Investigation of the Molecular Basis for Transcriptional Regulation of Tn*916* and Macrolide Resistance in *Bacillus subtilis*

Thesis submitted by

Norashirene Binti Mohamad Jamil

For the degree of

DOCTOR OF PHILOSOPHY

University College London

Department of Microbial Diseases

UCL Eastman Dental Institute

256 Gray's Inn Road

London WC1X 8LD

UK

- 2020 -

Declaration

I, Norashirene Binti Mohamad Jamil, confirm that the work presented in this thesis is my own. Where information has been derived from other sources, I confirm that this has been indicated in the thesis.

Abstract

Antibiotic resistance (AR) is one of the most serious threats to modern healthcare today. To understand how resistance spreads, we need to investigate the genetic basis of transferable AR. Conjugative transposons (CTns) have acquired the vast majority of resistance genes we currently know about which makes them one of the major vectors involved in their spread. This study aims to investigate how Tn*916* and Tn*916*-like elements maintain their stability following insertion into a bacterial genome.

We identified putative rho-independent terminators upstream of the conjugation genes of Tn*2010,* Tn*5397,* Tn*6000,* Tn*6002*, Tn*6003,* Tn*6087* and Tn*916* and hypothesised that their role is to prevent transcriptional readthrough into the conjugation genes upon integration into a new insertion site. To verify this experimentally, the terminator was cloned in between the *tet*(M) promoter and a *gusA* reporter in pHCMC05. We demonstrated the level of β-glucuronidase enzyme activity decreased, confirming termination activity. We have for the first time, identified and verified a group of conserved terminators in the conjugation region of the Tn*916*-like family of CTns. Further data supports our hypothesis that the terminator efficiency is modulated upon excision and circularisation of Tn*916*, which is the exact time when Tn*916* would require expression of its conjugation genes.

A fundamental understanding of the current antibiotic resistance mechanisms employed by bacteria is also essential to minimise the emergence of resistance and to devise effective resistance-control strategies. Another aim

of this study is to investigate the molecular mechanism underlying macrolide resistance in *Bacillus subtilis*. Macrolide-resistant *B. subtilis* were generated as part of the project and analysis revealed a new genetic mutation to be responsible for the macrolide resistance phenotype. Comparative genome analysis revealed 21 bp and 54 bp duplication in the *rplV* of these mutants in comparison to the wild type strain. The *rplV* encodes the large ribosomal subunit protein, L22. Alteration in L22 has led to a predicted alteration in the C-terminal loop of the protein, predicted to change the shape of the exit tunnel within the ribosome. Ectopic expression of the *rplV* mutants containing the 21 bp and 54 bp duplication in *B. subtilis* BS34A confers resistance to macrolides. This is the first observation of macrolide resistance due to 54 bp duplication in the *B. subtilis rplV* gene.

Impact Statement

Tn*916* and Tn*916*-like elements are responsible for the spread of AR genes and therefore research on the molecular basis of transcriptional regulation and mobilisation of these elements is essential. The work in this study has shown the presence of a previously unknown group of terminators located upstream of the conjugation module of Tn*916* and Tn*916*-like elements. These are structurally conserved across multiple elements suggesting their important role in regulating the transcription of conjugation genes responsible for transfer. I hypothesise that these terminators are biologically important in preventing transcription of conjugation genes in order to maintain their stability within a genome. Therefore, if we could interfere with the terminators then there is a possibility we could destabilise the element and induce their loss.

Many Gram-positive bacteria, including *B. subtilis* undergo ribosomal target site alteration to disrupt the interaction between the macrocyclic ring of the macrolide and its binding pocket within the nascent peptide exit channel (NPET). There is a diverse spectrum of mutations occurring directly or indirectly at the ribosome that confers macrolide resistance. However, there are layers of complexity that contribute to this resistance mechanism. Therefore, it is important to understand all the mechanisms responsible for resistance as this is crucial infromation to design new effective antimicrobial agents. In this study, *B. subtilis* mutants that are resistant to various macrolides were investigated. A novel resistance mechanism against erythromycin conferred by tandem duplication in *rplV* that encodes ribosomal protein L22 was identified.

To tackle the antibiotic resistance problems effectively, the understanding of the molecular basis of factors that cause resistance through horizontal gene transfer and chromosomal mutations is crucial as it may provide insight into novel approaches to prevent the dissemination of AR genes and strategies to minimise the emergence of AR to new antimicrobial agents. We have provided information regarding functional conserved terminators among various Tn*916*-like family of CTns and this will lead to a better understanding of the nature of these elements. The unravelling of the novel mutation in *rplV* of *B. subtilis* that confers resistance to macrolides has contributed to added knowledge on a newly discovered mechanism employed by bacteria in response to selective pressures. The knowledge of the resistance mechanism will provide an essential key in the development of new and improved antibiotics in tackling this global issue. In this study, we have investigated resistance against clinically used (erythromycin and tylosin A) and proprietary macrolides (tylosin A analogues) which are still in development. Understanding the resistance mechanism against these analogues is likely to provide insight into their exact mode of action. This data is also useful to predict how quickly resistance to these antibiotics is likely to evolve.

Acknowledgements

Finally, I wish to thank my beloved parents, brother and sisters for the unceasing love and support. My heartfelt gratitude to my husband, Syahrom H. Hassani, for his remarkable patience and understanding. This accomplishment would not have been possible without the dedicated support of my family.

To those who indirectly contributed in this research, your kindness means a lot to me. Thank you very much.

Table of Contents

List of Tables

List of Figures

Abbreviations

ACT	Acetyltransferase
AMR	Antimicrobial resistance
ANT	Adenyltransferase
AR	Antibiotic resistance
ARG	Antimicrobial resistance gene
ATP	Adenosine Tri-phosphate
BLAST	Basic Local Alignment Search Tool
bp	Base pair
°C	Degree Celsius
CAT	Chloramphenicol acetyltransferase
CIAP	Calf Intestinal Alkaline Phosphatase
CTAB	Cetyltrimethylammonium bromide
DNA	Deoxyribonucleic acid
EDTA	Ethylenediaminetetraacetic acid
ESBL	Extended-spectrum beta-lactamases
ESKAPE	*Enterococcus faecium*, *Staphylococcus aureus*, *Klebsiella pneumoniae*, *Acinetobacter baumannii*, *Pseudomonas aeruginosa*, and *Enterobacter* species
g	Gravitational force
HGT	Horizontal Gene Transfer.
hr	Hour
IPTG	Isopropyl-β-D-thiogalactopyranoside
kb	Kilobase

λ	Bacteriophage lambda
LB	Luria-Bertani
LPS	Lipopolysacharide
µg	Microgram
µl	Microlitre
MAMs	Macrolide arrest motifs
MCS	Multiple cloning site
MDR	Multiple drug resistance
MIC	Minimum inhibitory concentration
ml	Millimetre
MLS$_B$	Macrolide-lincosamide-streptogramin B
µm	Micrometre
µM	Micromolar
min	Minute
MVs	Membrane vesicles
nm	Nanometre
NPET	Nascent peptide exit tunnel
OD	Optical density
ONPG	o-nitrophenyl-β-D-galactopyranoside
ORF	Open reading frame
PBS	Phosphate-buffered saline
PCR	Polymerase chain reaction
PTC	Peptidyl transferse centre

PNPG	ρ-nitrophenyl-β-D-glucuronide
RBS	Ribosome binding site
RNA	Ribonucleic acid
rpm	Revolutions per minute
RPP	Ribosomal Protection Proteins
rRNA	Ribosomal Ribonucleic Acid
sec	Second
UV	Ultraviolet
V	Volt
v/v	Volume percent
w/v	Weight/volume percent
TAE	Tris Acetate EDTA
Tn	Transposon
WHO	World Health Organization
X-Gal	5-bromo-4-chloro-3-indolyl-β-D-galactopyranoside

1 General Introduction

1.1 Antibiotic Resistance

1.1.1 Antibiotic resistance

Antibiotics have played a pivotal role in treating and preventing bacterial infections, but evolution by natural selection along with the overuse and misuse of antibiotics have led to the emergence of antibiotic resistance (AR) (Read & Woods, 2014). AR is the ability of bacteria to resist the effects of antibiotics that were designed to kill them or inhibit their growth (Ventola, 2015, Yelin & Kishony, 2018). In 1940, the first case of penicillin resistance was reported in *Escherichia coli*, which can produce penicillinase (Abraham & Chain, 1988, Sengupta *et al.*, 2013). From then on, the emergence of AR-bacteria was continuous. Resistance to chloramphenicol, tetracycline and streptomycin were reported just within a few years after they were introduced (Crofton & Mitchison, 1948, Chopra & Roberts, 2001). Similarly, the first case of methicillin resistant *Staphylococcus aureus* (MRSA) was identified shortly after the introduction of methicillin in 1959 (Jevons, 1961). Today, the number or multiple drug resistance (MDR) bacteria keeps on increasing (Ligon, 2004, Ventola, 2015, Castro-Sánchez *et al.*, 2016, Munita & Arias, 2016, Yelin & Kishony, 2018).

In 2014, the World Health Organization (WHO) published the first global report on surveillance of antimicrobial resistance. Data from 114 countries showed steady emergence and spread of antibiotic resistance among bacteria of major public health importance such as *Escherichia coli, Klebsiella pneumoniae, Staphylococcus aureus, Streptococcus pneumoniae,*

Salmonella spp., *Shigella* spp. and *Neisseria gonorrhoeae* (World Health Organization, 2014). The report also acknowledged the lack of data due to poor surveillance and non-standardised monitoring of antibacterial resistance. Despite the limitations, the report demonstrates that AR is a global problem requiring a global solution (World Health Organization, 2014). In 2015, the WHO published a global action to provide a framework for individual countries to develop national action plans to tackle AMR. It is estimated that deaths caused by antimicrobial resistance can reach up to 10 million people per year by 2050 along with huge economic burden ($100 trillion per year), if no action is taken (O'Neill, 2016). However, the estimations from this report could be inaccurate as it was based on incomplete data from the European Antimicrobial Resistance Network (EARS-Net), which track the proportion of isolates within a given species that are resistant to an antibiotic, and not the number of infections caused by antimicrobial-resistant organisms, which are much harder data to collect (de Kraker *et al.*, 2016). Quantifying the global morbidity, mortality and the economic burden caused by AR is challenging as there are currently limited data on the prevalence and geographical distribution of AR particularly in low- and middle-income countries (Hay *et al.*, 2018).

AR in bacteria can occur naturally due to mutations or the acquisition of genetic material through horizontal gene transfer. These types of resistances are referred as "acquired resistance". Mutations result from errors during DNA replication or induced by mutagens. In response to environmental challenges or selective pressures such as the presence of antibiotic, bacteria with beneficial mutation will be able to survive via natural selection (Woodford &

Ellington, 2007). In contrast to acquired resistance, "intrinsic resistance" is the innate ability of the bacteria to resist the activity of specific antimicrobial agents due to their inherent structural or functional characteristics. It can be mediated by efflux pumps, impermeability of the outer membrane or lack of drug targets, as described in section 1.1.2.3 (Cox & Wright, 2013, Zhang & Feng, 2016).

The acceleration of AR can be driven by multiple factors that include; misuse and overuse of antibiotics in human, extensive use of antibiotics in agriculture, environmental contamination by waste products from antibiotics manufacturing, disposal and sewage waters (Alvarez-Martinez & Christie, 2009), nosocomial infections in healthcare transmission and sub-optimal dosing of antibiotic (Castro-Sánchez *et al.*, 2016). In order to understand how antibiotic resistance spreads, we need to investigate the mechanism of resistance as well as the genetic basis of transferable antibiotic resistance. This study focusses on acquired antibiotic resistance in *Bacillus subtilis*.

1.1.2 Antibiotics mode of action and resistance mechanisms

Antibiotics interfere with bacterial cellular processes, and the components or systems they affect differs with each class of antibiotic. They can be classified based on their mode of action; Aminoglycosides, chloramphenicol, macrolides, oxazolidinones and tetracyclines that inhibits protein synthesis, β-lactams and glycopeptides that inhibit cell wall synthesis, polymyxins that disrupt cell membrane function, quinolones and rifampin that inhibit nucleic acid synthesis and sulfonamides that inhibits folate synthesis (Table 1.1).

Table 1-1 Mode of action and resistance mechanism of antibiotics

Antibiotic class	Site or mode of action	Resistance mechanisms	Reference
Aminoglycosides	Inhibition of protein synthesis (30S ribosomal subunit).	• Target alteration: methylation of 16S rRNA by rRNA methyltransferases (RMTs). • Enzymatic modification: aminoglycoside modifying enzymes (AMEs). • Efflux pumps: RND family efflux transporter. • Decreased uptake: changes in outer membrane.	(Krause *et al.*, 2016)
β-lactams	Inhibition of cell wall biosynthesis. Form an acyl-enzyme complex with PBP, interfering terminal transpeptidation process.	• Target alteration: low affinity PBP 2a. • Enzymatic degradations: β-lactamases and extended-spectrum-β-lactamases (ESBL).	(Frère & Page, 2014, Kumar *et al.*, 2014, Ealand *et al.*, 2018)
Chloramphenicol	Inhibition of protein synthesis (50S ribosomal subunit).	• Enzymatic modification: chloramphenicol acetyltransferases (CATs). • Efflux pumps: MFS, RND family efflux transporter.	(Schwarz *et al.*, 2004)
Glycopeptides	Inhibition of cell wall biosynthesis. Bind to precursor of peptidoglycan, preventing cross-linking of the peptidoglycan layer.	• Target alteration: Formation of peptidoglycan receptors with reduced glycopeptide affinity; D-alanyl-D-lactate or D-ala-D-serine on the cell wall of vancomycin-resistant strains.	(Sujatha & Praharaj, 2012, Zeng *et al.*, 2016)
Macrolides	Inhibition of protein synthesis (50S ribosomal subunit).	• Target alteration: amino acid changes in the ribosonal protein L3, L4 and L22. • Target alteration: methylation of rRNA by *erm*-encoded methylases. • Enzymatic modification: macrolide phosphotransferases and macrolide esterases. • Efflux pumps: Mef and Msr transporter.	(Wekselman *et al.*, 2017, Vázquez-Laslop & Mankin, 2018)

Oxazolidinones	Inhibition of protein synthesis (50S ribosomal subunit). Suppress 70S inhibition and interact with peptidyl-t-RNA.	• Target alteration: methylation of rRNA by *cfr*-encoded methylases. • Efflux pumps: ABC-F family of ATP-binding cassette proteins encoded by *optr*A and *poxt*A.	(Aoki *et al.*, 2002, Wang *et al.*, 2015, Antonelli *et al.*, 2018)
Polymyxins	Alteration of cell membrane function by electrostatically bind to the negatively-charged LPS.	• Target alteration: LPS modifiations by cationic substitution of the phosphate groups. • Efflux pumps encoded by *AcrAB* and *KpnEF* in *Klebsiella pneumoniae.* • Overexpression of the outer membrane protein OprH in *Pseudomonas aeruginosa.*	(Young *et al.*, 1992, Moffatt *et al.*, 2010, Padilla *et al.*, 2010, Srinivasan & Rajamohan, 2013, Olaitan *et al.*, 2014, Poirel *et al.*, 2017)
Quinolones	Inhibition of DNA synthesis. Interfere with DNA replication and transcription.	• Target alteration: mutations in DNA gyrase and topoisomerase IV that reduce binding affinity. • Efflux pumps encoded by *oqxAB and qepA*. • Downregulation of porin expression.	(Strahilevitz *et al.*, 2009, Aldred *et al.*, 2014, Naeem *et al.*, 2016)
Rifampin	Inhibition of mRNA synthesis.	• Mutation within *rpoB* that encodes β-subunit of RNA polymerase.	(Nicholson & Maughan, 2002, Hellmark *et al.*, 2009, Goldstein, 2014)
Sulfonamides	Inhibition of folic acid metabolism (Competitively inhibits DHPS).	• Production of DHPS with low affinity for sulfonamides (encoded by *sul*1, *sul*2, *sul*3 and *sul*4). • Overproduction of PABA. • Efflux pump: p-aminobenzoyl-glutamate transporter (AbgT) family.	(Sköld, 2000, Delmar & Yu, 2016, Griffith *et al.*, 2018, Kim *et al.*, 2019)
Tetracyclines	Inhibition of protein synthesis (30S ribosomal subunit).	• Target protection: production of ribosomal protection proteins (RPPs) such as Tet(M) and Tet(O). • Enzymatic modification: Tet X, Tet 37. • Efflux pumps encoded by *tet* genes: Tet(A), Tet(B), Tet(K).	(Connell *et al.*, 2003, Yang *et al.*, 2004, Forsberg *et al.*, 2015, Chukwudi, 2016)

Abbreviations: RND - resistance-nodulation-cell division family, ABC - ATP binding cassette family, MFS - major facilitator superfamily, PBP - penicillin-binding protein, DHPS - dihydropteroate synthase, PBP - penicillin-binding protein, LPS – Lipopolysaccharides.

Bacteria have evolved sophisticated mechanisms of antibiotic resistance in order to survive in an environment where antimicrobial drug is present. Generally, bacteria can become resistance through a number of mechanisms that includes; i) changes in membrane permeability to restrict the access of antibiotics, ii) active efflux mechanisms to prevent the accumulation, iii) enzymatic modification or degradation of the antibiotics, iv) modification of the target sites, v) overproduction of the targets, vi) protection of the target site and vii) bypass; an acquisition of alternative metabolic pathways to those inhibited by the antibiotic (Spratt, 1994, Mc Dermott *et al.*, 2003, Munita & Arias, 2016, Alav *et al.*, 2018, Yelin & Kishony, 2018) (Figure 1-1).
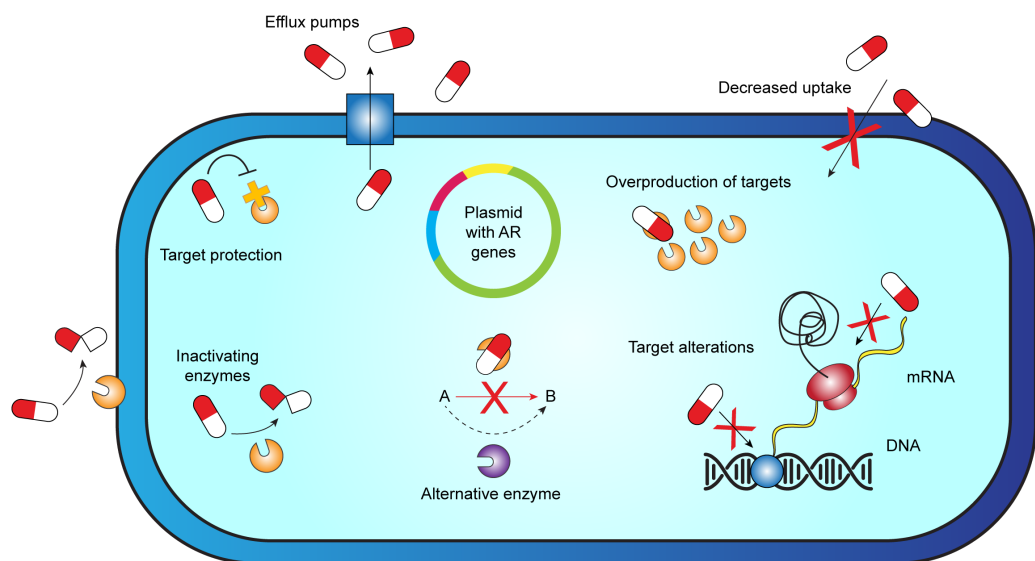
Figure 1-1 Mechanisms of antibiotics resistance in bacteria

The figure is drawn and adapted from Yelin and Kishony (2018) and Alav *et al.*, (2018).

7

1.1.2.1 Modification and protection of the target site

The interaction between antibiotic molecules and their targets are very specific. Therefore, modifications of the target site may decrease the affinity or prevent the antibiotic binding. Target modification can occur through enzymatic alterations of the binding site or point mutation in the gene encoding it (Munita & Arias, 2016). Alternatively, binding of the target site by a protective protein and overexpression of the target (change in target abundance) are other mechanisms of how resistance can be achieved (Wright, 2005).

A point mutation in *rpoB* gene is one of the classic examples of target modification that confers resistance to rifampicin. Rifampicin acts by binding to the β subunit (encoded by *rpoB*) of DNA-dependent RNA polymerase, therefore inhibiting the bacterial transcription (Hartmann *et al.*, 1967). High-level of rifampicin resistance occur by a single-step point mutation resulting in amino acid substitutions in *rpoB* causing a decreased affinity of rifampicin to its binding site (Floss & Yu, 2005). Point mutation that occurred within *gyrA*, *fusA* and *rpsL* are other examples of target modification that confer resistance towards, ciprofloxacin, fusidic acid and streptomycin, respectively.

Target modification can also be achieved by enzymatic activity. Erythromycin ribosomal methylation (*erm*) genes confer resistance to erythromycin that gives cross-resistance to other macrolides, lincosamide and streptogramin B (MLS_B). This is because *erm*-methyltransferases dimethylate the A2058 residue, in the conserved region of 23S rRNA, which is the target site for the MLS_B antibiotics (Skinner *et al.*, 1983, Maravic, 2004). Macrolide resistance

may also result from various alterations within the ribosomal proteins L4 and L22, described in section 1.3.2 (Davydova *et al.*, 2001, Zaman *et al.*, 2007).

Target protection by ribosomal protective proteins (RPPs) is another example of a defence mechanism employed by both Gram-positive and negative bacteria to block the binding of the antibiotic. The best studied RPP-mediated resistance is by Tet(O) and Tet(M), which confers tetracycline resistance (Chopra & Roberts, 2001). Tet(O) and Tet(M) acts by dislodging tetracycline from the ribosome by structural rearrangement. Subsequently, this will release the ribosome from the inhibitory effects of tetracycline, allowing the binding of aa-tRNA to continue protein synthesis (Connell *et al.*, 2003).

1.1.2.2 Modification or degradation of antibiotics

One of the major resistance mechanisms employed by bacteria is to produce enzymes that can structurally modify or degrade antibiotics, preventing them from interacting with their target. Typical targets of these modifying enzymes are antibiotics that inhibit protein synthesis such as aminoglycosides (streptomycin, kanamycin and gentamycin), phenicols and β-lactams. Chemical modifications include adenylation by *O*-adenyltransferase (ANT), acetylation by *N*-acetyltransferase (ACT) and phosphorylation by *O*-phosphotransferase which all result in steric hindrance, decreasing the affinity of the antibiotics to their target sites (Wright, 2005).

Aminoglycoside-modifying enzymes catalyse the modification at the amino or hydroxyl groups of the antibiotic molecule. AMEs are coded by *aac, aad* or *aph* genes which can be found in the chromosome, plasmid or transposons of Gram-positive and -negative bacteria such as *E. coli, P. aeruginosa, Salmonella* spp., *E. faecalis, S. aureus* and *S. pneumoniae* strains (Tolmasky, 2000, Ramirez & Tolmasky, 2010). Another example of drug modification is by chloramphenicol acetyltransferase (CAT) that acetylates hydroxyl groups of chloramphenicol. CAT is encoded by *cat* genes that have been reported to be prevalent in clinical strains (Schwarz *et al.*, 2004).

Antibiotic degradation is observed with β-lactamases, which hydrolyse the amide bond present in the β-lactam ring, deactivating the antimicrobial properties. β-lactamases can be mediated by chromosomal or plasmid encoded genes. Plasmid bearing the β-lactamases encoded genes can be transferred and shared by conjugation. Genes coding for β-lactamases are commonly found in the ESKAPE pathogens (*Enterococcus faecium*, *Staphylococcus aureus*, *Klebsiella pneumoniae*, *Acinetobacter baumannii*, *Pseudomonas aeruginosa*, and *Enterobacter* species) (Jacoby & Carreras, 1990).

1.1.2.3 Decreased permeability and antibiotic efflux

Another mechanism of resistance is to control the intracellular concentration of antibiotics by reduced permeability and efflux of antibiotics. In Gram-

negative bacteria, the outer membrane act as the first line of defence against the antimicrobial molecule. For the hydrophilic antibiotics (e.g.; β-lactams, quinolone and tetracycline) to traverse through the outer membrane, water-filled diffusion channels formed by porins are needed. Therefore, porin-mediated antibiotic resistance can be achieved by modification (function disruption) or low-level expression of porin encoded gene (Munita & Arias, 2016).

Antibiotic efflux is the capability of bacteria to extrude toxic compounds out of the cell in order to maintain their low-intracellular concentration (Li & Nikaido, 2004). Many classes of efflux systems have been described and divided into seven major families based on their structural conformation, energy source, range of substrates they are able to pump out and the type of bacteria they are found in, which are; the major facilitator superfamily (MFS), resistance-nodulation-division (RND) family, ATP-binding cassette superfamily (Carneiro *et al.*), multidrug and toxic compound extrusion (MATE) family, small multidrug resistance (SMR) family (Piddock, 2006, Alcalde-Rico *et al.*, 2016, Pasqua *et al.*, 2019), and two new families known as the proteobacterial antimicrobial compound efflux family that transport biocides such as acriflavne and chlorhexidine (Hassan *et al.*, 2018) and the p-aminobenzoyl-glutamate transporter (AbgT) family which is the sulfonamide antimetabolite transporter (Delmar & Yu, 2016).

In terms of structural conformation, the efflux system can be divided into a single-component or multiple-components system. One classic example of single-component efflux system are the Tet efflux pumps, which belong to the

11

major facilitator superfamily (MFS) (McMurry *et al.*, 1980). This type of efflux system is limited to specific profile of substrates where it can only extrude certain drugs or multiple drugs belonging to the same class. The multiple-components system is composed of tripartite membranes which are the inner membrane transporter, the outer membrane channel and the periplasmic adaptor protein. The resistance-nodulation-division (RND) efflux pump is an example of multiple-components system, capable of pumping out a more extensive range of antibiotics class such as macrolides, tetracyclines and fluoroquinolones making it clinically significant (Li & Nikaido, 2004). The RND system is predominant in Gram-negative bacteria (for example; AcrAB in *E coli* and MexB in *P. aeruginosa*) while other efflux systems are widely distributed in both Gram-positive and -negative bacteria (Puzari & Chetia, 2017).

The genes encoding efflux pumps can be found in chromosomes or mobile genetic elements such as transposon or plasmid. The existence of these genes in mobile genetic elements has significantly contributed to multidrug resistance in pathogens (Li & Nikaido, 2004).

1.2   Tetracycline

1.2.1   Mode of action

Tetracyclines are bacteriostatic agents; commonly used to treat diseases related to infections of respiratory, urogenital and gastrointestinal tracts. It is

a broad-spectrum antimicrobial agent that inhibits the growth of various Gram-negative and positive bacteria including chlamydiae, mycoplasma, rickettsiae, protozoan parasites and even viruses (Chopra *et al.*, 1992, Zink *et al.*, 2005, Michaelis *et al.*, 2007). In bacteria, the interaction between tetracycline and the 30S ribosomal subunit promotes translation arrest as it inhibits tRNA docking at the A-site during the elongation process (Maxwell, 1967, Brodersen *et al.*, 2000, Chopra & Roberts, 2001, Connell *et al.*, 2003, Connell *et al.*, 2003).

The mode of action of tetracycline primarily involves the uptake of this drug into bacterial cells and depending on whether the susceptible bacteria are Gram-positive or negative, it needs to traverse through one or two membrane systems. In Gram-negative bacteria such as *E. coli*, tetracycline diffuses passively through the outer membrane porin channels; OmpC and OmpF (Piddok & Mortimer, 1993) and passes across the cell wall in Gram-positive bacteria via an energy-dependent active transporter (Levy, 1992).

The binding sites for tetracycline have been identified through two independent crystallographic structure studies of the *Thermus thermophilus* 30S ribosomal subunit (Brodersen *et al.*, 2000, Pioletti *et al.*, 2001). The major binding pocket; Tet-1, plus the other five minor sites denoted as Tet 2 -Tet 6 were identified in small ribosomal subunit 16S rRNA (Brodersen *et al.*, 2000, Pioletti *et al.*, 2001). The Tet-1 binding site is referred as the primary binding site, where high-occupancy of tetracycline binding activity occurred. This Tet-1 binding pocket is in close proximity to the ribosomal A site (between helices h34 and h31) where tetracycline anchored itself by forming a binding complex

with two magnesium ions (Brodersen *et al.*, 2000). Occupancy of tetracycline at the A site prevents the binding of aminoacyl-tRNA via steric hindrance, consistent with the known, previously reported mode of action (Grossman, 2016). However, the relevance of the other five identified binding sites remains unclear.

1.2.2 Mechanisms of resistance to tetracycline

The resistance mechanisms to tetracycline include; RPPs (Burdett, 1991), efflux pumps and enzymatic alteration of the antibiotic (Nguyen *et al.*, 2014). Bacterial resistance to tetracycline is majorly due to the acquisition of resistance determinants rather than chromosomal mutations. The first reported tetracycline resistance occurred in 1953, conferred by *tet* gene carried on a conjugative R-plasmid that encodes tetracycline efflux protein in *Shigella dysenteriae* (Watanabe, 1963, Roberts, 1996). Tetracycline-specific efflux mechanism is mediated by MFS of transporters including Tet(A) and Tet(B) efflux pumps commonly found in Gram-negative isolates and Tet(K) and Tet(L) in Gram-positive clinical isolates. These pumps work by extruding the tetracycline molecules out of the bacterial cells by exchanging a proton with a tetracycline cation complex (Chopra & Roberts, 2001).

Out of the three tetracycline resistance mechanisms, antibiotic inactivation was initially reported as the rarest mechanism. However, enzymatic inactivation has emerged to be an alarming threat for the next-generation tetracyclines such as tigecycline (third generation), omadacycline and eravacycline (fourth generation) (Markley & Wencewicz, 2018). The

inactivation of tetracycline occurred through covalent modification by tetracycline destructases encoded by *tet*(X). This modification may lead to lower binding affinity, blocking of cellular uptake, and increased efflux thus lowering the intracellular and intercellular tetracycline concentrations (Yang *et al.*, 2004, Forsberg *et al.*, 2015). Another tetracycline modifying enzymes that possess similar activity with Tet X is Tet 37, encoded by *tet*(37). This gene was isolated from the metagenomic DNA. Despite their similar enzymatic action, there is no homology observed in the amino acid sequences of Tet 37 and Tet X (Diaz-Torres *et al.*, 2003).

The RPPs confers tetracycline resistance by binding to the ribosome that leads to blocking of the tetracycline target site, dislodging the bounded tetracycline from the ribosome or distorting the structure of ribosome to allow double binding of both tetracycline and tRNA without disrupting the protein translation process (Dönhöfer *et al.*, 2012). RPPs, which were originally described in *Campylobacter jejuni* and *Streptococcus* spp., possessed a sequence similarity to the ribosomal elongation factors, EF-G and EF-Tu (Sanchez-Pescador *et al.*, 1988). To date, there are 12 classes of reported ribosomal protection genes which can be disseminated among bacteria through horizontal transfer of mobile genetic elements. These include; *tet*(M), (O), (Q), (S), (T), (W), (32), (36), (44), *B*(P), *otr*(A) and *tet* (Roberts, 2005). Among these genes, *tet*(O) and *tet*(M) are the most common with *tet*(M) usually associated with the promiscuous Tn*916*/Tn*1545* family of conjugative transposon (Rice, 1998, Chopra & Roberts, 2001). Also, a subgroup of mosaic RPP genes has been identified, composed of multiple sections of different classes of characterized RPP genes due to recombination. For example;

tet(O/W/O) in *Megasphaera elsdenii* is a mosaic RPP gene that encodes RPP with a central section that shared 98.1% identity with Tet(W) flanked with the C- and N-terminal that showed 99.3% and 100% identity respectively with Tet(O) (Stanton & Humphrey, 2003). Other reported mosaic RPPs are Tet(O/32/O) found in *Clostridium saccharolyticum* K10, *Campylobacter coli* 202/04, and *C. coli* 317/04 (Warburton *et al.*, 2016) and Tet(S/M) found in *Streptococcus equinis* 1357 (Barile *et al.*, 2012) and *Streptococcus intermedius* (Lancaster *et al.*, 2004, Novais *et al.*, 2012).

1.3 Macrolides

1.3.1 Mode of action

Macrolides are a family of natural, synthetic and semisynthetic antibiotics which are clinically relevant. They are commonly used to treat infectious diseases caused by Gram-positive cocci such as *Staphylococcus aureus*, *Streptococcus pneumoniae* and *Streptococcus pyogenes* as well as Gram-negative pathogens such as *Neisseria gonorrhoeae, Haemophilus influenzae, Bordetella pertussis and Neisseria meningitis* (Dinos, 2017). Structurally, macrolides are comprised of 14-16 membered lactone ring carrying one or more sugar residues (Fig 1-2). Erythromycin A, a 14-membered lactone ring is the first generation of macrolide. However, it is not stable, making the delivery of this drug a problem (Hassanzadeh *et al.*, 2007). Therefore, to overcome the acid instability problem, an improved generation of macrolides that belongs to 15 and 16-membered ring was designed. These include the

15-membered ring azithromycin and the 16-membered ring macrolides such as tylosin, carbomycin A and spiramycin (Alvarez-Elcoro & Enzler, 1999). To date, the latest generation of macrolides, the ketolides, have been developed to combat the emergence of macrolide-resistant strains (Denis *et al.*, 1999, Felmingham, 2001).
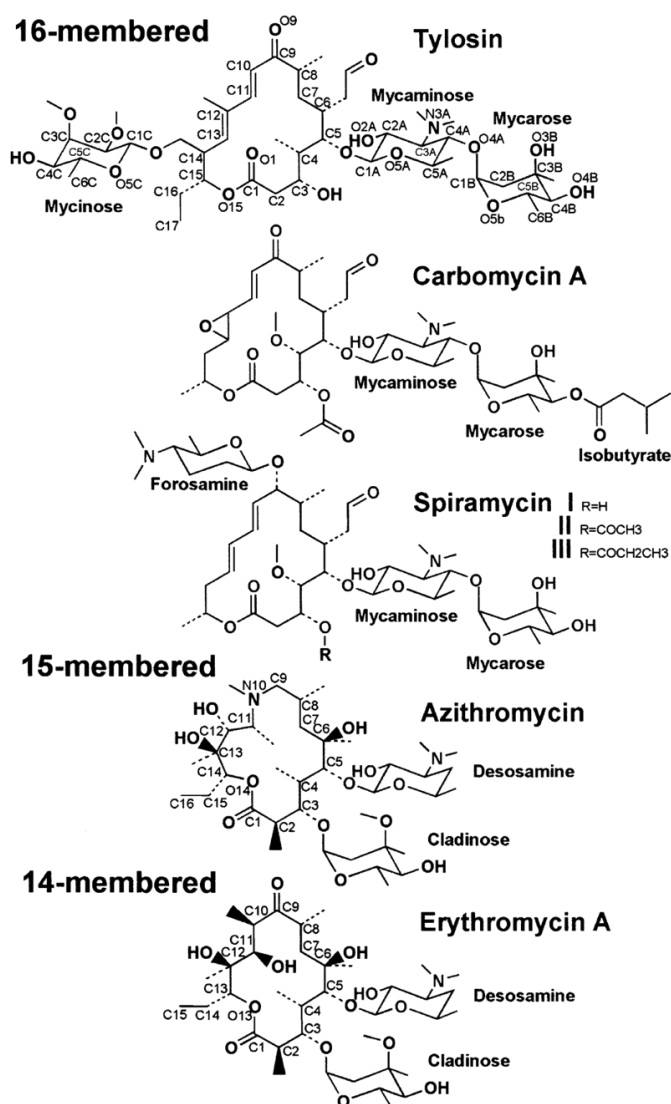
Figure 1-2 Chemical structures of the macrolide antibiotics.

First row: 16-membered ring macrolides; Tylosin, Carbomycin A and Spiramycin. Second row: 15-membered ring azithromycin. Third row: first generation, 14-membered ring erythromycin. Figure is reproduced with permission (Hansen *et al.,* 2002).

The specific structure of each macrolide is one of the key factors that differentiate the inhibitory action in various classes. This depends on the nature of their side chains and their molecular interactions with the ribosome (Moazed & Noller, 1987, Tu *et al.*, 2005). In general, there are four modes of inhibition in macrolides that have been described; i) Interference of 50S protein assembly (Chittum & Champney, 1995, Champney *et al.*, 1998, Usary & Champney, 2001), ii) Inhibition of peptide bond formation; iii) Inhibition of the nascent peptide chain progression into the NPET (Schlunzen *et al.*, 2001, Tu *et al.*, 2005, Bulkley *et al.*, 2010, Dunkle *et al.*, 2010) and iv) Premature dissociation of the peptidyl-tRNA from the ribosome (Otaka & Kaji, 1975, Menninger & Otto, 1982). These modes of action could occur independently or sequentially as one event could lead to another event, disrupting mRNA translation.

Macrolides are protein synthesis inhibitors that target the nascent peptide exit tunnel (NPET) around the peptidyl transferase center (PTC) where it interacts with 23S rRNA at the A2058 residue (Gabashvili *et al.*, 2001). NPET Is a tunnel or a passageway, which is 100 Å in length and 10 - 20 Å wide, where the polypeptide chain is released (Zhang *et al.*, 2014). Within the NPET, there is a constricted region formed by the extended loops of ribosomal proteins L4 and L22 (Tu *et al.*, 2005, Bulkley *et al.*, 2010) (Figure 1-3). It was long thought that the macrolides simply inhibits the release of proteins by physically blocking this passageway, or causing clogged polypeptides that disrupt the protein synthesis apparatus once the polypeptides reached 3-10 amino acids long (Tenson *et al.*, 2003). With the inhibition of peptide progression, this will

subsequently lead to peptidyl-tRNA drop-off from the ribosome (Menninger & Otto, 1982, Menninger, 1985, Tenson *et al.*, 2003).
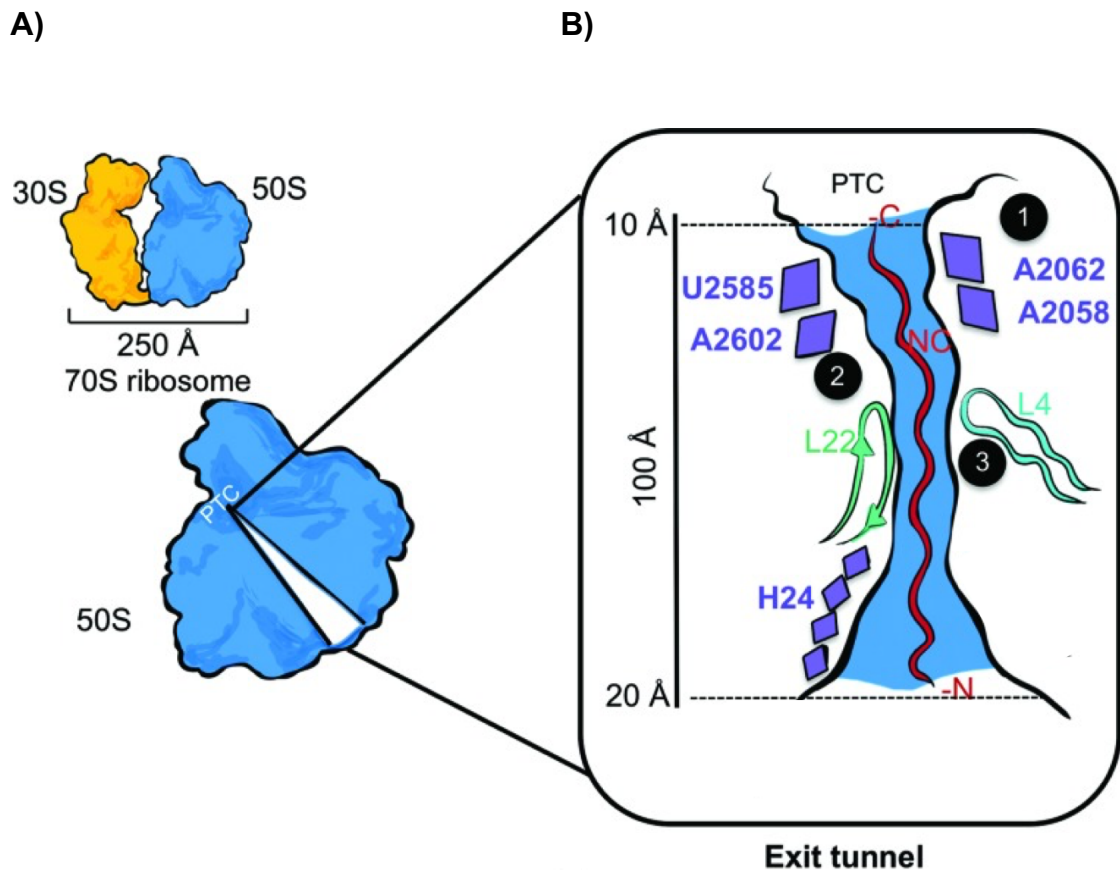
Figure 1-3 The structure of nascent peptide exit tunnel (NPET) with nascent chain (NC) shown in red.

Panel A shows a schematic diagram of the ribosome consist of small 30S (yellow) and large 50S (blue) subunits. Location of NPET are shown adjacent to PTC, expanding to the middle section of the large subunit (50S). Panel B shows the enlarged structure of NPET wall lined with functional regions; 23S rRNA nucleotides (purple) and the two ribosomal protein L4 and L22 loops (cyan and green), that forms a constricted region. 23S rRNA nucleotides marked as region 1 and 2, along with constricted part marked as region 3 interacts with the nascent peptide. Figure is adapted from (Javed *et al.*, 2017).

The clogged hypothesis is supported by the narrowing of the tunnel which is referred to as the 'plug-in-the-bottle' model (Yanouri *et al.*, 1993, Hansen *et al.*, 2002). However, recent findings showed that macrolides are not merely a protein plug and the mechanism of action of these drugs are more complicated (Kannan *et al.*, 2012, Sothiselvam *et al.*, 2014, Vázquez-Laslop & Mankin, 2018). Instead of being general translation inhibitors, macrolide selectively inhibit translation of a distinct subset of proteins.

Crystallographic studies of the macrolide-bounded ribosome structure shown that the tunnel is not totally blocked and there is a sufficient room left that allows the passage of some nascent peptides through the NPET (Kannan *et al.*, 2012, Arenz *et al.*, 2014, Arenz *et al.*, 2014, Kannan *et al.*, 2014, Dinos, 2017). Some peptides that manage to slip through bypassing the macrolide will either continue to be elongated resulting in long polypeptides on macrolide-bound ribosome or to be interrupted at the later stage of translation (Kannan *et al.*, 2012, Kannan *et al.*, 2014, Dinos, 2017). Hypothetically, these peptides may induce a 'drug-eviction mechanism', where it co-translationally pushed and dislocated the erythromycin molecule from its binding site within the NPET. When this occurs, the macrolides will be dislodged from the ribosome and will not be able to rebound as the NPET is now occupied by the elongated polypeptides (Tenson & Mankin, 2001).

The NPET is not merely an inert conduit but can interact with the specific sequence of the nascent peptide and modulate the translation process (Tenson & Ehrenberg, 2002). This interaction may lead to a translational arrest in a macrolide bound ribosome. Through ribosome profiling (ribo-seq), several

specific sequence motifs of nascent peptides where the macrolide-bound ribosome stalling occurred have been identified. This specific sequence motif is referred as macrolide arrest motifs (MAMs) (Davis *et al.*, 2014, Kannan *et al.*, 2014, Vázquez-Laslop & Mankin, 2018). One of the significant MAMs identified is the tripeptide Arg/Lys-X-Arg/Lys, where the middle X represents any amino acid, hence this MAM is identified as the +X+ motif. Stalling at this specific motif is due to the positive charge of Arginine and Lysine that interfere with the peptidyl transfer reaction (Davis *et al.*, 2014, Kannan *et al.*, 2014, Sothiselvam *et al.*, 2014). Moreover, the size of the Arg and Lys side chains also contributed to stalling, as there are among the longest amino acids in comparison to other 20 amino acids. Sothiselvam and co-workers demonstrated that the macrolide inhibitory effect is reduced, when the Lys residue within the MAM is replaced with Ala, an amino acid with a shorter side chain. The key to this mode of action is the interactions between the macrolide molecule (bounded at domain V of the NPET) and the nascent peptide containing the +X+ MAM that mediates the allosteric changes of the PTC, subsequently hindering the peptide bond formation (Sothiselvam *et al.*, 2014).

MAMs are present in the leader peptides of macrolide resistance genes (including the well-studied *ermCL* and *ermD*) and this is in line with the action of these leader peptides as a regulator of these resistance genes. Therefore, it leads to a conclusion that, protein synthesis is inhibited not because the macrolide molecule is obstructing it passageway, but rather because the macrolides prevent the ribosome from catalyzing the peptide bond formation between the MAM residues (Vazquez-Laslop *et al.*, 2008, Arenz *et al.*, 2014, Johansson *et al.*, 2014, Sothiselvam *et al.*, 2016). Therefore as mentioned

above, instead of being merely a plug in a tunnel, macrolide is protein-specific translation modulators.

1.3.2 Mechanisms of resistance to macrolides

Macrolide resistance can occur due to several different biochemical routes, that may coexist simultaneously in the same bacteria cells. This includes; i) macrolide inactivation via phosphorylation or hydrolysis of the lactone ring, ii) reduced intracellular concentration of macrolide by altering the bacterial cell membrane permeability or efflux pumps, and iii) modification of the ribosomal 23S rRNA via methylation or alteration in ribosomal protein loops L4 or L22.

Chemical alteration of macrolides results in impaired binding to their ribosomal target site. Two major classes of enzymes responsible for this alteration are macrolide phosphotransferases and macrolide esterases. Macrolide phosphotransferases, which is also known as macrolide protein kinases catalyse the transfer of phosphate group to the 2'-hydroxyl group of macrolides (14-, 15-, and 16-membered ring) thus disrupting the key contact site of macrolides with A2058 of 23S rRNA (Fig. 1-3). In contrast, the macrolide esterases are only capable of using 14- and 15-membered macrolides as their substrate but not the 16-membered macrolides. It acts in the reverse ring opening mode by cleaving the macrocyclic ester (of the 15-membered macrolide) (Wright, 2005).

Two major subfamilies of efflux pumps used for macrolide extrusion are Mef and Msr that belongs to major facilitator superfamily (MFS) and ATP-binding cassette superfamily, respectively. Mef proteins are antiporters driven by proton motive force with *mef*(A) and *mef*(E) being the two major subclasses. However, as they are 90% identical, they are now collectively referred as *mef*(A) (Roberts *et al.*, 1999). Although it was long thought that the Msr proteins are merely ATP-dependent active transporters, recent finding showed that its mode of action expands as it has been demonstrated to act similarly as the RPPs; Tet(M) and Tet(O) in protecting the ribosome (Fyfe *et al.*, 2016). Through interaction with the ribosome, it works by dislodging and displacing the bounded macrolide from its binding site (Wilson, 2016). Both Msr and Mef subfamily of proteins are capable of pumping out 14- and 15- membered macrolides efficiently out of the cell (Fyfe *et al.*, 2016).

Modification of the ribosomal target site is the major mode of resistance employed by bacteria and confers a broader spectrum of resistance in comparison to efflux and inactivation mechanism. The first case of erythromycin resistance occurred in 1956, soon after the drug was introduced. It was found in staphylococci due to the methylation of 23S rRNA at nucleotide A2058 by erythromycin ribosomal methyltransferase (Erm) (Weisblum, 1995). Erm enzymes are encoded by *erm* and can catalyse either a mono-methylation or a dimethylation reaction that confers low to moderate or high resistance to macrolides, respectively (Poehlsgaard & Douthwaite, 2005). A total of 38 *erm* genes have been reported with *erm*(B) and *erm(*C) being the most common (Fyfe *et al.*, 2016).

Apart from alteration that occurred within the 23S rRNA, alterations in L4 and L22 ribosomal proteins could also render resistance or reduced susceptibility towards macrolides. The type of mutation includes various deletion, insertion or substitution within the genes encoding it (*rplV* and *rplD*). These mutations have been observed in *E. coli* and *B. subtilis* laboratory isolates as well as clinical isolates such as *S. pneumoniae, S. pyogenes, S. aureus* and *S. oralis* (Bingen *et al.*, 2002, Canu *et al.*, 2002, Malbruny *et al.*, 2002, Doktor *et al.*, 2004, Zaman *et al.*, 2007, Chiba *et al.*, 2009). Alteration in L4 prevents the binding of the drugs while alteration in L22 neutralises the effects of binding (Wittmann *et al.*, 1973). The L4 mutant was also found to be functionally defective as the rate of its peptidyl transferase activity was greatly reduced. Interestingly, peptidyl transferase activity of the L22 mutant was not affected and observed to be close to regular rates (Wittmann *et al.*, 1973). Based on cryo-EM study of erythromycin-resistant *Escherichia coli* 70S ribosomes, alterations in L4 and L22 caused structural changes within the NPET. In the L4 mutant, narrowing of the tunnel was observed preventing the entry and binding of erythromycin as the size of the opening is now reduced to be smaller than the erythromycin A molecule. In contrast, alteration in L22 caused a widening of the tunnel, neutralising the effect of erythromycin binding (Gabashvili *et al.*, 2001). These observations suggest that the ability of the ribosome to bind macrolide molecule is correlated with the width of the tunnel entrance. The increased width of tunnel explains how the nascent polypeptide chain can egress through the NPET of macrolide-bounded ribosome. Interestingly, it has been observed in many cases of L22 mutants where the

macrolides are still able to bind with high affinity despite the conformational changes of the target site (Gabashvii *et al.*, 2001, Davydova *et al.*, 2002).

The study of erythromycin-bounded 50S ribosome structure revealed that the distance of the erythromycin binding pocket with L22 loop is too far for direct interaction to occur (~9 Å) (Schlunzen *et al.*, 2001, Tu *et al.*, 2005, Bulkley *et al.*, 2010, Dunkle *et al.*, 2010). Therefore, L22 alteration is considered as an indirect effect factor of macrolide resistance. In a more recent X-ray crystal structure study of erythromycin-resistant *Deinococcus radiodurans* 50S subunit (containing L22 mutant with three insertion residue; Dr-Ins3), β hairpin of the altered L22 were shown to be shifted towards the inner part of the exit tunnel, subsequently triggering a cascade of structural rearrangement within 23S rRNA nucleotides that propagates towards the binding pocket of erythromycin (Wekselman *et al.*, 2017). This is in agreement with the crystal structure study of the *Thermus thermophilus* L22 mutant containing triplet deletion at the residue 82–84 (Leu-Lys-Arg), where inwards shifting of the L22 β hairpin caused a destabilisation of the macrolide-binding pocket (Davydova *et al.*, 2002).

1.4 *Bacillus subtilis*

1.4.1 *B. subtilis* - model organism of Gram-positive bacteria

Bacillus subtilis is an aerobic, endospore-forming Gram-positive bacterium, under the genus of *Bacillus* (Harwood & Wipat, 1996). It is one of the best-

characterised bacterium, often used as a model system for cell differentiation (Piggot & Hilbert, 2004) and chromosome replication (Jameson & Wilkinson, 2017). It was discovered by Christian Gottfried Ehrenberg in 1835 whom formerly named it as *Vibrio subtilis.* Later*,* in 1872 it was renamed as *Bacillus subtilis* by a German botanist, Ferdinand Cohn. The word *Bacillus* refers to the rod-shape of this bacterium and *subtilis* means fine and slender. Their cells are typically about 4-10 μm long and 0.25-1.0 μm in diameter. *B. subtilis* is an environmental bacterium commonly found in soil, water, air, food and the rhizosphere. It forms endospores under stressful conditions which are highly resistant to heat and desiccation. This feature enables them to survive in extreme conditions.

Due to its amenability to genetic manipulation, *B. subtilis* became a reference for Gram-positive microorganisms. Important characteristics of *B. subtilis* includes*;* (i) an efficient natural genetic transformation system, the first genetic transformation system discovered in a non-pathogenic microorganism; (ii) cell factory, producing commercially important hydrolytic enzymes and bioactive compounds (the ability of *B. subtilis* to secrete proteins into the medium has been exploited for the production of industrially relevant bioproducts such as protease, amylase and riboflavin (Harwood, 1992, Cao *et al.*, 2017)) and (iii) the ability to differentiate into heat, desiccation and chemical resistant endospores (Errington, 1993).

1.4.2 Natural competence in *B. subtilis*

B. subtilis is advantageous in terms of their rapid growth, high natural competency and DNA uptake (Tosato & Bruschi, 2004). In *B. subtilis*, the state of natural competence is expressed at the transition between exponential growth and stationary phase. At this state, a subpopulation of the cells is capable to efficiently bind, process and take up extracellular DNA (Dubnau, 1991, Solomon & Grossman, 1996). The development of competence involves three different regulatory modes; (i) nutritional (medium constituents are important factors), (ii) growth stage specific and (iii) cell type specific. In *B. subtilis*, competence develops only in a small subpopulation of starving cells (less than 20%) (Kidane & Graumann, 2005). The expression of genes responsible for competency state of *B. subtilis* is controlled by the competence factor ComK. The regulatory pathway is complex involving more than 40 genes encoding both regulatory and structural components. The transcription factor ComK is directly responsible for the expression of *comK* itself as well as the genes involved in DNA binding, uptake and recombination, making it a master regulator for the establishment of competence state (van Sinderen *et al.*, 1995).

The major strain used in the study of *B. subtilis* is 168 (focusing on the physiology and sporulation properties). The genome of *B. subtilis* 168 is 4 173 719 bp in size with an average G+C content of 43% and consists of 4,244 coding sequence (covering 89,7% of total size), 30 rRNAs and 86 tRNAs. It is an auxotrophic bacterium (requires tryptophan) and is highly competent (Kunst *et al.*, 1997).

Strain 168 originated from *B. subtilis* Marburg strain that has been mutagenized with x-rays by Paul Burkholder and Norman Giles in Yale (Burkholder & Giles, 1947). With sublethal doses of x-rays or UV, they have developed many *B. subtilis* auxotrophic mutants that could survive with single nutrient supplementation. Among these many mutants, only five were preserved and kept in the possession of Charles Yanofsky including strain 23 (auxotrophs requiring threonine), strain 122 (auxotrophs requiring nicotinic acid and strains 160, 166, and 168 (auxotrophs requiring tryptophan). In the 1950s, three of these strains (122, 166 and 168) were transformed to prototrophy when exposed to DNA from strain 23 (Spizizen, 1958). Soon after, the highly transformable strain,168, became the favorite for further research on transformation (Anagnostopoulos & Spizizen, 1961, Young & Spizizen, 1963). Many different mutants have been developed since then and most of the derivatives of strain 168 are maintained at the Bacillus Genetic Stock Center (BGSC), Ohio State University (USA).

In this study, apart from strain 168, another two *B. subtilis* strains namely as CU2189 and BS34A strain were used. The CU2189 strain is the laboratory strain with the genotype *metB5 hisAl thr-5* (Christie *et al.*, 1987) and is commonly used as a recipient in the study of conjugative transposon transfers (Christie *et al.*, 1987, Mullany *et al.*, 1990, Mullany *et al.*, 1991, Mullany *et al.*, 2012). The BS34A strain was originally derived from CU2189 that has received a single copy of Tn*916* conjugative transposon (Roberts *et al.*, 2003).

1.5 Horizontal Gene Transfer

Horizontal gene transfer (HGT) is lateral transmission of genetic material between genomes and it can occur between more or less distantly related organisms; among bacteria or even between bacteria and eukaryotic cells (Burmeister, 2015). It is one of the adaptation mechanisms in bacteria that generates genome plasticity and drives evolution (speciation and sub-speciation) in bacteria. One well documented example of convergent evolution via HGT can be seen in the *Shigella* spp. which evolved from the non-pathogenic *E. coli* due to the acquisition of virulence factors (pathogenicity islands (PAIs) and virulence plasmid) (Schroeder & Hilbi, 2008). HGT also contributes to the dissemination of antibiotic resistance genes (ARGs) and gene clusters encoding biodegradation genes, thus contributing to the emergence of multidrug resistance (MDR) and virulent pathogens (Frost *et al.*, 2005).

HGT differs from vertical transmission by which genetic information is passed from parent to offspring. HGT is mediated through three classical modes of DNA transfer; conjugation, transformation and transduction (Figure 1-4 and Figure 1-7). There are other mechanisms of HGT that do not fit well into any of these three classical modes, including Membrane Vesicles (MVs) transfer (Mashburn-Warren & Whiteley, 2006), Gene Transfer Agent (GTAs) trafficking (Solioz *et al.*, 1975) and nanotubes transfer (Dubey & Ben-Yehuda, 2011) which will be further discussed below (Figure 1-4).
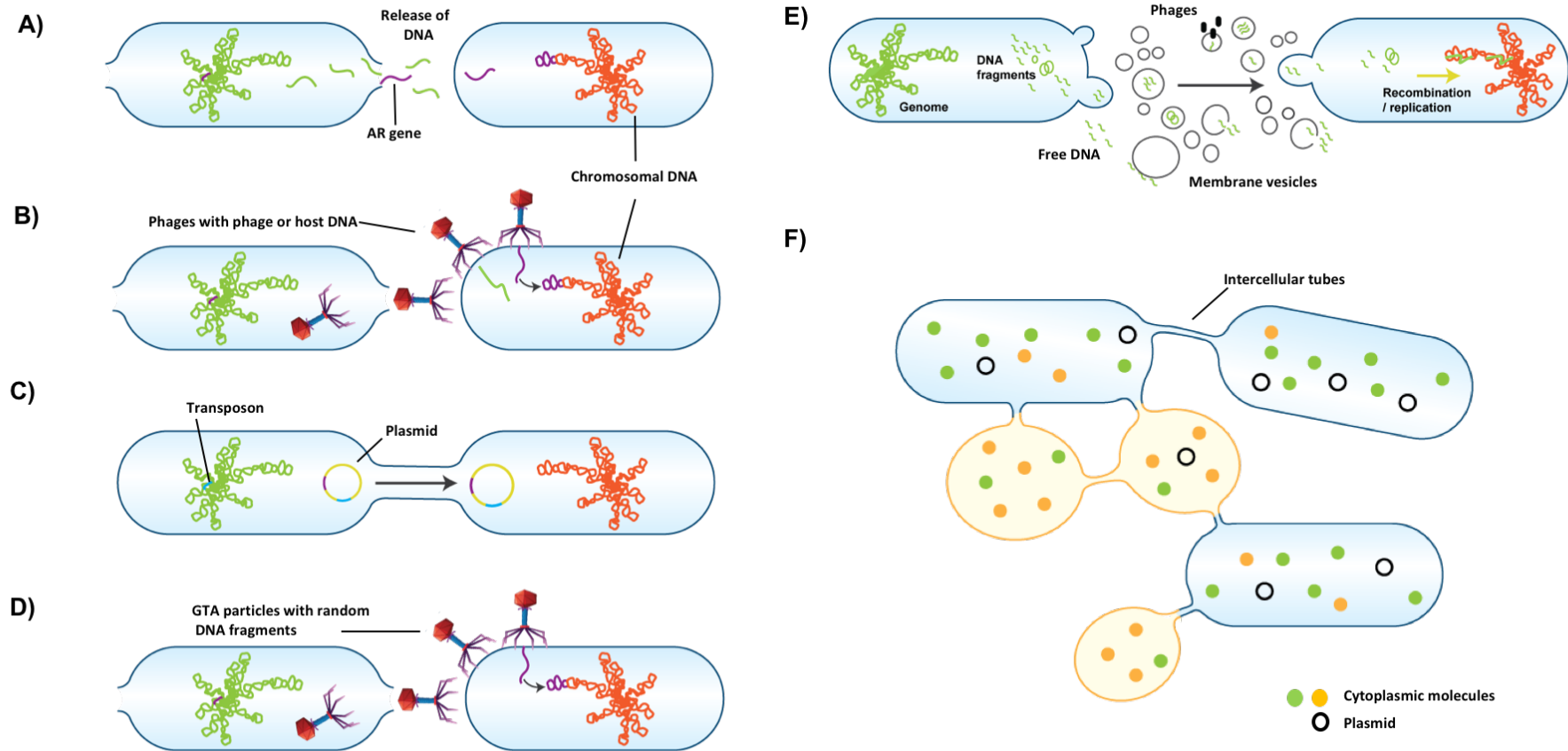
Figure 1-4 Mechanisms of horizontal gene transfer.

(A) Transformation is the uptake, integration, and expression of naked DNA from the cytoplasm, (B) Transduction is a process where bacteriophages act as a vector to transfer bacterial DNA from a previously infected donor to recipient cell, (C) Conjugation is a DNA transfer

process from a donor to recipient cell via cell surface pili or adhesins, (D) Gene transfer agents (GTAs) are bacteriophage-like particles that package unspecific segments of the bacterial genome and incomplete copies of their own genome. GTA particles are released through cell lysis. (E) Membrane vesicles (MVs) are lipid-bilayers spheres containing proteins, metabolites, DNA, RNA and/or signalling molecules produced by the donor cell and can be transferred to recipient cell. (F) Nanotubes are membranous intercellular bridges that mediate the transfer of cytoplasmic molecules between the same or different species of bacteria. Figure is drawn and adapted from (Dubey & Ben-Yehuda, 2011, Lang *et al.*, 2012, von Wintersdorff *et al.*, 2016, Chiang *et al.*, 2019).

1.5.1 Transformation

Transformation was the first mode of HGT to be discovered and regarded as an ancient beneficial adaptation mechanism in prokaryotes to repair DNA damage (Johnston *et al.*, 2014). It was first discovered in the Gram-positive bacteria, *Streptococcus pneumoniae* that became virulent due to the direct uptake and incorporation of exogenous DNA containing 'transforming principle' from the cytoplasm (Griffith, 1928). The ability of bacteria to uptake, integrate and express this extracellular DNA depend on several factors which are; the availability of the naked DNA within the cytoplasm, the competence state of the recipient bacteria and the stabilization of the translocated DNA via integration within the recipient's chromosome or by recircularization of the (autonomously replicating) plasmid. Internalizations of the DNA is assisted by a set of conserved multiprotein DNA uptake apparatus (Hahn *et al.*, 2005, Thomas & Nielsen, 2005). The competence state in bacteria may be induced differently, however the proteins involved in the uptake machinery are conserved in both Gram-positive and negative bacteria including *B. subtilis, S. pneumoniae, Neisseria gonorrhoeae, Haemophilus influenzae, and Vibrio cholerae* except in *Helicobacter pylori* that applied a conjugation-like system instead (Dubnau, 1999, Smeets & Kusters, 2002, Claverys & Martin, 2003).

Although the protein machinery is conserved, there are slight differences in the DNA uptake mechanism between the Gram-positive and negative bacteria as Gram-positive bacteria possessed a thicker peptidoglycan layer on their

cell membrane while Gram-negative bacteria have both an outer and inner membrane. In Gram-negative bacteria, the double-stranded DNA (dsDNA) is pulled to the outer membrane by the retraction of pseudopili resulting from the assembly of pseudopilin multimers and consequently translocated via pore-forming outer membrane proteins, PilQ (Chen & Dubnau, 2004, Laurenceau *et al.*, 2013). The polytopic membrane protein PilG and PilF (traffic NTPase) participate in this process. In the periplasm, the DNA is bound to a substrate binding protein ComE. Then, the ssDNA traverse through the inner membrane via a translocation channel formed by Com A (a ComEC ortholog) where simultaneous degradation activity of the other strand of DNA occurred. The internalised ssDNA is then protected by DprA and will be the substrate for homologous recombination by RecA (Figure 1-5) (Chen & Dubnau, 2004, Salzer *et al.*, 2016, Sun, 2018).
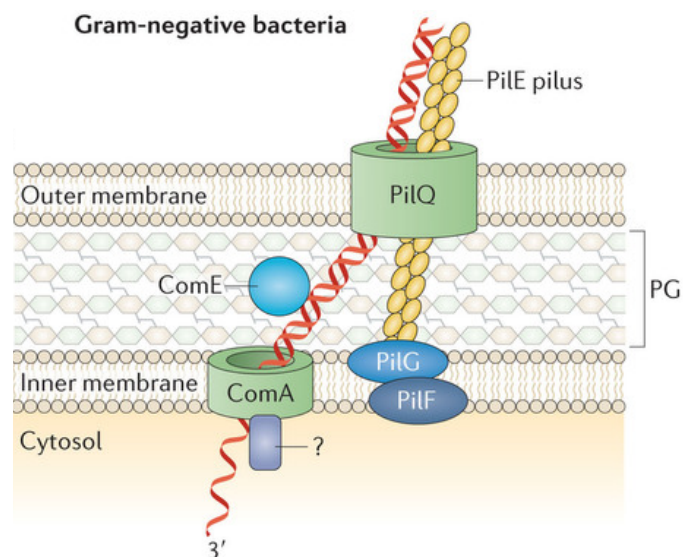
Figure 1-5 Mechanisms of DNA uptake during natural transformation of Gram-negative bacteria.

Figure is reproduced with permission from Johnston *et al.*, (2014).

In Gram-positive bacteria, the dsDNA traverse through the thick peptidoglycan layer before translocation across inner membrane occurs. To overcome the dense peptidoglycan layer, the dsDNA will be captured by the ComGC pilus that will retract and deliver the bound dsDNA to the cell surface reporter, ComEA. In *B. subtilis*, the ComGC pilus is homologous to the type IV pilus proteins assembled by the assistance of polytopic membrane protein (ComGB) and the traffic NTPase (ComGA) (Provvedi & Dubnau, 1999, Chen *et al.*, 2006). The internalisation of dsDNA is done similarly to the Gram-negative through the translocation channel (ComEC) with the assistance by ATPase ComFA (Figure 1-6).
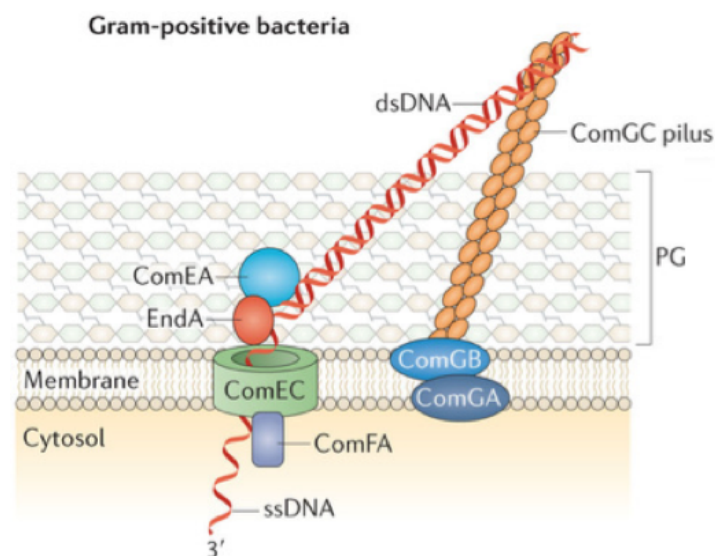
Figure 1-6 Mechanisms of DNA uptake during natural transformation of Gram-positive bacteria.

Figure is reproduced with permission from Johnston *et al.*, (2014).

1.5.2 Transduction

Transduction is a process in which genetic material is transferred from one bacterium to another by bacteriophages. This phenomenon was first described in *Salmonella typhimurium* which has undergone a recombination process due to the infection of phage P_{22} (Zinder & Lederberg, 1952). Unlike conjugation, no physical cell-to-cell contact is required for transduction to occur. There are three types of transduction; generalized, specialized and lateral transduction.

Generalised transduction occurs when nonspecific portion of bacterial DNA is packaged in one of the viral capsids due to an erroneous recognition of sequences, instead of the phage DNA and subsequently released through lysis of the bacterial cell. It is a rare event which occurs in about 1 out of 10,000 phages and mediated by lytic phages such as P_1 and P_{22} (Ikeda & Tomizawa, 1965). These transducing phages will adsorb to the surface of a new host and injects its DNA into the host cell. Once inside, subsequent integration of these DNA segments may occur into the chromosome of the recipient cell by homologous recombination (Figure 1-7). Alternatively, if the genetic material is a plasmid, it may remain in the cytoplasm and replicate autonomously where it will be pass on to daughter cells (Masters, 2000, Thierauf *et al.*, 2009).

A) Generalised

Phage DNA inserted into host cell. → Enzymatic breakdown of host DNA and copied phage DNA. → Random host DNA packaging. → Phage inserts bacterial DNA into a new host.

B) Specialised

Phage DNA spliced into host chromosome. → Imprecise excision of prophage, taking adjoining host DNA with it. → Packaging of recombined excised DNA. → Phage inserts recombined DNA into a new host.

C) Lateral

Phage DNA spliced into host chromosome. → Late excision. In-situ bidirectional replication and packaging. → Replicated DNA fills all the new phages. → Phage inserts bacterial DNA into a new host.

Figure 1-7 Types of transduction.

(A) Generalised transduction is the transfer of random fragments of bacterial chromosome. (B) Specialised transduction is the transfer of phage DNA (purple) plus flanking chromosomal DNA (green) (C) Lateral transduction is the transfer of phage DNA plus large size of adjacent chromosomal DNA as a result of atypical late excision, in situ replication and packaging. Figure is drawn and adapted from Chiang *et al.,* (2019).

Specialized transduction occurs when phage DNA excised imprecisely from the host chromosomal DNA taking adjoining bacterial DNA with it. This occurs in phage lambda which occasionally packages the *E. coli* biotin metabolism gene; *bio* or galactose metabolism gene; *gal* as these genes are located adjacent to the phage integration site (Morse *et al.*, 1956, Del Campillo-Campbell *et al.*, 1967). This recombined excised DNA then will be packaged into the capsid forming a specialized transducing phage. Stable inheritance of the donor genetic material may be established via site-specific recombination or homologous recombination once it enters a new host cell (Canchaya *et al.*, 2003).

Recently, a third type of transduction mechanism was described in the temperate phages of *S. aureus* referred to as lateral transduction and unlike the previously described mechanism, it does not occur due to erroneous process, but seems more like a natural part of phage life cycle (Chen *et al.*, 2018). Lateral prophage has atypical program where they excise late in their life cycle causing in situ bidirectional replication and packaging while the prophage is still integrated within the bacterial chromosome. The atypical order of lateral transduction is described as follows; (i) integration; (ii) replication; (iii) packaging and (iv) excision. This delay in excision results in the simultaneous replication and packaging of the adjacent bacterial chromosome together with the phage DNA. Lateral transduction captures larger sizes (>100 kb) of the bacterial chromosome in comparison to the specialised transduction. It also transfers at a higher frequencies (>1000-fold) than other mode of transductions (Chen *et al.*, 2018, Chiang *et al.*, 2019).

1.5.3 Gene Transfer Agents (GTAs)

Gene Transfer Agents (GTAs) are phage-like particles that packaged random DNA fragments of the cell producing it and transfer it to another bacterial cell. GTA genes are related to phage genes but have distinctive properties that differentiate it from phage. It generally carries a lesser amount of DNA which is insufficient to encode the phage-like structure itself (Lang *et al.*, 2012). The GTA-encoding genes are found in the genome of the host cell and are not transferable to another cell (Figure 1-8). These genes are thought to be ancestrally derived from altered bacteriophage DNA, encoding defective phages that are unable to produce phage particles but still possessed some phage-like structural characteristics (head and tail) (Yen *et al.*, 1979, Hendrix *et al.*, 1999). Once produced, the GTA particles are released by bacterial cell lysis and attached itself to recipient cell probably via tail-receptor interaction (Lang *et al.*, 2012, Lang *et al.*, 2017).

The best characterised GTA is RcGTA originates from a purple photosynthetic bacterium *Rhodobacter capsulatus* (Marrs, 1974, Solioz & Marrs, 1977, Yen *et al.*, 1979). It was the first GTA discovered when two strains of *R. capsulatus* with different antibiotic resistance phenotype were co-cultured resulting in a new strain with a double resistance phenotype. The genes encoding RcGTA are divided into two separate clusters; the 14 kb gene cluster encoding the protein needed for the head and tail morphogenesis and another structural gene cluster residing at another region (of the host chromosome) encoding the head spikes and tail fibers proteins. The RcGTA particles carry about 4kb of DNA in their approximately 30-nm spiked-capsid.
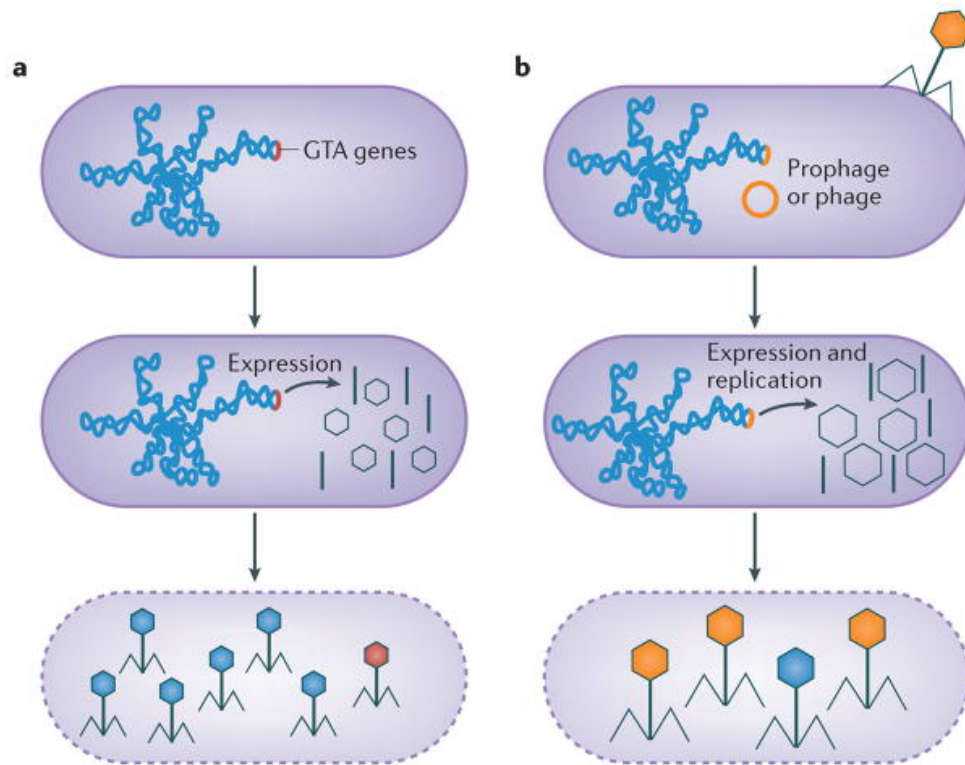
Figure 1-8 Differences between GTA and transducing bacteriophages.

a) production of GTA particles starts with the expression of GTA genes which are integrated within the host chromosome. The tiny-head GTA particles carry random segments of DNA derived from the host chromosome (blue particles) and in a rare event, it may carry a small amount of GTA genes (red particle). GTAs are released by cell lysis (dashed lines). b) production of transducing phages starts when bacteriophages inject their DNA into the host cell, followed by the expression and replication of the phage particles. The capsid may carry the complete phage genome (orange particles) or the host DNA (blue particles) that can be transferred to another cell. Figure is reproduced with permission (Lang *et al.*, 2012).

1.5.4 Membrane Vesicles (MVs)

MVs are lipid-bilayer spheres with lumen structure produced by blebbing of living cells or from cell lysis (Domingues & Nielsen, 2017). There are few types of MVs. The outer-membrane vesicles (OMVs) and outer-inner membrane vesicles (OIMVs) are produced by the Gram-negative bacteria, while the cytoplasmic membrane vesicles (CMVs) is produced by Gram-positive bacteria (Toyofuku *et al.*, 2019). MVs contain nucleic acids, polysaccharides or proteins. When carrying genetic material, MVs can mediate HGT upon exposure to new host cells. These genetic materials include both chromosomal and plasmid DNA, as well as different types of RNA including of phage origin (Domingues & Nielsen, 2017).

1.5.5 Nanotubes

HGT can also be mediated by membranous intercellular bridges termed nanotubes. Unlike conjugative pili, nanotube formation does not rely on a conjugative element and the transfer is bidirectional (Dubey & Ben-Yehuda, 2011). The nanotube structures were first identified in *B. subtilis* which has been demonstrated to transfer cytoplasmic green fluorescent protein (GFP) molecules to the adjacent cells. By using electron microscopy (EM), the GFP molecules were localized in the nanotubular protrusions that bridge the intercellular connections (Dubey & Ben-Yehuda, 2011). Additionally, the nanotubes can be formed between different bacterial species; such as *B. subtilis* and *S. aureus* and even to Gram-negative *E. coli* (Benomar et al.,

2015, Pande et al., 2015). Nanotube can mediate cytoplasmic exchange of proteins, metabolites and non-conjugative plasmids (Baidya *et al.*, 2017).

1.5.6 Conjugation

Bacterial conjugation is a process by which genetic material is transferred from a donor to a recipient cell through a cell-to-cell contact via sex pili or adhesins (Babic *et al.*, 2011). It is a multi-step process that relies on the conjugative machinery encoded by integrative conjugative elements (ICEs) within the chromosome or self-replicating plasmids in the cytoplasm (Burrus *et al.*, 2002, Smillie *et al.*, 2010, Wozniak & Waldor, 2010). Of all the horizontal gene transfer mechanisms, conjugation is the most efficient in transferring mobile genetic elements as it provides protection from the surrounding environment and often having a broader host range in comparison to bacteriophage transduction.

Conjugative DNA transport relies on a membrane-spanning multiprotein secretion apparatus, the Type IV Secretion System (T4SS). This system differs slightly in Gram-positive and negative bacteria in terms of their cell contact mechanism in order to initiate the conjugal transfer. The best-characterised T4SS is that of the alpha-proteobacteria, *Agrobacterium tumefaciens*. Generally, there are three main steps in conjugation starting with; (i) processing of the mobile genetic elements to form the relaxosome and nicking at *oriT* by relaxase aided by accessory proteins , (ii) recruitment of the ssDNA transfer-strand (T-strand) to the Type IV Coupling Protein (T4CP) and

lastly, (iii) the transfer of T-strand through the T4SS mating channel in response to cell-to-cell contact initiation (Alvarez-Martinez & Christie, 2009, Zechner *et al.*, 2012).

Direct contact between the donor and the recipient cells is essential for conjugative transfer. In Gram-negative bacteria, this is established by forming the extracellular filaments known as the sex pili. In Gram-positive bacteria, it is still unclear how this is established. However, among the enterococci, the transfer process of conjugative plasmid can be initiated by releasing a family of heat-stable peptide pheromones by the recipient cells. These pheromones will bind to the pheromone-sensor receptors that can be found on the cell surface of the donors carrying the plasmids (Chandler & Dunny, 2004). The donor cells bearing these pheromone-responding plasmids will synthesize an adhesin which in turn will induce mating aggregates with nearby recipient cells. The induced surface adhesin on the donor and recipient cell is named as aggregation substance and binding substance, respectively. The sex pheromone seem to be confined to enterococci and not use by most Gram-positive organisms (Grohmann *et al.*, 2003). The predictive T4SS model for the transfer of conjugative transposon Tn9*16* will be discussed further in section 1.6.1.3

1.6   Conjugative Transposons

Conjugative transposons are self-transferable elements that are able to integrate into bacterial chromosomes, and to excise from the chromosome,

and to transfer themselves from one bacterium to another by conjugation. (Flannagan *et al.*, 1994). They have the broadest host range and play a critical role in the dissemination of antibiotic resistance genes among pathogens (Clewell & Gawron-Burke, 1986, Clewell *et al.*, 1995, Partridge *et al.*, 2018). There are many different families of conjugative transposons but one of the most well studied is the Tn*916*/Tn*1545* family (Franke & Clewell, 1981, Clewell *et al.*, 1995, Ciric *et al.*, 2011, Santoro *et al.*, 2014).

1.6.1 Tn*916* and Tn*916*-like elements

Tn*916* is an 18 kb conjugative transposon that was first discovered in the late 1970s in *Enterecoccus faecalis* DS16 (Franke & Clewell, 1981, Flannagan *et al.*, 1994). After mating occurred between the donor *E. faecalis* DS16 and the recipient *E. faecalis* strain JH2-2, transconjugants were found to be resistant to tetracycline. It appeared to be a non-plasmid resistance transfer and further analysis revealed that tetracycline resistance was conferred by an acquired conjugative transposon designated as Tn*916* (Franke & Clewell, 1981). It has a vast host range, which enables it to conjugate into diverse species and genera of bacteria under the phyla of Deinococcus-Thermus, Actinobacteria, Firmicutes, Fusobacteria, Betaproteobacteria and Gammaproteobacteria (Roberts & Mullany, 2009).

The genes within Tn*916* and Tn*916*-like elements can be divided into four functional modules that are involved in conjugation, regulation, recombination and accessory functions such as antimicrobial resistance determinants (Figure 1-9) (Roberts & Mullany, 2009). The entire Tn*916*/Tn*1545* family

members share conserved sequence in their core region, comprising the conjugation and the regulation module (Figure 1-10). However, variation occurs in genes encoding for integrases, excisionases or recombinases and the accessory functions that they carry. For example, Tn*916* encodes a different subfamily of tyrosine recombinase from Tn*6000* (Roberts *et al.*, 2006). Sequence alignment revealed that the integrase from Tn*6000*; Int*6000* is much more closely related to tyrosine integrases from the *Staphylococcus aureus* pathogenicity islands, namely Int from SaPIbov (42% identical) and Sip from SapIbov2 (41% identical) (Roberts *et al.*, 2006). In Tn*5397*, a Tn*916*-like element from *Clostridiales difficile,* a large serine recombinase replaces the function of the excisionase (Xis) and integrase (Int) resulting in a different mechanism of recombination (Figure 1-10) (Wang *et al.*, 2000). Nearly all members of the Tn*916* and Tn*916*-like elements carry the tetracycline resistance gene *tet*(M) (Burdett, 1990, Scott & Churchward, 1995). This encodes resistance to both tetracycline and minocycline. However, some other elements also carry other antibiotic resistance genes such as *tet*(S/M) in Tn*916S* (Novais *et al.*, 2012, Warburton *et al.*, 2016)*, tet*(S) in Tn*6000* (Roberts *et al.*, 2006, Brouwer *et al.*, 2011), *aphA-3* in Tn*1545* (Courvalin & Carlier, 1986) and Tn*6003* (Cochetti *et al.*, 2008), as well as other accessory genes such as mercury resistance gene; *mer*(A) in Tn*6009* (Soge *et al.*, 2008).
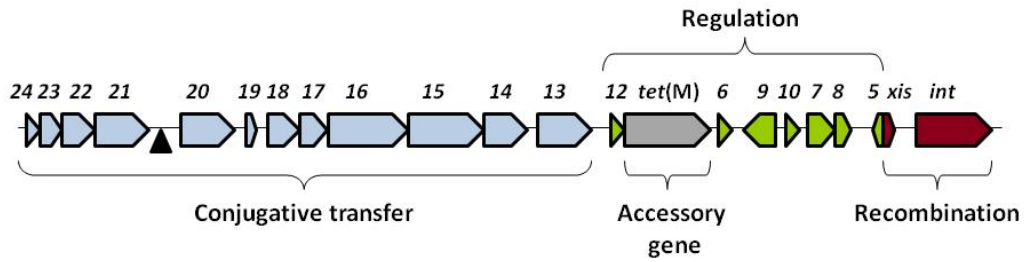
Figure 1-9 A schematic representation of Tn*916* showing the four functional modules.

The four modules: conjugation (blue); recombination (red); regulation (green) and the accessory gene *tet*(M) (grey). Arrow boxes represent the open reading frames (orfs) and the orientation of the genes. Filled triangle represents the position of the *oriT* (origin of transfer), which is the conjugation-nick-site. Figure is reproduced with permission (Roberts & Mullany, 2009).

SPnRi3*erm* (*Streptococcus pneumoniae*) *242322 21* *20* *erm*(B) *aphA-3* *erm*(B) *910 78 5 xis int* [27]

1 kb

Tn*1116* (*Streptococcus pyogenes*) *erm*(B) *IS1216 tet*(M) *6 910 78 tndX* [39]

Tn*5397* (*Clostridium difficile*) *242322 21* *20 191817 16* *15* *14* intron *14/13 25/26 tet*(M) *6 910 78 tndX* [47] M

Tn*916S* (*Streptococcus intermedius*) *242322 21* *20 191817 16* *15* *14 13 12 tet*(S) *6 910 78 5 xis int* [24] M

Tn*916* (*Enterococcus faecalis*; see also Table 1) *242322 21* *20 191817 16* *15* *14 13 12 tet*(M) *6 910 78 5 xis int* [23] M

Tn*6002* (*Streptococcus cristatus*) *242322 21* *20* *erm*(B) *191817 16* *15* *14 13 12 tet*(M) *6 910 78 5 xis int* [37] M

Tn*6003* (*Streptococcus pneumoniae*) *242322 21* *20* *erm*(B) *aphA-3 erm*(B) *191817 16* *15* *14 13 12 tet*(M) *6 910 78 5 xis int* [27] M

Tn*1545* (*Streptococcus pneumoniae*) *242322 21* *20* *erm*(B) *aphA-3 erm*(B) *191817 16* *15* *14 13 IS1239 12 tet*(M) *6 910 78 5 xis int* [27] M

Tn*3872* (*Streptococcus pneumoniae*) *242322 21* *20 191817 16* *15* *14 13 12 tet*(M) *6 9* *erm*(B) *9 10 78 5 xis int* [31, 39] M

Tn*2009* (*Streptococcus pneumoniae*) *242322 21* *20 191817 16* *15* *14 13 12 tet*(M) *6* *mef*(E) *910 78 5 xis int* [32] M

Tn*2010* (*Streptococcus pneumoniae*) *242322 21* *erm*(B) *20 191817 16* *15* *14 13 12 tet*(M) *6* *mef*(E) *910 78 5 xis int* [43]

Tn*2017* (*Streptococcus pneumoniae*) *242322 21* *20 191817 16* *15* *14 13 12 tet*(M) *6* *mef*(E) *6 9 erm*(B) *910 78 5 xis int* [69]

Tn*6009* (*Klebsiella pneumoniae*) *merA* *242322 21* *20 191817 16* *15* *14 13 12 tet*(M) *6 910 78 5 xis int* [26] M

Tn*5386* (*Enterococcus faecium*) col adhesion *2322* *21* *20 19 18 1716* intron *16 15* *14 13* putative *spa* operon *7 8 xis int5386* [19] M

46

Figure 1-10 Genetic structure of several Tn*916*/Tn*916*-like elements.

The name of each element is listed at the left end with the species in which it was first isolated from shown in bracket. Arrow boxes represent the orfs and their transcriptional orientation. Key modules are depicted in specific colour as shown above. Each member contains at least two conserved modules; conjugation module (consist of conjugation related genes shown in light blue) and the regulation module (consist of regulation related genes shown in green). Most of the members carry tetracycline resistance gene (grey). Figure is reproduced with permission (Roberts & Mullany, 2009).

1.6.1.1 Regulation of Tn*916*

Excision of Tn*916* from the chromosome can be induced by tetracycline. Tetracycline inhibits protein synthesis and as a result, promotes the transfer of Tn*916* by de-repressing transcription of the regulatory genes. These include o*rf7*, *orf8*, *orf9* and *orf12* (Su *et al.*, 1992, Celli & Trieu-Cuot, 1998, Roberts & Mullany, 2009). The organisation of the Tn*916* regulatory module is conserved among other Tn*916*-like elements. This strongly suggests an essential role in conjugative transposon transcriptional regulation (Roberts & Mullany, 2009).

Su *et al.*, (1992) proposed a Tn*916* regulatory model where the downregulation or upregulation of gene expression is *tet*(M)-dependent (Su *et al.*, 1992, Celli & Trieu-Cuot, 1998) (Figure 1-11). In the absence of tetracycline, *tet*(M) transcript that is initiated by the promoter upstream of *tet*(M) (P*tet*(M)), will terminate at the palindromic sequence located within the *orf*12. When this occurs, P*orf*9 will transcribe *orf*9 efficiently to produce a repressor that will downregulate P*orf*7. Consequently, this will lead to a basal level of transcription of *orf*7, *orf*8 and downstream genes. In the presence of tetracycline, *tet*(M) transcript will continue to *orf7* and extend to *orf8*. Overexpression of *orf7* and *orf8* will produce proteins that will upregulate the *orf7* promoter (P*orf7*). Transcripts originating from *orf7* are then elongated through the recombination region that will activate the transcription of *xisTn* and *intTn* to promote excision. Excision will initiate the re-circularisation of Tn*916*, and subsequently, transcription will extend through conjugation genes to promote transfer (Su *et al.*, 1992, Celli & Trieu-Cuot, 1998).
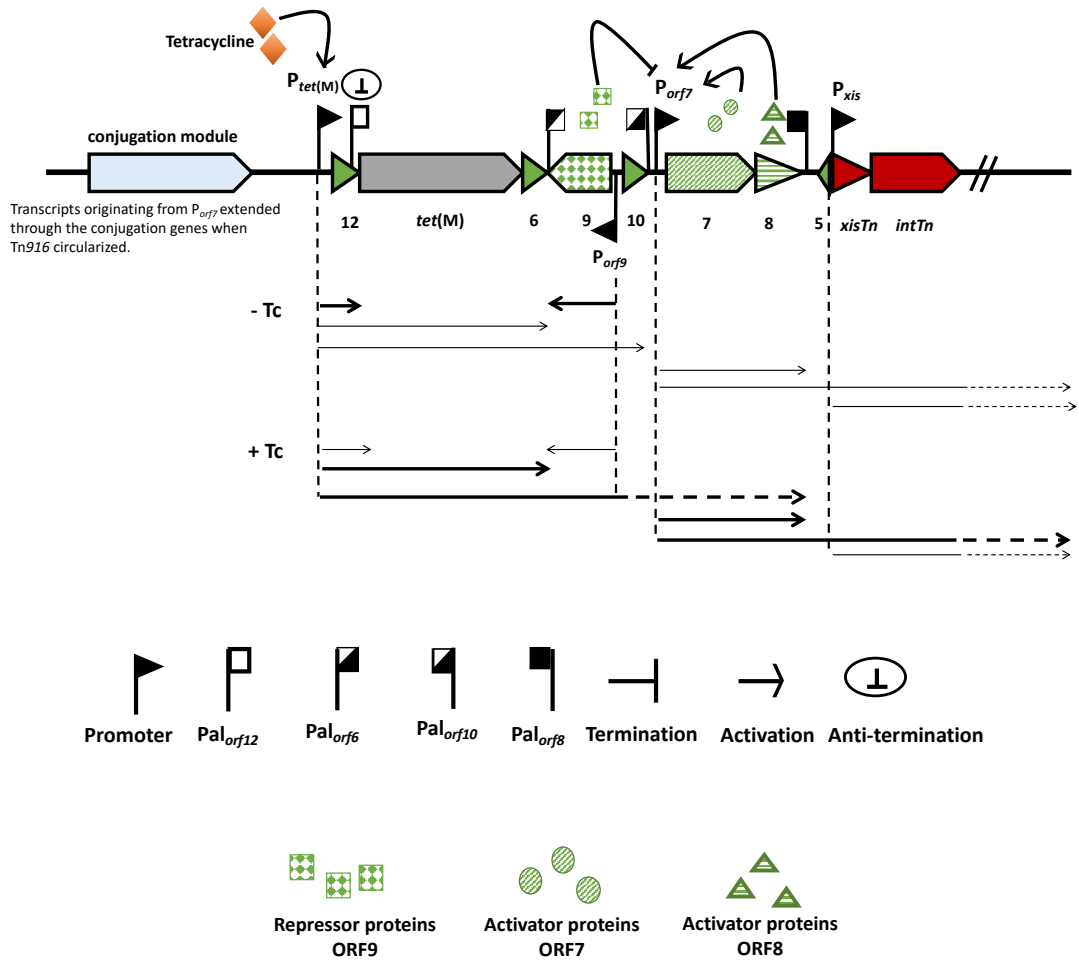
Figure 1-11 Regulation control of Tn*916*.

Orfs are depicted by coloured block arrows, pointing towards transcriptional direction. The thick black arrows represent high transcription level, while the thin black arrows represent low transcription level. Regulation activity are shown in the absence (-Tc) or in the presence of tetracycline (+Tc). Figure is drawn and adapted from Celli and Trieu-Cuot (1998).

It was hypothesised that the regulation of Tn*916* is controlled by the amount of charged tRNA molecules and any cell-damaging factors that could increase the level of these molecules, will trigger the upregulation of the element (Roberts & Mullany, 2009). In another study by (Seier-Petersen *et al.*, 2014), it was revealed that the regulation of Tn*916* does not necessarily depend on the presence of tetracycline. A *gusA* reporter construct was fused to a 450 bp fragment of Tn*916*, which includes the P*tet*(M), *orf12* and the terminator sequences. By measuring the GusA activity of the *B. subtilis* containing the constructs, it was observed that the expression of *gusA* increased upon exposure to cell-damaging biocides; hydrogen peroxide, ethanol, sodium hypochlorite and chlorhexidine digluconate (Seier-Petersen *et al.*, 2014). This is in line with a recent study by Scornec *et al.,* (2017) where the transfer of Tn*916* can be induced not exclusively by tetracycline but also to a variety of other antibiotics mostly belonging to the MLS group (Scornec *et al.*, 2017). These data support the hypothesis of Roberts and Mullany (2009) that any damage caused to the biological systems of bacteria, which affect translation could result in the accumulation of charged tRNA molecules. Upon sensing this stress condition, the element responds by upregulating its transcriptional activity.

1.6.1.2 Replication of Tn*916*

When Tn*916* is induced, a site-specific recombinase (Int*Tn)* will catalyse excision to form a circular intermediate (CI), which also acts as a substrate for

conjugative transfer (Marra & Scott, 1999). In this condition, Tn*916* encoded conjugative relaxase will recognise the *oriT(916)* and nick one strand by attaching itself to the 5' end of the DNA which will unwind, forming a single strand DNA called a transfer-DNA. Through a type IV secretion system (T4SS), the linear single-stranded transfer-DNA will be transported to the recipient cell. Here the transfer-DNA will be recircularised (Fig. 1-12) (Jaworski & Clewell, 1995, Johnson & Grossman, 2015).
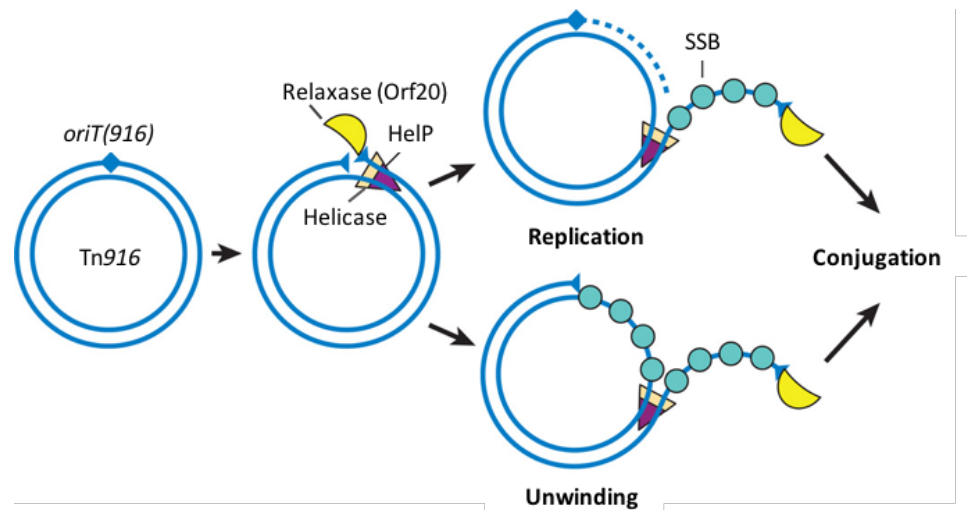
Figure 1-12 A schematic representation of double stranded circularised Tn*916* DNA (double blue lines) nicked at *oriT(916)* by the relaxase Orf20.
The relaxase is covalently attached to the free 5' end of the nicked strand that will be transferred during conjugation. At the same time, a helicase together with two putative helicase processivity factors (HelP; Orf22 and Orf23) interact to form a complex with single-stranded binding protein (turquoise circle; SSB) to unwind the single-stranded DNA. Unwinding of Tn*916* is required for replication and conjugation of Tn*916.* Adapted from Johnson and Grossman (2015).

Wright and Grossman (2016), have proved that replication depends on the same relaxase encoded by the *orf20* of Tn*916,* and it recognised the same *oriT(916)* as an origin of replication. Furthermore, *orf22* and *orf23* have also been demonstrated to encode two putative helicase processivity factors homologues that interact with relaxase to form a complex that facilitates the unwinding of the DNA strands after relaxase nicking. This processive unwinding mechanism is analogous to the mechanism used in ICE*Bs1* (Fig. 1-12) (Lee *et al.*, 2010). Similar to previously described rolling circle mechanism, a functional single strand origin of replication (*sso916*) has also been identified in Tn*916* (Fig. 1-13) (Wright & Grossman, 2016).
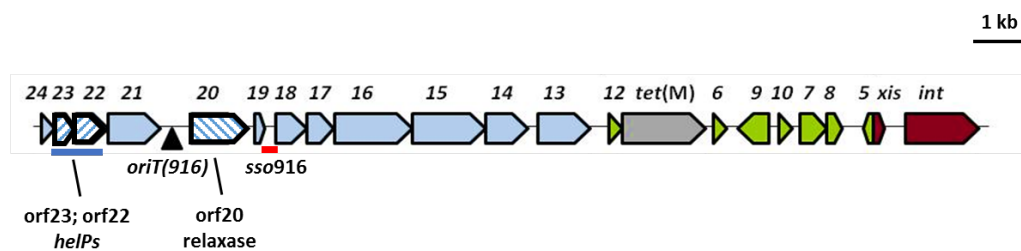
Figure 1-13 A schematic representation of Tn*916* genetic map.

The rectangular arrows represent the genes in each module and the direction of transcription. The black triangles and thick red line indicate the position of *oriT(916)* and *sso*916, respectively. The *orf20* encodes for relaxase, while *orf23* and *orf22* encode homologues of the helicase processivity factor(s) (HelP from ICE*Bs1*). Figure is drawn and adapted from Roberts & Mullany (2009).

It was long thought that Tn*916* is incapable of autonomous replication. However, (Wright & Grossman, 2016) reported that Tn*916* can replicate autonomously using a similar rolling circle mechanism as some phages and plasmids. The latest study by (Lunde *et al.*, 2019) support this finding where

the result of their digital droplet-PCR (ddPCR) assay showed that the percentage of Tn*916* circular intermediate (CI) was higher than the number of detected *B. subtilis* BS34A genomes in the presence of 10 ug/mL tetracycline. In the absence of tetracycline, the percentage of the circularisation ratio was 0.4% and in the presence of tetracycline at 5 ug/mL and 10 ug/mL, it increased to 9.8% and 113%, respectively. Therefore, it was suggested that at least, at this range of tetracycline concentrations, autonomous replication of Tn*916* occurred (Lunde *et al.*, 2019). This is in line with a previous study by (Scornec *et al.*, 2017) where it was shown that sub inhibitory concentrations of tetracycline can act as an inducer for the excision, recircularization, replication and transfer of Tn*916* .

1.6.1.3 Conjugation of Tn*916*

Very little information is available regarding the specific T4SS of Tn*916*. However, Tn*916* encodes proteins that are homologous to those in other genetic elements whose T4SS have been well characterised; such as the *B. subtilis* conjugative transposon; ICE*Bs1*. The T4SS system in ICE*Bs1* requires these three major components to enable its transfer activity; (i) Cell wall hydrolase (CwlT), (ii) Coupling protein and ATPase (Con Q and ConE) and (iii) Membrane channel proteins (ConB, ConC, ConD, and ConG) (DeWitt & Grossman, 2014, Leonetti *et al.*, 2015, Auchtung *et al.*, 2016). In Fig. 1-14, the working model of ICE*Bs1* T4SS together with the homologs of these conserved proteins with Tn*916* proteins (Orf13, Orf14, Orf15, Orf16, Orf17, Orf19 and Orf21) as well as with the well characterised T4SS proteins of *A.*

tumefaciens pTi plasmid (VirB1, VirB3, VirB4, VirB6, VirB8 and VirD4) are shown.

The cell wall hydrolase (CwlT) contains two catalytic domains; muramidase and peptidase that makes it efficient to degrade the thick cell wall of Gram-positive bacteria. Also, CwlT is needed in assisting the assembly of the translocation channel (DeWitt & Grossman, 2014). During conjugation, a coupling protein (Con Q) will transfer the relaxase-bounded transfer-DNA (T-DNA) through the translocation channel into the recipient cell. This is assisted by Con E, the conserved ATPase of ICE*Bs1* T4SS system. Con E energise DNA transfer by forming a doughnut-shaped hexamer through ATP hydrolysis. In relation to Tn*916*, Con Q and ConE is homologous to Orf21 and Orf16, respectively.

The ICE*Bs1* T4SS translocation channel is a model composed of four main putative integral membrane proteins which are; ConB, ConC, ConD and ConG (Fig. 1-14). Although these are likely to be the major ones, there is a possibility that there are yet to be identified proteins that make up the membrane channel component (Leonetti *et al.*, 2015). Similar proteins are found in Tn*916* and Tn*916*-like elements and other ICEs of Gram-positive bacteria (Alvarez-Martinez & Christie, 2009, Berkmen *et al.*, 2010). ConB is homologous to conjugation protein Orf13 in Tn*916*. ConC is a putative integral membrane protein with two predicted transmembrane helices that is homologous to Orf19 in Tn*916*. Alvarez-Martinez and Christie (2009) have reported that ConC and its homologs to be Gram-positive specific T4SS protein. Moreover, ConD and ConG are analogous to Orf 17 and Orf 15 of Tn*916*, respectively (Alvarez-

Martinez & Christie, 2009). The exact roles of these putative integral proteins within T4SS have not been tested biochemically. But, based on the function of their analogues of T4SS proteins (VirB3, VirB6, and VirB8) in Gram-negative bacteria, ConG and ConB, might be the major components of ICE*Bs1* membrane complex (Bhatty *et al.*, 2013). ConG is a very large polytopic protein (815 aa) followed by ConB in a moderate size of 354 aa. In contrast, ConC and ConD, which are smaller in size, are predicted to be involved in scaffolding and/or assembly factors (Leonetti *et al.*, 2015).

	Coupling protein and ATPase		Membrane channel proteins				Cell wall hydrolase
Vir homolog or counterpart	VirD4	VirB4 ATPase	VirB8-like	Vir B3-like	VirB6-like	?	VirB1
Tn*916* homolog or counterpart	Orf21	Orf16	Orf13	Orf 17	Orf15	Orf 19	Orf14
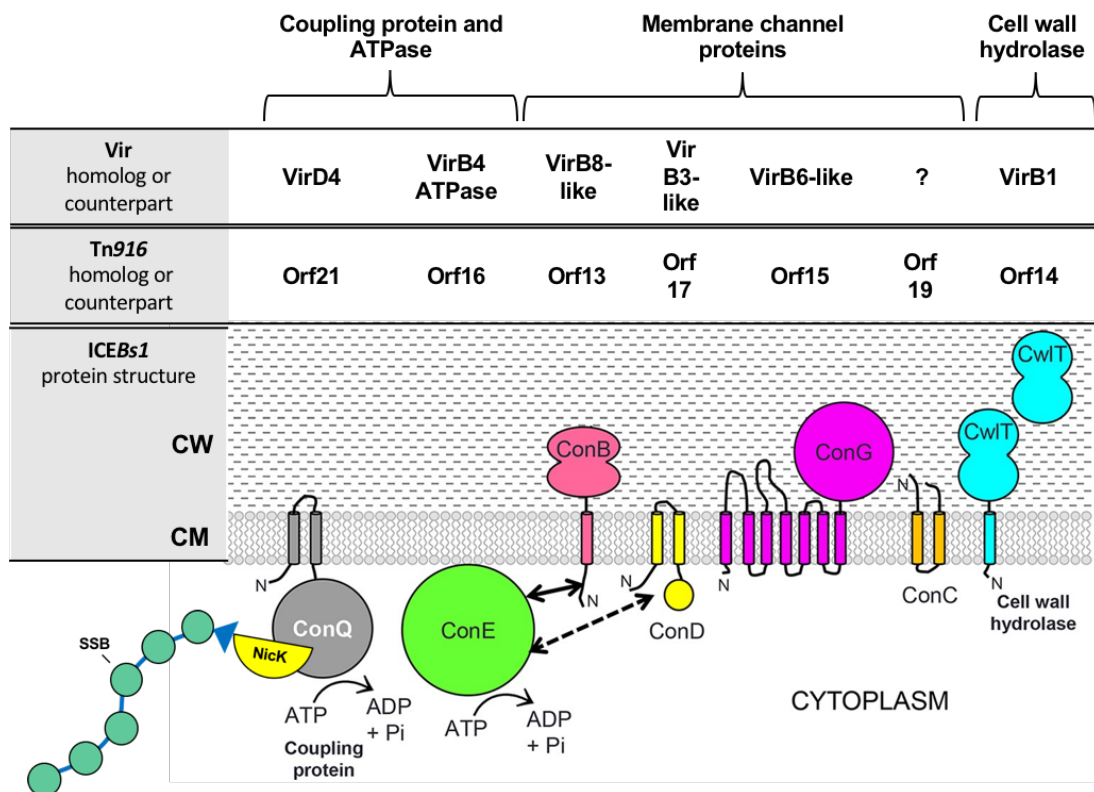
Figure 1-14 Working model of the ICE*Bs1* T4SS.

This speculative model relies on data from other T4SSs. The single-stranded conjugative DNA is shown in blue (forming a complex with single stranded binding protein (green circle)), covalently attached to the NicK relaxase in yellow. The

presumed coupling protein ConQ likely delivers NicK and associated T-DNA to the membrane-associated T4SS. ConB and ConG may make up the bulk of the membrane channel. The ConE ATPase may provide energy for T4SS assembly and/or DNA transfer. The Tn*916* and the *A. tumefaciens* pTi T4SS homologs or counterpart protein are listed above of each ICE*Bs1* illustrated T4SS protein structure. CW; cell wall, CM; cell membrane. Figure is drawn and adapted from Auchtung *et al.* (2016).

1.6.1.4 Recombination of Tn*916*

The transposition of Tn*916* begins with the excision of the transposon from the donor mediated by the staggered cleavages on both ends of the element. These staggered cleavages require two transposon-encoded proteins, the integrase (Int*Tn*) and excisionase (Xis*Tn*) (Poyart-Salmeron *et al.*, 1990, Rudy *et al.*, 1997). The open reading frames of *intTn* and *xisTn* are located in the recombination module approximately 4 kb extending from the stop codon of the *tet*(M) gene to the right terminus of Tn*916* (Fig. 1-9) (Jaworski *et al.*, 1996). The function of these open reading frames has been determined based on their homology to those of lambdoid phages (Poyart-Salmeron *et al.*, 1990). Post-excision, a circular intermediate with a mismatched joint region termed the coupling sequence is formed. These coupling sequences variously consist of five to seven bp heteroduplex originated from the donor DNA bases that flanked the transposon (Caparon & Scott, 1989, Scott & Churchward, 1995, Manganelli *et al.*, 1996). Insertion or integration of Tn*916* involves the reverse of this process where the heteroduplex is either resolved or undergoes repair in the recipient cell (Manganelli *et al.*, 1997). However, studies of the joint of Tn*916* termini in circular intermediates (CI) formed in both *E. faecalis* and *E.*

coli demonstrated that it is not always a heteroduplex. It has been shown that in *E. faecalis*, only homoduplex joint has been found in the CI (Manganelli *et al.*, 1997). While in *E. coli,* half of the CI contained a heteroduplex joint while the other half had a homoduplex joint (Manganelli *et al.,* 1997).

Although both Int*Tn* and Xis*Tn* are required for excision, only Int*Tn* is needed for integration (Storrs *et al.*, 1991, Marra & Scott, 1999). The integration of Tn*916* and Tn*916*-like family members does not generate duplication or replication of the target sequence (Caparon & Scott, 1989). Tn*916* has multiple target sites, which is reported to be an AT-rich region (Scott *et al.*, 1994). Mullany *et al.* (2012) have demonstrated in *C. difficile* strains 630 and R20291, Tn*916* preferentially integrates into the genome at an intergenic region, with a consensus motif sequence of 5'-TTTTA[AT][AT][AT][AT]AAAA-3' (Mullany *et al.*, 2012).

1.7   Rho-independent Terminators and the Mechanism of Intrinsic Termination

Transcription occurs in three major stages; initiation, strand elongation and termination. During the initiation process, RNA polymerase binds to the specific promoter sequence to form a small open complex. Elongation is a process where the core polymerase will catalyse the polymerisation of ribonucleoside 5'-triphosphates (NTPs) into RNA. RNA synthesis continues until RNA polymerase encounters a signal that tells it to stop, or terminate,

transcription. In prokaryotes, this signal can take two forms, rho-independent and rho-dependent (Ray-Soni *et al.*, 2016).

The rho-dependent signal relies on the rho factor that will bind at the specific rho-binding site within the mRNA known as the Rho utilization site (*rut*). Rho factor is the member of the family of ATP dependent hexameric helicases. It acts by unwinding the RNA transcript from the DNA template in 3' to 5' direction. Once the rho factor bounded to the binding site, it will start to translocate along the nascent transcripts towards the RNA polymerase. When it catches up with the RNA polymerase at the transcription elongation complex (TEC), the rho factor will initiate the separation of RNA transcripts and the template strand, that eventually will lead to the releasing of RNA molecule, ending the transcription process (Boudvillain *et al.*, 2013). Rho-dependent terminators are often found in bacterial genomes, comprising about 20-30% of transcription terminators. In *E.coli*, 50% of the transcription terminators are rho-dependent (Ciampi, 2006).

The rho-independent signal is found on the DNA template strand and consists of a region that contains a section that can form a secondary structure known as an intrinsic or rho-independent terminator. The rho-independent terminator is defined as a palindromic sequence that can form a hairpin or stem-loop structure followed by a stretch of thymidine residues (Lynn *et al.*, 1988, d'Aubenton Carafa *et al.*, 1990, Ermolaeva *et al.*, 2000, Lesnik *et al.*, 2001). For efficient transcriptional termination, both of these structural elements are required, where the Gibbs free energy (ΔG) (Lynn *et al.*, 1988) and the properties of T-stretch (Christie *et al.*, 1981), reflects the hairpin stability. The

Gibbs free energy (∆G) of an RNA secondary structure is the sum of individual energy contributions from loops, stacked base pairs and bulges (Zuker, 2003).

rho-independent terminators are commonly found at the end of a transcript but can also play a role as transcriptional attenuators when located between the genes of transcriptional units. When situated at the regulatory sites, these intrinsic terminators can affect the relative rate of translational activity (Wilson & von Hippel, 1995). Although the stability of the RNA secondary structure is usually proportional to a high amount of the G-C content in the stem followed by higher numbers of uridine, d'Aubenton Carafa *et al*. (1990) reported that this is not necessarily the case. They have identified RNA hairpin structures with equal thermodynamic stability that shows a very different termination efficiency. They also suggested that the sequences upstream and downstream of the terminator might contribute to differences in termination efficiency (d'Aubenton Carafa *et al.*, 1990).

The role of rho-independent terminators in intrinsic terminations has been proposed initially from the kinetic view of elongation: RNA polymerase will transcribe the inverted repeat region into RNA, and the inverted repeats in RNA will fold back on itself to form the hairpin loop structure (Figure 1-15) (Farnham & Platt, 1980, Platt, 1986).The formation of the hairpin structure will cause the RNA polymerase to pause or stall and eventually, this will create an opportunity for termination to occur (von Hippel & Yager, 1992). This is because the poly-U residue in the nascent transcript forms a very weak base-paired structure with the template DNA. This, coupled with stalled polymerase, will cause instability in the RNA-DNA hybrid resulting in the release of the

mRNA transcript and disassociation of RNA polymerase, terminating the transcription (Farnham & Platt, 1980, Martin & Tinoco, 1980, von Hippel & Yager, 1992).

The major gaps in our understanding of Tn*916* and the members of its family lie within the regulation and conjugation modules of the element. As mentioned above, Tn*916* can be found integrated into the bacterial chromosome or excised as circularised intermediate element. Whilst the element is integrated in the bacterial chromosome, any transcription read through from the promoter upstream of the element could affect the transcription of the conjugation genes, and if that happens, unregulated transfer may occur which may be deleterious to the cell. Therefore, we hypothesize that there must be a control mechanism to regulate the transcriptional activity of these conjugation genes; and this control mechanism could be a rho-independent terminator. Therefore, in this study, we aimed to search for terminators located upstream of the conjugation module of Tn*916* and Tn*916*-like elements as stabilisation in the genome demands that there is one there.
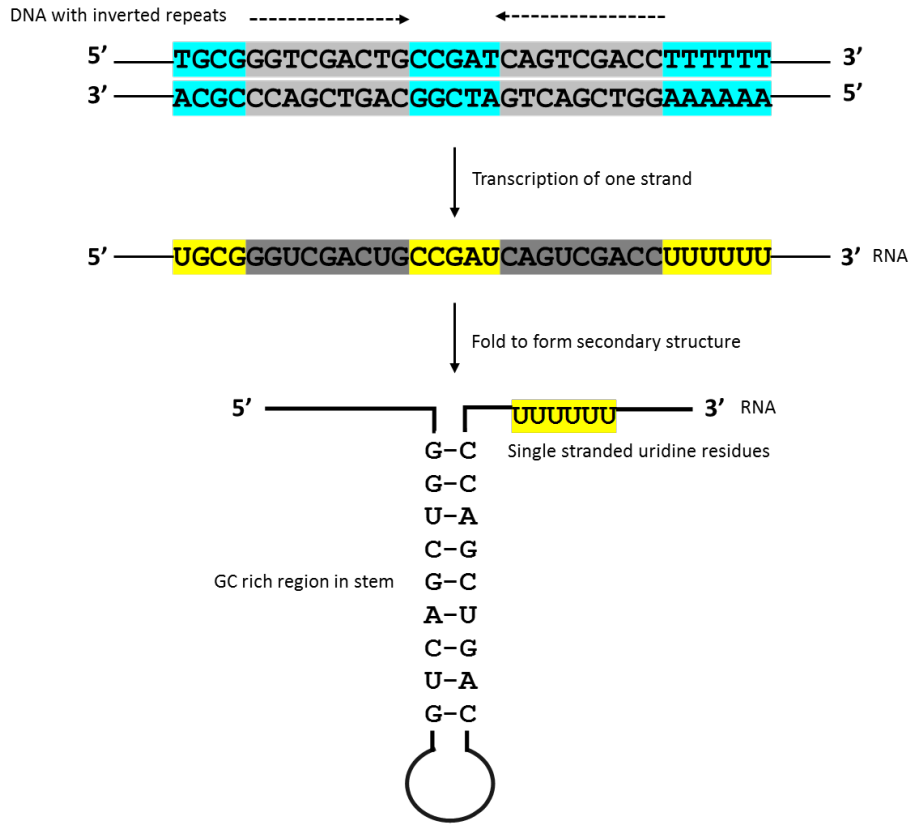
Figure 1-15 Intrinsic rho-independent terminator structure.

The stable hairpin structure is consisting of a GC rich region in stem and the unstable transcript of the poly-U residue.

1.8 Aims of the Study

To address the concerns described above, we pursued the following three specific aims in this study:

1. To identify and investigate the activity of the putative transcriptional terminators located within the conjugative module of Tn*916*, Tn*5397* and Tn*6000* by using an *in vitro* reporter system.

2. To investigate the biological function of the Tn*916* terminator by generating the Tn*916*△Term.

3. To determine the molecular mechanisms of macrolide resistance in *B. subtilis* mutants.

2 Materials and Methods

2.1 Sources of media, enzymes and reagents

Luria-Bertani (LB) broth and agar were obtained from Difco (Oxford, UK) or Sigma-Aldrich (Dorset, UK). Brain Heart Infusion (BHI) agar and broth were obtained from Oxoid Ltd (Basingstoke, UK). Cation-adjusted Mueller-Hinton Broth 2 (CA-MHB) was obtained from Sigma-Aldrich (Dorset, UK). All restriction enzymes were obtained from New England Biolabs (Hitchin, UK). All antibiotics were obtained from Sigma-Aldrich (Dorset, UK) except for TylAMac™ '469 and TylAMac™ '4083 (new analogues of Tylosin A). Tylosin Analogues Macrofilaricides (TylAMac™) is a macrolide-based antibiotic supplied by The Anti-Wolbachia Consortium (A·WOL) of Liverpool School of Tropical Medicine (LSTM) (Liverpool, UK) (von Geldern *et al.*, 2019).

2.2 Bacterial strains, conjugative transposons, plasmids and growth conditions

All bacterial strains, conjugative transposons and plasmids used in this study are listed in Table 2-1 and Table 2-2. *Escherichia coli* was grown LB broth or on LB agar. *Bacillus subtilis* strains, *Enterococcus casseliflavus* 664.1H1 and *Enterococcus faecium* JH2-2 were grown in BHI broth or BHI agar. All bacterial cultures were supplemented where necessary with tetracycline, chloramphenicol, nalidixic acid and / or rifampicin at 10 µg/ml, ampicillin at 100 µg/ml and fusidic acid at 5 µg/ml unless stated otherwise. All cultures were incubated at 37°C under shaking condition (200 rpm).

Table 2-1 Bacterial strains used in this study

Bacterial strains	Characteristics	Resistance marker	Source or Reference
B. subtilis BS34A	Contains Tn*916*	Tc^R	(Roberts *et al.*, 2003)
B. subtilis BS85A	Contains Tn*916* and pHCMC05/Tn*916* P*tet*(M) PO	Tc^R, Cm^R	(Jasni, 2013)
B. subtilis BS80A	Contains Tn*916* and pHCMC05/Tn*916* P*tet*(M) WT	Tc^R, Cm^R	(Jasni, 2013)
B. subtilis BS34A Tn*916.*T	Contains Tn*916* and pHCMC05/Tn*916.*T P*tet*(M) *gusA*	Tc^R, Cm^R	This study
B. subtilis BS34A Tn*5397.*T	Contains Tn*916* and pHCMC05/Tn*5397.*T P*tet*(M) *gusA*	Tc^R, Cm^R	This study
B. subtilis BS34A Tn*6000.*T	Contains Tn*916* and pHCMC05/Tn*6000.*T P*tet*(M) *gusA*	Tc^R, Cm^R	This study
B. subtilis BS34A A	Contains Tn*916* and pHCMC05/EJC P*tet*(M) *gusA*	Tc^R, Cm^R	This study
B. subtilis BS34A B	Contains Tn*916* and pHCMC05/CTGJ P*tet*(M) *gusA*	Tc^R, Cm^R	This study
B. subtilis BS34A ΔA	Contains Tn*916* and pHCMC05/EJCSub P*tet*(M) *gusA*	Tc^R, Cm^R	This study
B. subtilis BS34A ΔB	Contains Tn*916* and pHCMC05/CTGJSub P*tet*(M) *gusA*	Tc^R, Cm^R	This study
B. subtilis BS34A PO	Contains Tn*916* and pHCMC05/Tn*916* P*tet*(M) PO *Age*I-*Spe*I	Tc^R, Cm^R	This study
B. subtilis BS34A VO	Contains Tn*916* and pHCMC04	Tc^R, Cm^R	This study
B. subtilis BS34A *rplV*^{WT}	Contains Tn*916* and pHCMC04/*rplV*^{WT}	Tc^R, Cm^R	This study

B. subtilis BS34A *rplV*^{54D}	Contains Tn*916* and pHCMC04/*rplV*^{54D}	Tc^R, Cm^R, Erm^R, T469^R, T4083^R	This study
B. subtilis BS34A *rplV*^{21}	Contains Tn*916* and pHCMC04/*rplV*^{21}	Tc^R, Cm^R, Erm^R, T469^R, T4083^R	This study
B. subtilis CU2189	Recipient strain	none	(Christie *et al.*, 1987)
B. subtilis CU2189 Tn*916*.T	Containsp HCMCO5/Tn*916*.T P*tet*(M) *gusA*	Cm^R	This study
B. subtilis CU2189 Tn*5397*.T	Contains Tn*916* and pHCMCO5/Tn*5397*.T P*tet*(M) *gusA*	Cm^R	This study
B. subtilis CU2189 Tn*6000*.T	Contains pHCMCO5/Tn*6000*.T P*tet*(M) *gusA*	Cm^R	This study
B. subtilis CU2189 B	Contains pHCMCO5/EJC P*tet*(M) *gusA*	Cm^R	This study
B. subtilis CU2189 A	Contains pHCMCO5/CTGJ P*tet*(M) *gusA*	Cm^R	This study
B. subtilis CU2189 ΔB	Contains pHCMCO5/EJCSub P*tet*(M) *gusA*	Cm^R	This study
B. subtilis CU2189 ΔA	Contains pHCMCO5/CTGJSub P*tet*(M) *gusA*	Cm^R	This study
B. subtilis CU2189 PO	Contains pHCMCO5/Tn*916* P*tet*(M) PO AgeI-SpeI	Cm^R	This study
B. subtilis CU2189 Rif^r Nal^r	Recipient strain for filter mating experiments	Rif^R Nal^R	This study
B. subtilis 168	Wild type strain	None	(Kunst *et al.*, 1997)
B. subtilis 168 T469^r	T469 resistant mutant strain	T469^R	This study
B. subtilis 168 T4083^r	See above Mutant strain	T4083^R	This study

B. subtilis 168 Erm^r	Recipient strain	Erm^R	This study
B. subtilis BS34A Tn916ΔTerm	Contains pGEM-T/Tn916ΔTerm	Amp^R,Tc^R	This study
E. faecalis JH2-2	Recipient strain for filter mating experiments	Fa^R, Rif^R	Jacob and Hobbs, 1974
E.coli F1+F2	Contains pGEM-T/UPS+*cat*P	Amp^R, Cm^R	This study
E. coli F3+F4	Contains pGEM-T/DS1+DS2	Amp^R	This study
E.coli Tn916ΔTerm	Contains pGEM-T/Tn916ΔTerm	Amp^R, Cm^R, Tc^R	This study
E.coli pGEM-T/*rplV^{WT}*	Contains pGEM-T/*rplV^{WT}*	Amp^R	This study
E.coli pGEM-T/*rplV^{21D}*	Contains pGEM-T/*rplV^{21D}*	Amp^R	This study
E.coli pGEM-T/*rplV^{54D}*	Contains pGEM-T/*rplV^{54D}*	Amp^R	This study
E.coli pHCMC04/*rplV^{WT}*	Contains pGEM-T/*rplV^{WT}*	Amp^R	This study
E.coli pHCMC04/*rplV^{21D}*	Contains pGEM-T/*rplV^{21D}*	Amp^R	This study
E.coli pHCMC04/*rplV^{54D}*	Contains pGEM-T/*rplV^{54D}*	Amp^R	This study
E. coli (α-select silver efficiency)	Competent cells	Amp^R	Bioline (London, UK)

Abbreviations: Tc^R, tetracycline-resistant; Cm^R, chloramphenicol-resistant; Erm^R, erythromycin-resistant; Amp^R, ampicillin-resistant; T469^R, TylAMac™ '469-resistant; T4083^R, TylAMac™ '4083-resistant; Rif^R, rifampicin-resistant; Fa^R, fusidic acid-resistant; P*tet*(M) WT, wild type promoter construct; P*tet*(M) PO, promoter only construct; Tn916.T, Tn916 terminator construct; Tn5397.T, Tn5397 terminator construct; Tn6000.T, Tn6000 terminator construct; *rplV^{WT}*, wild

type *rpIV*; *rpIV*[54D], *rpIV* with 54 bp duplication; *rpIV*[21], *rpIV* with 21 bp duplication; EJC, joint ends of Tn*916* plus terminator construct (B); CTGJ, Tn*916* terminator region plus flanking DNA construct (A); EJCSub, joint ends of Tn*916* plus mutated terminator construct (ΔB); CTGJSub, mutated Tn*916* terminator region plus flanking DNA construct (ΔA) ; P*tet*(M) PO AgeI-SpeI, promoter only construct with added restriction sites of *Age*I and *Spe*I; Tn*916*ΔTerm, Tn*916* with deleted terminator; UPS+*cat*P, upstream region which is homologous to the BS34A genome (UPS) and chloramphenicol-resistance gene (*catP*); DS1+DS2, downstream regions flanking the Tn*916* terminator (homologous to the BS34A genome and Tn*916*).

Table 2-2 Plasmids and conjugative transposon used in this study.

Plasmids or Transposons	Characteristics	Resistance marker	Source / Reference
pRPF185	$gusA^+$, Contains inducible tetracycline promoter and *catP*.	Cm^R, Tm^r	Supplied by Dr Haitham Hussain (Eastman Dental Institute, UCL); Fagan *et al.,* 2011
pGEM-T® Easy	Cloning vector	Amp^R	Promega (Southampton, UK)
pGEM-T® /UPS+*cat*P	Plasmid contains mutant cassette	Amp^R, Cm^R	This study
pGEM-T® /DS1+DS2	Plasmid contains mutant cassette	Amp^R	This study
pGEM-T® /Tn*916*△Term	Plasmid contains mutant cassette	Amp^R, Cm^R	This study
pGEM-T® /*rplV*^{WT}	Plasmid contains *rplV*^{WT}	Amp^R	This study
pGEM-T® /*rplV*^{21D}	Plasmid contains *rplV*^{21D}	Amp^R	This study
pGEM-T® /*rplV*^{54D}	Plasmid contains *rplV*^{54D}	Amp^R	This study
pHCMC04	Shuttle vector, contains xylose (*xyl*) inducible promoter	Amp^R, Cm^R	BGSC (Ohio, US); (Nguyen *et al.*, 2005)
pHCMC04/*rplV*^{WT}	Shuttle vector, contains *rplV*^{WT}	Amp^R, Cm^R	This study
pHCMC04/*rplV*^{54D}	Shuttle vector, contains *rplV*^{21D}	Amp^R, Cm^R	This study
pHCMC04/*rplV*^{21}	Shuttle vector, contains *rplV*^{54D}	Amp^R, Cm^R	This study

pHCMC05	Shuttle vector	Amp^R, Cm^R	BGSC (Ohio, US); (Nguyen *et al.*, 2005)
pHCMC05/Tn*916* P*tet*(M) WT	Shuttle vector	Amp^R, Cm^R	(Jasni, 2013)
pHCMC05-P*tet*(M)-PO	Shuttle vector	Amp^R, Cm^R	(Jasni, 2013)
pHCMC05/Tn*916*.T P*tet*(M) *gusA*	Shuttle vector	Amp^R, Cm^R	This study
pHCMC05/Tn*5397*.T P*tet*(M) *gusA*	Shuttle vector	Amp^R, Cm^R	This study
pHCMC05/Tn*6000*.T P*tet*(M) *gusA*	Shuttle vector	Amp^R, Cm^R	This study
pHCMC05/EJC P*tet*(M) *gusA*	Shuttle vector	Amp^R, Cm^R	This study
pHCMC05/CTGJ P*tet*(M) *gusA*	Shuttle vector	Amp^R, Cm^R	This study
pHCMC05/EJCSub P*tet*(M) *gusA*	Shuttle vector	Amp^R, Cm^R	This study
pHCMC05/CTGJSub P*tet*(M) *gusA*	Shuttle vector	Amp^R, Cm^R	This study
pHCMC05-P*tet*(M)-PO-*Age*I-*Spe*I	Shuttle vector	Amp^R, Cm^R	This study
Tn*916*	Conjugative transposon	Tc^r	(Franke & Clewell, 1981)
Tn*916*ΔTerm	Mutant conjugative transposon (with terminator deletion)	Tc^r	This study

Abbreviations: BGSC, Bacillus Genetic Stock Center; Cm^R, chloramphenicol-resistant; Amp^R, ampicillin-resistant; Tc^R, tetracycline-resistant; P*tet*(M) WT, wild type promoter construct; P*tet*(M) PO, promoter only construct; Tn*916*.T, Tn*916* terminator construct; Tn*5397*.T, Tn*5397* terminator construct; Tn*6000*.T, Tn*6000* terminator construct; *rplV*^{WT}, wild type *rplV*; *rplV*^{54D}, *rplV* with 54 bp duplication; *rplV*^{21}, *rplV* with 21 bp duplication; EJC, joint ends of Tn*916* plus terminator construct (B); CTGJ, Tn*916* terminator region plus flanking DNA construct (A); EJCSub, joint ends of Tn*916* plus mutated terminator construct (ΔB); CTGJSub,

mutated Tn*916* terminator region plus flanking DNA construct (ΔA) ; P*tet*(M) PO *Age*I-*Spe*I, promoter only construct with added restriction sites of *Age*I and *Spe*I; Tn*916*△Term, Tn*916* with deleted terminator; UPS+*cat*P, upstream region which is homologous to the BS34A genome (UPS) and chloramphenicol-resistance gene (*catP*); DS1+DS2, downstream regions flanking the Tn*916* terminator (homologous to the BS34A genome and Tn*916*).

2.3 Storage of bacterial strains

All bacterial stocks were maintained in 1 ml aliquots of 20% (v/v) sterilised glycerol in LB at -80°C.

2.4 Molecular biology techniques

2.4.1 Genomic DNA purification

Genomic DNA purification was carried out using the Gentra Puregene Yeast/Bact Kit (Qiagen, UK) following protocol specifically for Gram-positive bacteria with slight modifications (Page 51- 52). For efficient isolation of genomic DNA from Gram-positive bacteria, lysozyme (Sigma Aldrich, UK) and mutanolysin (Sigma Aldrich, UK) were added to assist the cell lysis at step 6. Gram-positive bacteria are more resistant to lysis due to the thick peptidoglycan layer in their cell wall. Mutanolysin is a muramidase that hydrolyze the β-1,4 glycoside linkages between GlcNAc and MurNAc of the glycan backbone. The use of the enzyme mutanolysin in addition to lysozyme leads to increased yield of bacteria lysis (Gill *et al.*, 2016). Bacterial culture was grown overnight, and 500 µl of it was aliquoted into sterile 1.5 ml microcentrifuge tubes. The tubes containing overnight culture were placed on ice for 1 min. The cells were then subjected to centrifugation at 14680 x *g* (13000 rpm) by using table-top microcentrifuge (Eppendorf 5415 D) for 1 min to form a tight pellet. The supernatant was discarded carefully using a pipette, and 300 µl of Cell Suspension Solution was added to the pellet. For modification, a mixture of 150 µl of TE Buffer, 10 µl of 100 U mutanolysin and 40 µl of lysozyme (25 mg/mL) were added to the cell suspension on top of 1.5

μl of Lytic Enzyme Solution. The mixture was mixed by inverting the tube for 25 times and incubated at 37°C for 30 min to damage the cell walls. The cells were then centrifuged for 1 min at the same previous speed and the supernatant discarded.

The cell pellet was resuspended in 300 μl of Cell Lysis Solution and heated at 80°C for 5 min to complete the cell lysis. For RNase treatment, 1.5 μl of RNase A Solution was added to the cell lysate. The sample was mixed by inverting the tube 25 times and further incubated at 37°C for 15-60 min. Subsequently, the sample was cooled on ice for a minute before 100 μl of Protein Precipitation Solution was added. The sample was vortexed vigorously at high speed for 20 sec before centrifugation for 3 min. At this point, the precipitated proteins should be formed (tight white pellet). By pouring, the supernatant was transferred into 1.5 mL microcentrifuge tube pre-added with 300 μl of 100% isopropanol. The sample was mixed by inverting the tube gently for 50 times and centrifuged for 1 min. After centrifugation, a small white pellet (DNA) should be visible. The supernatant was carefully discarded without dislodging the white pellet. The tube was drained on a clean absorbent paper and washed with 300 μl of 70% ethanol. The tube was inverted several times and centrifuged for 1 min. The supernatant was discarded, and the DNA pellet was left to air dry for 5-10 mins. Lastly, the DNA pellets were resuspended in 50-100 μl of molecular grade water (Sigma-Aldrich, UK) and kept at -20 °C until further usage.

2.4.2 Plasmid DNA purification

The Plasmid DNA purification was carried out using the QIAprep Spin Miniprep Kit (Qiagen, UK). Culture of bacteria harbouring plasmids was grown in sterile tubes containing 5 mL of LB or BHI broth supplemented with appropriate antibiotics and incubated overnight for 16-18 hrs at 37°C, 200 rpm. The bacterial cells were pelleted to centrifugation at 4500 x g (5000 rpm) using table-top centrifuge (Eppendorf 5804 R) for 15 min at 4°C. The supernatant was discarded, and the pellet was resuspended with 300 µl of Buffer P1 and transferred to a 1.5 mL microcentrifuge tube. Then, 300 µl of buffer P2 was added, and the tube was mixed by inverting the tube for 4-6 times. Once mixed, 350 µl of buffer N3 was added, and again, the tube was inverted for 4-6 times. Then, the mixture was centrifuged at 14680 x g (13000 rpm) using mini centrifuge (Eppendorf 5415 D). By using a pipette, the supernatant was carefully transferred from the 1.5 mL microcentrifuge tube into the QIAprep spin column. The spin column was centrifuged for 1 min, and the flow-through was discarded. The spin column was added with 500 µl of Buffer PB and centrifuged for 1 min. The flow-through was discarded, and the spin column is washed with 700 µl of Buffer PE and centrifuged again for 1 min. The flow-through was discarded and additional centrifugation was then performed for a minute to remove the residual Buffer PE. The QIAprep spin column was then placed in a new sterile 1.5 ml microcentrifuge tube and the DNA was eluted by adding 30-50 µl of molecular biology grade water (Sigma-Aldrich, UK) to the centre of the membrane. The column was left to stand for 2 min and centrifuged for 1 min. The extracted DNA was kept in -20°C freezer until further usage.

2.4.3 PCR product purification

PCR purification was carried out using the QIAquick PCR Purification Kit (Qiagen, UK). This protocol was done to clean-up PCR products from impurities such as primers, nucleotides, enzymes, and salts. This protocol was designed to purify up to 10 µg of PCR products ranging from 100 bp to 10 kb in size. All of the centrifugation steps were carried out at 14680 x g (13000 rpm) using a mini centrifuge (Eppendorf 5415 D). Five volumes of Buffer PB were added to 1 volume of the PCR sample and mixed by pipetting it up and down. The mixture was then transferred and applied to the QIAquick spin column that has been placed into a provided collection tube. The spin column was centrifuged for 1 min, and flow-through was discarded. Then the column is washed with 700 µl of Buffer PE and centrifuged for a minute. The flow-through was discarded from the collection tube, and centrifuged for an additional 1 min to remove any remaining residual. The spin column was then placed into a new sterile 1.5 ml microcentrifuge tube, and the DNA was eluted by adding 30 µl - 50 µl of molecular biology grade water (Sigma-Aldrich, UK) to the centre of the membrane. The column was left to stand for 2 min and centrifuged for 1 min. The purified DNA was kept in -20°C freezer until further usage.

2.4.4 Agarose gel electrophoresis

DNA fragments were mixed with 6x loading and were separated on 1-2% (w/v) agarose (Bioline, London, UK) gel electrophoresis. Gels were prepared to the relevant percentage in 1X tris-acetate-EDTA buffer, run for 45-60 min at 80-

100V. Nucleic acid was stained with ethidium bromide (Promega, Southampton, UK) at a concentration of 0.5 µg/mL and GelRed™ (Biotium Inc, Cambridge, UK) or GelGreen™ (Biotium Inc, Cambridge, UK) with a 1:10,000 dilution. The HyperLadder 1kb (Bioline, UK) and 1 kb extend marker (NEB, UK) was used as a size reference. The gel was visualised either using a UV transilluminator (320 nm) via Alpha Imager (Innotech Corporation, UK) or LED transilluminator (468 nm) via Pearl Biotech Blue Box (Cambridge Bioscience Ltd, UK). Gel images were produced using Alpha View software (Innotech Corporation, UK).

2.4.5 Gel extraction

DNA extraction from agarose gel was carried out using the QIAquick Gel Extraction Kit from Qiagen (Crawley, UK). It was done to isolate and purify specific DNA fragment(s) based on the size of amplicons. The DNA sample was subjected to the agarose gel electrophoresis, as described in section 2.4.4. The specific DNA band of interest was excised with a clean and sharp scalpel by visualising under UV light. The sliced fragments were then transferred into a 1.5 ml microcentrifuge tube and weighed. Buffer QG was then added following a ratio of 1 volume gel (100 mg ~ 100 µl) to 3 volumes of buffer. The tube was incubated at 50°C and occasionally vortexed until the gel slice(s) completely dissolved. Then, 1 volume of isopropanol was added to the mixture and inverted for a few times. The mixture was transferred and applied to the QIAquick spin column that has been placed into a provided collection tube. The spin column was centrifuged for 1 min, and flow-through was discarded. The column is washed with 700 µl of Buffer PE and centrifuged

again for a minute. The flow-through was discarded from the collection tube and centrifuged for an additional 1 min to remove any remaining residual. The spin column was then placed into a new sterile 1.5 ml microcentrifuge tube and the DNA was eluted by adding 30 µl - 50 µl of molecular biology grade water (Sigma-Aldrich, UK) to the centre of the membrane. The column was left to stand for 2 min and centrifuged for 1 min. The purified DNA was kept in -20°C freezer until further usage.

2.4.6   Restriction endonuclease reactions

All Restriction enzymes were obtained from New England Biolabs (NEB, Hitchin, UK) and used according to the manufacturer's instructions. The standard digestion reactions contain; 1 µl of restriction enzymes (20 U), 1 µl of 10X digestion buffer, 1-5 µl of DNA samples and molecular biology grade water to a total of 10 µl. The reactions were incubated at 37°C for 1 hr. For the High-Fidelity (HF®) Restriction Enzymes (NEB, Hitchin, UK), the reactions were incubated at 37°C for 5-15 mins or overnight. For double digestion, CutSmart buffer or any other compatible buffers were used. For sequencing purposes, the digestion reactions with a total volume of 50 µl were prepared, followed by a clean-up using QIAquick PCR Purification Kit (Qiagen, UK).

2.4.7   Dephosphorylation reaction

Dephosphorylation was carried out to remove the 5' phosphate groups from the end of the digested vector before ligation. This will prevent the self-ligation of the vector that has been previously digested by a single enzyme during a ligation reaction. After the completion of digestion reaction, 1 µl of Calf

Intestinal Alkaline Phosphatase (CIAP) (1 U/μl), 10X reaction buffer and molecular biology grade water was added to the digestion mixture. The mixture was incubated at 37 °C for 30 min in a block heater (Thermo Scientific, UK). Another 1 μl of CIAP (1 U/μl) was then added to the reaction mixture and further incubated for another 30 min. The reaction was purified using the QIAquick PCR Purification Kit (Qiagen, UK), as described in section 2.4.3.

2.4.8 DNA ligation reactions

Ligation of DNA was carried out using T4 DNA ligase (NEB, Hertfordshire, UK) with a molar ratio of 1:3 vector to insert, according to the manufacturer's instructions. The ligation mixture was incubated for 1hr at room temperature or overnight at 4°C and heat-inactivated at 65 °C for 10 min in a block heater (Thermo Scientific, UK).

2.4.9 Preparation of *B. subtilis* competent cells

The *B. subtilis* competent cells were prepared according to the protocol described by Hardy (1985). A single colony of *B. subtilis* was inoculated into 10 ml of SPI broth (Appendix A) and grown at 30°C overnight with shaking (200 rpm) in a 200 ml conical flask. A 10 ml of the overnight culture was then transferred to the 100 ml fresh pre-warmed SPI medium in a 1 L flask to give an OD_{600} reading of about 0.1. The culture was further incubated 37°C with vigorous aeration, and periodic OD readings (OD_{600}) were taken every 30 min to assess cell growth. When the rate of cell growth is seen to depart from exponential, a 10 ml of this culture was transferred to the 90 ml SPII broth (Appendix A) and continue to be incubated for 1.5 h at 37°C, 200 rpm. The

bacterial suspension was centrifuged at 3,000 x *g* (5,000 rpm) for 10 min at 20°C. The bacterial pellet was resuspended carefully in 10 ml supernatant, to which sterile glycerol had been added to 10% (v/v). Aliquots of 500 µl were dispensed on ice and the cells frozen and stored at -80°C until required.

2.4.10 Transformation of *B. subtilis*

For transformation, 500 µl competent cells were thawed at 37°C and mixed with 1-5 µg DNA (vol 5-10 µl) in a sterile 50 ml conical tube. The tube was swirled and incubated at 37°C for 1 h under gentle agitation (50 rpm). Then, 5 ml of fresh LB broth was added, followed by a further incubation at 37°C for 1.5 h with shaking at 200 rpm. For positive selection, 200 µl of transformation mixture were plated onto BHI agar supplemented with appropriate antibiotics. The remaining cells were spun down at 3,000 x *g* (5,000 rpm) for 10 min and supernatant discarded. The pelleted cells were resuspended in 100 µl of fresh LB broth and spread onto BHI agar supplemented with appropriate antibiotics. All plates were incubated at 37°C and checked for growth at 24-48 h.

2.4.11 Transformation of *E. coli*

E. coli transformation was carried out using competent cells "α-select silver efficiency" (Bioline, UK) according to the standard heat-shock transformation protocol. Transformants were selected on LB agar supplemented with appropriate antibiotics depending on the plasmid marker and IPTG/X-gal plates for blue-white screening. All plates were incubated overnight at 37°C.

2.5 Primers synthesis

Sequences of all primers used in this study were listed in Table 2-3. All primers were synthesised by Sigma-Aldrich (Poole, UK) or by Integrated DNA Technologies, Inc. (IDT, UK).

Table 2-3 Primers used in this study

Primer name	Sequence (5'-3')[ab]	Target region	Reference
Primers designed for the construction of terminator constructs			
pHCMC05_P*tet*M_GusA_F	AAAATCGTCTCCCTCCGTTT	pHCMC05/P*tet*M GusA PO construct	This study
pHCMC05_PtetM_GusA_R	TCCGAGCTTCGTCCAAAATA	pHCMC05/P*tet*M GusA PO construct	This study
For*916* PO	ATGGAGGAAAATCACGAATTCCTGC	Tn*916* promoter only construct	(Jasni, 2013)
Rev*916* PO	ACAGATATTCTCCGGATACTTTAGA	Tn*916* promoter only construct	(Jasni, 2013)
For Tn*5397*-T*SpeI*	***ACTAGT*CCATTTGATTTTTCATATCAAGTGGTTTT TGTTATGGAGGAAAATCACGAATTCCTGC**	Q5SDM-Insertion of Tn*5397* terminator sequence	This study
For Tn*6000*-T*SpeI*	***ACTAGT*GACACTTCCAAAGTTGAGGTGTCTTTTT TATGGAGGAAAATCACGAATTCCTGC**	Q5SDM-Insertion of Tn*6000* terminator sequence	This study
For Tn*916*-T*SpeI*	***ACTAGT*GACACTTCAAAAAATGAGGTGTCTATTT TTTTATGGAGGAAAATCACGAATTCCTGC**	Q5SDM-Insertion of Tn*916* terminator sequence	This study
For Tn*916*-RE-*SpeI*	GCGC***ACTAGT*GAAGCAACAGGAGCGTCTTG**	Ligated ends of Tn*916*	This study

Rev Tn*916*-LE-*Age*I	GCGC***ACCGGT***ACTTCCTTTCAAAATCGGGT	Ligated ends of Tn*916* and BS34A genome::Tn*916* junction	This study
REGJ-*Spe*I	GCGC***ACTAGT***CCCGGTCATGAATTGAAAGA	BS34A genome::Tn*916* junction	This study
For*916* PO-*Spe*I-*Age*I	***ACTAGTACCGGT***ATGGAGGAAAATCACGAATTCCTGC	Tn*916* promoter only construct	This study
***916*Sub_F**	GGGGACCCGATTTTGAAAGGAAG	Q5SDM – substitution on T stretch of Tn*916* terminator	This study
***916*Sub_R**	CCCCCGACACCTCATTTTTTGAAGTG	Q5SDM – substitution on T stretch of Tn*916* terminator	This study
***916*Sub2_F**	GGGACCCGATCTCGAGAGGAAGTACC	Q5SDM – nucleotide substitution on Tn*916* terminator	This study
***916*Sub2_R**	CCCCCCGACACCTCATTT	Q5SDM – nucleotide substitution on Tn*916* terminator	This study

Primers designed for SOEing PCR

UPS_F	AGCCAGTAAGGGAACAAAAA	BS34A genomic DNA, fragment 1	This study
UPS-catP-R3	TTGATTTAAGCCCCG*AAAAG*AGTCAAATAGCCATTTCTAC	BS34A genomic DNA, fragment 1 20 bp overlapping with *catP* (fragment 2)	This study

catP_F2	*GTAGAAATGGCTATTTGACT*TTTTAGTTACAGAC AAACCT	*catP* (fragment 2) from pRPF185, 20 bp overlapping with fragment 1	This study
catP_R3_*xhoI*	GCGC***CTCGAG***ACCCGGCAGTTTTTCTTTTT	*catP* (fragment 2) from pRPF185	This study
BSA_F_*xhoI*	GCGC***CTCGAG***GTCTGAGGATTAATGGCTGTGT	BS34A genomic DNA, fragment 3	This study
UPS_BR	ACTTCCTTTCAAAATCGGGTTACCTATTAAATATT CAAATTT	BS34A genomic DNA, fragment 3 overlapping with 22 bp of fragment 4	This study
DS_BF	*AAATTTGAATATTTAATAGGTA*ACCCGATTTTGAA AGGAAGT	Conjugation region of Tn*916*, fragment 4 overlapping with 22 bp of fragment 3	This study
DS_BR	AGACAATCCAGCAGATCAAC	Conjugation region of Tn*916*, fragment 4	This study

Primers for the verification of mutant cassette

UPS_LF	AGCAAAATCTCCAGACGCATA	Mutant cassette	This study
UPS_L2F	CAGGAAAGATAAATAAGAAGCAAAAA	Mutant cassette	This study
HR_F	CGTAGTAAATGCCAACCGAAT	Mutant cassette	This study
HR_R	TGCTGGTCGTACAAAGGAAA	Mutant cassette	This study

catPLR	TGAATGGCGGTTTACAATCA	Mutant cassette	This study
catPLF	CCCTCTCAAATTCAAGTTTATCG	Mutant cassette	This study

Primers for Tn*916*

P*tet*(M)F	ACCAAAGCAACGCAGGTATCT	P*tet*(M) of Tn*916*	This study
P*tet*(M)R	GTGATTTTCCTCCAT	P*tet*(M) of Tn*916*	This study
xisF	ATGAAGCAGACTGACATTCC	*xis*Tn of Tn*916*	This study
xisR	CTAGATTGCGTCCAATGTA	*xisT*n of Tn*916*	This study
ETS_F	ATGGCGGAGCGAAATATCAT	Tn*916* empty target site in BS34A	This study
ETS_R	AGAACGGAATGGCCAGAATA	Tn*916* empty target site in BS34A	This study
916CE_F	AAAAGTGGCGAACGTCAAGT	Tn*916* circular form joint ends	This study
916CE_R	GAATCATGCGTCCTTGTCCT	Tn*916* circular form joint ends	This study
916REO_R	AATTGCCACACATCACTCCA	Tn*916* right end genome junction	This study

Primers for macrolide resistance study

***rplv*_F**	GACGCTAAGAGAGGAGGCTTT	*rplV* of *Bacillus subtilis* strain 168	This study
***rplv*_R**	CCGACTGGATTTACCTTTTGA	*rplV* of *Bacillus subtilis* strain 168	This study
***rplV*F_*Spe*I**	GCGC***ACTAGT***GACGCTAAGAGAGGAGGCTTT	*rplV* of *Bacillus subtilis* strain 168 added with *Spe*I restriction site	This study
***rplV*R_*BamH*I**	GCGC***GGATCC***CCGACTGGATTTACCTTTTGA	*rplV* of *Bacillus subtilis* strain 168 added with *BamH*I restriction site	This study
p-gyrA-f	CAG TCA GGA AAT GCG TAC GTC CTT	*gyrA*	(Chun & Bae 2000)
p-gyrA-r	CAA GGT AAT GCT CCA GGC ATT GCT	*gyrA*	(Chun & Bae 2000)
CMC04F	TCCTTTGTTTATCCACCGAAC	pHCMC04 insert region	This study
CMC04R	TTTCAACCATTTGTTCCAGGT	pHCMC04 insert region	This study
GusF1	TCCATCGCAGCGTAATGCTC	pHCMC05 insert region, within *gus*A	This study
GusF2	TTTTAACGATCAGTTCGCCG	pHCMC05 insert region, within *gus*A	This study

Primers for sequencing

M13F	GTAAAACGACGGCCAG	M13 forward sequencing (pGEM-T easy vector)	Universal
M13R	CAGGAAACAGCTATGAC	M13 reverse sequencing (pGEM-T easy vector)	Universal
27F	AGAGTTTGATCCTGGCTCAG	16S rRNA gene	Universal
1392R	GGTTACCTTGTTACGACTT	16S rRNA gene	Universal

[a]Restriction sites is indicated in bold and italic styles.

[b]Overlappping sequence (for SOEing PCR) is highlighted in grey.

2.6 Standard PCR protocol

PCR amplification was carried out using Biometra T3000 Thermocycler (Biometra, Netherlands) or ProFlex PCR System (Applied Biosystem, UK). For BioMix Red (Bioline, UK), the following thermal profile was used: initial denaturation at 94°C for 2 min, 30-35 PCR cycles [94°C, 1 min; 53-60°C, 1 min (dependent on primer annealing temperature); 72°C, 1-4 min (dependent on expected amplicon size, usually 1 min for each 1 kb is allowed); 1 cycle of 4 min at 72°C and preservation at 4°C until the sample was analysed]. All of the temperature and times are variable. For Q5 High Fidelity 2X Master Mix (NEB, UK), the following thermal profile was used: initial denaturation at 98°C for 30 sec, 25-30 PCR cycles [98°C, 10 sec; 53-60°C, 30 sec (dependent on primer annealing temperature); 72°C, (20-30 sec/kb); 1 cycle of 2 min at 72°C and preservation at 4°C until the sample was analysed. The PCR reaction mixture contains 25 µl of 2X BioMix Red (Bioline, UK) or Q5 High-Fidelity 2X Master Mix (NEB, UK), 2.5 µl of each forward and reverse primers (at a final concentration of 0.5 µM) , and 1 µl of DNA template (approximately 100 ng/µl). The total volume of PCR mixture was made up to 50 µl using distilled water.

2.7 Site-Directed Mutagenesis (SDM)

Site-specific mutagenesis was carried out using the Q5® Site-Directed Mutagenesis Kit from New England Biolabs (NEB) Ltd (Hertfordshire, UK). Specific non-overlapping primers were designed using NEBaseChanger to incorporate insertions, deletions or substitutions in the mutant construct. Primer design for the Q5® Site-Directed Mutagenesis Kit was done manually

or using the NEB online primer design software, NEBaseChanger at http://nebasechanger.neb.com.

2.7.1 Polymerase Chain Reaction (PCR)

The following thermal profile was used for the amplification: Initial denaturation at 96 °C for 30 seconds, denaturation at 96 °C for 10 seconds, primer annealing at 50 -72 °C (based on primer used) for 30 sec and extension at 72 °C for 20-30 seconds/kb. The final cycle included an extension for 2 mins at 72°C to ensure the full extension of the products.

2.7.2 Treatment and enrichment (kinase, ligase and *Dpn*I)

The following KLD reaction mixture was used: 1 µl of PCR product, 5 µl of 2X KLD Reaction Buffer, 1 µl 10X KLD Enzyme Mix and 3 µl nuclease-free water. The KLD mix was incubated for 5 mins at room temperature. The KLD mix contains a blend of kinase; for efficient phosphorylation of the amplicons, ligase; for intramolecular ligation or re-circularisation of the phosphorylated amplicons and *Dpn*I; for the removal of template DNA.

2.7.3 *E. coli* transformation

The transformation was carried out by adding 5 µl of KLD mix to 50 µl of chemically competent cells. The next subsequent step was carried out based on standard *E. coli* transformation protocols as described previously.

2.7.4 SDM product evaluation

Successful transformants harbouring the recombinant plasmid were screened through colony PCR and digestion using the standard protocol as described previously. The mutants were analysed by DNA sequencing. Successful transformants harbouring the recombinant plasmid were screened through colony PCR using specific primer pairs (pHCMCO5-Ptet(M)-GusA-F, pHCMCO5-Ptet(M)-GusA-R, GusA R1 and Gus A RL). These PCR products were subjected to digestion and finally analysed by DNA sequencing.

2.8 Filter-mating

The filter-mating experiments were conducted as described by Roberts *et al*. (2000). The donor and recipient strains were grown overnight on BHI agar supplemented with appropriate antibiotics. The donor and recipient strains were grown overnight at $37^{o}C$ in BHI broth supplemented with appropriate antibiotic(s). The cultures were diluted to OD_{600} ~ 0.1 and were left to grow at $37^{o}C$ until mid-exponential phase (OD_{600} = 0.5-0.6). The cultures were then harvested by centrifugation at 3,000 x g (5,000 rpm) for 10 min and the supernatant discarded. The pelleted cells were then resuspended in 1 mL of fresh broth where both donor and recipient cells were mixed gently and spread on 0.45 μm pore size sterilized nitrocellulose filter (Sartorius, UK) which had previously been placed on antibiotic-free BHI agar. Plates were incubated overnight at $37^{\circ}C$, and the next day, the filter containing the biofilm (of the mixed cells) is vortexed vigorously in a 1 mL BHI broth. The resultant bacterial suspension was aliquoted onto and spread over agar plates containing the

appropriate antibiotics to select for transconjugants, donors and recipient cells.

2.9   Determination of Minimum Inhibitory Concentrations (MICs)

The MIC values of *B. subtilis* strains were determined using the broth microdilution method following the Clinical and Laboratory Standards Institute (CLSI) guideline. Briefly, antibiotics were prepared by serial two-fold dilutions in Cation-adjusted Mueller-Hinton Broth 2 (CA-MHB) (Sigma-Aldrich, UK) in different ranges of concentration depending on the particular antibiotic. These are done in sterile U-bottom Costar® 96-well Clear Polystyrene Microplates 3367 (Corning, US) in triplicates. The media were inoculated with 50 µL of diluted overnight culture to obtain approximately 1×10^6 cfu/well in a 100 µL total volume. The plates were incubated at 37°C for 18-24 hrs. The MIC was defined as the lowest concentration of antibiotic which gives a complete inhibition of visible growth in comparison with inoculated and uninoculated antibiotic-free wells

2.10 Sequencing reactions

PCR products and plasmids were sent to Genewiz Inc. (Genewiz, United Kingdom) for DNA sequencing following the standard requirement of sample preparation; minimum template concentration prepared was ~ 50 ng/µl in 10 µl for plasmids that are less than 6 kb, ~ 4 ng/µl in 10 µl for purified PCR products that are in a range of size of 2-4 kb plus. The appropriate primers are supplied at a concentration of 5 µM in 5 µl. DNA quantitation was determined using NanoDrop™ 1000 Spectrophotometer (Thermo Scientific,

Surrey, UK). *B. subtilis* wild type and resistant mutant strains were sent to MicrobesNG (http://www.microbesng.uk) for whole genome sequencing. The strains to be sequenced were prepared on agar according to the protocol provided and sent in barcoded bead tube supplied by MicrobesNG (Birmingham, UK).

2.11 *In silico* analysis

Acquired sequences were aligned, assembled and manipulated by using BioEdit software version 7.2.0 (Hall, 1999), SnapGene 3.2.1 (GSL Biotech LLC, US) and Clustal Omega (http://www.ebi.ac.uk/Tools/msa/clustalo). The sequences were analysed by comparing the DNA sequences and translated amino acid sequences to National Center for Biotechnology Information (NCBI) databases with the Basic Local Alignment Search Tool (BLAST) (http://blast.ncbi.nlm.nih.gov/) (Altschul *et al.*, 1990).

ARNold finding terminators (http://rssf.i2bc.paris-saclay.fr/toolbox/arnold/) and BPROM (Softbery, Inc., NY) program were used for the identification of putative rho-independent transcription terminator and promoter sites, respectively (Macke *et al.*, 2001, Naville *et al.*, 2011). Genomic sequences were analysed using *Breseq* (http://barricklab.org/breseq) (Deatherage & Barrick, 2014). The protein structure was generated using SWISS-MODEL (Biasini *et al.*, 2014). Statistical analyses were conducted using GraphPad Prism version 6.0 (GraphPad Software, San Diego, CA, USA.

3 Identification and Characterisation of Terminators Located Upstream of the Tn*916*, Tn*5397* and Tn*6000* Conjugation Module

3.1 Introduction

Tn*916* can be found integrated into the bacterial chromosome or excised as circularised intermediate molecule. Conjugation of these elements usually occurs when they are in a circular form. Therefore, the conjugation genes must be regulated upon the excision and circularisation of the element (Celli & Trieu-Cuot, 1998). If the conjugation genes were expressed when the element is integrated into the bacterial host chromosome, this may lead to an unregulated transfer which can be deleterious to the cell. However, there is also a possibility that a conjugative transposon does not excise upon the expression of the conjugation genes, theoretically resulting in the co-transfer of the whole chromosomal DNA unless disruption of the mating pair or nicking of the incoming DNA strand occurs. This high-frequency recombination-like (Hfr-like) transfer mediated by conjugative transposon has been demonstrated in *Vibrio cholerae* (Hochhut *et al.*, 2000) and *Bacteroides* sp (Whittle *et al.*, 2006) where the conjugative transposons (SXT and CTnERL, respectively), are not excised and co-mobilised with the chromosomal DNA. From filter mating experiments in between the toxigenic strain *C. difficile* 630Δerm and non-toxigenic strain *C. difficile* CD37; co-transfer of Tn*5397* and variable-sized chromosome fragments containing the pathogenicity locus (PaLoc) has been demonstrated (Brouwer *et al.*, 2013). Nevertheless, as the conjugation of the Tn*916* element usually occurs upon excision and circularisation of the element, we hypothesize that the presence of terminator located upstream of the conjugation module is needed to prevent the transcription of their conjugation genes whilst they are integrated into the genome.

The Tn*916* and Tn*916*-like elements conjugative transposons can insert and are found in multiple sites within the genome. Tn*916* particularly has been shown to enter the genome of *C. diffiicile* at multiple AT-rich sites (Scott *et al.*, 1994, Roberts & Mullany, 2009, Mullany *et al.*, 2012). Furthermore, Mullany *et al.* (2012) have demonstrated in *C. difficile* strains 630 and R20291, Tn*916* preferentially integrates into the genome at intergenic regions, which may contain transcriptional regulatory elements (Mullany *et al.*, 2012).The fact that the conjugative transposons found in multiple sites within the genome (including the intergenic region that typically contain promoters), suggests that they are able to insert and be maintained in regions with differing transcriptional activity. Therefore, this demands a control mechanism to prevent variable transcription reaching the conjugation genes, and this control mechanism could be a rho-independent terminator. Initially in this study, we aimed to search for intrinsic rho-independent terminators located upstream of the conjugation module of Tn*916* and Tn*916*-like elements as stabilisation in the genome demands that there is one there.

The rho-independent terminator is defined as a palindromic sequence that can form a hairpin or stem loop structure followed by a stretch of thymine (T) residues (Figure 3-1) (Lesnik *et al.*, 2001, Naville *et al.*, 2011). These terminators are commonly found at the end of a transcript but can also play a role as transcriptional attenuators when located between the genes of different transcriptional units (Macke *et al.*, 2001).
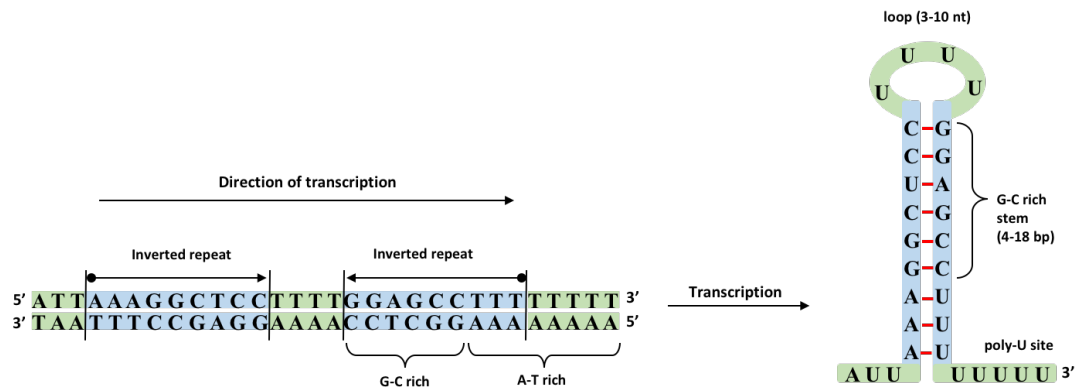
Figure 3-1 Rho-independent intrinsic terminator.

Basic characteristics include (5′ to 3′); a hairpin with a loop of 3–10 residues and stems of 4–18 base pairs, with or without bulges (usually the G-C rich regions) and poly-U site.

In this chapter, we identified a previously unknown group of terminators, structurally conserved across multiple elements upstream of the conjugation module. The termination efficiency of the putative terminators in Tn*916* and Tn*916*-like elements were evaluated based on a published algorithm (d'Aubenton Carafa *et al.*, 1990). Additionally, the termination efficiency activity (of Tn*916,* Tn*5397 and* Tn*6000* terminators) were investigated by using an *in vitro* reporter system. Further investigation was carried out on the Tn*916* terminator which showed the highest efficiency value.

3.2 Materials and methods

3.2.1 Prediction of putative terminator sequence and estimation of the termination efficiency

The putative rho-independent terminators were initially scanned by eye at the predicted region; upstream the conjugation module of Tn*916*/Tn*1545* family of conjugative transposons. Subsequently, the RNAMotif algorithm via ARNold program (Macke *et al.*, 2001) was used to validate the presence of the intrinsic rho-independent terminators in various conjugative transposons. The free energies (ΔG) of hairpin formation was predicted using Mfold program (Zuker, 2003). The termination efficiency of Tn*916*, Tn*6000* and Tn*5397* putative rho-independent terminators were evaluated based on an algorithm described previously by d'Aubenton Carafa *et al.* (1990). This algorithm was constructed based on two parameters; n_T and Y value.

Parameter 1; n_T

The n_T value is the number of thymine (T) residues with a weight decreasing in the 5' to 3' direction. The calculation used to calculate n_T is as follows:

$$\chi_n = \chi_{n-1} \times 0.9 \text{ if the } n^{th} \text{ nucleotide is a thymine (T)}$$

$$\chi_n = \chi_{n-1} \times 0.6 \text{ if the } n^{th} \text{ nucleotide is other than thymine (T)}$$

The value for the first T is: $\chi_1 = 0.9$, and χ_n is calculated as the sum of the T residues only, therefore;

$n_T = \Sigma \chi n$ for all T residues in the terminator.

For example, in hexanucleotide TTATTT, the n_T is calculated as follows:

T:	0.9
T:	0.9 x 0.9 = 0.81
A:	0.81 x 0.6 = 0.486
T:	0.486 x 0.9 = 0.437
T:	0.437 x 0.9 = 0.394
T:	0.394 x 0.9 = 0.354

Therefore, n_T = 0.9 + 0.81 + 0.437 + 0.394 + 0.354 = 2.895 (this stretch of T residues is considered as the worst possible for a terminator). Therefore, the predicted terminator must satisfy this minimal value of 2.895 to be considered as a real terminator and values less than 2.895 are therefore rejected (d'Aubenton Carafa *et al.*, 1990).

Parameter 2; *Y* value

The *Y* value is the function of the Gibbs free energy (ΔG) against the number of nucleotides between the 5' end of the stem and the first U in the stretch (LH). *Y* value is calculated as follows:

$Y = (-\Delta G)/L_H$

Based on these two parameters, d' Aubenton Carafa *et al.* (1990) plotted a two-dimensional diagram to separate the terminators from intracistronic structures. Line D was drawn to obtain the best separation of these two structures (Figure 3-2). All the structures were plotted as a point with *Y* [(-ΔG)/L_H] value on the x-coordinate and n_T on the y-coordinate. The final score;

d, represents the distance of the representative point to line D (indicated in the diagram). Based on the computational analysis, the following equation was derived;

Equation; *d* score

d = n_T x 18.16 + Y x 96.59 – 116.87, where the condition of *d* > 0 is applied to all terminator structures.
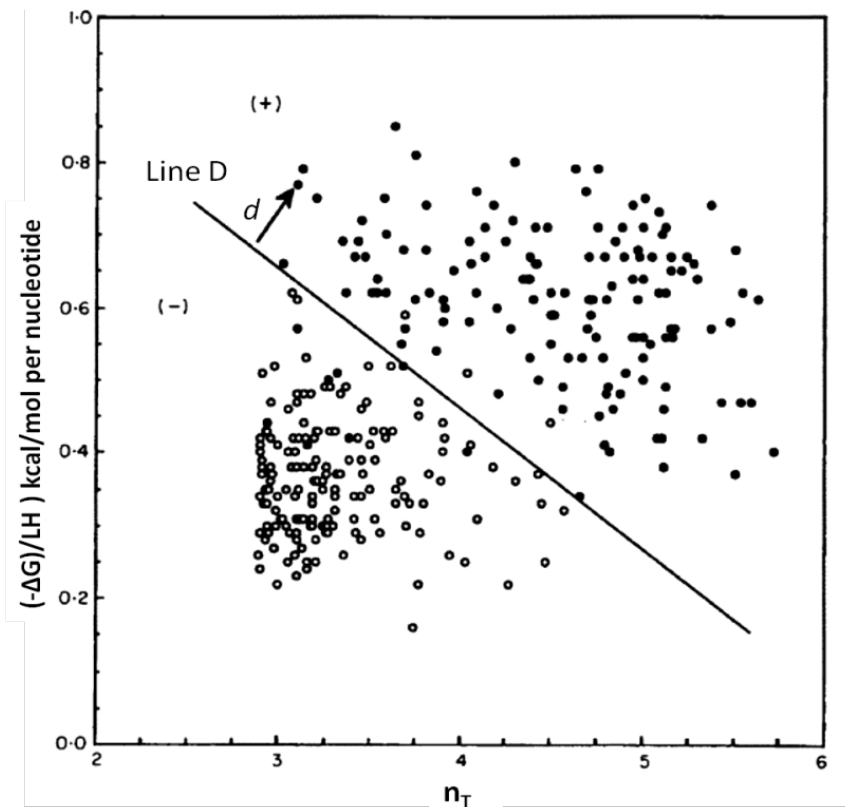
Figure 3-2 Two-dimensional diagram showing the separation of terminators from the intracistronic or random structures in *E. coli.*

The symbol (●) represents the real transcriptional terminators whereas (○) represents either the intracistronic or random structures. Figure is reproduced with permission (d'Aubenton Carafa *et al.,* 1990)**.**

To estimate the termination efficiency of the terminators, d' Aubenton Carafa *et al.* (1990), plotted the value of *d* against the *in vitro* termination efficiency (%) derived from of a set of *E. coli* rho-independent terminators. A curve is drawn to show the correlation between these two values (Figure 3-3) (d'Aubenton Carafa *et al.*, 1990).
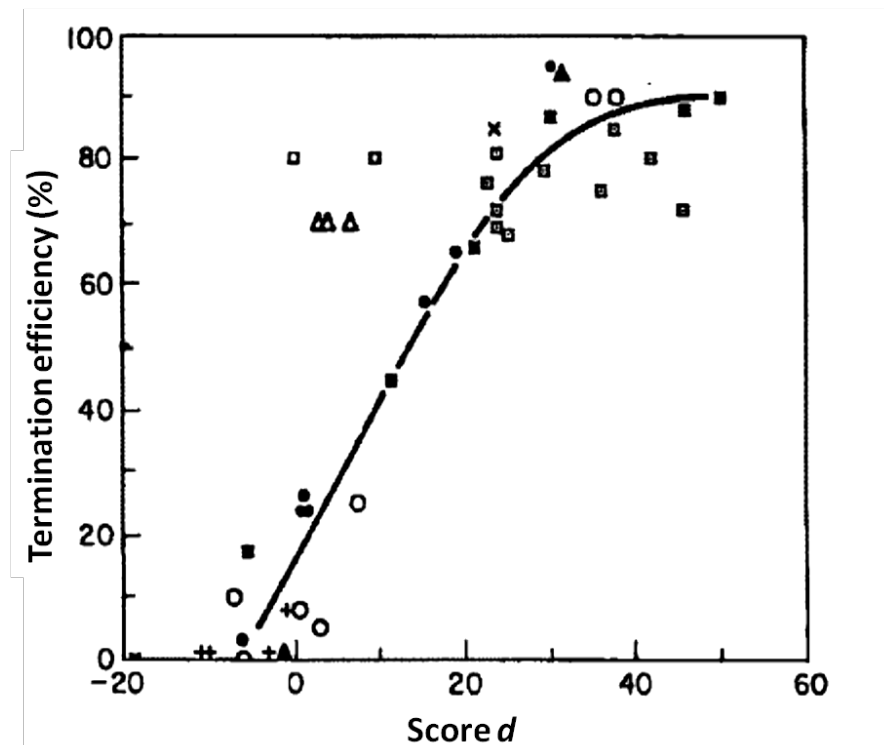
Figure 3-3 The correlation between the *d* score of some rho-independent terminators in *E. coli* and their efficiency *in vitro*.

The terminators are indicated based on the preceding gene or operon: (□) *rrnB T1*, bacteriophage T7 *Te*; (▲) *ampL* attenuator and *ampL35A* mutant; (○) *infC*, *pheS* attenuator, *his* attenuator, *trpt* and *trpC301* and *trpC302* mutants, bacteriophage T3 *Te*; (Δ) *tonB* (both directions), *rplT*; (●) *trp* attenuator, *trp a1419* and *trp a135* mutants, *trpL77*, *trpL78*, *trpL80*, *trpL153* mutants; (■) thr attenuator and *T2, T3, T4, T5, T6, T8* mutants in the poly (U) stretch; ( ) *thr* attenuator stem mutants *L135U, L138U, L139U, L140A, L151A, L151U, L153A, L153U, L153 + G, L153 −G, L156U*; (X) rnpB; (+) intracistronic signals in *cca*. Figure is reproduced with permission (d'Aubenton Carafa *et al.,* 1990).

3.2.2 Generation of Tn*916*, Tn*6000* and Tn*5397* terminator constructs

Terminator constructs of Tn*916*, Tn*6000* and Tn*5397* were generated using Q5® Site-Directed Mutagenesis Kit from New England Biolabs (NEB) Ltd (Hertfordshire, UK) as described in section 2.7. The terminators were cloned between P*tet*(M) and the reporter gene, *gusA* encoding for β-glucuronidase (Figure 3-4). Amplification was done by using either For Tn*916*-T*Spe*I, For Tn*5397*-T*Spe*I or For Tn*6000*-T*Spe*I forward primer, paired with the reverse primer; Rev*916* PO (Table 2-3). The pHCMCO5-P*tet*(M)-PO construct containing P*tet*(M) and *gusA* was used as a template (Table 2-2). Terminator sequence was incorporated into the 5' end of the forward primer while the reverse primer anneals back to back with the 5' end of the complementary region of the forward primer (Figure 3-4). *E. coli* chemically-competent cells were used for the transformation of terminator constructs (Section 2.7.3). Positive transformants were subjected to plasmid purification using the QIAprep Spin Miniprep Kit (Qiagen, UK) as described in section 2.4.2 and sent for sequencing for verification. Primer pair of pHCMC05_P*tet*M_GusA_F/R, GusF1 and GusF2 targeting the insert region was used for screening and sequencing (Table 2-3). The isolated plasmids containing the terminator were subsequently transformed into *B. subtilis* BS34A (Section 2.4.10) for the enzyme assay purpose.
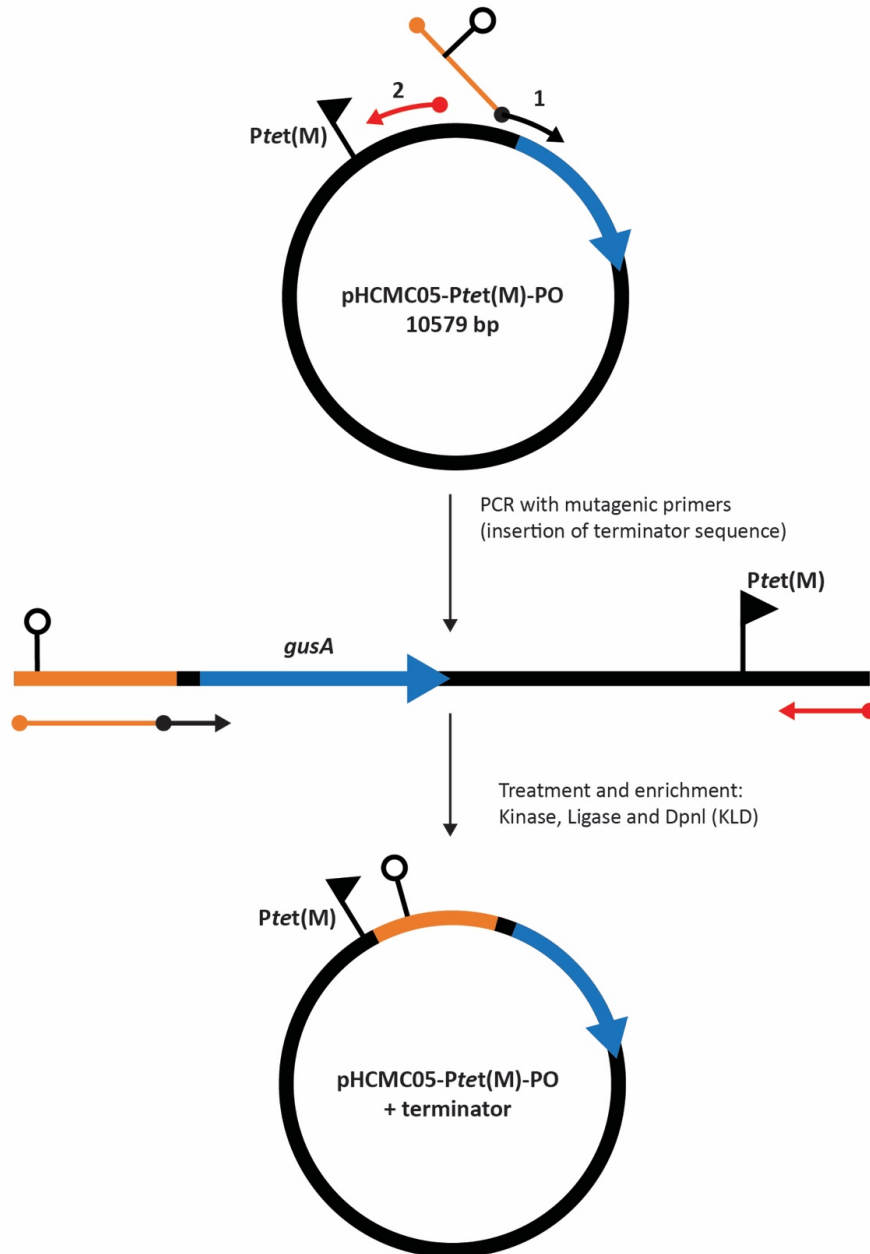
Figure 3-4 Schematic diagram of generation of the terminator reporter construct via site directed mutagenesis.

The pointed 'P' indicates the *tet*(M) promoter. The orange box represents either the terminator sequence of Tn*916*, Tn*6000* or Tn*5397*, blue arrow box represents the *gusA* gene, black box represents the plasmid backbone, arrows represent the annealing position of the primers and the direction of priming. The pHCMCO5-P*tet*(M)-PO (Jasni, 2013) was used as template. All PCR products were ligated and transformed into *E. coli* and subsequently into *B. subtilis* BS34A.

3.2.3 Generation of construct A (Tn*916* left end-BS34A genome junction) and construct B (Tn*916* joint ends region)

The terminator region is cloned in between the *tet*(M) promoter and a *gusA* reporter gene in a pHCMC05 shuttle vector, plus flanking chromosomal DNA of either the Tn*916* left end-BS34A genome junction (region A: representing the linear, integrated form) or the joint ends of Tn*916* (region B: representing the excised and circularised form) (Figure 3-5). Fragment A was amplified by using REGJ-*Spe*I (primer 3) and Rev Tn*916*-LE-*Age*I (primer 4) that were designed with added *Age*I and *Spe*I restriction site. Fragment B was amplified by using For Tn*916*-RE-*Spe*I (primer 5) and Rev Tn*916*-LE-*Age*I (primer 4) with added *Age*I and *Spe*I restriction site. This primer pair is able to amplify the joint ends region of Tn*916* if this element excises to form a circular intermediate. *B. subtilis* BS34A genomic DNA was used as a template for the amplification of both fragment A and B (Figure 3-5). All primer sequences are listed in Table 2-3 in Chapter 2.

The respective fragments were inserted into pHCMCO5-P*tet*(M)-PO construct (Jasni, 2013) via directional cloning. In order to do this, site directed mutagenesis on pHCMC05-P*tet*(M)-PO construct was done to insert the *Age*I and *Spe*I cutting sites in between the P*tet*(M) and the reporter gene*, gus*A using a primer pair of For*916* PO-*Spe*I-*Age*I and Rev*916* PO (Figure 3-6) (Table 2-3).

For the generation of the terminator constructs A and B, both fragments (A and B) and pHCMC05-P*tet*(M)-PO-*Age*I-*Spe*I vector were double digested

with *Age*I and *Spe*I and subsequently purified by using QIAquick PCR Purification Kit (Qiagen, UK). The respective purified fragments were then ligated into the digested pHCMC05-P*tet*(M)-PO-*Age*I-*Spe*I and transformed into *E. coli* (Figure 3-6). Transformants were screened by PCR and digested with *Age*I, *Spe*I and *BamH*I. Positive transformants were subjected to plasmid purification using the QIAprep Spin Miniprep Kit (Qiagen, UK) as described in section 2.4.2 and sent for sequencing for verification. Primer pair of pHCMC05_P*tet*M_GusA_F/R, GusF1 and GusF2 targeting the insert region was used for positive plasmid screening and sequencing (Table 2-3).

The isolated constructs were subsequently transformed into *B. subtilis* BS34A for the enzyme assay purpose. The BS34A strain carries a single copy of Tn*916* (Roberts *et al.*, 2003). This host is chosen because the Tn*916* is demonstrated to be stable in this site (Roberts *et al.*, 2003) and its complete genome sequence has been obtained (Browne *et al.*, 2015).
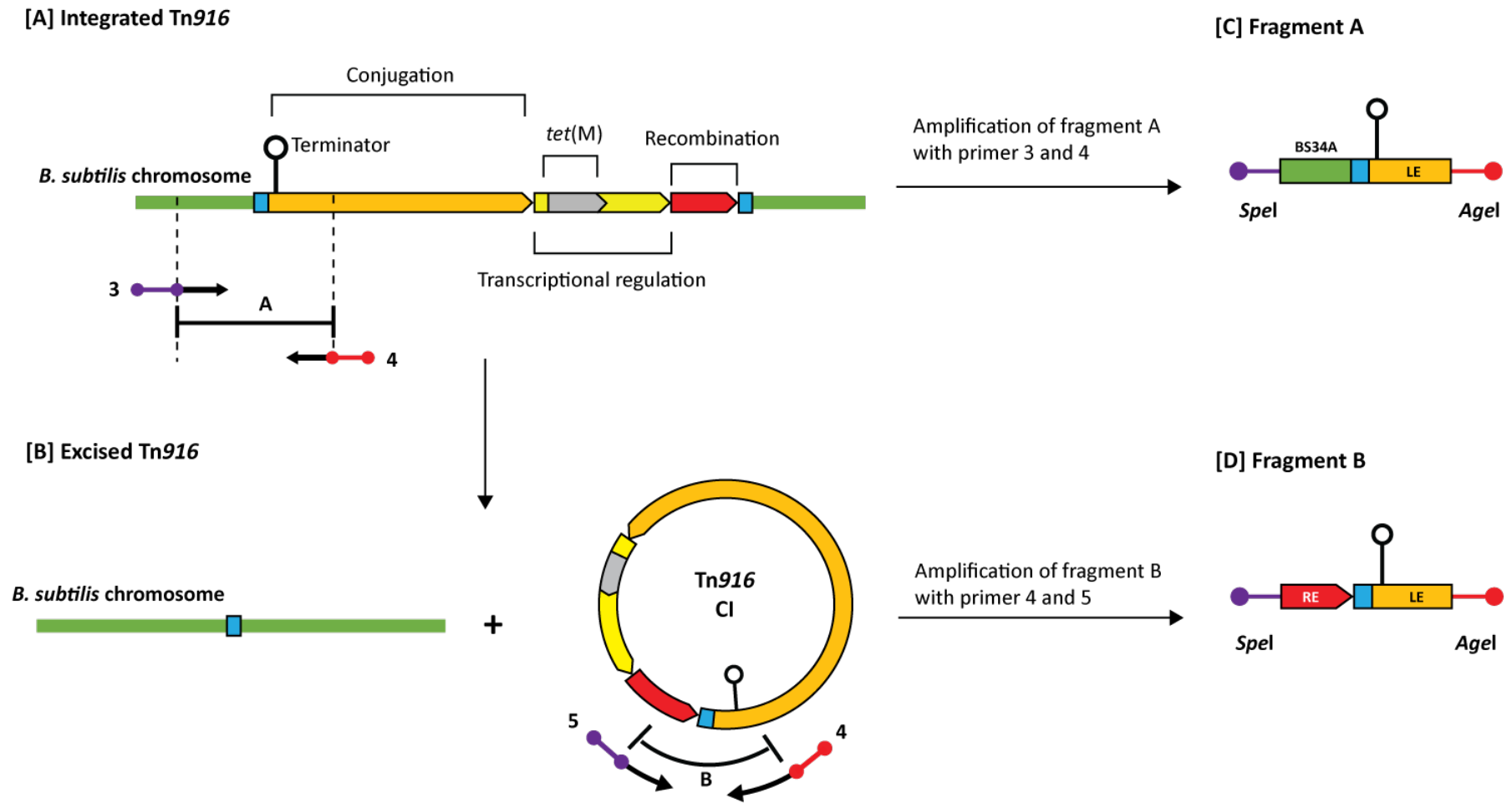
Figure 3-5 Schematic diagram of the integrated and excised Tn*916* conjugative transposon.

[A] Tn*916* integrated in the bacterial genome. Four functional modules of Tn*916* are shown: conjugation (orange); recombination (red); transcriptional regulation (yellow) and the accessory gene *tet*(M) (grey). Blue box represents the coupling sequence and green box represents the *B. subtilis* chromosome. [B] Tn*916* in an circular intermediate (CI) form and bacterial genome with excised Tn*916*, [C] Fragment A amplified with primer 3 (REGJ-*Spe*1) and 4 (Rev Tn*916*-LE-*Age*I) representing the integrated, linear form of Tn*916*, [D] Fragment B amplified with primer 4 (Rev Tn*916*-LE-*Age*I) and 5 (For Tn*916*-RE-*Age*I) representing the circularised form of Tn9*16*. CI; circular intermediate, RE; right end of Tn*916*, LE; left end of Tn*916*, BS34A; *B. subtilis* BS34A chromosomal fragment.
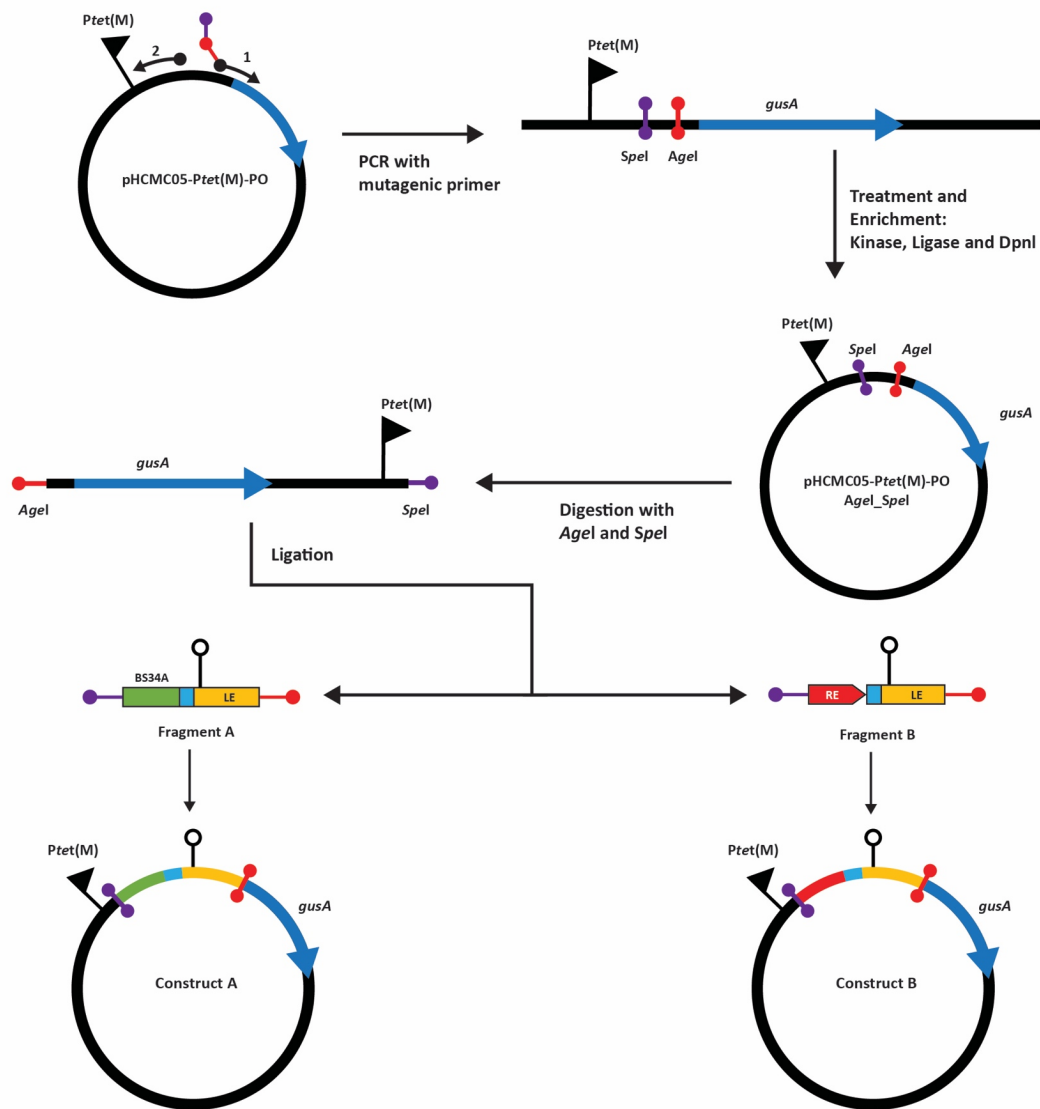
Figure 3-6 Generation of A and B constructs.

*Age*I and *Spe*I restriction sites were added to pHCMC05-P*tet*(M)-PO construct by site directed mutagenesis. The prepared plasmid was then digested with *Age*I and *Spe*I. Fragment A and B that were treated with the same restriction enzymes was cloned into pHCMC05-P*tet*(M)-PO plasmid. The generated construct A and B was then transformed into *E. coli* and subsequently into *B. subtilis* BS34A.

3.2.4 Generation of ΔSubA and ΔSubB constructs

Another two constructs similar to A and B with the substitution on the poly-T tail (ΔSubA & ΔSubB) of the terminator sequence were generated via site directed mutagenesis using construct A and B as the template (Figure 3-7). The poly-A tail of the Tn*916* terminator in ΔSubA and ΔSubB is substituted to 9 G(s) [GGGGGGGGGG] by using a primer pair; *916*Sub_F and *916*Sub_R (Table 2-3) and transformed into *E. coli.* Positive transformants were subjected to plasmid purification using the QIAprep Spin Miniprep Kit (Qiagen, UK) as described in section 2.4.2 and sequenced for verification. Primer pair of pHCMC05_PtetM_GusA_F/R, GusF1 and GusF2 targeting the insert region was used for positive plasmid screening and sequencing. The isolated ΔSubA and ΔSubB constructs were subsequently transformed into *B. subtilis* BS34A for the enzymatic reporter assay.
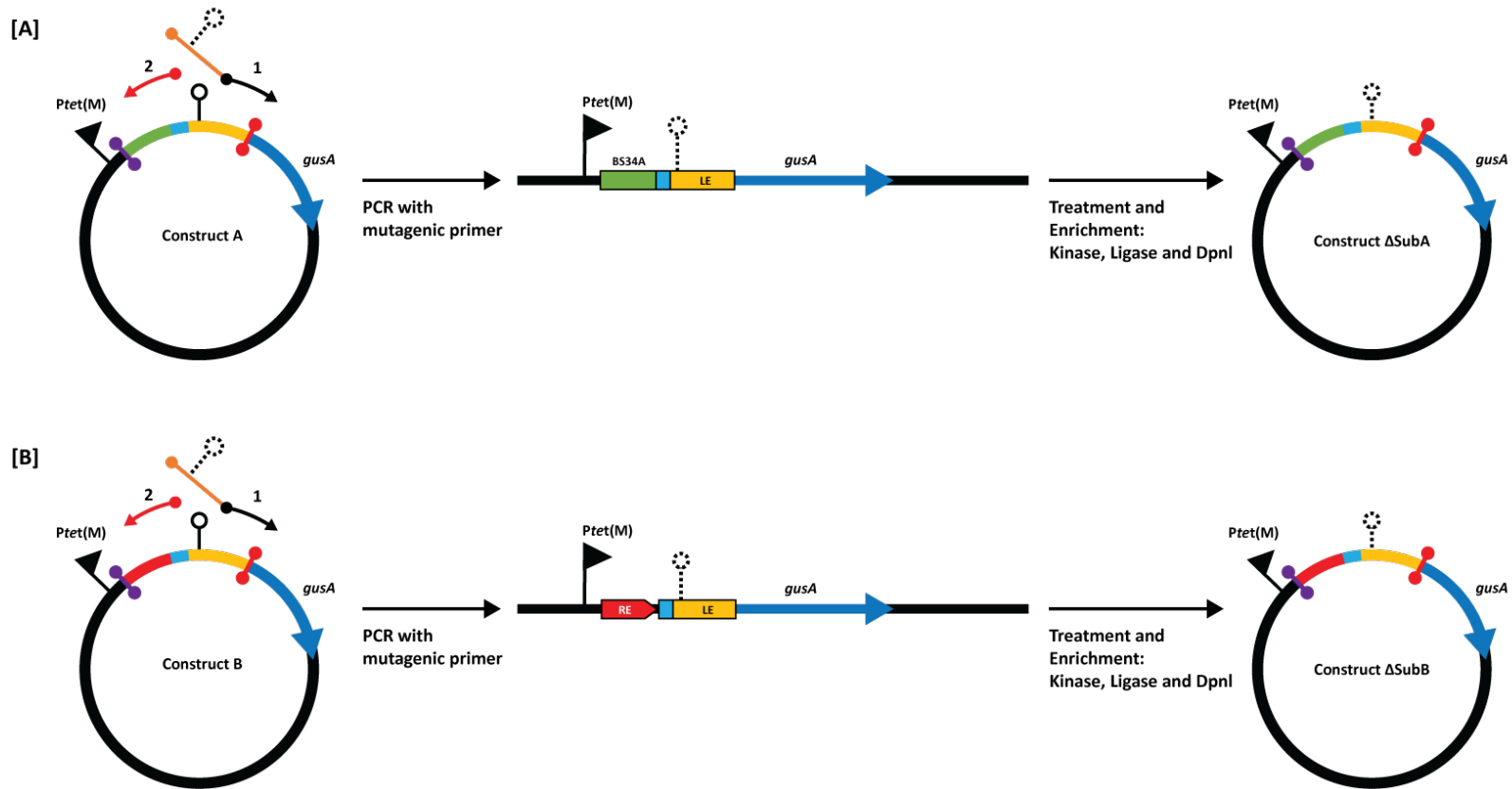
Figure 3-7 Generation of ΔSubA and ΔSubB constructs.

[A] Construction of ΔSubA via SDM using construct A as the template. [B] Construction of ΔSubB via SDM using construct B as the template. The green box represents the BS34A chromosome fragment, red box represents the right end of Tn*916* (RE), orange box represents left end of

Tn*916* (LE), blue box represents the *gusA* gene, light blue box represents the coupling sequence and black box represents the plasmid backbone.

The structure (⚲) represents the terminator, while this (⚬) structure represents the mutant terminator.

3.2.5 Spectrophotometric measurement of *gus*A expression in cell lysates

The β-glucuronidase activity was measured based on the method developed for *B. subtilis* (Belitsky *et al.*, 1995) with slight modifications. Each strain containing the terminator constructs to be analysed was inoculated into BHI broth supplemented with chloramphenicol (10 µg/ml), while the plasmid-less *B. subtilis* BS34A was inoculated into BHI broth supplemented with tetracycline (10 µg/ml). All strains were incubated overnight at 37°C, 200 rpm. The optical density of the overnight culture was measured at 600 nm and 0.5 ml culture was then harvested by centrifugation (3000 x *g*, 25°C, 10 min). The cell pellets were kept in the freezer (-80°C) for 1 h. The pellet was then thawed at room temperature and resuspended in 0.8 ml of Z buffer and 0.8 µl of toluene. The mixture was transferred to sterilised 1.5 ml tube containing unwashed glass beads (150-212 µm in diameter) and treated in a Ribolyser at setting 6.5 for 25 sec or vortexer at setting 7 for 5 mins bursts to lyse the cells. Treatment was repeated twice. The lysates were placed on ice for 1 min and to remove the glass beads, it was subjected to centrifugation at 3000 x *g*, 4°C for 3 minutes. Then, 30 uL of the supernatant was carefully transferred to a new microcentrifuge tube containing 0.77 ml Z buffer and incubated at 37°C for 5 min. The enzyme reaction was initiated by adding 0.16 ml of 6 mM ρ-nitrophenyl-β-D-glucuronide and incubated for 5 min at 37°C. The enzyme reaction was terminated by adding 0.4 ml of 1 M Na_2CO_3. Next, the lysates were centrifuged (3000 x *g*, 25°C, 10 min) to remove the cell debris. The optical density was measured at 405 nm (OD_{405}) using the spectrophotometer.

The β-glucuronidase units were calculated using the following equation: $(OD_{405} \times 1000) / [OD_{600} \times time\ (min) \times 1.25 \times volume\ (ml)]$ (Miller, 1972).

3.3 Results

3.3.1 *In silico* analysis of putative terminator in Tn*916* and Tn*916*-like elements

We hypothesise that a terminator is needed to prevent the transcriptional activity of the Tn*916* and Tn*916*-like conjugation genes when inserted at variable sites in variable genomes. This is important for these elements to maintain their stability in the bacterial genome. Therefore, rho-independent terminators were initially scanned by eye at the predicted region; upstream of the conjugation module of Tn*916*/Tn*1545* family of conjugative transposons. Upon identification of the intrinsic rho-independent terminator sequence in various conjugative transposons, the ARNold and Mfold program were used to further analyse the secondary structures. Seven putative terminators were found at the upstream region of the conjugation module of Tn*2010*, Tn*5397*, Tn*6000*, Tn*6002*, Tn*6003*, Tn*6087* and Tn*916,* respectively (Table 3-1). All predicted terminators match a descriptor of rho-independent terminators that constitute a hairpin with a loop of 3-10 residues, stems with or without bulges and thymine-rich region (Table 3-2). Five out of seven predicted terminators (from Tn*2010*, Tn*6002*, Tn*6003*, Tn*6087* and Tn*916*) share the same sequence as shown in multiple sequence alignment below (Figure 3-8). This is because these transposons are similar in DNA sequences starting from *orf*20 to *orf*24 of Tn*916*. However, the Tn*5397* and Tn*6000* terminators varied in terms of size of the loop, G-C content in the stem and the numbers of thymine in comparison to the other five terminators (Figure 3-8).

Table 3-1 Predicted rho-independent terminators via ARNold program.

Predicted transcriptional terminators derived from upstream sequence of Tn*2010*, Tn*5397*, Tn*6000*, Tn*6002*, Tn*6003*, Tn*6087* and Tn*916* conjugation module. Each predicted terminator contains starting position; strand direction; color-coded terminator sequence (blue: stem, red: loop, black, bold and underlined: T-stretch) and predicted free energy of terminator hairpin* (Kcal/mol).

Tn*916* and Tn*916*-like elements	Predicted terminators	Accession no.	Characteristics	Size (kb)	Original host	Reference(s)
Tn*2010*	70 Rnamotif + TTTAATAGGTAGACACTTC AAAAAATGAGGTGTCTA**TTTTTTT**ACC - 12.00*	AB426620.1	Confers tetracycline [*tet*(M)], macrolides [*mef*(E)] and MLS [*erm*(B)] resistance	26.3	*Streptococcus pneumoniae*	(Del Grosso *et al.*, 2009)
Tn*5397* **	88 Rnamotif + CATATAAAGAGCCATTTGA TTTTTCATATCAAGTGG**TTTTT**GTTATGT - 9.40*	AF333235.1	Confers tetracycline [*tet*(M)] resistance	20.6	*Clostridium difficile*	(Mullany *et al.*, 1996, Roberts *et al.*, 2001)
Tn*6000* **	437 Rnamotif + AAAAGGTTTTAGACACTT CCAAAGTTGAGGTGTC**TTTTTT**GAATAA - 10.10*	FN555436.1|	Confers tetracycline [*tet*(S)] resistance	33.3	*Enterococcus casseliflavus*	(Roberts *et al.*, 2006, Brouwer *et al.*, 2010)
Tn*6002*	170 Rnamotif + TTTAATAGGTAGACACTT CAAAAAATGAGGTGTC**TATTTTTTT**ACC - 12.00*	AY898750.1	Confers tetracycline [*tet*(M)] and MLS [*erm*(B)] resistance	20.8	*Streptococcus cristatus*	(Warburton *et al.*, 2007)
Tn*6003*	170 Rnamotif + GGCAAAAGGTGGACACTT CAAAAAATGAGGTGTC**TATTTTTTT**ACC - 12.00 *	AM410044.5	Confers tetracycline [*tet*(M)], MLS [*erm*(B)], and kanamycin [*aphA*-3] resistance	25.1	*Streptococcus pneumoniae*	(Cochetti *et al.*, 2008)

Tn6087	170 Rnamotif + TTTAATAGGTAGACACTTC AAAAAATGAGGTGTCTATTTTTTTACC - 12.00 *	HQ663849.2	Confers tetracycline [tet(M)], and CTAB [qrg] resistance	21.2	Streptococcus oralis	(Ciric et al., 2011)
Tn916 **	170 Rnamotif + TTTAATAGGTAGACACTTC AAAAAATGAGGTGTCTATTTTTTTACC - 12.00 *	U09422.1	Confers tetracycline [tet(M)] resistance	18.3	Enterococcus faecalis	(Franke & Clewell, 1981)

* Predicted free energy of terminator hairpin (Kcal/mol).

** Conjugative transposons chosen for the β-glucuronidase enzyme assay are highlighted in yellow.

[A]

```
Tn2010    GACACTTCAAAAAAATGAGGTGTCTATTTTTTT    32
Tn6002    GACACTTCAAAAAAATGAGGTGTCTATTTTTTT    32
Tn6003    GACACTTCAAAAAAATGAGGTGTCTATTTTTTT    32
Tn6087    GACACTTCAAAAAATGAGGTGTCTATTTTTTT     32
Tn916     GACACTTCAAAAAATGAGGTGTCTATTTTTTT     32
Tn6000    GACACTTCCAAAGTTGAGGTGTCTTTTTT---     29
          ******** ***   ********** ****
```

[B]

```
Consensus    GACACTTCAA----AAAATGAGGTGTCTATTTTTTT    32
Tn5397       CC-ATTTGATTTTTCATATCAAGTGGT--TTTTGTT    33
             *  ** *       * ** * ***    **** **
```

Figure 3-8 Multiple sequence alignment of putative terminators from Tn*916*/Tn*1545* family of conjugative transposons.
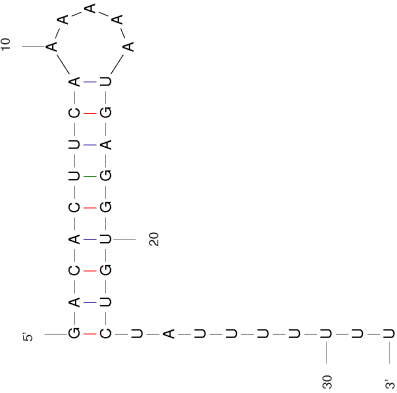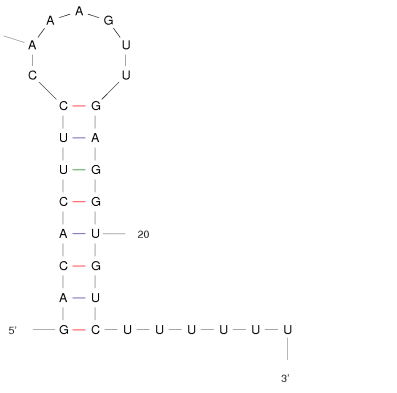
Panel A: Multiple sequence alignment of putative terminators from Tn*2010*, Tn*6002*, Tn*6003*, Tn*6087*, Tn*916*, and Tn*6000*. **Panel B:** Pairwise sequence alignment of putative terminator from Tn*5397* and consensus sequence derived from initial alignments of terminators in panel A. The terminator sequences are color-coded as follows: stem in blue, loop in red and the thymidine stretch in black, bold and underline. An * (asterisk) indicates positions which have a single, fully conserved residue.
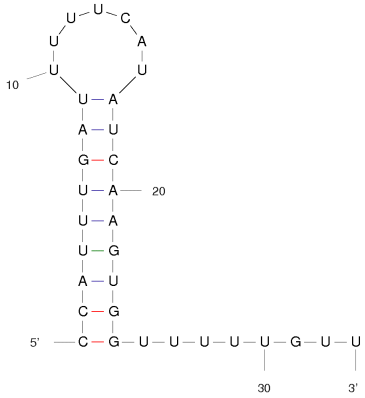
3.3.2 Prediction of the termination efficiency and secondary structure

Based on the multiple sequence alignment of the predicted terminators, Tn*2010*, Tn*6002*, Tn*6003*, Tn*6087* and Tn*916* are shown to have the same sequence (GACACTTCAAAAAATGAGGTGTC**TATTTTTTT)** and predicted free energy (Figure 3-8). Therefore, Tn*916* was chosen as a representative from this group for further analysis, assuming that the result will be the same for all five terminators. Variation in terms of size of the loop, G-C content in the stem, numbers of thymidine stretch and -ΔG value however were observed among Tn*916,* Tn*6000* and Tn*5397* and therefore, these three putative rho-independent terminators were chosen for further analysis. The predicted secondary structure of Tn*916*, Tn*6000* and Tn*5397* terminators, the value of $n_T,$ - ΔG, $L_H,$ *Y, d* and their respective termination efficiency are shown in Table 3-2 below.

Using the defined parameters, the *d* values were plotted on the diagram in Figure 3-9, which shows the correlation between the score *d* of some rho-independent terminators and their efficiency *in vitro* (d'Aubenton Carafa *et al.,* 1990) and the termination efficiency were estimated to be approximately 21% in Tn*6000* and 31% in Tn*916*. The efficiency score of Tn*5397* gave a negative *d* value although having the common structural characteristics similar to the reported terminators. To further analyse this discrepancy, terminator constructs were generated by PCR mediated site directed mutagenesis followed by enzymatic assay using the *gusA* reporter gene system to investigate the termination activity of these putative structures.

Table 3-2 Termination efficiency and secondary structure of Tn*916*, Tn*6000* and Tn*5397* putative rho-independent terminators.

Terminator	Sequences	Predicted structure by Mfold	n_T	- ΔG (kcal/mol)	L_H	Y	d	Estimation of termination efficiency
Tn*916*	+GACACTTCAAAAAATG AGGTGTC**TATTTTTTT**		3.974	- 12.00	23	0.52	5.53	31 %
Tn*6000*	+ GACACTTCCAAAGTTG AGGTGTC**TTTTTT**		4.216	- 10.10	23	0.439	2.09	21 %

| Tn*5397* | + CCATTTGATTTTTCAT ATCAAGTGG**TTTTTGTT** | | 4.004 | - 9.40 | 25 | 0.376 | - 7.84 | - |

n_L, number of bases in the loop

L_H, number of nucleotides between the 5' end of the stem and the first U in the stretch

$Y = (-\Delta G)/LH$

$d = n_T \times 18.16 + Y \times 96.59 - 116.87$

118
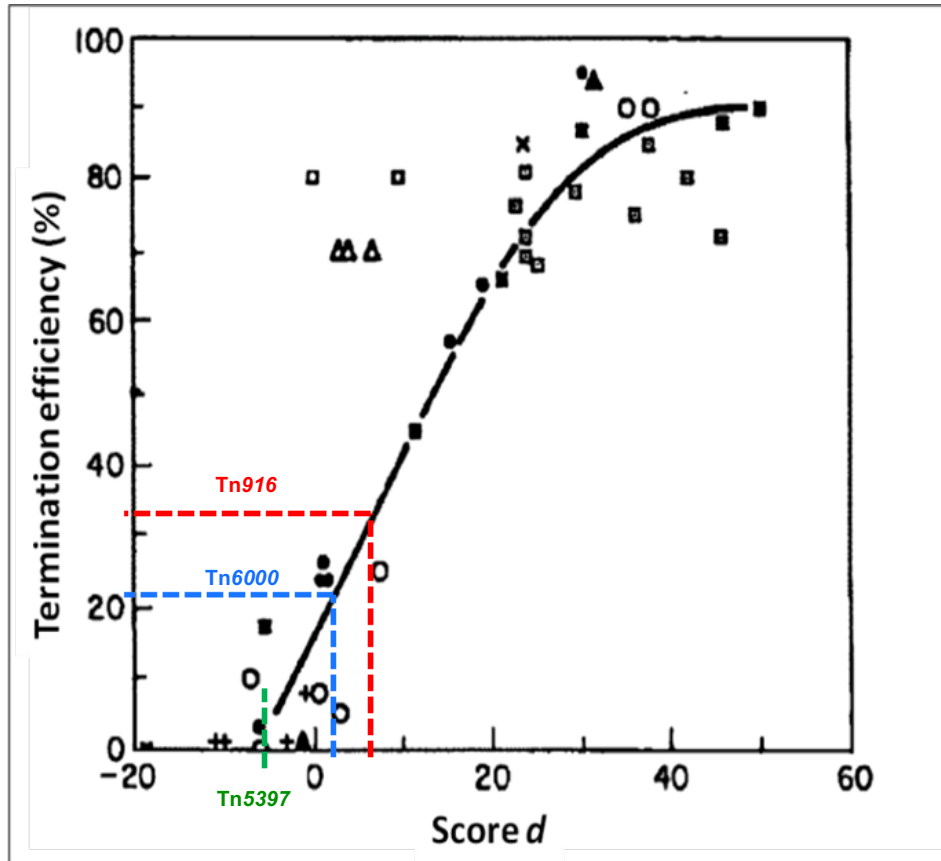
Figure 3-9 Correlation between the score _d_ of the putative rho-independent terminators and their efficiency _in vitro_ (d' Aubenton Carafa et al., 1990).

The _d_ values of Tn_916_ (5.53) and Tn_6000_ (2.09) putative terminators were plotted on the x-axis and correlated with 31% (red dotted line) and 21% (blue dotted line) of termination efficiencies, respectively. The negative _d_ value of Tn_5397_ (-7.84) putative terminator was plotted only on the x-axis, represented by the green dotted line.

3.3.3 Tn*916,* Tn*6000* and Tn*5397* terminator constructs

The plasmid-based terminator constructs of Tn*916*, Tn*6000* and Tn*5397* were generated and verified by DNA sequencing. The transformants harbouring respective terminator constructs were screened by amplifying the mutagenesis region of the terminator constructs (Figure 3-10), followed by sequencing analysis. Each of the three constructs contain a fusion of a *tet*(M) promoter as indicated by -35 and -10 (Su *et al.*, 1992), *gusA* and the terminator sequence respectively as shown in Figure 3-11. The schematic diagram of the transcriptional terminator constructs is shown in Figure 3-12.

[A]   [B]
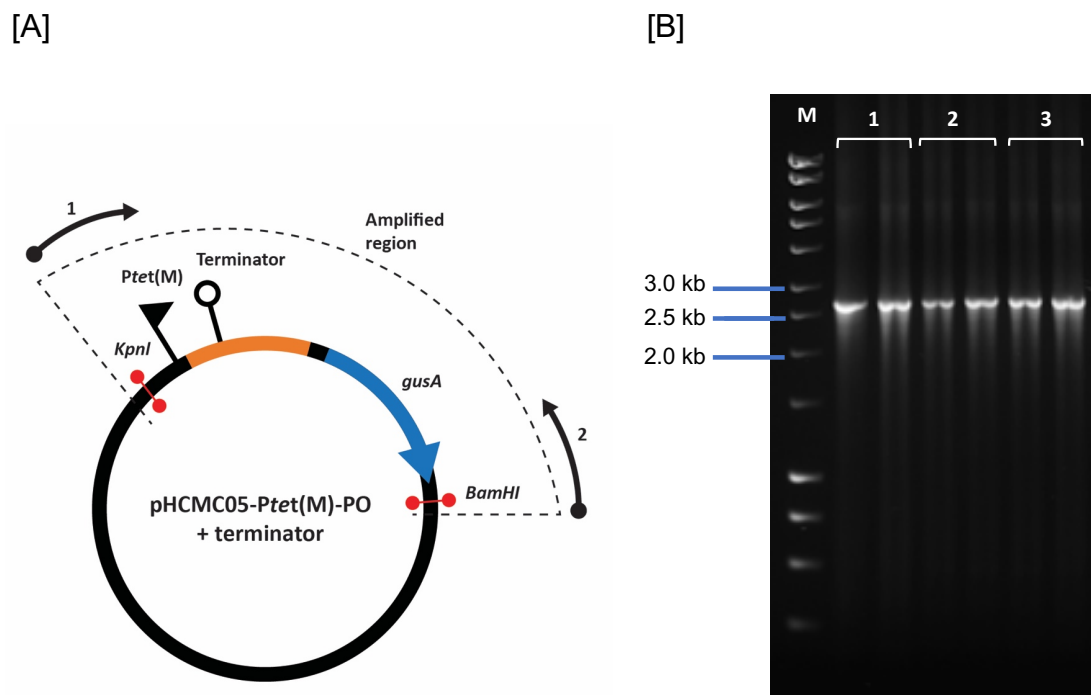
Figure 3-10 Schematic representation of the amplified region of terminator constructs.

Panel A: Schematic diagram of the amplified region using primer pair; 1 and 2 (pHCMC05_P*tet*M_GusA_F/R). **Panel B:** agarose gel electrophoresis of the amplicons. Lane M; HyperLadder™ 1kb, [1]; Tn*916* terminator construct amplicon, [2]; Tn*6000* terminator construct amplicon and [3]; Tn*5397* terminator construct amplicon.

```
                                                    - 35                      - 10
Tn916.T    TCTCTTTGATAAAAAAATTGGAGATTCCTTTACAAATATGCTCTTACGTGCTATTATTTAA    359
Tn6000.T   TCTCTTTGATAAAAAAATTGGAGATTCCTTTACAAATATGCTCTTACGTGCTATTATTTAA    356
Tn5397.T   TCTCTTTGATAAAAAAATTGGAGATTCCTTTACAAATATGCTCTTACGTGCTATTATTTAA    353
           ************************************************************


Tn916.T    GTATCTATTTAAAAGGAGTTAATAAATATGCGGCAAAGTATTATTAAATAAACTGTCAAT     419
Tn6000.T   GTATCTATTTAAAAGGAGTTAATAAATATGCGGCAAAGTATTATTAAATAAACTGTCAAT     416
Tn5397.T   GTATCTATTTAAAAGGAGTTAATAAATATGCGGCAAAGTATTATTAAATAAACTGTCAAT     413
           ************************************************************


Tn916.T    TTGATAGCGGGAACAAATAATTGGATGTCCTTTTTTAGGAGGGCTTAGTTTTTTGTACCC     479
Tn6000.T   TTGATAGCGGGAACAAATAATTGGATGTCCTTTTTTAGGAGGGCTTAGTTTTTTGTACCC     476
Tn5397.T   TTGATAGCGGGAACAAATAATTGGATGTCCTTTTTTAGGAGGGCTTAGTTTTTTGTACCC     473
           ************************************************************

                                                                       <---
Tn916.T    AGTTTAAGAATACCTTTATCATGTGATTCTAAAGTATCCGGAGAATATCTGTACTAGTGA     539
Tn6000.T   AGTTTAAGAATACCTTTATCATGTGATTCTAAAGTATCCGGAGAATATCTGTACTAGTGA     536
Tn5397.T   AGTTTAAGAATACCTTTATCATGTGATTCTAAAGTATCCGGAGAATATCTGTACTAGTCC     533
           ********************************************************

           ---------------------------------------------------------->
Tn916.T    CACTTCAAAA-AATGAGGTGTCTATTTTTTTATGGAGGAAAATCACGAATTCCTGCAGTA     598
Tn6000.T   CACTTCCAAA-GTTGAGGTGTCTT---TTTTATGGAGGAAAATCACGAATTCCTGCAGTA     592
Tn5397.T   ATTTGATTTTTCATATCAAGTGGTTTTTGTTATGGAGGAAAATCACGAATTCCTGCAGTA     593
             *       *       *   **      *  ****************************
           RBS       gusA
Tn916.T    AAGGAGAAAATTTTATGTTACGTCCTGTAGAAACCCCAACCCGTGAAATCAAAAAACTCG    658
Tn6000.T   AAGGAGAAAATTTTATGTTACGTCCTGTAGAAACCCCAACCCGTGAAATCAAAAAACTCG    652
Tn5397.T   AAGGAGAAAATTTTATGTTACGTCCTGTAGAAACCCCAACCCGTGAAATCAAAAAACTCG    653
           ************************************************************


Tn916.T    ACGGCCTGTGGGCATTCAGTCTGGATCGCGAAAACTGTGGAATTGATCAGCGTTGGTGGG    718
Tn6000.T   ACGGCCTGTGGGCATTCAGTCTGGATCGCGAAAACTGTGGAATTGATCAGCGTTGGTGGG    712
Tn5397.T   ACGGCCTGTGGGCATTCAGTCTGGATCGCGAAAACTGTGGAATTGATCAGCGTTGGTGGG    713
           ************************************************************


Tn916.T    AAAGCGCGTTACAAGAAAGCCGGGCAATTGCTGTGCCAGGCAGTTTTAACGATCAGTTCG    778
Tn6000.T   AAAGCGCGTTACAAGAAAGCCGGGCAATTGCTGTGCCAGGCAGTTTTAACGATCAGTTCG    772
Tn5397.T   AAAGCGCGTTACAAGAAAGCCGGGCAATTGCTGTGCCAGGCAGTTTTAACGATCAGTTCG    773
           ************************************************************
```

Figure 3-11 Sequence alignment of Tn*916*, Tn*6000* and Tn*5397* terminator constructs.

The -35 and -10 sequences of *tet*(M) promoter are in grey boxes. The terminator sequence is shown in red. The ribosome binding site and the start codon of *gusA* are shown in blue. Tn9*16*.T; contains Tn*916* terminator, Tn*6000*.T; contains Tn*6000* terminator and Tn*5397*.T; contains Tn*5397* terminator. An * (asterisk) indicates positions which have a single, fully conserved residue.

Terminator Constructs

A. Tn*916*-T

B. Tn*6000*-T

C. Tn*5397*-T

D. Promoter only (PO)

Figure 3-12 Schematic diagram of the transcriptional terminator constructs set. All constructs were cloned into pHCMC05-P*tet*(M)-PO and transformed into *B. subtilis* BS34A. The pointed 'P' indicates the *tet*(M) promoter. The red-shaded area represents the loop and uridine-rich region of the terminator sequence of Tn*916,* Tn*6000* and Tn*5397.* Panel D; Promoter only (PO) construct is a construct without any terminator sequence cloned in it and used as a positive control of *gus*A expression activity.

3.3.4 *In vitro* reporter gene assay of Tn*916*, Tn*6000* and Tn*5397* terminator constructs

Figure 3-13 shows the enzyme activity for terminator constructs of Tn*916,* Tn*6000* and Tn*5397* in comparison to the promoter only (PO) construct (comprised only the P*tet*(M)) and the plasmid-less *B. subtilis* (BS34A). By measuring the enzyme activity in *B. subtilis* BS34A, the level of enzyme activity decreased by 88.4% in Tn*916*.T construct and 77.9% in Tn*6000*.T construct as compared to the PO construct, confirming their termination activity. In the Tn*5397*.T construct, there is no significance difference observed when compared with the PO construct. The plasmid-less *B. subtilis* BS34A that serves as the negative control gave the lowest or presumably zero enzyme activity. Further investigation was done on the putative rho-independent terminator of Tn*916* that showed the highest termination activity.
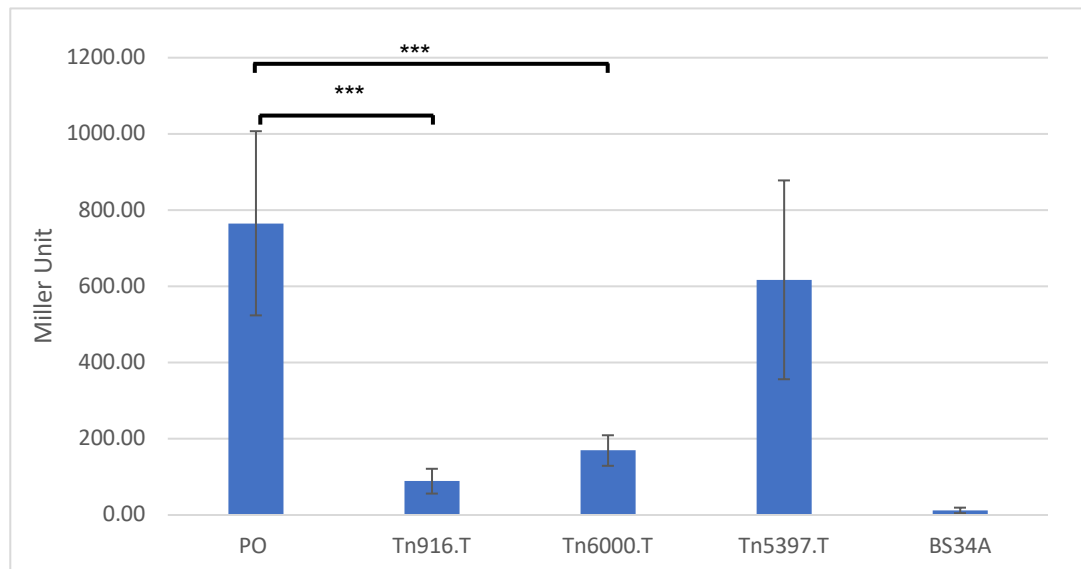
Figure 3-13 β-glucuronidase enzyme activity in cell lysates of *B. subtilis* BS34A containing various conjugative transposons terminator constructs.

The enzyme activity was measured after overnight growth. Error bars indicate the standard deviation of three independent experiments. The * (asterisk) indicate the constructs were statistically significantly different from the control group (PO) with the ***p≤0.0001 by using ordinary one-way ANOVA followed by Dunnett's multiple comparison test.

3.3.5 Generation of construct A (Tn*916* left end-BS34A genome junction region), construct B (Tn*916* joint ends region) and its mutated terminator variants (ΔSubA & ΔSubB constructs)

The Tn*916* terminator is hypothesized to prevent transcription of the conjugation genes when Tn*916* is integrated in the host genome. To test this, the terminator region is cloned in between the *tet*(M) promoter and a *gusA* reporter gene in a pHCMC05 shuttle vector, plus flanking chromosomal DNA of either the Tn*916* left end-BS34A genome junction (consruct A) or the joint ends of Tn*916* (construct B) (Figure 3-5). Construct A (representing the linear, integrated form of Tn*916*) and construct B (representing the excised and circularized form of Tn9*16*) were generated and verified by DNA sequencing. Extracted plasmids were also subjected to digestion with *Age*I, *Spe*I and *Bam*HI restriction enzymes. The results showed the expected size of DNA bands, confirming the successful directional cloning of each fragment (Figure 3-14). Sequence alignment of each construct with PO construct is shown in Figure 3-15 and 3-16, respectively. Construct A contains P*tet*(M) as indicated by -35 and -10, fragment A (plus the Tn*916* terminator) and the *gusA* gene. Construct B contains similar components, with fragment B replacing the fragment A.

The distance between P*tet*(M) and *gusA* might also be another factor that could affect the efficiency of the terminator. Therefore, another two constructs similar to A and B with the substitution on the poly-T tail (ΔSubA & ΔSubB) of the terminator sequence were also generated to disrupt the function of the terminators. The ΔSubA & ΔSubB were verified by DNA sequencing where

the poly-A tail of the Tn*916* terminator was successfully substituted to 9 G(s) [GGGGGGGGGG]**.** A schematic diagram of all the transcriptional terminator constructs is shown in Figure 3-17.
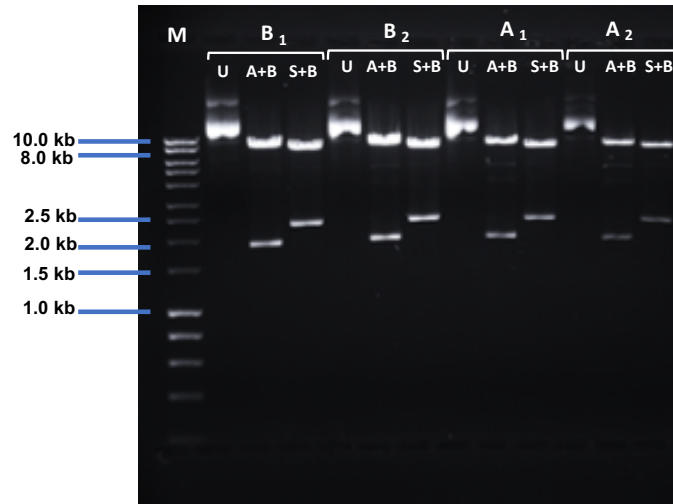
Figure 3-14 Agarose gel electrophoresis of the digestion analysis of the extracted A and B construct.

Lane M; HyperLadder™ 1kb, B_1; Construct B clone 1, B_2; Construct B clone 2, A_1; Construct A clone 1, A_2; Construct A clone 2. The (U) indicates undigested plasmid, (A+B); digestion with *AgeI* and *BamHI* and (S+B); double digestion with *SpeI* and *BamHI*. For construct A, digestion with (A+B); 8365 and 1890 bp fragments and digestion with (S+B); 8023 and 2232 bp fragments. For construct B, digestion with (A+B); 8365 and 1890 bp fragments and digestion with (S+B); 8025 and 2230 bp fragments.

```
                                                      - 35                              - 10
A     TACTCTCTTTGATAAAAAATTGGAGATTCCTTTACAAATATGCTCTTACGTGCTATTATT     357
PO    TACTCTCTTTGATAAAAAATTGGAGATTCCTTTACAAATATGCTCTTACGTGCTATTATT     420
      ************************************************************

A     TAAGTATCTATTTAAAAGGAGTTAATAAATATGCGGCAAAGTATTATTAAATAAACTGTC     417
PO    TAAGTATCTATTTAAAAGGAGTTAATAAATATGCGGCAAAGTATTATTAAATAAACTGTC     480
      ************************************************************

                                                              SpeI
A     AATTTGATAGCGGGAACAAATAATTGGATGTCCTTTTTTAGGAGGGCTTAGTTTTACTAG     477
PO    AATTTGATAGCGGGAACAAATAATTGGATGTCCTTTTTTAGGAGGGCTTAGTTTTACTAG     538
      ************************************************************

A     TCCCGGTCATGAATTGAAAGAACGGAATGGCCAGAATAGTTTATGTTATAAGTCCAACCC     537
PO    T-----------------------------------------------------------     538
      *

A     TACATACATATATCAATACAGGAAAGATAAATAAGAAGCAAAAATAGAGAAGCTTTCAAC     597
PO    ------------------------------------------------------------     538

A     CGGAGTAGAAATGGCTATTTGACTGTCTGAGGATTAATGGCTGTGTTAAACACTATGATT     657
PO    ------------------------------------------------------------     538

A     TTTCCTTCAAACTTATTTTCTAAGAAAAATAGCATAAAAATCTAGTTATCCGCATAAAAA     717
PO    ------------------------------------------------------------     538

A     CTGGACTTATCACACTTTATCAAGGTCAAAACCACTCAATTTACTACTAATTTACTACTT     777
PO    ------------------------------------------------------------     538

A     ATGAATGAGCTTTGATACGACGATTTATCCTTGAAAAGTGAAGATATAAAGATACTTCCA     837
PO    ------------------------------------------------------------     538

                                 Tn916.T
A     ATAAAATTTGAATATTTAATAGGTAGACACTTCAAAAAATGAGGTGTCTATTTTTTTTACC     897
PO    ------------------------------------------------------------     538

                      AgeI                                      RBS
A     CGATTTTGAAAGGAAGTACCGGTATGGAGGAAAATCACGAATTCCTGCAGTAAAGGAGAA     957
PO    ----------------ACCGGTATGGAGGAAAATCACGAATTCCTGCAGTAAAGGAGAA     584
                      ****************************************

          gusA
A     AATTTTATGTTACGTCCTGTAGAAACCCCAACCCGTGAAATCAAAAAACTCGACGGCCTG     1017
PO    AATTTTATGTTACGTCCTGTAGAAACCCCAACCCGTGAAATCAAAAAACTCGACGGCCTG     644
      ************************************************************
```

Figure 3-15 Sequence alignment of A and PO (promoter only) constructs.

The -35 and -10 sequences of *tet*(M) promoter are in green and highlighted in grey. The inserted fragment A is highlighted in yellow flanked by the *Spe*I and *Age*I restriction sites. The Tn*916* terminator (Tn*916*.T) sequence is shown in red. The ribosome binding site and the start codon of *gusA* are shown in blue. An * (asterisk) indicates positions which have a single, fully conserved residue.

```
                                                  - 35                    -10
B    TACTCTCTTTGATAAAAAATTGGAGATTCCTTTACAAATATGCTCTTACGTGCTATTATT        356
PO   TACTCTCTTTGATAAAAAATTGGAGATTCCTTTACAAATATGCTCTTACGTGCTATTATT        420
     **********************************************************

B    TAAGTATCTATTTAAAAGGAGTTAATAAATATGCGGCAAAGTATTATTAAATAAACTGTC        416
PO   TAAGTATCTATTTAAAAGGAGTTAATAAATATGCGGCAAAGTATTATTAAATAAACTGTC        480
     ***********************************************************

                                                                    SpeI
B    AATTTGATAGCGGGAACAAATAATTGGATGTCCTTTTTTAGGAGGGCTTAGTTTTACTAG        476
PO   AATTTGATAGCGGGAACAAATAATTGGATGTCCTTTTTTAGGAGGGCTTAGTTTTACTAG        533
     ***********************************************************

B    TGAAGCAACAGGAGCGTCTTGTTGCTTAGTAGTACAAATGAATTTACTACTTATTTACCA        536
PO   T-----------------------------------------------------------        533
     *

B    CTTCTGACAGCTAAGACATGAGGAAATATGCAAAGAAACGTGAAGTATCTTCCTACAGTA        596
PO   -----------------------------------------------------------        533

B    AAAATACTCGAAAGCACATAGAATAAGGCTTTACGAGCATTTAAGAAAATATAAAAAGAT        656
PO   -----------------------------------------------------------        533

B    AATTAGAAATTTATACTTTGTTTCTAAGAAAAATAGCATAAAAATCTAGTTATCCGCATA        716
PO   -----------------------------------------------------------        533

B    AAAACTGGACTTATCACACTTTATCAAGGTCAAAACCACTCAATTTACTACTAATTTACT        776
PO   -----------------------------------------------------------        533

B    ACTTATGAATGAGCTTTGATACGACGATTTATCCTTGAAAAGTGAAGATATAAAGATACT        836
PO   -----------------------------------------------------------        533

                                           Tn916.T
B    TCCAATAAAATTTGAATATTTAATAGGTAGACACTTCAAAAAATGAGGTGTCTATTTTTT        896
PO   -----------------------------------------------------------        533

                       AgeI                                    RBS
B    TACCCGATTTTGAAAGGAAGTACCGGTATGGAGGAAAATCACGAATTCCTGCAGTAAAGG        956
PO   -------------------ACCGGTATGGAGGAAAATCACGAATTCCTGCAGTAAAGG        580
                         ***************************************

        gusA
B    AGAAAATTTTATGTTACGTCCTGTAGAAACCCCAACCCGTGAAATCAAAAAACTCGACGG       1016
PO   AGAAAATTTTATGTTACGTCCTGTAGAAACCCCAACCCGTGAAATCAAAAAACTCGACGG        640
     ***********************************************************
```

Figure 3-16 Sequence alignment of B and PO (promoter only) constructs.

The -35 and -10 sequences of *tet*(M) promoter are in green and highlighted in grey. The inserted fragment B is highlighted in yellow flanked by the *Spe*I and *Age*I restriction sites. The Tn*916* terminator sequence is shown in red. The ribosome binding site and the start codon of *gusA* are shown in blue. An * (asterisk) indicates positions which have a single, fully conserved residue.

Terminator Constructs

A. Construct A
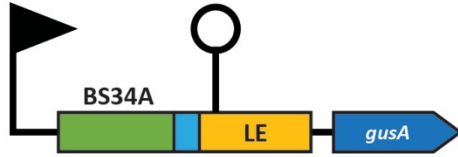
B. Construct B
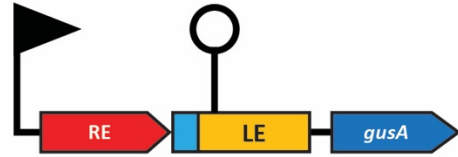
C. Construct ΔSubA
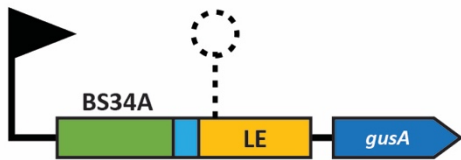
D. Construct ΔSubB
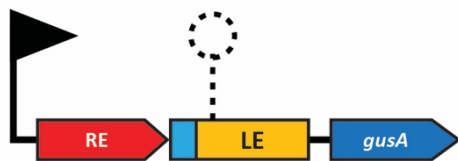
E. Tn*916*.T
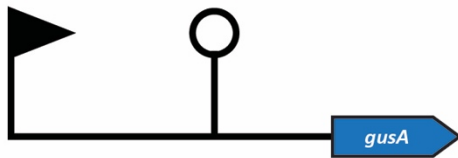
F. PO

Figure 3-17 Schematic diagram of the transcriptional terminator constructs set. All constructs were cloned into pHCMC05-P*tet*(M)-PO and transformed into *B. subtilis* BS34A. The pointed 'P' indicates the *tet*(M) promoter. The construct is color-coded as follows: *B. subtilis* genome fragment in green, Tn*916* left end fragment in orange, Tn*916* right end fragment in red and *gusA* in blue. The structure (⌀) represents the terminator, while the structure (⌀) represents the mutant terminator.

3.3.6 *In vitro* reporter gene assay of Tn*916* joint-ends and genome junction terminator constructs

As shown in Figure 3-18, ΔSubA and ΔSubB constructs which comprises the mutated T-residues following the stem loop structure demonstrated an increase in the enzyme activity as compared to A and B construct, but still indicates a lower enzyme activity than the PO construct. Interestingly, the enzyme activity observed is two-fold higher in the construct representing the circularised form (B) compared to the construct representing the linear (A), integrated form of Tn*916* (Figure 3-19). Construct A showed the lowest or presumably zero enzyme activity as it shares similar level of enzyme activity as the plasmid-less BS34A (negative control).
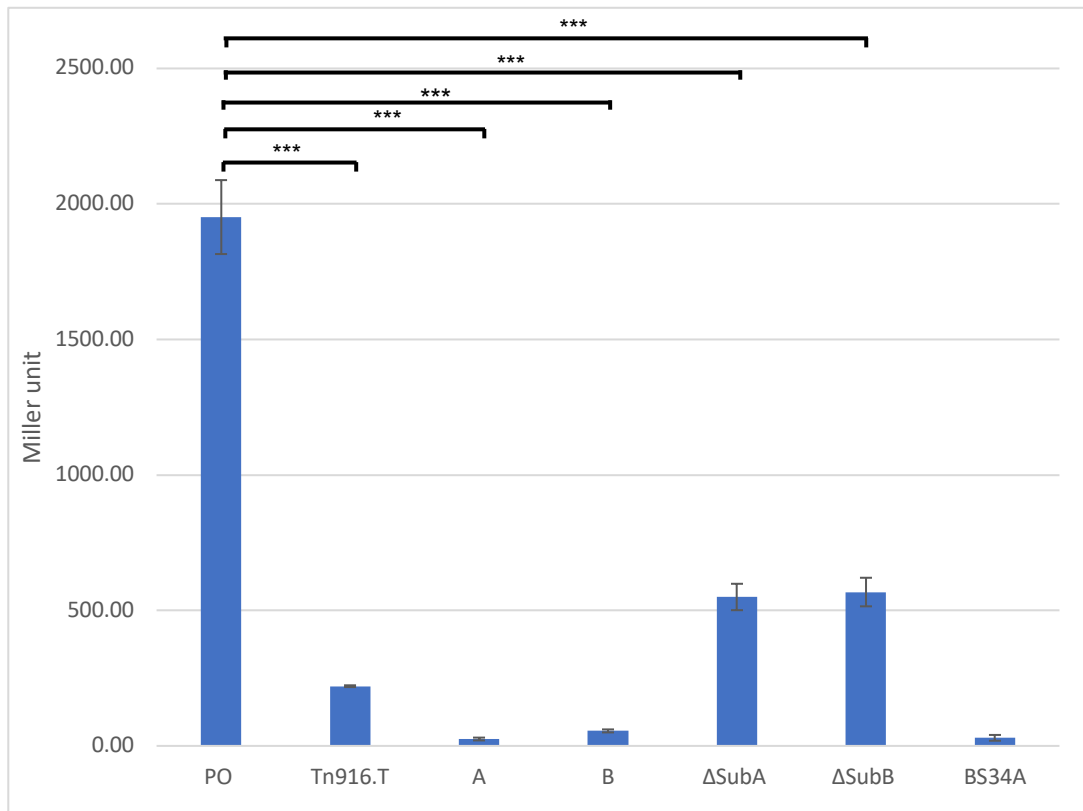
Figure 3-18 β-glucuronidase enzyme activity of Tn*916* terminator constructs.

The enzyme activity was measured after overnight growth. Error bars indicate the standard deviation of three independent experiments. The * (asterisks) indicate the constructs were statistically significantly different from the control group (PO) with the ***p≤0.0001 by using ordinary one-way ANOVA followed by Dunnett's multiple comparison test.
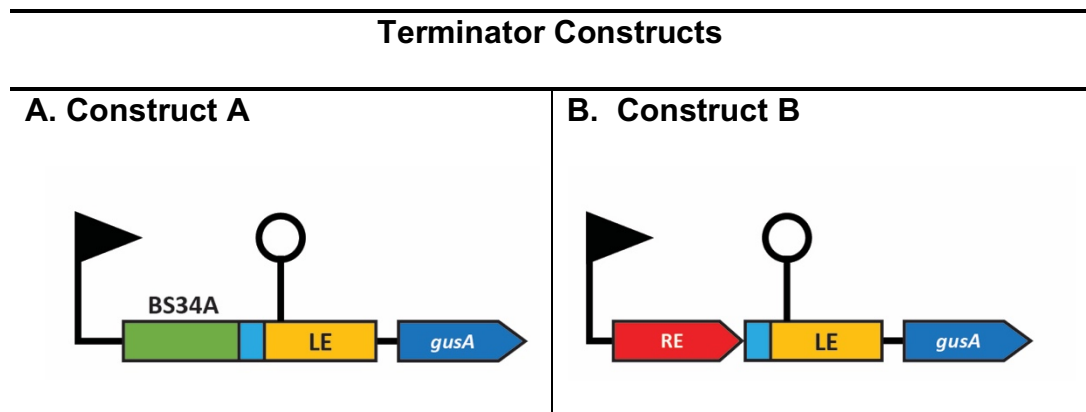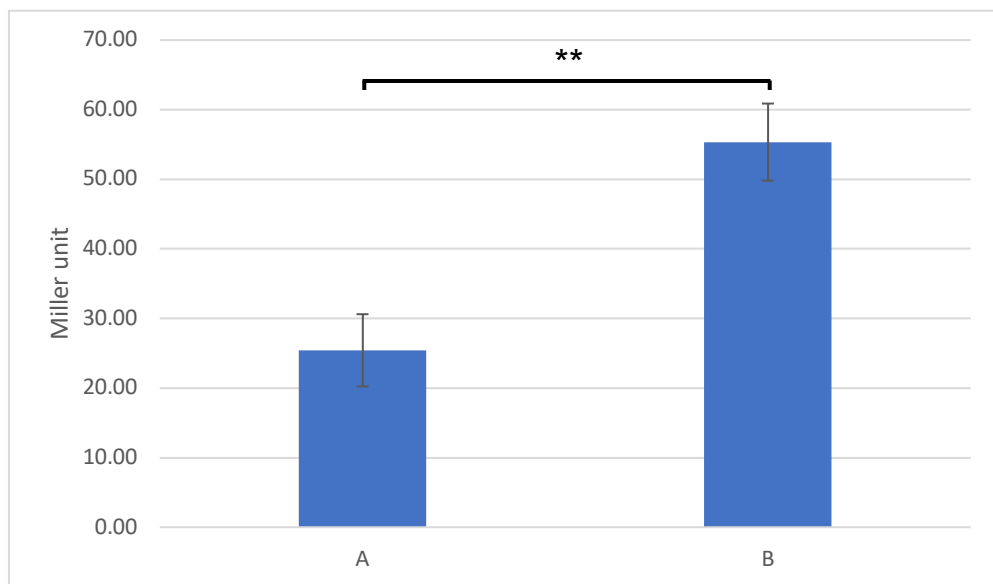
Figure 3-19 Comparison of β-glucuronidase enzyme activity in A and B constructs.

Panel A: Schematic diagram of the terminator construct A and B. The pointed 'P' indicates the *tet*(M) promoter. The construct is color-coded as follows: *B. subtilis* genome fragment in green, Tn*916* left end fragment in orange, Tn*916* right end fragment in red and *gusA* in blue. The structure (⟨) represents the terminator. **Panel B:** The β-glucuronidase enzyme activity in B construct (representing the circularised form of Tn*916*) is two-fold higher in comparison to A construct (representing the linear, integrated form of Tn*916*). Error bars indicate the standard deviation of three independent experiments. The * (asterisks) indicate that the construct A and B were statistically significantly different with the **p≤0.01 by using an unpaired student's t-test.

3.4 Discussion

All of the predicted terminators from Tn*916* and Tn*916*-like conjugative transposons were shown to match a canonical descriptor of rho-independent terminators consisting of a G-C rich dyad symmetry stem, a loop, a T-stretch and a 0-2 nt length spacer in between the stem and the T-stretch region (Lynn *et al.*, 1988, d'Aubenton Carafa *et al.*, 1990, Lesnik *et al.*, 2001, Macke *et al.*, 2001). The hairpin structure and the T-stretch are the standard features of intrinsic terminators, but differences in exact sequence and secondary structure formation cause variations in the efficiency of termination and in mechanisms (Wilson & von Hippel, 1995). This is observed in this study where the Tn*916* terminator that contains the highest number of G-C pair and thymine, have the lowest transcriptional read-through followed by the Tn*6000* and the Tn*5397* terminator.

A study on a relationship between hairpin stability and termination efficiency by altering the length of λt_{R2} hairpin stem of *E. coli* has been done (Wilson & von Hippel, 1995). Deletion of one G-C pair from the top of the hairpin stem resulted in a significant decrease in termination efficiency. Conversely, an addition of a G-C pair at the same position slightly increased the termination efficiency (Wilson & von Hippel, 1995). Moreover, G-C pairings that provide stronger hydrogen bonds in comparison to A-U pairings enhanced the overall stability of the stem loop structure.

Although Tn*916* and Tn*6000* terminators have a similar GC content, the number of thymine residue in the Tn*916* T-stretch region is higher, making it

a more efficient terminator. Reduction of a thymine number has been shown to weaken the termination activity (Christie *et al.*, 1981, Stroynowski *et al.*, 1983). The role of the uridine stretch in the RNA is to provide a weak A:U bond which hybrid melting at the proximal region (U_1-U_5) is prerequisite for hairpin formation (Martin & Tinoco, 1980, Gusarov & Nudler, 1999). The distal portion of the uridine stretch has also been reported to be responsible for transcriptional pausing by slowing down a ternary elongation complex (TEC) at the termination point, thus giving the hairpin extra time to be formed (Gusarov & Nudler, 1999). In this study, intrinsic terminators identified at the same region from Tn*6002*, Tn*6003*, Tn*2010* and Tn*6087* share similar DNA sequence with Tn*916*, therefore it is assumed that the termination efficiency will be the same if tested in a reporter construct.

Overall, the *in vitro* experimental data is consistent with the predicted termination efficiency estimated via an algorithm including the Tn*5397* terminator. The *d* score of the Tn*5397* putative terminator was calculated to give a negative value and when tested *in vitro*, it demonstrates no significant difference to the PO construct. This result is expected, as it clearly did not possess characteristics of a strong terminator. Although Tn*916* and Tn*5397* are very closely related, the left end region; upstream of the conjugation genes where both tested terminators are found, are non-identical. Tn*5397* does not contain 201 bp of the Tn*916* left end segment which is replaced by 180 bp of unrelated sequence, resulting in seven bp deletion of *orf*24 (Wang *et al.*, 2000). Sequence alignment shows that these two putative terminators (from Tn*916* and Tn*5397*) are non-identical (Figure 3-8).

On another note, integration of Tn*5397* into a genome is far more site specific in comparison to Tn*916* that have a variety of target sites and hosts. Tn*5397* inserts into two specific target sites in *Clostridium difficile* CD37 (with a strong preference for a region designated as *attB*$_{Cd}$) and multiple sites in *B. subtilis* (Mullany *et al.*, 1990, Wang *et al.*, 2006). The target sites of Tn*5397* always contains a conserved central GA dinucleotide (Wang *et al.*, 2000). In contrast, Tn*916* inserts into multiple target sites in almost all hosts. These target sites include regions with differing transcriptional activity such as the intergenic regions which typically contain promoters. Therefore, we hypothesized the presence of rho-independent terminator upstream of the conjugation module is needed as a control mechanism to prevent transcriptional read through from reaching the conjugation genes. This way, it will be able to maintain its stability in the host chromosome.

Whereas, due to the site specificity of Tn*5397*, the presence of this terminator in Tn*5397* is not essential and consequently, it will be lost by genetic drift. It might be that the evolutionary selective pressure for the maintenances of terminator is no longer there with Tn*5397* because of its target site specificity. Most of the target sites of Tn*5397* has been identified to be within an open reading frame (ORF). For example, in *C. difficile*, Tn*5397* inserted into an ORF that was predicted to encode a protein that has limited homology to *pivNM-2* (pilin gene inverting protein) (Wang *et al.*, 2000) and the *fic* gene (Wang *et al.*, 2006, Mullany *et al.*, 2015). In *E. faecalis* JH2-2, Tn*5397* inserts into an ORF encoding a IIA component of a mannose/sorbose-specific sugar phosphotransferase system (NCBI accession number NP_814245) (Jasni *et al.*, 2010). Our hypothesis is that as target site selection has become more

specific, the need to be able to cope with differering levels of transcription from upstream regions has decreased.

The Tn*916* and Tn*6000* terminators satisfy the minimal value of the parameter (n_T) which is ≥ 2.895, although the Tn*916* terminator efficiency were calculated by excluding one of the minimal conditions for the T-stretch region described by d' Aubenton Carafa *et al.* (1990), where the stretch must begin by two consecutive T residues. However, the Tn*916* terminator was initially searched by using an improved algorithm called RNAMotif (Macke *et al.*, 2001) based on a descriptor developed by Lesnik *et al.* (2001). The descriptor is similar to that of d' Aubenton de Carafa *et al.* (1990), but additional sequence constraints on the T-stretch region does not require the proximal T-stretch to begin with at least two T residues as long as it contains at least three T residues, no more than one G, and no 5'-TVVTT stretches (V is A, C or G).

In this study, a second set of Tn*916* terminator construct variants was generated to investigate its role in preventing the transcription of the conjugation genes. We hypothesized that when Tn*916* is integrated in its host chromosome, the conjugation genes will not be expressed. This is important in order to keep it stabilized so it remains integrated in the genome. Conversely, when Tn*916* is excised, a transcription read-through from Porf7 (and possibly Pxis and Pint) past the joint of the circular form is expected to occur allowing the transcription of the conjugation genes (Celli & Trieu-Cuot, 1998). Since there are no promoters within the transposon upstream of the conjugation region identified, it is likely that the promoters from the regulatory region are responsible for the transcription of the genes in the conjugation

module (Celli & Trieu-Cuot, 1998). However, in order for this to occur, circularization of Tn*916* and a transcriptional read through beyond the terminator is required.

This prediction is relevant to our results which demonstrated that the enzyme activity observed is twofold higher in the construct representing the circularized form compared to the construct representing the linear, integrated form, although the mechanism of transcriptional read-through remains elusive. Despite the fact that the actual increase of the enzyme activity in construct B (representing the circularized form) is low in comparison to the promoter only construct, it may be enough for conjugation to occur. Biologically, that small amount of increase maybe significant within the cell as overexpression of the conjugation genes may be detrimental for the cell. Furthermore, the conjugation frequency of Tn*916* is relatively lower in comparison to conjugative plasmid. For example, the transfer frequency of broad-host range RP4 plasmid ranges from 10^{-3} to 10^{-6} in comparison to Tn*916* which is within the range of 10^{-4} to $<10^{-9}$ per donor cell (Bertram *et al.*, 1991, Marra *et al.*, 1999, Grohmann *et al.*, 2003).

Sequences situated upstream of the stem loop structure and downstream of the T-stretch region might play a role in determining the efficiency of a terminator (d'Aubenton Carafa *et al.*, 1990). In this study, a large difference in termination activity observed in construct A and B in comparison to the Tn*916*.T construct. Since construct A and B includes the flanking chromosomal DNA of either the Tn*916* left end-BS34A genome junction (region A: representing the linear, integrated form) or the joint ends of Tn*916*

(region B: representing the excised and circularised form) (Figure 3-5), this could be the factor contributing to the decreased enzyme activity.

The distance between P*tet*(M) and *gusA* might also be another factor that could affect the efficiency of the terminator. Therefore, construct ΔSubA and ΔSubB with the substitution on the poly-T tail were generated. In these mutant constructs, the whole T-region consisting of 5'-TATTTTTTT-3' was substituted to 5'-GGGGGGGGG-3' eliminating the function of both proximal and distal T-region. As expected, an increase in the enzyme activity in comparison to A, B and Tn*916*.T construct was observed confirming the importance of the T-stretch region in playing its role in the transcriptional pausing (d'Aubenton Carafa *et al.*, 1990, Gusarov & Nudler, 1999, Lesnik *et al.*, 2001) and the disruption of RNA:DNA hybrid duplex (Martin & Tinoco, 1980). However, the termination activity is not completely eliminated as the enzyme activity is still lower than the PO construct. This may be due to the fact that the stem-loop structure was not disrupted leading to partial intrinsic termination. The nucleation of the hairpin is crucial for a complete destabilization and dissociation of ternary elongation complex (TEC) (Wilson and Hippel, 1995; Gusarov and Nudler, 1999). Further investigation is therefore necessary to validate the distance factor in between the promoter and the reporter gene.

3.5 Conclusions

In conclusion, we have for the first time, identified and experimentally verified a group of conserved terminators in the conjugation region of Tn*916* and Tn*916*-like genetic elements. The termination efficiencies are correlated with the number of thymine residue and the GC content of the stem. Further analysis on Tn*916* demonstrates that the enzyme activity observed is two-fold higher in the construct representing the circularized form compared to the construct representing the linear, integrated form of Tn*916*. This data supports our hypothesis that the terminator efficiency is modulated upon excision and circularization of Tn*916*, which is the exact time when Tn*916* would require expression of its conjugation genes. This terminator is biologically important to keep the element stabilised and remaining integrated into the genome. Therefore, unravelling the function of the terminator is important for fuller understanding of the transcriptional and translational operators of this element.

4 Investigation into the Role of Tn*916* terminator

4.1 Introduction

A group of structurally conserved terminators within the conjugation module of Tn*916* and Tn*916*-like genetic elements have been identified and experimentally verified (Chapter 3). From our initial *in vitro* study, the highest termination efficiency was observed for the Tn*916* terminator (and other terminators from Tn*916*-like elements that shares the same sequence and predicted secondary structure). However, the terminator-like structure from Tn*5397* was demonstrated to show no significant difference in their activity when compared to the positive control or the activity could be very subtle that it is unable to be detected via *in vitro* assay (Chapter 3).

The integration of Tn*5397* within the host genome is more site specific in comparison to Tn*916* which inserts into multiple sites within a wide variety of Gram-positive and negative bacteria (Bertram *et al.*, 1991, Poyart *et al.*, 1995, Roberts *et al.*, 2003, Mullany *et al.*, 2012). For example, in both *C. difficile* 630 and *E. faecalis* genomes, Tn*5397* inserts only into one single site (Wang *et al.*, 2006, Jasni *et al.*, 2010). All of the analysed sites have a central GA dinucleotide sequence (Wang & Mullany, 2000, Wang *et al.*, 2000, Jasni *et al.*, 2010). In *C. difficile*, Tn*5397* inserts into a *fic* gene that encodes a domain termed Fic (filamentation processes induced by cAMP) (Wang *et al.*, 2006, Mullany *et al.*, 2015). Specifically within the genome of *B. subtilis,* it has been demonstrated that Tn*5397* inserts into multiple sites (Wang *et al.*, 2000). However, when the original target site (*fic* DNA) from *C. difficile* is introduced into the *B. subtilis* genome, Tn*5397* always inserted into this site. This shows

that the Tn*5397* has a strong preference for this particular target site and the target selection is not related to obvious host factors (Wang *et al.*, 2006).

In contrast, Tn*916* inserts into multiple target sites in *C. difficile* 630 with a consensus motif sequence of 5'-TTTTA[AT][AT][AT][AT]AAAA-3' (Mullany *et al.*, 2012). An exception to this is in the nontoxigenic *C. difficile* strain CD37 where Tn*916* inserts into only one target site, which is an intergenic and AT-rich region (Wang *et al.*, 2000). We suggested that the presence of the terminator (located upstream of the conjugation module) allows Tn*916* to use multiple target sites. Tn*916* has been shown to integrate into the genome in variable AT-rich sites in a wide variety of bacteria, suggesting the ability of this element to insulate itself within regions of differing transcriptional activity. Therefore, we hypothesised that the presence of the terminator within the conjugation module of Tn*916* is needed to protect the element when it is integrated at the region with high transcriptional activity. The terminator may act as a control mechanism to prevent variable transcriptional read-through derive from the host genome from reaching the conjugation genes of Tn*916*. Whereas for Tn*5397*, the presence of terminator may not be essential due to their target site preferences.

Another notable difference between Tn*916* and Tn*5397* is their recombination module, where instead of *int* and *xis* in Tn*916*, Tn*5397* contains only *tndX* encoding large serine recombinase (Wang & Mullany, 2000, Roberts *et al.*, 2001). Divergence of Tn*5397* from Tn*916* is observed at the last 180 bp end which includes the absence of first seven nucleotides of *orf24* within Tn*5397*. In Tn*916*, the first eleven bp of the *orf24* is part of the terminator sequence.

Possibly, Tn*5397* is a result of recombination event between two ancestral elements; one contained *tndX* and another one containing a Tn*916*-like conjugation system. The TndX is a large serine recombinase, related to TnpX from the mobilizable transposons Tn*4451* and Tn*4453* (Bannam *et al.*, 1995). The site selectivity of large serine recombinase is more specific in comparison to tyrosine recombinase (Curcio & Derbyshire, 2003). The choice of Tn*5397* to be inserted into an open reading frame is probably based on target site selection mediated by TndX.

To investigate the biological function of the Tn*916* terminator, a mutant (Tn*916* with a deletion of the terminator) was generated and denoted as Tn*916*ΔTerm. It is hypothesised that with deletion of the terminator, it will alter the conjugation activity of the element. In this chapter, the mutant *B. subtilis* Tn*916*ΔTerm was constructed and investigated for their conjugal transfer activity using filter mating experiments.

4.2   Materials and methods

4.2.1   Bacterial strains and plasmids

E. coli α-select (silver efficiency) was used for cloning the mutant cassette (Figure 4-1 to 4-4). *B. subtilis* BS34A that carries a single copy of wild type Tn*916* was used as the host for the development of BS34A mutant strain (BS34A Tn*916*ΔTerm). In filter mating experiments, *B. subtilis* BS34A and BS34A Tn*916*ΔTerm were used as the donor strains of wild type Tn*916* and mutant Tn*916*, respectively. Two other strains of *B. subtilis* that are resistant

to specific antibiotics were selected and used as recipients in filter mating experiments; *B. subtilis* CU2189 Rif[R] Nal[R] (resistant to rifampicin and nalidixic acid) and *B. subtilis* BS168 Erm[R] (resistant to erythromycin). *E. faecalis* JH2-2 was also used as a recipient in filter mating experiment. All strains were grown on BHI agar or broth at 37°C for 16-24 hr with shaking at 200 rpm with appropriate concentrations of antibiotic(s) as listed in Table 4-1. All constructs and strains used in this study are listed in Table 2-1 and Table 2-2 in Chapter 2.

4.2.2 Generation of mutant cassette by Splicing Overlap Extension PCR (SOE-PCR)

To develop the mutant, a mutant cassette carrying a selectable marker (*cat*P) flanked by regions homologous to the target locus was generated by SOE-PCR method. The mutant cassette was constructed by splicing four fragments; Fragment 1 is the upstream region which is homologous to the BS34A genome and denoted as upstream sequence (UPS), Fragment 2 is *cat*P; chloramphenicol resistance gene chosen as the selective marker for the mutant. Fragment 3 and 4 are amplified from *B. subtilis* BS34A genome targeting the left end junction region of Tn*916* where the terminator structure is found. It is separated into two fragments so that the deletion of the terminator sequence (32 bp) can be generated (Figure 4-1).
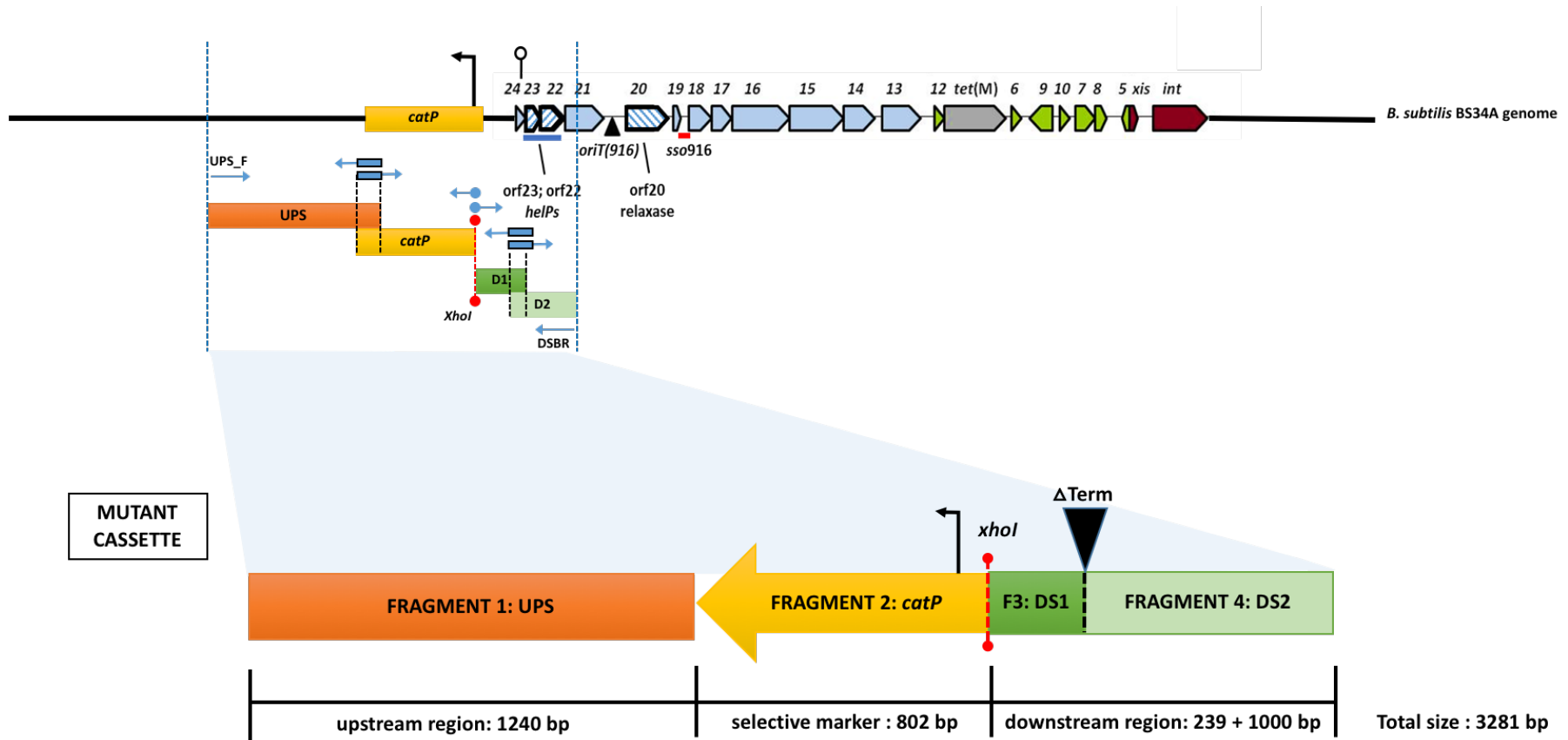
Figure 4-1 The structure of mutant cassette.

Schematic figure (top) showing the Tn*916* conjugative transposon, integrated into the *B. subtilis* BS34A chromosome with the recombined mutant cassette. Coloured arrow boxes represent the open reading frames (ORFS) and the orientation of the genes in Tn*916* (conjugation (blue);

recombination (red); regulation (green) and the accessory gene *tet*(M) (grey)). The enlarged figure (bottom) showing the detail structure of the mutant cassette consist of fragment 1:UPS (orange), fragment 2: *cat*P in (yellow), fragment 3: DS1 (dark green) and fragment 4: DS2 (light green). The structure (⟶) represents a primer; the structure (▭⟶) represents a Phusion primer (contains overlapping region); structure (

●⟶) represents primer with added restriction enzyme site (*xho*I); structure (⌐) represents the promoter of *cat*P, structure (⸰) represents the rho-independent terminator located within ORF24 of Tn*916*, structure (▼) represents the deleted terminator region labelled as ΔTerm.

Fragment 1 was amplified from the BS34A genome using the primer pair; UPS_F and UPS-catP-R3 (contains 20 bp overlapping with Fragment 2). Fragment 2 was amplified from the pRPF185 plasmid using a primer pair; catP_F2 (20 bp overlapping with Fragment 1) and catP_R3_*xho*I. Fragment 1 and 2 were spliced by using Splicing Overlap Extension-PCR (SOE-PCR) method with a primer pair UPS_F and catP_R3_*xho*I generating a 2042 bp SOE-PCR product denoted as Fragment [1+2] (Figure 4-2).
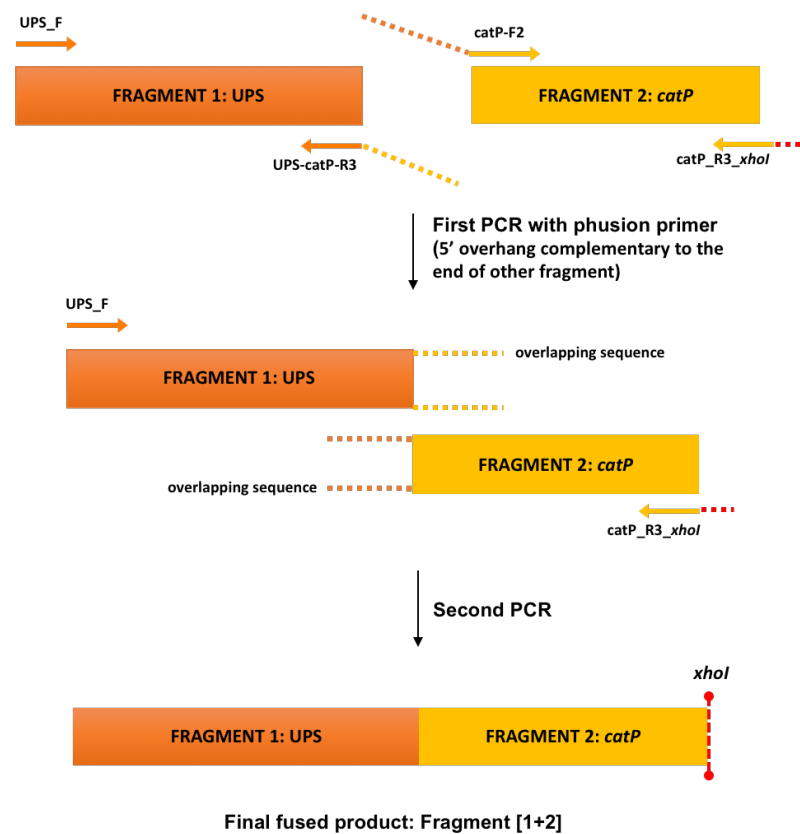
Figure 4-2 Construction of Fragment [1+2] by SOE-PCR.

To splice two DNA fragments, Phusion primers are used at the ends that are to be joined. The Phusion primer is designed such that it has a 5' overhang complementary to the end of other fragment. Fragment 1: UPS (in orange box) and fragment 2: *catP* (in yellow box). The structure (→) represents a primer; structure ( ) and ( ) represents a Phusion primer pair; structure ( ) represents primer with added restriction enzyme site (*xho*I).

147

Fragment 3 (DS1) was amplified from the BS34A genome using BSA_F_*xho*I and UPS_BR (contains 22 bp overlapping with fragment 4), while fragment 4 (DS2) was amplified from the BS34A genome using DS_BF (contains 22 bp overlapping with fragment 3) and DS_BR. Fragment 3 and 4 were spliced by SOE-PCR using the primer pair BSA_F_*xho*I and DS_BR generating a 1239 bp SOE-PCR product denoted as Fragment [3+4] (Figure 4-3).
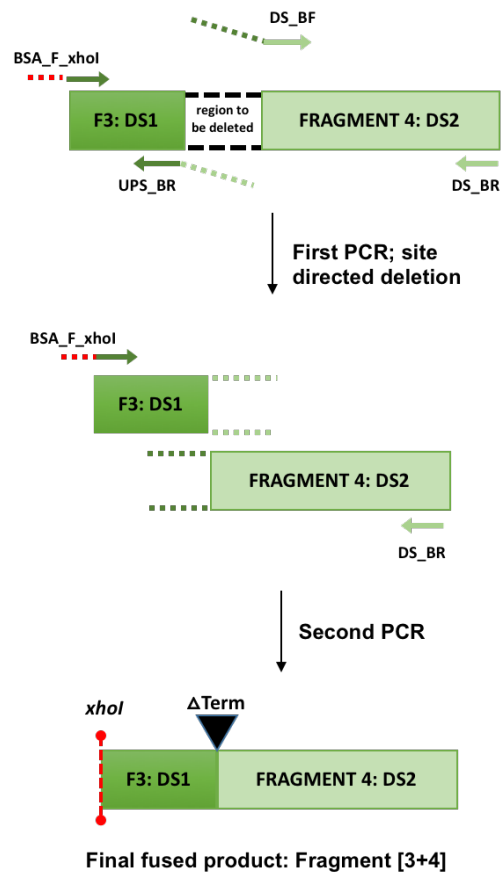
Figure 4-3 Construction of Fragment [3+4] by SOE-PCR.

Deletion of the terminator sequences (32 bp) was done by SOE-PCR using site directed deletion method as shown above. Fragment 3: DS1 (in dark green box) and fragment 4: DS2 (in light green box). The structure (➡) represents a primer; structure (⇢) and (⇠) represents a Phusion primer (contains overlapping region); structure (•••➡) represents primer with added restriction enzyme site (*xho*I); structure (♀) represent the rho-independent terminator, structure (▼) represents the deleted terminator region labelled as ΔTerm.

The two spliced fragments (Fragment [1+2] and Fragment [3+4]) were then cloned into pGEM-T Easy vector (Promega, UK) respectively, and later digested with *xho*I (a restriction site which has been added at the ends where fragment 2 and 3 are to be joined), to create sticky ends followed by ligation to produce a final product of fusion fragments [1+2+3+4]. Finally, this mutant cassette of 3281 bp in size (consisting of all four fused fragments) is cloned into pGEM-T Easy vector, screened and verified by sequencing. This construct is denoted as pGEM-T/Tn*916*ΔTerm; a mutant construct with a deleted Tn*916* terminator with *cat*P as the selective marker with the total size of 6296 bp (Figure 4-4).

SOEing PCR of F1 & F2

FRAGMENT 1: UPS

FRAGMENT 2: catP

xhoI

Cloning of [UPS+catP] into
pGEM-T Easy vector

FRAGMENT 1: UPS

FRAGMENT 2: catP

xhoI

pGEM-T /UPS+catP

SOEing PCR of F3 & F4

xhoI

ΔTerm

F3: DS1

FRAGMENT 4: DS2

Cloning of [DS1+DS2] into
pGEM-T Easy vector

xhoI

F3: DS1

FRAGMENT 4: DS2

pGEM-T /DS1+DS2
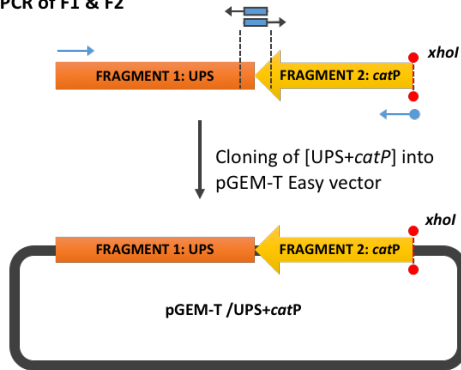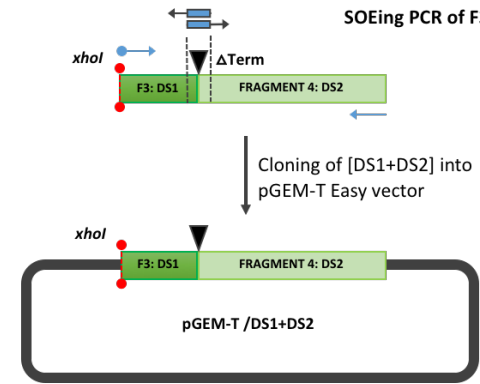
Digestion with *EcoRI* HF
and dephosphorylation of
5' ends of DNA using CIP

Restriction digest with *xhoI*
and ligation

Digestion with *EcoRI* HF
and dephosphorylation of
5' ends of DNA using CIP

FRAGMENT 1: UPS

FRAGMENT 2: catP

xhoI

+

xhoI

ΔTerm

F3: DS1

FRAGMENT 4: DS2

MUTANT CASSETTE

FRAGMENT 1: UPS

FRAGMENT 2: catP

ΔTerm

F3: DS1

FRAGMENT 4: DS2

A-tailing with Taq Polymerase and
cloning of [**MUTANT CASSETTE**]
into pGEM-T Easy vector

FRAGMENT 1: UPS

FRAGMENT 2: catP

ΔTerm

F3: DS1

FRAGMENT 4: DS2

pGEM-T /Tn*916* ΔTerm

150
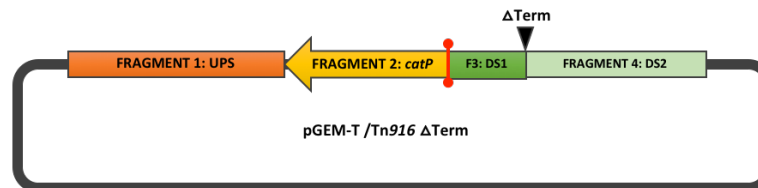
Figure 4-4 Construction of mutant cassette by SOE-PCR and ligation.

The structure ($\longrightarrow$) represents a primer; structure ($\blacksquare\!\!\rightarrow$) represents a Phusion primer (contains overlapping region); structure ($\bullet\!\!\rightarrow$) represents primer with added restriction enzyme site (*xho*I); structure (⌐) represents the promoter of *cat*P, structure (⦵) represent the rho-independent terminator located within ORF24, structure (▼) represents the deleted terminator region labelled as ΔTerm.

4.2.3 Transformation of *E. coli* with pGEM-T/Tn*916*ΔTerm

E. coli transformation with pGEM-T/Tn*916*ΔTerm was carried out using competent cells α-select silver efficiency (Bioline, UK) according to the standard heat-shock transformation protocol as described in section 2.4.11. Transformants were selected on LB agar supplemented with ampicillin (100 µg/mL), chloramphenicol (10 ug/mL) and IPTG/X-gal for blue-white screening after 18-24 hrs of incubation at 37°C. Plasmid extraction was carried out according to protocol described in section 2.4.2, followed by restriction digest of the pGEM-T/Tn*916*ΔTerm using *Sca*I to linearise the plasmid. The linearised and the non-linearised plasmid were subsequently transformed into *B. subtilis* BS34A.

4.2.4 Preparation of *B. subtilis* BS34A competent cells

The *B. subtilis* competent cells were prepared according to the protocol by Hardy (1985) as describe in section 2.4.9.

4.2.5 Transformation of *B. subtilis* BS34A and homologous recombination of the mutant cassette (pGEM-T/Tn*916*ΔTerm)

B. subtilis BS34A (that carries a single copy of wild type Tn*916*) transformation with linearised and non-linearised pGEM-T/ Tn*916*ΔTerm was carried out according to the protocol describe in section 2.4.10. Transformants were selected on BHI agar supplemented with chloramphenicol (10 ug/mL). The homologous regions that flanked the terminator structure and *cat*P is expected

to integrate into the target locus of *B. subtilis* BS34A chromosome via double recombination (either at the upstream or downstream of homologous regions).

4.2.6 Validation of the integrated mutant cassette into the BS34A chromosome

Validation of the integrated mutant cassette into the BS34A chromosome was done via PCR using three pairs of primers targeting three different regions denoted as R1, R2 and R3. Region 1 (R1) includes the area that expands outside the target integration site of the mutant cassette. Region 2 (R2) and 3 (R3) were amplified using outward primer from the mutant cassette (catP_LR and catP_LF) and the inward primer (HR_F and HR_R) from the genomic region just outside the homologous region, respectively (Figure 4-5).
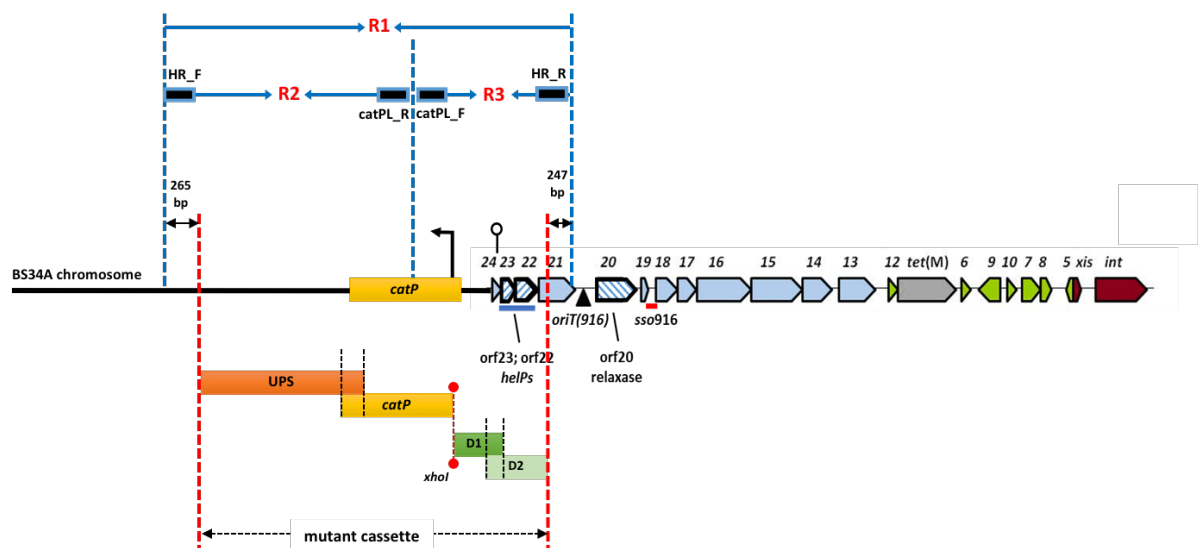
Figure 4-5 Schematic diagram showing the amplification region of R1, R2 and R3 for the validation of the mutant cassette integration.

4.2.7 Selection of rifampicin and nalidixic acid resistant *B. subtilis* CU2189 and erythromycin resistant *B. subtilis* BS168 as recipients for filter mating experiments

A triplicate of 50 mL antibiotic free brain heart infusion (BHI) broth was inoculated with a single colony of *B. subtilis* strain CU2189 or *B. subtilis* strain BS168. The inoculated broth was incubated overnight at 37°C, 200 rpm for 16-18 hrs. The overnight culture was spun down at 4500 x *g* (5000 rpm) for 15 mins and the pellet was resuspend in 1 mL fresh BHI broth (10 µl of the suspension is added to 90 µL of 1X Phosphate Buffered Saline (PBS) and used for serial dilution to determine the number of cells). An aliquot of 100 µL of the suspension was plated onto BHI agar plates that were supplemented with rifampicin and nalidixic acid at a concentration of 25 µg/mL and 10 µg/mL, or with erythromycin at a concentration of 10 µg/mL. These plates were incubated at 37°C and observed every day (for five days) for colony growth. The experiment was repeated in three biological replicates. The rifampicin and nalidixic acid resistant isolates (occurred through point mutation) and erythromycin resistant isolates were individually picked and streaked onto fresh selected plates. Bacterial stocks are maintained in 1 ml aliquots of 20% (v/v) sterilised glycerol in Luria Bertani (LB) broth at - 80°C. The selected resistant isolates were initially identified by 16S rDNA amplification using 27F and 1392R primers. Identification to the species level was done by *gyrA* sequencing using p-gyrA-f and p-gyrA-r primer pair (Chun & Bae, 2000).

4.2.8 Whole genome sequencing and *in silico* analysis of the *B. subtilis* BS34A Tn*916* ΔTerm, *B. subtilis* CU2189 Rif^R Nal^R and *B. subtilis* BS168 Erm^R

B. subtilis BS34A Tn*916*ΔTerm (carrying the mutant Tn*916*), *B. subtilis* CU2189 Rif^R Nal^R and *B. subtilis* BS168 Erm^R strains were sent to MicrobesNG (http://www.microbesng.uk) for whole genome sequencing using 2 × 250 bp paired-end reads on the Illumina platform. *De novo* assembly of each of the genomes was carried out with SPAdes (Bankevich *et al.*, 2012) via MicrobesNG (Birmingham). An automated annotation of the assembled genomes was carried out using Prokka (Seemann, 2014). Mutations predictions (single-nucleotide polymorphisms (SNPs), insertion, deletion and duplications) were performed by using Breseq (Deatherage & Barrick, 2014). Sequence alignment was carried out by using BioEdit software version 7.2.0 (Hall, 1999), SnapGene 3.2.1 (GSL Biotech LLC, US) and Clustal Omega (http://www.ebi.ac.uk/Tools/msa/clustalo).

4.2.9 Transfer experiments and transconjugants selection

Transfer of Tn*916* and Tn*916*ΔTerm (contains Tn*916* with a deleted terminator) between respective donor and various recipient strains were carried out via filter-mating based on a protocol described in section 2.8. The flow chart of the filter-mating experiment is illustrated in Figure 4-6. The donors, recipients and transconjugants were grown overnight in BHI broth

155

supplemented with appropriate antibiotic(s) with the concentration as listed in

Table 4-1 below;

Table 4-1 List of donors and recipients used in filter-mating experiment and their respective antibiotic concentration.

Strains	Antibiotic (µg/mL)
Donors	
BS34A Tn*916* WT Tc^R	Tetracycline (10 µg/mL)
BS34A Tn*916* ΔTerm Tc^R Cm^R	Tetracycline (10 µg/mL), Chloramphenicol (10 µg/mL)
Recipients	
B. subtilis CU2189 Rif^R Nal^R	Rifampicin (10 µg/mL), Nalidixic acid (5 µg/mL)
B. subtilis BS168 Erm^R	Erythromycin (10 µg/mL)
E. faecalis JH2-2 Rif^R Fa^R	Rifampicin (25 µg/mL), Fusidic acid (5 µg/mL)
Transconjugants	
B. subtilis CU2189 Rif^R Nal^R Tc^R	Rifampicin (10 µg/mL), Nalidixic acid (5 µg/mL), Tetracycline (10 µg/mL)
B. subtilis BS168 Erm^R Tc^R	Erythromycin (10 µg/mL), Tetracycline (10 µg/mL)
E. faecalis JH2-2 Rif^R Fa^R Tc^R	Rifampicin (25 µg/mL), Fusidic acid (5 µg/mL), Tetracycline (10 µg/mL)

Abbreviations: Tc^R, tetracycline-resistant; Cm^R, chloramphenicol-resistant; Erm^R, erythromycin-resistant; Rif^R, rifampicin-resistant; Fa^R, fusidic acid-resistant.
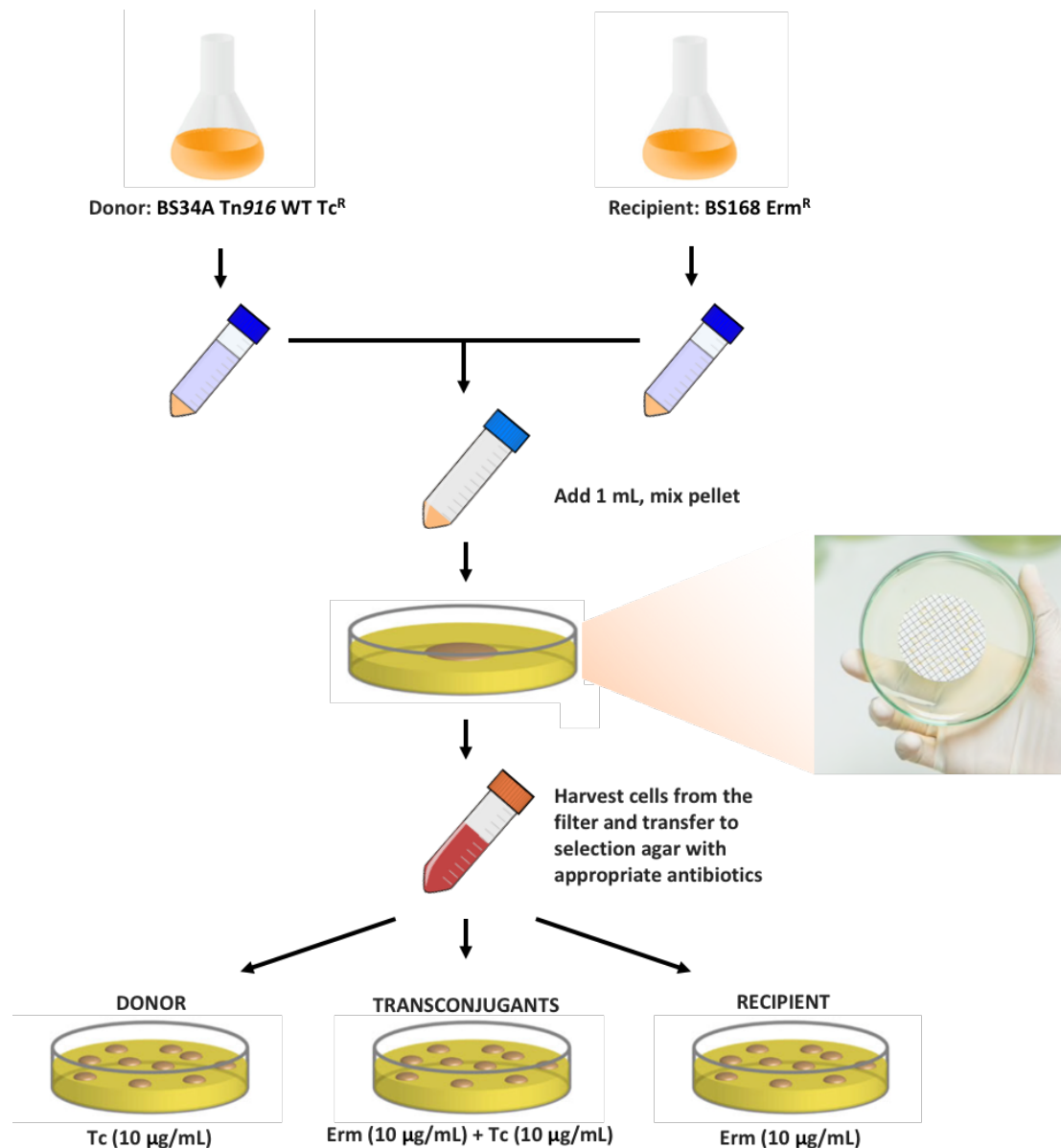
Figure 4-6 Schematic overview of filter mating experiment.

The donor and recipient strains were grown overnight in BHI broth supplemented with appropriate antibiotics. The cultures were then harvested by centrifugation and supernatant discarded. The pellets were resuspended in 1 mL of fresh broth and both donor and recipient cells were mixed and spread on 0.45 μm pore size sterilized nitrocellulose filter which had been previously been placed on antibiotic free BHI agar. Plates were incubated overnight and cells from the filter were harvested and spread over agar plates containing the appropriate antibiotics to select for transconjugants, donor and recipient cells. Then, diagnostic PCR was carried out on the putative transconjugants.

4.2.10 PCR analysis of the transconjugants

Transconjugants obtained from the filter mating experiment were selected based on their phenotypic resistance profile by testing their capability of growth on transconjugants-specific agar plates supplemented with tetracycline and other appropriate antibiotics that can differentiate them from donors or recipients. The presence of Tn*916* or Tn*916*ΔTerm in the transconjugants were determined by the amplification of *intTn* using a primer pair of CTn1670F and IntR. The detection of *tet*(M) was done by using primers tetM-1 and tetM-2. Amplification of other Tn*916*-derived sequences; joint-ends region of Tn*916* was carried out using two sets of primer pair; End Tn916-F/HR_R primers and 916 REO/HR_R primers. Amplification of the joint-ends region was carried out specifically to detect the presence or absence of the terminator within Tn*916* and Tn*916*ΔTerm. The amplification of right end regions of Tn*916* was carried out using a primer pair 916 REO and ETS_F primers. The amplification of the left end regions of Tn*916* and Tn*916*ΔTerm was carried out using a primer pair catPL_F and HR_R. The PCR mixture and condition were carried out as described in section 2.6.

4.3 Results

4.3.1 Generation of the mutant cassette

To investigate the biological importance of the terminator in preventing the transcription of conjugation genes in the Tn*916* conjugation module, *B. subtilis* BS34A mutant with a deletion of the Tn*916* terminator was generated via homologous recombination. The mutant cassette was successfully constructed by splicing four different DNA fragments (Appendix III). The UPS and *cat*P fragments were spliced together using an overlapping phusion primers generating amplicon with an expected size of 2042 bp [UPS+*cat*P] (Figure 4-7). The *cat*P was fused in an opposite transcriptional direction of the ORFs in the conjugation module of the Tn*916*. As *cat*P contains its own promoter, fusing it in the opposite direction will prevent the transcriptional read through from the *cat*P promoter into the Tn*916* (Figure 4-7(A))*.*

The DS1 and DS2 fragments were spliced together in order to delete the 32 bp terminator structure within the Tn*916* conjugation module. The spliced product [DS1+DS2] was obtained at the expected size of 1239 bp (Figure 4-7). The ligation of two independently spliced fragments; [UPS+*cat*P] and [DS1+DS2] generates four different types of ligation products as each fragment contain *xho*I sticky ends. The four combinations of ligation products with their expected size are shown in Figure 4-7. The ligation product of interest; [UPS+*cat*P] + [DS1+DS2] with the size of 3281 bp (consists of all four fragments) was chosen and extracted from the agarose gel. The purified mutant cassette [UPS+*cat*P+DS1+DS2] was cloned pGEM-T Easy vector.
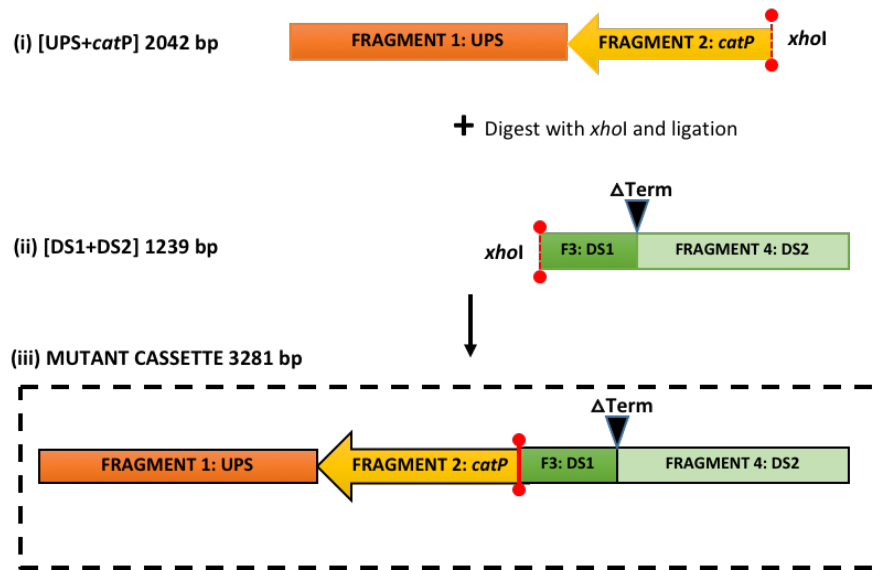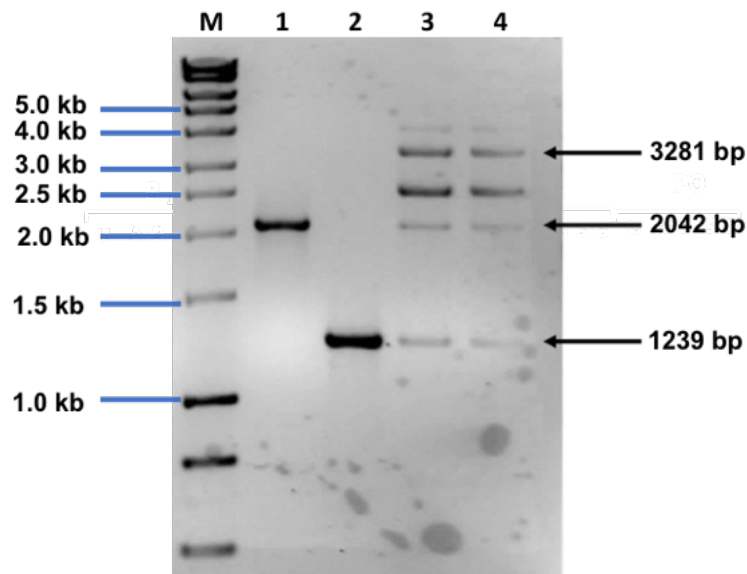
Figure 4-7 Agarose gel electrophoresis of the [UPS+catP] and [DS1+DS2] amplicons and their ligation products.

Panel A: Schematic diagram showing the ligation of [UPS+*cat*P] and [DS1+DS2]. The structure (⦙) represents the sticky ends with *Xho*I restriction site. Panel B: Lane M; HyperLadder™ 1kb, Lane 1: UPS + *cat*P [F1+F2] with the size of 2042 bp, Lane 2: DS1 +DS2 [F3 +F4] with the size of 1239 bp, Lane 3 and 4; ligation products of three different combinations: [F1+F2] [F1+F2] = 4084 bp, [F3+F4] [F3+F4] = 2478 bp, [F1+F2][F3+F4] = 3281 bp, [F1+F2] only = 2042 bp, [F3+F4] only = 1239 bp.

From the selective plates, four positive clones harbouring the pGEM-T/Tn*916*ΔTerm were selected. The *E. coli*::pGEM-T/Tn*916*ΔTerm isolates were then subjected to plasmid purification and digested with *EcoR*I and *Xho*I. Digestion products of the positive clones produced three fragments at expected sizes as shown below (Figure 4-8). Sequencing result showed that the mutant cassette was successfully cloned into the pGEM-T Easy vector (Appendix III).

[A]   [B]
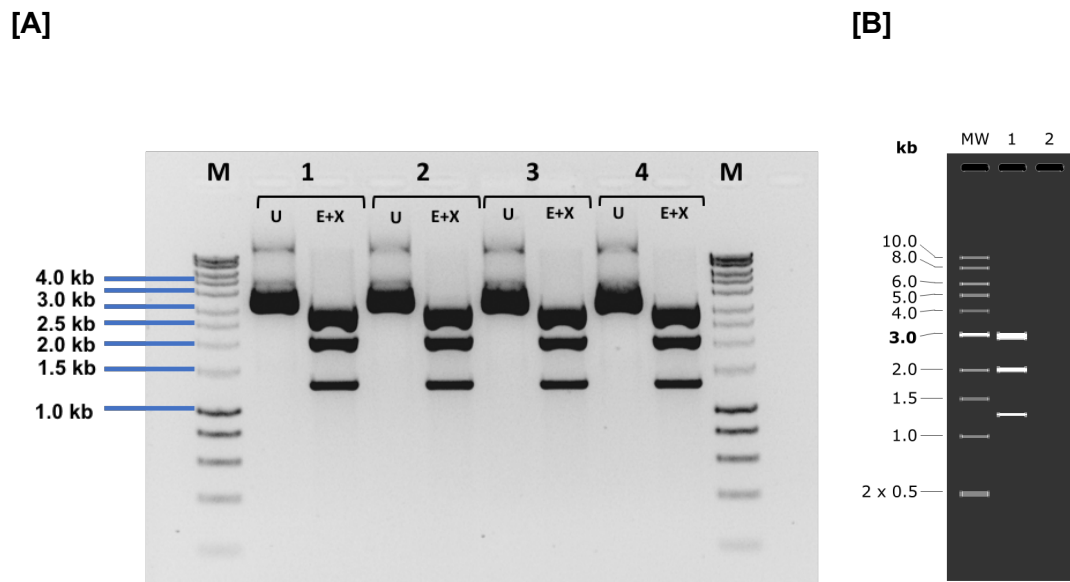
Figure 4-8 Agarose gel electrophoresis of the extracted pGEM-T/Tn*916*ΔTerm and digestion products.

Panel A: Lane M; HyperLadder™ 1kb, Lane 1-4: U; Undigested pGEM-T/Tn*916*ΔTerm clone 1-4, E+X; *EcoR*I and *Xho*I digestion of pGEM-T/Tn*916*ΔTerm clone 1-4. Panel B: *EcoR*I and *Xho*I digestion of pGEM-T/Tn*916*ΔTerm simulated by SnapGene software.

4.3.2 Generation of *B. subtilis* BS34A::Tn*916*ΔTerm

Both linearised and non-linearised pGEM-T/Tn*916*ΔTerm were transformed into *B. subtilis* BS34A and selected on BHI supplemented with chloramphenicol agar plates. Five clones were obtained from BS34A transformed with non-linearised pGEM-T/Tn*916*ΔTerm but none with the linearised pGEM-T/Tn*916*ΔTerm. These five clones were designated as HR_C1, HR_C2, HR_C3, HR_C7 and HR_C8. All clones were subjected to diagnostic PCR for the validation of mutant cassette integration into the BS34A chromosome. The amplification region of R1, R2 and R3 are shown in Figure 4-9(A). DNA bands with expected sizes were observed for R2 (1988 bp) and R3 (1644 bp) amplicons (Figure 4-9(B) and (C)). However, there are no DNA band observed for R1 amplicon on all samples.

Sequencing result of R3 amplicon showed deletion of the Tn*916* terminator in three out of five clones; HR_C1, HR_C3 and HR_C8 (Figure 4-10). For R2 amplicon, only a partial sequencing read (from catPL_R) was obtained. As the the BS34A transformation was done with the non-linearised pGEM-T/Tn*916*ΔTerm, co-integration of the pGEM-T backbone together with the mutant cassette might have occurred. This could be the reason why R1 region could not be amplified and the R2 amplicon could only be partially sequenced on single direction (Figure 4-11). It is predicted that a single crossover might have occurred within the DS2 region resulting in mutant with the targeted deletion of the Tn*916* terminator, but with the integration of the rest of the mutant cassette plus pGEM-T backbone followed by the original copy of Tn*916* left end region that contains the terminator (Figure 4-11).

[A]
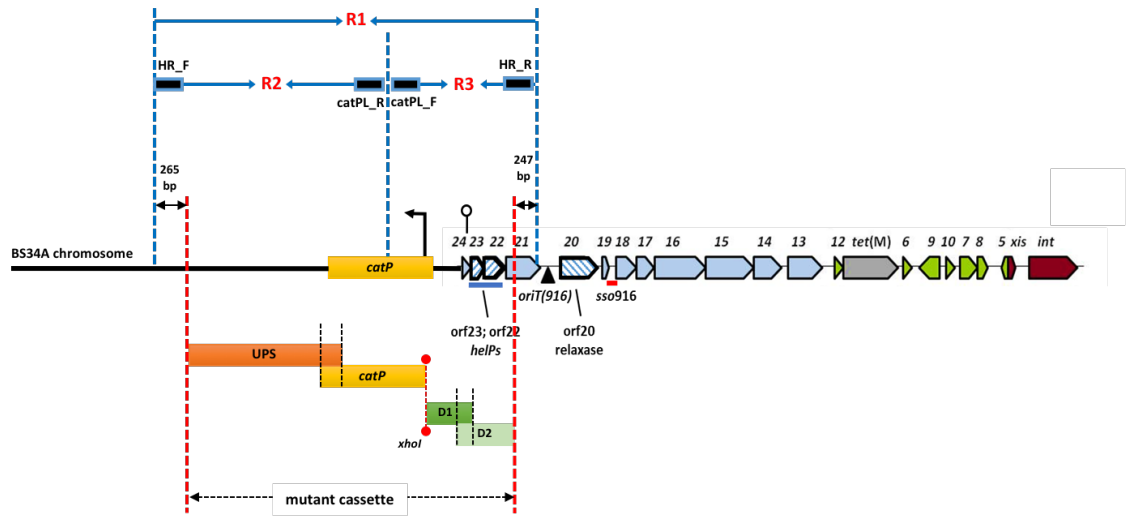
[B]
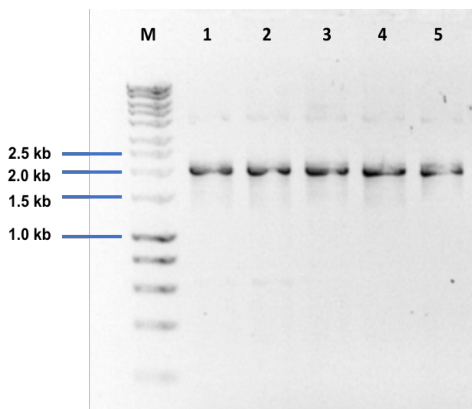
[C]
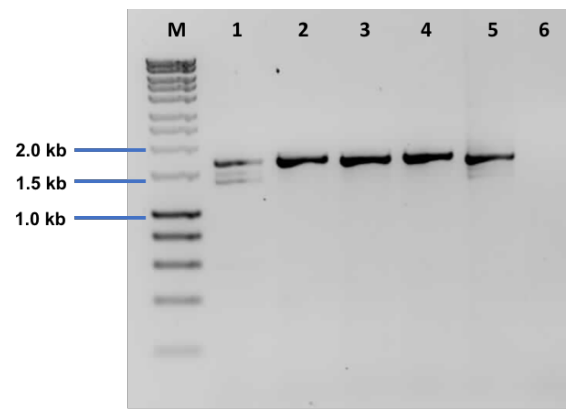
Figure 4-9 Gel electrophoresis of R2 and R3 amplicons.

Panel A: Schematic diagram of the integrated Tn*916*ΔTerm showing region R1, R2 and R3. Arrows represent the primers and direction of priming. Amplicon of primer pair HR_F/HR_R is expected to be 3793 bp in size, HR_F/catPL_R is 1988 bp in size and catPL_F/HR_R is 1644 bp in size. Panel B: The PCR products of R2 amplifications. Lane M: HyperLadder™ 1kb; Lane 1-5: R2 amplicons from the genomic DNA of *B. subtilis::*Tn*916*ΔTerm of five isolates. Panel C: The PCR products of R3 amplifications. Lane M: HyperLadder™ 1kb; Lane 1-5: R3 amplicons from the genomic DNA of *B. subtilis::*Tn*916*ΔTerm of five isolates; Lane 6: R3 amplicon from genomic DNA of *B. subtilis* BS34A (negative control).

```
HR_C1   ACACTTGCCGAAAAAGAAAAACTGCCGGGTCTCGAGGTCTGAGGATTAATGGCTGTGTTA    112
HR_C2   ACACTTGCCGAAAAAGAAAAACTGCCGGGTCTCGAGGTCTGAGGATTAATGGCTGTGTTA    107
HR_C3   ACACTTGCCGAAAAAGAAAAACTGCCGGGTCTCGAGGTCTGAGGATTAATGGCTGTGTTA    119
HR_C7   ACACTTGCCGAAAAAGAAAAACTGCCGGGTCTCGAGGTCTGAGGATTAATGGCTGTGTTA    118
HR_C8   ACACTTGCCGAAAAAGAAAAACTGCCGGGTCTCGAGGTCTGAGGATTAATGGCTGTGTTA    120
        ************************************************************

HR_C1   AACACTATGATTTTTCCTTCAAACTTATTTTCTAAGAAAAATAGCATAAAAATCTAGTTA    172
HR_C2   AACACTATGATTTTTCCTTCAAACTTATTTTCTAAGAAAAATAGCATAAAAATCTAGTTA    167
HR_C3   AACACTATGATTTTTCCTTCAAACTTATTTTCTAAGAAAAATAGCATAAAAATCTAGTTA    179
HR_C7   AACACTATGATTTTTCCTTCAAACTTATTTTCTAAGAAAAATAGCATAAAAATCTAGTTA    178
HR_C8   AACACTATGATTTTTCCTTCAAACTTATTTTCTAAGAAAAATAGCATAAAAATCTAGTTA    180
        ************************************************************

HR_C1   TCCGCATAAAAACTGGACTTATCACACTTTATCAAGGTCAAAACCACTCAATTTACTACT    232
HR_C2   TCCGCATAAAAACTGGACTTATCACACTTTATCAAGGTCAAAACCACTCAATTTACTACT    227
HR_C3   TCCGCATAAAAACTGGACTTATCACACTTTATCAAGGTCAAAACCACTCAATTTACTACT    239
HR_C7   TCCGCATAAAAACTGGACTTATCACACTTTATCAAGGTCAAAACCACTCAATTTACTACT    238
HR_C8   TCCGCATAAAAACTGGACTTATCACACTTTATCAAGGTCAAAACCACTCAATTTACTACT    240
        ************************************************************

HR_C1   AATTTACTACTTATGAATGAGCTTTGATACGACGATTTATCCTTGAAAAGTGAAGATATA    292
HR_C2   AATTTACTACTTATGAATGAGCTTTGATACGACGATTTATCCTTGAAAAGTGAAGATATA    287
HR_C3   AATTTACTACTTATGAATGAGCTTTGATACGACGATTTATCCTTGAAAAGTGAAGATATA    299
HR_C7   AATTTACTACTTATGAATGAGCTTTGATACGACGATTTATCCTTGAAAAGTGAAGATATA    298
HR_C8   AATTTACTACTTATGAATGAGCTTTGATACGACGATTTATCCTTGAAAAGTGAAGATATA    300
        ************************************************************

HR_C1   AAGATACTTCCAATAAAATTTGAATATTTAATAGGTA-----------------------    329
HR_C2   AAGATACTTCCAATAAAATTTGAATATTTAATAGGTAGACACTTCAAAAAATGAGGTGTC    347
HR_C3   AAGATACTTCCAATAAAATTTGAATATTTAATAGGTA-----------------------    336
HR_C7   AAGATACTTCCAATAAAATTTGAATATTTAATAGGTAGACACTTCAAAAAATGAGGTGTC    358
HR_C8   AAGATACTTCCAATAAAATTTGAATATTTAATAGGTA-----------------------    337
        *************************************

HR_C1   ---------ACCCGATTTTGAAAGGAAGTGAACTTATGAAAACAAAAAATCAAGAATCAA    380
HR_C2   TATTTTTTTACCCGATTTTGAAAGGAAGTGAACTTATGAAAACAAAAAATCAAGAATCAA    407
HR_C3   ---------ACCCGATTTTGAAAGGAAGTGAACTTATGAAAACAAAAAATCAAGAATCAA    387
HR_C7   TATTTTTTTACCCGATTTTGAAAGGAAGTGAACTTATGAAAACAAAAAATCAAGAATCAA    418
HR_C8   ---------ACCCGATTTTGAAAGGAAGTGAACTTATGAAAACAAAAAATCAAGAATCAA    388
        **************************************************
```

Figure 4-10 Sequence alignment of the R3 amplicons amplified from BS34A::pGEM-T/Tn*916*ΔTerm clones; HR_C1, HR_C2, HR_C3, HR_C7 and HR_C8 (shown in partial sequence).

The alignment shows deletion of the 32 bp of terminator sequence in clone HR_C1, HR_C3 and HR_C8. In contrast, terminator sequence is detected in clone HR_C2 and HR_C7. Sequences highlighted in yellow; partial *cat*P sequence, sequences in green; Fragment 3:DS1 of the mutant cassette, sequences highlighted in blue; 32 bp of terminator sequence, sequences highlighted in neon green; Fragment 4:DS2. An * (asterisk) indicates positions which have a single, fully conserved residue.
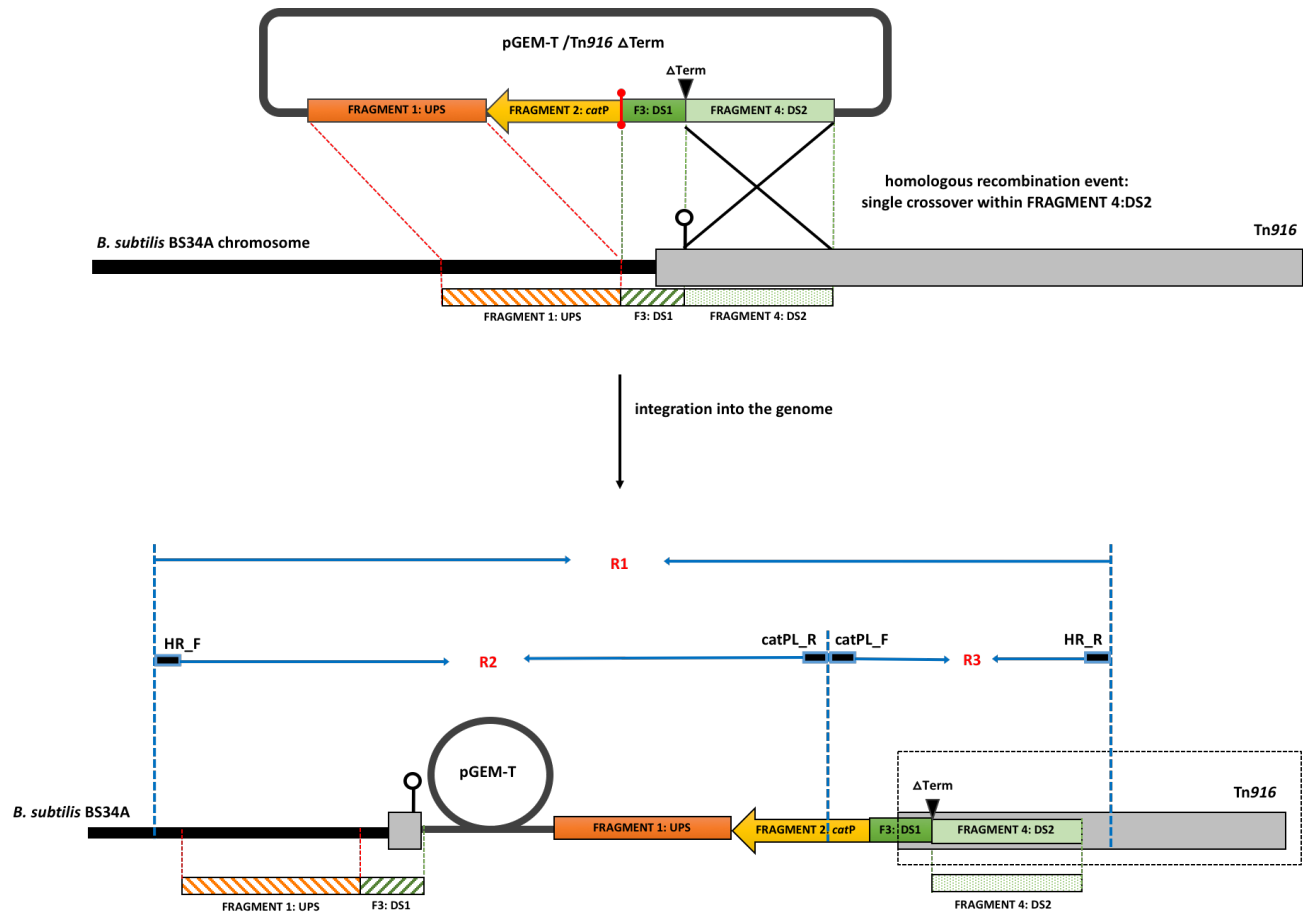
Figure 4-11 The predicted homologous recombination event showing the co-integration of the circular pGEM-T/Tn*916*ΔTerm followed by the original left end of Tn*916* that carry the terminator. The mutant was denoted as *B. subtilis* BS34A Tn*916*ΔTerm.

4.3.3 Diagnostic PCR of the *B. subtilis* BS34A Tn*916*ΔTerm

Amplification of the Tn*916* joint-ends of the circular intermediate (CI) and empty target site in BS34A Tn*916*ΔTerm was done. The results showed that both amplicons at the expected size were amplified (Figure 4-12). However, as the sequencing results of the joint-ends region of Tn*916* were poor, the deletion of the terminator sequence was not able to be validated in the excised CI of Tn*916*ΔTerm.

The amplification of the right end genome junction of the Tn*916* was carried out (based on the original position of Tn*916* in *B. subtilis* BS34A) to check if the Tn*916*ΔTerm remain integrated at the same site after the deletion of the terminator. PCR products at an expected size of 580 bp were obtained (Figure 4-13). The right end amplicons were validated by sequencing. This result suggest that the Tn*916*ΔTerm remain intact at the original position.
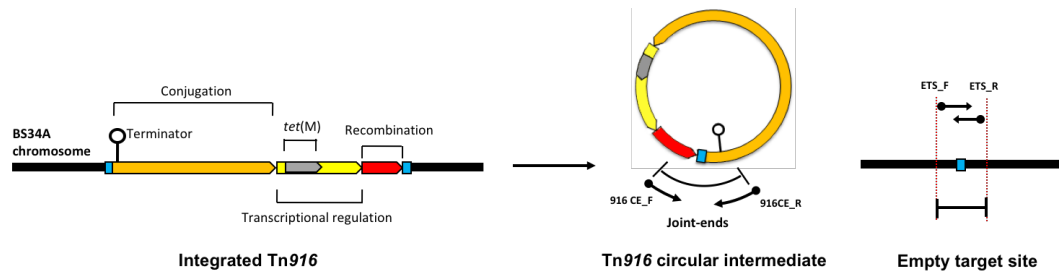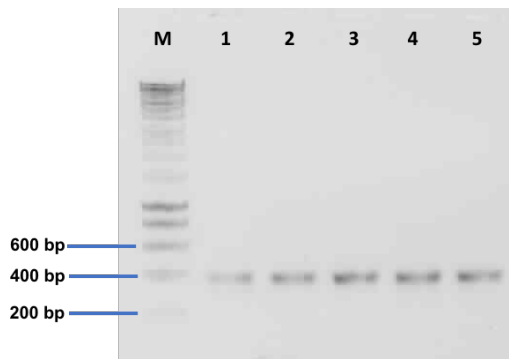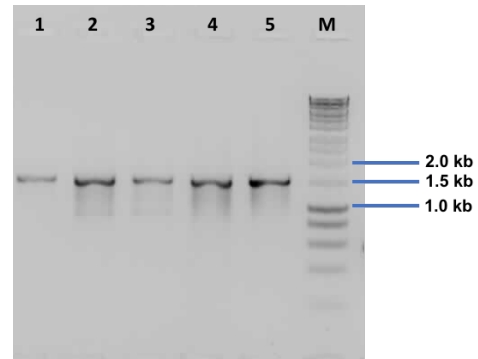
Figure 4-12 Amplification of the joint-ends of Tn*916* circular intermediate (Panel B) and empty target site (Panel C) in BS34A Tn*916*ΔTerm.

Panel A: Schematic diagram of the Tn*916* conjugative transposon showing the amplification region of Tn*916* joint-ends. Excision of Tn*916* from the chromosome produces the circular intermediates of Tn*916* containing the joint-ends (blue) and empty target site. The empty target site of Tn*916* was amplified using ETS_F and ETS_R. The Tn*916* joint-ends region of CI was amplified using primers 916CE_F and 916CE_R. Arrows represent the primers and direction of priming. **Panel B:** Gel electrophoresis of Tn*916* empty target site amplicons. Lane M: HyperLadder™ 1kb; Lane 1-5: empty target site amplicons with the expected size of 383 bp. **Panel C:** Gel electrophoresis of Tn*916* joint-ends region amplicons. Lane M: HyperLadder™ 1kb; Lane 1-5: joint-ends amplicons with the expected size of 1435 bp.

[A]
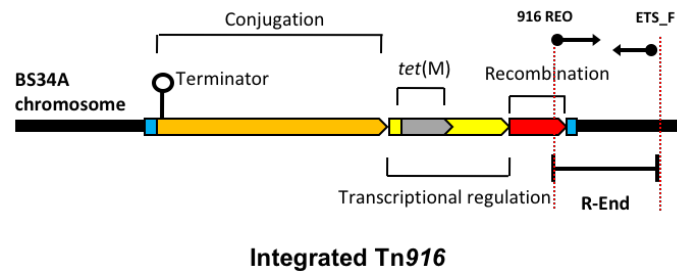
[B]
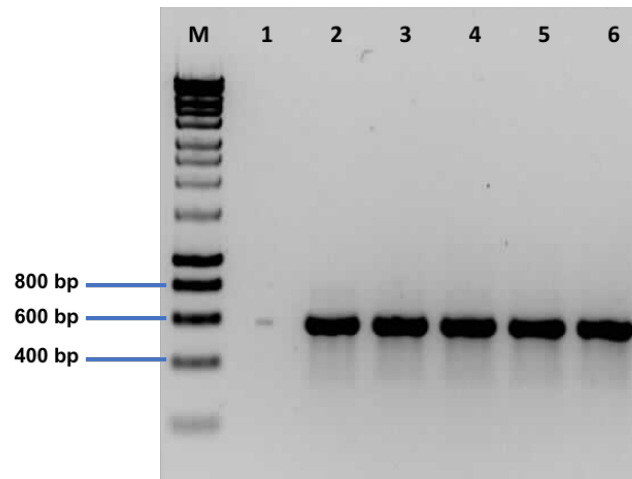
Figure 4-13 Gel electrophoresis of the right end amplicons of integrated Tn*916*ΔTerm.

Panel A: Schematic diagram of Tn*916* conjugative transposon showing the amplification region of Tn*916* right end using 916 REO (located within the Tn*916*) and ETS_F (located within the BS34A chromosome) primers. Arrows represent the primers and direction of priming. **Panel B:** Lane M: HyperLadder™ 1kb; Lane 1: negative control. Lane 2-6: right end amplicons with the expected size of 580 bp.

4.3.4 Genomic sequence analysis of *B. subtilis* mutant strain BS34A Tn*916*ΔTerm

A *de novo* genome assembly of the BS34A Tn*916*ΔTerm Illumina sequence data was obtained from MicrobesNG, UK. The assembly was performed using SPAdes, resulting in a total of 70 contigs, with 14 contigs larger than 1,000 bp. The draft genome of BS34A Tn*916*ΔTerm is 4 196 838 bp, with 56.85 mean coverage, 43.40% G+C content, encoding 4 232 predicted coding sequences (CDS), 11 rRNAs, 84 tRNAs, 1 tmRNA and 60 miscellaneous RNAs. Taxonomic distribution showed the sequence reads map to the family Bacillaceae at 93.64% and genus Bacillus at 93.61%.

The 18 032 bp sequence of Tn*916* element from *Enterococcus faecalis* DS16 (GenBank U09422.1) was located in CONTIG 3 (9 18961 bp). However, the Tn*916* sequence detected within CONTIG 3 is truncated, by 210 bp of the Tn*916* left end sequences. We failed to validate the deletion of the 32 bp terminator sequence within the BS34A Tn*916*ΔTerm as it is located within this missing region (Figure 4-14). Another two contigs, denoted as CONTIG 44 (286 bp) and CONTIG 47 (254 bp) were also observed to match the left end of Tn*916* around the targeted deletion region. For CONTIG 44 (286 bp); the presence of 32 bp terminator sequence are detected, flanked by 127 bp of sequence (which are homologous to F3:DS1, of the mutant cassette) and another 127 bp sequence (which are homologous to F4:DS2) (Figure 4-15). The F3:DS1 and F4:DS2 are the two fragments spliced together to delete the terminator sequence in the generated mutant cassette (Figure 4-15). For CONTIG 47 (254 bp); the presence of 32 bp terminator sequence are not

detected. Therefore 127 bp sequences of F3:DS1 and another 127 bp sequence of F4:DS2 region are aligned next to each other (Figure 4-15).

Further analysis of the truncated Tn*916* element within CONTIG 3 (derived from BS34A Tn*916*ΔTerm) in comparison to the *Enterococcus* Tn*916* element (Genbank U09422.1) showed three single nucleotide variations. These variations are the same as the BS34A Tn*916* element reported by Browne *et al.* (2015). The three SNPs observed within the Tn*916* element of BS34A Tn*916*ΔTerm are; i) a substitution mutation of Guanine (G) to Thymine (T) within ORF9 resulting in a conversion of Glysine to Lysine, ii) a single nucleotide insertion (Cytosine) within ORF12, resulting in a frameshift mutation of the protein with 725 amino acids. This insertion also caused a formation of a recognition site for restriction enzyme *Sty*I, iii) a single nucleotide deletion (Guanine) in the *oriT* region between ORF21 and ORF20. The mutation does not affect the nick site (TGGTGTGG) (Figure 4.14).
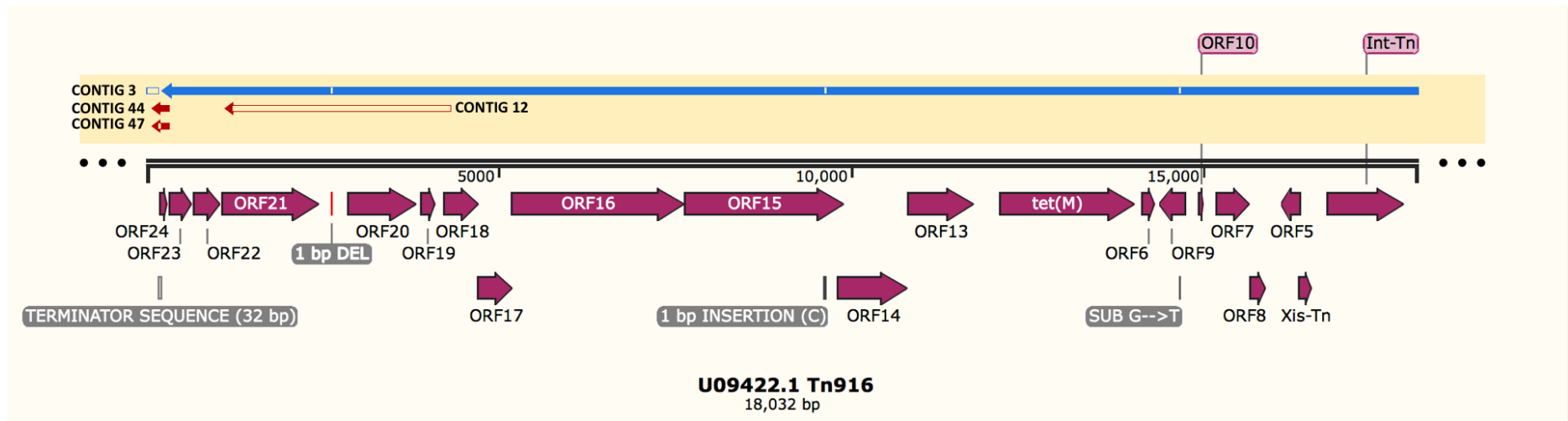
Contig	CONTIG 3	CONTIG 12	CONTIG 44	CONTIG 47
Length	918 961 bp	3 271 bp	286 bp	254 bp
Coverage	23.0214	1004.46	270.667	693.929

Figure 4-14 Alignment of the Tn*916* element from *Enterococcus faecalis* DS16 (GenBank U09422.1) with various BS34A Tn*916*ΔTerm whole genome sequence contigs.

The alignment resulted in various truncated matches with four different contigs; CONTIG 3 (blue fragment), CONTIG 12, CONTIG 44 and CONTIG 47 (shown in maroon fragments). Filled box represent the aligned sequence, non-filled box represents unaligned sequence within the contigs. The burgundy block arrows underneath the double black lines represent the ORFs within the Tn*916* element, including *tet*(M), *xis*Tn and *int*Tn. The position of 32 bp terminator sequence, 1 bp deletion in between ORF21 and ORF20, 1 bp insertion within ORF15 and nucleotide substitution (G → T) is labelled in grey box.

171
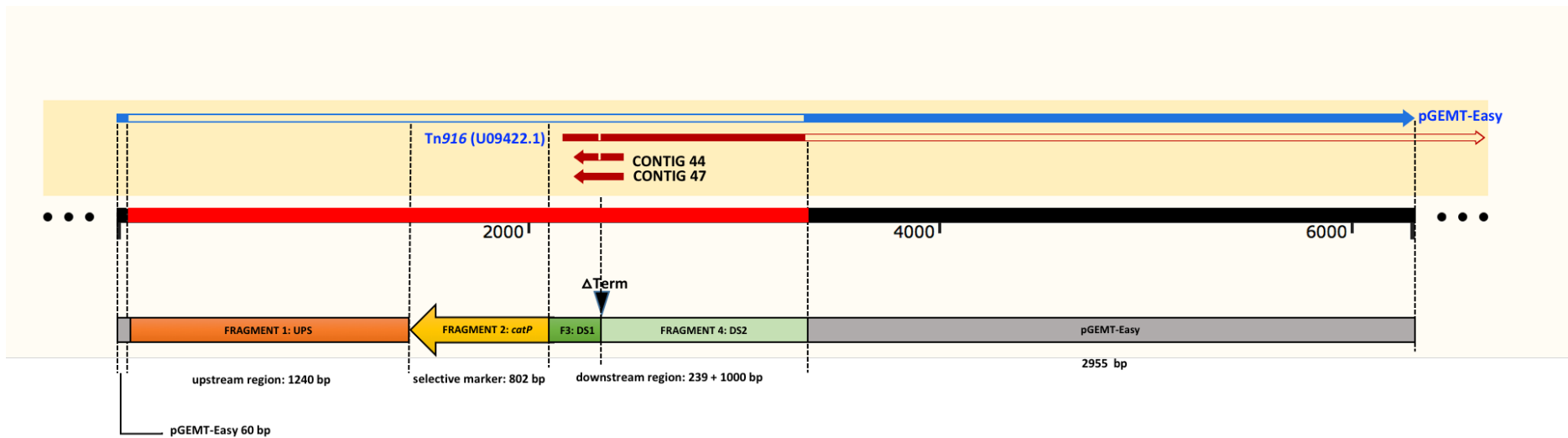
Figure 4-15 Alignment of the pGEM-T/Tn*916*ΔTerm mutant cassette with Tn*916* (GenBank U09422.1), CONTIG 44 and CONTIG 47 of BS34A Tn*916*ΔTerm whole genome sequence.

Filled box represent the aligned sequence, non-filled box represents unaligned sequence within the contigs. The red fragment represents the Tn*916*ΔTerm mutant cassette consist of; fragment 1:UPS (orange), fragment 2: *cat*P in (yellow), fragment 3: DS1 (dark green) and fragment 4: DS2 (light green). The structure (▼) represents the deleted terminator region labelled as ΔTerm. The black fragment represents the pGEM-T Easy plasmid sequences.

The 6296 bp sequence of pGEM-T/Tn*916*ΔTerm mutant cassette was aligned to all obtained contigs and detected in CONTIG 12, CONTIG 15 and CONTIG 16. The CONTIG 12 sequence were identical with the reference pGEM-T Easy sequence (3015 bp) but also contains a partial 69 bp of F1:UPS sequence and a partial 127 bp of F4: DS2 (Figure 4-16). CONTIG 15 is fully aligned with 1240 bp of F1: UPS (which is a homologous region in the BS34A chromosome). For CONTIG 16 (1170 bp), the presence of *catP* are detected flanked by F3:DS1 and partial F1:UPS sequence. The CONTIG 16 is truncated at the connection between F3:DS1 and F4:DS2, thus deletion of the terminator sequence failed be detected (Figure 4-16).

High coverage of reads usually indicates that the sequence reads exist in multiple copies and might be of plasmid origin (Antipov *et al.*, 2016, Roosaare *et al.*, 2018), or in this case in circular intermediate form. As high coverage value is observed in CONTIG 12, 15 and 16, it suggests that it may exist in a circular form (Figure 4-16). However, this does not exclude the possibility that at least some part of the pGEM-T/Tn*916*ΔTerm has inserted into the chromosome or the probability that it exists in both forms. The coverage value of CONTIG 44 is low and detected without the 32 bp terminator sequence (ΔTerm) suggesting that at least this part of the mutant cassette has inserted into the chromosome via homologous recombination. However, the existence of a similar contig designated CONTIG 47, but with 32 bp terminator sequence, suggest that it is likely that the pGEM-T/Tn*916*ΔTerm was detected in both forms. Due to single recombination event, there might be two copies of Tn*916* left ends (Fragment F3:DS1); one with the terminator and one without the terminator as depicted in Figure 4-11. This resulted in some

contigs failing to be assembled due to mismatch to the reference sequence and therefore were discarded and placed at the end of the genome sequence instead.
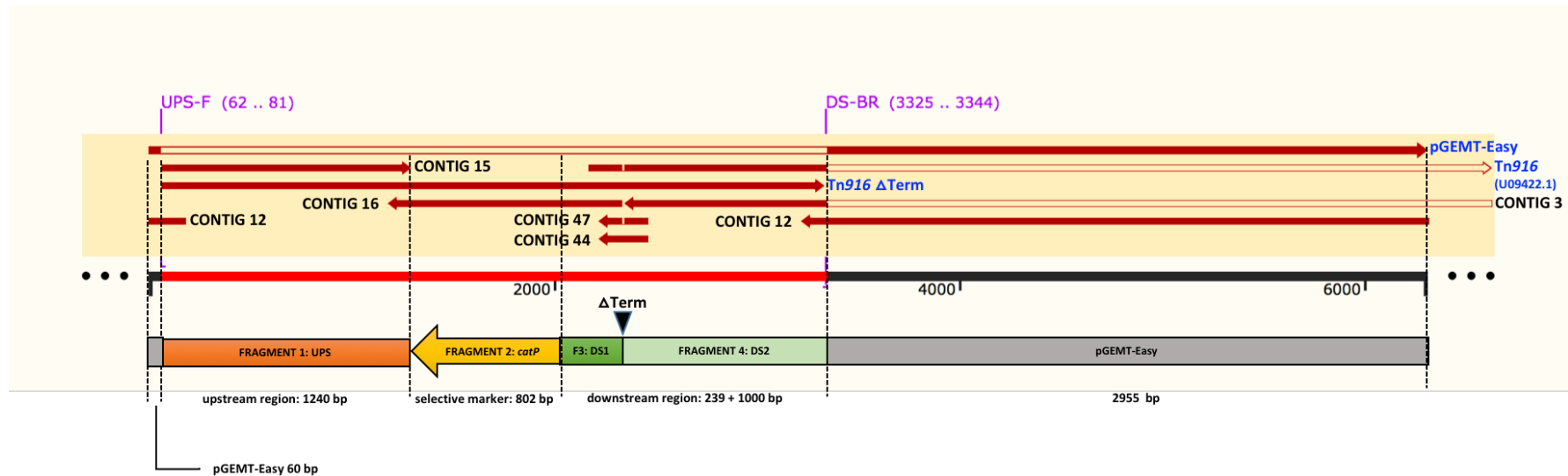
Contig	CONTIG 3	CONTIG 12	CONTIG 15	CONTIG 16	CONTIG 44	CONTIG 47
Length	918 961 bp	3 271 bp	1 240 bp	1 170 bp	286 bp	254 bp
Coverage	23.0214	1004.46	1091.55	824.708	270.667	693.929

Figure 4-16 Alignment of the pGEM-T/Tn*916*ΔTerm mutant cassette with various BS34A Tn*916*ΔTerm whole genome sequence contigs.

The alignment resulted in various truncated matches with six different contigs; CONTIG 3, CONTIG 12, CONTIG 15, CONTIG 16, CONTIG 44 and CONTIG 47 (shown in maroon fragments). Filled box represent the aligned sequence, non-filled box represents unaligned sequence within the contigs. The red fragment represents the Tn*916*ΔTerm mutant cassette consist of; fragment 1:UPS (orange), fragment 2: *cat*P in (yellow), fragment 3: DS1 (dark green) and fragment 4: DS2 (light green). The structure (▼) represents the deleted terminator region labelled as ΔTerm. The black fragment represents the pGEM-T Easy plasmid sequences.

4.3.5 Genomic sequence analysis of *B. subtilis* mutant strain CU2189 Rif[R] Nal[R] and BS168 Erm[R]

B. subtilis strain CU2189 isolates that are resistant to rifampicin and nalidixic acid were selected and denoted as CU2189 Rif[R] Nal[R]. Growth were observed in BHI plates supplemented with rifampicin at 25 μg/mL and nalidixic acid at 10 μg/mL. These isolates were verified by 16S rRNA and *gyrA* sequencing, showing 100% identity with various strains of *B. subtilis* including Bacillus sp. genome assembly BS34ACh (Accession no: LN680001.1). *B. subtilis* BS34A is the CU2189 strain containing a single copy of Tn*916* (Roberts *et al.*, 2003). The genomic sequence alignments of CU2189 Rif[R] Nal[R] and *B. subtilis* BS34A showed there is a substitution mutation of Thymine (T) to Guanine (G) in the *gyrA* of CU2189 Rif[R] Nal[R] (Table 4-2). Point mutation that caused single amino acid changes in DNA gyrase subunit A has been reported to confer resistance towards quinolones including nalidixic acid (Price *et al.*, 2003, Morgan-Linnell *et al.*, 2009, Aldred *et al.*, 2014).

Another substitution mutation within the of CU2189 Rif[R] Nal[R] strain was observed in the *rpoB* which encodes the β subunit of RNA polymerase (RNAP). Mutations within the *rpoB* have been reported to confer resistance to rifampicin as a result of decreased affinity of rifampicin to its binding site (Xu *et al.*, 2005). Further analysis on CU2189 Rif[R] Nal[R] showed that the identified mutation occurred in the Cluster I region of *rpoB* (Goldstein, 2014).

B. subtilis BS168 isolates that are resistant to erythromycin was successfully selected and denoted as BS168 Erm[R]. Growth was observed on LB agar

supplemented with erythromycin at a concentration of 10 µg/mL. The average value of cfu/mL of the total cells plated is at 2.84 x 10^{10}. The isolates were verified by 16S rRNA and *gyrA* sequencing, showing 100% identity with various strain of *B. subtilis* including *B. subtilis* strain 168 16S ribosomal RNA complete sequence (Accession no: NR_102783.2). The genomic sequence alignments of BS168 Erm^R and the wild type strain BS168 showed there is a 54 bp duplication in the *rplV* conferring erythromycin resistance. Another substitution mutation of Cytosine (C) to Thymine (T) in a hypothetical protein was also observed (Table 4-3). Further analysis of the BS168 Erm^R strain was carried out and is reported in Chapter 5. Both *B. subtilis* mutant strain CU2189 Rif^R Nal^R and BS168 Erm^R were used as a recipient in the filter mating experiments

Table 4-2 Breseq output for genome alignments of CU2189 Rif^R Nal^R and BS34A (CU2189::Tn*916*)

Strain	Position in chromosome	mutation	annotation	gene	description
CU2189 Rif^R Nal^R	7,243	T → G	substitution S84A* (**T**CA-**G**CA)	*gyrA* →	DNA gyrase subunit A
	123,374	C → T	substitution S487L* (T**C**A-T**T**A)	*rpoB* →	DNA-directed RNA polymerase subunit beta
	1,886,555	Δ18,038 bp		*xerC_2* – BS34A_19440	This deletion is an artefact because BS34A is a derivative of CU2189 which contains Tn*916* conjugative transposon (17 genes)

Table 4-3 Breseq output for genome alignment of BS168 Erm^R and BS168

Strain	Position in chromosome	mutation	annotation	gene	description
BS168 Erm^R	523,669	C → T	substitution W261* (T**G**G-T**A**G)	JAMKKJHE_01890 ←	Hypothetical protein
	22,267	54 bp x 2	duplication (ACCAATAAAGAGTTCGTAA GCAACTGCTTCCGGGATGC AATTTTTCTAAGGCGG)$_{1 \to 2}$	*rplV* ←	50S ribosomal protein L22

An → (arrow) to the right indicates a substitution of nucleotide (substituted nucleotide is bold and underlined in red); a symbol (x 2) and ($_{1 \to 2}$) indicates a tandem duplication of nucleotides; an ← (arrow) to the left indicates the orientation of the gene.

4.3.6 Transfer of Tn*916* WT and Tn*916*ΔTerm from *B. subtilis* BS34A and *B. subtilis* Tn*916*ΔTerm to *B. subtilis* CU2189 Rif^R Nal^R

To investigate the ability of the mutant *B. subtilis* Tn*916*ΔTerm to transfer Tn*916*ΔTerm via conjugation, filter mating experiment was carried out using *B. subtilis* CU2189 Rif^R Nal^R as a recipient. The experiment was carried out simultaneously with the transfer of the wild type Tn*916* from BS34A to CU2189 Rif^R Nal^R. However, during the experiment, it was observed that the donor cells; *B. subtilis* BS34ATn*916*ΔTerm grows numerously on control plates (BHI agar supplemented with rifampicin, nalidixic acid and tetracycline), indicating that they have likely undergone spontaneous mutation to be resistant to rifampicin and nalidixic acid. Therefore, this experiment is halted as it was deemed that the recipient is not suitable.

Nevertheless, diagnostic PCR were carried out to see if we could distinguish between the donor and the putative transconjugants obtained. Four putative transconjugants of CU2189 Rif^R Nal^R::Tn*916* ΔTerm (F2ΔC1, F3ΔC3, F3ΔC9 and F5ΔC14) and three putative transconjugants of CU2189 Rif^R Nal^R::Tn*916* (WTC5, WTC6 and WTC8) were chosen for further verification. For the mating pair of BS34A Tn*916* WT and CU2189 Rif^R Nal^R, amplification of the R3 region in the donor, recipient and putative transconjugants (WTC5, WTC6 and WTC8) showed no visible DNA band. This is expected as the forward primer was designed to anneal within the *cat*P sequence and none of them carry the *cat*P. Interestingly, amplification result of the left end region of Tn*916* ΔTerm (R3) in all of the CU2189 Rif^R Nal^R::Tn*916*ΔTerm transconjugants (F2ΔC1, F3ΔC3, F3ΔC9 and F5ΔC14) showed amplicons at the expected size of 1635

bp. The same size of amplicon was observed for the donor cell; BS34A::Tn*916*

ΔTerm, although some other non-specific bands were also observed (Figure

4-17).

Sequencing result of these R3 amplicons showed that deletion of the

terminator was detected in all of the putative transconjugants; F2ΔC1, F3ΔC3,

F3ΔC9 and F5ΔC14. This result suggests that *cat*P has co-transferred with

the Tn*916*ΔTerm as depicted in Figure 4-17 (A) or as mentioned above, it may

be that the donor cells have undergone spontaneous mutation during the

selection process, generating the same phenotype. We failed to verified if the

putative transconjugants obtained are genuine and therefore, the filter mating

experiment was repeated using a different recipient, which is *B. subtilis* strain

BS168 Erm[R].
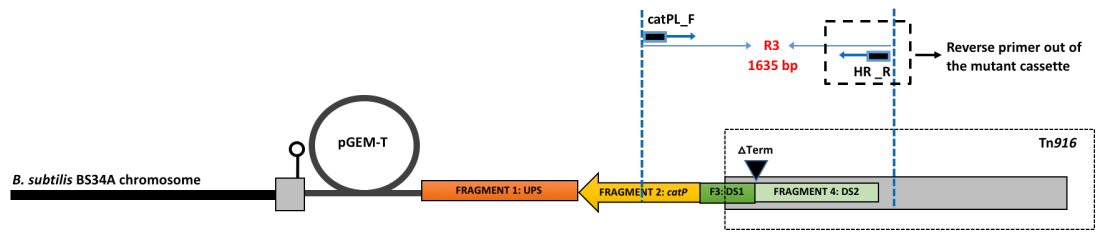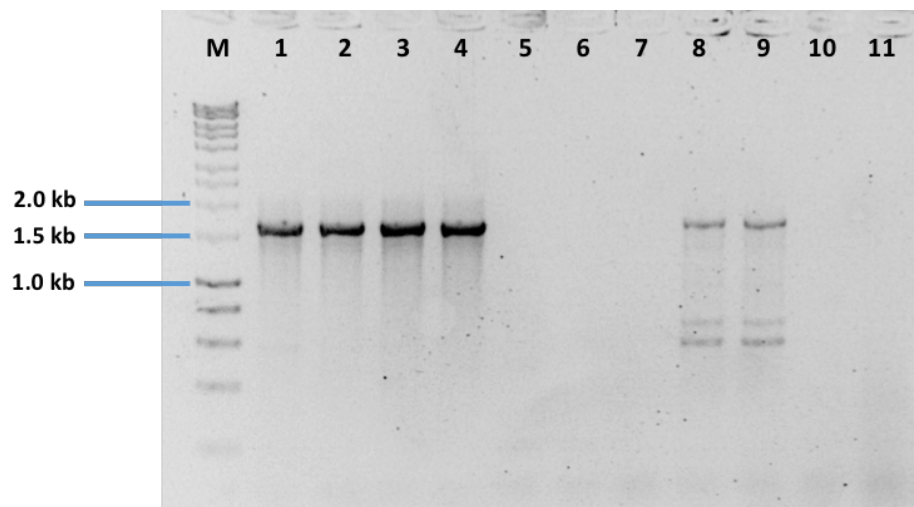
Figure 4-17 Amplification of the left end region (R3) of the Tn*916*ΔTerm and Tn*916* WT within the putative transconjugants.

Lane M: HyperLadder™ 1kb; Lane 1-4: left end amplicons from F2ΔC1, F3ΔC3, F3ΔC9 and F5ΔC14 putative transconjugants; Lane 5-7: left end amplicons from WTC5, WTC6 and WTC8 putative transconjugants; Lane 8-9: BS34A Tn*916*ΔTerm (positive control); Lane 10: BS34A Tn*916* (negative control); Lane 11: CU2189 Rif^R Nal^R (negative control).

4.3.7 Transfer of Tn*916* WT and Tn*916*ΔTerm from *B. subtilis* BS34A to *B. subtilis* BS168 Erm^R

Conjugal transfer of Tn*916* WT and Tn*916*ΔTerm from *B. subtilis* BS34A to *B. subtilis* BS168 erm^R did not generate any transconjugants. This might be due to the fact that the recipient cells which is the *B. subtilis* BS168 Erm^R is an unfit strain (Chapter 5) or a poor recipient. After three unsuccessful attempts, the filter mating experiments were repeated using a different recipient, which is *E. faecalis* JH2-2.

4.3.8 Transfer of Tn*916* WT and Tn*916*ΔTerm from *B. subtilis* BS34A to *E. faecalis* JH2-2

Conjugal Transfer of Tn*916*ΔTerm from *B. subtilis* BS34A to *E. faecalis* JH2-2 was carried out simultaneously with the transfer of the Tn*916* WT. A total of 53 putative transconjugants (*E. faecalis* JH2-2::Tn*916*ΔTerm) were obtained and reinoculated onto fresh BHI agar plates supplemented with rifampicin, fusidic acid and tetracycline. Mating in between the *B. subtilis* BS34A::Tn*916* with the same recipients (*E. faecalis* JH2-2) resulted in five putative transconjugants (*E. faecalis* JH2-2::Tn*916* WT). A total of 11 isolates of JH2-2::Tn*916*ΔTerm and three isolates of JH2-2::Tn*916* WT were chosen for further analysis. These putative transconjugants were selected from different filters to exclude the possibility of analysing siblings that may be presence in the transconjugant pool of the same filter.

The *intTn* fragment were successfully amplified from 11 isolates of JH22::Tn*916*ΔTerm and three isolates of JH22::Tn*916* WT, validating the

presence of Tn*916* or Tn*916*ΔTerm in these putative transconjugants (Figure 4-18). Transconjugants harbouring the *intTn* were further subjected to PCR detection of *tet*(M) fragment using *tet*M_1 and *tet*M_2 primers. The expected DNA bands of *tet*(M) amplicons at 740 bp were observed (Figure 4-18 (C)). All amplicons were verified by sequencing.

As the integration site(s) of both Tn*916* and Tn*916*ΔTerm within JH2-2 transconjugants are unknown, amplification of the joint-ends of the circular intermediates of both Tn*916* and Tn*916*ΔTerm were carried out. The sequence result of these amplicons will validate the presence or deletion of the terminator sequence within the transferred Tn*916* and Tn*916*ΔTerm, respectively. Simultaneously, it was carried out to investigate the ability of Tn*916*ΔTerm to excise and form circular intermediates. The expected size of the joint-ends amplicons was successfully obtained at 1436 bp and verified by sequencing (Figure 4-19). Interestingly, sequencing analysis of all the joint-ends amplicons derived from putative transconjugants of *E. faecalis* JH2-2::Tn*916*ΔTerm were detected to contain the 32 bp terminator sequence which has been originally deleted in the donor cell; BS34A Tn*916*ΔTerm. Gram staining result of the transconjugants appear as Gram-positive cocci in contrast to the donor cells which appear as Gram-positive rods. Furthermore, the transconjugants are chloramphenicol sensitive. The pGEMT fragment were successfully amplified within the donor cells *B. subtilis* BS34A Tn*916*ΔTerm and not in the *E. faecalis* JH2-2::Tn*916*ΔTerm transconjugants.
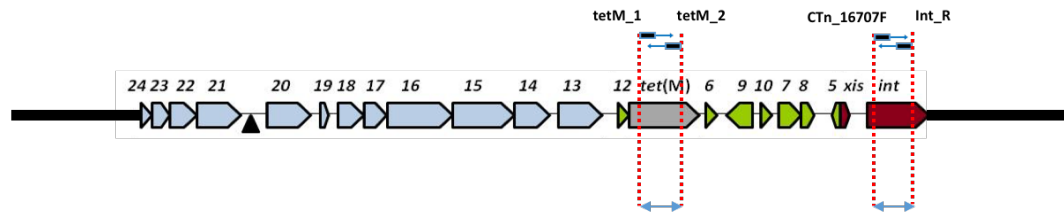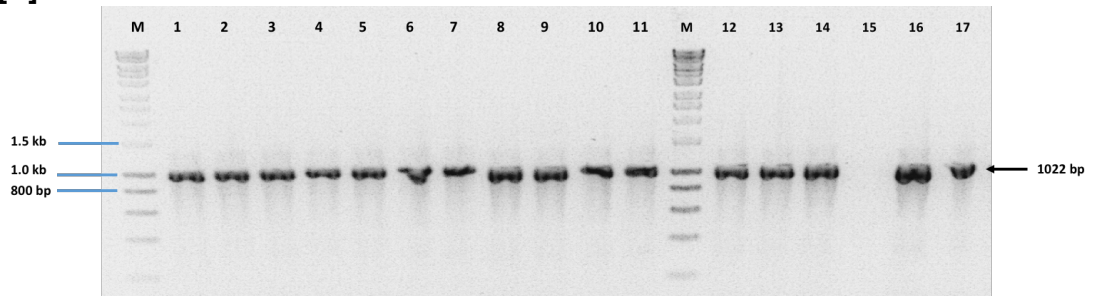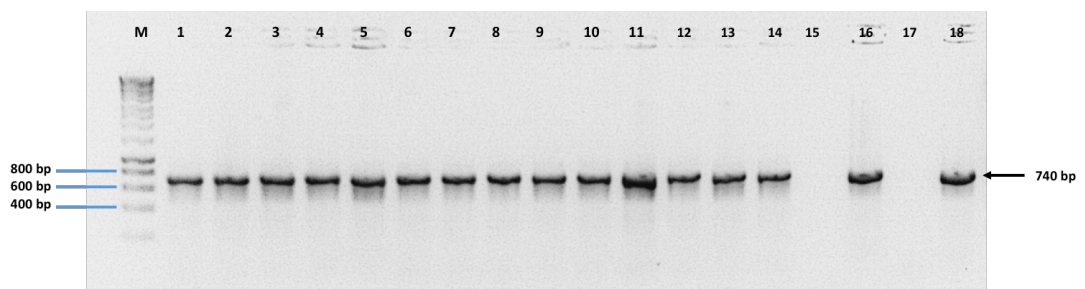
Figure 4-18 The amplification of *tet*(M) and *int*Tn fragments from the transconjugants; JH2-2::Tn*916* WT and JH2-2::Tn*916*ΔTerm.

Panel A: Schematic diagram of the Tn*916* conjugative transposon showing the amplification region of *tet*(M) and *intTn*. **Panel B:** Gel electrophoresis of *intTn* amplification product. Lane M: HyperLadder™ 1kb; Lane 1-11: *intTn* amplicons from JH2-2::Tn*916*ΔTerm of Δ1- Δ11; Lane 12-14: *intTn* amplicons from JH2-2::Tn*916* WT of WT1, WT2 and WT5; Lane 15: JH2-2 (negative control); Lane 16: BS34A::Tn*916* (positive control); Lane 17: BS34A::Tn*916*ΔTerm (positive control). **Panel C:** Gel electrophoresis of *tet*(M) amplicons. Lane M: HyperLadder™ 1kb; Lane 1-11: *tet*(M) amplicons from JH2-2::Tn*916*ΔTerm of Δ1- Δ11; Lane 12-14: *tet*(M) amplicons from JH2-2::Tn*916* WT of WT1, WT2 and WT5; Lane 15: JH2-2 (negative control); Lane 16: BS34A::Tn*916* (positive control); Lane 17: empty; Lane 18: BS34A::Tn*916*ΔTerm (positive control).
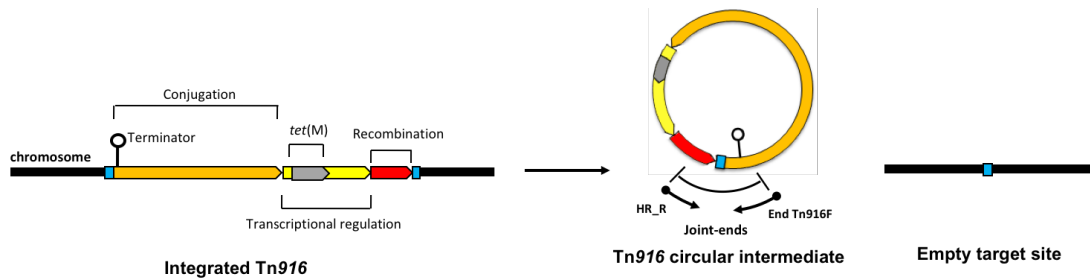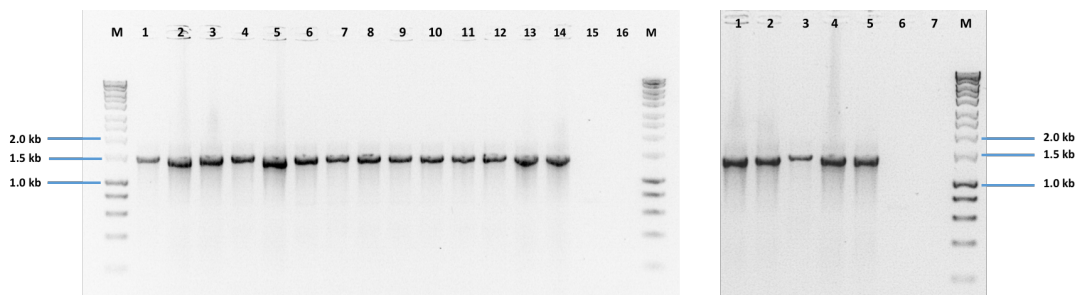
Figure 4-19 The amplification of Tn*916* joint-ends fragments from the transconjugants; JH2-2::Tn*916*ΔTerm (Panel B) and JH2-2:: Tn*916* WT (Panel C).

Panel A: Schematic diagram of the Tn*916* conjugative transposon showing the amplification region of Tn*916* joint-ends. Excision of Tn*916* from the chromosome produces the circular intermediate form of Tn*916* containing the joint-ends (blue) and empty target site. The products are detected using EndTn*916*F and HR_R primers. Arrows represent the primers and direction of priming. **Panel B:** Gel electrophoresis of joint-ends region amplicons from JH2-2::Tn*916*ΔTerm transconjugants. Lane M: HyperLadder™ 1kb; Lane 1-11: joint-ends amplicons from JH2-2:: Tn*916*ΔTerm of Δ1-Δ11; Lane 12: BS34A::Tn*916*ΔTerm (positive control); Lane 13-14: BS34A::Tn*916* (positive control); Lane 15-16: JH2-2 (negative control). **Panel C:** Gel electrophoresis of joint-ends region amplicons from JH2-2::Tn*916* WT transconjugants. Lane M: HyperLadder™ 1kb; Lane 1-3: joint-ends amplicons from JH2-2::Tn*916* WT of WT1- 3; Lane 4: BS34A::Tn*916*ΔTerm (positive control); Lane 5: BS34A::Tn*916* (positive control); Lane 6-7: JH2-2 (negative control).

4.4 Discussion

The role of the terminator which is located at the upstream of the Tn*916* conjugation module was investigated by generating the BS34A Tn*916*ΔTerm. The mutant was successfully generated by homologous recombination. The development of the mutant cassette was carried out using a novel technique where one of the homology arms is overlapping at the end of the Tn*916* element and the other arm is homologous to the chromosomal sequence of the host.

To develop the BS34A Tn*916*ΔTerm mutant, the circularised pGEM-T/Tn*916*ΔTerm mutant cassette was used for the transformation of BS34A. The integration of DNA fragments into the *B. subtilis* chromosome relies on plasmids that do not replicate in the host (Brigidi *et al.*, 1990, Vojcic *et al.*, 2012). As pGEM-T Easy vector contains pUC ORI that replicates efficiently in *E. coli*, it is assumed that the vector will not be able to replicate when transformed in *B. subtilis*. By using the circularised pGEM-T/Tn*916*ΔTerm, we have successfully generated three mutant clones with the targeted deletion of the Tn*916* terminator. However, the single crossover event of the mutant cassette had occurred with the co-integration of the pGEM-T/Tn*916*ΔTerm (Figure 4-11).

Conjugation of Tn*916* element is primarily likely to occur when it is in a circular form. The conjugation of Tn*916* begins with the excision of the transposon from the donor mediated by the staggered cleavages on both ends of the element. These staggered cleavages require two transposon-encoded

proteins, the integrase (Int*Tn*) and excisionase (Xis*Tn*) that will bind specifically at the two transposon ends (Caparon & Scott, 1989, Storrs *et al.*, 1991, Rudy *et al.*, 1997, Celli & Trieu-Cuot, 1998). These binding sites are referred as inverted repeats right (IR_R) and inverted repeats region left (IR_L) (Lu & Churchward, 1994). Upon excision from the chromosome, Tn*916* forms a circular intermediate (CI) structure (Caparon & Scott, 1989). In our *B. subtilis* BS34A Tn*916*ΔTerm mutant, additional IR_L had been introduced as a result of co-integration of the pGEM-T/Tn*916*ΔTerm followed by the original copy of the Tn*916* terminator. This additional IR_L was denoted as IR_L[1] (Figure 4-20). With the additional binding site for the recombinase, two possible forms (Type A or B) of Tn*916* circular intermediates may be generated as shown in Figure 4-20. Recombinase activity on IR_L[2] and IR_R will create a circular intermediate type A: Tn*916*ΔTerm CI that carry the targeted deletion of the terminator. Alternatively, if the recombinase activity occurred on IR_L[1] and IR_R, circular intermediate type B might be generated. The circular intermediate type B will contain the whole Tn*916* with deleted terminator, plus the integrated pGEM-T vector backbone and the rest of the mutant cassette followed by the Tn*916* terminator (Figure 4-20).
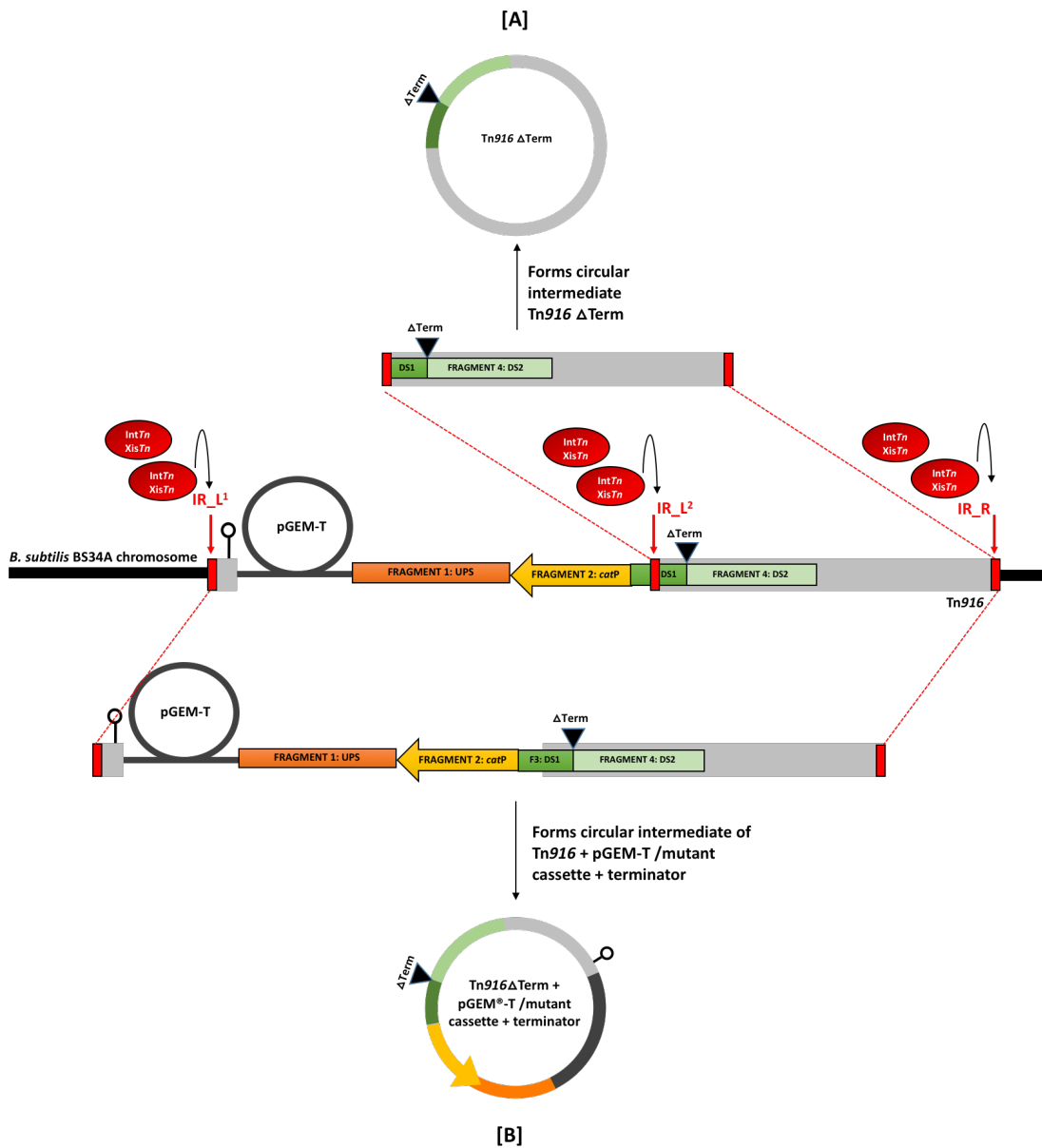
Figure 4-20 Possible forms of circular intermediates generated based on the recombinase activity on IR_R paired with IR_L^1 or IR_L^2.

Structure ( ) represent the rho-independent terminator, structure ( ) represents the deleted terminator region labelled as ∆Term, structure ( ) represents the recombinase Int*Tn* and Xis*Tn* depicted to act on their binding sites; IR_R, IR_L^1 or IR_L^2 represented by the red box.

To further investigate on the integration of the pGEM-T/Tn*916*ΔTerm into the BS34A Tn*916*ΔTerm chromosome, a whole genome sequence analysis was done. We failed to detect any sequence within contigs that validated the integration of the pGEM-T/Tn*916*ΔTerm into the BS34A chromosome as the sequence reads obtained from the Illumina sequencing are relatively shorts. However the presence of two contigs (CONTIGS 44 and 47) containing Tn*916* left end sequences with and without the terminator is in line with the predicted structure in Figure 4-11. We acknowledged the limitation obtained with the Illumina sequencing data and a better insight may be achieved with long read sequencing data.

Although our diagnostic PCR results suggest that the homologous recombination had occurred together with the co-integration of *cat*P and the pGEM-T plasmid backbone, the targeted deletion had occurred and therefore, the mutant was tested for their ability to transfer the Tn*916*ΔTerm by filter mating experiments. The transfer of the Tn*916*ΔTerm was also done in attempt to obtain a scar-less mutant with the targeted terminator deletion within the conjugation module of Tn*916.* However, no obvious transconjugants were obtained from the mating experiment with CU2189 Rif^R Nal^R as recipients. Another recipient strain was selected to be resistant to erythromycin (BS168 Erm^R). However, the mutation in the *rplV* of BS168 that confers erythromycin resistance seems to impart a fitness cost on the mutant. A delayed growth rate of the BS168 Erm^R was observed in comparison to the wild type; affecting its ability to be used as an efficient recipient for the filter mating experiment. The erythromycin resistance in BS168 Erm^R strain is due to a novel mutation and further investigated in Chapter 5.

With *E. faecalis* JH2-2 Rif^R Fus^R as the recipient cells, transconjugants were obtained but the sequencing result of their CI joint-ends amplicon revealed that it contains the terminator sequence (which was deleted in the Tn*916*ΔTerm). Our interpretation of this data is that the co-integration of the pGEM-T vector backbone together with another copy of homologous region adds another substrate for crossing over to occur, thus creating a complex recombination event (Figure 4-21). With the co-integration of pGEM-T vector backbone and the rest of the mutant cassette followed by the Tn*916* terminator in the donor; two pairs of additional homologous regions were introduced; (i) A^1 and A^2, which are the sequence of Fragment 1:UPS of the mutant cassette (represented by striped orange and orange box, respectively) and (ii) B^1 and B^2, which are the sequence of Fragment 3:DS1 of the mutant cassette (represented by striped green and dark green box, respectively). The B^2 is the fragment originated from the mutant cassette with the deleted terminator sequence. While B^1 is the fragment originated from the wild type Tn*916* of BS34A that contains the terminator sequence (Figure 4-21). Homologous recombination in between B^1 and B^2 regions within *B. subtilis* BS34A Tn*916*ΔTerm may regenerate the wild type Tn*916* containing the terminator. This wild type Tn*916* then will be excised, forming a CI that was subsequently transferred to the recipient cell of *E. faecalis* JH2-2 (Figure 4-22).
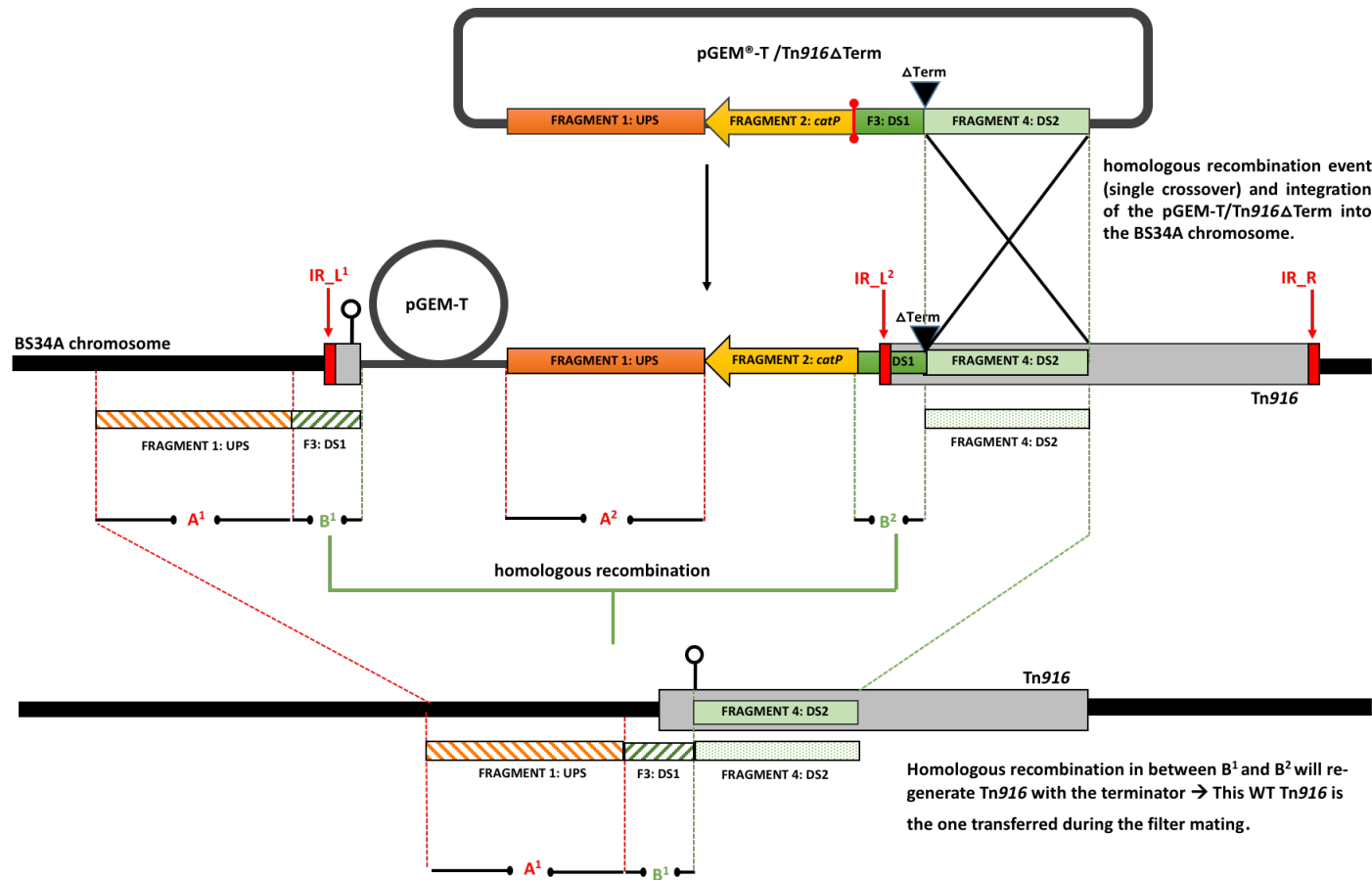
Figure 4-21 Additional homologous recombination substrates that have been introduced in *B. subtilis* BS34A::Tn*916*ΔTerm.

The two pairs of homologous regions are labelled as A[1] (striped orange box; Fragment 1:UPS) and A[2] (dark orange box; Fragment 1:UPS), B[1] (striped green box; Fragment 3: DS1 with the terminator) and B[2] (dark green box; Fragment 3: DS1 with deleted terminator).
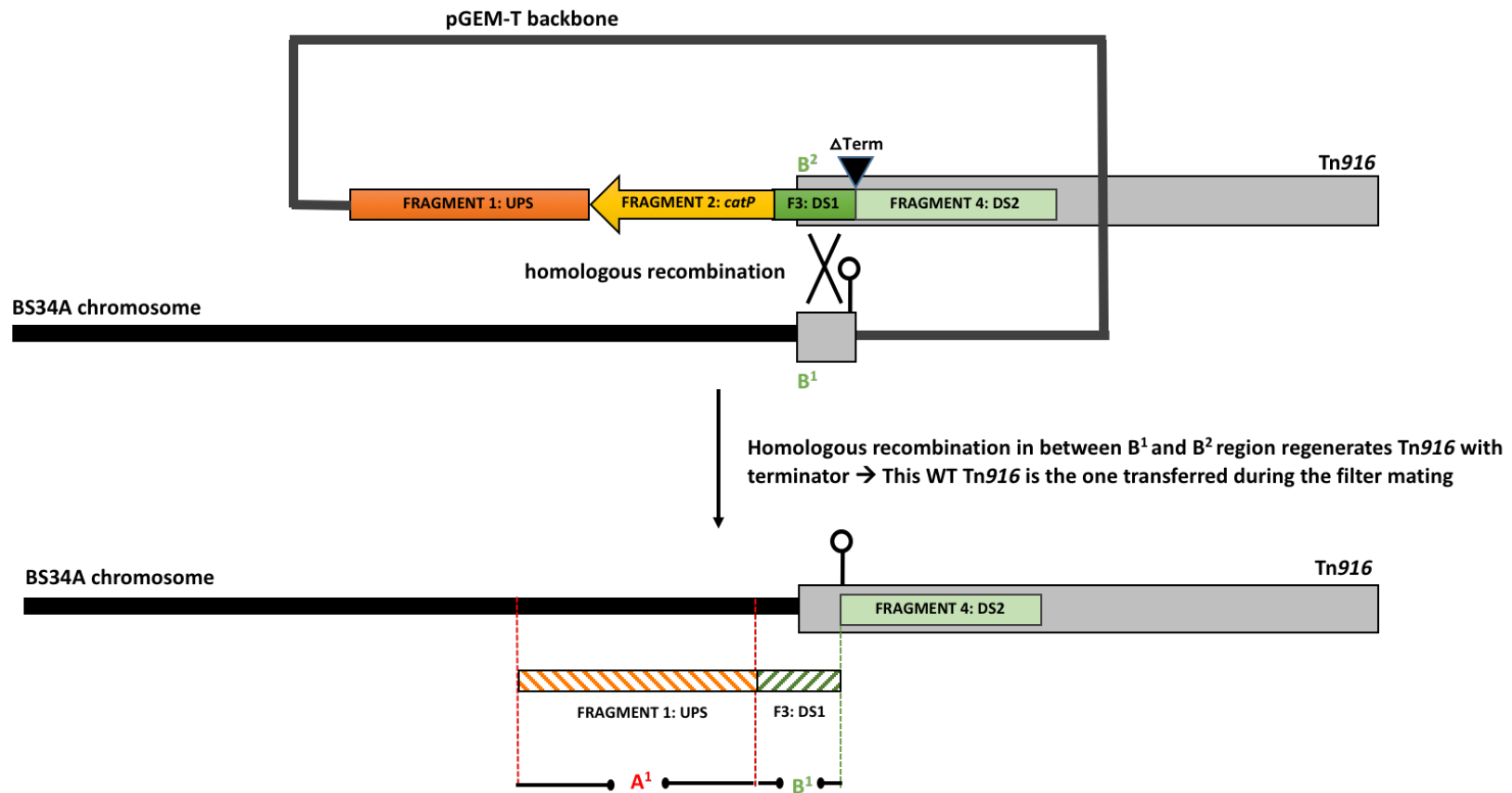
Figure 4-22 Regeneration of the Tn*916* with terminator as a result of homologous recombination in between the B^1 and B^2 regions within *B. subtilis* BS34A::Tn*916*ΔTerm.

The B^1 (striped green box; Fragment 3: DS1 with the terminator) and B^2 (dark green box; Fragment 3: DS1 with deleted terminator).

4.5   Conclusions

The *B. subtilis* BS34A mutant with a deletion of the Tn*916* terminator was successfully generated via SOEing PCR and homologous recombination, denoted as BS34A Tn*916*∆Term. The mutant cassette was generated using a novel technique where one of the homology arms is overlapping at the end of the Tn*916* element. To our knowledge, this is the first time that it has been done. From the PCR and genomic data analysis, the co-integration of the pGEM-T vector backbone followed by the rest of the mutant cassette (that carry another copy of homologous regions) within the generated BS34A Tn*916*∆Term mutant was observed. The introduction of these additional copies of homologous regions had resulted in the re-generation of Tn*916* with terminator. In our filter-mating experiments, we cannot detect the transfer of Tn*916*∆Term, and therefore, further investigation regarding the transfer activity of the element was not carried out.

5 Analysis of *Bacillus subtilis* Erythromycin and Tylamac Resistant Strains

5.1 Introduction

The ribosome is the site of protein synthesis and is known to be a main target of antibiotics. Macrolides are one example of protein synthesis inhibitors that target the large subunit of ribosome (50S) specifically around the peptidyl transferase center (PTC) and the nascent peptide exit tunnel (NPET) (Gabashvili *et al.*, 2001). NPET is a passageway that extends across the large subunit of the ribosome where the elongated polypeptide chains are released (Figure 5-1) (Nissen *et al.*, 2000). Macrolides were thought to hamper the progression and release of these nascent peptides simply by blocking this passageway, or by causing a clogged polypeptide that disrupt the protein synthesis apparatus once the polypeptides reached 3-10 amino acids long (Tenson *et al.*, 2003). However, the current view of mode of action of macrolides is more complicated, instead of being global protein synthesis inhibitor, macrolides have been shown to play a role as a modulator of translation where it selectively interferes the production of a subset of proteins. As discussed in Chapter 1, the macrolide-bound ribosome stalls when it needs to polymerize the amino acid sequence of MAM (Chiba *et al.*, 2011, Davis *et al.*, 2014, Kannan *et al.*, 2014, Wekselman *et al.*, 2017, Halfon *et al.*, 2019).

The middle section of the NPET is gated by long extensions of protein L4 and L22 extending to the core from the globular surface domain of the ribosome (Figure 5-1) (Ban *et al.*, 2000, Nissen *et al.*, 2000). These L4 and L22 proteins are encoded by *rplV* and *rplD*, respectively and studies have shown that mutations that occur within these genes cause resistance to macrolides (Wittmann *et al.*, 1973, Chittum & Champney, 1995, Unge *et al.*, 1998,

Gregory & Dahlberg, 1999, Davydova *et al.*, 2001, Gabashvili *et al.*, 2001, Bingen *et al.*, 2002, Canu *et al.*, 2002, Canu *et al.*, 2002, Davydova *et al.*, 2002, Franceschi *et al.*, 2004, Rolain & Raoult, 2005, Cagliero *et al.*, 2006, Zaman *et al.*, 2007, Gentry & Holmes, 2008, Moore & Sauer, 2008, Lovmar *et al.*, 2009).
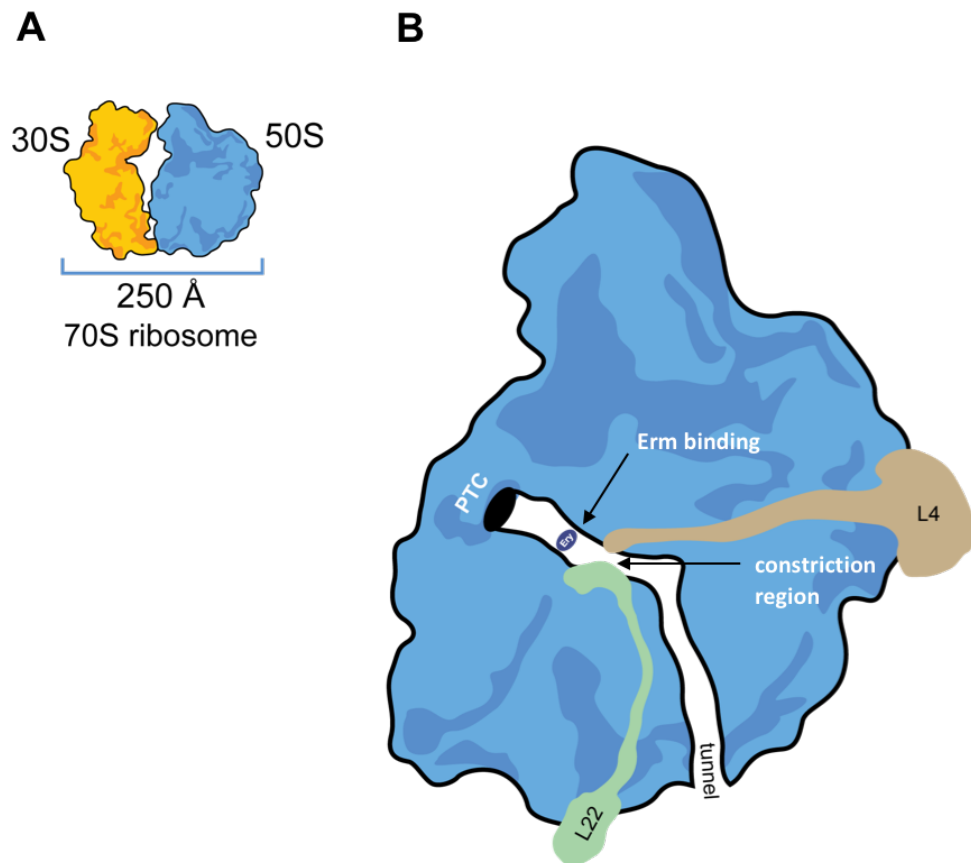
Figure 5-1 The structure of nascent peptide exit tunnel (NPET).

Panel A shows a schematic diagram of the ribosome consist of small 30S (yellow) and large 50S (blue) subunits. Panel B shows the cross section of the large subunit (50S). The location of NPET (labelled as tunnel) is adjacent to the peptidyl-transferase centre (PTC) (black). The erythromycin binding site is shown in dark blue (arrow). The loops of ribosomal proteins L4 (light brown) and L22 (green) form the narrowest constriction region within the NPET.

L22 protein is one of the core loops that forms the narrow-constricted part in the NPET. Near this constricted region, lies the binding pocket for macrolides (Figure 5-1). Conformational changes in this protein loop can alter the size and shape of NPET that may neutralise the effect of macrolide binding that in turn will mediate the macrolide resistance (Gabashvili *et al.*, 2001, Davydova *et al.*, 2002, Moore & Sauer, 2008, Lovmar *et al.*, 2009, Wekselman *et al.*, 2017). This alteration could occur due to single point mutation, deletions or insertions within *rplV* that encodes L22 protein. For example, in *E.coli,* nine bp deletions within *rplV* had led to the removal of three amino acids (Met-Lys-Arg at the residue 82-83-84) in L22 that caused an increased width of the tunnel allowing nascent polypeptide to progress and pass through the NPET of erythromycin bounded ribosome (Wittmann *et al.*, 1973, Chittum & Champney, 1994, Zaman *et al.*, 2007). A similar effect has also been reported due to the three amino acids deletion (Δ82–84 mutation) in the L22 of *Thermus thermophilus* and *Haloarcula marismortui* (Davydova *et al.*, 2002, Tu *et al.*, 2005, Wekselman *et al.*, 2017).

Most of the early reported changes in L22 are due to single amino acid substitution or three amino acids deletions. More recently, amino acid duplications within the L22*,* majorly affecting the carboxy region of L22 have been described in various macrolide and ketolide resistant isolates of *B. subtilis*, *Staphylococcus aureus*, *Streptococcus pneumoniae, E. coli* and *Campylobacter jejuni* (Doktor *et al.*, 2004, Hisanaga *et al.*, 2005, Cagliero *et al.*, 2006, Zaman *et al.*, 2007, Gentry & Holmes, 2008, Chiba *et al.*, 2009, Han *et al.*, 2018). Only two other cases of spontaneous macrolide resistance in *B. subtilis* strain 168 have been reported although this resistance was shown to

be mediated by the alteration of ribosomal protein L17 encoded by *rplQ* and not by protein L22 (Tipper *et al.*, 1977, Sharrock *et al.*, 1981).

In this study, erythromycin and TylAMac (Tylosin A analogue) resistance have been selected for in *B. subtilis* 168 that resulted in mutant strains with duplications in their ribosomal gene; *rplV*. The erythromycin resistant *B. subtilis* BS168 was initially selected to be used as a recipient in the filter mating experiment for the horizontal gene transfer study (Chapter 4). Discussion with the Anti-Wolbachia Consortium (A·WOL) research group of Liverpool School of Tropical Medicine (LSTM), who were trying to determine the target site of their novel macrolide based antibiotics; Tylosin Analogues Macrofilaricides (TylAMac™) has led us to test these novel antibiotics against our erythromycin resistant *B. subtilis* BS168 strain. TylAMac™ is a macrofilaricidal agent that was designed to target the *Wolbachia* endosymbiont of filarial nematodes to treat lymphatic filariasis (Johnston *et al.*, 2017, Hong *et al.*, 2019). Anti-*Wolbachia* therapy has been clinically validated with doxycycline (Hoerauf *et al.,* 2008). However, treatment with doxycycline requires long-term therapy and is unsafe for pregnant women and children below 9 years. This initiated the development of new anti-*Wolbachia* compounds with superior profiles; reduced treatment duration, improved oral absoprtion and increased potency (Johnston *et al.*, 2017, Hong *et al.,* 2019).

In this chapter we aimed to identify the resistance determinant in *B. subtilis* BS168 Erm[R], BS168 T469[R] and BS168 T4083[R] and to determine the target site of the proprietary tylosin A analogues; TylAMac™ '469 and '4083. This study therefore set out to investigate the resistance mechanism of macrolide-

resistant isolates against TylAMac™ '469 and '4083. *B. subtilis* is a fully sequenced and genetically tractable organism which is susceptable to macrolides. We also have extensive experience of using genetic systems in *B. subtilis* within our laboratory and the generation of macrolide-resistant strains of this bacterium was carried out for a previous chapter described in Section 4.2.7. Therefore, despite being found to target the Gram-negative *Wolbachia*, analysis of the *B. subtilis* resistant isolates provides insights into mode of action and binding site of TylAMac™. Additionally, as *Wolbachia* cannot be grown without intracellular culture, a macrolide susceptible *B. subtilis* was considered a suitable model organism to investigate the mechanism of action of the tylosin analogues. Susceptibility testing was also done to determine the MIC for macrolides; erythromycin, tylosin A, and the two Tylosin Analogues Macrofilaricides (TylAMac™) '469 and '4083 in *B. subtilis*.

5.2 Materials and methods

5.2.1 Selection of erythromycin , Tylosin A and TylAMac™ resistant *B. subtilis*

A triplicate of 50 mL antibiotic free BHI broth was inoculated with a single colony of *B. subtilis* strain 168. The inoculated broth was incubated overnight at 37°C, 200 rpm for 16-18 hrs. The overnight culture was spun down at 4500 x *g* for 15 mins and the pellet was resuspend in 1 mL fresh BHI broth (10 µl of the suspension is added to 90 µL of 1X Phosphate Buffered Saline (PBS) and used for serial dilution to determine the number of cells). An aliquot of 100 µL of the suspension was plated onto LB agar plates that were supplemented with either erythromycin at a concentration of 10 µg/mL, Tylosin A (commercially available drug used in veterinary markets), TylAMac™ '469, TylAMac™ '4083 (new analogues of Tylosin A, active against *Wolbachia*) (Taylor *et al.*, 2019) or Doxycycline (used as a negative control) at a concentration of 4.0 µg/mL. These plates were incubated at 37°C and observed every day (for five days) for colony growth. The experiment was repeated in three biological replicates. The obtained erythromycin and TylAMac resistant isolates were individually picked and streaked on fresh selective plates. Bacterial stocks were maintained in 1 ml aliquots of 20% (v/v) sterilised glycerol in Luria Bertani (LB) broth at - 80°C. The erythromycin resistant isolates were confirmed to the species level by *gyrA* sequencing using p-gyrA-f and p-gyrA-r primer pair (Chun & Bae, 2000) as listed in Table 2-3.

5.2.2 Bacterial genomic DNA and plasmid extraction

Genomic DNA purification was carried out using the Gentra PUREGENE®
Yeast/Bact Kit (Qiagen, UK) with slight modifications as described in section
2.4.1. Plasmid DNA purification was carried out using QIAprep Spin Miniprep
Kit (Qiagen, UK) as described in section 2.4.2.

5.2.3 Amplification of *rplV* from *B. subtilis* 168, BS168 Erm^R, BS168 T469^R and BS168 T4083^R

The primers for the amplification of *rplV* was designed based on the genomic
sequence data of *B. subtilis* strain 168; *rplV*_F and *rplV*_R (as listed in Table
2-3). PCR amplification was carried out using ProFlex PCR System (Applied
Biosystem, UK) with the following thermal profile: initial denaturation at 98°C
for 30 sec, 25-30 cycles [98°C, 10 sec; 56.1°C, 30 sec (annealing
temperature); 72°C, (20-30 sec/kb); 1 cycle of 2 min at 72°C and preservation
at 4°C until the sample were analysed. The PCR reaction mixture contains 25
µl of Q5 High-Fidelity 2X Master Mix (NEB, UK), 2.5 µl of each forward and
reverse primers (at a final concentration of 0.5 µM) , and 1 µl of DNA template
(approximately 100 ng/µl). The total volume of PCR mixture was made up to
50 µl using distilled water.

5.2.4 Sequence analysis of *rplV* derived from *B. subtilis* 168, BS168 ErmR, BS168 T469^R and BS168 T4083^R

PCR products and plasmids were sent to Genewiz Inc. (Genewiz, United
Kingdom) for DNA sequencing. Sequences were aligned, assembled and

manipulated by using BioEdit software version 7.2.0 and Clustal Omega (http://www.ebi.ac.uk/Tools/msa/clustalo). The sequences were analyzed by comparing the DNA sequences and translated amino acid sequences to National Center for Biotechnology Information (NCBI) databases with the Basic Local Alignment Search Tool (BLAST) (http://blast.ncbi.nlm.nih.gov/) (Altschul *et al.*, 1990).

5.2.5 Whole genome sequencing of *B. subtilis* 168, BS168 Erm[R], BS168 T469[R] and BS168 T4083[R]

B. subtilis 168 (parental strain) and the mutant strains (BS168 Erm[R], T469[R] and T4083[R]) were sent to MicrobesNG (http://www.microbesng.uk) for whole genome sequencing using 2 × 250 bp paired-end reads on the Illumina platform. *De novo* assembly of each of the genomes was carried out with SPAdes (Bankevich *et al.*, 2012) via MicrobesNG (Birmingham). An automated annotation of the assembled genomes was carried out using Prokka (Seemann, 2014).

5.2.6 Analysis of whole genome sequence data of *B. subtilis* 168, BS168 Erm[R], BS168 T469[R] and BS168 T4083[R]

Identification of mutation(s) and any genetic variations (single-nucleotide polymorphisms (SNPs), insertion, deletion and duplications) was done by comparing the genomic data of the mutants strains with the parental strain by using *Breseq* (http://barricklab.org/breseq) (Deatherage & Barrick, 2014). The L22 protein structure modelling was carried out by using SWISS-MODEL (Biasini *et al.*, 2014), which searched and built the target structure based on

related evolutionary structures in the protein database. The L22_7D and L22_18D were superimposed with the wild type L22 using The PyMol Molecular Graphics System (Schrodinger, LLC). Screening for AMR genes were performed using ResFinder (Zankari, 2014) and alternatively, sequence of known resistance genes commonly found in *B. subtilis* were blasted against the acquired genomic data of the strains.

5.2.7 Determination of Minimum Inhibitory Concentrations (MICs) of erythromycin, TylAMac™ '469, TylAMac™ '4083 and Tylosin A for *B. subtilis* 168, BS168 Erm^R, BS168 T469^R and BS168 T4083^R

The MIC values of *B. subtilis* 168 (*rplV*^{WT}), BS168 Erm^R, BS168 T469^R and BS168 T4083^R strains were determined for erythromycin, Tylosin A, TylAMac™ '469 and TylAMac™ '4083 using broth microdilution method following the Clinical and Laboratory Standards Institute (CLSI) guideline. Erythromycin and Tylosin A were obtained in powdered form from Sigma-Aldrich (UK). The TylAMac™ '469 and TylAMac™ '4083 compounds were supplied by the Anti-Wolbachia (A.WOL) consortium (Liverpool School of Tropical Medicine, UK) (Taylor *et al.*, 2019). Briefly, antibiotics were prepared by serial two-fold dilutions in Cation adjusted Mueller-Hinton Broth 2 (CA-MHB) (Sigma-Aldrich ,UK) in different ranges of concentration depending on the particular antibiotic. These are done in sterile U-bottom Costar® 96-well Clear Polystyrene Microplates 3367 (Corning, US) in triplicates. The media were inoculated with 50 μL of diluted overnight culture to obtain approximately 1×10^6 cfu/well in a 100 μL total volume. The plates were incubated at 37°C for 18-24 hrs. The MIC was defined as the lowest concentration of antibiotic

which gives a complete inhibition of visible growth in comparison with inoculated and uninoculated antibiotic-free wells.

5.2.8 Forward genetics

5.2.8.1 Cloning of *rplV*^WT^, *rplV*^21D^ and *rplV*^54D^ into pGEM-T Easy vector and directional cloning of *rplV*^WT^, *rplV*^21D^ and *rplV*^54D^ into pHCMC04 vector

The *rplV* genes from parental and mutant strains of *B. subtilis* 168 (*rplV*^{WT}, *rplV*^{21D} and *rplV*^{54D}) were amplified using a primer pair of *rplV*F_*Spe*I and *rplV*R_*BamH*I (Table 2-3). The amplified region includes the ribosomal binding site, the open reading frame (ORF) plus the *BamH*I and *Spe*I restriction sites. The *rplV* genes were cloned into pGEM-T Easy vector by TA-cloning generating three different constructs denoted as pGEM-T/*rplV*^{WT}, pGEM-T/*rplV*^{21D} and pGEM-T/*rplV*^{54D} (section 2.4.11). These constructs were extracted from *E. coli* transformants (section 2.4.2), and the inserts (*rplV*^{WT}, *rplV*^{21D} and *rplV*^{54D}) were digested with *BamH*I and *Spe*I (section 2.4.6), gel purified (section 2.4.5) and cloned into pHCMC04 shuttle vector by directional cloning method generating pHCMC04/*rplV*^{WT}, pHCMC04/*rplV*^{21D} and pHCMC04/*rplV*^{54D} constructs.

5.2.8.2 Transformation of *E. coli* and *B. subtilis* BS34A with *rplV* mutant constructs

E. coli transformation was carried out using competent cells α-select silver efficiency (Bioline, UK) according to the standard heat-shock transformation

protocol as described in section 2.4.11. Transformants were selected on LB agar supplemented with ampicillin (100 µg/mL) and IPTG/X-gal plates for blue-white screening. All plates were incubated overnight at 37°C. The transformants were subjected to plasmid extraction and digestion with *EcoR*I to select for positive clones. The clones carrying the correct insert size were sent for sequencing to verify the insert sequence by using M13-F primer (Table 2-3). *E. coli* transformants were denoted as *E.coli* pGEM-T/*rplV*^{WT}, *E.coli* pGEM-T/*rplV*^{21D} and *E.coli* pGEM-T/*rplV*^{54D} (Table 2.1).

E. coli transformation with pHCMC04/*rplV*^{WT}, pHCMC04/*rplV*^{21D} and pHCMC04/*rplV*^{54D} constructs were also done with the same protocol as described in section 2.4.11. Transformants were selected on LB agar supplemented with chloramphenicol (5 µg/mL) and all plates were incubated overnight at 37°C. Screening for positive transformants was done by amplifying the insert region using a primer pair; CMC04F and CMC04R (Table 2-3). The clones carrying the correct insert size were sent for sequencing to verify the insert sequence. *E. coli* transformants were denoted as *E.coli* pHCMC04/*rplV*^{WT}, *E.coli* pHCMC04/*rplV*^{21D} and *E.coli* pHCMC04/*rplV*^{54D} (Table 2.1) and subsequently subjected to plasmid purification (section 2.4.2). The purified pHCMC04/*rplV*^{WT}, pHCMC04/*rplV*^{21D} and pHCMC04/*rplV*^{54D} constructs plus pHCMC04 without any insert (pHCMC04 vector only (VO)) were subsequently cloned into *B. subtilis* BS34A.

B. subtilis transformation was carried out using *B. subtilis* BS34A competent cells according to the protocol describe in section 2.4.10. The *B. subtilis* BS34A competent cells were prepared as described in section 2.4.9. The

transformants were selected on BHI agar supplemented with chloramphenicol (5 µg/mL). All plates were incubated overnight at 37°C. The transformants were subjected to plasmid purification (section 2.4.2) and screened by amplifying the insert region using a primer pair; CMC04F and CMC04R (Table 2-3). The clones carrying the correct insert size were sent for sequencing to verify the insert sequence by using CMC04F primer (Table 2-3). Positive transformants were denoted as *B. subtilis* BS34A VO, *B. subtilis* BS34A *rplV*[WT], *B. subtilis* BS34A *rplV*[54D] and *B. subtilis* BS34A *rplV*[21D].

5.2.8.3 Ectopic expression of the *rplV* mutant constructs in *B. subtilis* BS34A

The MIC values of *B. subtilis* BS34A strains (BS34A, BS34A VO, BS34A *rplV*[WT], BS34A *rplV*[54D], BS34A *rplV*[21]) were determined for erythromycin, TylAMac™ '469 and TylAMac™ '4083 using broth macrodilution method. Antibiotics are prepared in a serial two-fold dilutions as described above in sterile 50 mL tubes. The media were inoculated with 5 mL of diluted overnight culture in a 10 mL total volume. The tubes were incubated at 37°C, with shaking at 200 rpm for 18-24 hrs (Andrews, 2001). Bacterial growth was determined by reading the optical density at 600 nm (OD_{600}) before and after the incubation period.

5.3 Results

5.3.1 Erythromycin and TylAMac resistance strains of *B. subtilis* BS168

A total of three *B. subtilis* 168 mutants which are resistant to erythromycin, TylAMac™ '469 and '4083 were selected and denoted as BS168 Erm^R, T469^R and T4083^R, respectively. There was only one colony obtained during the selection of BS168 Erm^R. To determine whether the spontaneous mutation of BS168 Erm^R mutant can be reselected, repetition of the resistance selection experiment was performed. After five days incubation, no growth observed on any of the 30 plates. The average value of cfu/mL of the total cells plated is at 2.84×10^{10}, suggesting that the spontaneous mutation in *rplV* of *B. subtilis* is a rare event.

The MICs results of erythromycin, TylAMac™ '469, TylAMac™ '4083 and Tylosin A for *B. subtilis* BS168 (parental strain), Erm^R, T469^R and T4083^R strain is shown in Table 5-1. Overall, the MICs value of all tested antibiotics for BS168 Erm^R strain is higher than BS168 T469^R and T4083^R strains with a very large difference (two to three folds) observed for Tylosin A and both TylAMac™ (Table 5-1). The MICs value of BS168 T469^R and T4083^R strains against all tested antibiotics are within a similar range determined in three independent experiments. The MIC for erythromycin is at the lowest value in comparison to other antibiotics in all mutant strains.

Table 5-1 Minimum Inhibitory Concentrations (MICs) of erythromycin, TylAMac™ '469, TylAMac™ '4083 and Tylosin A for *B. subtilis* BS168 (parental strain), Erm^R, T469^R and T4083^R. The MIC is determined from three biological replicates where a range of MICs value is given.

Strain	MIC (µg/ml)			
	TylAMac™ '469	TylAMac™ '4083	Tylosin A	Erythromycin
BS168	0.5	0.25 - 0.5	0.125 - 0.5	0.015 - 0.03
BS168 Erm^R	32.0	16.0 - 32.0	16.0 - 32.0	4.0 - 8.0
BS168 T469^R	8.0	4.0	4.0	1.0 - 2.0
BS168 T4083^R	8.0	4.0 - 8.0	4.0 – 8.0	2.0

All of the mutant strains demonstrated cross-resistance towards TylAMac™ '469, TylAMac™ '4083, Tylosin A and erythromycin indicating that the mutations might have occurred around the same sites directly or indirectly affecting the binding effect of macrolides.

5.3.2 Growth of the mutant strains; BS168 Erm^R, T469^R and T4083^R in comparison to the parental strain BS168

The growth curves of the mutants and parental strain were determined in antibiotic-free CA-MHB broth media at 37°C. There are growth differences observed among the mutants (BS168 Erm^R, T4083^R and T469^R) and the wild type strain (Figure 5-2).

BS168 Erm^R exhibits a different colony morphology in comparison to the parental strain. It has a smooth, glistening surface with entire and rounded

edge (Figure 5-3). But this distinctive colony morphology was only seen on agar plates supplemented with erythromycin. On antibiotic free agar plates, the colony morphology was the same as wild type *B. subtilis* BS168, which have a rough surface with irregular and undulated edge (Figure 5-3). To ensure that the colonies observed are BS168 Erm[R] strain, both the BS168 parental strain and the BS168 Erm[R] mutant strain were identified to the species level by *gyrA* sequencing. The sequence alignment of partial *gyrA* derived from BS168 and BS168 erm[R] amplicons showed 100% identity. The other two mutants strains; BS168 T469[R] and T4083[R] do not exhibit a different colony morphology in comparison to the parental strain.
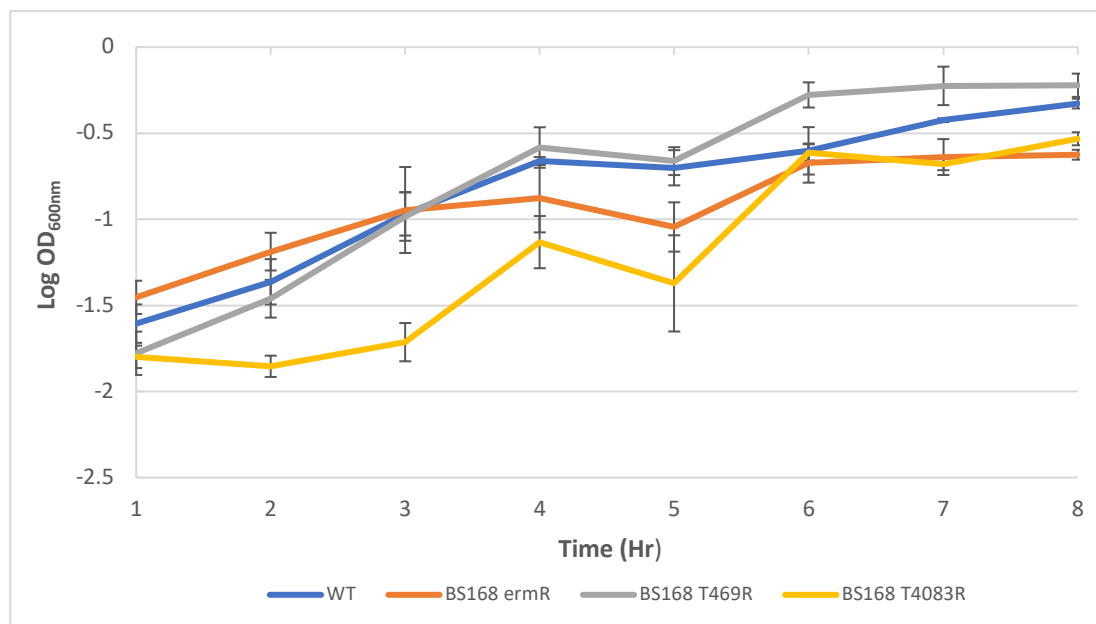
Figure 5-2 Comparative growth curves of *B. subtilis* BS168 WT, BS168 Erm[R], BS168 T469[R] and BS168 T4083[R].

The growth curve was carried out three times in triplicate. Error bars indicate the standard errors of three independent experiment.

A. 24 hrs growth on LB abf
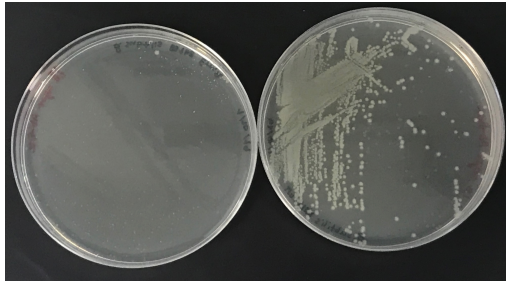
BS168 Erm^R   **BS168 WT**

E. 24 hrs growth on LB erm
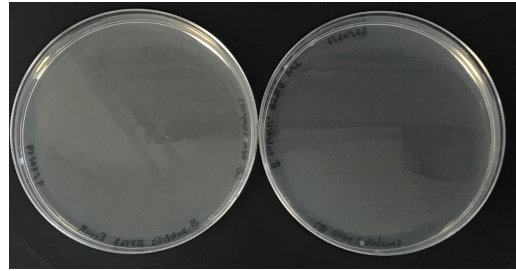
BS168 Erm^R   **BS168 WT**

B.
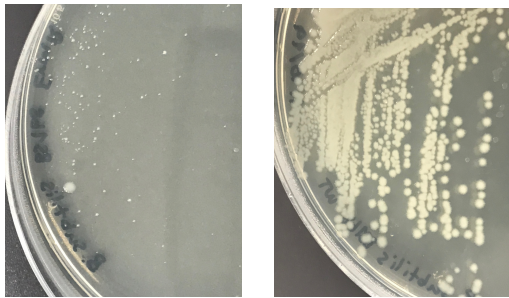
F.
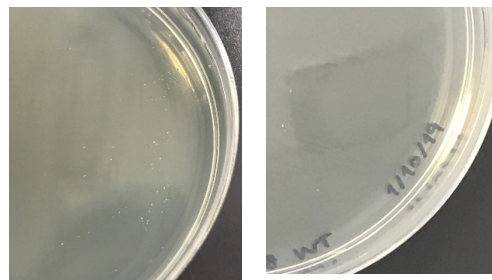
C. 48 hrs growth on LB abf
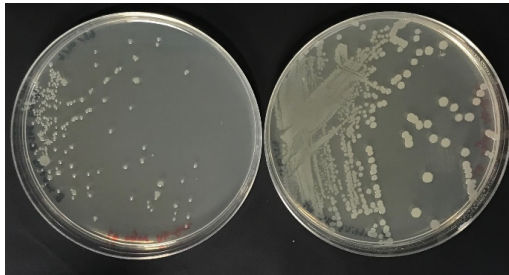
BS168 Erm^R   **BS168 WT**

G. 48 hrs growth on LB erm
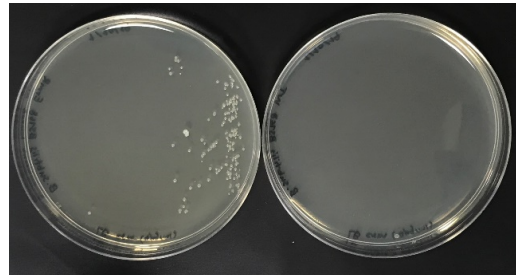
BS168 Erm^R   **BS168 WT**

D.
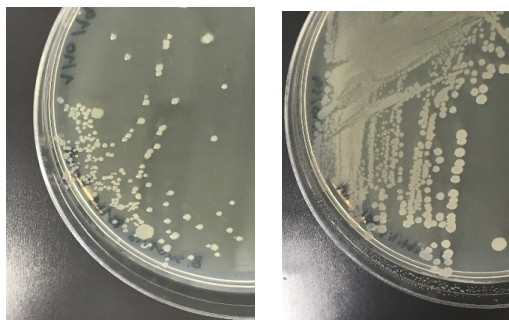
G.
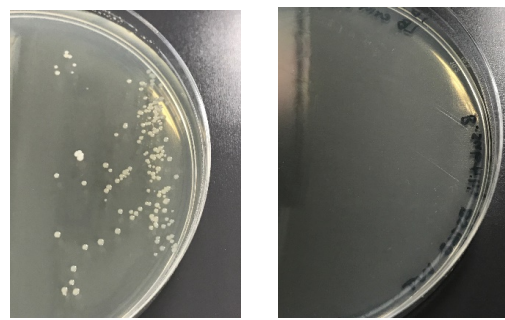
Figure 5-3 Growth comparison of erythromycin resistant *B. subtilis* mutant (BS168 Erm^R) and the parental strain BS168 after 24 and 48 hrs incubation time. Panel A: *B. subtilis* BS168 Erm^R (left) and BS168 parental strain WT (right) on abf LB agar after 24 hrs. **Panel B:** left; a close up of BS168 Erm^R colonies; right; a close up of BS168 WT colonies after 24 hrs of incubation time. **Panel C:** *B. subtilis* BS168 Erm^R (left) and BS168 parental strain WT (right) on abf LB agar after 48 hrs. **Panel**

D: left; a close up of BS168 Erm^R colonies; right; a close up of BS168 WT colonies after 48 hrs of incubation time. **Panel E:** *B. subtilis* BS168 Erm^R (left) and BS168 parental strain WT (right) on LB agar supplemented with erythromycin (4.0 ug/mL) after 24 hrs. **Panel F:** left; a close up of BS168 Erm^R colonies; right; a close up of BS168 WT after 24 hrs of incubation time. **Panel G:** *B. subtilis* BS168 Erm^R (left) and BS168 parental strain WT (right) on LB agar supplemented with erythromycin (4.0 ug/mL) after 48 hrs. **Panel H:** left; a close up of BS168 Erm^R colonies; right; a close up of BS168 WT after 48 hrs of incubation time.

5.3.3 Amplification of *rplV* from *B. subtilis* BS168 WT, Erm^R, T469^R and T4083^R

Generally, resistance against macrolides can occur due to several mechanisms; (i) ribosomal modification by erythromycin ribosomal methylase (*erm*); (ii) macrolide efflux pump (*mef*) and (iii) alteration in ribosomal proteins L4 and L22 (encoded by *rplD* and *rplV*, respectively) (Golkar *et al.*, 2018). In *B. subtilis*, spontaneous macrolide resistance has been reported due to the alteration of ribosomal protein L22 encoded by *rplV* (Tipper *et al.*, 1977, Sharrock *et al.*, 1981, Chiba *et al.*, 2009). Based on this, the *rplV* from the genome of parental strain BS168 and mutant strains were amplified and sequenced.

Figure 5-4 shows the gel electrophoresis of *rplV* amplicons from BS168, Erm^R, T469^R and T4083^R. The DNA bands showed the size of the *rplV* derived from BS168 Erm^R (*rplV*^54D) is bigger than *rplV* derived from both BS168 T469^R and T4083^R strains (*rplV*^21D), and both *rplV* derived from both BS168 T469^R and T4083^R strains are bigger than *rplV* derived from the parental strain (*rplv*^WT).

Sequence alignment analysis revealed that the *rplV* amplicon of BS168 Erm[R] contains 54 duplication and both *rplV* amplicon of BS168 T469[R] and T4083[R] contains 21 bp duplication, which is subsequently verified by genomic data (Figure 5-5).
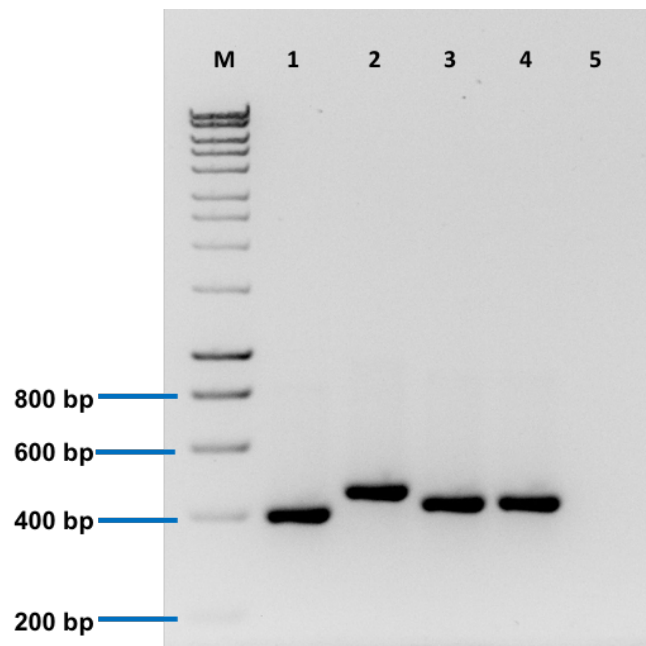
Figure 5-4 Amplification of *rplV* from *B. subtilis* BS168, Erm[R], T469[R] and T4083[R]. The PCR products of *rplV*. Lane M: HyperLadder™ 1kb (Bioline, United Kingdom); Lane 1: *rplV*[WT] of BS168 (415 bp), Lane 2: *rplV*[54D] of BS168 Erm[R] (469 bp), Lane 3: *rplV*[21D] of BS168 T469[R] (436 bp); Lane 4: *rplV*[21D] of BS168 T4083[R] (436 bp); Lane 5: negative control.

```
rplV_WT    atgcaagctaaagctgttgcaagaacagtccgtattgctcctcgtaaagcacgtctagta 60
rplV_54D   atgcaagctaaagctgttgcaagaacagtccgtattgctcctcgtaaagcacgtctagta 60
rplV_21D   atgcaagctaaagctgttgcaagaacagtccgtattgctcctcgtaaagcacgtctagta 60
           ************************************************************


rplV_WT    atggacctgattcgaggcaagcaagtaggtgaggcagtatcaatcttgaaccttacacca 120
rplV_54D   atggacctgattcgaggcaagcaagtaggtgaggcagtatcaatcttgaaccttacacca 120
rplV_21D   atggacctgattcgaggcaagcaagtaggtgaggcagtatcaatcttgaaccttacacca 120
           ************************************************************


rplV_WT    agagctgcttctccaattatcgagaaagtattaaaatccgctattgcaaatgctgagcat 180
rplV_54D   agagctgcttctccaattatcgagaaagtattaaaatccgctattgcaaatgctgagcat 180
rplv_21D   agagctgcttctccaattatcgagaaagtattaaaatccgctattgcaaatgctgagcat 180
           ************************************************************


rplV_WT    aactatgaaatggacgctaacaacctggttatttctcaagcattcgttgacgaaggccct 240
rplV_54D   aactatgaaatggacgctaacaacctggttatttctcaagcattcgttgacgaaggccct 240
rplV_21D   aactatgaaatggacgctaacaacctggttatttctcaagcattcgttgacgaaggccct 240
           ************************************************************


rplV_WT    acgttaaaaagattccgcc----------------------------------------- 219
rplV_54D   acgttaaaaagattccgcctggttatttctcaagcattcgttgacgaaggccctacgtta 300
rplV_21D   acgttaaaaagattccgcc----------------------------------------- 259
           *******************


rplV_WT    -------------cacgtgctatgggacgtgcgagccaaatcaacaaacgtacgagcc-- 283
rplV_54D   aaaagattccgcccacgtgctatgggacgtgcgagccaaatcaacaaacgtacgagcc-- 358
rplV_21D   -------------cacgtgctatgggacgtgcgagccaaatcaacaaacgtacgagccaa 306
                        ******************************************


rplV_WT    ------------------acattacaatcgttgtatcagaaaagaaggagggataaatc 324
rplV_54D   ------------------acattacaatcgttgtatcagaaaagaaggagggataaatc 399
rplV_21D   atcaacaaacgtacgagccacattacaatcgttgtatcagaaaagaaggagggataaatc 366
                            ******************************************
```

213

Figure 5-5 Sequence alignment of the wild type *rplV* (*rplV*_WT) against the mutant *rplV* from *B. subtilis* Erm^R (*rplV*_54D), T469^R (*rplV*_21D) and T4083^R (*rplV*_21D).

The nucleotides in red and italics are coincided with the tandem 54 bp duplication highlighted in cyan. The nucleotides in blue and italics are coincided with the tandem 21 bp duplication highlighted in green. The nucleotides highlighted in yellow are the start and the stop codon of the *rplV*. An * (asterisk) indicates positions which have a single, fully conserved residue.

5.3.4 Analysis of whole genome sequencing data identifies expected mutations

To fully investigate genetic variations and validate the gene mutation that confers erythromycin and TylAMac™ resistance, we compared the genome sequence of parental and mutant strains. This validates our amplicon sequence data, showing 54 bp duplication occurred within the *rplV* of BS168 Erm[R] and 21 bp duplication in both BS168 T469[R] and BS168 T4083[R] strains. For BS168 Erm[R], BS168 T469[R] and BS168 T4083[R] mutant strains, the duplication of the bases occurred at the position [22,267 to 22,320], [22,210 to 22,231] and [22,246 to 22,267] respectively. Table 5-2 summarized the *Breseq* output of the genome alignments for BS168 Erm[R], BS168 T469[R] and BS168 T4083[R] against the parental strain *B. subtilis* 168. There are two other mutations identified in BS168 T4083[R]; a point mutation in *fusA* gene and a 147 bp deletion in the *dinG_2* gene.

Table 5-2 *Breseq* output of for genome alignments of BS168 Erm^R, BS168 T469^R and BS168 T4083^R with BS168.

Strain	Position in chromosome	mutation	annotation	gene	description
BS168 Erm^R					
	523,669	C → T	substitution W261* (T**G**G-T**A**G)	JAMKKJHE_01890 ←	Hypothetical protein
	22,267	54 bp x 2	duplication (ACCAATAAAGAGTTCGTAAGCAACTGCTT CCGGGATGCAATTTTTCTAAGGCGG)$_{1 \to 2}$	*rplV* ←	50S ribosomal protein L22
BS168 T469^R					
	22,210	21 bp x 2	duplication (TACGTTTGTTGATTTGGCTCG)$_{1 \to 2}$	*rplV* ←	50S ribosomal protein L22
BS168 T4083^R					
	22,246	21 bp x 2	duplication (TACGTTTGTTGATTTGGCTCG)$_{1 \to 2}$	*rplV* ←	50S ribosomal protein L22
	29,093	G → A	substitution S415L (T**C**A→T**T**A)	*fusA* ←	Elongation factor G
	372,273	Δ 147 bp	coding (454-600/2796 nt)	*dinG_2* ←	putative ATP-dependent helicase DinG
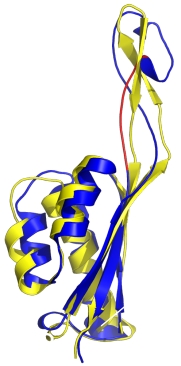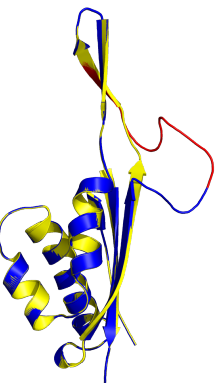
An → (arrow) to the right indicates a substitution of nucleotide (substituted nucleotide is bold and underlined in red); a symbol (x 2) and ($_{1 \to 2}$) indicates a tandem duplication of nucleotides; a symbol Δ indicates a deletion; an ← (arrow) to the left indicates the orientation of the gene.

5.3.5 Protein sequence alignment and modelling of altered L22

The *rplV* encodes for ribosomal protein L22. The mutant *rplV* [54D] of BS168 Erm[R] encodes L22 with 18 amino acid duplication (Leucine[69] to Arginine[86]) and the *rplV*[21D] of both BS168 T469[R] and BS168 T4083[R] encodes L22 with seven amino acids duplication (Serine[94] to Threonine[100]). These altered L22 are denoted as L22_18D (*rplV* [54D]) and L22_7D (*rplV* [21D]). Protein sequence alignment of the wild type L22 against L22_18D and L22_7D, showed that the duplications occurred at the region encoding the carboxyl-terminus of L22 (Figure 5-6). The duplication does not alter the reading frame and results in an insertion of amino acids with no other change.

Tandem duplications within L22 around the same region have been reported in macrolide and ketolide-resistant of *B. subtilis* and *S. aureus* (Gentry & Holmes, 2008, Chiba *et al.*, 2009, Han *et al.*, 2018). Although *rplV* mutation in *B. subtilis* correlated with macrolide resistance is infrequently described, a spontaneous erythromycin-resistant strain derived from *B. subtilis* SCB610 has been described (Chiba *et al.*, 2009). This contains altered L22 with the same seven amino acids duplication ([94]SQINKRT[100]) as our BS168 T469[R] and BS168 T4083[R] strains (L22_7D). In addition, the duplication observed in our L22_7D overlaps with mutations in L22 of *S. aureus* telithromycin-resistant mutants; KT04 (INKRTSHIT), KT05 (**R**SAINKRT), KT06 (SAINKRT) and KT09 (SRASAIN) isolated by Gentry and Holmes (2008) and *S. aureus* macrolide-resistant mutant L22[indel] (KRTSHITIV) isolated by Han *et al* (2018) (Figure 5-7).

A

B

```
L22_WT    MQAKAVARTVRIAPRKARLVMDLIRGKQVGEAVSILNLTPRAASPIIEKVLKSAIANAEH    60
L22_7D    MQAKAVARTVRIAPRKARLVMDLIRGKQVGEAVSILNLTPRAASPIIEKVLKSAIANAEH    60
          ************************************************************

L22_WT    NYEMDANNLVISQAFVDEGPTLKRFRPRAMGRASQINKRT-------SHITIVVSEKKEG   113
L22_7D    NYEMDANNLVISQAFVDEGPTLKRFRPRAMGRASQINKRTSQINKRTSHITIVVSEKKEG   120
          *****************************************         ************
```

C

D

```
L22_WT    MQAKAVARTVRIAPRKARLVMDLIRGKQVGEAVSILNLTPRAASPIIEKVLKSAIANAEH    60
L22_18D   MQAKAVARTVRIAPRKARLVMDLIRGKQVGEAVSILNLTPRAASPIIEKVLKSAIANAEH    60
          ************************************************************

L22_WT    NYEMDANNLVISQAFVDEGPTLKRFR------------------PRAMGRASQINKRTSH   102
L22_18D   NYEMDANNLVISQAFVDEGPTLKRFRLVISQAFVDEGPTLKRFRPRAMGRASQINKRTSH   120
          *************************                    ***************

L22_WT    ITIVVSEKKEG   113
L22_18D   ITIVVSEKKEG   131
          ***********
```

Figure 5-6 Protein modelling and protein sequence alignment of L22 (rpIV).

(A) Protein structure of L22 with seven amino acids duplication ($rpIV^{21D}$) (in blue) superimposed with the wild type L22 ($rpIV^{WT}$) (in yellow). The duplication region (loop) is highlighted in red. (B) amino acids sequence alignment of the wild type L22 (L22_WT) against the mutant L22_7D revealed seven amino acids repetition (^{94}SQINKRT^{100}), (C) Protein structure of L22_18D with 18 amino acids duplication ($rpIV^{54D}$) (in blue) superimposed with the wild type L22 ($rpIV^{WT}$) (in yellow). The duplication region (loop) is highlighted in red. (D) amino acids sequence alignment of the wild type L22 (L22_WT) against the mutant L22_18D revealed 18 amino acids repetition (^{69}LVISQAFVDEGPTLKRFR^{86}). An * (asterisk) indicates positions which have a single, fully conserved residue.

```
L22_7D     MQAKAVARTVRIAPRKARLVMDLIRGKQVGEAVSILNLTPRAASPIIEKVLKSAIANAEH     60
L22_KT04   MEAKAVARTIRIAPRKVRLVLDLIRGKNAAEAIAILKLTNKASSPVIEKVLMSALANAEH     60
L22_indel  MEAKAVARTIRIAPRKVRLVLDLIRGKNAAEAIAILKLTNKASSPVIEKVLMSALANAEH     60
L22_KT05   MEAKAVARTIRIAPRKVRLVLDLIRGKNAAEAIAILKLTNKASSPVIEKVLMSALANAEH     60
L22_KT06   MEAKAVARTIRIAPRKVRLVLDLIRGKNAAEAIAILKLTNKASSPVIEKVLMSALANAEH     60
L22_KT09   MEAKAVARTIRIAPRKVRLVLDLIRGKNAAEAIAILKLTNKASSPVIEKVLMSALANAEH     60
           *:*******:******.***:******:...**::**:** :*:**:***** **:*****


L22_7D     NYEMDANNLVISQAFVDEGPTLKRFRPRAMGRASQINKRTSQ--INKRTSHITIVVSEKK    118
L22_KT04   NYDMNTDELVVKEAYANEGPTLKRFRPRAQGRASAINKRTSHITINKRTSHITIVVSDGK    120
L22_indel  NYDMNTDELVVKEAYANEGPTLKRFRPRAQGRASAINKRTSHITIVKRTSHITIVVSDGK    120
L22_KT05   NYDMNTDELVVKEAYANEGPTLKRFRPRAQGRASAINKRTR-SAINKRTSHITIVVSDGK    119
L22_KT06   NYDMNTDELVVKEAYANEGPTLKRFRPRAQGRASAINKRT--SAINKRTSHITIVVSDGK    118
L22_KT09   NYDMNTDELVVKEAYANEGPTLKRFRPRAQGRASAINSRA--SAINKRTSHITIVVSDGK    118
           **:*::::**::*:.:*********** **** **.*:    * ***********: *


L22_7D     EG---- 120
L22_KT04   EEAKEA 126
L22_indel  EEAKEA 126
L22_KT05   EEAKEA 125
L22_KT06   EEAKEA 124
L22_KT09   EEAKEA 124
           *
```

Figure 5-7 Protein sequence alignment of our *B. subtilis* BS168 L22_7D (⁹⁴SQINKRT¹⁰⁰) (highlighted in yellow) with other *S. aureus* macrolide-resistant L22 mutants; KT04 (INKRTSHIT), KT05 (RSAINKRT), KT06 (SAINKRT) and KT09 (SRASAIN) (Gentry and Holmes, 2008) and L2_indel (KRTSHTIV) (Han *et al*., 2018).

All of the duplications and or insertions overlaps with each other within the conserved C-terminus region of L22 ribosomal protein. The duplicated amino acid sequences are coloured in red and highlighted in grey. The single amino acid insertion is bold and underlined in red. An * (asterisk) indicates positions which have a single, fully conserved residue; a : (colon) indicates conservation between groups of strongly similar properties; a . (period) indicates conservation between groups of weakly similar properties.

Duplication of our L22_18D ([69]LVISQAFVDEGPTLKRFR[86]) is observed to partially overlap with the duplication in other *S. aureus* telithromycin-resistant mutants; KT10 (EGPTL) and KT11 (**V**RPR) (Figure 5-8) (Gentry and Holmes, 2008). These mutants also give cross-resistance to erythromycin with MICs of 4 ug/mL and 16 ug/mL for KT10 and KT11, respectively (Gentry & Holmes, 2008). Most of the reported L22 alterations occurred within the conserved C-terminal of L22 ribosomal protein (Figure 5-9) and are invariably located at the end of the loop, near the macrolide binding site (Gregory & Dahlberg, 1999, Gabashvili *et al.*, 2001). In the macrolide-resistant *E. coli* that carry L22 with triplet deletions ([82]M**KR**[84]), two of the residues; Lys[83] and Arg[84] are found to be conserved among Gram-positive and negative bacterial L22 protein sequences (Davydova *et al.*, 2002). Interestingly, these two conserved residues are identified in part of duplicated residues ([69]LVISQAFVDEGPTL**KR**FR[86]) of our L22_18D (*rplV[54D]*) mutant, BS168 Erm[R] strain (Figure 5-9). In *E.* coli, erythromycin-resistant mutants can occur without reducing the binding affinity but neutralising the macrolides effect by increasing the width of the NPET tunnel (Gabashvili *et al.*, 2001).

```
L22_18D    MQAKAVARTVRIAPRKARLVMDLIRGKQVGEAVSILNLTPRAASPIIEKVLKSAIANAEH    60
L22_KT10   MEAKAVARTIRIAPRKVRLVLDLIRGKNAAEAIAILKLTNKASSPVIEKVLMSALANAEH    60
L22_KT11   MEAKAVARTIRIAPRKVRLVLDLIRGKNAAEAIAILKLTNKASSPVIEKVLMSALANAEH    60
           *:*******:****** .***:******:..**::**:** :*:**:***** **:*****

L22_18D    NYEMDANN**LVISQAFVDEGPTLKRFR**LVISQAFVDEGPTLKRFRPRAMGRASQINKRTSH    120
L22_KT10   NYDMNTDELVVKEAYAN**EGPTL**-------------EGPTLKRFRPRAQGRASAINKRTSH    107
L22_KT11   NYDMNTDELVVKEAYANEGPTL------------KRF-**RPRV**RPRAQGRASAINKRTSH    106
           **:*:::**..*:.:****          :       *.**** **** *******

L22_18D    ITIVVSEKKEG----    131
L22_KT10   ITIVVSDGKEEAKEA    122
L22_KT11   ITIVVSDGKEEAKEA    121
           ******: **
```

Figure 5-8 Protein sequence alignment of our L22_18D ([69]LVISQAFVDEGPTLKRFR[86]) with other previously described *S. aureus* macrolides-resistant L22 mutants; KT10 (EGPTL) and KT11 (VRP) (Gentry and Holmes, 2008).

All of the duplications and or insertions are overlapping with each other within the conserved C-terminus region of L22 ribosomal protein. The duplicated amino acid sequences are coloured in red and highlighted in grey. The single amino acid insertion is bold and underlined in red. An * (asterisk) indicates positions which have a single, fully conserved residue; a : (colon) indicates conservation between groups of strongly similar properties; a . (period) indicates conservation between groups of weakly similar properties.

```
L22_E.coli      METIAKHRHARSSAQKVRLVADLIRGKKVSQALDILTYTNKKAAVLVKKVLESAIANAEH        60
L22_BS168       MQAKAVARTVRIAPRKARLVMDLIRGKQVGEAVSILNLTPRAASPIIEKVLKSAIANAEH        60
L22_S.aureus    MEAKAVARTIRIAPRKVRLVLDLIRGKNAAEAIAILKLTNKASSPVIEKVLMSALANAEH        60
                *::  *   *   *  : :*.*** ******:...*: **. *  : :: :::*** **:*****

L22_E.coli      NDGADIDDLKVTKIFVDEGPSMKRIMPRAKGRADRILKRTSRITVVVSDR-------        110
L22_BS168       NYEMDANNLVISQAFVDEGPTLKRFRPRAMGRASQINKRTSHITIVVSEKKEG----        113
L22_S.aureus    NYDMNTDELVVKEAYANEGPTLKRFRPRAQGRASAINKRTSHITIVSDGKEEAKEA        117
                *    : ::*  :.: :.:***::**: *** ***. *  ****:**:***:

                     ___ ___ ___ ___        ___ ___ ___
                     EGPTLKRFRPRA  GRA   I  KRTSHITIVVS

                [-----------------------------------------------------------------------------------------------------------]
                                          C-terminal region
```

Figure 5-9 Comparison of the wild type L22 protein sequences of *B. subtilis* BS168, *E. coli* and *S. aureus*.

Sequences highlighted in yellow and in bold are the mutation sites of each bacteria (with or without overlap), around the conserved carboxyl-terminus region of L22 (underlined). An * (asterisk) indicates positions which have a single, fully conserved residue; a : (colon) indicates conservation between groups of strongly similar properties; a . (period) indicates conservation between groups of weakly similar properties.

The predicted protein structures of L22_18D and L22_7D were generated and compared with the wild type L22, which consists of a single domain with three alpha-helices packed against three-stranded antiparallel beta-sheet. Two of the beta sheets formed a beta-hairpin that protrudes from the core of the protein (Figure 5-10 and 5-12) (Unge *et al.*, 1998). The tip of this beta-hairpin reaches the lumen of the NPET to form part of its lining (region marked with red dotted ovoid in Figure 5-11 and 5-13). Together with the tip of L4 protein, it creates a constricted part within the NPET (Figure 5-11 and 5-13) (Ban *et al.*, 2000, Nissen *et al.*, 2000, Davydova *et al.*, 2001). Most of the *rplV* mutation that renders macrolide resistance are mapped to the beta hairpin loop (Unge *et al.*, 1998). Protein structure of L22_7D with seven amino acids duplication (*rplV*[21D]) resulted in an altered structure specifically at the extended beta hairpin loop (Figure 5-10). Superimposed structures of the wild type L22 and L22_7D showed that beta hairpin loop structure is distorted where the tip of the loop of the L22-7D is shifted towards the right of the hairpin (Figure 5-10 and 5-11). The L22_18D show high structural homology with its wild type counterpart, except for the extra 18 amino acids residues that cause a bulge in one of the beta sheet arms (labelled in red in Figure 5-12 and 5-13).
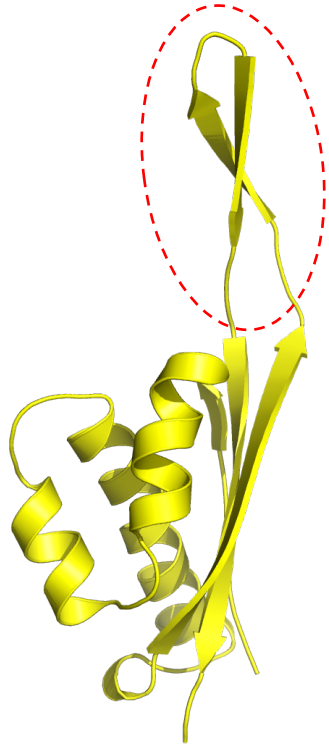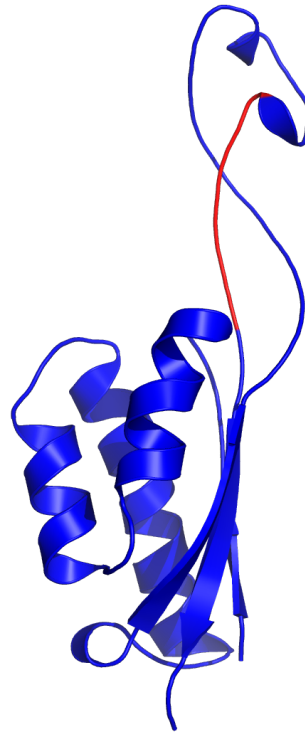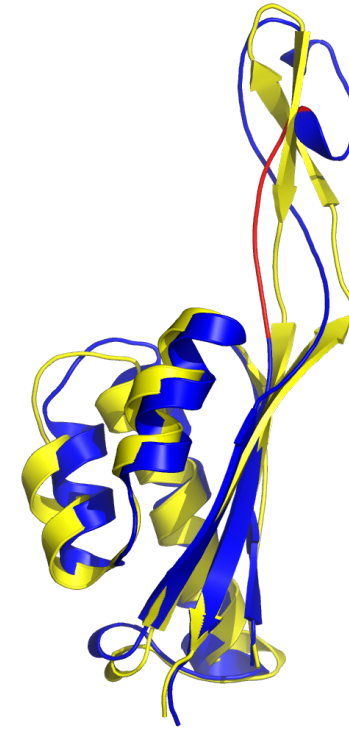
Figure 5-10 Protein structure comparison of mutant L22_7D with wild type L22.

(A) Protein structure of a wild type form L22 ($rpIV^{WT}$) (in yellow). (B) Protein structure of the mutant L22_7D with seven amino acids duplication ($rpIV^{21D}$) (in blue). (C) Protein structure of L22_7D ($rpIV^{21D}$) (in blue) superimposed with the wild type L22 ($rpIV^{WT}$) (in yellow). The duplication region (loop) is highlighted in red. The tip of this beta-hairpin that reach to the lumen of the NPET to form part of its lining is marked with red dotted ovoid.
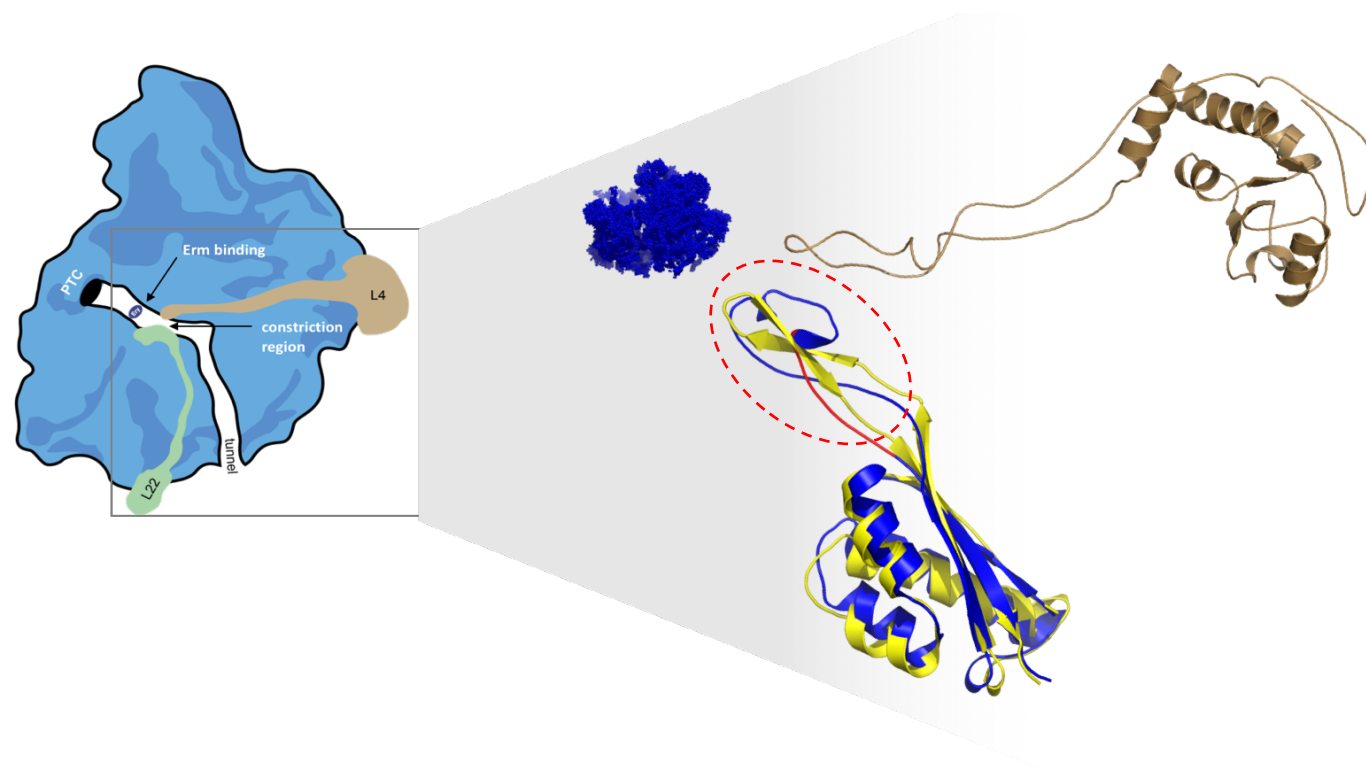
Figure 5-11 Schematic diagram showing relative position of the constricted region L4 and L22_7D within the NPET and the erythromycin binding site (blue).

Altered confirmation of L22 loop is observed due to seven amino acids repetition (^{94}SQINKRT^{100}). The altered region is shown in red and the wild type in yellow. The red dotted ovoid shows the tip of L22 beta-hairpin that reaches the lumen of the NPET to form part of its lining.
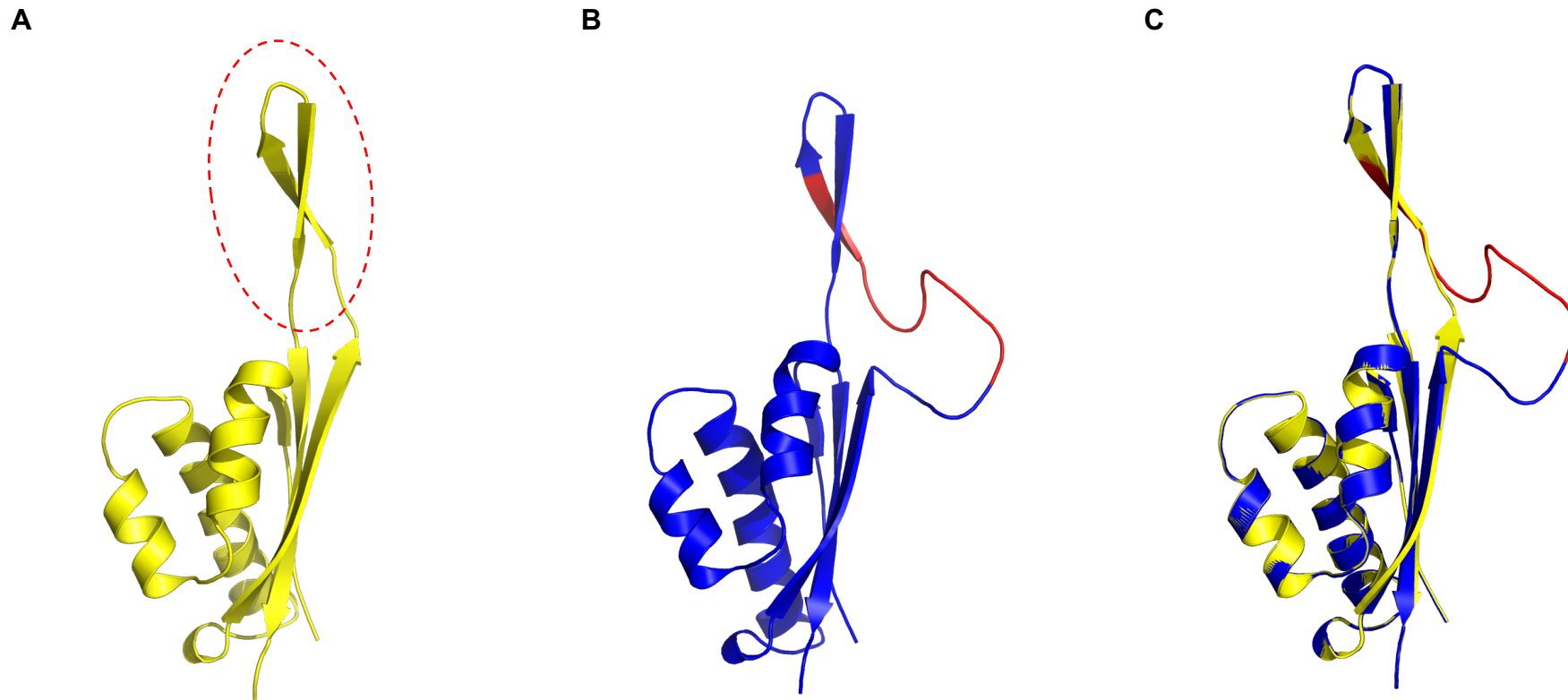
Figure 5-12 Protein structure comparison of mutant L22_18D with wild type L22.

(A) Protein structure of a wild type form L22 ($rplV^{WT}$) (in yellow). (B) Protein structure of the mutant type L22_18D with 18 amino acids duplication ($rplV^{54D}$) (in blue). (C) Protein structure of L22_18D ($rplV^{54D}$) (in blue) superimposed with the wild type L22 ($rplV^{WT}$) (in yellow). The duplication region (loop) is highlighted in red. The tip of this beta-hairpin that reach to the lumen of the NPET to form part of its lining is marked with red dotted ovoid.
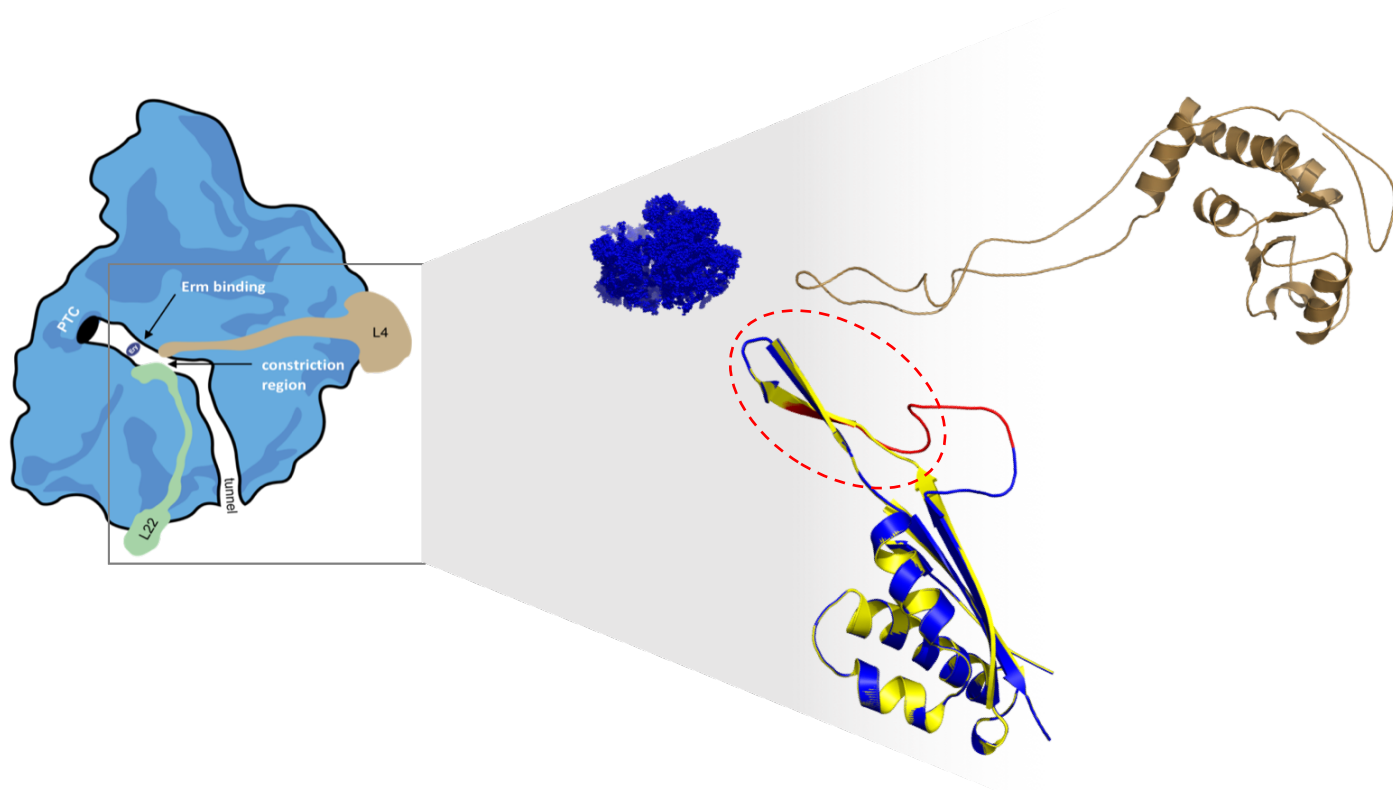
Figure 5-13 Schematic diagram showing relative position of the constricted region L4 and L22_18D within the NPET and the erythromycin binding site (blue).

Altered confirmation of L22 loop is observed due to the 18 amino acid seven amino acids repetition (^{69}LVISQAFVDEGPTLKRFR^{86}). The altered region is shown in red and the wild type in yellow. The red dotted ovoid shows the tip of L22 beta-hairpin that reaches the lumen of the NPET to form part of its lining.

5.3.6 Ectopic expression of *rplv²¹ᴰ* and *rplv⁵⁴ᴰ* confers erythromycin and tylamac resistance in *B. subtilis*

Analysis of genomic data revealed that none of the mutant isolates harbour any other known *erm* determinants or mutations in other ribosomal protein (L4), suggesting that the resistance to macrolides is likely solely attributed to the duplication occurred within the *rplV* gene. However, as there are two other mutations identified in BS168 T4083[R] and one other in BS168 Erm[R] (Table 5-2), ectopic expression of *rplv²¹ᴰ* and *rplv⁵⁴ᴰ* was done to rule out these possible factors and validate that the macrolide resistance is due to mutations in *rplV*.

The *rplV* amplicons with 21 bp duplication, 54 bp duplication and the wild type *rplV* were cloned into pHCMC04 (Nguyen *et al.*, 2005). The *B. subtilis* BS34A transformants obtained with their relevant resistance determinant were listed in Table 5-3.

Table 5-3 The *B. subtilis* BS34A transformants and their relevant resistance determinant.

Strain	Relevant resistance determinant
B. subtilis BS34A VO	Contains pHCMC04 without insert (negative control)
B. subtilis BS34A *rplV^{WT}*	Contains pHCMC04 with wild type *rplV*
B. subtilis BS34A *rplv²¹ᴰ*	Contains pHCMC04 with mutant *rplV* (21 bp duplication)
B. subtilis BS34A *rplv⁵⁴ᴰ*	Contains pHCMC04 with mutant *rplV* (54 bp duplication)

All of the plasmids were verified by sequencing. The transformants were tested for their susceptibility to erythromycin and Tylosin A analogues; TylAMac™ '469 and '4083. The transformants that carry $rplv^{21D}$ and $rplv^{54D}$ were able to grow in medium containing the antibiotics in comparison to *B. subtilis* BS34A (with wild type *rplV*). No growth was observed in all negative controls (BS34A and BS34A VO) (Figure 5-14). In comparison to BS168 Erm^R, BS168 T469^R and BS168 T4083^R mutant strains, the level of resistance in BS34A transformants (ectopic expression), particularly for BS34A $rplV^{54D}$ are lower. However, this might be due to the fact that the ectopic expression was performed with a background of ribosomes containing chromosome-encoded wild type L22. The MICs of BS34A transformants were determined by broth macrodilution method and the results is shown in Table 5-4 below. These observations confirm that the duplication in *rplV* confers erythromycin and TylAMac™ resistance in *B. subtilis.*

A Erythromycin

OD600

Strain	OD600
BS34A	0.03
BS34A::pHCMC04 VO	0.04
BS34A::pHCMC04/rplVWT	0.05
BS34A::pHCMC04/rplv21D	1.02
BS34A::pHCMC04/rplV54D	0.88
abf + 21D	2.08
abf + 54D	2.27

B TylAMac 'T4083

OD600

Strain	OD600
BS34A	0.05
BS34A::pHCMC04 VO	0.03
BS34A::pHCMC04/rplVWT	0.03
BS34A::pHCMC04/rplv21D	0.39
BS34A::pHCMC04/rplV54D	0.29
abf + 21D	1.98
abf + 54D	2.33

231
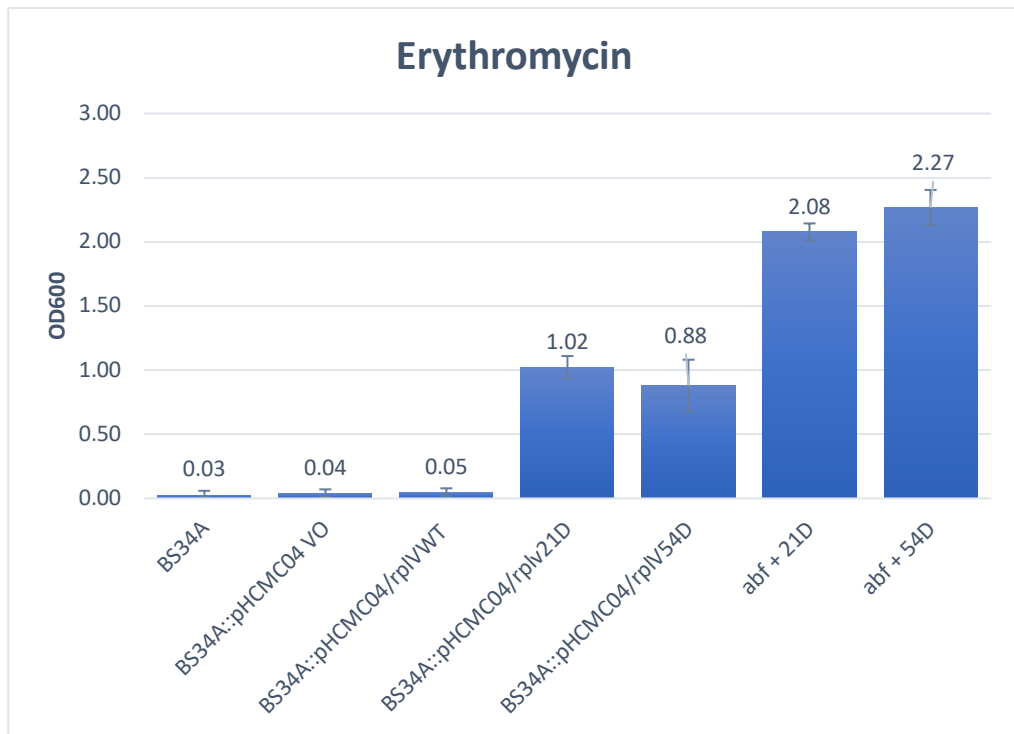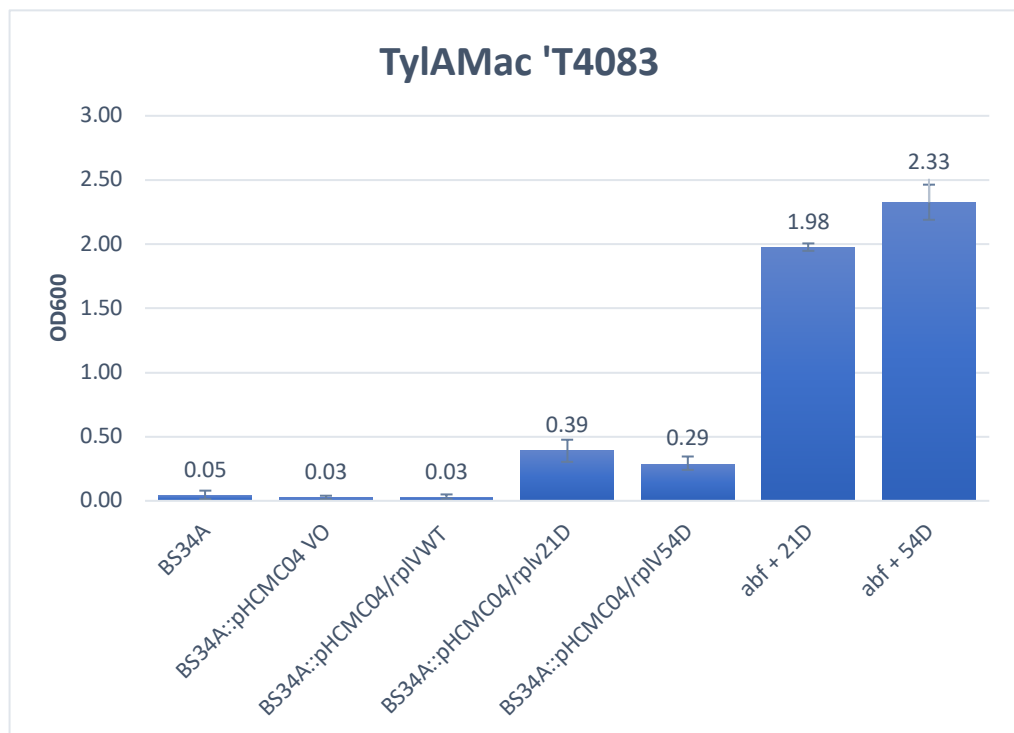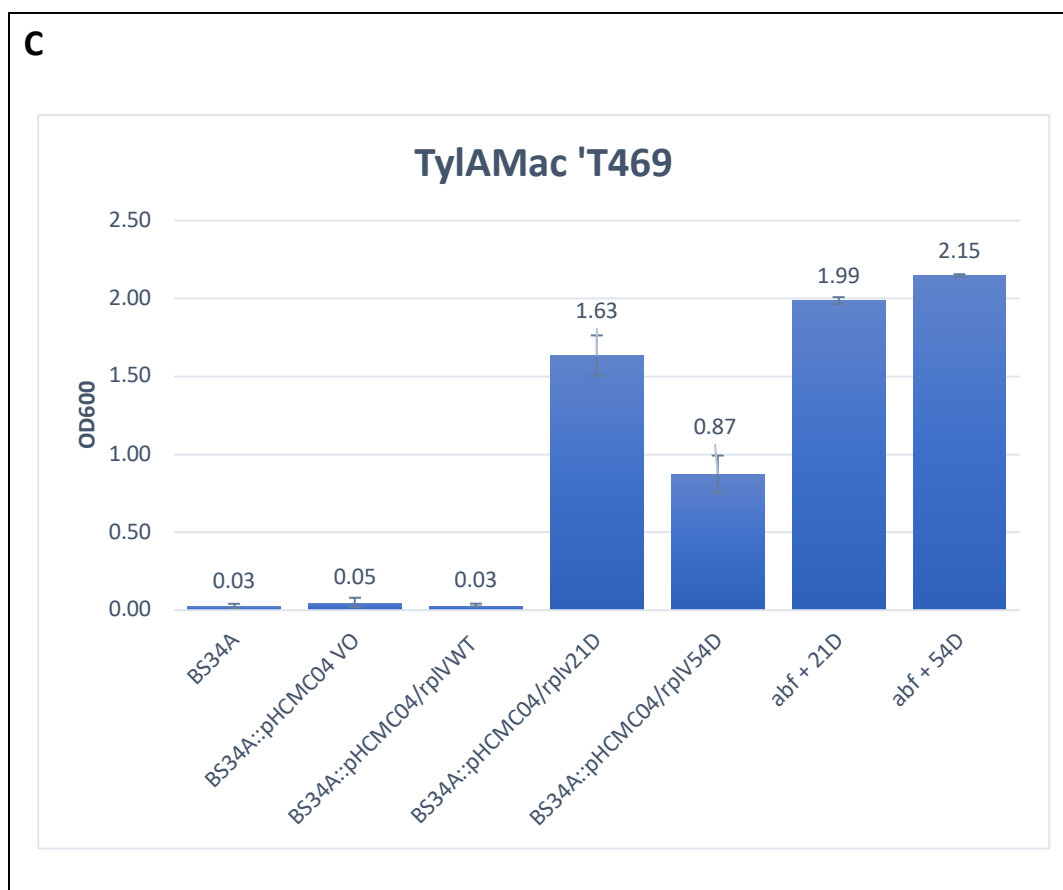
Figure 5-14 Ectopic expression of BS34A *rplV*²¹ᴰ (*rplV* with 21 bp duplication) and BS34A *rplV*⁵⁴ᴰ (*rplV* with 54 bp duplication) in comparison to BS34A *rplV*ᵂᵀ (wild type *rplV*).

All transformants were grown in CA-MHB media supplemented with (A) erythromycin (4.0 ug/mL), (B) TylAMac™ '469 (4.0 ug/mL) and (C) TylAMac™ '4083 (4.0 ug/mL), respectively. BS34A (without any vector) and BS34A VO (containing only the vector without insert) were used as a negative control. Antibiotic free (abf) broth inoculated with BS34A *rplv*²¹ᴰ (abf + 21D), and BS34A *rplV*⁵⁴ᴰ (abf + 54D) were used as a positive control. Bacteria growth was determined by reading the optical density (OD) at 600 nm after 16-18 hours incubation period.

Table 5-4 Minimum Inhibitory Concentrations (MICs) of erythromycin, TylaMac '469, TylaMac '4083 and Tylosin A for *B. subtilis* BS34A VO (containing only the vector as a negative control), BS34A *rplV*^{WT} (wild type *rplV*), BS34A *rplV*^{21D} (*rplV* with 21 bp duplication), and BS34A *rplV*^{54D} (*rplV* with 54 bp duplication). Determined from three independent experiments using broth macrodilution techniques with a range of antibiotics set at 0.5 – 8.0 µg/mL.

Strain	Relevant resistance determinant	MIC (µg/ml)		
		TylAMac™ '469	TylAMac™ '4083	Erythromycin
BS34A: VO	Contains pHCMC04 without insert (negative control)	0.5	0.5	0.5
BS34A *rplV*^{WT}	Contains pHCMC04 with wild type *rplV*	0.5	0.5	0.5
BS34A *rplv*^{21D}	Contains pHCMC04 with mutant *rplV* (21 bp duplication)	8.0	4.0	8.0
BS34A *rplV*^{54D}	Contains pHCMC04 with mutant *rplV* (54 duplication)	8.0	8.0	8.0

5.4 Discussion

Our study demonstrates a novel mutation in macrolide-resistant *B. subtilis* due to 54 bp tandem duplication in *rplV* encoding the ribosomal protein L22 (BS168 Erm[R]). Another two of our selected macrolide-resistant mutants BS168 T469[R] and BS168 T4083[R] (L22_7D) contain a 21 bp duplication as described previously in *B. subtilis* erythromycin resistant-SCB610 mutant strain (Chiba *et al.*, 2009).

The BS168 Erm[R] mutants formed distinctive heteromorphic colonies in comparison to the parental strain when grown on plates supplemented with antibiotics. This is not the case for the other two mutants, BS168 T469[R] and BS168 T4083[R]. It is also interesting to note that, similar colony morphology is observed in the BS34A strain that carry the mutant construct; pHCMC04/*rplV*[54D], and not in the ones carrying the pHCMC04/*rplV*[21D] or the pHCMC04/*rplV*[WT]. This suggest that the distinctive changes are attributed to the L22_18D (Figure 5-15).
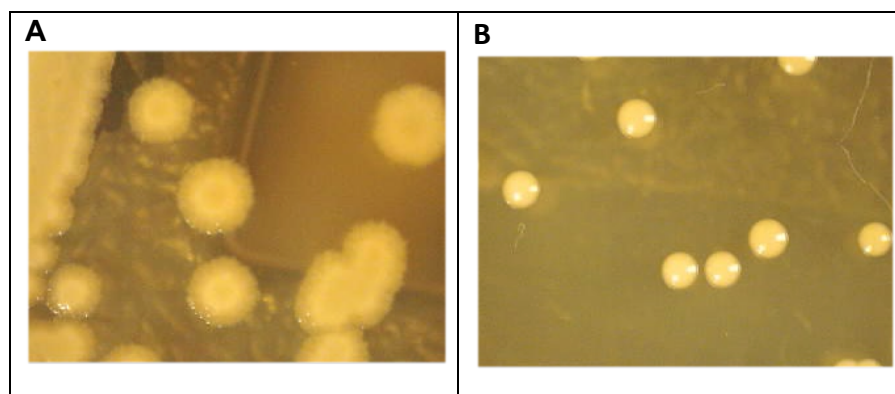
Figure 5-15 Colony morphology is altered in erythromycin resistant *B. subtilis* BS34A *rplV*[54D].

(A) BS34A *rplV*[WT] on antibiotic free LB agar, (B) BS34A *rplV*[54D] on LB agar supplemented with erythromycin.

Based on genomic data analysis, two additional mutations were identified within the genome of BS168 T4083[R] which are in *fusA* and *dinG_2* gene. The *fusA* encodes the elongation factor G (EF-G) that functions in shifting the nascent polypeptide chain from the A site to the P site within 30S subunit during protein synthesis (Fernandes, 2016). Fusidic acid inhibits protein synthesis by binding to EF-G resulting in the inhibition of peptide translocation. Mutation in *fusA* caused alteration in EF-G that confer resistance to fusidic acid (Farrell *et al.*, 2011, Fernandes, 2016). The *dinG_2* gene encodes ATP-dependent DNA helicase. In *E.coli*, dinG is a damage-inducible SOS-regulated gene encoding for a superfamily 2 DNA helicase (Voloshin & Camerini-Otero, 2007).

Our study revealed that differences in the size and sites of mutations might be correlated with the MICs. This is supported by the fact that both BS168 T469[R] and T4083[R] strains that carry the same duplication (21 bp) at exactly the same site shared similar MICs range to all antibiotics. While the BS168 erm[R] strain that carry 54 duplication at a different site within their *rplV* showing a higher MIC in comparison to BS168 T469[R] and T4083[R] strains. A Cryo-EM study of the ribosome structure of erythromycin-resistant *E.coli* has revealed that the ability of ribosome to bind erythromycin is correlated with the width of the NPET (Chittum and Champney, 1994, Gabashvii et al., 2001). Based on the predicted protein structure of L22_7D and L22_18D, different types of mutations caused different conformational changes in the L22 β-hairpin loop, thus explaining the possible reason why the 54 duplication and 21 duplication in *rplV* shows a different value of MIC against the same antibiotics (Figure 5-10- Fgure 5-13).

The L22 mutations observed in our *in vitro* selected mutants; BS168 Erm[R], BS168 T469[R] and BS168 T4083[R] are all located at the highly conserved region at the 3' end of *rplV*. Mutations at these sites have been described to render macrolide resistance despite the variants in the type of mutations; insertions, single amino acid substitutions, triplet deletions or tandem duplications (Chittum & Champney, 1994, Hisanaga *et al.*, 2005, Zaman *et al.*, 2007, Gentry & Holmes, 2008, Chiba *et al.*, 2009, Han *et al.*, 2018). The protein structure of both L22_7D and L22_18D were altered at the beta-hairpin loop region. Although these predictive structures of L22_7D and L22_18D is not sufficient for accurate determination of their placement within the NPET, based on the conserved region where the mutations are mapped, it can be proposed that the altered conformation of the beta hairpin loop may trigger structural rearrangements within the wall of NPET that directly or indirectly renders macrolide resistance.

Ectopic expression studies were undertaken by transforming a susceptible *B. subtilis* BS34A strain with mutated and non-mutated *rplV*. The result confirms that the seven ([94]SQINKRT[100]) and 18 ([69]LVISQAFVDEGPTLKRFR[86]) amino acids duplication of L22 renders erythromycin and TylAMac™ resistance in *B. subtilis*. The mechanism of erythromycin resistance due to L22 mutations in *B. subtilis* have not been studied but based on described structure *of Deinococcus radiodurans* and *T. thermophilus* L22 mutant, it can be suggested that at least it involves in the positional shifting of the L22 β-hairpin loop towards the inner part of NPET, triggering a cascade of changes at 23 rRNA nucleotides that leads to destabilisation of the erythromycin binding pocket affecting its binding affinity (Davydova *et al.*, 2001, Gabashvili *et al.*,

2001, Davydova *et al.*, 2002, Wekselman *et al.*, 2017). To our knowledge, the 54 bp tandem duplication found within the *rplV* of our *in vitro* selected BS168 Erm[R] mutant strain is unique and different to any other previous reported mutations related to macrolide resistance.

5.5   Conclusions

Erythromycin and TylAMac-resistant mutants have been selected in *B. subtilis*. Comparative genome analysis using BRESEQ revealed 54 bp and 21 bp duplication within *rplV* of these mutants in comparison to the parental strain, *B. subtilis* 168. The *rplV* encodes a large ribosomal subunit protein, L22. Alignment of the L22_7D and L22_18D with the wild type L22 protein showed seven amino acids and 18 amino acids repetition, respectively. Based on our predicted structural protein modelling, tandem duplication within L22 has led to rearrangement in *rplV* β-hairpin loop, leading to conformational changes in exit tunnel that renders macrolide resistance. The BS168 Erm[R] and BS168 T4083[R] mutant strains grew slower than the wild type and the 54 duplication in *rplV* have also resulted in changes of the colony morphology when grown on agar supplemented with erythromycin. Ectopic expression of the *rplV* mutant constructs containing 21 or 54 duplication in *B. subtilis* BS34A confers resistance against erythromycin and TylaMac to its host**.** This is the first observation of macrolide resistance due to 54 nucleotide duplication within the *rplV* in *B. subtilis*. This will also be the key to determine the mode of action of the new Tylosin A analogues; TylAMac™ '469 and '4083.

6   Final Conclusions and Future Work

This study was conducted to provide better insights into the molecular mechanisms of resistance in *B. subtilis*. The first two chapters focused on AR mediated by Tn*916* and Tn*916*-like elements. Almost all Tn*916* and Tn*916*-like elements carry *tet*(M) that confers tetracycline resistance. They are common which makes them primary vectors in spreading *tet*(M) among a broad host range of bacteria including clinically relevant pathogens. Therefore, investigation into the molecular basis of regulation and movement of these mobile elements is crucial to gain an insight for controlling the dissemination of antibiotic resistance genes.

In Chapter 3, we have identified putative rho-independent terminators at the end, and upstream, of the conjugation genes of Tn*2010,* Tn*5397,* Tn*6000,* Tn*6002*, Tn*6003,* Tn*6087* and Tn*916*. Using an *in vitro* reporter system, we demonstrated that these terminators are functional except for the Tn*5397* terminator-like structure. The Tn*2010,* Tn*6000,* Tn*6002*, Tn*6003,* and Tn*6087* terminator possess the same sequence and predicted secondary structure as the Tn*916* terminator. The conserved sequence of the terminators suggest their important role in regulating the conjugation genes of the element. The efficiency of the Tn*6000* terminator is lower than Tn*916* terminator and it was observed that the differences in their efficiency is correlated with the numbers of GC pairs in the stem and uridine residues, consistent with the previously reported data. To our knowledge, this is the first time a group of conserved terminators were identified and experimentally verified.

The conjugation genes of Tn*916* are up-regulated upon the excision of the element. Therefore, we hypothesized that the presence of a terminator located

239

upstream of the conjugation module is needed as a control mechanism; preventing the transcription of their conjugation genes whilst Tn*916* is integrated in the host genome. To test this, the terminator region was cloned in the reporter construct plus either; the flanking DNA (region A: representing the linear, integrated form) or the ligated ends of Tn*916* (region B: representing the excised and circularised form). Our data showed that the enzyme activity observed is twofold higher in the end-joint construct (B), as compared to Tn*916*::BS34A genome junction construct (A) supporting our hypothesis that the efficiency of the terminator is modulated upon excision and circularization of Tn*916*, which is the exact time when Tn*916* would require expression of its conjugation genes. However, the mechanism of transcription read-through from the P*tet*(M) to the conjugation genes past the terminator remain undetermined and should be further investigated.

In order to investigate the biological function of the terminator in-vivo; a Tn*916* mutant with a deletion of the terminator was generated using *B. subtilis* BS34A as a host. The mutant cassette was generated using a novel technique where one of the homology arms is overlapping at the end of the Tn*916* element. However, we have detected a co-integration of the pGEM-T/Tn*916*ΔTerm mutant cassette plus the original copy of the Tn*916* terminator within the chromosome of our mutant cell. This contributes to the additional binding site for recombinase and the re-generation of the wild type Tn*916*. As a result, the Tn*916*ΔTerm failed to be transferred to recipient cells. Another factor that might contribute to the failure of Tn*916*ΔTerm transfer is the deletion of 11 bp of *Orf24* which is part of the terminator sequence. The specific function of

Orf24 in Tn*916* is currently unknown. However Tn*5* mutagenesis indicated that Orf24 to Orf13 involved in the conjugation machinery of the element (Senghas *et al.*, 1988). The final question remains to be answered is the biological function of the terminator and their relation to the stability of the Tn*916* and Tn*916*-like elements within their host chromosome. Therefore, future work on this subject should focus on generating a marker-less mutant, perhaps with a partial deletion of the Tn*916* terminator (excluding the sequence that overlaps with *Orf24*).

Several molecular mechanisms leading to macrolide resistance have been revealed. Spontaneous mutations can occur in the 23S rRNA leading to the decreased affinity of macrolide binding to the ribosome or in the gene encoding the ribosomal proteins L4 and L22 or by the acquisition of methyltransferases. In this study, we revealed a novel mutation that confers macrolide resistance in *B. subtilis*. The novel mutation is caused by a 54 bp duplication within *rplV* that does not caused a frame-shift mutation. Based on the predicted protein structure of the mutant L22_18D, the duplication is likely to induce conformational changes specifically at the beta-hairpin loop region. Future studies will be directed to build a crystal structure of the *B. subtilis* 50S with the duplication within the beta-hairpin of their L22 in order to determine the exact effect of the structural rearrangement that confers macrolide resistance. By generating the L22_18D-50S ribosome and erythromycin complex, perhaps the binding site of erm and the specific resistance mechanism induced by this duplication can be identified.

References

Abraham EP & Chain E (1988) An enzyme from bacteria able to destroy penicillin. 1940. *Reviews of infectious diseases* **10**: 677-678.

Alav I, Sutton JM & Rahman KM (2018) Role of bacterial efflux pumps in biofilm formation. *Journal of Antimicrobial Chemotherapy* **73**: 2003-2020.

Alcalde-Rico M, Hernando-Amado S, Blanco P & Martínez JL (2016) Multidrug efflux pumps at the crossroad between antibiotic resistance and bacterial virulence. *Frontiers in Microbiology* **7**: 1483-1483.

Aldred KJ, Kerns RJ & Osheroff N (2014) Mechanism of quinolone action and resistance. *Biochemistry* **53**: 1565-1574.

Altschul SF, Gish W, Miller W, Myers EW & Lipman DJ (1990) Basic local alignment search tool. *Journal of Molecular Biology* **215**: 403-410.

Alvarez-Elcoro S & Enzler MJ (1999) The macrolides: erythromycin, clarithromycin, and azithromycin. *Mayo Clinic proceedings* **74**: 613-634.

Alvarez-Martinez CE & Christie PJ (2009) Biological diversity of prokaryotic type IV secretion systems. *Microbiology and Molecular Biology Reviews* **73**: 775.

Anagnostopoulos C & Spizizen J (1961) Requirements for transformation in *Bacillus subtilis*. *Journal of Bacteriology* **81**: 741-746.

Andrews JM (2001) Determination of minimum inhibitory concentrations. *Journal of Antimicrobial Chemotherapy* **48 Suppl 1**: 5-16.

Antipov D, Hartwick N, Shen M, Raiko M, Lapidus A & Pevzner PA (2016) plasmidSPAdes: assembling plasmids from whole genome sequencing data. *Bioinformatics (Oxford, England)* **32**: 3380-3387.

Antonelli A, D'Andrea MM, Brenciani A, Galeotti CL, Morroni G, Pollini S, Varaldo PE & Rossolini GM (2018) Characterization of *poxtA*, a novel phenicol–oxazolidinone–tetracycline resistance gene from an MRSA of clinical origin. *Journal of Antimicrobial Chemotherapy* **73**: 1763-1769.

Aoki H, Ke L, Poppe SM, Poel TJ, Weaver EA, Gadwood RC, Thomas RC, Shinabarger DL & Ganoza MC (2002) Oxazolidinone antibiotics target the P site on *Escherichia coli* ribosomes. *Antimicrobial Agents and Chemotherapy* **46**: 1080-1085.

Arenz S, Meydan S, Starosta AL, Berninghausen O, Beckmann R, Vazquez-Laslop N & Wilson DN (2014) Drug sensing by the ribosome induces translational arrest via active site perturbation. *Molecular cell* **56**: 446-452.

Arenz S, Ramu H, Gupta P, Berninghausen O, Beckmann R, Vazquez-Laslop N, Mankin AS & Wilson DN (2014) Molecular basis for erythromycin-dependent ribosome stalling during translation of the ErmBL leader peptide. *Nature Communications* **5**: 3501.

Auchtung JM, Aleksanyan N, Bulku A & Berkmen MB (2016) Biology of ICE*Bs1*, an integrative and conjugative element in *Bacillus subtilis*. *Plasmid* **86**: 14-25.

Babic A, Berkmen MB, Lee CA & Grossman AD (2011) Efficient gene transfer in bacterial cell chains. *mBio* **2**.

Ban N, Nissen P, Hansen J, Moore PB & Steitz TA (2000) The complete atomic structure of the large ribosomal subunit at 2.4 Å resolution. *Science* **289**: 905-920.

Bankevich A, Nurk S, Antipov D*, et al.* (2012) SPAdes: a new genome assembly algorithm and its applications to single-cell sequencing. *Journal of Computational Biology* **19**: 455-477.

Bannam TL, Crellin PK & Rood JI (1995) Molecular genetics of the chloramphenicol-resistance transposon Tn*4451* from *Clostridium perfringens*: the TnpX site-specific recombinase excises a circular transposon molecule. *Molecular Microbiology* **16**: 535-551.

Barile S, Devirgiliis C & Perozzi G (2012) Molecular characterization of a novel mosaic *tet*(S/M) gene encoding tetracycline resistance in foodborne strains of *Streptococcus bovis*. *Microbiology* **158**: 2353-2362.

Belitsky BR, Janssen PJ & Sonenshein AL (1995) Sites required for GltC-dependent regulation of *Bacillus subtilis* glutamate synthase expression. *Journal of Bacteriology* **177**: 5686-5695.

Benomar S, Ranava D, Cardenas ML, Trably E, Rafrafi Y, Ducret A, Hamelin J, Lojou E, Steyer JP & Giudici-Orticoni MT (2015) Nutritional stress induces exchange of cell material and energetic coupling between bacterial species. *Nature Communications* **6**: 6283.

Berkmen MB, Lee CA, Loveday EK & Grossman AD (2010) Polar positioning of a conjugation protein from the integrative and conjugative element ICE*Bs1* of *Bacillus subtilis*. *Journal of Bacteriology* **192**: 38-45.

Bertram J, Stratz M & Durre P (1991) Natural transfer of conjugative transposon Tn*916* between gram-positive and gram-negative bacteria. *Journal of Bacteriology* **173**: 443-448.

Bhatty M, Laverde Gomez JA & Christie PJ (2013) The expanding bacterial type IV secretion lexicon. *Research in Microbiology* **164**: 620-639.

Biasini M, Bienert S, Waterhouse A*, et al.* (2014) SWISS-MODEL: modelling protein tertiary and quaternary structure using evolutionary information. *Nucleic Acids Research* **42**: W252-258.

Bingen E, Leclercq R, Fitoussi F, Brahimi N, Malbruny B, Deforche D & Cohen R (2002) Emergence of group A Streptococcus strains with different mechanisms of macrolide resistance. *Antimicrobial Agents and Chemotherapy* **46**: 1199.

Boudvillain M, Figueroa-Bossi N & Bossi L (2013) Terminator still moving forward: expanding roles for Rho factor. *Current Opinion in Microbiology* **16**: 118-124.

Brigidi P, De Rossi E, Bertarini ML, Riccardi G & Matteuzzi D (1990) Genetic transformation of intact cells of *Bacillus subtilis* by electroporation. *FEMS Microbiology Letters* **55**: 135-138.

Brodersen DE, Clemons WM, Jr., Carter AP, Morgan-Warren RJ, Wimberly BT & Ramakrishnan V (2000) The structural basis for the action of the antibiotics tetracycline, pactamycin, and hygromycin B on the 30S ribosomal subunit. *Cell* **103**: 1143-1154.

Brouwer MS, Mullany P & Roberts AP (2010) Characterization of the conjugative transposon Tn*6000* from *Enterococcus casseliflavus* 664.1H1 (formerly *Enterococcus faecium* 664.1H1). *FEMS Microbiology Letters* **309**: 71-76.

Brouwer MS, Warburton PJ, Roberts AP, Mullany P & Allan E (2011) Genetic organisation, mobility and predicted functions of genes on integrated, mobile genetic elements in sequenced strains of *Clostridium difficile*. *PLOS One* **6**: e23014.

Brouwer MSM, Roberts AP, Hussain H, Williams RJ, Allan E & Mullany P (2013) Horizontal gene transfer converts non-toxigenic *Clostridium difficile* strains into toxin producers. *Nature Communications* **4**: 2601.

Browne HP, Anvar SY, Frank J, Lawley TD, Roberts AP & Smits WK (2015) Complete genome sequence of BS49 and draft genome sequence of BS34A, *Bacillus subtilis* strains carrying Tn*916*. *FEMS Microbiology Letters* **362**: 1-4.

Bulkley D, Innis CA, Blaha G & Steitz TA (2010) Revisiting the structures of several antibiotics bound to the bacterial ribosome. *Proceedings of the National Academy of Sciences* **107**: 17158-17163.

Burdett V (1990) Nucleotide sequence of the *tet*(M) gene of Tn*916*. *Nucleic Acids Research* **18**: 6137.

Burdett V (1991) Purification and characterization of Tet(M), a protein that renders ribosomes resistant to tetracycline. *Journal of Biological Chemistry* **266**: 2872-2877.

Burkholder PR & Giles NH, Jr. (1947) Induced biochemical mutations in *Bacillus subtilis*. *American Journal of Botany* **34**: 345-348.

Burmeister AR (2015) Horizontal gene transfer. *Evolution, Medicine, and Public Health* **2015**: 193-194.

Burrus V, Pavlovic G, Decaris B & Guedon G (2002) Conjugative transposons: the tip of the iceberg. *Molecular Microbiology* **46**: 601-610.

Cagliero C, Mouline C, Cloeckaert A & Payot S (2006) Synergy between efflux pump CmeABC and modifications in ribosomal proteins L4 and L22 in conferring macrolide resistance in *Campylobacter jejuni* and *Campylobacter coli*. *Antimicrobial Agents and Chemotherapy* **50**: 3893-3896.

Canchaya C, Fournous G, Chibani-Chennoufi S, Dillmann ML & Brussow H (2003) Phage as agents of lateral gene transfer. *Current Opinion in Microbiology* **6**: 417-424.

Canu A, Malbruny B, Coquemont M, Davies TA, Appelbaum PC & Leclercq R (2002) Diversity of ribosomal mutations conferring resistance to macrolides, clindamycin, streptogramin, and telithromycin in *Streptococcus pneumoniae*. *Antimicrobial Agents and Chemotherapy* **46**: 125-131.

Cao H, van Heel AJ, Ahmed H, Mols M & Kuipers OP (2017) Cell surface engineering of *Bacillus subtilis* improves production yields of heterologously expressed alpha-amylases. *Microbial Cell Factories* **16**: 56.

Caparon MG & Scott JR (1989) Excision and insertion of the conjugative transposon Tn*916* involves a novel recombination mechanism. *Cell* **59**: 1027-1034.

Carneiro L, Yu L, Dupree P & Ward RJ (2018) Characterization of a beta-galactosidase from *Bacillus subtilis* with transgalactosylation activity. *International Journal of Biological Macromolecules* **120**: 279-287.

Castro-Sánchez E, Moore LSP, Husson F & Holmes AH (2016) What are the factors driving antimicrobial resistance? Perspectives from a public event in London, England. *BMC Infectious Diseases* **16**: 465-465.

Celli J & Trieu-Cuot P (1998) Circularization of Tn*916* is required for expression of the transposon-encoded transfer functions: characterization of long tetracycline-inducible transcripts reading through the attachment site. *Molecular Microbiology* **28**: 103-117.

Champney WS, Tober CL & Burdine R (1998) A comparison of the inhibition of translation and 50S ribosomal subunit formation in *Staphylococcus aureus* cells by nine different macrolide antibiotics. *Current Microbiology* **37**: 412-417.

Chandler JR & Dunny GM (2004) Enterococcal peptide sex pheromones: synthesis and control of biological activity. *Peptides* **25**: 1377-1388.

Chen I & Dubnau D (2004) DNA uptake during bacterial transformation. *Nature Reviews Microbiology* **2**: 241-249.

Chen I, Provvedi R & Dubnau D (2006) A macromolecular complex formed by a pilin-like protein in competent *Bacillus subtilis*. *Journal of Biological Chemistry* **281**: 21720-21727.

Chen J, Quiles-Puchalt N, Chiang YN, Bacigalupe R, Fillol-Salom A, Chee MSJ, Fitzgerald JR & Penades JR (2018) Genome hypermobility by lateral transduction. *Science* **362**: 207-212.

Chiang YN, Penadés JR & Chen J (2019) Genetic transduction by phages and chromosomal islands: The new and noncanonical. *PLOS Pathogens* **15**: e1007878.

Chiba S, Lamsa A & Pogliano K (2009) A ribosome-nascent chain sensor of membrane protein biogenesis in *Bacillus subtilis*. *The EMBO Journal* **28**: 3461-3475.

Chiba S, Kanamori T, Ueda T, Akiyama Y, Pogliano K & Ito K (2011) Recruitment of a species-specific translational arrest module to monitor different cellular processes. *Proceedings of the National Academy of Sciences* **108**: 6073.

Chittum HS & Champney WS (1994) Ribosomal protein gene sequence changes in erythromycin-resistant mutants of *Escherichia coli*. *Journal of Bacteriology* **176**: 6192-6198.

Chittum HS & Champney WS (1995) Erythromycin inhibits the assembly of the large ribosomal subunit in growing *Escherichia coli* cells. *Current Microbiology* **30**: 273-279.

Chopra I & Roberts M (2001) Tetracycline antibiotics: mode of action, applications, molecular biology, and epidemiology of bacterial resistance. *Microbiology and Molecular Biology Reviews* **65**: 232-260

Chopra I, Hawkey PM & Hinton M (1992) Tetracyclines, molecular and clinical aspects. *Journal of Antimicrobial Chemotherapy* **29**: 245-277.

Christie GE, Farnham PJ & Platt T (1981) Synthetic sites for transcription termination and a functional comparison with tryptophan operon termination sites *in vitro*. *Proceedings of the National Academy of Sciences* **78**: 4180-4184.

Christie PJ, Korman RZ, Zahler SA, Adsit JC & Dunny GM (1987) Two conjugation systems associated with *Streptococcus faecalis* plasmid pCF10: identification of a conjugative transposon that transfers between *S. faecalis* and *Bacillus subtilis*. *Journal of Bacteriology* **169**: 2529-2536.

Chukwudi CU (2016) rRNA binding sites and the molecular mechanism of action of the tetracyclines. *Antimicrobial Agents Chemotherapy* **60**: 4433-4441.

Chun J & Bae KS (2000) Phylogenetic analysis of *Bacillus subtilis* and related taxa based on partial *gyrA* gene sequences. *Antonie Van Leeuwenhoek* **78**: 123-127.

Ciampi MS (2006) Rho-dependent terminators and transcription termination. *Microbiology* **152**: 2515-2528.

Ciric L, Mullany P & Roberts AP (2011) Antibiotic and antiseptic resistance genes are linked on a novel mobile genetic element: Tn*6087*. *Journal of Antimicrobial Chemotherapy* **66**: 2235-2239.

Claverys JP & Martin B (2003) Bacterial "competence" genes: signatures of active transformation, or only remnants? *Trends in microbiology* **11**: 161-165.

Clewell DB & Gawron-Burke C (1986) Conjugative transposons and the dissemination of antibiotic resistance in streptococci. *Annual Review of Microbiology* **40**: 635-659.

Clewell DB, Flannagan SE & Jaworski DD (1995) Unconstrained bacterial promiscuity: the Tn*916*-Tn*1545* family of conjugative transposons. *Trends in Microbiology* **3**: 229-236.

Clewell DB, Jaworski DD, Flannagan SE, Zitzow LA & Su YA (1995) The conjugative transposon Tn*916* of *Enterococcus faecalis*: structural analysis and some key factors involved in movement. *Developments in Biological Standardization* **85**: 11-17.

Cochetti I, Tili E, Mingoia M, Varaldo PE & Montanari MP (2008) *erm*(B)-carrying elements in tetracycline-resistant pneumococci and correspondence between Tn*1545* and Tn*6003*. *Antimicrobial Agents and Chemotherapy* **52**: 1285-1290.

Connell SR, Tracz DM, Nierhaus KH & Taylor DE (2003) Ribosomal protection proteins and their mechanism of tetracycline resistance. *Antimicrobial Agents and Chemotherapy* **47**: 3675-3681.

Connell SR, Trieber CA, Dinos GP, Einfeldt E, Taylor DE & Nierhaus KH (2003) Mechanism of Tet(O)-mediated tetracycline resistance. *The EMBO Journal* **22**: 945-953.

Courvalin P & Carlier C (1986) Transposable multiple antibiotic resistance in *Streptococcus pneumoniae*. *Molecular & General Genetics* **205**: 291-297.

Cox G & Wright GD (2013) Intrinsic antibiotic resistance: mechanisms, origins, challenges and solutions. *International Journal of Medical Microbiology* **303**: 287-292.

Crofton J & Mitchison DA (1948) Streptomycin resistance in pulmonary tuberculosis. *British medical journal* **2**: 1009-1015.

Curcio MJ & Derbyshire KM (2003) The outs and ins of transposition: from mu to kangaroo. *Nature Reviews Molecular Cell Biology* **4**: 865-877.

d'Aubenton Carafa Y, Brody E & Thermes C (1990) Prediction of rho-independent *Escherichia coli* transcription terminators. A statistical analysis of their RNA stem-loop structures. *Journal of Molecular Biology* **216**: 835-858.

Davis AR, Gohara DW & Yap MN (2014) Sequence selectivity of macrolide-induced translational attenuation. *Proceedings of the National Academy of Sciences* **111**: 15379-15384.

Davydova N, Fedorov R, Streltsov V, Liljas A & Garber M (2001) Crystals of a mutant form of ribosomal protein L22 rendering bacterial ribosomes resistant

to erythromycin. *Acta Crystallographica Section D: Biological Crystallography* **57**: 1150-1152.

Davydova N, Streltsov V, Wilce M, Liljas A & Garber M (2002) L22 ribosomal protein and effect of its mutation on ribosome resistance to erythromycin. *Journal of Molecular Biology* **322**: 635-644.

de Kraker ME, Stewardson AJ & Harbarth S (2016) Will 10 million people die a year due to antimicrobial resistance by 2050? *PLoS medicine* **13**: e1002184.

Deatherage DE & Barrick JE (2014) Identification of mutations in laboratory-evolved microbes from next-generation sequencing data using breseq. *Methods in Molecular Biology* **1151**: 165-188.

Del Campillo-Campbell A, Kayajanian G, Campbell A & Adhya S (1967) Biotin-requiring mutants of *Escherichia coli* K-12. *Journal of Bacteriology* **94**: 2065-2066.

Del Grosso M, Camilli R, Libisch B, Fuzi M & Pantosti A (2009) New composite genetic element of the Tn*916* family with dual macrolide resistance genes in a *Streptococcus pneumoniae* isolate belonging to clonal complex 271. *Antimicrobial Agents and Chemotherapy* **53**: 1293-1294.

Delmar JA & Yu EW (2016) The AbgT family: A novel class of antimetabolite transporters. *Protein Science* **25**: 322-337.

Denis A, Agouridas C, Auger JM*, et al.* (1999) Synthesis and antibacterial activity of HMR 3647 a new ketolide highly potent against erythromycin-resistant and susceptible pathogens. *Bioorganic and Medicinal Chemistry Letters* **9**: 3075-3080.

DeWitt T & Grossman AD (2014) The bifunctional cell wall hydrolase CwlT is needed for conjugation of the integrative and conjugative element ICE*Bs1* in *Bacillus subtilis* and *B. anthracis. Journal of Bacteriology* **196**: 1588-1596.

Diaz-Torres ML, McNab R, Spratt DA, Villedieu A, Hunt N, Wilson M & Mullany P (2003) Novel tetracycline resistance determinant from the oral metagenome. *Antimicrobial Agents and Chemotherapy* **47**: 1430-1432.

Dinos GP (2017) The macrolide antibiotic renaissance. *British Journal of Pharmacology* **174**: 2967-2983.

Doktor SZ, Shortridge VD, Beyer JM & Flamm RK (2004) Epidemiology of macrolide and/or lincosamide resistant *Streptococcus pneumoniae* clinical isolates with ribosomal mutations. *Diagnostic Microbiology and Infectious Disease* **49**: 47-52.

Domingues S & Nielsen KM (2017) Membrane vesicles and horizontal gene transfer in prokaryotes. *Current Opinion in Microbiology* **38**: 16-21.

Dönhöfer A, Franckenberg S, Wickles S, Berninghausen O, Beckmann R & Wilson DN (2012) Structural basis for TetM-mediated tetracycline resistance. *Proceedings of the National Academy of Sciences* **109**: 16900.

Dubey GP & Ben-Yehuda S (2011) Intercellular nanotubes mediate bacterial communication. *Cell* **144**: 590-600.

Dubnau D (1991) Genetic competence in *Bacillus subtilis*. *Microbiological Reviews* **55**: 395-424.

Dubnau D (1999) DNA uptake in bacteria. *Annual Review of Microbiology* **53**: 217-244.

Dunkle JA, Xiong L, Mankin AS & Cate JH (2010) Structures of the *Escherichia coli* ribosome with antibiotics bound near the peptidyl transferase center explain spectra of drug action. *Proceedings of the National Academy of Sciences* **107**: 17152-17157.

Ealand CS, Machowski EE & Kana BD (2018) β-lactam resistance: The role of low molecular weight penicillin binding proteins, β-lactamases and ld-transpeptidases in bacteria associated with respiratory tract infections. *IUBMB Life* **70**: 855-868.

Ermolaeva MD, Khalak HG, White O, Smith HO & Salzberg SL (2000) Prediction of transcription terminators in bacterial genomes. *Journal of Molecular Biology* **301**: 27-33.

Errington J (1993) *Bacillus subtilis* sporulation: regulation of gene expression and control of morphogenesis. *Microbiological Reviews* **57**: 1-33.

Farnham PJ & Platt T (1980) A model for transcription termination suggested by studies on the *trp* attenuator *in vitro* using base analogs. *Cell* **20**: 739-748.

Farrell DJ, Castanheira M & Chopra I (2011) Characterization of global patterns and the genetics of fusidic acid resistance. *Clinical infectious Diseases : An Official Publication of the Infectious Diseases Society of America* **52 Suppl 7**: S487-492.

Felmingham D (2001) Microbiological profile of telithromycin, the first ketolide antimicrobial. *Clinical Microbiology and Infection* **7 Suppl 3**: 2-10.

Fernandes P (2016) Fusidic Acid: A bacterial elongation factor inhibitor for the oral treatment of acute and chronic Staphylococcal infections. *Cold Spring Harbor Perspectives in Medicine* **6**: a025437.

Flannagan SE, Zitzow LA, Su YA & Clewell DB (1994) Nucleotide sequence of the 18-kb conjugative transposon Tn*916* from *Enterococcus faecalis*. *Plasmid* **32**: 350-354.

Floss HG & Yu TW (2005) Rifamycin-mode of action, resistance, and biosynthesis. *Chemical Reviews* **105**: 621-632.

Forsberg KJ, Patel S, Wencewicz TA & Dantas G (2015) The tetracycline destructases: a novel family of tetracycline-inactivating enzymes. *Chemistry and Biology* **22**: 888-897.

Franceschi F, Kanyo Z, Sherer EC & Sutcliffe J (2004) Macrolide resistance from the ribosome perspective. *Current Drug Targets - Infectious Disorders* **4**: 177-191.

Franke AE & Clewell DB (1981) Evidence for a chromosome-borne resistance transposon (Tn*916*) in *Streptococcus faecalis* that is capable of "conjugal" transfer in the absence of a conjugative plasmid. *Journal of Bacteriology* **145**: 494-502.

Franke AE & Clewell DB (1981) Evidence for conjugal transfer of a *Streptococcus faecalis* transposon (Tn*916*) from a chromosomal site in the absence of plasmid DNA. *Cold Spring Harbor Symposia on Quantitative Biology* **45 Pt 1**: 77-80.

Frère J-M & Page MGP (2014) Penicillin-binding proteins: evergreen drug targets. *Current Opinion in Pharmacology* **18**: 112-119.

Frost LS, Leplae R, Summers AO & Toussaint A (2005) Mobile genetic elements: the agents of open source evolution. *Nature Reviews Microbiology* **3**: 722-732.

Fyfe C, Grossman TH, Kerstein K & Sutcliffe J (2016) Resistance to macrolide antibiotics in public health pathogens. *Cold Spring Harbor Perspectives in Medicine* **6**.

Gabashvili IS, Gregory ST, Valle M, Grassucci R, Worbs M, Wahl MC, Dahlberg AE & Frank J (2001) The polypeptide tunnel system in the ribosome and its gating in erythromycin resistance mutants of L4 and L22. *Molecular cell* **8**: 181-188.

Gentry DR & Holmes DJ (2008) Selection for high-level telithromycin resistance in *Staphylococcus aureus* yields mutants resulting from an *rplB*-to-*rplV* gene conversion-like event. *Antimicrobial Agents and Chemotherapy* **52**: 1156-1158.

Gill C, van de Wijgert JH, Blow F & Darby AC (2016) Evaluation of lysis methods for the extraction of bacterial DNA for analysis of the vaginal microbiota. *PLOS One* **11**: e0163148.

Goldstein BP (2014) Resistance to rifampicin: a review. *Journal of Antibiotics (Tokyo)* **67**: 625-630.

Golkar T, Zieliński M & Berghuis AM (2018) Look and outlook on enzyme-mediated macrolide resistance. *Frontiers in Microbiology* **9**: 1942-1942.

Gregory ST & Dahlberg AE (1999) Erythromycin resistance mutations in ribosomal proteins L22 and L4 perturb the higher order structure of 23 S ribosomal RNA. *Journal of Molecular Biology* **289**: 827-834.

Griffith EC, Wallace MJ, Wu Y, *et al.* (2018) The structural and functional basis for recurring sulfa drug resistance mutations in *Staphylococcus aureus* dihydropteroate synthase. *Frontiers in Microbiology* **9**: 1369.

Griffith F (1928) The significance of pneumococcal types. *Journal of Hygiene (London)* **27**: 113-159.

Grohmann E, Muth G & Espinosa M (2003) Conjugative plasmid transfer in Gram-positive bacteria. *Microbiology and Molecular Biology Reviews* **67**: 277-301.

Grossman TH (2016) Tetracycline antibiotics and resistance. *Cold Spring Harbor Perspectives in Medicine* **6**: a025387.

Gusarov I & Nudler E (1999) The mechanism of intrinsic transcription termination. *Molecular Cell* **3**: 495-504.

Hahn J, Maier B, Haijema BJ, Sheetz M & Dubnau D (2005) Transformation proteins and DNA uptake localize to the cell poles in *Bacillus subtilis*. *Cell* **122**: 59-71.

Halfon Y, Matzov D, Eyal Z, Bashan A, Zimmerman E, Kjeldgaard J, Ingmer H & Yonath A (2019) Exit tunnel modulation as resistance mechanism of *S. aureus* erythromycin resistant mutant. *Scientific Reports* **9**: 11460.

Han D, Liu Y, Li J, Liu C, Gao Y, Feng J, Lu H & Yang G (2018) Twenty-seven-nucleotide repeat insertion in the *rplV* gene confers specific resistance to macrolide antibiotics in *Staphylococcus aureus*. *Oncotarget* **9**: 26086-26095.

Hansen JL, Ippolito JA, Ban N, Nissen P, Moore PB & Steitz TA (2002) The structures of four macrolide antibiotics bound to the large ribosomal subunit. *Molecular Cell* **10**: 117-128.

Hartmann G, Honikel KO, Knusel F & Nuesch J (1967) The specific inhibition of the DNA-directed RNA synthesis by rifamycin. *Biochimica et Biophysica Acta* **145**: 843-844.

Harwood CR (1992) *Bacillus subtilis* and its relatives: molecular biological and industrial workhorses. *Trends in Biotechnology* **10**: 247-256.

Harwood CR & Wipat A (1996) Sequencing and functional analysis of the genome of *Bacillus subtilis* strain 168. *FEBS letters* **389**: 84-87.

Hassan KA, Liu Q, Elbourne LDH, *et al.* (2018) Pacing across the membrane: the novel PACE family of efflux pumps is widespread in Gram-negative pathogens. *Research in Microbiology* **169**: 450-454.

Hassanzadeh A, Barber J, Morris GA & Gorry PA (2007) Mechanism for the degradation of erythromycin A and erythromycin A 2'-ethyl succinate in acidic aqueous solution. *Journal of Physical Chemistry A* **111**: 10098-10104.

Hay SI, Rao PC, Dolecek C, Day NPJ, Stergachis A, Lopez AD & Murray CJL (2018) Measuring and mapping the global burden of antimicrobial resistance. *BMC Medicine* **16**: 78.

Hellmark B, Soderquist B & Unemo M (2009) Simultaneous species identification and detection of rifampicin resistance in staphylococci by sequencing of the *rpoB* gene. *European Journal of Clinical Microbiology & Infectious Diseases* **28**: 183-190.

Hendrix RW, Smith MC, Burns RN, Ford ME & Hatfull GF (1999) Evolutionary relationships among diverse bacteriophages and prophages: all the world's a phage. *Proceedings of the National Academy of Sciences* **96**: 2192-2197.

Hisanaga T, Hoban DJ & Zhanel GG (2005) Mechanisms of resistance to telithromycin in *Streptococcus pneumoniae*. *Journal of Antimicrobial Chemotherapy* **56**: 447-450.

Hochhut B, Marrero J & Waldor MK (2000) Mobilization of plasmids and chromosomal DNA mediated by the SXT element, a constin found in *Vibrio cholerae* O139. *Journal of Bacteriology* **182**: 2043-2047.

Hoerauf A, Specht S, Büttner M*, et al.* (2008) *Wolbachia* endobacteria depletion by doxycycline as antifilarial therapy has macrofilaricidal activity in onchocerciasis: a randomized placebo-controlled study. *Medical Microbiology and Immunology* **197**: 295-311.

Hong WD, Benayoud F, Nixon GL*, et al.* (2019) AWZ1066S, a highly specific anti-*Wolbachia* drug candidate for a short-course treatment of filariasis. *Proceedings of the National Academy of Sciences* **116**: 1414-1419.

Ikeda H & Tomizawa J-i (1965) Transducing fragments in generalized transduction by phage P1: I. Molecular origin of the fragments. *Journal of Molecular Biology* **14**: 85-109.

Jacoby GA & Carreras I (1990) Activities of beta-lactam antibiotics against *Escherichia coli* strains producing extended-spectrum beta-lactamases. *Antimicrobial Agents and Chemotherapy* **34**: 858-862.

Jameson KH & Wilkinson AJ (2017) Control of initiation of DNA Replication in *Bacillus subtilis* and *Escherichia coli*. *Genes* **8**: 22.

Jasni A (2013) Investigation into the regulation and transfer of conjugative transposons of the Tn*916*-like family. Thesis, University College London, UCL, London.

Jasni AS, Mullany P, Hussain H & Roberts AP (2010) Demonstration of conjugative transposon (Tn*5397*)-mediated horizontal gene transfer between *Clostridium difficile* and *Enterococcus faecalis*. *Antimicrobial Agents and Chemotherapy* **54**: 4924-4926.

Javed A, Christodoulou J, Cabrita LD & Orlova EV (2017) The ribosome and its role in protein folding: looking through a magnifying glass. *Acta Crystallographica. Section D, Structural Biology* **73**: 509-521.

Jaworski DD & Clewell DB (1995) A functional origin of transfer (*oriT*) on the conjugative transposon Tn*916*. *Journal of Bacteriology* **177**: 6644-6651.

Jaworski DD, Flannagan SE & Clewell DB (1996) Analyses of *traA*, *int-Tn*, and *xis-Tn* mutations in the conjugative transposon Tn*916* in *Enterococcus faecalis*. *Plasmid* **36**: 201-208.

Jevons MP (1961) "Celbenin" - resistant Staphylococci. *British medical journal* **1**: 124-125.

Johansson M, Chen J, Tsai A, Kornberg G & Puglisi JD (2014) Sequence-dependent elongation dynamics on macrolide-bound ribosomes. *Cell Reports* **7**: 1534-1546.

Johnson CM & Grossman AD (2015) Integrative and conjugative elements (ICEs): what they do and how they work. *Annual Review of Genetics* **49**: 577-601.

Johnston C, Martin B, Fichant G, Polard P & Claverys JP (2014) Bacterial transformation: distribution, shared mechanisms and divergent control. *Nature Reviews Microbiology* **12**: 181-196.

Johnston KL, Cook DAN, Berry NG*, et al.* (2017) Identification and prioritization of novel anti-*Wolbachia* chemotypes from screening a 10,000-compound diversity library. *Science Advances* **3**: eaao1551.

Kannan K, Vázquez-Laslop N & Mankin Alexander S (2012) Selective protein synthesis by ribosomes with a drug-obstructed exit tunnel. *Cell* **151**: 508-520.

Kannan K, Kanabar P, Schryer D, Florin T, Oh E, Bahroos N, Tenson T, Weissman JS & Mankin AS (2014) The general mode of translation inhibition by macrolide antibiotics. *Proceedings of the National Academy of Sciences* **111**: 15958-15963.

Kidane D & Graumann PL (2005) Intracellular protein and DNA dynamics in competent *Bacillus subtilis* cells. *Cell* **122**: 73-84.

Kim D-W, Thawng CN, Lee K, Wellington EMH & Cha C-J (2019) A novel sulfonamide resistance mechanism by two-component flavin-dependent monooxygenase system in sulfonamide-degrading actinobacteria. *Environment International* **127**: 206-215.

Krause KM, Serio AW, Kane TR & Connolly LE (2016) Aminoglycosides: An overview. *Cold Spring Harbor Perspectives in Medicine* **6**: a027029.

Kumar KM, Anbarasu A & Ramaiah S (2014) Molecular docking and molecular dynamics studies on β-lactamases and penicillin binding proteins. *Molecular BioSystems* **10**: 891-900.

Kunst F & Ogasawara N & Moszer I*, et al.* (1997) The complete genome sequence of the gram-positive bacterium *Bacillus subtilis*. *Nature* **390**: 249-256.

Lancaster H, Roberts AP, Bedi R, Wilson M & Mullany P (2004) Characterization of Tn*916S*, a Tn*916*-like element containing the tetracycline resistance determinant *tet*(S). *Journal of Bacteriology* **186**: 4395-4398.

Lang AS, Zhaxybayeva O & Beatty JT (2012) Gene transfer agents: phage-like elements of genetic exchange. *Nature Reviews Microbiology* **10**: 472-482.

Lang AS, Westbye AB & Beatty JT (2017) The distribution, evolution, and roles of gene transfer agents in prokaryotic genetic exchange. *Annual review of virology* **4**: 87-104.

Laurenceau R, Pehau-Arnaudet G, Baconnais S*, et al.* (2013) A type IV pilus mediates DNA binding during natural transformation in *Streptococcus pneumoniae*. *PLOS Pathogens* **9**: e1003473.

Lee CA, Babic A & Grossman AD (2010) Autonomous plasmid-like replication of a conjugative transposon. *Molecular Microbiology* **75**: 268-279.

Leonetti CT, Hamada MA, Laurer SJ, Broulidakis MP, Swerdlow KJ, Lee CA, Grossman AD & Berkmen MB (2015) Critical components of the conjugation machinery of the integrative and conjugative element ICE*Bs1* of *Bacillus subtilis*. *Journal of Bacteriology* **197**: 2558-2567.

Lesnik EA, Sampath R, Levene HB, Henderson TJ, McNeil JA & Ecker DJ (2001) Prediction of Rho-independent transcriptional terminators in *Escherichia coli*. *Nucleic Acids Research* **29**: 3583-3594.

Levy SB (1992) Active efflux mechanisms for antimicrobial resistance. *Antimicrobial Agents and Chemotherapy* **36**: 695-703.

Li XZ & Nikaido H (2004) Efflux-mediated drug resistance in bacteria. *Drugs* **64**: 159-204.

Ligon BL (2004) Penicillin: its discovery and early development. *Seminars in Pediatric Infectious Diseases* **15**: 52-57.

Lovmar M, Nilsson K, Lukk E, Vimberg V, Tenson T & Ehrenberg M (2009) Erythromycin resistance by L4/L22 mutations and resistance masking by drug efflux pump deficiency. *The EMBO Journal* **28**: 736-744.

Lu F & Churchward G (1994) Conjugative transposition: Tn*916* integrase contains two independent DNA binding domains that recognize different DNA sequences. *The EMBO Journal* **13**: 1541-1548.

Lunde TM, Roberts AP & Al-Haroni M (2019) Determination of copy number and circularization ratio of Tn*916*-Tn*1545* family of conjugative transposons in oral streptococci by droplet digital PCR. *Journal of Oral Microbiology* **11**: 1552060.

Lynn SP, Kasper LM & Gardner JF (1988) Contributions of RNA secondary structure and length of the thymidine tract to transcription termination at the *thr* operon attenuator. *Journal of Biological Chemistry* **263**: 472-479.

Macke TJ, Ecker DJ, Gutell RR, Gautheret D, Case DA & Sampath R (2001) RNAMotif, an RNA secondary structure definition and search algorithm. *Nucleic Acids Research* **29**: 4724-4735.

Malbruny B, Canu A, Bozdogan B, Fantin B, Zarrouk V, Dutka-Malen S, Feger C & Leclercq R (2002) Resistance to quinupristin-dalfopristin due to mutation of L22 ribosomal protein in *Staphylococcus aureus*. *Antimicrobial Agents and Chemotherapy* **46**: 2200-2207.

Manganelli R, Ricci S & Pozzi G (1996) Conjugative transposon Tn*916*: evidence for excision with formation of 5'-protruding termini. *Journal of Bacteriology* **178**: 5813-5816.

Manganelli R, Ricci S & Pozzi G (1997) The joint of Tn*916* circular intermediates is a homoduplex in *Enterococcus faecalis*. *Plasmid* **38**: 71-78.

Maravic G (2004) Macrolide resistance based on the Erm-mediated rRNA methylation. *Current Drug Targets - Infectious Disorders* **4**: 193-202.

Markley JL & Wencewicz TA (2018) Tetracycline-inactivating enzymes. *Frontiers in Microbiology* **9**: 1058-1058.

Marra D & Scott JR (1999) Regulation of excision of the conjugative transposon Tn*916*. *Molecular Microbiology* **31**: 609-621.

Marra D, Pethel B, Churchward GG & Scott JR (1999) The frequency of conjugative transposition of Tn*916* is not determined by the frequency of excision. *Journal of Bacteriology* **181**: 5414-5418.

Marrs B (1974) Genetic recombination in *Rhodopseudomonas capsulata*. *Proceedings of the National Academy of Sciences* **71**: 971-973.

Martin FH & Tinoco I, Jr. (1980) DNA-RNA hybrid duplexes containing oligo(dA:rU) sequences are exceptionally unstable and may facilitate termination of transcription. *Nucleic Acids Research* **8**: 2295-2299.

Mashburn-Warren LM & Whiteley M (2006) Special delivery: vesicle trafficking in prokaryotes. *Molecular Microbiology* **61**: 839-846.

Masters M (2000) Transduction: host DNA transfer by bacteriophages. *The Encyclopedia of Microbiology* **4**: 637-650.

Maxwell IH (1967) Partial removal of bound transfer RNA from polysomes engaged in protein synthesis *in vitro* after addition of tetracycline. *Biochimica et Biophysica Acta* **138**: 337-346.

Mc Dermott PF, Walker RD & White DG (2003) Antimicrobials: modes of action and mechanisms of resistance. *International Journal of Toxicology* **22**: 135-143.

McMurry L, Petrucci RE, Jr. & Levy SB (1980) Active efflux of tetracycline encoded by four genetically different tetracycline resistance determinants in *Escherichia coli*. *Proceedings of the National Academy of Sciences* **77**: 3974-3977.

Menninger JR (1985) Functional consequences of binding macrolides to ribosomes. *Journal of Antimicrobial Chemotherapy* **16 Suppl A**: 23-34.

Menninger JR & Otto DP (1982) Erythromycin, carbomycin, and spiramycin inhibit protein synthesis by stimulating the dissociation of peptidyl-tRNA from ribosomes. *Antimicrobial Agents and Chemotherapy* **21**: 811-818.

Michaelis M, Kleinschmidt MC, Doerr HW & Cinatl J, Jr. (2007) Minocycline inhibits West Nile virus replication and apoptosis in human neuronal cells. *Journal of Antimicrobial Chemotherapy* **60**: 981-986.

Miller JH (1972) *Experiments in molecular genetics*. Cold Spring Harbor Laboratory Press, Cold Spring Harbor, NY.

Moazed D & Noller HF (1987) Chloramphenicol, erythromycin, carbomycin and vernamycin B protect overlapping sites in the peptidyl transferase region of 23S ribosomal RNA. *Biochimie* **69**: 879-884.

Moffatt JH, Harper M, Harrison P*, et al.* (2010) Colistin resistance in *Acinetobacter baumannii* is mediated by complete loss of lipopolysaccharide production. *Antimicrobial Agents and Chemotherapy* **54**: 4971.

Moore SD & Sauer RT (2008) Revisiting the mechanism of macrolide-antibiotic resistance mediated by ribosomal protein L22. *Proceedings of the National Academy of Sciences* **105**: 18261-18266.

Morgan-Linnell SK, Becnel Boyd L, Steffen D & Zechiedrich L (2009) Mechanisms accounting for fluoroquinolone resistance in *Escherichia coli* clinical isolates. *Antimicrobial Agents and Chemotherapy* **53**: 235-241.

Morse ML, Lederberg EM & Lederberg J (1956) Transduction in *Escherichia coli* K-12. *Genetics* **41**: 142-156.

Mullany P, Wilks M & Tabaqchali S (1991) Transfer of Tn*916* and Tn*916* delta E into *Clostridium difficile*: Demonstration of a hot-spot for these elements in the *C. difficile* genome. *FEMS Microbiology Letters* **63**: 191-194.

Mullany P, Allan E & Roberts AP (2015) Mobile genetic elements in *Clostridium difficile* and their role in genome function. *Research in Microbiology* **166**: 361-367.

Mullany P, Pallen M, Wilks M, Stephen JR & Tabaqchali S (1996) A group II intron in a conjugative transposon from the Gram-positive bacterium, *Clostridium difficile*. *Gene* **174**: 145-150.

Mullany P, Wilks M, Lamb I, Clayton C, Wren B & Tabaqchali S (1990) Genetic analysis of a tetracycline resistance element from *Clostridium difficile* and its conjugal transfer to and from *Bacillus subtilis*. *Journal of General Microbiology* **136**: 1343-1349.

Mullany P, Williams R, Langridge GC*, et al.* (2012) Behavior and target site selection of conjugative transposon Tn*916* in two different strains of toxigenic *Clostridium difficile*. *Applied and Environmental Microbiology* **78**: 2147-2153.

Munita JM & Arias CA (2016) Mechanisms of antibiotic resistance. *Microbiology Spectrum* **4**.

Naeem A, Badshah SL, Muska M, Ahmad N & Khan K (2016) The current case of quinolones: Synthetic approaches and antibacterial activity. *Molecules (Basel, Switzerland)* **21**: 268-268.

Naville M, Ghuillot-Gaudeffroy A, Marchais A & Gautheret D (2011) ARNold: a web tool for the prediction of Rho-independent transcription terminators. *RNA biology* **8**: 11-13.

Nguyen F, Starosta AL, Arenz S, Sohmen D, Donhofer A & Wilson DN (2014) Tetracycline antibiotics and resistance mechanisms. *Biological Chemistry* **395**: 559-575.

Nguyen HD, Nguyen QA, Ferreira RC, Ferreira LC, Tran LT & Schumann W (2005) Construction of plasmid-based expression vectors for *Bacillus subtilis* exhibiting full structural stability. *Plasmid* **54**: 241-248.

Nicholson WL & Maughan H (2002) The spectrum of spontaneous rifampin resistance mutations in the *rpoB* gene of *Bacillus subtilis* 168 spores differs from that of vegetative cells and resembles that of *Mycobacterium tuberculosis*. *Journal of Bacteriology* **184**: 4936-4940.

Nissen P, Hansen J, Ban N, Moore PB & Steitz TA (2000) The structural basis of ribosome activity in peptide bond synthesis. *Science* **289**: 920-930.

Novais C, Freitas AR, Silveira E, Baquero F, Peixe L, Roberts AP & Coque TM (2012) A *tet*(S/M) hybrid from CTn*6000* and CTn*916* recombination. *Microbiology* **158**: 2710-2711.

O'Neill J (2016) Tackling drug-resistant infections globally: final report and recommendations. *The Review on Antimicrobial Resistance*.

Olaitan AO, Diene SM, Kempf M*, et al.* (2014) Worldwide emergence of colistin resistance in *Klebsiella pneumoniae* from healthy humans and patients in Lao PDR, Thailand, Israel, Nigeria and France owing to inactivation of the PhoP/PhoQ regulator mgrB: an epidemiological and molecular study. *International Journal of Antimicrobial Agents* **44**: 500-507.

Otaka T & Kaji A (1975) Release of (oligo) peptidyl-tRNA from ribosomes by erythromycin A. *Proceedings of the National Academy of Sciences* **72**: 2649-2652.

Padilla E, Llobet E, Doménech-Sánchez A, Martínez-Martínez L, Bengoechea JA & Albertí S (2010) *Klebsiella pneumoniae* AcrAB efflux pump contributes to antimicrobial resistance and virulence. *Antimicrobial Agents and Chemotherapy* **54**: 177.

Pande S, Shitut S, Freund L, Westermann M, Bertels F, Colesie C, Bischofs IB & Kost C (2015) Metabolic cross-feeding via intercellular nanotubes among bacteria. *Nature Communications* **6**: 6238.

Partridge SR, Kwong SM, Firth N & Jensen SO (2018) Mobile genetic elements associated with antimicrobial resistance. *Clinical Microbiology Reviews* **31**: e00088-00017.

Pasqua M, Grossi M, Zennaro A, Fanelli G, Micheli G, Barras F, Colonna B & Prosseda G (2019) The varied role of efflux pumps of the MFS Family in the interplay of bacteria with animal and plant cells. *Microorganisms* **7**: 285.

Piddock LJ (2006) Clinically relevant chromosomally encoded multidrug resistance efflux pumps in bacteria. *Clinical Microbiology Reviews* **19**: 382-402.

Piddok LJV & Mortimer PGS (1993) The accumulation of five antibacterial agents in porin-deficient mutants of *Escherichia coli*. *Journal of Antimicrobial Chemotherapy* **32**: 195-213.

Piggot PJ & Hilbert DW (2004) Sporulation of *Bacillus subtilis*. *Current Opinion in Microbiology* **7**: 579-586.

Pioletti M, Schlunzen F, Harms J*, et al.* (2001) Crystal structures of complexes of the small ribosomal subunit with tetracycline, edeine and IF3. *The EMBO Journal* **20**: 1829-1839.

Platt T (1986) Transcription termination and the regulation of gene expression. *Annual Review of Biochemistry* **55**: 339-372.

Poehlsgaard J & Douthwaite S (2005) The bacterial ribosome as a target for antibiotics. *Nature Reviews Microbiology* **3**: 870.

Poirel L, Jayol A & Nordmann P (2017) Polymyxins: antibacterial activity, susceptibility testing, and resistance mechanisms encoded by plasmids or chromosomes. *Clinical Microbiology Reviews* **30**: 557.

Poyart C, Celli J & Trieu-Cuot P (1995) Conjugative transposition of Tn*916*-related elements from *Enterococcus faecalis* to *Escherichia coli* and *Pseudomonas fluorescens*. *Antimicrobial Agents and Chemotherapy* **39**: 500-506.

Poyart-Salmeron C, Trieu-Cuot P, Carlier C & Courvalin P (1990) The integration-excision system of the conjugative transposon Tn*1545* is structurally and functionally related to those of lambdoid phages. *Molecular Microbiology* **4**: 1513-1521.

Price LB, Vogler A, Pearson T, Busch JD, Schupp JM & Keim P (2003) *In vitro* selection and characterization of *Bacillus anthracis* mutants with high-level resistance to ciprofloxacin. *Antimicrobial Agents and Chemotherapy* **47**: 2362-2365.

Provvedi R & Dubnau D (1999) ComEA is a DNA receptor for transformation of competent *Bacillus subtilis*. *Molecular Microbiology* **31**: 271-280.

Puzari M & Chetia P (2017) RND efflux pump mediated antibiotic resistance in Gram-negative bacteria *Escherichia coli* and *Pseudomonas aeruginosa*: a major issue worldwide. *World Journal of Microbiology and Biotechnology* **33**: 24.

Ramirez MS & Tolmasky ME (2010) Aminoglycoside modifying enzymes. *Drug Resistance Updates : Reviews and Commentaries in Antimicrobial and Anticancer Chemotherapy* **13**: 151-171.

Ray-Soni A, Bellecourt MJ & Landick R (2016) Mechanisms of bacterial transcription termination: all good things must end. *Annual Review of Biochemistry* **85**: 319-347.

Read AF & Woods RJ (2014) Antibiotic resistance management. *Evolution, Medicine, and Public Health* **2014**: 147.

Rice LB (1998) Tn*916* family conjugative transposons and dissemination of antimicrobial resistance determinants. *Antimicrobial Agents and Chemotherapy* **42**: 1871-1877.

Roberts AP & Mullany P (2009) A modular master on the move: the Tn*916* family of mobile genetic elements. *Trends in Microbiology* **17**: 251-258.

Roberts AP, Johanesen PA, Lyras D, Mullany P & Rood JI (2001) Comparison of Tn*5397* from *Clostridium difficile*, Tn*916* from *Enterococcus faecalis* and the CW459*tet*(M) element from *Clostridium perfringens* shows that they have similar conjugation regions but different insertion and excision modules. *Microbiology* **147**: 1243-1251.

Roberts AP, Davis IJ, Seville L, Villedieu A & Mullany P (2006) Characterization of the ends and target site of a novel tetracycline resistance-encoding conjugative transposon from *Enterococcus faecium* 664.1H1. *Journal of Bacteriology* **188**: 4356-4361.

Roberts AP, Hennequin C, Elmore M, Collignon A, Karjalainen T, Minton N & Mullany P (2003) Development of an integrative vector for the expression of antisense RNA in *Clostridium difficile*. *Journal of Microbiological Methods* **55**: 617-624.

Roberts MC (1996) Tetracycline resistance determinants: mechanisms of action, regulation of expression, genetic mobility, and distribution. *FEMS Microbiology Reviews* **19**: 1-24.

Roberts MC (2005) Update on acquired tetracycline resistance genes. *FEMS Microbiology Letters* **245**: 195-203.

Roberts MC, Sutcliffe J, Courvalin P, Jensen LB, Rood J & Seppala H (1999) Nomenclature for macrolide and macrolide-lincosamide-streptogramin B resistance determinants. *Antimicrobial Agents and Chemotherapy* **43**: 2823.

Rolain JM & Raoult D (2005) Prediction of resistance to erythromycin in the genus Rickettsia by mutations in L22 ribosomal protein. *Journal of Antimicrobial Chemotherapy* **56**: 396-398.

Roosaare M, Puustusmaa M, Mols M, Vaher M & Remm M (2018) PlasmidSeeker: identification of known plasmids from bacterial whole genome sequencing reads. *PeerJ* **6**: e4588.

Rudy C, Taylor KL, Hinerfeld D, Scott JR & Churchward G (1997) Excision of a conjugative transposon *in vitro* by the Int and Xis proteins of Tn*916*. *Nucleic Acids Research* **25**: 4061-4066.

Rudy CK, Scott JR & Churchward G (1997) DNA binding by the Xis protein of the conjugative transposon Tn*916*. *Journal of Bacteriology* **179**: 2567-2572.

Salzer R, Kern T, Joos F & Averhoff B (2016) The *Thermus thermophilus* comEA/comEC operon is associated with DNA binding and regulation of the DNA translocator and type IV pili. *Environmental Microbiology* **18**: 65-74.

Sanchez-Pescador R, Brown JT, Roberts M & Urdea MS (1988) Homology of the TetM with translational elongation factors: implications for potential modes of *tet*M-conferred tetracycline resistance. *Nucleic Acids Research* **16**: 1218.

Santoro F, Vianna ME & Roberts AP (2014) Variation on a theme; an overview of the Tn*916*/Tn*1545* family of mobile genetic elements in the oral and nasopharyngeal streptococci. *Frontiers in Microbiology* **5**: 535.

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

Schroeder GN & Hilbi H (2008) Molecular pathogenesis of *Shigella* spp.: controlling host cell signaling, invasion, and death by type III secretion. *Clinical Microbiology Reviews* **21**: 134-156.

Schwarz S, Kehrenberg C, Doublet B & Cloeckaert A (2004) Molecular basis of bacterial resistance to chloramphenicol and florfenicol. *FEMS Microbiology Reviews* **28**: 519-542.

Scornec H, Bellanger X, Guilloteau H, Groshenry G & Merlin C (2017) Inducibility of Tn*916* conjugative transfer in *Enterococcus faecalis* by

subinhibitory concentrations of ribosome-targeting antibiotics. *Journal of Antimicrobial Chemotherapy* **72**: 2722-2728.

Scott JR & Churchward GG (1995) Conjugative transposition. *Annual Review of Microbiology* **49**: 367-397.

Scott JR, Bringel F, Marra D, Van Alstine G & Rudy CK (1994) Conjugative transposition of Tn*916*: preferred targets and evidence for conjugative transfer of a single strand and for a double-stranded circular intermediate. *Molecular Microbiology* **11**: 1099-1108.

Seemann T (2014) Prokka: rapid prokaryotic genome annotation. *Bioinformatics (Oxford, England)* **30**: 2068-2069.

Seier-Petersen MA, Jasni A, Aarestrup FM, Vigre H, Mullany P, Roberts AP & Agerso Y (2014) Effect of subinhibitory concentrations of four commonly used biocides on the conjugative transfer of Tn*916* in *Bacillus subtilis*. *Journal of Antimicrobial Chemotherapy* **69**: 343-348.

Senghas E, Jones JM, Yamamoto M, Gawron-Burke C & Clewell DB (1988) Genetic organization of the bacterial conjugative transposon Tn*916*. *Journal of Bacteriology* **170**: 245-249.

Sengupta S, Chattopadhyay MK & Grossart HP (2013) The multifaceted roles of antibiotics and antibiotic resistance in nature. *Frontiers in Microbiology* **4**: 47.

Sharrock RA, Leighton T & Wittmann HG (1981) Macrolide and aminoglycoside antibiotic resistance mutations in the *Bacillus subtilis* ribosome resulting in temperature-sensitive sporulation. *Molecular and General Genetics MGG* **183**: 538-543.

Skinner R, Cundliffe E & Schmidt FJ (1983) Site of action of a ribosomal RNA methylase responsible for resistance to erythromycin and other antibiotics. *Journal of Biological Chemistry* **258**: 12702-12706.

Sköld O (2000) Sulfonamide resistance: mechanisms and trends. *Drug Resistance Updates* **3**: 155-160.

Smeets LC & Kusters JG (2002) Natural transformation in *Helicobacter pylori*: DNA transport in an unexpected way. *Trends in microbiology* **10**: 159-162; discussion 162.

Smillie C, Garcillan-Barcia MP, Francia MV, Rocha EP & de la Cruz F (2010) Mobility of plasmids. *Microbiology and Molecular Biology Reviews* **74**: 434-452.

Soge OO, Beck NK, White TM, No DB & Roberts MC (2008) A novel transposon, Tn*6009*, composed of a Tn*916* element linked with a *Staphylococcus aureus mer* operon. *Journal of Antimicrobial Chemotherapy* **62**: 674-680.

Solioz M & Marrs B (1977) The gene transfer agent of *Rhodopseudomonas capsulata*. Purification and characterization of its nucleic acid. *Archives of Biochemistry and Biophysics* **181**: 300-307.

Solioz M, Yen HC & Marris B (1975) Release and uptake of gene transfer agent by *Rhodopseudomonas capsulata*. *Journal of Bacteriology* **123**: 651-657.

Solomon JM & Grossman AD (1996) Who's competent and when: regulation of natural genetic competence in bacteria. *Trends in Genetics* **12**: 150-155.

Sothiselvam S, Neuner S, Rigger L, Klepacki D, Micura R, Vazquez-Laslop N & Mankin AS (2016) Binding of macrolide antibiotics leads to ribosomal selection against specific substrates based on their charge and size. *Cell Reports* **16**: 1789-1799.

Sothiselvam S, Liu B, Han W*, et al.* (2014) Macrolide antibiotics allosterically predispose the ribosome for translation arrest. *Proceedings of the National Academy of Sciences* **111**: 9804-9809.

Spizizen J (1958) Transformation of biochemically deficient strains of *Bacillus subtilis* by deoxyribonucleate. *Proceedings of the National Academy of Sciences* **44**: 1072-1078.

Spratt BG (1994) Resistance to antibiotics mediated by target alterations. *Science* **264**: 388.

Srinivasan VB & Rajamohan G (2013) KpnEF, a new member of the *Klebsiella pneumoniae* cell envelope stress response regulon, is an SMR-type efflux pump involved in broad-spectrum antimicrobial resistance. *Antimicrobial Agents and Chemotherapy* **57**: 4449.

Stanton TB & Humphrey SB (2003) Isolation of tetracycline-resistant *Megasphaera elsdenii* strains with novel mosaic gene combinations of *tet*(O) and *tet*(W) from swine. *Applied and Environmental Microbiology* **69**: 3874-3882.

Storrs MJ, Poyart-Salmeron C, Trieu-Cuot P & Courvalin P (1991) Conjugative transposition of Tn*916* requires the excisive and integrative activities of the transposon-encoded integrase. *Journal of Bacteriology* **173**: 4347-4352.

Strahilevitz J, Jacoby GA, Hooper DC & Robicsek A (2009) Plasmid-mediated quinolone resistance: a multifaceted threat. *Clinical Microbiology Reviews* **22**: 664-689.

Stroynowski I, Kuroda M & Yanofsky C (1983) Transcription termination *in vitro* at the tryptophan operon attenuator is controlled by secondary structures in the leader transcript. *Proceedings of the National Academy of Sciences of the United States of America* **80**: 2206-2210.

Su YA, He P & Clewell DB (1992) Characterization of the *tet*(M) determinant of Tn*916*: evidence for regulation by transcription attenuation. *Antimicrobial Agents and Chemotherapy* **36**: 769-778.

Sujatha S & Praharaj I (2012) Glycopeptide resistance in Gram-positive cocci: A review. *Interdisciplinary Perspectives on Infectious Diseases* **2012**: 781679.

Sun D (2018) Pull in and push out: mechanisms of horizontal gene transfer in bacteria. *Frontiers in Microbiology* **9**: 2154.

Taylor MJ, von Geldern TW, Ford L*, et al.* (2019) Preclinical development of an oral anti-*Wolbachia* macrolide drug for the treatment of lymphatic filariasis and onchocerciasis. *Science Translational Medicine* **11**.

Tenson T & Mankin AS (2001) Short peptides conferring resistance to macrolide antibiotics. *Peptides* **22**: 1661-1668.

Tenson T & Ehrenberg M (2002) Regulatory nascent peptides in the ribosomal tunnel. *Cell* **108**: 591-594.

Tenson T, Lovmar M & Ehrenberg M (2003) The mechanism of action of macrolides, lincosamides and streptogramin B reveals the nascent peptide exit path in the ribosome. *Journal of Molecular Biology* **330**: 1005-1014.

Thierauf A, Perez G & Maloy AS (2009) Generalized transduction. *Methods in Molecular Biology* **501**: 267-286.

Thomas CM & Nielsen KM (2005) Mechanisms of, and barriers to, horizontal gene transfer between bacteria. *Nature Reviews Microbiology* **3**: 711-721.

Tipper DJ, Johnson CW, Ginther CL, Leighton T & Wittmann HG (1977) Erythromycin resistant mutations in *Bacillus subtilis* cause temperature sensitive sporulation. *Molecular and General Genetics MGG* **150**: 147-159.

Tolmasky ME (2000) Bacterial resistance to aminoglycosides and beta-lactams: the Tn*1331* transposon paradigm. *Frontiers in Bioscience* **5**: D20-29.

Tosato V & Bruschi CV (2004) Knowledge of the *Bacillus subtilis* genome: impacts on fundamental science and biotechnology. *Applied Microbiology and Biotechnology* **64**: 1-6.

Toyofuku M, Nomura N & Eberl L (2019) Types and origins of bacterial membrane vesicles. *Nature Reviews Microbiology* **17**: 13-24.

Tu D, Blaha G, Moore PB & Steitz TA (2005) Structures of MLSBK antibiotics bound to mutated large ribosomal subunits provide a structural explanation for resistance. *Cell* **121**: 257-270.

Unge J, berg A, Al-Kharadaghi S, Nikulin A, Nikonov S, Davydova N, Nevskaya N, Garber M & Liljas A (1998) The crystal structure of ribosomal protein L22 from *Thermus thermophilus*: insights into the mechanism of erythromycin resistance. *Structure* **6**: 1577-1586.

Usary J & Champney WS (2001) Erythromycin inhibition of 50S ribosomal subunit formation in *Escherichia coli* cells. *Molecular Microbiology* **40**: 951-962.

van Sinderen D, Luttinger A, Kong L, Dubnau D, Venema G & Hamoen L (1995) *comK* encodes the competence transcription factor, the key regulatory protein for competence development in *Bacillus subtilis. Molecular Microbiology* **15**: 455-462.

Vazquez-Laslop N, Thum C & Mankin AS (2008) Molecular mechanism of drug-dependent ribosome stalling. *Molecular Cell* **30**: 190-202.

Vázquez-Laslop N & Mankin AS (2018) How macrolide antibiotics work. *Trends in Biochemical Sciences* **43**: 668-684.

Ventola CL (2015) The antibiotic resistance crisis: part 1: causes and threats. *P & T: A Peer-Reviewed Journal for Formulary Management* **40**: 277-283.

Vojcic L, Despotovic D, Martinez R, Maurer KH & Schwaneberg U (2012) An efficient transformation method for *Bacillus subtilis* DB104. *Applied Microbiology and Biotechnology* **94**: 487-493.

Voloshin ON & Camerini-Otero RD (2007) The DinG protein from *Escherichia coli* is a structure-specific helicase. *Journal of Biological Chemistry* **282**: 18437-18447.

von Geldern TW, Morton HE, Clark RF*, et al.* (2019) Discovery of ABBV-4083, a novel analog of Tylosin A that has potent anti-*Wolbachia* and anti-filarial activity. *PLOS Neglected Tropical Diseases* **13**: e0007159.

von Hippel PH & Yager TD (1992) The elongation-termination decision in transcription. *Science* **255**: 809-812.

von Wintersdorff CJ, Penders J, van Niekerk JM, Mills ND, Majumder S, van Alphen LB, Savelkoul PH & Wolffs PF (2016) Dissemination of antimicrobial resistance in microbial ecosystems through horizontal gene transfer. *Frontiers in Microbiology* **7**: 173.

Wang H & Mullany P (2000) The large resolvase TndX is required and sufficient for integration and excision of derivatives of the novel conjugative transposon Tn*5397*. *Journal of Bacteriology* **182**: 6577-6583.

Wang H, Roberts AP & Mullany P (2000) DNA sequence of the insertional hot spot of Tn*916* in the *Clostridium difficile* genome and discovery of a Tn*916*-like element in an environmental isolate integrated in the same hot spot. *FEMS Microbiology Letters* **192**: 15-20.

Wang H, Smith MCM & Mullany P (2006) The conjugative transposon Tn*5397* has a strong preference for integration into its *Clostridium difficile* target site. *Journal of Bacteriology* **188**: 4871-4878.

Wang H, Roberts AP, Lyras D, Rood JI, Wilks M & Mullany P (2000) Characterization of the ends and target sites of the novel conjugative transposon Tn*5397* from *Clostridium difficile*: excision and circularization is mediated by the large resolvase, TndX. *Journal of Bacteriology* **182**: 3775-3783.

Wang Y, Lv Y, Cai J*, et al.* (2015) A novel gene, optrA, that confers transferable resistance to oxazolidinones and phenicols and its presence in *Enterococcus faecalis* and *Enterococcus faecium* of human and animal origin. *Journal of Antimicrobial Chemotherapy* **70**: 2182-2190.

Warburton PJ, Amodeo N & Roberts AP (2016) Mosaic tetracycline resistance genes encoding ribosomal protection proteins. *Journal of Antimicrobial Chemotherapy* **71**: 3333-3339.

Warburton PJ, Palmer RM, Munson MA & Wade WG (2007) Demonstration of in vivo transfer of doxycycline resistance mediated by a novel transposon. *Journal of Antimicrobial Chemotherapy* **60**: 973-980.

Watanabe T (1963) Infective heredity of multiple drug resistance in bacteria. *Bacteriological reviews* **27**: 87-115.

Weisblum B (1995) Insights into erythromycin action from studies of its activity as inducer of resistance. *Antimicrobial Agents and Chemotherapy* **39**: 797-805.

Wekselman I, Zimmerman E, Davidovich C*, et al.* (2017) The ribosomal protein uL22 modulates the shape of the protein exit tunnel. *Structure* **25**: 1233-1241.e1233.

Whittle G, Hamburger N, Shoemaker NB & Salyers AA (2006) A *Bacteroides* conjugative transposon, CTnERL, can transfer a portion of itself by conjugation without excising from the chromosome. *Journal of Bacteriology* **188**: 1169-1174.

Wilson DN (2016) The ABC of ribosome-related antibiotic resistance. *mBio* **7**: e00598-00516.

Wilson KS & von Hippel PH (1995) Transcription termination at intrinsic terminators: the role of the RNA hairpin. *Proceedings of the National Academy of Sciences of the United States of America* **92**: 8793-8797.

Wittmann HG, Stöffler G, Apirion D, Rosen L, Tanaka K, Tamaki M, Takata R, Dekio S, Otaka E & Osawa S (1973) Biochemical and genetic studies on two different types of erythromycin resistant mutants of *Escherichia coli* with altered ribosomal proteins. *Molecular and General Genetics MGG* **127**: 175-189.

Woodford N & Ellington MJ (2007) The emergence of antibiotic resistance by mutation. *Clinical Microbiology and Infection* **13**: 5-18.

World Health Organization (2014) Antimicrobial resistance: global report on surveillance. *WHO: Geneva, Switzerland*.

Wozniak RA & Waldor MK (2010) Integrative and conjugative elements: mosaic mobile genetic elements enabling dynamic lateral gene flow. *Nature Reviews Microbiology* **8**: 552-563.

Wright GD (2005) Bacterial resistance to antibiotics: enzymatic degradation and modification. *Advanced Drug Delivery Reviews* **57**: 1451-1470.

Wright LD & Grossman AD (2016) Autonomous replication of the conjugative transposon Tn*916*. *Journal of Bacteriology* **198**: 3355-3366.

Xu M, Zhou YN, Goldstein BP & Jin DJ (2005) Cross-resistance of *Escherichia coli* RNA polymerases conferring rifampin resistance to different antibiotics. *Journal of Bacteriology* **187**: 2783-2792.

Yang W, Moore IF, Koteva KP, Bareich DC, Hughes DW & Wright GD (2004) TetX is a flavin-dependent monooxygenase conferring resistance to tetracycline antibiotics. *Journal of Biological Chemistry* **279**: 52346-52352.

Yanouri A, Daniel RA, Errington J & Buchanan CE (1993) Cloning and sequencing of the cell division gene *pbpB,* which encodes penicillin-binding protein 2B in *Bacillus subtilis*. *Journal of Bacteriology* **175**: 7604-7616.

Yelin I & Kishony R (2018) Antibiotic resistance. *Cell* **172**: 1136-1136.e1131.

Yen HC, Hu NT & Marrs BL (1979) Characterization of the gene transfer agent made by an overproducer mutant of *Rhodopseudomonas capsulata*. *Journal of Molecular Biology* **131**: 157-168.

Young FE & Spizizen J (1963) Incorporation of deoxyribonucleic acid in the *Bacillus subtilis* transformation system. *Journal of Bacteriology* **86**: 392-400.

Young ML, Bains M, Bell A & Hancock RE (1992) Role of *Pseudomonas aeruginosa* outer membrane protein OprH in polymyxin and gentamicin resistance: isolation of an OprH-deficient mutant by gene replacement techniques. *Antimicrobial Agents and Chemotherapy* **36**: 2566.

Zaman S, Fitzpatrick M, Lindahl L & Zengel J (2007) Novel mutations in ribosomal proteins L4 and L22 that confer erythromycin resistance in *Escherichia coli*. *Molecular Microbiology* **66**: 1039-1050.

Zankari E (2014) Comparison of the web tools ARG-ANNOT and ResFinder for detection of resistance genes in bacteria. *Antimicrobial Agents and Chemotherapy* **58**: 4986.

Zechner EL, Lang S & Schildbach JF (2012) Assembly and mechanisms of bacterial type IV secretion machines. *Philosophical Transactions of the Royal Society of London Series B Biological Sciences* **367**: 1073-1087.

Zeng D, Debabov D, Hartsell TL, Cano RJ, Adams S, Schuyler JA, McMillan R & Pace JL (2016) Approved glycopeptide antibacterial drugs: mechanism of action and resistance. *Cold Spring Harbor perspectives in medicine* **6**: a026989.

Zhang G & Feng J (2016) The intrinsic resistance of bacteria. *Yi chuan = Hereditas* **38**: 872-880.

Zhang XZ, You C & Zhang YH (2014) Transformation of *Bacillus subtilis*. *Methods in Molecular Biology* **1151**: 95-101.

Zinder ND & Lederberg J (1952) Genetic exchange in Salmonella. *Journal of Bacteriology* **64**: 679-699.

Zink MC, Uhrlaub J, DeWitt J, Voelker T, Bullock B, Mankowski J, Tarwater P, Clements J & Barber S (2005) Neuroprotective and anti-human immunodeficiency virus activity of minocycline. *JAMA* **293**: 2003-2011.

Zuker M (2003) Mfold web server for nucleic acid folding and hybridization prediction. *Nucleic Acids Research* **31**: 3406-3415.

Appendices

Appendix I: Composition of media and solutions

Media/solutions	Composition
SP4X	28.0 g dipotassium phosphate (K_2HPO_4), 12.0 g potassium phosphate (KH_2PO_4), 4.0 g ammonium sulphate (($NH_4)_2SO_4$), 2.0 g trisodium citrate dihydrate ($C_6H_9Na_3O_9$), 0.4 magnesium sulphate heptahydrate ($MgSO_4.7H_2O$), 2.0g casamino acids and 2.0 g yeast extract. Add dH_2O to 500 mL, adjust to pH 7.2 with NaOH
Glucose 20%	4.0 g glucose. Add in 20 mL dH_2O.
Thymine 3.5 mg/mL	0.08 g Thymine. Add in 22.8 mL dH_2O.
SPI	25 mL SP4X, 2.5 mL glucose (20%), 2.86 mL thymine (35 mg/mL) and 5.0 mL amino acids solution [histidine, threonine and methionine (1 mg ml^{-1})]. Add dH_2O to 100 ml.
SPII	22.5 mL SP4X, 2.25 mL glucose (20%) and 2.6 mLthymine (35 mg/mL). Add dH_2O to 90 ml.
Z buffer	8.0 g disodium hydrogen phosphate heptahydrate (Na2HPO4.7H2O), 2.75 g sodium dihydrogen phosphate monohydrate (NaH2PO4.H2O), 0.375 g Potassium chloride (KCl), 0.125 g Magnesium sulphate heptahydrate (MgSO4.7H2O). Add dH_2O to 500 ml, and adjust to pH7. Add 0.14 ml 2-mercaptoethanol in 50 mL Z buffer prior to usage (50 mM)
β-glucuronidase enzyme assay stop solution (1 M Na_2CO_3)	5.3 g Na2CO3. Add dH_2O to 50 ml.

Appendix II: Antibiotics working and stock concentration

Antibiotic	Solvent	Stock concentration	Working concentration
TylAMac™ '469	Sterile water	10 mg/mL	4 - 8 µg/mL
TylAMac™ '4083	Sterile water	10 mg/mL	4 - 8 µg/mL
Tylosin A	Sterile water	10 mg/mL	4 - 8 µg/mL
Doxycycline	Sterile water	10 mg/mL	4 - 8 µg/mL
Tetracycline	70% Ethanol	10 mg/mL	10 µg/mL
Ampicillin	70% Ethanol	100 mg/mL	100 µg/mL
Chloramphenicol	70% Ethanol	25 mg/mL	1 2.5 µg/mL
Erythromycin	Ethanol	10 mg/mL	10 µg/mL
Kanamycin	Sterile water	50 mg/mL	20 µg/mL
Fusidic acid	Sterile water	10 mg/mL	5 µg/mL
Nalidixic acid	Sterile water (pH to 11 with NaOH)	30 mg/mL	10 ug/mL
Rifampicin	DMSO	25 mg/mL	100 ug/mL

Appendix III: Sequence verification of the mutant cassette; pGEM-T/Tn*916*ΔTerm.

The nucleotides in blue; Fragment 1:UPS, nucleotides highlighted in yellow; *cat*P in opposite transcriptional direction of the ORFs in the conjugation module of Tn*916*, nucleotides highlighted in green; Fragment 3:DS1 and nucleotides in turquoise; Fragment 4:DS2. The *XhoI* restriction site is highlighted in red. An * (asterisk) indicates positions which have a single, fully conserved residue.

```
                                              |------------- UPS_F----------|
Tn916ΔT.MC   GCTCCCGGCCGCCATGGCGGCCGCGGGAATTCGATTAGCCAGTAAGGGAACAAAAAATGC
REF          ------------------------------------AGCCAGTAAGGGAACAAAAAATGC
                                                 ***********************

Tn916ΔT.MC   AACGACAAGTCCCCCAAGATTCGCAAATAACATGCGAACAGAAGTTATACTAACAACTTC
REF          AACGACAAGTCCCCCAAGATTCGCAAATAACATGCGAACAGAAGTTATACTAACAACTTC
             ************************************************************

Tn916ΔT.MC   TTGATTATTTCTAGTCATTGCAGAAGTTAACGCGCCATATGGAACGTTTATCGTTGTGTA
REF          TTGATTATTTCTAGTCATTGCAGAAGTTAACGCGCCATATGGAACGTTTATCGTTGTGTA
             ************************************************************

Tn916ΔT.MC   TGTAAGCGACAAGCCAACATAGGTTATATAGGCATATATTAATTTCCCCATATCCGAAAA
REF          TGTAAGCGACAAGCCAACATAGGTTATATAGGCATATATTAATTTCCCCATATCCGAAAA
             ************************************************************

Tn916ΔT.MC   GTCTGGTGTTGTAAAACAGAGTATTGCCAGTATGACAAATGGAAAAGCTCCGAATAAAAG
REF          GTCTGGTGTTGTAAAACAGAGTATTGCCAGTATGACAAATGGAAAAGCTCCGAATAAAAG
             ************************************************************

Tn916ΔT.MC   ATACGGTCTAAAGCGTCCAAATCTGCTGTTCGTTCTGTCAACTATTGTTCCGATAAAAGG
REF          ATACGGTCTAAAGCGTCCAAATCTGCTGTTCGTTCTGTCAACTATTGTTCCGATAAAAGG
             ************************************************************

Tn916ΔT.MC   ATCAGCGAGAGCGTCGATTATTCTAACCACTAAAAACATAGTACCGGCTGCTGCTGCCGA
REF          ATCAGCGAGAGCGTCGATTATTCTAACCACTAAAAACATAGTACCGGCTGCTGCTGCCGA
             ************************************************************

Tn916ΔT.MC   TAAACCAAAAACATCTGTATAGAAGAACAAAAGATACGTAGACACTGTTGCATAAATTAA
REF          TAAACCAAAAACATCTGTATAGAAGAACAAAAGATACGTAGACACTGTTGCATAAATTAA
             ************************************************************

Tn916ΔT.MC   ATTACAAGCAAAATCTCCAGACGCATATCCAACTTTTTCAACCATGCTAATCTTCTTTAC
REF          ATTACAAGCAAAATCTCCAGACGCATATCCAACTTTTTCAACCATGCTAATCTTCTTTAC
             ************************************************************

Tn916ΔT.MC   ATTTTCACTGAGCATGATTTCTCCCCTTTTCGATCCTTTAATATGGAACAGTGTGAAACG
REF          ATTTTCACTGAGCATGATTTCTCCCCTTTTCGATCCTTTAATATGGAACAGTGTGAAACG
             ************************************************************
```

```
Tn916ΔT.MC   GTATAACAATTTCTCGAATTTACGGTCAGTAAACTTTTATAAAAGCGCTTTCAAAACAGT
REF          GTATAACAATTTCTCGAATTTACGGTCAGTAAACTTTTATAAAAGCGCTTTCAAAACAGT
             ************************************************************


Tn916ΔT.MC   GCAAAAAAATATGGTGATCAGCAAGCAAAAAATTCATCTTTTTTCGTCATACCTTATATA
REF          GCAAAAAAATATGGTGATCAGCAAGCAAAAAATTCATCTTTTTTCGTCATACCTTATATA
             ************************************************************


Tn916ΔT.MC   CCTGGAAAAGGATAAAAGCCTCATTCTTTCGTTTCTTATATTGTTTAGCTCCATTAAATC
REF          CCTGGAAAAGGATAAAAGCCTCATTCTTTCGTTTCTTATATTGTTTAGCTCCATTAAATC
             ************************************************************


Tn916ΔT.MC   GACCCTCTTATCCCCCTTCCATCCATATTAGTAAGATCTTATCTTGTGGATAATGAAAGT
REF          GACCCTCTTATCCCCCTTCCATCCATATTAGTAAGATCTTATCTTGTGGATAATGAAAGT
             ************************************************************


Tn916ΔT.MC   TCATTTCATTAGTTTGTTGATCAAACAAACTAACTTGACTTAATATTATAAAATCCTTTC
REF          TCATTTCATTAGTTTGTTGATCAAACAAACTAACTTGACTTAATATTATAAAATCCTTTC
             ************************************************************


Tn916ΔT.MC   GACTTTTTCAACATCTTTTTCAGAAAGTTTAAACAGAATGAATTAAATATTGATTTTTGA
REF          GACTTTTTCAACATCTTTTTCAGAAAGTTTAAACAGAATGAATTAAATATTGATTTTTGA
             ************************************************************


Tn916ΔT.MC   CTGTTTTGTTATATTCAGTGTCTCAGTCCAAAACAGATTCTTGTGTAAACACATGACAAA
REF          CTGTTTTGTTATATTCAGTGTCTCAGTCCAAAACAGATTCTTGTGTAAACACATGACAAA
             ************************************************************


Tn916ΔT.MC   GATGATCTATGCGAAATAAAGAATTGGTTAGTGGTAAGTTGCATTTAAGAGTTTGATAAA
REF          GATGATCTATGCGAAATAAAGAATTGGTTAGTGGTAAGTTGCATTTAAGAGTTTGATAAA
             ************************************************************


Tn916ΔT.MC   GGTGATCACAGGATGTATGGAGACTGTTAGCCTAGATTTATGCTGATGGCAAGCCCGGTC
REF          GGTGATCACAGGATGTATGGAGACTGTTAGCCTAGATTTATGCTGATGGCAAGCCCGGTC
             ************************************************************


Tn916ΔT.MC   ATGAATTGAAAGAACGGAATGGCCAGAATAGTTTATGTTATAAGTCCAACCCTACATACA
REF          ATGAATTGAAAGAACGGAATGGCCAGAATAGTTTATGTTATAAGTCCAACCCTACATACA
             ************************************************************

                                                              |------
Tn916ΔT.MC   TATATCAATACAGGAAAGATAAATAAGAAGCAAAAATAGAGAAGCTTTCAACCGGAGTAG
REF          TATATCAATACAGGAAAGATAAATAAGAAGCAAAAATAGAGAAGCTTTCAACCGGAGTAG
             ************************************************************
             ----------------------- catP_F2 --------------------------|     stop
Tn916ΔT.MC   AAATGGCTATTTGACTTTTTAGTTACAGACAAACCTGAAGTTAACTATTTATCAATTCCT
REF          AAATGGCTATTTGACTttttagttacagacaaacctgaagttaactatttatcaattcct
             ************************************************************


Tn916ΔT.MC   GCAATTCGTTTACAAAACGGCAAATGTGAAATCCGTCACATACTGCGTGATGAACTTGAA
REF          gcaattcgtttacaaaacggcaaatgtgaaatccgtcacatactgcgtgatgaacttgaa
             ************************************************************


Tn916ΔT.MC   TTGCCAAAGGAAGTATAATTTTGTTATCTTCTTTATAATATTTCCCCATAGTAAAAATAG
REF          ttgccaaaggaagtataattttgttatcttcttttataatatttccccatagtaaaaatag
             ************************************************************


Tn916ΔT.MC   GAATCAAATAATCATATCCTTTCTGCAAATTCAGATTAAAGCCATCGAAGGTTGACCACG
REF          gaatcaaataatcatatcctttctgcaaattcagattaaagccatcgaaggttgaccacg
             ************************************************************


Tn916ΔT.MC   GTATCATAGATACATTAAAAATGTTTTCCGGAGCATTTGGCTTTCCTTCCATTCTATGAT
REF          gtatcatagatacattaaaaatgttttccggagcatttggctttccttccattctatgat
```

```
                    ************************************************************

Tn916ΔT.MC         TGTTTCCATACCGTTGCGTATCACTTTCATAATCTGCTAAAAATGATTTAAAGTCAGACT
REF                tgtttccataccgttgcgtatcactttcataatctgctaaaaatgatttaaagtcagact
                    ************************************************************


Tn916ΔT.MC         TACACTCAGTCCAAAGGCTGGAAAATGTTTCAGTATCATTGTGAAATATTGTATAGCTTG
REF                tacactcagtccaaaggctggaaaatgtttcagtatcattgtgaaatattgtatagcttg
                    ************************************************************


Tn916ΔT.MC         GTATCATCTCATCATATATCCCCAATTCACCATCTTGATTGATTGCCGTCCTAAACTCTG
REF                gtatcatctcatcatatatccccaattcaccatcttgattgattgccgtcctaaactctg
                    ************************************************************


Tn916ΔT.MC         AATGGCGGTTTACAATCATTGCAATATAATAAAGCATTGCAGGATATAGTTTCATTCCCT
REF                aatggcggtttacaatcattgcaatataataaagcattgcaggatatagtttcattccct
                    ************************************************************


Tn916ΔT.MC         TTTCCTTTATTTGTGTGATATCCACTTTAACGGTCATGCTGTAGGTACAAGGTACACTTG
REF                tttcctttatttgtgtgatatccactttaacggtcatgctgtaggtacaaggtacacttg
                    ************************************************************


Tn916ΔT.MC         CAAAGTAGTGGTCAAAATACTCTTTTCTGTTCCAACTATTTTTATCAATTTTTTCAAATA
REF                caaagtagtggtcaaaatactcttttctgttccaactattttttatcaatttttttcaaata
                    ************************************************************
                    start                                             - 10
Tn916ΔT.MC         CCATCTAAGTTCCCTCTCAAATTCAAGTTTATCGCTCTAATGAACAAAGATATTATACCA
REF                ccatctaagttccctctcaaattcaagtttatcgctctaatgaacaaagatattatacca
                    ************************************************************
                         - 35
Tn916ΔT.MC         CATTTTTGTGAATTTTTCAACTTGCCCACTTCGACTGCACTCCCGACTTAATAACTTCTT
REF                catttttgtgaattttttcaacttgcccacttcgactgcactcccgacttaataacttctt
                    ************************************************************
                                                     |----------------BSA_F_xhol--------------|
Tn916ΔT.MC         GAACACTTGCCGAAAAAGAAAAACTGCCGGGTCTCGAGGTCTGAGGATTAATGGCTGTGT
REF                gaacacttgccgaaaaagaaaaactgccgggtCTCGAGGTCTGAGGATTAATGGCTGTGT
                    ************************************************************


Tn916ΔT.MC         TAAACACTATGATTTTTCCTTCAAACTTATTTTCTAAGAAAAATAGCATAAAAATCTAGT
REF                TAAACACTATGATTTTTCCTTCAAACTTATTTTCTAAGAAAAATAGCATAAAAATCTAGT
                    ************************************************************


Tn916ΔT.MC         TATCCGCATAAAAACTGGACTTATCACACTTTATCAAGGTCAAAACCACTCAATTTACTA
REF                TATCCGCATAAAAACTGGACTTATCACACTTTATCAAGGTCAAAACCACTCAATTTACTA
                    ************************************************************


Tn916ΔT.MC         CTAATTTACTACTTATGAATGAGCTTTGATACGACGATTTATCCTTGAAAAGTGAAGATA
REF                CTAATTTACTACTTATGAATGAGCTTTGATACGACGATTTATCCTTGAAAAGTGAAGATA
                    ************************************************************
                                                 |------------------------------DS_BF--------------------------------|
Tn916ΔT.MC         TAAAGATACTTCCAATAAAATTTGAATATTTAATAGGTAACCCGATTTTGAAAGGAAGTG
REF                TAAAGATACTTCCAATAAAATTTGAATATTTAATAGGTAACCCGATTTTGAAAGGAAGTG
                    ************************************************************


Tn916ΔT.MC         AACTTATGAAAACAAAAAAATCAAGAATCAAAAGGTCGTTCCCCACTCTTTAAGACCATCA
REF                AACTTATGAAAACAAAAAAATCAAGAATCAAAAGGTCGTTCCCCACTCTTTAAGACCATCA
                    ************************************************************


Tn916ΔT.MC         AACATTCATTCAGCCAATAAAAAAAGAAAGGATAGGTAAAAATATGGAACTTAAATTTGTG
REF                AACATTCATTCAGCCAATAAAAAAAGAAAGGATAGGTAAAAATATGGAACTTAAATTTGTG
                    ************************************************************
```

274

```
Tn916ΔT.MC    ATTCCCAACATGGAAAAAACATTCGGCAATTTAGAATTTGCTGGCGAGGATAAAGTCGTT
REF           ATTCCCAACATGGAAAAAACATTCGGCAATTTAGAATTTGCTGGCGAGGATAAAGTCGTT
              ************************************************************

Tn916ΔT.MC    CAGCGAAGAATCAACGGACGGCTAACTGTCTTATCAAGAAGCTATAATCTCTATTCTGAT
REF           CAGCGAAGAATCAACGGACGGCTAACTGTCTTATCAAGAAGCTATAATCTCTATTCTGAT
              ************************************************************

Tn916ΔT.MC    GTTCAAAGAGCAGATGATATTGTGGTGGTGCTTCCTGCTGAAGCTGGCGAAAAACATTTC
REF           GTTCAAAGAGCAGATGATATTGTGGTGGTGCTTCCTGCTGAAGCTGGCGAAAAACATTTC
              ************************************************************

Tn916ΔT.MC    GGCTTTGAGGAACGTGTGAAGTTAGTCAATCCACGTATTACCGCAGAGGGCTACAAAATC
REF           GGCTTTGAGGAACGTGTGAAGTTAGTCAATCCACGTATTACCGCAGAGGGCTACAAAATC
              ************************************************************

Tn916ΔT.MC    GGCACTCGTGGTTTTACAAATTACCTTTTACATGCTGACGACATGATAAAAGAATAAAGA
REF           GGCACTCGTGGTTTTACAAATTACCTTTTACATGCTGACGACATGATAAAAGAATAAAGA
              ************************************************************

Tn916ΔT.MC    AAGAGAGGAAAAATGATGAGATTAGCAAATGGCATTGTATTAGATAAAGACACGACTTTT
REF           AAGAGAGGAAAAATGATGAGATTAGCAAATGGCATTGTATTAGATAAAGACACGACTTTT
              ************************************************************

Tn916ΔT.MC    GGAGAATTGAAATTCTCTGCTCTACGTCGTGAAGTGAGAATCCAAAATGAAGACGGGTCG
REF           GGAGAATTGAAATTCTCTGCTCTACGTCGTGAAGTGAGAATCCAAAATGAAGACGGGTCG
              ************************************************************

Tn916ΔT.MC    GTTTCAGATGAAATCAAGGAACGTACCTATGACTTAAAATCCAAAGGACAAGGACGCATG
REF           GTTTCAGATGAAATCAAGGAACGTACCTATGACTTAAAATCCAAAGGACAAGGACGCATG
              ************************************************************

Tn916ΔT.MC    ATTCAAGTAAGTATTCCTGCCAGCGTGCCTTTGAAAGAGTTTGATTATAACGCACGGGTG
REF           ATTCAAGTAAGTATTCCTGCCAGCGTGCCTTTGAAAGAGTTTGATTATAACGCACGGGTG
              ************************************************************

Tn916ΔT.MC    GAACTTATCAATCCCATTGCGGACACCGTTGCTACTGCCACCTATCAAGGAGCAGATGTT
REF           GAACTTATCAATCCCATTGCGGACACCGTTGCTACTGCCACCTATCAAGGAGCAGATGTT
              ************************************************************

Tn916ΔT.MC    GACTGGTATATCAAGGCAGACGATATTGTGCTGACAAAGGATTCTAGTTCATTCAAAGCT
REF           GACTGGTATATCAAGGCAGACGATATTGTGCTGACAAAGGATTCTAGTTCATTCAAAGCT
              ************************************************************

Tn916ΔT.MC    CAACCACAAGCAAAGAAAGAACCGACACAAGACAAATAGTCGCTAGGTAGAAAGGAGACT
REF           CAACCACAAGCAAAGAAAGAACCGACACAAGACAAATAGTCGCTAGGTAGAAAGGAGACT
              ************************************************************

Tn916ΔT.MC    TTTTCGCATGAAACAGCGTGGTAAAAGGATTCGCCCATCTGGTAAAGATTTAGTCTTTCA
REF           TTTTCGCATGAAACAGCGTGGTAAAAGGATTCGCCCATCTGGTAAAGATTTAGTCTTTCA
              ************************************************************

                                                                        |-
Tn916ΔT.MC    TTTTACGATAGCGTCACTCCTGCCTGTTTTCCTGCTGGTTGTCGGACTGTTTCATGTGAA
REF           TTTTACGATAGCGTCACTCCTGCCTGTTTTCCTGCTGGTTGTCGGACTGTTTCATGTGAA
              ************************************************************

              ------ DS_BR---------------|
Tn916ΔT.MC    GACAATCCAGCAGATCAACAAATCACTAGTGAATTCGCGGCCGCCTGCAGGTCGACCATA
REF           GACAATCCAGCAGATCAAC-----------------------------------------
              *******************
```

275