------------------------------------------------          589
87_(reversed)                   ------------------------------------------------------------          641
expected_knockout               CATTTGCTTCTTAAAAACAATCCAAACATCTGGATGACCCTGATATTCTTCAGGATAATA          3077
82a_(reversed)                  ------------------------------------------------------------          677
83b                             ------------------------------------------------------------          641


PTS18CBA_START_1(reversed)      ------------------------------------------------------------          745
PTS18CBA_START_2(reversed)      ------------------------------------------------------------          561
CAT_F_(reversed)                ------------------------------------------------------------          809
P1R                             ------------------------------------------------------------          815
81_(reversed)                   ------------------------------------------------------------          690
80_term                         ------------------------------------------------------------          687
85                              ------------------------------------------------------------          671
CAT_R                           ------------------------------------------------------------          589
87_(reversed)                   ------------------------------------------------------------          641
expected_knockout               AATATTAATTGGAACGACATGGTGATTTTGATGATCACGCGCCACATCAACGCCGCCCTG          3137
82a_(reversed)                  ------------------------------------------------------------          677
83b                             ------------------------------------------------------------          641


PTS18CBA_START_1(reversed)      ------------------------------------------------------------          745
PTS18CBA_START_2(reversed)      ------------------------------------------------------------          561
CAT_F_(reversed)                ------------------------------------------------------------          809
P1R                             ------------------------------------------------------------          815
81_(reversed)                   ------------------------------------------------------------          690
80_term                         ------------------------------------------------------------          687
85                              ------------------------------------------------------------          671
CAT_R                           ------------------------------------------------------------          589
87_(reversed)                   ------------------------------------------------------------          641
expected_knockout               CTCACGGACCTTTGCGACGTTTGGTTCCCACGTGTCGATAAAATCAACGTCAACACCGGC          3197
82a_(reversed)                  ------------------------------------------------------------          677
83b                             ------------------------------------------------------------          641


PTS18CBA_START_1(reversed)      ------------------------------------------------------------          745
PTS18CBA_START_2(reversed)      ------------------------------------------------------------          561
CAT_F_(reversed)                ------------------------------------------------------------          809
P1R                             ------------------------------------------------------------          815
81_(reversed)                   ------------------------------------------------------------          690
80_term                         ------------------------------------------------------------          687
85                              ------------------------------------------------------------          671
CAT_R                           ------------------------------------------------------------          589
87_(reversed)                   ------------------------------------------------------------          641
expected_knockout               GTTTTCTTGTAACATTACACCATACCGATATCCCATTGCGCCTGCGCCAATAATGCCATA          3257
82a_(reversed)                  ------------------------------------------------------------          677
83b                             ------------------------------------------------------------          641


PTS18CBA_START_1(reversed)      ----------------------------------          745
PTS18CBA_START_2(reversed)      ----------------------------------          561
CAT_F_(reversed)                ----------------------------------          809
P1R                             ----------------------------------          815
81_(reversed)                   ----------------------------------          690
80_term                         ----------------------------------          687
85                              ----------------------------------          671
CAT_R                           ----------------------------------          589
87_(reversed)                   ----------------------------------          641
expected_knockout               TTTCATAAGTAAAACCATCCTTTCGTCTTGCCCA          3291
82a_(reversed)                  ----------------------------------          677
83b                             ----------------------------------          641
```

Appendix 17 Multiple alignment of PTS22CBA

CLUSTAL OMEGA (http://www.ebi.ac.uk/Tools/msa/clustalo/) multiple sequence alignment performed with default settings on PTS22CBA from *L. plantarum* NC8, *L. plantarum* ATCC 8014 and *L. plantarum* subsp. *plantarum* ATCC 14917

```
8014        --MKNYLQRMGRSLQLPVAVLPAAALLVGIGNWWASASNDIVAHFLQAGGNAVLSQLPLL
14917       --MKNYLQRMGRSLQLPVAVLPAAALLVGIGNWWASASNDIVAHFLQAGGNAVLSQLPLL
NC8         MYMKNYLQRMGRSLQLPVAVLPAAALLVGIGNWWASASNDIVAHFLQAGGNAVLSQLPLL
              ************************************************************

8014        FAVGLALGMSKDKDGAAALAGLVAFELPTNVLKPESVATLLNIKVATVNPAFGQIGNVFI
14917       FAVGLALGMSKDKDGAAALAGLVAFELPTNVLKPESVATLLNIKVATVNPAFGQIGNVFI
NC8         FAVGLALGMSKDKDGAAALAGLVAFELPTNVLKPESVATLLNIKVATVNPAFGQIGNVFI
            ************************************************************

8014        GIISGLIAAALYNRFHETKLPMALSFFSGKRLVPILAALTMLLVSVALLIVWPPVYDALV
14917       GIISGLIAAALYNRFHETKLPMALSFFSGKRLVPILAALTMLLVSVALLIVWPPVYDALV
NC8         GIISGLIAAALYNRFHETKLPMALSFFSGKRLVPILAALTMLLVSVALLIVWPPVYDALV
            ************************************************************

8014        TFGKFIVGLGAVGAGLYGFFNRLLIPTGLHQALNSVFWFDVAGINDIGKFLAHKGIKGVT
14917       TFGKFIVGLGAVGAGLYGFFNRLLIPTGLHQALNSVFWFDVAGINDIGKFLAHKGIKGVT
NC8         TFGKFIVGLGAVGAGLYGFFNRLLIPTGLHQALNSVFWFDVAGINDIGKFLAHKGIKGVT
            ************************************************************

8014        GMYQAGFFPVMMFGLPAGAYAIYRNALPERKVETASLMMAGAFASFFTGVTEPLEFSFMF
14917       GMYQAGFFPVMMFGLPAGAYAIYRNALPERKVETASLMMAGAFASFFTGVTEPLEFSFMF
NC8         GMYQAGFFPVMMFGLPAGAYAIYRNALPERKVETASLMMAGAFASFFTGVTEPLEFSFMF
            ************************************************************

8014        VAWPLYVLHAIFTGLSLAFAAFMHWTAGFAFSAGLVDYILSLKNPIANQPLMLILQGLVL
14917       VAWPLYVLHAIFTGLSLAFAAFMHWTAGFAFSAGLVDYILSLKNPIANQPLMLILQGLVL
NC8         VAWPLYVLHAIFTGLSLAFAAFMHWTAGFAFSAGLVDYILSLKNPIANQPLMLILQGLVL
            ************************************************************

8014        AVIYYFGFDFAIKKFHLMTPGREPVTADDADDLAIATDADDDKYTRQAKQIYAALGGADN
14917       AVIYYFGFDFAIKKFHLMTPGREPVTADDADDLAIATDADDDKYTRQAKQIYAALGGADN
NC8         AVIYYFGFDFAIKKFHLMTPGREPVTADDADDLAIATDADDDKYTRQAKQIYAALGGADN
            ************************************************************

8014        LTVVDNCTTRLRLQLADTSTINEAAIKHSGAAGINKLDDHNLQIIIGTEVQFVADVLSHL
14917       LTVVDNCTTRLRLQLADTSTINEAAIKHSGAAGINKLDDHNLQIIIGTEVQFVADALSHL
NC8         LTVVDNCTTRLRLQLADTSTINEAAIKHSGAAGINKLDDHNLQIIMGTEVQFVADALSHL
            ********************************************:*********.****

8014        KLTNAPITKTTTSDEAIATAPTATNITSGVTTDFYSVANGQYVDIENVADNTFAKKMLGD
14917       KLTNAPITKTTTSNEAVATAPTATNITSGVTTDFYSVANGQYVDIENVDDDTFAKKMLGD
NC8         KLTNAPITKTTTSNEAVATAPTATNITSGVTTDFYSVANGQYVDIENVDDDTFAKKMLGD
            *************:**:****************************** *:*********

8014        GFAIDPADGTITTPVDGTVSTVFPTKHAIGFKTDSGLEILLHMGIDTVELNGAPFEVLVQ
14917       GFAIDPADGTITAPVDGTVSTVFPTKHAIGFKTDSGLEILLHMGIDTVELNGAPFEVLVQ
NC8         GFAIDPADGTITAPVDGTVSTVFPTKHAIGFKTDSGLEILLHMGIDTVELNGAPFEVLVQ
            ************:***********************************************

8014        TGQPVHHGEVVAKADLAAIKAAGKATTMMVIITNMDAVNLMKFKMLTNQVKVSDEIMLVT
14917       TDQPVHHGEVVAKADLAAIKAAGKATTMMVIITNMDAVNLMKFKMLTNQVKVSDEIMLVT
NC8         TDQPVHHGEVVAKADLAAIKAAGKATTMMVIITNMDAVNLMKFKMLTNQVKVSDEIMLVT
            * ************************************************************
8014        TK
14917       TK
NC8         TK
            **
```

Appendix 18 Plasmid map of expression vector pET-21b(+)



a) Schematic diagram of Novagen® T7 expression plasmid pET-21b(+) (EMD Millipore; Billerica, USA). Restriction enzyme recognition sites used in this study are shown (dark blue). The selectable marker ampicillin resistance (Amp^R) encoded by *amp* (dark gray) is under the control of the amp promoter (light green). b) Multiple cloning site where cloning was performed. Restriction digestion with *Not*I and *Nde*I result in the loss of the T7 tag and eleven downstream codons. This allows the *pts18CBA* coding sequence to be cloned directly into the vector start codon (position 238) and results in a five amino acid gap between the C-terminus of the PST18CBA and the histidine tag.

Appendix 19 Results of Mascot Database Search

The two replicates are provided in two folders in electronic material. Replicate 1 with the peptide score greater than the threshold is folder 'KD_InGel_1L_(Mascot_Search_Results)'

Appendix 20 Amino acid sequence of rPTS18CBA

>L_plantarum_8014_PTS18CBA_hex_histag_c_term|pred_pI_6.46_|pred_avg_Mw_71601.63

MKTYFQKIGQSLMLPIATLPAAAILVGIGNYLPKQWLFANYLIQGGNVVLNNLALLFAVGLAIGMSVNKDGAAAIAGLIAF
EVPVMVLKPATLATMLNVKVSQINPAFSALDNNVLIGISAGLIAAALYNRFHEVKLPMALSFFSGKRLVPIMAAFVMLIVT
AVLYLVWPFVYDAIVLFATGISKLGFVGAGLYGFFNRLLIPTGLHHALNSVFWYNVAGINDIGNFWASHGVKGITGMYEA
GFFPIMMFGLPAGAYAIYRNARPERKKEVGSLMLAGAFASFFTGVTEPLEFSFMFVAWPLYLLHAVFMGLSLGFAALMH
WTASFSFSGGLVDYLLSFRMPLANQPYMLLVQGLVMAVIYYFGFDFAIKRFNLKTPGREVVSADVDGVGAPASPAVAVA
ATDDKYMRQAKQIYAAIGGHDNISVINNCTTRLRLQ**L**KDTEKVDQPAVMAAGVPGLNVLDVHNIHIVIGTEVQFVAEALQ
KLFSGQVATTPASDAESKAPVEPQTATVTEAPVTTILRAPATGQLMPISTVADETFAGKLLGDGYAVEPEDGEVVAPVSGT
VTSVFPTKHAIGLKTTSGLEVLLHMGINTVEMNGTPFKLHVATGDEIAAGTAVATVDLAAIKSAGKATTMMVVITNMDH
VNKLTLNPTGHVTSGDLIGAAEAAALEHHHHHH

Appendix 21 Nucleotide sequence of *pts18CBA* from *L. plantarum* ATCC 8014

>Lactobacillus_plantarum_ATCC_8014_EV52_01314|1989_bp

ATGAAGACATATTTTCAGAAAATCGGTCAGTCATTGATGCTACCGATTGCAACGTTGCCCGCAGCAGCGATTTTGGTT
GGGATTGGGAATTACCTGCCAAAACAGTGGCTGTTTGCAAATTACTTGATTCAGGGTGGTAACGTTGTCCTGAATAAC
TTAGCCTTGCTGTTTGCAGTCGGCTTAGCTATCGGGATGTCAGTCAATAAGGACGGCGCGGCAGCTATTGCTGGTTTG
ATCGCGTTTGAAGTTCCGGTGATGGTTTTAAAACCAGCGACATTAGCGACGATGCTGAATGTTAAAGTGAGTCAGAT
CAATCCAGCATTTAGCGCGTTAGATAACAATGTGCTGATTGGAATCAGTGCTGGACTGATTGCGGCTGCGCTCTATAA
TCGGTTCCACGAAGTAAAATTACCAATGGCGCTCTCATTCTTTAGTGGTAAACGCCTAGTGCCAATTATGGCTGCTTT
TGTGATGCTGATTGTGACGGCGGTATTGTACCTTGTTTGGCCGTTTGTCTACGATGCCATCGTCCTCTTTGCAACGGGA
ATTTCAAAATTAGGTTTCGTGGGGGCTGGTCTATATGGCTTTTTCAATCGTTTATTGATTCCGACTGGTTTGCACCACG
CCTTAAATTCCGTATTCTGGTATAACGTTGCGGGGATCAATGATATTGGCAATTTCTGGGCCAGCCATGGTGTTAAAG
GAATCACAGGGATGTATGAAGCTGGATTCTTCCCAATTATGATGTTTGGCTTGCCAGCTGGTGCGTATGCAATTTATC
GTAATGCACGACCGGAACGAAAAAAAGAAGTGGGTTCATTGATGTTAGCTGGTGCGTTTGCGTCGTTCTTTACGGGG
GTGACCGAACCGCTTGAATTTTCATTCATGTTCGTGGCATGGCCGCTCTACTTATTACATGCCGTTTTTATGGGACTAT
CATTAGGATTCGCCGCGTTAATGCACTGGACCGCCAGCTTTTCGTTTAGTGGCGGGTTAGTCGATTATTTATTGAGTTT
TCGGATGCCGTTAGCCAACCAACCCTATATGTTACTGGTTCAAGGGCTGGTGATGGCCGTCATTTATTACTTTGGTTTT
GACTTTGCAATCAAACGGTTTAATTTGAAGACGCCTGGTCGTGAAGTTGTGTCTGCAGATGTCGATGGGGTGGGTGCA
CCTGCTAGTCCAGCGGTGGCCGTTGCAGCAACGGATGATAAGTATATGCGCCAAGCAAAGCAAATTTATGCAGCTAT
TGGTGGTCATGACAATATTAGTGTGATCAACAATTGTACGACGCGGCTGCGGTTACAACTTAAGGATACGGAAAAAG
TCGATCAGCCGGCCGTAATGGCTGCTGGCGTGCCTGGTTTGAACGTACTGGATGTTCATAACATTCACATTGTGATCG
GCACGGAAGTCCAGTTCGTCGCGGAGGCTTTACAAAAATTATTTTCTGGTCAGGTGGCGACGACCCCGGCATCTGAT
GCTGAATCAAAGGCCCCGGTTGAACCGCAAACTGCTACTGTAACGGAAGCGCCGGTGACAACGATTTTGCGGGCACC
GGCAACGGGACAATTAATGCCGATTAGTACGGTTGCGGATGAGACGTTTGCTGGTAAACTCTTAGGTGATGGTTACG

CTGTTGAACCCGAAGATGGTGAAGTCGTTGCTCCGGTTAGTGGGACAGTGACAAGTGTCTTTCCGACCAAACACGCT
ATCGGTTTGAAGACGACGAGTGGTTTAGAAGTCTTATTACACATGGGGATCAATACCGTGGAAATGAATGGTACGCC
GTTCAAGTTACACGTGGCAACGGGTGATGAAATTGCGGCCGGTACTGCGGTAGCGACGGTTGACTTGGCTGCTATCA
AGTCAGCCGGTAAAGCGACAACCATGATGGTCGTTATCACCAATATGGATCACGTTAACAAACTAACCCTTAATCCA
ACTGGTCATGTGACTAGCGGTGATTTGATTGGCGCAGCTGAATAA

Appendix 22 Amino acid sequence of *L. plantarum* ATCC 8014 PTS18CBA

>L_plantarum_ATCC_8014|MGS_EV52_01314

MKTYFQKIGQSLMLPIATLPAAAILVGIGNYLPKQWLFANYLIQGGNVVLNNLALLFAVGLAIGMSVNKDGAAAIAGLIAF
EVPVMVLKPATLATMLNVKVSQINPAFSALDNNVLIGISAGLIAAALYNRFHEVKLPMALSFFSGKRLVPIMAAFVMLIVT
AVLYLVWPFVYDAIVLFATGISKLGFVGAGLYGFFNRLLIPTGLHHALNSVFWYNVAGINDIGNFWASHGVKGITGMYEA
GFFPIMMFGLPAGAYAIYRNARPERKKEVGSLMLAGAFASFFTGVTEPLEFSFMFVAWPLYLLHAVFMGLSLGFAALMH
WTASFSFSGGLVDYLLSFRMPLANQPYMLLVQGLVMAVIYYFGFDFAIKRFNLKTPGREVVSADVDGVGAPASPAVAVA
ATDDKYMRQAKQIYAAIGGHDNISVINNCTTRLRLQLKDTEKVDQPAVMAAGVPGLNVLDVHNIHIVIGTEVQFVAEALQ
KLFSGQVATTPASDAESKAPVEPQTATVTEAPVTTILRAPATGQLMPISTVADETFAGKLLGDGYAVEPEDGEVVAPVSGT
VTSVFPTKHAIGLKTTSGLEVLLHMGINTVEMNGTPFKLHVATGDEIAAGTAVATVDLAAIKSAGKATTMMVVITNMDH
VNKLTLNPTGHVTSGDLIGAAE

Appendix 23 Nucleotide sequence of *pts18CBA* from *L. plantarum* subsp. *plantarum* ATCC 14917

>L_plantarum_subsp_plantarum_ATCC_14917_EV48_02088|1989_bp

ATGAAGACATATTTTCAGAAAATCGGTCAGTCATTGATGCTGCCGATTGCAACGTTGCCCGCAGCAGCGATTTTGGTT
GGGATTGGTAATTACCTGCCAAAACAGTGGCTGTTTGCAAATTACTTGATTCAGGGTGGTAACGTTGTCCTGAATAAC
TTAGCCTTGCTGTTTGCAGTCGGCTTAGCTATCGGGATGTCAGTCAATAAGGACGGCGCGGCAGCTATTGCTGGTTTG
ATCGCGTTTGAAGTTCCGGTGATGGTGTTAAAACCAGCGACATTAGCGACGATGCTGAATGTTAAAGTGAGTCAGAT
CAATCCAGCATTTAGCGCGTTAGATAACAATGTGCTGATTGGGATCAGTGCTGGACTGATTGCGGCTGCGCTCTATAA
TCGGTTCCACGAAGTAAAATTACCAATGGCGCTCTCATTCTTTAGTGGTAAACGCCTAGTGCCAATTATGGCTGCTTT
TGTGATGCTGATTGTGACGGCGGTATTGTACCTTGTTTGGCCGTTTGTCTACGATGCCATTGTCCTCTTTGCAACGGGG
ATTTCAAAATTAGGTTTCGTGGGGGCTGGTCTATATGGCTTTTTCAATCGTTTATTGATTCCAACTGGTTTGCACCACG
CCTTAAATTCCGTATTCTGGTATAACGTTGCTGGTATCAATGATATTGGCAATTTCTGGGCCAGCCATGGTGTTAAAG
GAATCACAGGGATGTATGAAGCTGGATTCTTCCCAATTATGATGTTTGGCTTGCCAGCTGGTGCGTATGCAATTTATC
GTAATGCACGACCGGAACGAAAAAAAGAAGTGGGTTCATTGATGTTAGCGGGTGCGTTTGCGTCGTTCTTTACGGGG
GTGACCGAACCGCTTGAATTTTCATTCATGTTCGTGGCATGGCCGCTCTACTTATTACATGCCGTTTTTATGGGACTAT
CATTAGGATTCGCCGCGTTAATGCACTGGACCGCCAGCTTTTCGTTTAGTGGCGGGTTAGTCGATTATTTATTGAGTTT
TCGGATGCCGTTAGCCAACCAACCCTATATGTTACTGGTTCAAGGGCTGGTGATGGCCGTCATTTATTACTTTGGTTTT
GACTTTGCAATCAAGCGGTTTAATTTGAAGACGCCTGGTCGTGAAGTTGTGTCTGCAGATGTCGATGGGGTGGGTGCA
CCTGCTAGTCCAGCGGTGGCCGTTGCAGCAACGGATGATAAATATATGCGCCAAGCAAAGCAAATTTATGCGGCTAT
TGGTGGTCATGACAATATTAGTGTGATCAACAATTGTACGACGCGGCTGCGGTTACAACTTAAGGATACGGAAAAAG
TTGATCAGCCGGCCGTAATGGCTGCTGGCGTGCCTGGTTTGAACGTACTGGATGTTCATAACATCCACATTGTGATCG
GTACGGAAGTCCAGTTCGTTGCGGAGGCTTTACAAAAATTATTTTCTGGTCAGGTGGCGACGACCCCGGCATCTGATG

CTGAATCAAAGGCCCCGCTTGAACCGCAAACTGCTACTGTAACGGAAGCGCCGGTGACAACGATTTTGCGGGCACCG
GCAACGGGACAATTAATGCCGATTAGTGCGGTTGCGGATGAGACGTTTGCTGGTAAACTCTTAGGTGATGGTTACGC
TGTTGAACCCGAAGATGGTGAAGTCGTTGCTCCGGTTAGTGGGACAGTGACAAGTGTCTTTCCGACCAAACACGCTA
TCGGTTTGAAGACGACGAGTGGTTTAGAAGTCTTATTACACATGGGGATCAATACCGTGGAAATGAATGGTACGCCG
TTCAAGTTACACGTGGCAACGGGTGATGAAATTGCGGCCGGTACTGCGGTAGCGACGGTTGACTTGGCTGCTATCAA
GTCAGCCGGTAAAGCGACAACCATGATGGTCGTTATCACCAATATGGATCACGTTAACAAACTAACCCTTAATCCAA
CTGGTCATGTGACTAGCGGTGATTTGATTGGCGCAGCTGAATAA

Appendix 24 Amino acid sequence of *L. plantarum* subsp. *plantarum* ATCC 14917 PTS18CBA

>L_plantarum_subsp_plantarum_ATCC_14917_PTS18CBA| MGS_EV48_02088

MKTYFQKIGQSLMLPIATLPAAAILVGIGNYLPKQWLFANYLIQGGNVVLNNLALLFAVGLAIGMSVNKDGAAAIAGLIAF
EVPVMVLKPATLATMLNVKVSQINPAFSALDNNVLIGISAGLIAAALYNRFHEVKLPMALSFFSGKRLVPIMAAFVMLIVT
AVLYLVWPFVYDAIVLFATGISKLGFVGAGLYGFFNRLLIPTGLHHALNSVFWYNVAGINDIGNFWASHGVKGITGMYEA
GFFPIMMFGLPAGAYAIYRNARPERKKEVGSLMLAGAFASFFTGVTEPLEFSFMFVAWPLYLLHAVFMGLSLGFAALMH
WTASFSFSGGLVDYLLSFRMPLANQPYMLLVQGLVMAVIYYFGFDFAIKRFNLKTPGREVVSADVDGVGAPASPAVAVA
ATDDKYMRQAKQIYAAIGGHDNISVINNCTTRLRLQLKDTEKVDQPAVMAAGVPGLNVLDVHNIHIVIGTEVQFVAEALQ
KLFSGQVATTPASDAESKAPLEPQTATVTEAPVTTILRAPATGQLMPISAVADETFAGKLLGDGYAVEPEDGEVVAPVSGT
VTSVFPTKHAIGLKTTSGLEVLLHMGINTVEMNGTPFKLHVATGDEIAAGTAVATVDLAAIKSAGKATTMMVVITNMDH
VNKLTLNPTGHVTSGDLIGAAE

Appendix 25 Nucleotide sequence of *purR* from *L. plantarum* ATCC 8014

>L_plantarum_ATCC_8014_ EV52_02355|837_bp

ATGAAAGTACGTAGAAGCGAACGTTTAATAGATATGACGCGCTATTTGTTGGAACGTCCCCATACATTAGTTTCATTG
ACATACTTCGCGGAACGGTACGCATCCGCAAAGTCGTCAATCAGTGAAGATTTAACGATTTTGAAGAAAGTATTTCA
GGCTCGGGGAACAGGAATCTTGGAAACCATTCCCGGTGCTGCTGGTGGTGCGCGGTTTATTCCTTATATTTTAAAGGA
AGAGGCCCAAGAATTTATTGACGAGATGACCACCAAGGTTGCCGACAAGAGTCGGGTGTTACCGGGTGGCTATGTTT
ATTTATCTGATATTTTAGGCCAACCAGATATTTTGCGCCAAGTGGGTCGGGTGATTGCGACTCAGTACTTGAATCAAA
AAATCGACGCTGTTATGACCGTTGCAACTAAAGGAATTCCGATTGCCCAAAGTGTTGCGACTTATTTAAACGTGCCGT
TTGTGATTGTCCGCAATGATTCTAAAATTACGGAAGGCTCGACCGTTAGTGTGAACTATGTCTCGGGTTCTTCAGAAC
GGATCAAGAAGATGGAACTTTCCAAGCGAAGCCTAACGGAAGGCGCCAACGTGCTCATCGTGGATGATTTCATGAAG
GGCGGCGGCACAATCAACGGGATGAAGACCTTGATTGAAGAATTTGATGCTAATTTAATTGGCATCACCGTTTTTGCC
GAAGCCAGTTTTGTTGGCAATCGTTTAATCGATGACTACACGTCGTTATTACGAGTAACCAACGTGAGTGATAGTGAC
AAGGCGATTAAAGTCGTTGCTGGTAACTATTTGGAGAAGGTTTTCGGTGCGGAAGTCTAG

Appendix 26 Amino acid sequence of PurR from *L. plantarum* ATCC 8014

>L_plantarum_ATCC_8014-PurR| MGS_EV52_02355

MKVRRSERLIDMTRYLLERPHTLVSLTYFAERYASAKSSISEDLTILKKVFQARGTGILETIPGAAGGARFIPYILKEEAQEFI
DEMTTKVADKSRVLPGGYVYLSDILGQPDILRQVGRVIATQYLNQKIDAVMTVATKGIPIAQSVATYLNVPFVIVRNDSKI
TEGSTVSVNYVSGSSERIKKMELSKRSLTEGANVLIVDDFMKGGGTINGMKTLIEEFDANLIGITVFAEASFVGNRLIDDYT
SLLRVTNVSDSDKAIKVVAGNYLEKVFGAEV

Appendix 27 List of chemicals and enzyme used in this study

Name	Source
3-(N-morpholino) propanesulfonic acid (MOPS)	Fisher Scientific: New Jersey USA
Acetic acid (glacial)	RCI Labscan; Bangkok, Thailand
Acetonitrile	Fisher Scientific: New Jersey USA
Acrylamide:Bis 29.1:0.9, 40 % (v/v)	Merck; Darmstadt
Agar Bacteriological (Agar No. 1)	Oxoid Ltd; Basingstoke, England
Ammonium bicarbonate	Sigma; Auckland, NZ
Ammonium persulfate	Sigma; Auckland, NZ
Ampicillin	USB; Ohio, USA
Boric acid	BDH ProLabo
Bromophenol blue	USB; Ohio, USA
Calcium Chloride dihydrate ($CaCl_2.2H_2O$)	Ajax UNIVAR; VIC, Australia
Chloramphenicol	Fisher Scientific: New Jersey USA
Coomassie Brilliant Blue R-250	Sigma; Auckland, NZ
Deoxyribonucleate 5'-oligonucleotidohydrolase (DNAse I)	Sigma; Auckland, NZ; #DN-25
Dithiothreitol (DTT)	Goldbio; St Louis, USA
*Ecl*136II	Fisher Scientific: New Jersey USA
*Eco*RV	Invitrogen™; Wilmington, DE, USA
EMSURE® ethanol	Merck Millipore; Billerica, USA
Erythromycin	Fisher Scientific: New Jersey USA
Ethanol	Ajax UNIVAR; VIC, Australia
Ethidium Bromide (EtBr)	USB; Ohio, USA
Ethylenediaminetetraacetic acid (EDTA)	Sigma; Auckland, NZ
Formic acid	Laboratory BDH Reagents; Poole, England
Glucose	Sigma; Auckland, NZ
Glycerol	Fisher Scientific: New Jersey USA
Glycine	Pure Science Limited
Hydrochloric Acid (HCL)	Merck; Darmstadt
HyAgarose™	HydraGene Co, Ltd

Invitrogen™ 1 kb Plus DNA Ladder	Invitrogen™; Wilmington, DE, USA
Iodoacetamide	GE
iProof™ High-Fidelity DNA Polymerase	Bio-Rad Laboratories; USA
Isopropanol	Scharlau; Sentmenat, Spain
Isopropyl-B-thio-galactoside (IPTG)	Goldbio; St Louis, USA
K_2HPO_4	Ajax UNIVAR; VIC, Australia
KH_2PO_4	Ajax UNIVAR; VIC, Australia
Lactobacilli de Man, Rogosa & Sharpe (MRS) broth	Acumedia® Neogen Corporation; Lansing, Michigan, USA
Lauria broth base	Invitrogen™; Wilmington, DE, USA
Lithium Acetate (LiAc)	Ajax UNIVAR; VIC, Australia
Magnesium Chloride ($MgCl_{2)}$	Ajax UNIVAR; VIC, Australia
Manganese Chloride tetrahydrate ($MnCl_2.4H_2O$)	May and Baker Ltd; England
Methanol	Ajax UNIVAR; VIC, Australia
N-acetylglucosamine (GlcNAc)	Sigma; Auckland, NZ
*Nde*I	New England Biolabs®; MA, USA
Orange G dye	Unknown Source
Phusion High Fidelity Polymerase	Thermo Scientific; Wilmington, DE, USA
*Pme*I	New England Biolabs®; MA, USA
Polyethylene Glycol 1500 (PEG-1500)	BDH ProLabo
Potassium Acetate (CH3CO2K)	Merck; Darmstadt
Roche® *Pst*I	Roche®; Basel, Switzerland
Roche® Taq DNA Polymerase	Roche®; Basel, Switzerland
Rubidium chloride (RbCl2)	Sigma; Auckland, NZ
*Sac*II	Invitrogen™; Wilmington, DE, USA
*Sma*I	Roche®; Basel, Switzerland
Sodium Chloride (NaCl)	Ajax UNIVAR; VIC, Australia
Sodium Dodecyl Sulfate (SDS)	Gibco BRL; Cergy Pontoise, Frace
Sucrose	Ajax UNIVAR; VIC, Australia
Tetramethylethylenediamine (TEMED)	Sigma; Auckland, NZ
Tris-Base	Pure Science Limited; Porirua, NZ
β-D-fructose	Sigma; Auckland, NZ