Kingston University London

The Antibiotic Resistance Growth Plate (ARGP) as an experimental evolution tool to explore the phenotypic and genotypic mutational pathways underlying the emergence of antimicrobial resistance in *Escherichia coli*.

A thesis submitted in partial fulfilment for the degree of

Doctor of Philosophy

By

Lucky Bonnie Lucia CULLEN

Faculty of Science, Engineering and Computing

February 2019

Declaration

This thesis entitled ‘The Antibiotic Resistance Growth Plate (ARGP)’ as an experimental evolution tool to explore the phenotypic and genotypic mutational pathways underlying the emergence of antimicrobial resistance in *Escherichia coli*’ is based upon the work conducted in the Faculty of Science, Engineering and Computing at Kingston University London and in collaboration with Dr Philip Aldridge at Newcastle University and Katie Hopkins and Neil Woodford from Public Health England. All the work described is the candidate’s own original work unless otherwise stated. None of the work presented has been submitted for another degree internally or externally.

Acknowledgements

At the start of this PhD I had a dream, a dream which could have never been accomplished without the support of many people. Firstly, I would like to express my sincere appreciation to my director of studies Professor Mark Fielder, your provision, guidance and unconditional belief in my potential has contributed significantly to the completion of this thesis. I would also like to acknowledge the overwhelming support received from, Dr Scott Lawton, Camilla Eldridge and Ben Jones, your wealth of knowledge and expertise in the field of evolutionary biology and bioinformatics has assisted greatly in concluding the work within this PhD.

I would also like to praise the incredible multidisciplinary support received at Kingston University, specifically that from Dr Gary-Forster Wilkins, Dr Adam Le Gresley, Dr Brian Rooney, Dr James Denholm-Price and Richard Giddens. I would also like to extend my gratitude to all those in the Microbiology laboratory including my lovely friends and technicians, who have made this experience enjoyable. A special thank you goes to Coral Brazier, from the Kingston University wellbeing team.

Finally, I would like to thank my wonderful family and friends who have believed in me endlessly, thank you all for being a part of my journey. To my sister Kaycee thank you for being my glimmer of sunshine during the darkest times. The greatest thank you goes to my mum, nobody knows how challenging this journey has been like you do and neither would this have been possible without you. Everything, I am and everything I aspire to be, I owe to you. Nanny Kathy although you will not remember, I dedicate this to you.

Publications

Antibiotic Resistance Growth Plate (ARGP): A simple tool for exploring the evolution of antimicrobial resistance.

Cullen LBL, Eldridge C, Forster-Wilkins G and Fielder MD (Article in preparation).

Presentations

Developing an experimental tool to explore the evolution of antimicrobial resistance? Oral presentation. Society for Applied Microbiology ECS Research Conference. The Royal Society of Medicine. October 2016.

Exploring the evolutionary pathways to antimicrobial resistance using experimental evolution. Oral presentation. Society for Applied Microbiology Antimicrobial Resistance Conference. One Great George Street. November 2016.

Characterising the emergence of antibiotic resistance through mutations using experimental evolution. Poster presentation. Science Engineering and Computing Conference. Kingston University London. April 2017.

Characterising the emergence of antibiotic resistance: mutation mapping of the resistome. Oral presentation. American Society for Microbiology (ASM) Microbe Conference. New Orleans. June 2017.

An alternative approach to tackling the antibiotic resistance crisis. Oral presentation. Three Minute Thesis winner and national semi-finalist. Kingston University London. 2017.

List of Acronyms

°	Degree
°C	Degrees Centigrade
µg	Micrograms
µl	Microliter
µM	Micromolar
π	Nucleotide Diversity
3’	3 prime end
5’	5 prime end
ABU	Asymptomatic Bacteriuria
AIC	Akaike Information criterion
ALE	Adaptive Laboratory Evolution
AACs	Aminoglycoside Acetyltransferases
AMEs	Aminoglycoside Modifying Enzymes
AMR	Antimicrobial Resistance
APHA	Animal and Plant Health Agency
ARGP	Antibiotic Resistance Growth Plate
ASM	American Society for Microbiology
BIC	Bayesian Information Criterion
BSAC	British Society for Antimicrobial Chemotherapy
CDC	Centre for Disease Control and Prevention
CDS	Coding Sequence
CFU/mL	Colony Forming Units Per mL

CLED	Cysteine-Lactose-Electrolyte Deficient
CPE	Carbapenemase Producing Enterobacteria
CRE	Carbapenem-Resistant *Enterobacteriaceae*
Cryo-EM	Cryo-Electron Microscopy
D_2O	Deuterium Oxide
dH_2O	Distilled Water
DMAB	Dimethylaminobenzaldehyde
dN/dS	Non-synonymous/Synonymous
dNTPs	Fluorescently Tagged Nucleotides
DT	Decision Theory
EARS-Net	European Antimicrobial Resistance Surveillance Network
E. coli	*Escherichia coli*
EDTA	Ethylenediaminetetraacetic Acid
EF-G	Elongation Factor-G
EF-Tu	Elongation Factor-Tu
EMBL-EBI	European Bioinformatics Institute
EPEC	Enteropathogenic *E. coli*
EUCAST	European Committee on Antimicrobial Susceptibility Testing
ExPEC	Extraintestinal Pathogenic *E. coli*
G. mellonella	*Galleria mellonella*
GF-Score	Goodness-of-Fit Score
GMQE	Global Model Quality Estimate
GTP	Guanosine Triphosphate
HGT	Horizontal Gene Transfer

-InL	Likelihood
ICL	Imperial College London
iNTPs	Initiating Nucleoside Triphosphate
IS	Insertion Elements
iTOL	Interactive Tree of Life
KU	Kingston University
LCBs	Locally Collinear Blocks
LC-MS	Liquid Chromatography- Mass Spectroscopy
MEA	Maximum Expected Accuracy
MEGA	Molecular Evolutionary Genetics Analysis
MFE	Minimum Free Energy
MH	Mueller Hinton
MIC	Minimum Inhibitory Concentration
ML	Maximum Likelihood
MLST	Multilocus Sequence Typing
mRNA	Messenger Ribonucleic Acid
MSW	Mutant Selection Window
MUSCLE	Multiple Sequence Comparison Log-Expectation
MVA	Multivariate Analyses
NCBI	National Centre for Biotechnology
NCTC	National Collection of Type Cultures
NJ	Neighbour Joining
NMR	Nuclear Magnetic Resonance
NNI	Nearest-neighbour Interchange

nt	Nucleotide
NU	Newcastle University
OD	Optical Density
O/N	Overnight
OPG	Osmoregulated Periplasmic Glucan
OPLS-SA	Orthogonal Partial Least Squares- Discriminant Analysis
P. mirabilis	*Proteus mirabilis*
PBS	Phosphate Buffered Saline
PCA	Principal Component Analysis
PCR	Polymerase Chain Reaction
PDB	Protein Data Bank
PHE	Public Health England
ppGpp	Guanosine5’-diphosphate 3’-diphosphate
QMEAN	Quality Model Energy
QSQE	Quaternary Structure Quality Estimate
RND	Resistance-Nodulation-Division
rrn	Ribosomal RNA
rRNA	Ribosomal Ribonucleic Acid
SAPS	Statistical Analysis of Protein Sequences
SEM	Scanning Electron Microscopy
SfAM	Society for Applied Microbiology
SMART	Simple Modular Architecture Research Tool
smFRET	Single-Molecule Fluorescence Resonance Energy Transfer
SNP	Single Nucleotide Polymorphism

Spp.	Species
TBE	Tris-Borate-EDTA
TDR	Totally Drug-Resistant
TMS	Tetramethylsilane
trGTPases	Translational GTPases
TTC	Triphenyltetrazolium Chloride
VSEPR	Valence Shell Electron Pair Repulsion
WGS	Whole Genome Sequencing
WT	Wild-Type

Table of Contents

Table of Figures

Table of Tables

Abstract

The increasing threat of an antimicrobial resistance crisis is a significant global concern. Antimicrobial treatment failures are worsened by the rapid evolution of resistance amongst bacterial pathogens. Therefore, in addition to developing novel antimicrobial agents, there is growing interest in exploring the underlying genotypic factors of resistance evolution. Traditionally, such studies have focused heavily on well-established mechanisms of acquired resistance involving horizontal gene transfer, yet the evolution of resistance through the acquisitions of mutations is yet to be fully elucidated.

Adaptive laboratory evolution studies have provided insights into the genetic basis of adaptation through the direct observation of the evolutionary process. Experimental evolution has advanced from serial passage in well mixed systems to the incorporation of spatiotemporal antibiotic concentration gradients. During this study the Antibiotic Resistance Growth Plate (ARGP) was developed as a simple experimental tool to explore the ability of bacteria to evolve resistance across an antibiotic landscape. The device (90mm × 15mm) facilitates the direct observation of the evolutionary trajectories of mutational lineages within a circular format supporting the radial growth and the exploration of phenotypic space within bacterial populations. Whole genome sequencing of the evolved resistant strains, identified key mutations in the 16s rRNA genes and the *fusA* gene encoding elongation factor-G, specific to antimicrobial agent gentamicin. Additional gene sequencing revealed parallel gentamicin resistant bacterial populations, evolved identical mutations within the *fusA* gene. Combined bioinformatic, phylogenetic and molecular docking analysis

uncovered the biological significance of the *fusA* gene in the mutational pathways of gentamicin resistance in *E. coli* MG1655 *in vitro*. In contrast, the observed biological fitness costs associated with the acquisition of resistance conferring mutations, emphasised why such costly resistant genotypes were unidentified in natural and clinical settings.

This study has established an experimental evolution model to explore the mutational pathways underlying antimicrobial resistance development *in vitro.* The ARGP offers a platform for the continued research into the acquisition of antimicrobial resistance through mutations, for more complex bacterial pathogens selected against a range of antimicrobial agents. As a tool, the ARGP can be utilised to inform therapeutic decisions based on the evolutionary risk management, provide new opportunities within the field of drug development and holds scope for its application within an educational setting.

Chapter 1: General Introduction

1.1 Antimicrobial resistance

The development of antimicrobial agents was an eminent discovery of the last century; as their introduction saw a drastic reduction in morbidity and mortality (Davies, 2006). However, the rapid emergence of single and multi-drug-resistant bacterial strains has diminished our repertoire of effective antimicrobial agents available (Davies and Davies, 2010). Antimicrobial chemotherapy acts as a selective pressure driving the development of resistance, however preceding the therapeutic use of Penicillin, an enzyme capable of destroying it was identified (Abraham and Chain, 1940; Aminov, 2010). Following the widespread therapeutic application of penicillin, resistant strains of *Staphylococcus aureus* began to disseminate, not dissimilar to the heightened situation seen today (Lowy, 2003; Foster, 2017). The inevitable trend of antimicrobial discovery, and concomitant resistance development has caused an antimicrobial discovery void, where despite the alarming rates of resistance development, no novel classes of antimicrobial agents have been discovered (Payne *et al.,* 2007; Silver, 2011; Brown, 2015).

1.1.1 The antimicrobial resistance threat

To appreciate the magnitude of the problem in regards to human health, in the year 2015, the European Antimicrobial Resistance Surveillance Network (EARS-Net) estimated that, 33,110 deaths were attributable to antibiotic resistant infections (Cassini *et al.,* 2019). Two years previous, the United States Centers for Disease Control and Prevention (CDC) projected that 2 million people developed drug resistant infections, which were expected to account for a minimum of 23,000 deaths

per year (CDC, 2013). Despite this, there are major inconsistencies in the global surveillance and reporting of antimicrobial resistant infections (Felmingham, 2002). Therefore, the extent of this global issue is likely to be more serious than the available data suggests (World Health Organisation, 2014). In fact, a review commissioned by the UK government on antimicrobial resistance, revealed that based on current trends the annual deaths caused by antimicrobial resistant infections could increase to 10 million per year by 2050 (O'Neil, 2016).

1.1.2 Antimicrobial resistance: a one health perspective

The extensive misuse of antimicrobial agents in the environment, human medicine and animal husbandry is a predominant cause of the global rise of antimicrobial resistance (AMR) (Hollis and Ahmed, 2013). The concept of 'One Health' accurately portrays the issue of AMR, as it refers to the ecological relationships between human, animal and environmental health (American Veterinary Medical Association, 2008; Robinson *et al.,* 2016). Therefore, tackling the multisectoral issue of AMR, requires interdisciplinary collaboration across a range of sectors including; human, animal, food, agriculture and pharmaceutics (Singh, 2017; WHO, 2018). This was acknowledged by the G20 health ministers in Berlin, who declared a commitment to develop future AMR actions plans, based on a one health approach (G20 Health Ministers Declaration, 2017).

1.1.3 Mechanisms of antimicrobial resistance

Antimicrobial agents are characterised based on their mechanisms of action and exert their activity through the disruption of cellular processes or functions essential to bacterial cell growth (bacteriostatic antibiotics) or viability (bactericidal antibiotics) (Walsh, 2003). The principal mechanisms of action include: inhibition of cell wall synthesis by β-lactam antibiotics (Tomasz, 1979), interference of deoxyribonucleic acid (DNA) synthesis through double stranded DNA breaks by quinolones (Drlica *et al.,* 2008), inhibition of ribosomal protein synthesis by a variety of antimicrobial agents including the aminoglycosides (Vakulenko and Mobashery*,* 2003) and the inhibition of metabolic pathways including folate synthesis by sulfonamides (Miller, Doukas and Seydel, 1972; Tenover, 2006). The antimicrobial mechanisms and drug-target interactions have been extensively reviewed by Kohanski, Dyer and Collins (2010). Microbial resistance has emerged through an abundance of mechanisms to overcome the antimicrobial actions of the various classes of antibiotics including; drug target modification, enzymatic modification, enzymatic degradation and active drug efflux (Wright, 2011), as illustrated in Figure (1.1).

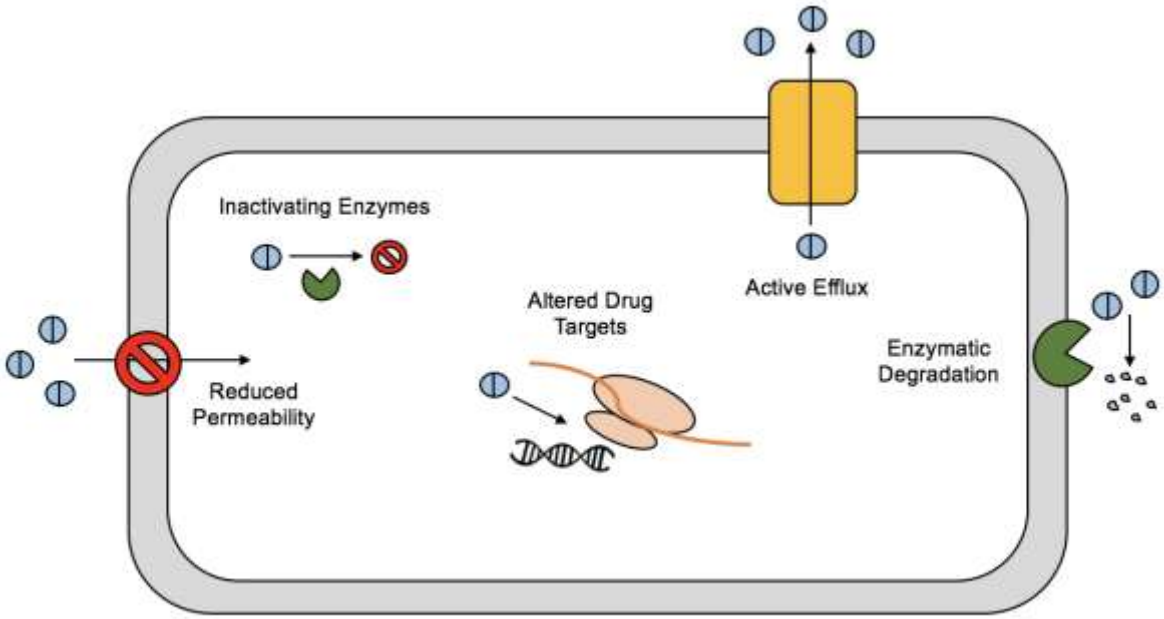


Figure 1.1 Bacterial antimicrobial resistance strategies. The image displays some of the prevalent mechanisms utilised by bacteria to overcome antimicrobial selective pressures. Examples include: enzymatic inactivation, altered drug uptake, target modification and active drug efflux. Other mechanisms including target overproduction, target mimicry and enzymatic bypass of drug inhibition pathways have not been included in this image.

1.1.3.1 Intrinsic mechanisms of antimicrobial resistance

Intrinsic resistance indicates bacterial species which are inherently resistant to a variety of classes of antibiotics, independent of the antimicrobial selective pressure, this mechanism of resistance has been reviewed extensively by Cox and Wright (2013). Intrinsic resistance is a growing concern in both clinical and environmental environments, owed to the prevalence of intrinsic resistance within multi-drug resistant (MDR) Gram-negative pathogens, and the abundance of intrinsically resistant organisms being isolated from soil samples (Nikaido, 1994; D'Costa *et al.,* 2006). This has motivated efforts to inhibit the key targets of the intrinsic resistome, to impart antimicrobial sensitivity on inherently resistant microorganisms (Liu *et al.*, 2010).

1.1.3.2 Acquired mechanisms of antimicrobial resistance

Antimicrobial resistance can be acquired by previously susceptible bacteria, through the horizontal acquisition of genetic material, referred to as horizontal gene transfer (HGT) (Davies, 1997). There are three mechanisms which facilitate HGT, which include: transduction, natural transformation and conjugation (Blakely, 2015; Sun, 2018). Extracellular mobile genetic elements (MGEs) including integrons and transposons promote HGT by natural transformation (Domingues, Nielsen and da Silva, 2012). Whereas, plasmids and bacteriophages which are widespread amongst bacterial populations, assist the process of HGT by conjugation and transduction respectively (Harrison *et al.,* 2015; Lerminiaux and Cameron, 2018). As HGT can occur between closely related and phenotypically diverse organisms, HGT is widely considered as the mechanism accountable for the global dissemination of resistance genes (Wiedenbeck and Cohan, 2011; Johnston *et al.,* 2014; von Wintersdorff *et al.,* 2016; Lerminiaux and Cameron, 2018).

1.1.3.3 Acquired antimicrobial resistance through mutations

In addition to the horizontal acquisition of resistance, bacteria can acquire resistance through spontaneous chromosomal mutations referred to as vertical evolution (Martinez and Baquero, 2000). The chromosomal resistance conferring mutations typically emerge within antimicrobial targets (Spratt, 1994). Although uncommon, a single mutation in a bacterial genome can significantly alter the susceptibility of an organism to antimicrobial agents, as demonstrated by the spontaneous mutations conferring clarithromycin resistance in *Helicobacter pylori*, rifampin resistance in *Staphylococcus aureus* and fluoroquinolone resistance in *Campylobacter jejuni*

(Aubry-Damon, Soussy and Courvalin, 1998; Engberg *et al.,* 2001; Wang *et al.,* 2001). Nevertheless, the development of high levels of antimicrobial resistance, in the absence of plasmids, has been directly linked to a stepwise accumulation of multiple resistance conferring mutations (Levy and Marshall, 2004; Lozovsky *et al.,* 2009). While advancements in laboratory evolution experiments, have suggested that the distinctive stepwise trend in mutational trajectories is highly dependent on the antimicrobial class (Toprak *et al.,* 2011).

To study the evolution of bacterial genomes through mutations, it is necessary to understand the determinants of spontaneous mutations. The greatest source of genetic variation within prokaryotic genomes occurs through DNA polymerase errors and through endogenous and exogenous DNA damage (Foster *et al.,* 2015). Spontaneous mutation rates in the absence of exogenous DNA damaging agents such as antibiotics, can vary significantly between bacterial genomes (Drake *et al.,* 1998; Rosche and Foster, 2000). The spontaneous mutation rates in the bacterium *E. coli* is estimated to be 1.0×10^{-3} per genome per generation; with a generation time of twenty minutes (Lee *et al.,* 2012). Antimicrobial agents are known to accelerate mutation rates in bacterial populations, however evolved mutations can have varying impacts depending on whether they are deleterious, neutral or beneficial to the organism (Denamur and Matic, 2006; Long *et al.,* 2016). Therefore, it is unsurprising that bacterial hypermutators, which display elevated mutation rates have an increased frequency of resistance conferring mutations (Giraud *et al.,* 2001). Typically, adaptive evolution is determined using likelihood estimates of non-synonymous and synonymous substitution rates (dN/dS), however when very few

mutations are present in a small number of genes, the likelihood approach lacks statistical power (Yang and Nielsen, 1998). Thus, in such cases, adaptive evolution is confirmed through sequencing parallel evolved bacterial strains for the existence of adaptive mutations (Palmer and Kishony, 2013). Sequentially, the functional impacts of the identified resistance conferring mutations are assessed based on positional conservation within the mutated gene, and any significant alterations in protein properties, structure and function (Toprak *et al.,* 2011; Stucki and Gagneux, 2013; Baym *et al.,* 2016b).

1.1.4 The significance of Gram-negative bacteria

The bacterial cell envelope differentiates Gram-positive and Gram-negative bacterial species (Gram, 1884). A unique feature of Gram-negative bacteria is the outer membrane, which decreases the permeability of the bacterial cell and is considered a causative factor of intrinsic bacterial resistance (Silhavy, Kahne and Walker, 2010). The outer membrane barrier is amongst several factors that has contributed to the lack of progress in the development of new antimicrobial agents effective against Gram-negative pathogens (Vaara, 1992; Livermore, 2011). This is of great concern, owing to the drastic increase in multidrug-resistant Gram-negative bacterial infections (Worthington and Melander, 2013; Exner *et al.,* 2017). As a result, previously withdrawn antimicrobial agents including the polymyxin colistin have re-emerged as therapeutic options (Stein and Raoult, 2002; Ouderkirk *et al.*, 2003; Falagas and Kasiakou, 2005).

The Gram-negative, *Enterobacteriaceae* family, pose a significant threat as in addition, to the multi-drug resistance described above, enteric bacteria are emerging with extremely drug-resistant (XDR) and totally drug-resistant phenotypes (TDR) (Rossolini *et al.,* 2014). The prevalence of highly resistant phenotypes amongst Gram-negative bacteria including, *Klebsiella pneumoniae*, *Pseudomonas aeruginosa* and *Enterobacter* spp. is extremely alarming as these bacteria form part of the ESKAPE pathogens which are the leading cause of nosocomial infections worldwide (Pendleton, Gorman and Gilmore, 2013; Santajit and Indrawattana, 2016). The emergence of Carbapenem-resistant *Enterobacteriaceae* (CRE) presents a major public health crisis, as Carbapenemase producing *Enterobacteria* (CPE) infections are associated with mortality rates as high as 70% and almost all CPE infections are of the XDR phenotype (Tzouvelekis *et al.,* 2012). The rapid evolution and dissemination of CPE, caused an international epidemic of colistin resistant *Klebsiella pneumoniae* carbapenemase (KPC) producing *K. pneumoniae* of the ST258 clonal lineage (Kitchel *et al.,* 2009; Samuelsen *et al.,* 2009; Bogdanovich *et al.*, 2011).

1.1.4.1 *Escherichia coli* and its application as a model organism

Escherichia coli (*E. coli*) is a Gram-negative, rod-shaped bacterium, of the *Enterobacteriaceae* family (Li *et al.*, 2013; Croxen *et al.,* 2013). The facultative anaerobe, is found commonly within the mammalian gastrointestinal tract, where it typically constitutes between 0.1%-5% of the gut microflora (Conway, Krogfelt and Cohen, 2004; Tenaillon *et al.,* 2010; Conway and Cohen, 2015). While commensal strains of *E. coli* are considered beneficial and generally harmless, pathogenic strains of *E. coli* contribute significantly to global morbidity and mortality (Moal Lievin-Le and Servin, 2006; Khalil *et al.,* 2018). The pathogenic strains of *E. coli* are

phylogenetically diverse, and can cause a variety of enteric and extraintestinal infections in humans, which vary according to their pathogenicity profile or pathotype (Duriez *et al.*, 2001; Kaper, Nataro and Mobley, 2004; Croxen and Finlay, 2009; Clements *et al.,* 2012). The extraintestinal pathogenic *E. coli* (ExPEC), are the most frequently detected Gram-negative pathogens, and are associated with uropathogenic infections, systemic infections including septicaemia and neonatal meningitis (Johnson and Russo, 2002; Russo, and Johnson, 2003; Johnson *et al.,* 2003).

Antimicrobial resistance in *E. coli* is a growing public health concern (Schwaber *et al.,* 2006). The management of hospital-acquired and community-onset ExPEC infections by β-lactam antibiotics, has been compromised by the production of extended-spectrum beta-lactamases (ESBLs) within the *Enterobacteriaceae* family (Livermore and Woodford, 2006; Pitout and Laupland, 2008). The widespread increase in CTX-M ESBL producing isolates of *E. coli*, is of major global concern seeing the CTX-M pandemic and the association of such strains with increased resistance to multiple classes of antibiotics (Canton and Coque, 2006; Pitout 2012; Woerther *et al.,* 2013). The global dissemination of the CTX-M genotypes is displayed in figure (1.2) (Hawkey and Jones, 2009).

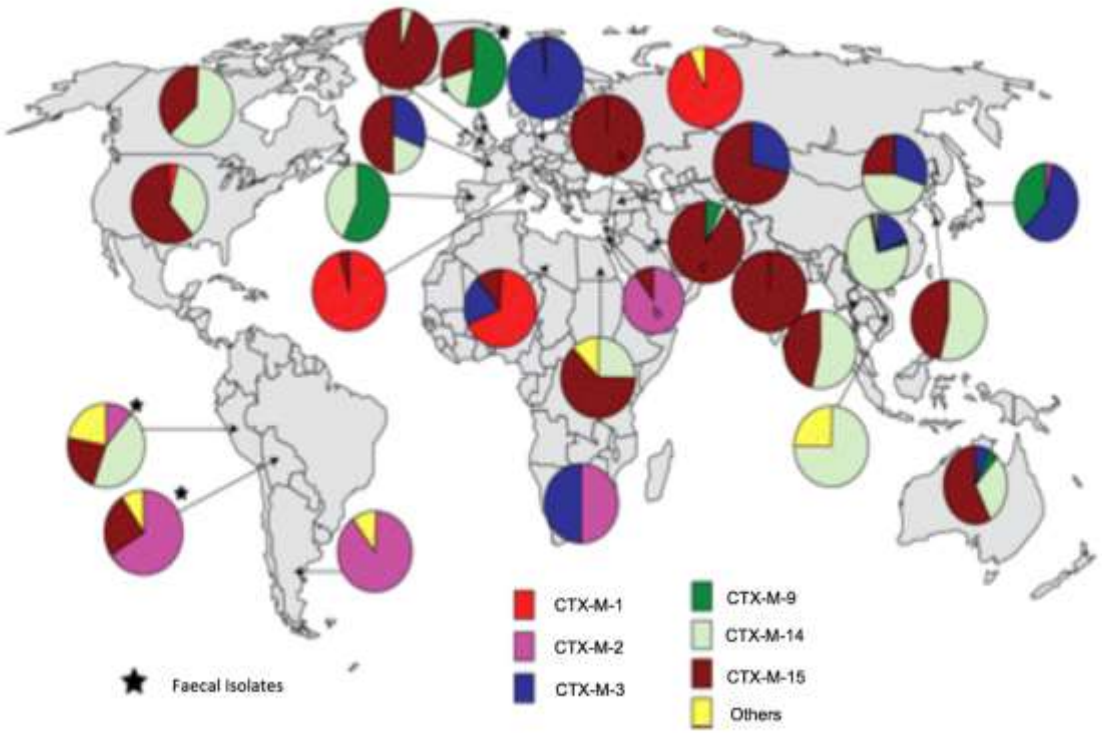


Figure 1.2 Global epidemiology of the CTX-M genotypes. The image displays the geographical locations of the most frequently isolated CTX-M genotypes worldwide. The CTX-M enzymes isolated from faecal samples are indicated using a star. Adapted from: Hawkey and Jones, 2009.

Since the discovery of the bacterium in 1885, *E. coli* has become one of the most comprehensively studied bacterial species (Escherich, 1988). This was enhanced, through the completion of the K-12 MG1655 strain whole genome sequence, which led to prominent discoveries in genetic recombination (Lederberg and Tatum, 1946; Blattner *et al.,* 1997; Riley *et al.,* 2006). Subsequently, the *E. coli* K-12 MG1655 strain was adopted widely as the wild-type (WT) in laboratory experiments, and became the favoured bacterial model organism in experimental evolution studies (Kawecki *et al.,* 2012; Blount, 2015). Consequently, the *E. coli* strain K-12 MG1655 was selected as the principle organism within this thesis.

1.1.4.2 The importance of aminoglycoside antibiotics

The first aminoglycoside antibiotic discovered, was streptomycin, from *Streptomyces griseus*, in 1944 (Schatz, Bugle and Waksman, 1944). The aminoglycoside antibiotics were originally referred to as aminoglycosidic aminocyclitols, as they are characterised by amino sugars attached to an aminocyclitol by glycosidic linkages (Begg and Barclay, 1995). Aminoglycosides exert their antimicrobial activity through binding to the bacterial ribosome and inhibiting ribosomal protein synthesis (Vakulenko and Mobashery, 2003). Additionally, they express synergistic bactericidal activity when in combination with antibiotics which inhibit cell wall synthesis (Iannini, Ehret and Eickhoff, 1976; Calderwood *et al.,* 1977). Aminoglycosides have broad-spectrum activity against a range of clinically significant bacterial pathogens including members of the *Enterobacteriaceae* family, *P. aeruginosa* and *Mycobacterium* spp. (Swenson *et al.,* 1985; Aggen *et al.,* 2010; Krause *et al.,* 2016).

The most recurrently administered aminoglycoside is gentamicin, which was derived from *Micromonospora purpurea*, in 1963 (Weinstein *et al.,* 1963) and is used widely in the empirical treatment of severe Gram-negative pathogens within neonatal departments (Kumana and Yuen, 1994; Germovsek, Barker and Sharland, 2016). The therapeutic application of aminoglycosides within clinical practice, is heavily compromised by the development of aminoglycoside resistance (Garneau-Tsodikova and Labby, 2016; Doi, Wachino and Arakawa, 2016). This is worsened by the diverse mechanisms of resistance to aminoglycosides, which comprise: enzymatic drug modification through aminoglycoside-modifying enzymes (AMEs) including aminoglycoside N-acetyltransferases (AACs) (Ramirez and Tolmasky, 2010), active

efflux primarily through the multidrug transporter AcrAD, of the resistance-nodulation-division (RND) family of efflux pumps (Nikaido,1996), decreased permeability through modifications in the barrier properties of the bacterial outer membrane (Hancock, 1981) and ribosomal/target site modification which could be enzymatic or mutational (Mingeot-Leclercq, Glupczynski and Tulkens 1999; Vakulenko and Mobashery, 2003; Krause *et al.,* 2016; Doi, Wachino and Arakawa, 2017). Consequently, the aminoglycoside gentamicin was the most extensively utilised antimicrobial agent within this thesis.

1.2 The history of microbial culture methods

Traditional bacterial culture methods were based on elective enrichment whereby favourable conditions were established to encourage the identification of naturally occurring organisms (van Niel, 1944). Subsequent developments in bacterial cultivation methods, saw the introduction of selective enrichment techniques, including replica plating, where precise organisms were selected based on adaptive changes under specified conditions (Lederberg and Lederberg, 1952; Lagier *et al.,* 2015). This direct method of selection, was used to demonstrate the development of resistance mutants to increased concentrations of penicillin (Eagle, Fleischman and Levy, 1951).

1.2.1 The application of adaptive laboratory evolution (ALE) in microbiology

Bacterial selection experiments, prompted the use of microorganisms in ecological studies of the evolutionary process of natural selection (Dykhuizen, 1990). Historically, this was inferred through the analysis of evolutionary outcomes, in

multicellular eukaryotic organisms (Darwin, 1859). Microorganisms are well-suited within this field, as the rapid-generation times allows researchers to explore evolution in action (Elena and Lenski, 2003). This stimulated long-term evolution experiments investigating the phenotypic adaptations of bacterial populations over an infinite number of generations (Lenski, 2004; Barrick *et al.,* 2009). The recent advancements in genomic technologies have transformed the field of laboratory evolution, as bacterial populations can be comparatively monitored in highly controlled environments to decipher the genetic basis of adaptation (Herring *et al.,* 2006; Hegreness and Kishony, 2007). Despite the genomic technologies in place, the application of evolutionary studies within the field of microbiology has been limited due to the lack of experimental approaches available.

1.2.2 Experimental approaches in the study of antimicrobial resistance

The study of resistance traditionally relied upon the application of strong selective pressures, to determine bacterial survival under fixed antibiotic concentrations (Pankey and Sabath, 2004; Michel *et al.,* 2008). The study of drug resistance by agar selection or disc diffusion has permitted the selection of single step resistant mutations with large phenotypic effects (Palmer and Kishony, 2013).

The implementation of laboratory evolution within this field, has provided valuable insights into the genetic mutations underpinning resistance development (Demerec, 1945; Albert *et al.,* 2005; Friedman, Alder and Silverman, 2006; Lee *et al.,* 2010). These experiments typically utilised the Mutant Selection Window (MSW), which is relatively passive and excludes the multiple mutations which appear under sustained drug selection (Dong *et al.*, 1999; Drlica, 2003; Toprak *et al.,* 2013). Whereby, when

a resistant mutation appears within a population, the selected antimicrobial concentration is no longer providing the selective pressures required to allow mutants to evolve higher levels of resistance (Toprak *et al.,* 2013). In addition, ALE experiments using serial passage, frequently include bottlenecks, which severely affect the evolutionary dynamics of bacterial populations (Wahl, Gerrish and Saika-Voivod, 2002; Barrick and Lenski, 2013).

To explore the multi-step evolution of antibiotic resistance, ALE experiments were adapted to incorporate continuous culture and spatial gradients (Palmer and Kishony, 2013). Continuous culture devices include, chemostats, which facilitate the external control of batch cultures, through the fixed addition of increased concentrations of antibiotics, to maintain the selective pressures on evolving bacterial populations (Dykhuizen and Hartl, 1983; Gresham and Hong, 2015). The morbidostat developed by Toprak *et al*. (2013), is an innovative selection device, which regulates cell growth by automatically adjusting antibiotic concentrations, in response to antibiotic inhibition (Figure 1.3) (Toprak *et al.*, 2013).

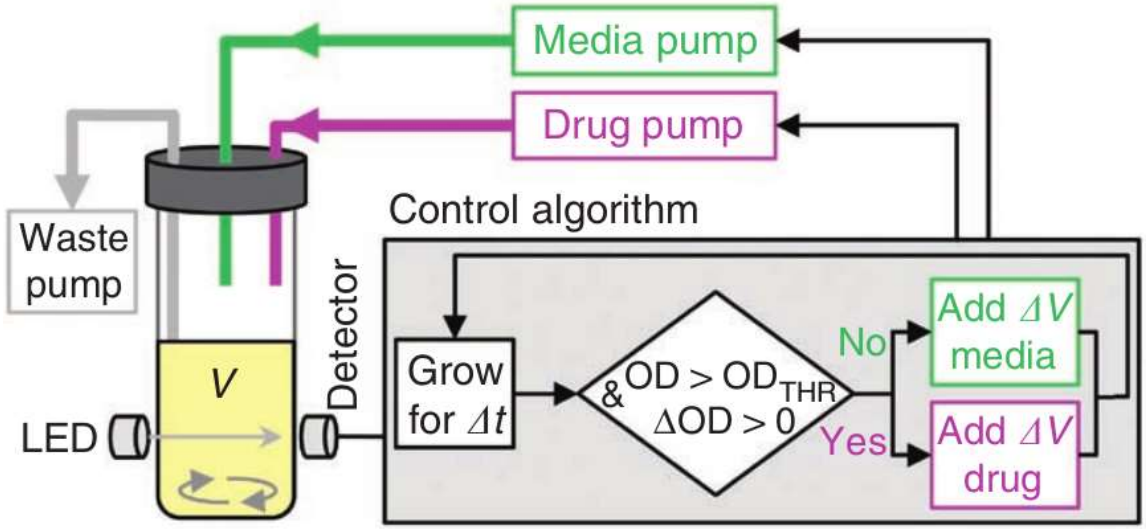


Figure 1.3 The Morbidostat algorithm. The process of bacterial growth inhibition of evolving bacterial populations within the morbidostat in response to resistance development through the external regulation of drug concentration. Image taken from: Toprak *et al.* (2013).

In addition to continuous culture, selective pressures can be maintained in spatial gradients using microfluidic devices, which contain microenvironments of increased antibiotic concentrations (Zhang *et al.,* 2011). Similarly, an experimental device known as the microbial evolution and growth arena (MEGA)-plate, permits the study of multi-step resistance evolution of distinct bacterial populations, using a highly structured spatiotemporal gradient of antibiotics (Figure 1.4) (Baym *et al.,* 2016b). Although the laboratory evolution studies varied in experimental approach, the findings have suggested that the mutational pathways to antibiotic resistance are highly structured and reproducible (Zhang *et al.,* 2011; Toprak *et al.,* 2013; Baym *et al.,* 2016b).

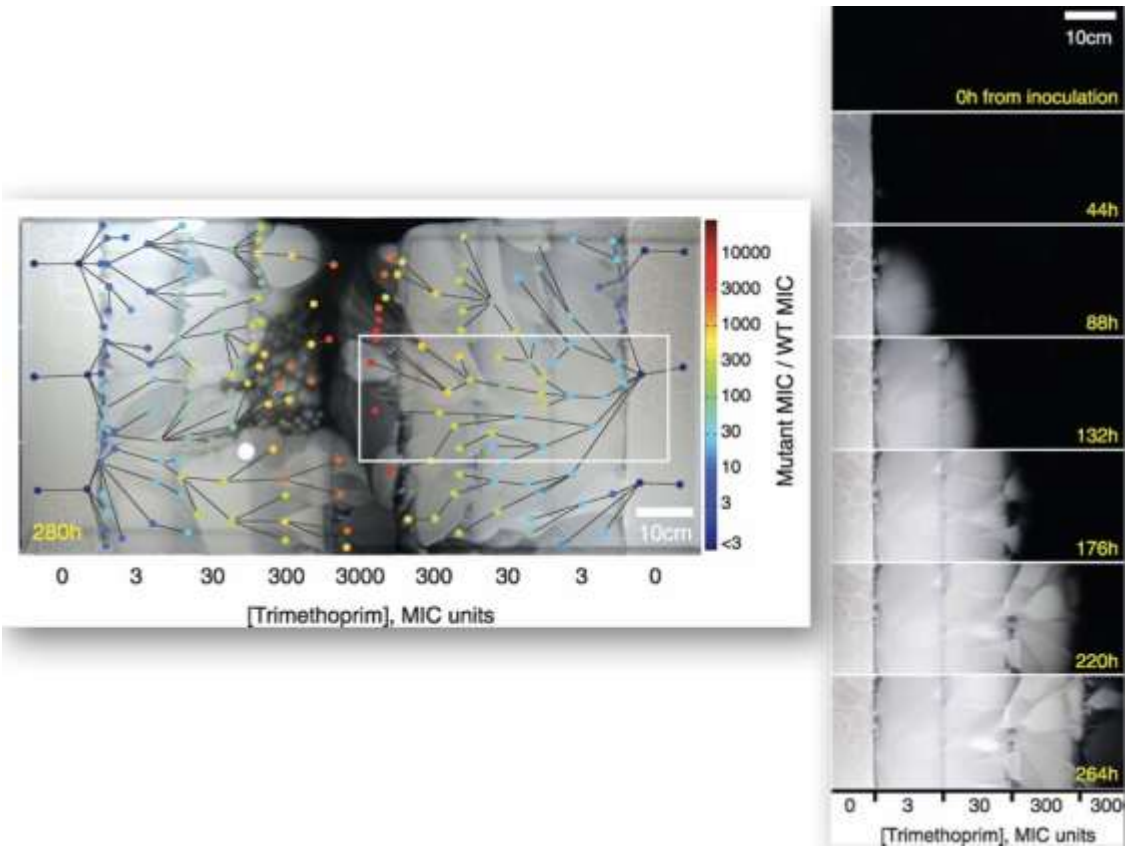


Figure 1.4 The MEGA-Plate experimental device. Displaying *E. coli* under trimethoprim selection over a period of 12-days, with time lapse imagining of the highlighted section. The coloured dots represent the sampling locations at given antimicrobial concentrations, connected by lines indicative of mutational ancestry. Image adapted from: Baym *et al.* (2016b).

In modern medicine, combinational therapy has been highly effective, through employing various inhibition mechanisms, in the treatment of drug resistant infections (Rybak and McGrath, 1996). Despite this success, the effects of chosen antibiotic combinations and their ensuing interactions on resistance evolution, remains unclear (Torella, Chait and Kishony, 2010). ALE investigations into the cross resistance and epistatic interactions between synergistic and antagonistic drug pairs, revealed that for both *E. coli* and *Staphylococcus aureus* synergistic combinations which are preferred clinically, may favour resistance evolution (Michel *et al.,* 2008; Hegreness *et al.,* 2008). Moreover, antagonistic interactions have been valued for their ability to

unfold biologically functions, and what holds even greater potential is their ability to offset the evolution of resistance (Yeh *et al.,* 2010). Further to combinational therapy, the intermittent rotation of drug treatments using alternative therapy, has been shown to constrain the evolution of resistance (Kim, Lieberman and Kishony, 2014). The findings from ALE experiments have challenged prescribing practices and have resulted in altered recommendations to slow resistance evolution (Jahn *et al.,* 2017).

1.2.3 The future of ALE Studies

There is considerable debate regarding the translation of evolutionary findings from *in vitro* to *in vivo*, as the evolved mutations may differ when the host environment is accounted for (Yang *et al.,* 2011). The immune system plays an imperative role during antimicrobial therapy and it is sensed that *in vitro* ALE studies heavily underestimate the immune system involvement (Geli *et al.,* 2012). However, the advancements in Whole Genome Sequencing (WGS), have enabled researchers to monitor the genomic evolution of bacterial pathogens during endemic infections (Harris *et al.,* 2010; Croucher *et al.,* 2011; Kennemann *et al.,* 2011). A problem with this approach is that it becomes difficult to decipher the adaptive mutations, from neutral mutations which have been fixed within a population by genetic drift (Goodarzi, Hottes and Tavazoie, 2009). To overcome this, comparative WGS has been combined with ALE experiments, to reveal the recurrent patterns of evolution implicated in resistance development *in vivo,* by sampling human subjects during infectious disease outbreaks and in the treatment of complicated bacterial infections (Mwangi *et al.,* 2007; Howden *et al.,* 2011; Lieberman *et al.,* 2011; Comas *et al.,* 2012).

There are various challenges to modelling bacterial evolution *in vivo* (Metcalf *et al.,* 2015). The within-host interactions and increased bacterial genomic plasticity, can have considerable effects on bacterial adaptation (Geli *et al.,* 2012; Woodcock *et al.,* 2017). An *in vivo* study monitoring the genome evolution of a prototype asymptomatic bacteriuria (ABU) strain of *E. coli* following bladder colonisation, revealed the bacterial evolution was distinct for each subject, but followed key changes in genes involved in metabolism and virulence (Zdziarski *et al.,* 2010). Studying the evolution of resistance in humans through the course of antimicrobial therapy, would provide invaluable information regarding the effects of host imprints on bacterial evolution, however ethical restrictions largely prevent such studies (Kouyos *et al.,* 2014). Moreover, there is a lack of a suitable animal model to explore the adaptive bacterial evolution *in vivo* (Giraud *et al.,* 2001). In antimalarial drug resistance studies, the *Plasmodium chabaudi* mouse malaria model, has provided valuable insights into the within host evolutionary dynamics during parasite co-infection (Huijben *et al.,* 2010).

Mathematical models have played a crucial role in increasing our understanding of antimicrobial resistance (Opatowski *et al.,* 2011). One of the earliest models in this field was produced by Lipsitch and Levin (1997), which evaluated the bacterial population dynamics and the likelihood of resistance mutations arising during antibiotic treatment. Since then, mathematical models have advanced significantly, and are now being used to investigate various aspects of AMR including: infection prevention and control, evolution, epidemiology and antimicrobial treatment protocols (Bonten, Austin and Lipsitch, 2001; Grundmann and Hellriegel, 2006; Consortium 2007; Temime *et al.,* 2008; Bonhoeffer, Liptsitch and Levin, 1997; D'Agata *et al.,*

2007, 2008). In the future, mathematical models could be combined with ALE, to interpret empirical laboratory observations and to assess how likely it is for a resistance mutation to occur, and whether the selected antimicrobial treatment regimen, minimises the risk of resistance evolution (Spicknall *et al.,* 2013; Ankomah and Levin, 2014).

1.3 Fitness costs associated with resistance development

Antibiotic resistance is frequently associated with bacterial fitness costs (Vogwill and MacLean, 2015). These fitness costs frequently manifest as reductions in growth, virulence and competitiveness within environments free of antibiotics (Enne *et al.,* 2004; Binet and Maurelli, 2005; Johnson *et al.,* 2005; Sjolund *et al.,* 2005; Andersson, 2006). The antimicrobial fitness costs are central to the evolutionary dynamics of resistance, by contributing to the rate and spread of resistance in the presence of antibiotics, but also, the rate of clearance or persistence in their absence (Andersson and Levin, 1999). Subsequently, fitness is used as a predictor of resistance development, as irrespective of rate, costly mutants will be selected against (Andersson, 2003). This was demonstrated in a study of rifampicin-resistance in *Staphylococcus aureus*, where the most frequently isolated resistant mutants in a clinical setting presented little or no fitness costs (O'Neil *et al.,* 2006). Therefore, when utilising ALE models, it is necessary to define the genotypes underlying antimicrobial resistance development and measure the corresponding phenotypic effects on bacterial fitness (Andersson, 2006; Palmer and Kishony, 2013).

1.4 Phylogenetic interpretation of ALE findings

Phylogenetic reconstruction can be performed alongside ALE studies, as the evolutionary history of a strain can provide valuable information regarding both the mutation and transmission events underlying resistance evolution (Baldauf, 2003; Palmer and Kishony, 2013). Traditionally in ALE the reoccurrence of a mutation in repeated experiments provides evidence of parallel evolution, however phylogenetics can be applied to determine strong signatures of adaptive evolution through the identification of mutations independently in related bacterial strains, this is referred to as evolutionary convergence (Didelot and Falush, 2007; Laehnemann *et al.,* 2014). The phylogenetic differentiation of neutral mutations from true resistance conferring mutations, allows researchers to establish whether the identified laboratory evolved polymorphisms can be utilised as markers of antimicrobial resistance (Feuerriegel, Köser and Niemann, 2014).

1.5 Treatment strategies informed by ALE

The desired bacterial inhibitory effects of antimicrobial treatment routines are heavily compromised by the evolution of resistance, which has emphasised the requirement for techniques to control this evolutionary process (Baym, Stone and Kishony, 2016a). ALE tools have provided great insights into the evolutionary dynamics of resistance during single, alternating and multi-drug environments, and have suggested the potential for the manipulation of these interactions to revert the evolution of resistance (Hegreness *et al.*, 2008; Kim, Lieberman and Kishony, 2014; Baym, Stone and Kishony, 2016a). For this to be accomplished ALE must be coupled with WGS, to not only identify mutated genes which confer resistance, but to decipher

the adaptive mutational pathways to resistance for range of antimicrobial agents, independently and in combination (Jansen, Barbosa and Schulenburg, 2013; Palmer and Kishony, 2013). The simplest application of ALE models, is within diagnostic experimental evolution, where susceptible strains could be isolated from patients and evolved in real-time to predict the potential for resistance, and in cases where resistance has already emerged, ALE could inform treatment strategies which minimise subsequent evolution (Lieberman *et al.,* 2011; Toprak *et al.,* 2012; Palmer *et al.,* 2015; Woods and Read, 2015).

1.6 Aims and objectives

The prevalence of antimicrobial resistance is an ever-increasing global concern. Antimicrobial treatment failures are exacerbated by the rapid evolution of resistance amongst bacterial pathogens. Therefore, in addition to developing novel antimicrobial agents, there is increased interest in exploring the underlying genotypic and phenotypic factors of resistance evolution. Several studies have suggested the potential of ALE, to uncover the adaptive mutational pathways to antimicrobial resistance. ALE models can be used to monitor the evolutionary potential of bacterial strains, against a range of antimicrobial agents and inform therapeutic decisions based on evolutionary risk management. In addition, the identification of underlying mutational mechanisms of resistance, will provide new opportunities for drug development. The development of a simplified ALE approach, offers a unique opportunity for the application of such tools in scientific teaching laboratories to observe evolution in real-time.

The aims of this study are:

1. To develop a simplified ALE tool, to explore the evolution of antimicrobial resistance through mutations.

2. Using comparative WGS, to identify SNPs associated with resistance development and through confirmatory PCR determine the reproducibility of the mutational pathways.

3. Using bioinformatic analysis explore the functional impacts of the identified adaptive mutations and establish through molecular docking whether the

binding of antimicrobial agents to their given antimicrobial targets is disrupted.

4. Explore the phylogeny of the mutated genes within the *Enterobacteriaceae* family and distinguish any signatures of convergent evolution.

5. Determine the fitness costs associated with antimicrobial resistance development.

Chapter 2: Materials and methods

2.1 Chemicals and reagents

All chemicals and reagents used within the study were purchased from Sigma-Aldrich, unless stated otherwise.

2.2 Bacterial strains

In this study, drug sensitive strains of *E. coli* (NCTC 11954) and *Proteus mirabilis* (*P. mirabilis*) (NCTC 13376) were used, as provided by Kingston University. The *E. coli* MG1655 strain was provided by Dr Nick Coldham from the Animal and Plant Health Agency (APHA). Motile clinical strains of *E. coli* and Enteropathogenic *E. coli* (EPEC) strains were kindly provided by Dr Philip Aldridge, Newcastle University (NU) and Dr Dan Brown, Imperial College London (ICL). Further Details of bacterial strains are provided in (Table 2.1).

Table 2.1 Bacterial strains used within this study. Table showing the bacterial ID of the *E. coli* and *P. mirabilis* strains utilised within this study and information regarding the strain origin and source.

Strain ID	Strain Type	Source
Escherichia coli 11954	Laboratory Strain	KU
Proteus mirabilis 13376	Laboratory Strain	KU
Escherichia coli MG1655	Laboratory Strain	APHA
Escherichia coli UTI89	Clinical Strain Class 2	NU
Escherichia coli CFT073	Clinical Strain Class 2	NU
Escherichia coli TPA2743	Clinical Strain EPEC	ICL
Escherichia coli TPA4792	Clinical Strain EPEC	ICL

2.3 Bacterial cultures and inoculum preparation

Frozen stocks of *P. mirabilis* and *E. coli* were isolated and grown on brain heart infusion agar and were incubated overnight (O/N) aerobically at 37°C for 12 hours. To prepare the inoculum for use, the bacteria were suspended in phosphate buffered saline (PBS), to obtain a cell density equivalent to that of a 0.5 McFarland standard, with an approximate cell count of 1.50×10^8 colony forming units/mL (CFU/mL). Bacterial strains sourced from outside KU, were transported on Nutrient Agar Plates and upon arrival sub-cultured and incubated O/N aerobically at 37°C for 12 hours prior to cryopreservation (see below section 2.3.1). Details of the media utilised within this study are provided in table (2.2).

Table 2.2 Media used within this study. Table showing the media used within the study, their composition and the working concentrations of each component.

Medium	Composition	Concentration (g/L)
Mueller Hinton (MH) Agar	Agar Beef Infusion Solids Casein Hydrolysate Starch	17.0 2.0 17.5 1.5
Mueller Hinton Broth	Beef Infusion Solids Casein Hydrolysate Starch	2.0 17.5 1.5
M9 Minimal Medium	Na_2HPO_4 KH_2PO_4 NH_4Cl NaCl D (+) -Glucose $MgSO_4$ $CaCl_2$ Thiamine	33.9 15.0 5.0 2.5 4.0 (20mL of 20% Solution) 0.2407 (2mL 1M Solution) 0.011 (0.1mL 1M Solution) 0.0005 (0.1mL 0.5% w/v Solution)
M9 Minimal Medium Agar	M9 Minimal Medium Bacteriological Agar	 30.0
Motility Agar	Pancreatic Digest of Casein NaCl Beef Extract Bacteriological Agar	10.0 5.0 3.0 4.0
Motility Agar + Triphenyltetrazolium Chloride (TTC)	Pancreatic Digest of Casein NaCl Beef Extract Bacteriological Agar TTC	10.0 5.0 3.0 4.0 0.5
Mueller Hinton Agar of Semi Solid Consistency (25%)	Agar Beef Infusion Solids Casein Hydrolysate Starch	4.25 0.5 4.375 0.375
Cysteine-Lactose-Electrolyte Deficient (CLED) Agar	Agar Beef Extract Bromo thymol blue Casein enzymatic hydrolysate L-cystine Lactose Peptic digest of animal tissue	15.0 3.0 0.002 4.0 0.128 10.0 4.0
Nutrient Agar	Agar Beef extract Peptic digest of animal tissue	15.0 3.0 5.0

2.3.1 Preparation of bacterial stocks

Microbank™ vials (Pro-Lab diagnostics) facilitated the long-term storage and retrieval of bacterial cultures. For long term storage, bacterial strains were transferred to Microbank™ vials and stored at -80°C. *E. coli* strains were used to inoculate MH plates and were incubated at 37°C within a static incubator overnight for 18-24 hours. Single colonies of a pure culture were used to inoculate the cryopreservative fluid, to emulsify, the vial was then inverted 4-5 times and immediately stored at -80°C.

2.3.2 Recovery of bacterial stocks

The Microbank™ vial containing the culture of interest was removed from storage at -80°C. The cryovial was asceptically opened and using a sterile swab a sample of the culture was taken. The cryovial was then closed immediately and returned to storage at -80°C. The swab was then used to inoculate a fresh MH plate which was incubated at 37°C O/N aerobically within a static incubator. The culture was then passaged a further three times before being utilised within any experiment.

2.4 Preparation of antimicrobial stock solutions

The antimicrobial agents utilised within this study are detailed below (Table 2.3). Stock solutions of each antimicrobial agent were prepared as per Sigma Aldrich supplier guidelines. An example calculation used in the preparation of antimicrobial stock solutions is shown in table (2.4).

Table 2.3 Antimicrobial agents used within this study. Table showing the antimicrobial agents utilised within this study and information regarding the antimicrobial class and minimal inhibitory concentration.

Antimicrobial Agent	Class of Antimicrobial	Minimal Inhibitory concentration (mg/L)
Chloramphenicol	Chloramphenicol	8
Meropenem	Carbapenem	2
Trimethoprim	Benzene and substituted derivative	2
Colistin	Polypeptide	2
Levofloxacin	Quinolone	1
Gentamicin	Aminoglycoside	2

Table 2.4 Calculation example for the preparation of antimicrobial stock solutions. Table showing an example calculation of the preparation of a 200mg/L trimethoprim stock solution.

Antimicrobial Agent and Stock Concentration	Stock Solution Calculation
Trimethoprim (200mg/L)	200mg/L ↓ ÷1000 0.20mg/mL ↓ ÷1000 0.0002g/mL (X10 to allow a scalable reading on the weighing scales) ↓ x10 0.002g/mL (Followed by a 1 in 10 dilution of the prepared solution)

Once prepared, stock solutions were either stored at 4°C or -20°C in accordance to manufacturer guidelines. The required volume of antimicrobial stock solution to be added to the molten MH agar preparation was calculated using the formula $C_1V_1=C_2V_2$.

2.4.1 Antibiotic sensitivity testing methods

The first sensitivity test conducted was the disc diffusion method. MH plates were prepared and seeded with fresh cultures of both organisms with a cell density of 1.50 x 10^8 CFU/mL. Commercially obtained discs with the standard concentration of each antimicrobial agent, were lightly pressed against the agar surface and incubated O/N aerobically at 37°C for 12 hours. After incubation, the zones were measured and compared to the standard interpretation chart provided by the European Committee on Antimicrobial Susceptibility Testing (EUCAST, 2018).

To determine the minimum inhibitory concentration (MIC) the microbroth dilution method was conducted within a microtiter plate. A total MH broth volume of 100µL was placed into used wells, with 200µL of the maximum concentration of antimicrobial agent placed into its corresponding well, to give a final concentration of 128mg/L. Subsequently, two-fold serial dilutions were performed to produce a series of concentrations of each given antimicrobial agent (0.25, 0.50, 1.0, 2.0, 4.0, 8.0, 16.0, 32.0, 64.0 and 128mg/L). MH broth containing no antibiotic was inoculated with the bacterial strains as the positive control and uninoculated MH broth as the negative control. The inoculum was prepared as stated above (2.3) and 100µL of each bacterium was added to the corresponding wells. The plates were then incubated aerobically at 37°C and the MIC was identified.

2.5 Characterisation of bacterial strains

The initial confirmatory test for bacterial identification was a Gram-Stain. This staining technique was used to differentiate between Gram-positive and Gram-negative microorganisms and to provide further details regarding bacterial morphology (Gregersen, 1978). Here, bacterial colonies were heat fixed onto microscope slides followed by staining with a series of reagents using the traditional Gram-stain procedure outlined by Bartholomew and Mittwer (1952). All reagents were applied for a duration of one minute, excluding acetone which was left for ten seconds. Each stage of the staining process was followed by a wash step where slides were flooded with distilled water (dH_2O). Subsequently, slides were blotted dry and analysed using light microscopy.

In addition to the primary identification using the Gram procedure, differential tests were required to distinguish *E. coli* from *P. mirabilis*, and for the identification of potential contaminants. Below is a description of the media and reagent components used for bacterial identification (Table 2.5). Details of the expected results are stipulated in table (2.6).

Table 2.5 Media and reagents used in bacterial identification. Table displaying the media and reagent components required for the bacterial identification tests utilised within this study.

Bacterial Identification Test	Reagents/Media	Components
Catalase	Catalase Reagent	Hydrogen Peroxide (H_2O_2)
Oxidase	Oxidase Reagent	0.01g Tetramethyl-p-phenylenediamine (TMPD) in aqueous dH_2O
Lactose Fermentation	MacConkey Agar	17.0 g/L Peptone 3.0g/L Proteose Peptone 10.0g/L Lactose 1.5g/L Bile Salts No. 3 5.0g/L Sodium Chloride (NaCl) 13.5g/L Agar 0.03g/L Neutral Red 0.001g/L Crystal Violet
Indole Test	Tryptone Broth Kovacs Reagent	10g Tryptone 5g Sodium Chloride (NaCl) in aqueous dH_2O 150mL Amyl alcohol 10g p-dimethylaminobenzaldehyde (DMAB) 50mL Hydrogen Chloride (HCl)
Urease test	Urea Broth	9g/L Potassium dihydrogen orthophosphate (KH_2PO_4) 0.100g/L Yeast Extract 9.5g/L Anhydrous disodium hydrogen phosphate (Na_2HPO_4) 20g/L Urea 0.01g Phenol red

2.5.1 Catalase test

A catalase test is used predominantly in the identification of *Enterobacteriaceae*. Bacteria which possess the enzyme catalase can be identified using H_2O_2 and observing the formation of bubbles (Taylor and Achanzar, 1972). Bacterial colonies were transferred to a microscope slide using a sterile loop. A drop of 3% H_2O_2 was

added and apparent bubbling was indicative of a catalase positive organism.

2.5.2 Oxidase test

The oxidase test is a technique which works again through the identification of a bacterial enzyme, in this case Cytochrome C Oxidase which is essential during the final stage of aerobic respiration (Kovacs, 1956). The bacterial inoculum was obtained using Whatman filter paper, by scrapping bacterial colonies from a freshly prepared O/N culture. This was followed by the addition of a drop of 1% TMPD to the filter paper and the observation of a dark blue colour within ten seconds was indicative of an oxidase positive organism.

2.5.3 Lactose fermentation

Bacterial identification was also achieved using MacConkey agar which contains the pH indicator neutral red which turns red in acidic environments through lactose fermentation (MacConkey, 1905). Lactose fermenting organisms were identified as pink/red colonies on MacConkey agar following O/N aerobic incubation at 37°C for 12 hours.

2.5.4 Indole test

Following the identification of lactose fermenting organisms on MacConkey agar. Gram-negative organisms were differentiated using an Indole test (Vracko and Sherris, 1963). Bacterial strains were inoculated in sterile tubes containing tryptone water and incubated O/N aerobically at 37°C for 12 hours. Following incubation, a drop of Kovac's reagent was added (McWilliams, 2009). A positive result was indicted

by a red/pink alcohol layer on the surface of the solution.

2.5.5 Urease test

The final bacterial identification technique used within this study was the urease test, utilised explicitly for the identification of *P. mirabilis* (Christensen, 1946). Bacteria were inoculated in 10mL of urea broth containing the indicator phenol red and were incubated at 37°C. A positive result for *Proteus* spp. was identified by a rapid change in broth colour to bright pink, which can occur within 1-2 hours.

Table 2.6 Bacterial identification techniques for *E. coli* and *P. Mirabilis*. Table displaying the expected results from the bacterial identification tests to differentiate *E. coli* from *P. mirabilis* based upon several distinctive characteristics.

Characteristic	***Escherichia coli***	***Proteus mirabilis***
Gram	Negative Rod	Negative Rod
Catalase	Positive	Positive
Oxidase	Negative	Negative
Lactose	Positive	Negative
Indole	Positive	Negative
Urease	Negative	Positive

Genotypic confirmation of bacterial strains was achieved through whole genome sequencing, discussed further in Chapter 3.

2.6 Enumeration of bacteria

Bacterial enumeration was achieved through quantification using CFU/mL calculations. When enumeration was required, bacterial suspensions were transferred to a serial dilution plate containing PBS. Serial dilutions of the sample were prepared to produce a dilution series of 10^{-1} to 10^{-8}. Ten µL of diluted suspensions were then plated onto MH agar and incubated at 37°C O/N for 12 hours. Plates with dilutions containing 20-200 bacterial colonies were counted and the CFU/mL was determined using the formula: CFU/mL = (no. of colonies x dilution factor) x (1000 / volume plated on MH (µL)).

2.7 Method development and optimisation

2.7.1 Preliminary experiment: production of Bioplates

The preliminary test was to devise a simple model, which allowed the study of the evolution of antimicrobial resistance under antimicrobial exposure. To do this, diphasic plates were produced by the insertion of Bioplates (see two prototypes with dimensions below figure 2.1).

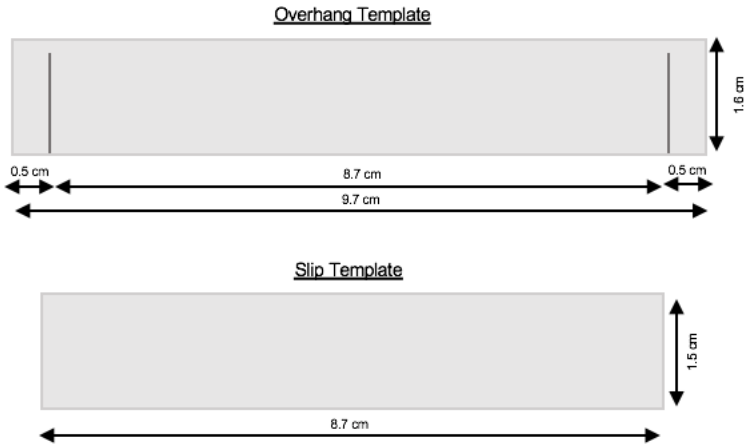


Figure 2.1 Template of the prototype Bioplates. Figure displaying the “Overhang” and “Slip” Bioplate templates with dimensions.

2.7.1.1 Preliminary diphasic plate experiment: MH agar

Diphasic plates were produced by the sterile insertion of autoclaved Bioplates, the experiment was repeated to test both the overhang and slip Bioplates. MH agar was prepared and autoclaved at 121ºC at 115psi for 15 minutes. Bioplates were sealed into place using sterilised MH agar and allowed to set. The two phases contained non-antibiotic and antibiotic containing agar at MIC concentration. The antibiotic was added to the media post-sterilisation. Once the two phases were poured and set, the Bioplates were aseptically removed. The gap was then filled using sterile MH agar. The bacteria were inoculated onto the non-antibiotic phase as 3 separate 10µL drops. The diphasic plates were then incubated at 37 ºC over a period of 7 days (Figure 2.2). Control plates containing MH agar plus the MIC concentration of the selected antimicrobial agent were produced and incubated over the duration of the study, to test the effectiveness of the antimicrobial agent.

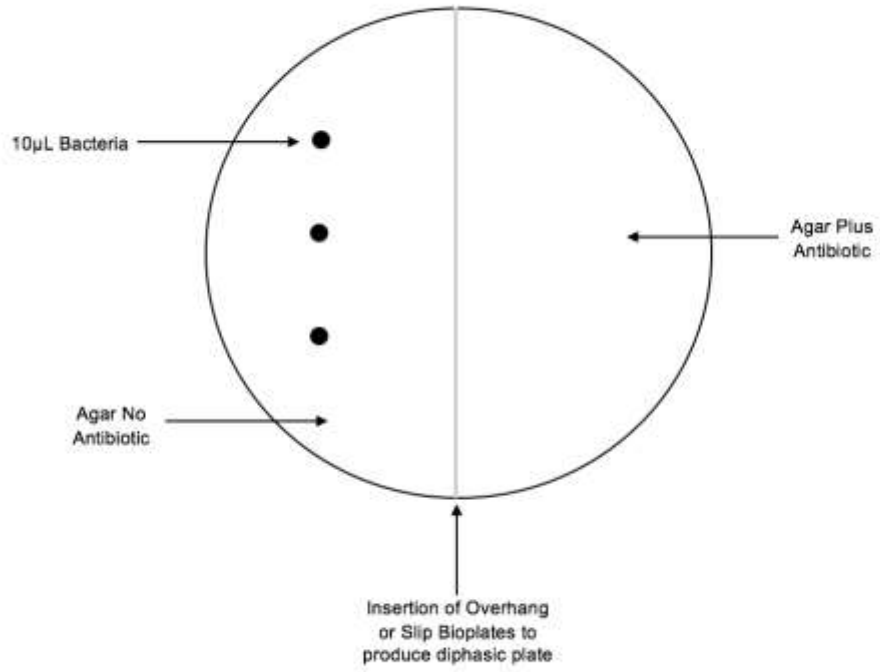


Figure 2.2 The preliminary diphasic plate layout. Diagram displaying the plate set-up used in the production of the diphasic plate. Plates were divided using Bioplates to produce a gradient of antimicrobial agent. Plates were inoculated with three 10µL (1.50×10^8 CFU/mL) drops of drug sensitive bacterial suspensions on the phase containing no antimicrobial agent.

2.7.2 Preliminary filter tip experiment

The filter tip experimental method was produced to explore the flow of microorganisms through a controlled laboratory system, to study the evolution of antimicrobial resistance. The initial experiments were conducted primarily to see whether organisms could flow through the filter tip.

To prepare for the preliminary filter tip experiment P1000 micropipette tips and filter tips were sterilised within glass petri dishes. The filter tips were divided into two separate dishes representing the two different types of agar. The P1000

micropipettes were removed asceptically from the glass petri dish and placed within a micropipette tip box. Nutrient and MacConkey Agar was poured over the filter tips in their respective glass petri dishes and the agar was left to set. Following on from this, 600µL of Nutrient agar was used to plug the base of the micropipette tip. Once the plugged agar had set, sterile forceps were used to transfer the filter tips containing MacConkey agar to the micropipette tip, followed by the placement of the Nutrient agar filter tips directly on top. The remaining gap was filled with Nutrient agar using a Pasteur pipette (Figure 2.3). Bacterial suspensions were produced as described in (2.3), and 10µL (1.50×10^8 CFU/mL) of each suspension was used to inoculate the surface of tip via a stab culture. Micropipette tip boxes were then placed within a Tupperware box, which had been layered with a water covered absorbent pad. The lid of the box contained two 0.2µM filters, inserted to facilitate the passage of sterile air, the box was then placed in the static aerobic incubator at 37ºC and analysed daily for a period of 7 days. To confirm the presence of bacteria, filter tips were aseptically removed and dotted onto Nutrient and CLED agar.

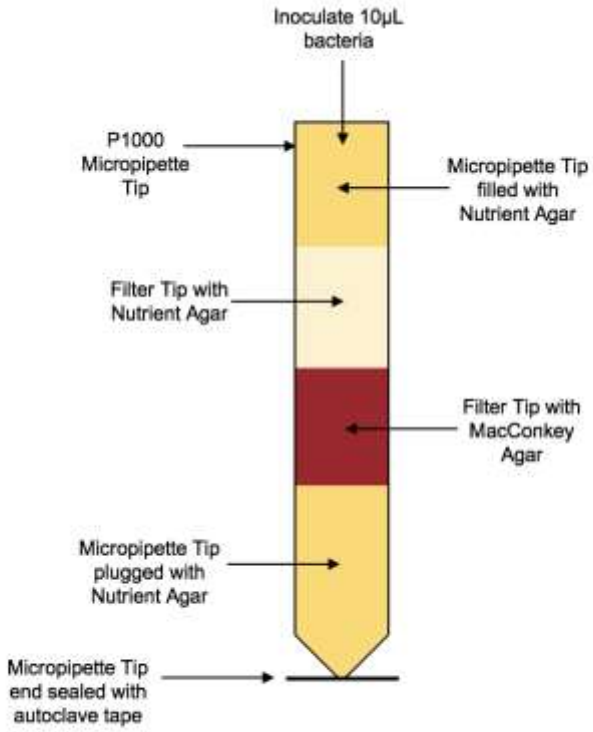


Figure 2.3 The preliminary filter tip experiment with MacConkey agar. Diagram showing the set-up of the filter tip experiment with MacConkey agar. P1000 micropipette tips were sealed with autoclave tape and sterilised. The micropipette tips were then plugged with Nutrient agar. Filter tips were covered in both MacConkey and Nutrient agar and allowed to solidify. The filter tips were then asceptically transferred to the micropipette tips. Nutrient agar was then used to fill the remainder of the micropipette tip. The micropipette tip was then inoculated with 10µL (1.50×10^8 CFU/mL) of bacterial suspension.

2.7.2.1 Filter tip experiment

The experiment described in (2.7.2) was repeated with the inclusion of MH agar. Here the filter tips were divided into two separate dishes representing the differing antimicrobial concentrations. Agar containing both the MIC of the selected antibiotic and MH with no antibiotic was poured over the filter tips respectively. The P1000 micropipette tips were plugged and any remaining gaps were filled with MH agar (Figure 2.4).

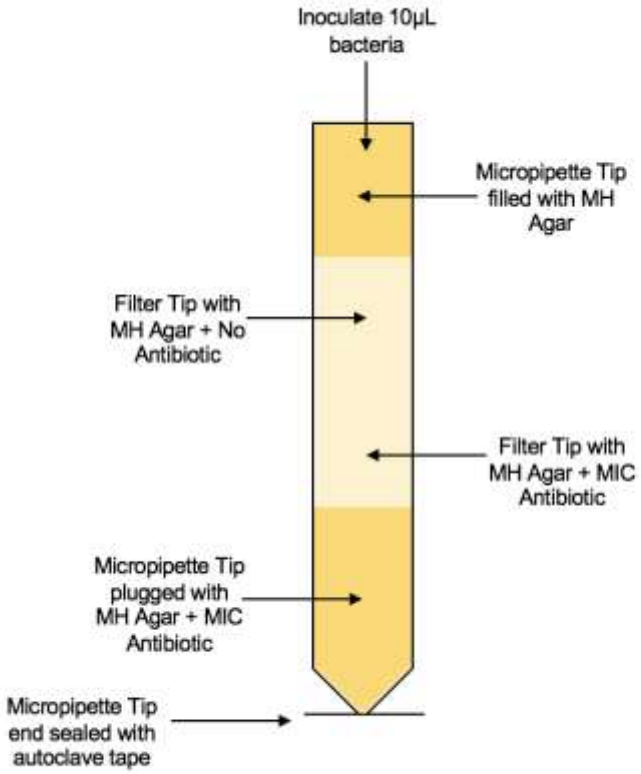


Figure 2.4 The modified filter tip experiment with antimicrobial agents. Diagram displaying the set-up of the filter tip experiment with antimicrobial agents. P1000 micropipette tips were sealed with autoclave tape and sterilised. The micropipette tips were then plugged with MH agar. Filter tips were covered in both MH agar and MH agar plus antimicrobial agent and allowed to solidify. The filter tips were then asceptically transferred to the micropipette tips. MH agar was then used to fill the remainder of the micropipette tip. The micropipette tip was then inoculated with 10µL (1.50×10^8 CFU/mL) of drug sensitive bacterial suspension.

Following incubation, daily triplicates were removed and analysed over a period of seven days. The analysis involved the aseptic removal of filter tips, the initial MH tip was discarded into a Virkon waste pot and the tip containing the antimicrobial agent and thus any potentially resistant organisms was plated onto MH agar plates of the selected antimicrobial concentration to maintain the selective pressure. After O/N incubation at 37ºC plates were analysed for growth, providing growth and confirmation of phenotypic resistance, isolates were cryopreserved.

2.7.3 Triphasic circle plate experiment

To facilitate the radial growth of the organisms the triphasic circle plate method was developed. Here within a glass petri-dish using concentric rings of different sizes, a concentration gradient of three phases of antimicrobial agent was produced representing: zero, MIC and x10 MIC. The bacteria were inoculated in the inner zone containing no antibiotic and plates were observed daily to see whether growth would proceed to the outer phases containing increasing concentrations of a selected antibiotic. To achieve this, a thin layer of Mueller Hinton agar was poured and two concentric rings were placed into the agar and allowed to set. Antimicrobial agents were added to the media once cooled and phases were poured from the inside out, with the final step being the removal of the rings. Plates were then inoculated with 10µL of a 0.5 McFarland standard of bacteria (1.50×10^8 CFU/mL) as described in (2.3) (Figure 2.5). Control plates were produced using the antibiotic containing media used in the preparation of the triphasic circle plate, which were tested daily over the experimental course. When a potentially resistant organism arises, samples were taken and plated onto MH plates of the selective antibiotic concentration and cryopreserved once resistance had been confirmed.

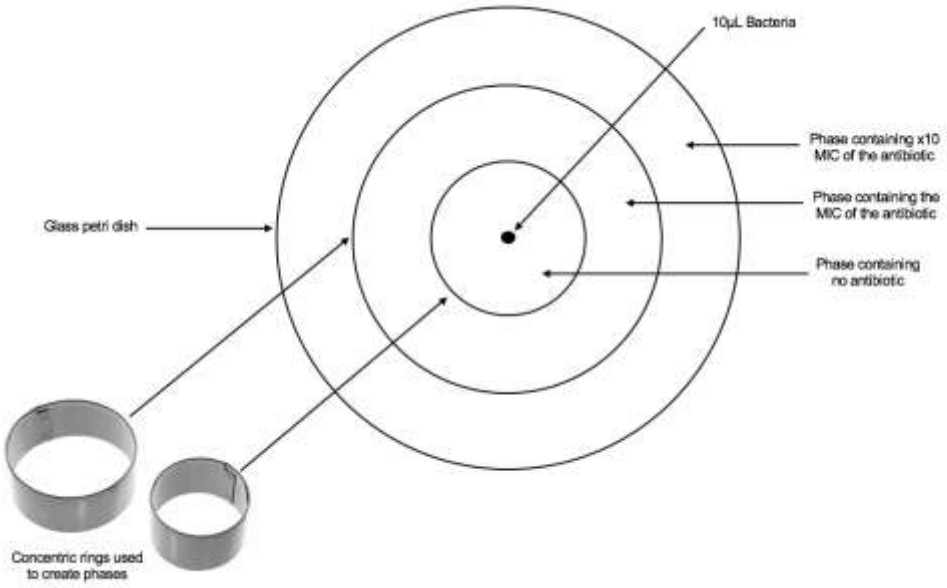


Figure 2.5 The triphasic circle plate. Diagram displaying the triphasic circle plate set-up and the concentric rings used to produce the antibiotic phases. The concentric rings were added to the triphasic plate to produce a three-phase antimicrobial gradient raging from zero to x10 the MIC of a given antimicrobial agent. The triphasic circle plates were inoculated with a 10µL (1.50×10^8 CFU/mL) drop of a drug sensitive bacterial suspension in the central phase containing no antimicrobial agent.

2.7.4 Broth method serial passage

Alongside the evolutionary experiments carried out using spatial temporal gradients, comparative experiments were conducted using serial passage. Here, MH broth was prepared and 4.5mL was transferred to wells of a 12-well Nunc Multidish (ThermoFisher Scientific) in triplicate, and plates were inoculated with 500µL of drug sensitive *E. coli* standardised to an OD_{600} of 0.5 OD units (1.50 x 10^8 CFU/mL). The plates were then covered with a breath-easier plate sealer and incubated aerobically O/N at 37ºC. Daily negative and positive controls were included, see plate set-up figure (2.6).

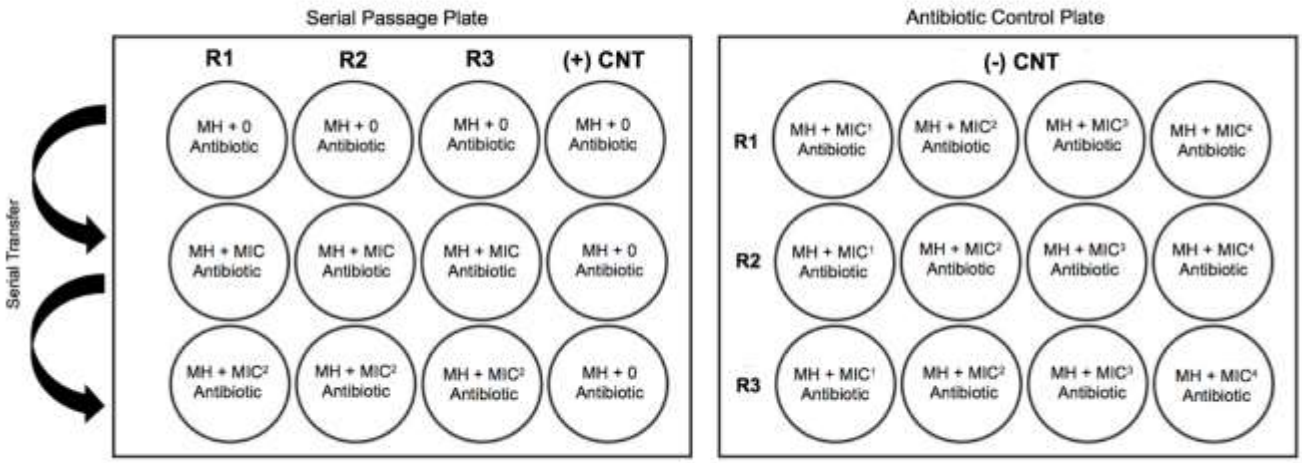


Figure 2.6 The serial passage and antimicrobial control plate set-up. Diagram showing the set-up of the two plates utilised within the broth method. The serial passage plate contains three replicates of isogenic bacterial populations of either drug sensitive *E. coli* or *P. mirabilis* and a positive control containing MH broth. Bacterial populations were inoculated originally into MH broth containing no antibiotic, and following O/N growth 500µL of each culture was transferred to increased antimicrobial concentrations. The antimicrobial concentrations followed the sequential doubling of the MIC for a given antimicrobial agent. The antibiotic control plate provided the experimental negative control. Each antimicrobial broth concentration prepared was inoculated with the originally sensitive bacterial replicates and was monitored for growth following O/N incubation.

Providing growth following O/N incubation, MH broth was prepared with the addition of selected antimicrobial agents, the initial antimicrobial concentration was pre-determined from the MIC microbroth dilution experiments described in section (2.4.1). Then 500µL of bacteria from the previous wells was serially transferred to the freshly prepared broth containing the MIC of the given antimicrobial. This passaging process was repeated every 24 hours with doubling antimicrobial concentrations. To confirm phenotypic resistance 10µL of daily broth cultures was transferred to MH agar plates of the selected antimicrobial concentration, and incubated O/N at 37ºC. Quantification was achieved by daily OD_{600} measurements and CFU/mL calculations.

2.8 Genome sequencing, assembly and alignment

2.8.1 Whole genome sequencing at Microbes NG

Whole genome sequencing was provided by Microbes NG (http://www.microbesng.uk) which is supported by the Biotechnology and Biological Sciences Research Council (BBSRC) (grant number BB/L024209/1). Strains for whole genome sequencing were placed into pre-barcoded bead tubes and were returned to Microbes NG sequencing facility at Birmingham University. Genomic DNA was extracted and purified before being transferred to a Hamilton Microlab STAR automated liquid handling systems for quantification using the Quantit dsDNA HS assay and genomic DNA library preparation using Nextera XT Library Prep Kit (Illumina, San Diego, USA). The pooled library DNA was then quantified prior to sequencing using the Kapa Biosystems Library Quantification Kit for Illumina on a Roche light cycle 96 qPCR machine (https://microbesng.uk/microbesng-faq/). Whole genome sequencing was carried out using Illumina HiSeq or MiSeq using 250bp paired end.

Figure (2.7) displays the four main stages behind the chemistry of illumina sequencing. The initial sample preparation step involves the fragmentation of genomic DNA, followed by the addition of adapters and supplementary motifs which are then thermally amplified and transferred to a flow cell. The templates are then simultaneously amplified by bridge amplification to form clusters. Sequencing by synthesis follows with the addition of fluorescently tagged nucleotides (dNTPs) alongside revisable terminators. After the addition of each nucleotide the clusters are excited and the signal emitted determines the base called. The samples are then

clustered based on sequence similarity and forward and reverse sequences are paired to form contiguous sequences for bioinformatic analysis using an automated pipeline (Illumina, 2018).

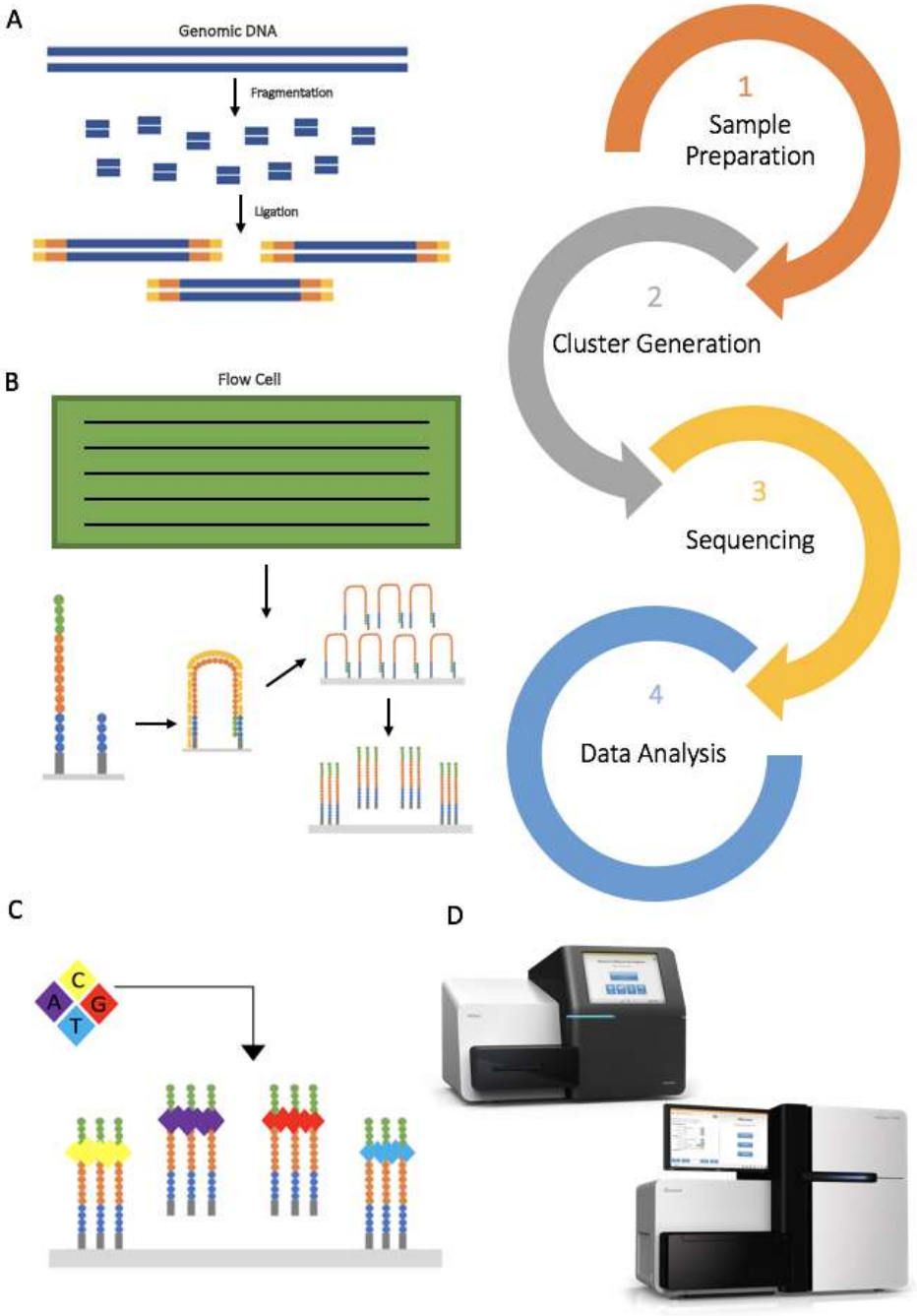


Figure 2.7 Overview of Illumina Next-Generation Sequencing. The image displays the four steps of illumina sequencing: (A) Library Sample Preparation. Here the genomic DNA is fragmented and ligated using specialised adapters. (B) Cluster Generation. The library generated in step 1 is loaded onto a flow cell and the fragments and hybridized into clusters using bridge amplification. (C) Sequencing. Fluorescently labelled nucleotides are added and the emission wavelength and intensity from each cluster is recorded throughout the sequencing cycle. (D) Data Analysis. Reads are then aligned to a reference genome and analysed using various bioinformatics software. The illumina MiSeq and HiSeq Series platform images (D) were obtained from Illumina. (2018).

2.8.2 Whole genome assembly CLC

Whole genome sequencing data retrieved from Microbes NG sequencing facility was analysed using CLC genomics workbench 7.7.5. Sequence data was imported into CLC in Genbank (. gbk) format and saved as DNA sequences. Quality reports were used to assess the quality of the sequences, and failed reads were removed accordingly. The remaining poor quality reads were trimmed based on quality (0.01), ambiguities (0) and length (<50) using the automated trimming tool to ensure consistency between sequences. After each stage during the whole genome assembly in CLC workbench a quality report was produced. Once sequences were refined to meet quality standards, sequence reads were assembled to the *E. coli* K-12 MG1655 reference genome (NP_CP027060.1). The reference sequence was assembled in contigs, and sequences were analysed for SNP variation amongst closely related strains of bacteria. The workflow utilised during whole genome assembly within CLC genomics workbench is detailed in figure (2.8).

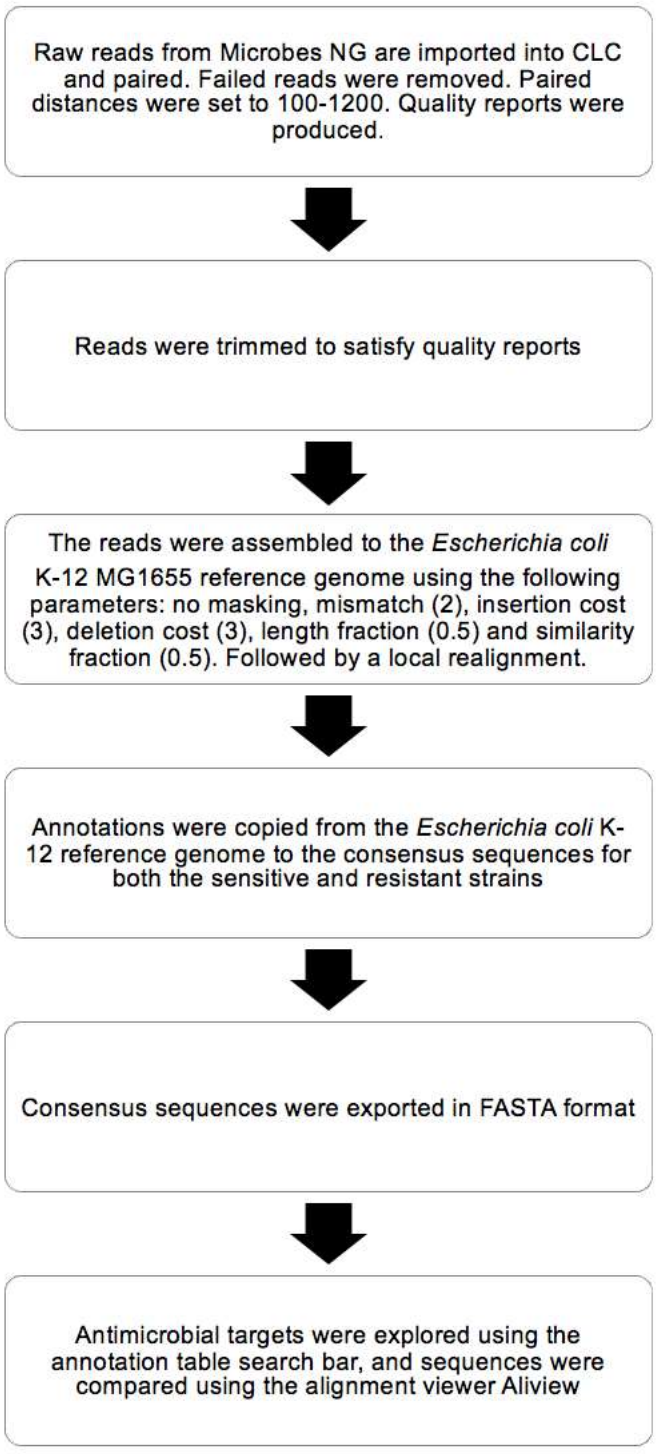


Figure 2.8 Whole genome assembly in CLC genomics workbench. Image detailing the workflow utilised to assembly the whole genome sequences obtained from microbes NG and the given parameters used during the quality control procedure.

2.8.3 Alignment of conserved genomic sequences with rearrangements using Mauve

Genomes sequencing data retrieved from microbes MG sequencing facility was of draft quality and therefore it was necessary to reorder contigs prior to alignment. Reordering was achieved using a mauve algorithm known as Mauve Contig Mover (MCM), by ordering contigs from each genome independently to the *E. coli* K-12 MG1655 reference genome and subsequently outputting multiple mauve alignments files reflecting stages within the reordering process (Rissman *et al.,* 2009). Genomes were aligned with progressive mauve under default parameters and visualised within the mauve alignment display. To identify the presence of single nucleotide polymorphisms on a genome-wide scale and to eliminate any uncertainty surrounding missed mutations using the antimicrobial target search approach within CLC workbench, the SNP export function within mauve was utilised.

2.9 Physiochemical analysis of mutagenesis

2.9.1 RNA consensus secondary structure XRNA

Initially to determine the generalised locality of all the mutations observed within two ribosomal ribonucleic acid (rRNA) variants, a consensus 16s RNA secondary structure was used to represent all seven rrn operons within *E. coli*. The XRNA java based suite version 1.2.0 (beta) was downloaded from http://rna.ucsc.edu/rnacenter/xrna/xrna.html [accessed July 17, 2018]. Following this the *E. coli* 16s.xrna secondary structure was exported http://rna.ucsc.edu/rnacenter/xrna/xrna_download.html [accessed July 18, 2018]. and viewed within xRNA. The 16s-secondary structure was then annotated to highlight the mutation positions and visualised in adobe illustrator version 22.1.0.

2.9.2 RNA secondary structure analysis using RNAfold

The folding of RNA molecules plays an important role in the determination of function (Mohsen *et al.,* 2009). Thus, it was necessary to identify whether the observed rRNA mutagenesis resulted in alterations in rna secondary structure. RNAfold (Gruber *et al.,* 2008 available at: http://rna.tbi.univie.ac.at/cgi-bin/RNAWebSuite/RNAfold.cgi [accessed September 3, 2018] was used for the computational prediction and graphical representation of rRNA secondary structures. Ribonucleic Acid (RNA) fold algorithmically predicts the optimal secondary structure based on the given nucleotide sequence using the dynamic programming technique (Lorenz *et al.,* 2016). The rRNA sequences were uploaded independently to RNAfold in Fasta format and secondary structures prediction were calculated using default parameters. Results including graphical outputs and predictions plots were downloaded from the RNAfold webserver. Direct visual comparisons were then made between the Minimum Free

Energy (MFE) and Centroid structures produced for the seven *E. coli* rRNA variants between the sensitive and resistant strains.

2.9.3 Phylogenetic analysis using MEGA

The phylogenies of the *fusA* and ribosomal RNA nucleotide sequences were resolved by comparing the sensitive and resistant sequences of the experimental strains to that of the corresponding *E. coli* sequences collected from the National Centre for Biotechnology (NCBI) available at: https://blast.ncbi.nlm.nih.gov/Blast.cgi [accessed October 19, 2017]. The fasta aligned sequences were then imported into MEGA version 7.0 (Tamura *et al*., 2016), for the construction of Maximum Likelihood (ML) Trees. Once the sequences were uploaded into MEGA a local alignment was carried out using the integrated alignment tool; Multiple Sequence Comparison Log-Expectation (MUSCLE) (Edgar, 2004). Using the Molecular Evolutionary Genetics Analysis (MEGA) alignment explorer, sequences were analysed to determine the presence of incomplete genomic sequences and to identify sequences which required reverse complementation prior to tree construction. Sequences were then realigned prior to model selection analysis, to determine the best fit DNA model for the given sequence data set. Following model selection, ML trees were constructed with a bootstrap of 1000.

2.9.4 EMBL SAPS protein structure analysis

The computational analysis tool for the Statistical Analysis of Protein Sequences (SAPS) provided by the European Bioinformatics Institute (EMBL-EBI), was used to evaluate the protein sequences to determine any alterations in the functional properties of proteins relating specifically to the amino acid type and charge

distribution observed at the site of mutagenesis. This is achieved through the clustering of negative, positive and mixed charged clusters relative to the overall charge of the protein (Brendel *et al.,* 1992). Protein sequences were uploaded into SAPS, *E. coli* K-12 was selected as the species and under defaults settings lysine and arginine was selected as positively charged prior to submission.

2.9.5 TreeSAAP selection of amino acid properties

To assess the effects of positive selection on the molecular evolutional of the functional domains of the *fusA* gene, it was necessary to determine any significant changes in the physiochemical properties of the amino acids within the protein Elongation Factor G. TreeSAAP (Woolley *et al.,* 2003) was used to analyse 31 structural and biochemical amino acid properties of Elongation Factor G from both the sensitive and resistant strains of *E. coli*. A prerequisite of TreeSAAP is the construction of a tree for phylogenetic analysis and this was achieved through the alignment of sequences followed by the production of a Neighbour joining (NJ) tree. As there was only a direct comparison being made between two sequence it was necessary to duplicate each sequence to avoid forcing tree topology. Input nucleotide sequences was compiled in NEXUS format and loaded into
TreeSAAP for analysis. The number of magnitude categories was set to eight with a sliding window of 15 with a step of one as per developer guidelines. In addition, the bacterial genetic code 11 was selected. The magnitude categories are hierarchical with the lowest category depicting non-synonymous mutations which conserve an amino acid property through to eight where the alteration causes a radical change in property (Datta *et al.*, 2010). Consequently, only highly significant results ($p<0.001$)

from categories six to eight were further analysed due to likely impact of the physiochemical change of an amino acid on protein function and thus could be associated to adaption. This value is determined by a goodness-of-fit score (GF-score) and a series of z-scores where a higher score is reflective of the intensity of selection (McClellan and McCracken, 2001). An identical analysis was performed on all genes with non-synonymous mutations identified between the sensitive and resistant strains of *E. coli* MG1655.

2.9.6 Homology modelling of protein structure using Swiss Model

Protein modelling was initially carried out using Swiss-Model available at: https://swissmodel.expasy.org [accessed August 26, 2018] as its automated ease requires no user participation apart from the submission of an initial input sequence (Guex and Peitsch, 1997). Swiss Model enabled the modelling of the experimentally determined sensitive and resistant protein structures based on templates of known evolutionary related protein structures (Waterhouse *et al.,* 2018). Fasta sequences were imported into Swiss Model and a subsequent search was carried out to identify a template, prior to modelling. Template results are ranked based on the expected model quality which is assessed by Global Model Quality Estimate (GMQE) and Quaternary Structure Quality Estimate (QSQE) scores. In addition to these two values, templates were selected based on their protein coverage and identity to the submitted sequences. The selected template was then used to build the protein models of protein sequences of interest. Once complete, the best quality models were chosen based on the Quality Model Energy Analysis (QMEAN) value which is a statistical potential score of the nativeness of structures centred around geometric

descriptors (Benkert *et al.,* 2007). Resulting, Protein Data Bank (PDB) files and model reports were downloaded and protein homology was compared.

2.10 Molecular genetics based on bioinformatics findings

2.10.1 Primer design

Primer sets specific to the intended *fusA* gene and ribosomal RNA targets were generated using Primer-Blast which employs both Primer3 and BLAST available at: https://www.ncbi.nlm.nih.gov/tools/primer-blast/ [accessed July 23, 2018] (Ye *et al*., 2012).

The PCR template was inputted as a DNA Sequence complied in FASTA format for the *fusA* gene and ribosomal RNA 123 and 154 respectively and the user adjustable parameters were modified accordingly indicated in table (2.7). To ensure the designed primers flanked the regions containing the SNPs, the sequence ranges were adapted.

Table 2.7 Primer parameters. Table detailing the primer parameters selected within Primer-Blast specific to the PCR template.

Primer Parameter	**Selection**
Database	nr
Organism	*Escherichia* (Taxid: 561)
Max T^m difference	5
Minimum primer melting temperature	50
Optimum primer melting temperature	55
Maximum primer melting temperature	60
Max target size	4000

2.10.2 DNA extraction

The bacterial DNA of *E. coli* MG1655 resistant and sensitive isolates was extracted through thermal lysis. Five individual colonies of overnight growth of gentamicin resistant and sensitive *E. coli* MG1655 were picked from MH agar plates. The individual colonies were transferred to sterile Eppendorf tubes containing 100µL of DNase/RNase free water, to obtain a suspension of bacterial cells approximately 1-2 x 10^9 cells ml^{-1}. The Eppendorf tubes were then placed in a heat block pre-set to 95°C for 10 minutes lyse the bacterial cells. Following lysis, the extracted bacterial DNA was immediately cooled on ice before being processed by Polymerase Chain Reaction (PCR).

2.10.3 Polymerase chain reaction

PCR was performed on all DNA samples extracted using the thermal lysis technique described above. Using gene specific primers (table 5.1) generated using Primer-Blast (Section 2.10.1), PCR reactions were conducted to amplify the specific gene sequences with the forward and reverse primer sets incorporating the mutated regions of the genes respectively. Primarily, a gradient PCR reaction was carried out to determine the optimal annealing temperature for all three primer sets used in the PCR reaction. Two microliters of bacterial template DNA was transferred to PCR reaction tubes containing 23 microliters of pre-aliquoted PCR reaction mix. The PCR reaction mix components and the corresponding quantities used for all PCR reactions are presented in table (2.8). The thermocycling conditions and the resultant durations of the PCR steps are shown in table (2.9). Through use of the gradient function on the thermocycler, an annealing gradient of 52-57°C was set. After identifying the optimal

annealing temperature, subsequent PCR reactions were conducted using conditions shown in table (2.10).

Table 2.8 PCR reaction mix components. Table displaying the PCR reaction mix components and the corresponding quantities and concentrations used.

Component	**Volume (µL)**	**Concentration**
Master Mix Dream Taq	12.5	2X
Forward primer	1.5	1 µM
Backwards primer	1.5	1 µM
Template DNA	2.0	~100 ng
Nuclease free water	7.5	-
Total volume (µL)	**25**	**-**

Table 2.9 Thermocycling PCR conditions for gradient PCR. Table displaying the thermocycling conditions used for all three primer sets in the initial gradient PCR reaction.

Step	**Temperature (°C)**	**Time (mins)**	**No. of Cycles**
Initial Denaturation	94	3.0	1
Denaturation	95	0.5	35
Annealing	52,53,54,55,56,57	0.5	
Extension	72	2.0	
Final Extension	72	10.0	1

Table 2.10 Thermocycling conditions for PCR reactions. Table displaying the thermocycling conditions used for all three primer sets, following the identification of the optimal annealing temperature using gradient PCR.

Step	**Temperature (°C)**	**Time (mins)**	**No. of Cycles**
Initial Denaturation	94	3.0	1
Denaturation	95	0.5	35
Annealing	54	0.5	
Extension	72	2.0	
Final Extension	72	10.0	1

2.10.4 Agarose gel electrophoresis

Agarose gel electrophoresis was used to analyse the resultant PCR amplicons prior to sequencing. All bacterial DNA and RNA samples were analysed on a 2% w/v Top Vision Agarose gel (Thermoscientific) with the addition of the nucleic acid stain Gel Red (Cambridge Bioscience) in 1 x Tris-Borate-EDTA (TBE) buffer (table 2.11). DNA gel loading dye (ThermoFisher) was added to each DNA sample for the visual tracking of samples within the gel during electrophoresis at a 1:5 ratio of loading dye to DNA. The DNA samples were then added to well of the agarose gel together with 2kb hyper ladder 2 50bp-2kb (FisherScientific) as a marker of molecular weight. Agarose gels were run at 80v (BIO-RAD Power-Pac 300) for 45 minutes in a gel electrophoresis flow chamber (Flowgen) filled with 1x TBE as the running buffer. Gels were then visualised and imaged using the G: Box Chemi XRQ (Syngene) gel documentation imaging system.

Table 2.11 Buffers used for agarose gel electrophoresis. Table showing the materials used for agarose gel electrophoresis. 1 x TBE was used as a runner buffer to make a 1% w/v agarose gel. DNA gel loading dye was added to the DNA samples at a ratio of 1:5.

Buffer	Composition	Concentration
1x TBE	Tris base Boric acid EDTA HCl	89mM 89mM 2mM pH 8.3
DNA Gel Loading Dye	Tris Bromophenol blue Xylene cyanol FF Glycerol Ethylenediaminetetraacetic Acid (EDTA) (pH 7.6 adjusted with Sodium Hydroxide (NaOH))	10mM 0.03% 0.03% 60% 60mM

2.11 Molecular docking analysis

2.11.1 Molecular structure analysis

The molecular drawing tool MoleculeSketch version 2.0 (download from the Apple App Store) was used to sketch the structures of the molecules which played an essential role in the antimicrobial docking of gentamicin to the 30s-ribosomal subunit. Molecular structures were drawn for the gentamicin complex, as well as the predicted binding sites which include single ribonucleotide bases of the 16s rRNA and the amino acid lysine residue of the 30s-ribosomal protein S12.

2.11.2 Preparation of PDB files for AutoDock

Computational protein-ligand docking and virtual drug screening was achieved using the AutoDock suite. Before docking the freely available open-source software AutoDock 4.2 and MGLTools 1.5.6 must be downloaded and installed (Morris *et al*., 2009). In addition, coordinate files for both the receptor and ligands must be obtained from the PDB or any other database or server providing experimental coordinates in a variety of formats including pdb, mol2, cif or sdf. As the protein complex required for this protein-ligand docking experiment was not available on the routine databases, the PDB files for the Pre-and Post 70s-ribosomal complex in combination with Elongation Factor G were supplied by Professor Jinzhong Lin at the State Key Laboratory of Genetic Engineering at Fudan University in Shanghai, China.
The ribosomal complexes were formed as previously described by Gagnon *et al*. (2012) with some additional modifications (Lin *et al*., 2015). X-ray diffraction data were collected from the ribosome crystals prepared by Lin *et al*. (2015) at beamlines X25 at Brookhaven National Laboratory and 24-ID-C at the Advanced Photon Source

at Argonne National Laboratory using 0.2° or 0.3° oscillations. Data were integrated and scaled using the XDS programme package (Kabsch, 1993). All crystals belong to orthorhombic $P2_12_12_1$ space group and all ligands were removed as a starting model. A summary of the crystal parameters and statics of data collection and refinement are detailed in the supplementary material of the Lin *et al*. (2015) paper. The two ribosomal structures provided were resolved at 50 – 2.8 Å and the R_{work}/R_{free} values were 20.2/25.2% and 22.1/26.6% for the pre- and post- ribosomal complexes respectively. The final step before starting Autodock tools was to collate all Autodock related files into a folder, including the pdb and Autodock and Autogrid .exe files.

2.11.3 Receptor preparation for docking

Once Autodock tools was opened, using preferences the Autodock tools folder was selected as the start-up directory and this was made default. The protein for docking was inputted using read molecule and selecting the receptor.pdb file. It is here error messages can be encountered due to input file formatting issues particularly concerning the spacing within a pdb file containing metal ions (magnesium and zinc). The next steps involved editing the receptor to ensure compatibility for docking, the first step being involved deleting water residues, followed by inspection of the hydrogens by adding hydrogens to polar atoms. The next step included the assignment of charges, to do this Gasteiger Charges were computed and the correctness of the assignment was checked and the deficit was spread across residues if necessary. The final step in the preparation of the receptor for docking was to merge the non-polar hydrogens.

2.11.4 Ligand preparation for docking

To prepare the ligand for docking firstly the receptor was hidden from view and the corresponding ligand.pdb was opened. Once the ligand has been opened in AutoDock several steps were followed to prepare a PDBQT file for docking. The first being to detect the root or the centre of the ligand which was determined by selecting torsion tree followed by detect root. In addition to the root detection, the next step was to choose torsions which displayed the number of fixed and rotatable bonds, and it was essential at this stage that all amide bonds were made rotatable. The final stage in the ligand preparation involved saving the ligand output file in PDBQT format.

2.11.5 Grid box preparation and docking with AutoDock

The final step before the docking simulation experiment within AutoDock was to generate the receptor coordinate file, by selecting the receptor as the macromolecule using the grid settings. The ligand PDBQT file previously prepared was selected as the ligand using the set map type option. Once the formatting of the receptor and ligand was completed the grid box was defined. The grid box centre was defined by providing explicit coordinates selected based on known binding regions for the designated the ligand, the grid output was then saved with a gpf extension. The grid map for the 70s pre-and post- ribosomal complex were of dimensions 60 x 60 x 60 points with a spacing of 0.465 Å, yielding a cylindrical box with a size of 1.5. The grid parameters were altered for docking stimulations for the ribosomal protein S12, with grid map dimensions of 40 x 40 x 40 points with a spacing of 0.375 Å.
Command line was then used to run Autogrid, by ensuring both Autogrid executable and autogrid4 were located within the same directory. Prior to running Autodock,

using the docking function, the macromolecule receptor pdbqt file and the ligand were selected and outputted with Lamarkian GA type with a dpf extension. Subsequently, autodock can was launched by selecting the newly generated dpf file, and upon completion a log file with a dlg extension was produced.

2.11.6 Single docking analysis and visualisation with PyMOL

Once autodock has completed its run, all posed docking was visualised by selecting the log generated files, the docked conformations were clustered using a tolerance of 2. All poses were saved in pdb. format and figures were generated using PyMOL.

2.12 The phylogenetic analysis of the *fusA* gene

To resolve the phylogeny of the *fusA* gene both within the *Escherichia* taxid and *Enterobacteriaceae* family, Maximum Likelihood trees were constructed. The sequences of the *fusA* gene were collected using a Nucleotide BLAST search of the sensitive *E. coli* MG1655 query sequence optimised for somewhat similar sequences with a maximum number of target sequences set to 500. The DNA sequence alignments for 100% query cover were downloaded from NCBI and aligned using MUSCLE (Edgar, 2004). The sequences were compiled with the *fusA* gene sequences of the experimental strains exploited in this study followed by a further realignment using MUSCLE.

2.12.1 The evolutionary model selection using JModel

The software JModelTest2 version 2.1.10 was used for the statistical selection of model based on the aligned *fusA* nucleotide sequences. The DNA alignment file was

loaded into JModel and the likelihood scores were computed using the following settings: Bayesian Information Criterion (BIC) clustering heuristics, base frequencies (+F) and rate variation (+I and +G). The number of processes requested within the search was 4 with 203 substitution schemes, giving a total of 1624 models for the comparison of best fit. The ML optimised option was selected as the base tree for likelihood calculations and the base tree search was Nearest-neighbour Interchange (NNI). Following the analysis of likelihood Akaike Information Criterion (AIC) and BIC calculations were computed under default settings (Guindon and Gascuel, 2003; Darriba *et al*., 2012).

2.12.2 Phylogenetic tree construction using PhyML

PHYML version 3.1 was downloaded from: www.atgc-montpellier.fr/download/binaries/phyml/PhyML-3.1.zip and was compiled using the command-line interface. The DNA alignment file was converted to the PhyML supported Phylip format (interleaved or non-interleaved) using EMBOSS Seqret available at: https://www.ebi.ac.uk/Tools/sfc/emboss_seqret/ [accessed July 2, 2018]. Following reformatting, the parameters for the phylogenetic analysis were determined. The selection of substitution model was predetermined using JModel_2.0 (Section 2.12.1). For all ML trees generated in PhyML, invariable sites models were utilised with fixed topology. The final selection parameter was the number of bootstraps and to increase the statistical reliability of the ML trees, the maximum number of bootstraps was selected (1000). The final step involved the configuration of a command line string to execute the phylogenetic ML tree construction within PhyML. The string included the number of bootstraps (–b), the model (-m), the input

file (-i) and finally the data type (-d) (Guindon *et al*., 2010). The output text file from the phylogenetic analysis using PHYML was then viewed and annotated using the online tree viewer iTOL (Interactive Tree of Life) version 4.2 available at: https://itol.embl.de [accessed October 4, 2018].

2.13 Analysis of bacterial fitness

2.13.1 Manual growth curve

To compare the growth of the sensitive and resistant *E. coli* MG1655 strains, manual growth curves were conducted. Initially, O/N cultures of sensitive and resistant strains were used to inoculate 50mL falcon tubes containing 20mL of MH broth, in triplicate. Broth cultures were incubated O/N aerobically at 37ºC 100 rpm. Following incubation, cultures were centrifuged at 4,500 x g for 5 minutes, and bacterial cell populations were set to a starting OD of 0.2 at 600nm. Next, 1mL of standardised cultures were transferred to pre-warmed conical flasks containing 100mL of MH broth, to achieve a starting OD_{600} of 0.002 OD units, corresponding to approximately 10^4-10^6 CFU/mL. Conical flasks including MH broth controls were incubated aerobically at 37ºC 100 rpm for the experimental course. The OD_{600} was monitored every 30 minutes for 12 hours using 1cm path length cuvettes in an Eppendorf Biophotometer. In addition, at each time point quantification was achieved through serial dilutions and later CFU/mL calculations, described previously in section (2.6).

2.13.2 *Galleria mellonella* virulence model

The *Galleria mellonella* (*G. mellonella*) toxicity model was used to assess bacterial virulence. This was achieved through cultivating the resistant and sensitive cultures

from frozen as previously described section (2.3.2) before use within the *Galleria* toxicity model. The resistant bacteria and sensitive control were transferred into universal tubes containing 20mL of MH broth and the cultures were incubated aerobically at 37°C shaking at 100 rpm (Heidolph Unimax 110).

Following overnight incubation, 1mL of each culture was used to inoculate pre-warmed Erlenmeyer flasks containing 100mL MH broth. The freshly inoculated broths were then transferred into the shaking incubator at 37°C 100 rpm for 240 minutes (exponential phase of growth). Subsequently, a 1mL aliquot from each Erlenmeyer flasks (duplicated) was transferred to a sterile Eppendorf and using a centrifuge (Eppendorf Minispin), the aliquot was centrifuged for 2 minutes at 13,000 x g. The supernatant was then discarded and the pellet was resuspended in 1mL sterile PBS. The cells were washed a further two times, before a final resuspension in 1mL PBS.

Assuming the duplicates are identical, 1mL of each test sample was transferred to a plastic cuvette and using a spectrophotometer (Thermo Spectronic Helios) the OD at 600nm was recorded. To achieve the desired starting OD_{600} of 0.8 OD units, either PBS was added to dilute the sample or alternatively to achieve a higher OD the samples were re-centrifuged and suspended in smaller volume of PBS.

A total of 60 *Galleria* larvae (Chessington Garden Centre, Surrey) were utilised within the toxicity model experiments. The larvae were divided into 6 control and experimental groups, with 10 larvae per group. The groups comprised: the non-injected control, the PBS control, the sensitive control, MIC resistant cultures, x10

MIC resistant cultures and x100 MIC resistant cultures. A syringe (Hamilton 700 series, bevelled tip) was used to inject each group, excluding the non-injected control. The syringe was sterilised between each inoculation using Industrial Methylated Spirit (IMS) followed by a final rinse with sterile PBS.

The *G. mellonella* were injected with 10µL of bacterial suspension (10^6-10^8 CFU/mL), through the last left pro-leg. The larvae were placed between the thumb and forefinger and held in place while injecting, the larvae were cautiously removed from the needle and placed into the respective petri dishes. Following inoculation, plates were placed in the static aerobic incubator at 37°C to mimic the mammalian environment for a period of 5 days (Figure 2.9).

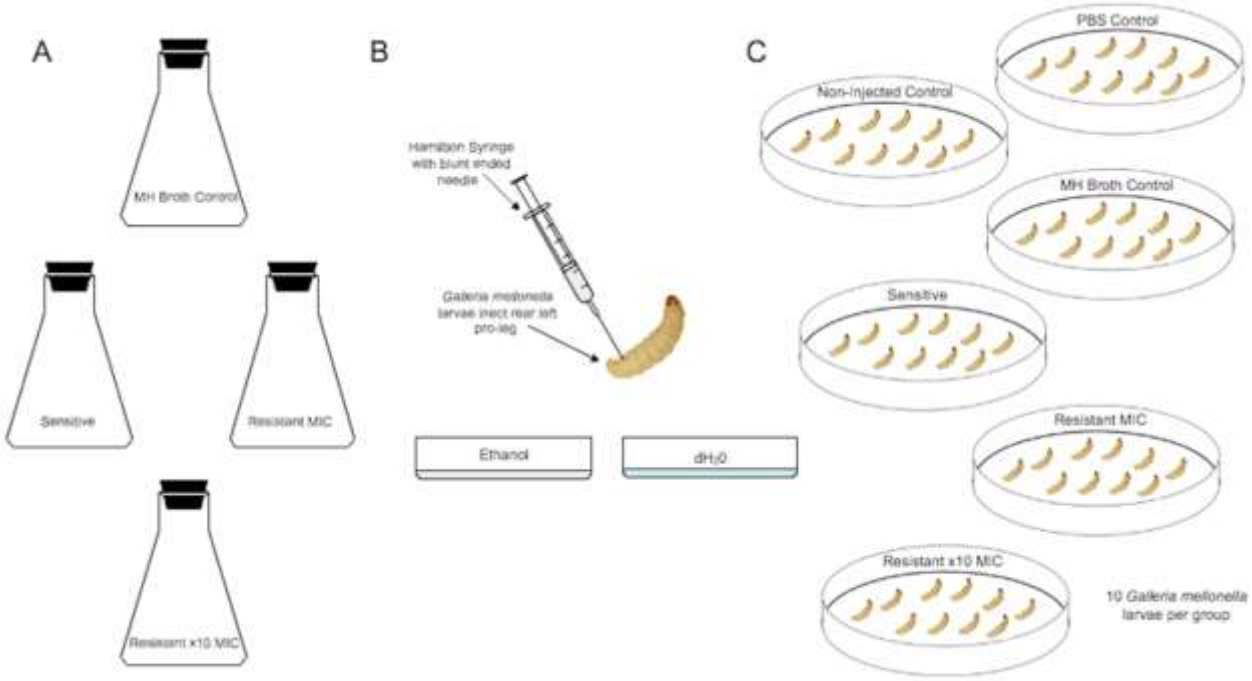


Figure 2.9 The *Galleria mellonella* virulence model. The image outlines the three steps in the preparation of the *galleria* larvae for use within the toxicity model: (A) Bacterial sample preparation. (B) Injection procedure. The Hamilton syringe utilised within the study was flushed with ethanol followed by dH_2O, prior to the inoculation of the injected larvae groups with 10µL of standardised bacterial samples or MH broth. The larvae were then placed within the respective petri dishes and incubated at 37°C.

The *Galleria* larvae were analysed daily for mortality. The larvae were assessed based on observations of mobility, melanisation and survival. The larvae were considered dead when inert and completely melanised, appearing black in colour.

2.13.3 Scanning electron microscopy

Scanning Electron Microscopy (SEM) was used to determine any morphological differences in *E. coli* MG1655 due to the acquisition of gentamicin resistance. Gentamicin sensitive and MIC resistant strains were cultivated from frozen as described in section (2.3.2).

Prior to biofixation, 12mm round coverslips were immersed in 100% ethanol overnight to ensure sterility, and left to air dry. Next, to facilitate the adherence of bacterial cells, a 10μL drop of dH_2O was placed on the coverslip. Single bacterial colonies were isolated and smeared onto the cover slip, forming a thin layer of bacterial cells, which were fixed by the succinctly passing through a Bunsen Burner flame.

The coverslips containing the fixed resistant and sensitive bacterial cells, were transferred to designated wells within a six-well polystyrene plate and covered with 0.1 M phosphate buffer. Unless otherwise stated, subsequent sample preparation was performed in a class two biosafety cabinet. Once in the fume hood, the phosphate buffer was removed and immediately replaced with 2.5% glutaraldehyde for two hours. After the primary fixation, bacterial samples were washed with phosphate buffer for 15 minutes and this is repeated four times.

Following washing, bacterial cells were secondary fixed with 1% osmium tetroxide for one hour. Ensuing secondary fixation, bacterial cells were dehydrated through a graded series of 50, 70, 80, 90 and 90% ethanol in increasing concentration, for a period of twenty minutes within each concentration. The final dehydration step

involved the emersion of bacterial cells in 100% ethanol for three successive periods of twenty minutes. Bacterial cells were dried by immersion in hexamethyledisilizane for five minutes, with plates covered, then allowed to air dry for 24 hours within the biosafety cabinet. During the sample preparation procedure, it was imperative to ensure that the coverslips were always in solution, and this was achieved by quickly decanting and replacing reagents.

The coverslips containing the fixed bacterial cells, were then placed on carbon discs of appropriate width, and mounted onto an aluminium stub. Flash dry silver paint, was placed on each of the four corners, to create contact between the filter and the stub. The samples were then transferred to Quorum (Quorum Technologies, East Sussex UK) and nano coated via sputtering with a conductive medium of Gold/Palladium alloy. The sample was pumped down, purged with argon and sputtered with gold for 120 seconds on a rotating stage. Once removed from the sputter coater it was essential that the samples were not touched, and non-conductive tape was added to secure the sample in place prior to analysis.

All samples were imaged in a JCM-5700 Scanning Electron Microscope contained inside a mobile biological containment enclosure. The gold coated specimens were imaged under a high vacuum at 6kV, with an 8mm working distance and a 30μm objective lens aperture. Images were collected using a secondary electron detector, with an acquisition time per image of 160 seconds and each image was 2560 x 1920 pixels. SEM images were recorded at magnifications ranging from 3,000x to 19,000x.

Chapter 3: Experimental evolution as a tool for exploring the evolution of antimicrobial resistance across an antibiotic landscape under laboratory conditions.

3.1 Introduction

Experimental evolution studies have facilitated the analysis of key evolutionary processes, under experimentally determined conditions, allowing the direct visualisation of the evolutionary process (Kawecki *et al.,* 2012). The employment of microorganisms to experimental evolution studies is permitted by the rapid generation times and mutation rates observed within bacterial populations (Elena and Lenski, 2003). However, the true potential of experimental evolution studies was uncovered when utilised together with WGS techniques (Barrick and Lenski, 2013). The sequencing of bacterial genomes allows the genotypic pathways underlying the evolutionary changes to be determined (Conrad, Lewis and Palsson, 2011).

Experimental evolution studies have been applied within the field of antimicrobial resistance, because in addition to characterising known resistance mechanisms such as the horizontal acquisition of resistance genes, it is essential to understand how resistance evolves through mutations (Lukacisinova and Bollenbach, 2017). This will not improve the understanding of the evolution of antimicrobial resistance within bacteria, but will provide insights into whether the genomic basis of adaptation can be quantitatively predicted (Furusawa, 2017).

An initial approach used to explore the evolution of antimicrobial resistance was through serial passage, where experiments were conducted for extensive periods

over a given number of generations (Lenski *et al.*,1991). Advancements within the field has seen the introduction of microbial selection devices which allowed the dynamic regulation of the experimental conditions influential to resistance evolution. Such studies have included distinct microenvironments, reflective of natural bacterial niches (Zhang *et al.*, 2010) and sustained antimicrobial selection, whereby evolving bacterial populations are continuously challenged through the external regulation of antibiotic concentration in response to bacterial growth and nutrient utilisation (Toprak *et al.,* 2011).

A limitation of serial passage systems and fluidic selection devices, is the capacity to explore the ability of bacterial populations to migrate between spatially distinct regions, essential within the natural environment (Matinez, 2009). Therefore, experimental evolution approaches were developed which incorporated spatial temporal antibiotic concentration gradients. Grounded previous work conducted by the Kishony Lab at Harvard University, saw the introduction of the Microbial Evolution Growth Arena, denoted the MEGA plate reflecting its large size (120×60cm) facilitated the study of microbial evolution over an antibiotic landscape (Baym *et al.,* 2016b). Owing to the success of the MEGA plate there was a need for similar more simplified approaches to facilitate the high throughput study of microbial evolution.

This chapter is non-hypothesis driven as it is based upon method development and optimisation. The aim of this chapter is to develop a simple reproducible tool to explore the evolution of antimicrobial resistance over a spatial concentration gradient

of antimicrobial agents. A comparative assessment will also be conducted with the more traditional evolutionary approach of serial passage.

3.2 Materials and methods

3.2.1 Preliminary diphasic plate experiment

The initial stage in the development of a model to explore the evolution of antibiotic resistance, was to devise a diphasic agar plate. Diphasic plates were produced using bioplates which allowed the generation of two phases of MH agar with and without antibiotic supplementation. Details of the production of bioplates and preliminary diphasic experiments are detailed in chapter 2 (Section 2.7.1).

3.2.1.1 Preliminary diphasic plate experiment: semi-solid MH agar

The protocol described in (2.7.1.2) was repeated with the following adaptation. The consistency of the MH agar was altered to produce semi-solid consistencies of 75% and 50% (See table 3.1).

Table 3.1 Constituents of semi-solid MH agar. Table showing the formulation used to produce semi-solid MH agar of 50% and 75% consistencies for use within the diphasic plate experiments.

Consistency	Phase	MH Broth (g)	Volume of dH_20 (mL)	Bacteriological Agar (g)
50%	1	2.6250	125.00	1.3125
	2	2.5988	123.75	1.2994
75%	1	2.6250	125.00	1.9688
	2	2.5988	123.75	1.9491

3.2.1.2 Preliminary diphasic plate experiment: M9 minimal media

The protocol described in (2.7.1.2) was repeated with the following modifications. The nutrient availability within the agar was reduced, by use of minimal media. Initially the first step was to make the M9 salts, which were to be added to the media post sterilisation (Table 3.2). Subsequently, the agar was prepared by sterilising 700mL dH_2O with 1.5% w/v bacteriological agar. The components added post sterilisation are detailed in table (3.3). Plates were inoculated with one 10µL drop of bacterial suspension (1.5 x 10^8 CFU/mL), which was streaked using a sterile swab (Figure 3.1).

Table 3.2 M9 salt composition. Table displaying the components required to produce M9 salts to be added post sterilisation for the preparation of M9 minimal media.

M9 salts	
Components	**Weight (g)**
Na_2HPO_4-$7H_2O$	64
KH_2PO_4	15
NaCl	2.5
NH_4Cl	5.0

Table 3.3 M9 minimal media composition. Table displaying the components required to be added post sterilisation for the preparation of M9 minimal media, with the inclusion of the M9 salts detailed previously in table 3.2.

Minimal media	
Components	**Volume (mL)**
M9 salts	200
1M $MgSO_4$	2
20% glucose	20
1M $CaCl_2$	0.1

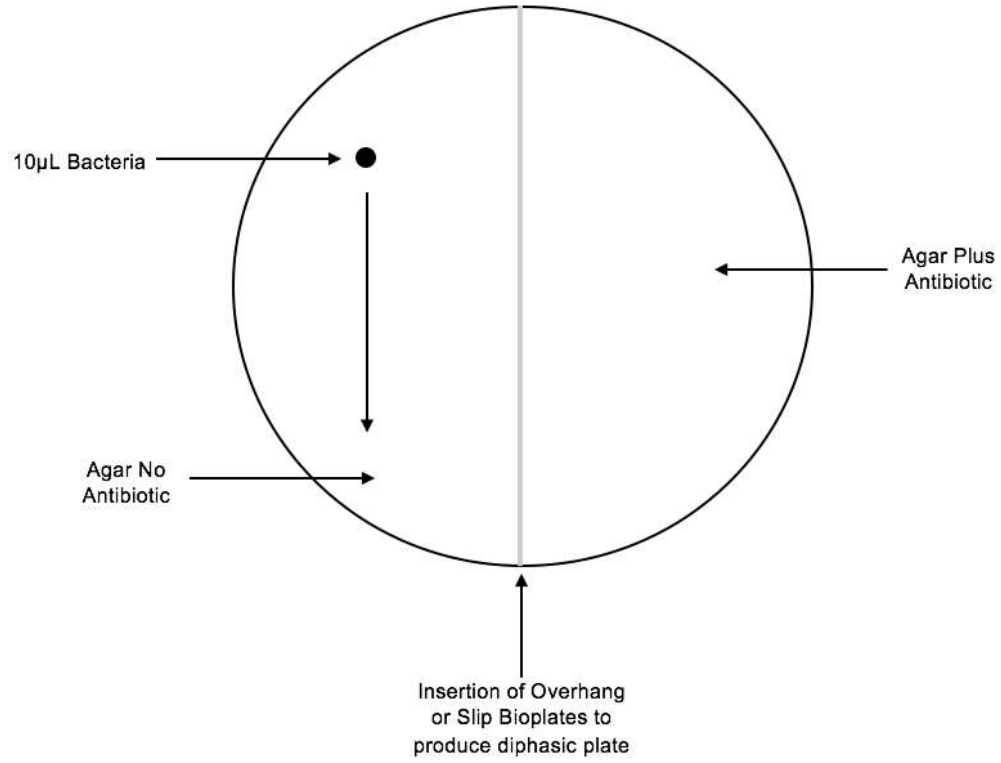


Figure 3.1 The modified diphasic plate layout. Diagram displaying the plate set-up used in the production of the modified diphasic plate. Plates were divided into two phases using the bioplates to produce a gradient of antimicrobial agent. Plates were then inoculated with a single 10µL (1.5×10^8 CFU/mL) drop of sensitive bacterial suspension on the phase containing no antimicrobial agents, which was streaked horizontally using a sterile swab.

3.2.2 Preliminary triphasic plate experiment

The experimental procedure described in (2.7.1.2) was repeated on a larger scale in a sterile plate with Teflon coating and dimensions of 24cmx24cmx5cm. Indentations were made in the plate to facilitate the insertion of the larger overhang bioplates. The larger dimensions of the plate enabled the set up shown in figure (3.2). Here the plate was sterilised and the two outer zones were poured first, followed by the removal of bioplates and finally the filling of the inner zone. The triphasic plates were inoculated with three 100µL (1.5 x 10^8 CFU/mL) drops of drug-sensitive bacterial suspension.

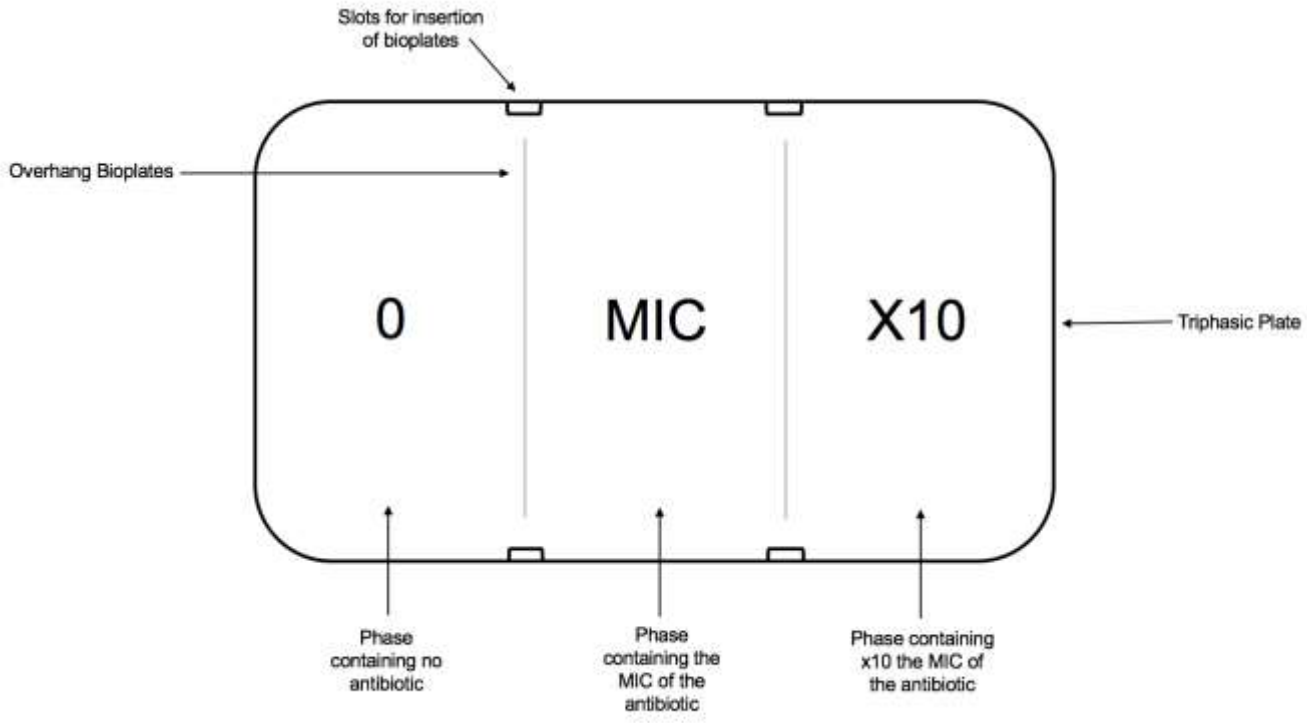


Figure 3.2 The triphasic plate layout. Diagram displaying the triphasic plate set-up. The plate was divided into three sections by inducing slots for the insertions of the overhang bioplates. This facilitated the production of a three-phase antimicrobial concentration gradient ranging from zero to x10 the MIC of a given antimicrobial agent. The triphasic plate diagram does not include bacterial inoculation.

3.2.2.1 Preliminary triphasic plate experiment: semi-solid MH agar

As described in (3.2.1) the experimental procedure was repeated with the inclusion of semi-solid media. Consistencies of 60% and 75% MH agar were tested (Table 3.4).

Table 3.4 Semi-solid MH agar composition for the triphasic plate experiments. Table showing the formulation used to produce semi-solid MH agar of 60% and 75% consistencies for the three phases within the triphasic plate experiments.

Consistency	Phase	MH Broth (g)	Bacteriological Agar (g)
60%	1	21.00	12.60
	2	20.79	12.47
	3	18.90	11.34
75%	1	21.00	15.75
	2	20.79	15.59
	3	18.90	14.18

3.2.2.2 Preliminary diphasic plate experiment: agar slope

To explore the trickling effects of antimicrobial agents, as opposed to the effect of a set concentration gradient in previous methods a slope gradient was produced. Here the tray was sterilised and placed on a block, initially the second phase was poured containing the antimicrobial agent and allowed to set. Following on from this the block was removed and MH agar containing no antimicrobial agent was poured and allowed to set (Figure 3.3). The same method of inoculation was used as described in section (3.2.2).

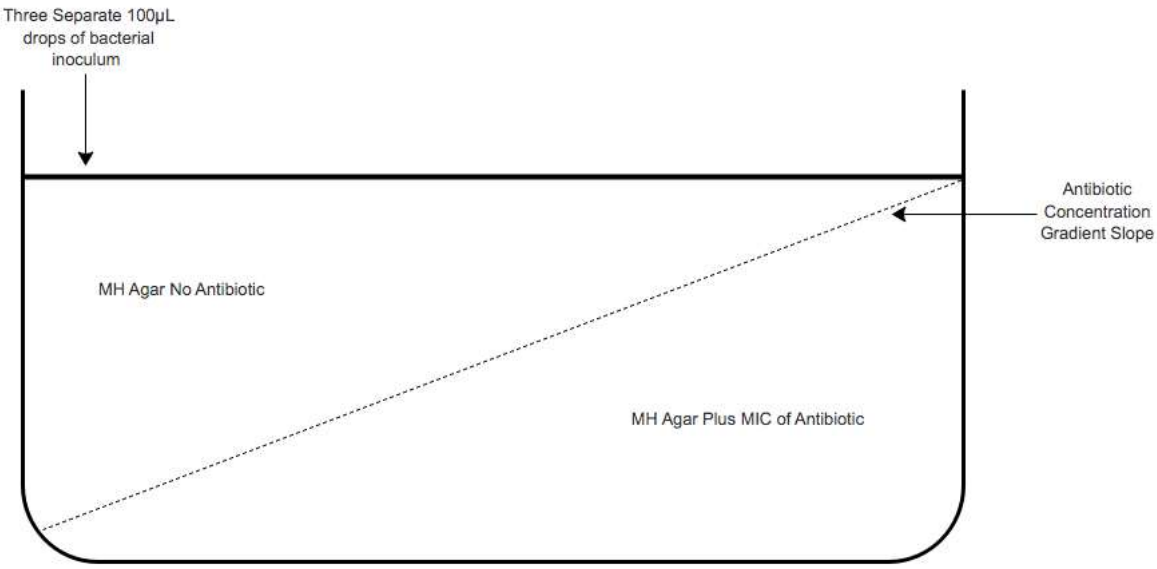


Figure 3.3 The diphasic slope plate layout. Diagram displaying the diphasic slope plate. The plate was divided into two antibiotic phases comprising zero antibiotic and the MIC of a selected antimicrobial agent. The concentration gradient slope was produced by pouring the antimicrobial containing media and placing the diphasic slope plate on a block whilst the media solidified. The block was subsequently removed and the media containing no antibiotic was poured. The diphasic slope plate was then inoculated with three separate 100µL drops of drug-sensitive bacterial suspension.

3.2.3 Preliminary filter tip experiment

The filter tip experiment was devised to see the flow of bacteria through a controlled system as an evolutionary model of antimicrobial resistance. To be able to observe this flow of organisms through the filter tips; two filter tips were used which had been impregnated in two different agars; nutrient agar and the semi-selective MacConkey agar (Figure 2.3). Therefore, the migration of the bacteria was identifiable by a media colour change. Details of the preliminary filter tip experiment are described in chapter 2 (Section 2.7.2).

3.2.3.1 The filter tip experiment

Once the flow of microorganisms through the filter tip system, had been confirmed. The experimental protocol described in (Section 2.7.2), was repeated with the inclusion of MH agar and filter tips were supplemented with antimicrobial agents. Experimental details are provided in (Section 2.7.2.1), with supplementary steps for the isolation of the potentially resistant organisms.

3.2.3.2 Quantification of the filter tip experiment

To quantify the filter tip experiment the experimental procedure described in section (3.2.3.1) was repeated, but here, the antimicrobial impregnated filter tip was used to perform colony counts. Samples were taken from the antibiotic containing tips via wooden swabs, which were immediately transferred to PBS and vortexed for 10 seconds to isolate the potentially resistant organisms. Two hundred microliters of the suspension was transferred to a 96-well plate and a total of six serial dilutions were performed per isolate. Subsequently, 10µL of each dilution was plated and incubated

aerobically O/N at 37ºC. Post incubation colonies were counted and CFU/mL were calculated (Figure 3.4).

3.2.3.3 The altered filter tip experiment

The method described in (3.2.3.2) was repeated with minor modifications. The P1000 micropipette tips were cut at the base to enlarge the exit hole prior to sterilisation. The media used to plug the tip, was change to antibiotic containing agar with a concentration of the MIC of the given antimicrobial agent; Chloramphenicol (MIC ≤8mg/L) and Trimethoprim (MIC ≤2mg/L). After incubation, instead of removing the tips using sterile forceps, here forceps are used to push the media through the micropipette tube. Thus, the sterile wooden swab is used here to sample the antibiotic containing agar as opposed to the tips, to obtain the potentially drug resistant microorganisms. Once in the PBS, the suspensions can be used to inoculate antibiotic agar plates or perform serial dilutions to quantify viable counts of potentially drug resistant organisms.

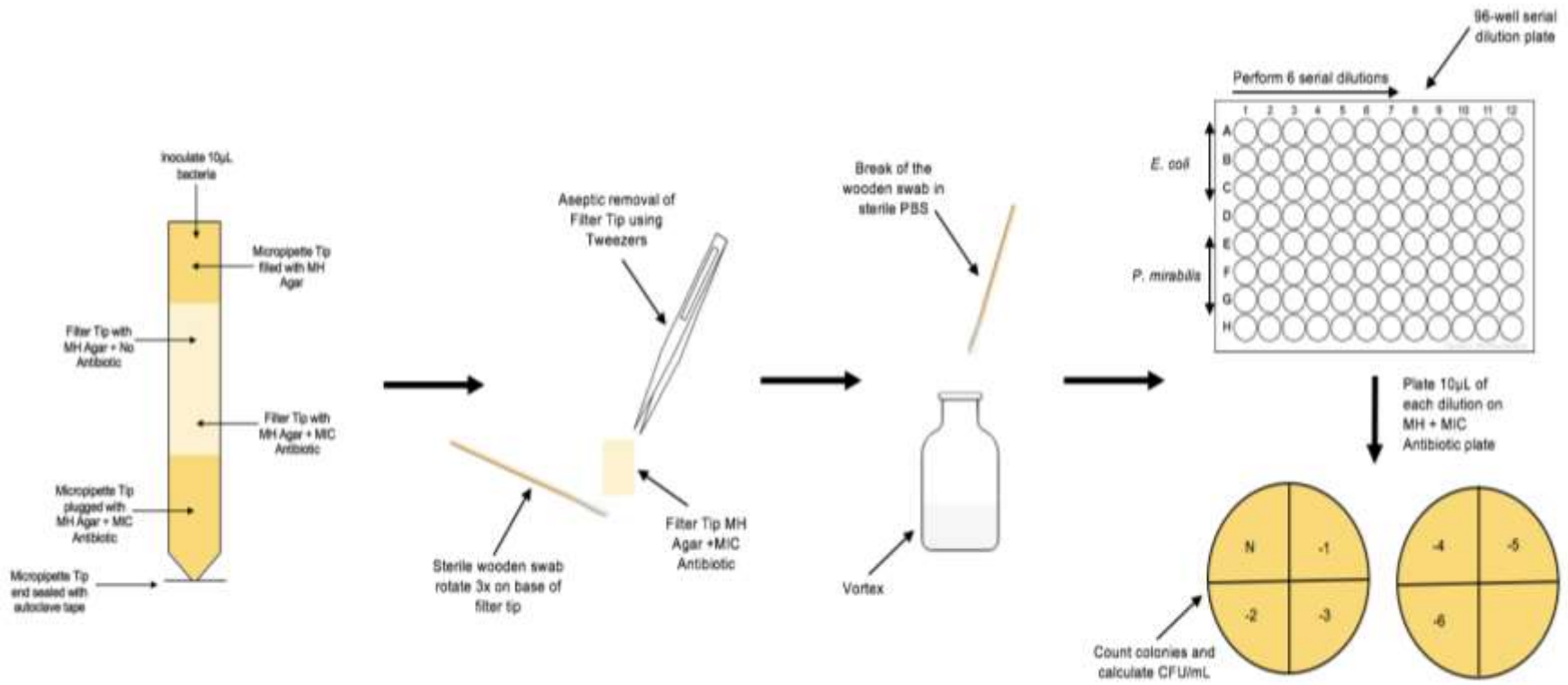


Figure 3.4 Quantification of the filter tip experiment. Diagram showing the method used for the quantification of the filter tip experiment. The P1000 micropipette tips were prepared using filter tips which were impregnated with both MH agar and MH plus the MIC of the selected antimicrobial agent. The micropipette tips were then inoculated with 10µL of a drug sensitive bacterial suspension. Following incubation, the filter tip containing antimicrobial agent was asceptically removed from the micropipette tip using tweezers. The filter tip was then swabbed using a sterile cotton swab, which was broken into a sterile bijou containing PBS and vortexed for 10 seconds. The bacterial suspension was then transferred to a 96-well plate and serial dilutions were performed before plating 10µL of each dilution onto MH plates for CFU/mL calculations.

3.2.4 The triphasic circle plate experiment

The Triphasic plate method was modified to facilitate the radial growth of microorganisms, which was hindered within original rectangular format. This revised method was named the Triphasic Circle Plate, and a comprehensive description of the protocol used in the production of the circular plate can be found in chapter 2 section (2.7.3).

3.2.4.1 The altered triphasic circle plate method

The Triphasic Circle Plate experimental procedure was repeated with a slight alteration. The method was prepared as described in (2.7.3), but modifications followed the initial pouring of a thin layer of Mueller Hinton agar to the base of the plate. After this the outer phase ring was positioned and the x10 MIC segment was poured and allowed to set. Subsequently the inner phase ring was placed in position and the outer ring was removed, followed by the pouring of the second phase containing MIC of the antimicrobial agent. Finally, the inner ring was removed and the primary phase was poured containing zero antimicrobial agent. Plates were placed inside a plastic box containing an absorbent pad covered in autoclaved water; the lid of the box contained two 0.2µM filters to facilitate the sterile passage of air. The boxes were then placed in the incubator at 37ºC, and observed daily for bacterial growth.

3.2.5 Motility experiments

An impeding factor concerning all the spatial temporal models described previously within this chapter, is the motility requirement. To explore the motility of the microorganisms further, both phenotypic and genotypic experiments were conducted.

3.2.5.1 Motility testing: stab test

The initial phenotypic motility test conducted was the motility stab test. Motility Media components (Table 3.5), was made up to 1 liter with dH_2O and heated by boiling to melt the agar. Post boiling 5mL of 1% v/v Triphenyl tetrazolium chloride (TTC) solution was added, to improve the visibility of diffusion which is facilitated by the bacterial reduction of TTC to form a red colour (Kelly and MacDonald, 1952). Following sterilisation, 5mL aliquots of media were decanted into 15mL falcon tubes, which were cooled upright within a tube racks. To test for motility a sterile needle was used to pick an isolated colony from an O/N MH plate, which was then stabbed into the motility agar to within 1 cm of the bottom of the falcon tube. After inoculation, racks were incubated at 35°C for 18 hours or until growth is evident. A positive motility test is confirmed by a red turbid area extending outwards into the media from the line of inoculation.

Table 3.5 Motility media components. Table displaying the components required to produce motility media for the stab test in the assessment of bacterial motility.

Motility Media	
Components	**Weight (g)**
Beef Extract	3.0
Pancreatic digest of casein	10
Sodium Chloride	5.0
Agar	4.0

3.2.5.2 Genomic analysis of bacterial motility

As well as the phenotypic tests described above, motility can be assessed genotypically. Motility within *E. coli* is largely dependent on the production of flagella, which is initiated by the activation of transcription factors flhD and flhC, referred to as the flhDC complex (Wang *et al.,* 2015). As the expression levels of flhDC plus the number of insertion (IS) elements, have been directly linked to motility in *E. coli* K-12, PCR reactions for flhD and yecG was conducted on all *E. coli* strains used within this study (Fahrner and Berg, 2015).

3.2.5.2.1 DNA extraction for the analysis of bacterial motility

DNA extraction was completed using the Gentra Puregene Yeast/Bacteria Kit. As per kit protocol. To prepare for DNA extraction isolates were inoculated into MH Broth for 24hrs at 37°C 100RPM in a shaking incubator, after revival from cryovials. Optical densities were taken at 600nm on spectrophotometer and bacterial numbers were confirmed via serial dilution. The overnight culture must contain approximately 1-2 x

10^8 cells. After extraction, DNA was incubated at room temperature (15-25°C) overnight with gentle shaking or stored at -20°C after concentration and purity checks were completed using the NanoVue.

3.2.5.2.2 PCR for the analysis of bacterial motility

Primer sets specific to the intended flhD and yecG gene targets (Table 3.6) were generated using Primer-Blast available at: https://www.ncbi.nlm.nih.gov/tools/primer-blast/ [accessed June 18, 2018]. The primer design parameters are detailed in chapter 2 (Table 2.7). The primer sets were tested against five different isolates of *E. coli* (UTI89, CFT073, TPA4792, TPA2743 and MG1655) in the assessment of bacterial motility.

Table 3.6 Primer sequences used in the genomic analysis of bacterial motility.
Table displaying the forward primer sequences of the flhD and yecG primers used in the genomic assessment of bacterial motility.

Primer Target	Forward Sequence (5'-3')	Reverse Sequence (3'-5')
flhD	TGTTTCAGCAACTCGGAGGTATGC	GCATACCTCCGAGTTGCTGAAACA
yecG	CAGCGACAAGAATATTGCTATAGC	GCTATAGCAATATTCTTGTCGCTG

3.2.5.2.3 DreamTaq Polymerase PCR for the analysis of bacterial motility

PCR was completed using DreamTaq DNA Polymerase (Thermoscientific). The PCR reaction mix components and corresponding quantities used for the motility PCR reactions are presented in table (3.7). The thermocycling conditions and durations are detailed in table (3.8). An annealing temperature of 54ºC was used here, with a 2-10-minute amplification time depending on the number of IS elements.

Table 3.7 PCR reaction mix for the bacterial motility PCR reactions. Table displaying the PCR reaction mix components and the corresponding quantities and concentrations used in the PCR reactions for the assessment of bacterial motility.

Component	Volume (µL)	Concentration
Dream Taq	25	2X
Forward primer	2	1µM
Backward primer	2	1µM
Template DNA	1	~ 100 ng
Nuclease free water	20	-
Total volume (µL)	**50**	**-**

Table 3.8 Thermocycling conditions used for the bacterial motility PCR reactions. Table displaying the thermocycling conditions and the corresponding durations used in the PCR reactions in the assessment of bacterial motility.

Step	Temperature (°C)	Time (mins)	No. of Cycles
Initial Denaturation	95	1.5	1
Denaturation	95	0.5	30
Annealing	54	0.5	
Extension	72	0.5	
Final Extension	72	2-10	1

3.2.5.2.4 Gel electrophoresis for the analysis of bacterial motility

PCR products were determined using Gel Electrophoresis. 2% w/v agarose gel stained with SYBR safe (Invitrogen, S33102). PCR products were loaded with loading buffer together with the GeneRuler 2Kb DNA ladder (Thermoscientific).

3.2.5.3 Motility testing: 6-well plate assay

In addition to the phenotypic stab test described above, a secondary phenotypic test was used to explore swimming motility using semi-solid motility agar, which would later be applied within the Triphasic Circle Plate Method. Here both tryptone and MH broth were prepared and solidified with 0.3% w/v bacteriological agar (Barker *et al.,* 2004). The semi-solid agars were then added to wells of a 6-well Multidish and inoculated with 2 µL of an O/N broth culture standardised to an OD of 0.5. Stationary semi-solid cultures were then incubated at 37°C and swarm diameters visually observed.

3.2.6 Finalised triphasic plate circle method renamed the Antibiotic Resistance Growth Plate (ARGP)

To provide a simple reproducible approach for studying bacterial evolution, the triphasic circle plate described in chapter 2 (Section 2.7.3) was downscaled to a polystyrene petri dish with dimensions of 90mm × 15mm, and the finalised experimental evolution model was named the ARGP. Here, the same experimental approach was followed as described in (Section 3.2.4.1), with an additional step. After the spatially distinct regions had been filled with molten agar of increasing antibiotic concentrations, this was overlaid by a layer of semi-solid agar, to enhance bacterial motility (Figure 3.5). The motility layer was prepared as described in (3.2.5.3), but here MH broth was solidified with 0.3% w/v to 3% w/v bacteriological agar to determine the optimal consistency. ARGP plates were inoculated and sealed, followed by upright aerobic incubation at 37ºC. Daily observations were captured using the G-Box Imaging system.

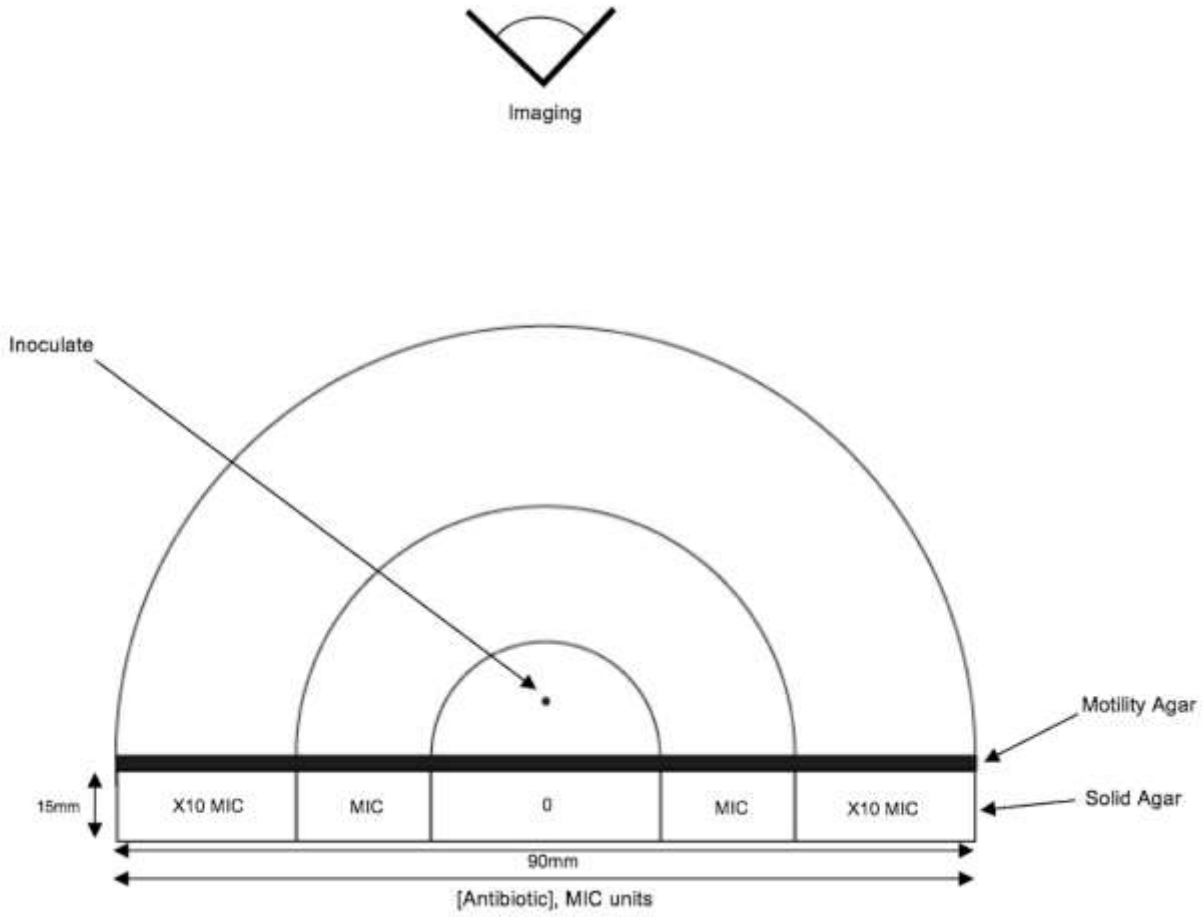


Figure 3.5 The finalised ARGP set-up. Diagram displaying the three-step antimicrobial concentration gradient utilised within the ARGP with antimicrobial concentration increasing exponentially by a factor of 10 of the MIC of the selected antimicrobial agent from the centre of the plate outwards. This was then covered with a thin motility layer and plates were inoculated in the central drug free region of the plate. Daily images are taken using the G-Box imaging system.

3.2.7 Quantification of antimicrobial diffusion and degradation within the ARGP

To investigate the extent of antimicrobial diffusion and degradation occurring within the ARGP, it was necessary to assess the effect of MH agar on antimicrobial diffusion by NMR spectroscopy. Aside from depicting how the agar composition will affect the diffusion kinetics, it was essential to decipher the concentration of antimicrobial agent present within each phase over the experimental course. The different behaviour of classes of antimicrobial agents within bacteriological agar, is largely dependent on the agar composition. The predominant antimicrobial used within this study is Gentamicin, and aminoglycosides are known to be heavily influenced by variations in chemical groups within agar matrices of which magnesium, acid sulphate and calcium are attributable to the greatest antibiotic interactions.

3.2.7.1 Quantification of antimicrobial diffusion and degradation using Nuclear Magnetic Resonance (NMR)

All antimicrobials employed within this study were high performance liquid chromatography (HPLC) grade. Gentamicin stock solutions of 2000mg/L were prepared using bidistilled water, and stored within a freezer at -20°C. MH agar was prepared per guidelines, as previously described, to a pH of 7.3 +/- 0.1. The varying gentamicin agar concentrations, were utilised to produce the antibiotic phases within the ARGP. Gentamicin agar extraction from the ARGP was achieved by harvesting agar cylinders (1-5μL) using calibrated glass microcapillaries, containing approximately 150mg per sample. Cylinders were removed at set time points over a given period of aerobic incubation at 37°C: 0, 0.5, 1, 2, 4, 6, 8, 16, 24, 48 and 72

hours. The extraction mapping trajectories within the ARGP for the NMR analysis of antimicrobial diffusion and degradation are shown in figure (3.6).

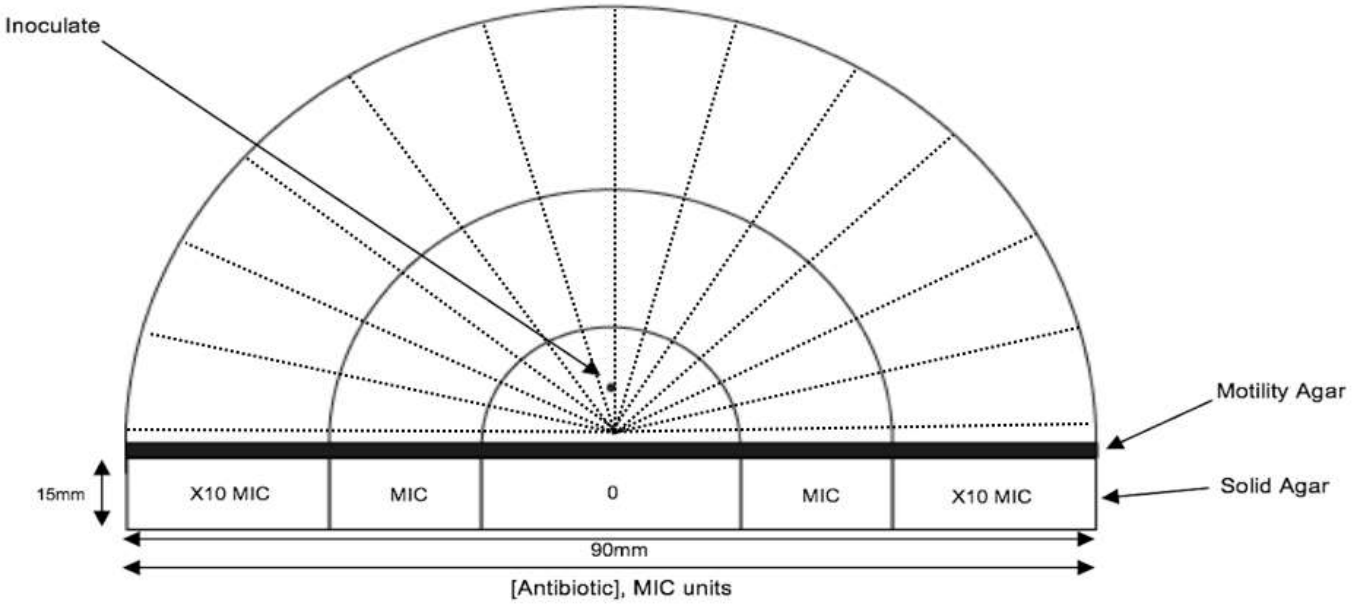


Figure 3.6 Agar extraction mapping within the ARGP. The diagram displays the ARGP set-up with the extraction mapping trajectories representing the extraction time points: 0, 0.5, 1, 2, 4, 6, 8, 16, 24, 48 and 72 hours. Agar extracts were harvested by agar cylinders using calibrated glass microcapillaries.

Derivatisation follows to improve the analytical capabilities of the agar gentamicin matrix. The agar cylinder removed at the given time point was placed in glass beaker with 20mL dH_2O and was boiled until the agar had dissolved. Samples were cooled, transferred to Eppendorf tubes and frozen at -20ºC until analysed using NMR. Prior to NMR analysis, 200µl of solution was taken up in 3mL NMR tubes. Followed by the addition of 5µl of 68mM Trimethylsilyl propionate (TSP), as the internal reference standard for quantification. NMR tubes were then placed in the 400MHz NMR

Machine, accounting for both the presence salt and water and 256 scans were performed.

The linearity of the method was tested in aqueous working solutions of the antimicrobial agent gentamicin, obtained by diluting the 2000mg/L antibiotic stock (0.5,1, 2, 4, 8, 16, 32 and 40mg/L). A slight modification was made to the derivatization method, in that 12μl of working solutions was placed in Eppendorf tubes; the subsequent steps were identical. Linearity of the agar extracts was evaluated by preparing plates (in triplicate) with known homogenous concentrations of gentamicin (2, 4, 8, 16, 32 and 40mg/L), agar cylinders were treated as previously described. Linear regression analysis was used to calculate the slope of the calibration curve, by plotting gentamicin area vs. agar concentration. The antibiotic diffusion was studied in agar as a function of time since the start of incubation (0, 0.5, 1, 2, 4, 6, 8, 16, 24, 48 and 72 hours). Agar samples were taken from the central region of each zone (approximately 3mm). Kinetics were performed in triplicate.

3.2.7.2 Quantification of antimicrobial diffusion and degradation using LC-MS

The gentamicin extraction and derivatisation steps described above (section 3.2.7.1) were repeated, followed by analysis using Liquid Chromatography-Mass Spectroscopy (LC-MS), to increase the sensitivity of the method with lower detection limits. Initial experiments were conducted to produce a calibration curve, with agar extracts comprising 1, 4, 10, 20 and 40mg/L gentamicin using the API 3000 (Applied Biosystems, Analytical Technologies) and the Analyst Classic Algorithm.

The quantitative analysis of gentamicin using LC-MS presented several challenges as gentamicin contains several components which lack chromophoric groups to facilitate spectroscopic detection (Heller *et al*., 2005). Additionally, gentamicin has both hydrophilic and hydrophobic properties and thus cannot be retained by a conventional C18 column (Blanchaert *et al*., 2017). Subsequently, to avoid extensive column optimisation and incorporating a cleaning phase during the agar extraction process, a more expensive fusion column was utilised. The Synergi 4µm Fusion-RP 80 Å, LC Column (100mm x 2mm) selectivity permitted balanced polar and hydrophobic compound retention and resolution. The stationary phase consisted of polar embedded C18 with Tetramethylsilane (TMS) endcapping with a porous silica solid support. The mobile phase was composed of water-acetonitrile (70:30) with a flow rate of 1mL/min and a detection wavelength of 281nm.

3.2.8 Serial passage

Alongside the spatial temporal models proposed within this chapter, the traditional experimental evolution method of serial passage was included as a comparative assessment of resistance evolution within a well-mixed system and solid bacterial growth media. The protocol used for serial passage was described in chapter 2 (Section 2.7.4) and no modifications to the original method were necessary.

3.3 Experimental Results

3.3.1 Antimicrobial susceptibility testing results

The initial antimicrobial agent used within the study was chloramphenicol which was then followed by trimethoprim. The antimicrobial agents colistin, gentamicin, levofloxacin and meropenem was used in subsequent experiments following the optimisation of experimental methods.

The initial susceptibility test carried out was the disc diffusion method described in chapter 2 (Section 2.4.1). The zone diameter breakpoints for the *Enterobacteriaceae* to the given antimicrobials agents utilised within this study were obtained from EUCAST (Table 3.9). The zone diameter breakpoint results from the susceptibility testing of *E. coli* and *P. mirabilis* against the antimicrobials; chloramphenicol and trimethoprim are shown in table (3.10).

Table 3.9 EUCAST susceptibility breakpoint guidelines. The EUCAST breakpoint table displays both the MIC and zone diameter breakpoints for the antimicrobial agents utilised within this study for the *Enterobacteriaceae* including *E. coli and P. mirabilis*.

EUCAST Susceptibility Testing: *Enterobacteriaceae*						
Antimicrobial Agent	**MIC Breakpoint (mg/L)**			**Zone Diameter Breakpoint (mm)**		
	S	I	R	S	I	R
Chloramphenicol	8	-	8	21	-	20
Trimethoprim	2	4	4	17	14-16	13
Colistin	2	-	2	-	-	-
Gentamicin	4	4	2	16	17-19	20
Levofloxacin	1	2	2	17	14-16	13
Meropenem	2	4-8	8	27	20-26	19

Table 3.10 Zone diameter breakpoints. Table displaying the zone diameter breakpoints (mm) for *E. coli* and *P. mirabilis* observed during the disc diffusion experiments with the antimicrobials chloramphenicol and trimethoprim.

Antimicrobial Agent	**Disc Content (µg)**	**Zone Diameter Breakpoint (mm)**	
		E. coli	***P. mirabilis***
Chloramphenicol	30	21	20
Trimethoprim	25	34	25

The results confirmed the susceptibility of *E. coli* and *P. mirabilis* laboratory strains to both chloramphenicol and trimethoprim. This was further supported by the results from the microbroth dilution assay where after 24 hours of incubation at 37 ºC the MIC breakpoints of chloramphenicol and trimethoprim were 8mg/L and 2mg/L respectively, which according to table (3.9) is indicative of bacterial susceptibility to the given antimicrobial agents. The results generated from both susceptibility tests, indicate that chloramphenicol and trimethoprim could be utilised within this the study.

3.3.2 The preliminary diphasic plate experiment

3.3.2.1 The production of Bioplates for the diphasic plate experiment

Prior to experimentation, it was hypothesised that the 'Overhang Bioplate' would be the most suitable Bioplate for use within this diphasic plate experiments, as it would produce a tighter fit within the plate, and the hang would facilitate an easier removal. The results however, disproved the initial hypothesis as a greater amount of contamination was observed with the overhang Bioplate and the agar was disrupted upon removal. This result meant the 'Slip Bioplate' was used in subsequent experiments.

3.3.2.2 The diphasic plate experiment: resistance evolution of laboratory strains of *E. coli* and *P. mirabilis* under chloramphenicol exposure on MH agar

The initial experiment was subject to contamination, which was identified as normal skin flora. In addition to this, the added step of removing the Slip Bioplate and the subsequent filling with non-antibiotic media resulted in contamination along the division line of the two phases within the diphasic plate. Asides from contamination, the 10µL drop of bacterial suspension failed to migrate within the proposed method.

3.3.2.3 The diphasic plate experiment: resistance evolution of laboratory strains of *E. coli* and *P. mirabilis* under chloramphenicol exposure on semi-solid MH agar

To increase the bacterial migratory rates, the consistency of the agar was modified to produce MH agar with consistencies of 50% w/v and 75% w/v. The semi-solid agar with a consistency of 50% w/v was too aqueous. After incubation, the bacteria produced a covering on the surface of the agar, and the antimicrobial agent had precipitated from the media. The semi-solid agar with a consistency of 75% w/v worked best, the streak was effective but care must be taken to avoid tearing of the agar. Again, precipitation of the antimicrobial agent was observed along with contamination. However, filling the divide with agar containing chloramphenicol, significantly reduced the amount of contamination present.

3.3.2.4 The diphasic plate experiment: resistance evolution of laboratory strains of *E. coli* and *P. mirabilis* under chloramphenicol exposure on M9 minimal media

Furthermore, M9 Minimal Media was tested to increase the migratory rates of the bacterial populations. After 24 hours incubation at 37°C the growth of both *E. coli* and *P. mirabilis* was insignificant, thus the media appeared to be having an adverse effect on the growth of the two organisms. There were contaminants present within the media, which indicated that a component of the minimal media was contaminated. Each component added post-sterilisation was tested and the contaminant identified was the 20% v/v glucose solution. Due to the restricted growth of the organism, M9 Minimal media was not used in subsequent diphasic plate experiments.

3.3.3 The triphasic plate experiment: resistance evolution of laboratory strains of *E. coli* and *P. mirabilis* under chloramphenicol exposure on MH agar

After one day of incubation at 37°C *P. mirabilis* had migrated across the entire first phase containing no antimicrobial agent. After two-days incubation, there was a slight discoloration of the agar within the second phase, suggestive of bacterial growth. To examine this further, three samples were taken from the triphasic plate (labelled A, B, C) and were plated onto MH plates containing the MIC of trimethoprim, which maintains the antimicrobial selective pressure whilst selecting for the growth of potential resistant organisms. The mass growth on the antimicrobial plates suggests the growth of potentially resistant organisms within the second phase of the triphasic plate. After one-week incubation, *P. mirabilis* had migrated across all three phases within the triphasic and evolved resistance to x10 the MIC of chloramphenicol, again signified by discoloration of the MH agar containing the antimicrobial agent. Again, three samples were taken and plated onto selective media with a concentration of x10 MIC. After O/N incubation growth was present indicating potentially resistant organisms. There was no contamination present on the triphasic plate after one-week incubation.

The growth of *E. coli* was much slower than that of *P. mirabilis* and the pattern of growth was significantly different. Here the inoculum appeared to bud out of the initial point of inoculation. After one week, the size of the initial inoculum had grown significantly within the initial phase, however the growth proceeded backwards and migration did not occur past the first boundary. After two weeks incubation, the three evolving populations had budded through phase one into the second phase

containing the MIC of chloramphenicol. Samples were obtained from the triphasic plate and cultured onto selective media containing MIC of the antimicrobial agent to maintain the selective pressure. However, following sampling of the triphasic plate a contaminant was identified within the MIC phase. To comprehend the duration for *E. coli* to migrate across the triphasic plate, the experiment was left to proceed. After two weeks incubation, the contaminant had increased in size and was gram stained for bacterial identification. Gram staining revealed a Gram-positive filamentous rod with spores, suggestive of the *bacillus* bacterial genus, which are known environmental contaminants.

3.3.3.1 The triphasic plate experiment: resistance evolution of laboratory strains of *E. coli* and *P. mirabilis* under chloramphenicol exposure in semi-solid MH agar

Due to the increased surface area of the triphasic plate and the large volumes of media required. It was extremely difficult to control contamination during this experimental procedure. As previously discussed in section (3.3.2.3), the reduced agar content failed to increase the migratory rates of the evolving bacterial populations. Moreover, the slime layer produced from the 50% w/v MH experiment was detrimental here as it covered the surface of all three phases, on the sterile tray. Here contamination was identified as Gram-positive rod, with the same morphology as the contaminant described in section (3.3.3). Comparable to findings that observed in section (3.3.2.3), where semi-solid media with a 75% w/v consistency was most effective but due to slow growth and contamination observed, semi-solid agar will not be used within the triphasic plate method.

3.3.3.2 The diphasic agar slope plate experiment: resistance evolution of laboratory strains of *E. coli* and *P. mirabilis* under chloramphenicol exposure

In theory studying the trickling effect of antimicrobial agents on the growth of an organism was very promising, however the large triphasic plate did not facilitate this. Again, as seen with the results obtained in section (3.3.3), the repeated occurrence of contaminants within the proposed experimental evolution models was a major issue.

3.3.4 The preliminary filter tip experimental results

The use of MacConkey agar permitted the differentiation of Gram-negative bacteria that ferment lactose to those that do not. The flow of organism was detected visibly within the micropipette tip, as *E. coli* is a lactose fermenter and appeared pink, whereas *P. mirabilis* is not and appeared colourless within the tip. The presence of growth was confirmed for all replicate samples obtained from filter tips within the micropipette tip for both organisms.

3.3.4.1 The filter tip experiment: resistance evolution of laboratory strains of *E. coli* and *P. mirabilis* resistance under trimethoprim exposure

After seven days, filter tips were removed and selectively sub cultured onto trimethoprim plates with a concentration of 2mg/L for each of the six replicates (Table 3.11). The confirmed resistant cultures were then cryopreserved and stored at -80°C.

Table 3.11 The filter tip experimental results. The table displays the presence or absence of growth of six replicates of *E. coli* and *P. mirabilis* obtained after one-week incubation within the filter tip experiment.

Replicate	Potential R^x *E. coli*	Potential R^x *P. mirabilis*
1	Mass Growth	Mass Growth
2	Mass Growth	Mass Growth
3	Minimal Growth	Mass Growth
4	No Growth	Mass Growth
5	No Growth	Mass Growth
6	Minimal Growth	Mass Growth

The results shown here, identified the presence or absence of potential resistant organisms after one-week incubation within the filter tip experiment. However, as the resistance evolution of bacterial populations can fluctuate daily, it was of interest to quantify the bacterial growth by performing CFU/mL calculations daily within the filter tip experiment.

3.3.4.2 Quantification of the filter tip experiment: resistance evolution of a laboratory strain of *E. coli* under trimethoprim exposure

The quantified filter tip experiment was arduous in terms of the materials and the pre-experimental preparation required, thus the experimental procedure was initially downscaled by using *E. coli* as the primary organism and the experiments were performed in duplicate. An observation was made on day 6 that the agar had desiccated within the P1000 micropipette, leaving the filter tip isolated at the top of the micropipette tip. This meant that the bacterial colony counts were not performed

on day 7 of the filter tip experiment. The daily CFU/mL counts for the two parallel populations of *E. coli* are shown below (Table 3.12).

Table 3.12 The quantification of the filter tip experiment. The table displays the daily colony-forming unit counts of two replicates of *E. coli* (A and B) under trimethoprim selection over a period of six days within the quantified filter tip experiment.

Day 1	**Colony Counts (CFU/mL)**
E. coli A	4.30×10^5
E. coli B	5.20×10^5
Day 2	
E. coli A	9.80×10^5
E. coli B	1.10×10^6
Day 3	
E. coli A	6.80×10^6
E. coli B	3.20×10^7
Day 4	
E. coli A	8.20×10^6
E. coli B	5.40×10^6
Day 5	
E. coli A	9.10×10^6
E. coli B	2.30×10^6
Day 6	
E. coli A	15.60×10^6
E. coli B	8.90×10^6

The results presented in table (3.12), suggested apparent fluctuations between the bacterial colony counts of the duplicate populations of *E. coli*. This could be due to compensated growth as a bacterial fitness cost associated with the acquisition of antimicrobial resistance or because of a fault within the experimental design. Irrespective of the potential cause of this, the experimental procedure must be repeated with the inclusion of a greater number of parallel evolving populations.

3.3.4.3 The altered filter tip experiment: resistance evolution of *E. coli* and *P. mirabilis* under trimethoprim exposure

Due to the fluctuations observed within table (3.12), the experimental method was scrutinised to identify potential caveats. The method of filter tip removal described in section (3.2.3) was identified as possible cause for such variability. When removing the filter tip from the micropipette tip the abrasion against the micropipette tip walls could result in the removal of resistant organisms and or the replacement of resistant organisms with susceptible bacterial populations. This altered method was implemented, however due to unsuccessful experimental controls the filter tip experiments were terminated.

3.3.5 The triphasic circle plate experiment: resistance evolution of *E. coli* under trimethoprim exposure

The preliminary triphasic circle plate experiment was conducted with *E. coli* only. Bacterial contamination was detected within the outer phase after one day of incubation and after day 3 significant dehydration was observed within the outer phase of the triphasic circle plate (Appendix Figure A.1). The small sized triphasic circle plate experiment was subject to substantial contamination, condensation and

agar separation within a condensed space, thus it was decided to repeat the experiment with the inclusion of the organism *P. mirabilis* within a bigger triphasic circle plate format.

3.3.5.1 The altered triphasic circle plate experiment: resistance evolution of *P. mirabilis* under trimethoprim exposure

The bacterium *P. mirabilis* was utilised within the altered triphasic plate experiment, as the swarming capabilities of *Proteus* spp. was expected to promote faster bacterial migration. After one day of incubation *P. mirabilis* had migrated and covered the whole of the first segment containing no antimicrobial agent. After two days of incubation, *P. mirabilis* had budded from a point around the circumference of the inner zone containing the MIC of the given antimicrobial agent. Within a total of three days *P. mirabilis* had grown across the surface of all three zones of the triphasic circle plate. The bacterial growth was made evident by a clear decolourisation of the media when the triphasic circle plate was positioned against a light box (Appendix Figure A.2).

The rates of resistance evolution appeared to differ significantly based on the organism utilised within the proposed experimental evolution model. After one day incubation *E. coli,* had migrated 25mm, and the daily growth of the bacterium within the triphasic circle plate was almost immeasurable. After two weeks, *E. coli* had migrated a total of 37mm within the initial phase containing no antibiotic. In addition, the growth of *E. coli* was morphologically dissimilar to that observed for *P. mirabilis* with sporadic growth within initial zone of the triphasic circle plate (Appendix Figure

A.3). Interestingly, no contamination was observed within the altered triphasic circle plate method for the two organisms.

3.3.5.2 The altered triphasic circle plate experiment with M9 minimal media: resistance evolution of *P. mirabilis* and *E. coli* under trimethoprim exposure

As with the previous methods detailed within this chapter, it was decided to repeat the experiment with the inclusion of M9 minimal media alongside MH agar. The results uncovered comparable results with that observed in section (3.3.5.1) using MH agar. Following day one of incubation *E. coli* had migrated 31mm within the initial phase of the plate containing no antimicrobial agent. Conversely, zonal separation was observed on day one, where the phases moved upon manipulation. This resulted in the termination of the experiment following day two of incubation. After two days, *E. coli* had migrated a further 1mm within the initial zone. In contrast, ensuing one-day incubation *P. mirabilis* had migrated 51mm within the initial zone containing no antimicrobial agent. After day two the *P. mirabilis* had evolved resistance to trimethoprim, through migrating 80mm and encompassing the entire phase containing the MIC of the antimicrobial agent.

In comparison to the triphasic circle plate experiments utilising M9 minimal media, after day one of incubation, the plates appeared moist and it was difficult to visualise the growth of the inoculum for both *E. coli* and *P. mirabilis*. The plates were subsequently placed against a light source and the outline of the bacterial inoculum was marked. Analysis of the control plates revealed no bacterial growth on the plates containing no antimicrobial agent. Therefore, it was concluded that the original growth marked on the triphasic circle plate, was likely to be a watermark and not bacterial

growth. Subsequently, the M9 minimal media triphasic circle plate experiments were terminated. A comparison of the bacterial growth of the two organisms on both MH and M9 minimal media within the triphasic circle plate over the two-day period of incubation can be found in the appendix (Figure A.4).

The triphasic circle method was amended to decrease the amount of agar separation between zones, which was impeding bacterial migration. The alterations made to the triphasic circle plate experiment described in section (3.2.4.1) was successful as zonal separation was not observed during a two-week trial experiment. Despite, the effective alteration in experimental approach, there was concerns regarding the requirement of bacterial motility within the proposed model.

3.3.6 Assessment of bacterial motility

3.3.6.1 The motility stab test

The motility stab test results indicated *E. coli* MG1655 is weakly motile (Figure 3.7), as shown by the marginal swarming from the initial line of inoculation into the surrounding media. This motile organism *P. mirabilis* was utilised within the motility test as the positive control. The results confirmed that *E. coli* MG1655 was considerably less motile than *P. mirabilis*, illustrated by the extensive diffusion throughout the medium by the motile organism. This provided a further explanation for the increased migratory rates of *P. mirabilis* observed within the triphasic circle plate experiments. Moreover, this result indicated that measures should be explored to enhance the bacterial motility of *E. coli* MG1655 for its application within the proposed experimental evolution model.

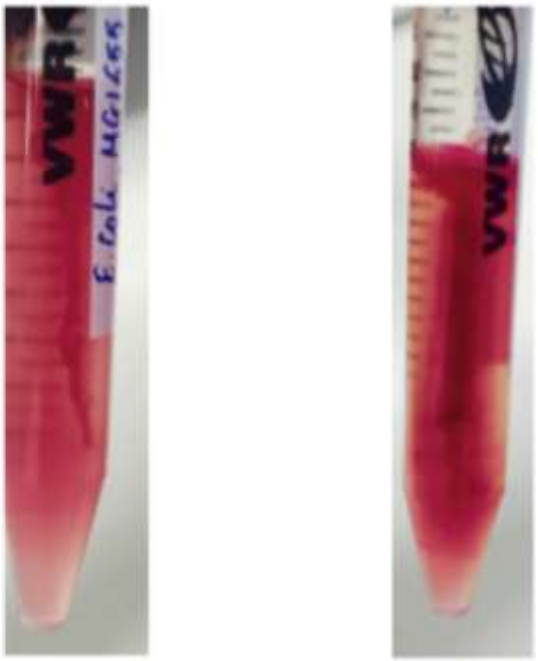


Figure 3.7 The motility stab test. The image shows the motility stab test results for both *E. coli* (left) and *P. mirabilis* (right). A positive motility test result was indicated by the diffusion of bacterial growth from the initial point of inoculation into the surrounding media. *P. mirabilis* was used in this test as the positive control.

3.3.6.2 Genomic analysis of motility

The genomic analysis of motility was conducted primarily to screen for the presence of IS elements within the flhD promoter region of the flhD operon. The expected sizes of the PCR products when IS elements were present within the flhDC promoter region was 0.85kb. Gel electrophoresis of the five strains of *E. coli* indicated prominent flhDC expression in *E. coli* strains; CFT073, TPA2743 and TPA4792 (Appendix Figure A.5). However, under the selected PCR thermocycling conditions IS elements were not detected, as no bands were identified between 0.75 and 1.0kb. This genomic results also suggested low expression of the flhDC operon within the *E. coli* strains MG1655 and UTI89 as no bands were present at 0.75kb. Therefore, the reduced motility observed in *E. coli* MG1655, may be attributable to low levels of flhDC expression,

essential to the initiation of flagellar synthesis. To increase the likelihood of IS element detection, PCR reactions were repeated with an increased amplification time of ten minutes. The extended amplification time facilitated the identification of IS elements for the *E. coli strains* CFT073, TPA2743 and TPA4792. Similarly, the repeated PCR experiments revealed undetectable levels of flhDC expression within *E. coli* MG1655 again supporting the previous results of reduced motility within the proposed evolutionary models.

3.3.6.3 Motility assay

When analysing the literature for studies investigating the genomics of bacterial motility, it was noticed that phenotypic motility assays were conducted using broth solidified with bacteriological agar raging from 0.1-3% w/v. The qualitative results from the motility assays suggested that motility layers containing 0.3%-3% w/v agar could be utilised within the Triphasic circle method to increase motility of the weakly motile strains of *E. coli* (Appendix Figure A.6).

3.3.7 The triphasic circle plate experiment: ARGP with 0.3% motility layer

The Triphasic plate experiment was modified through the addition of a motility layer, and renamed the ARGP. The preliminary ARGP experiments were conducted on the *E. coli* strain MG1655 under gentamicin selection. Initially a motility layer of 0.3% w/v agar was tested see below (Figure 3.8), however the liquid consistency made inoculation and incubator transfer extremely difficult. Following O/N incubation, the plates were concealed in bacterial growth, which reflected the consistency of the motility layer as opposed to the evolution of potentially drug resistant organisms. This

was confirmed by sampling organisms from the ARGP onto agar plates containing the MIC of gentamicin.

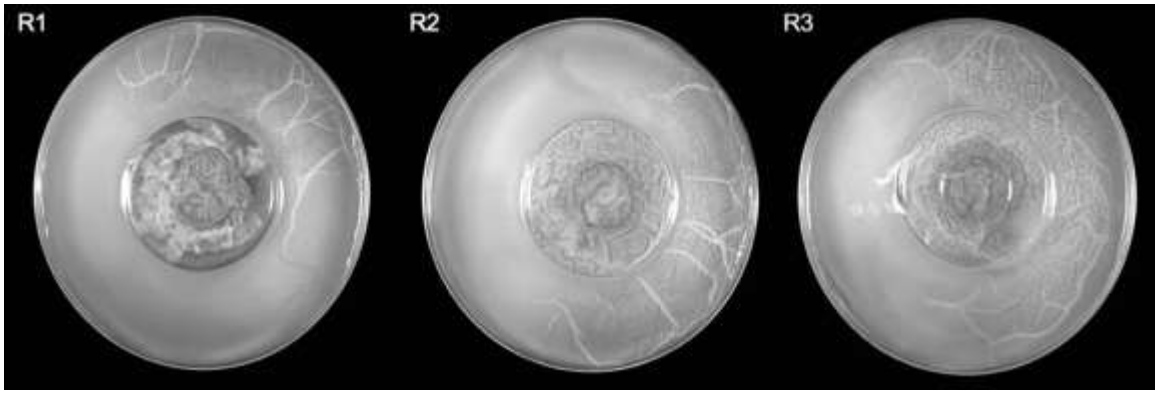


Figure 3.8 The ARGP 0.3% motility layer. Figure displaying three replicates of *E. coli* MG1655 under gentamicin section following O/N incubation on the ARGP using a 3% motility layer. Images were taken using the G-Box imaging system.

3.3.7.1 The triphasic circle plate experiment: ARGP with 3% Motility Layer

To overcome the issues faced above (Section 3.3.7), the consistency of the motility layer was increased to 3% w/v agar. The consistency of this layer facilitated the precise inoculation of the initial zone containing no antibiotic. Following O/N incubation, the clear budding of bacterial colonies from the central zone into the phase containing the MIC was evident. Following day two of incubation the bacteria had migrated into the third phase containing x10 the MIC of gentamicin. The 0.3% w/v motility layer, therefore successfully permitted the visualisation of the evolution of bacterial resistance across an antibiotic landscape (Figure 3.9). The distinct buds of bacterial colonies, were sampled daily and cryopreserved for later genomic analysis.

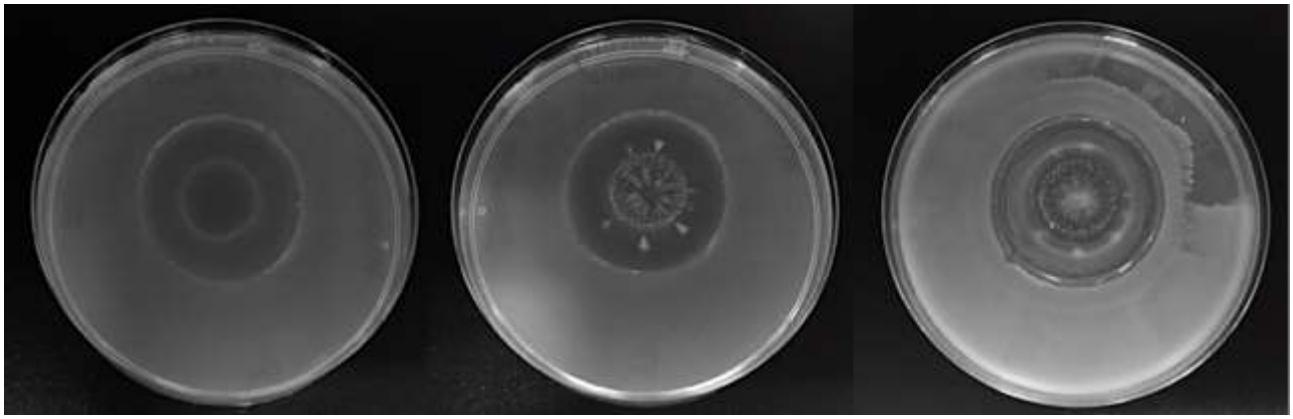


Figure 3.9 The ARGP 3% motility layer: *E. coli* MG1655. G–Box images taken displaying the resistance evolution of *E. coli* MG1665 under gentamicin selection in the ARGP over a period of three days. The contrast in these images has been slightly adjusted to improve visualisation but otherwise remain unedited.

To further test the consistency of the motility layer, the experiment was repeated with the inclusion of the uropathogenic strains of *E. coli* (UTI89, CFT073, TPA4792 and TPA2743). Prior to experimentation within the ARGP, antimicrobial susceptibility tests were conducted on all strains using the microbroth dilution assay (Section 2.4.1). The MIC of all four uropathogenic strains of *E. coli* under gentamicin selection was confirmed as 4mg/L.

The antimicrobial concentrations within the phases of the ARGP were modified accordingly, before experiments were repeated with the inclusion of the 3% w/v motility layer. The results confirmed that the 3% w/v agar consistency, enhanced bacterial motility and enabled the visualisation of resistance evolution. This can be seen below (Figure 3.10), where the uropathogenic strain of *E. coli* CFT073 evolved resistance to x10 the MIC of gentamicin in a period of two days as seen with *E. coli* MG1665. However, there are clear distinctions between the two bacterial strains in the patterns of resistance evolution across the antimicrobial landscape.

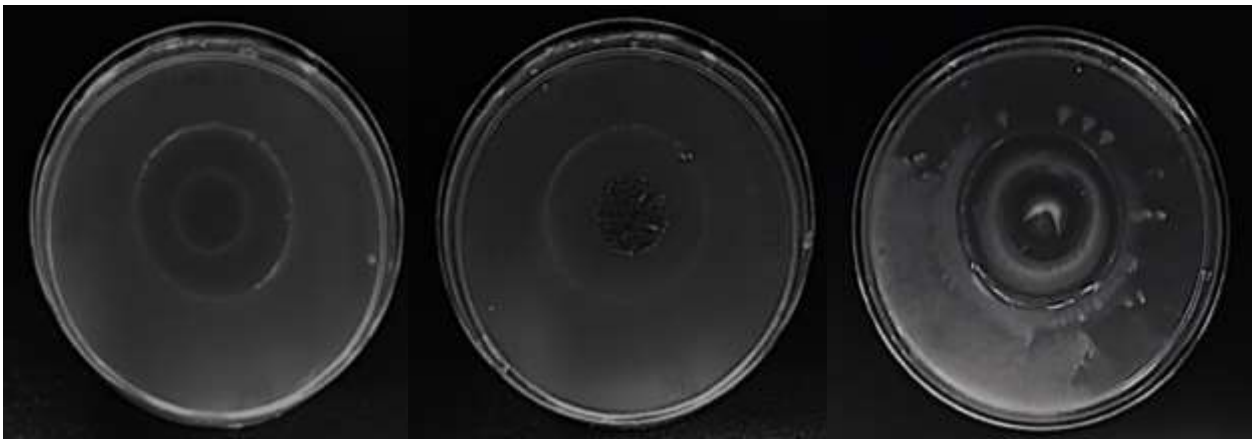


Figure 3.10 The ARGP 3% motility layer: *E. coli* CFT073. G–Box images taken displaying the resistance evolution of the uropathogenic strain of *E. coli* CFT073 under gentamicin selection in the ARGP over a period of three days. The contrast in these images has been slightly adjusted to improve visualisation but otherwise remain unedited.

3.3.8 Assessment of antimicrobial degradation and diffusion within the ARGP

3.3.8.1 NMR spectra of the selected antimicrobial agents and media utilised within this study

NMR was evaluated as a method to detect the presence of antimicrobial agents within the ARGP and to account for the degree of diffusion and degradation over the experimental time course. Initial NMR experiments were conducted on a range of antimicrobial agents and agarose in powdered form. Through the visual inspection of the stacked plot, the results were encouraging as distinctive peaks were observed specific to the given class of antimicrobial agents (Figure 3.11).

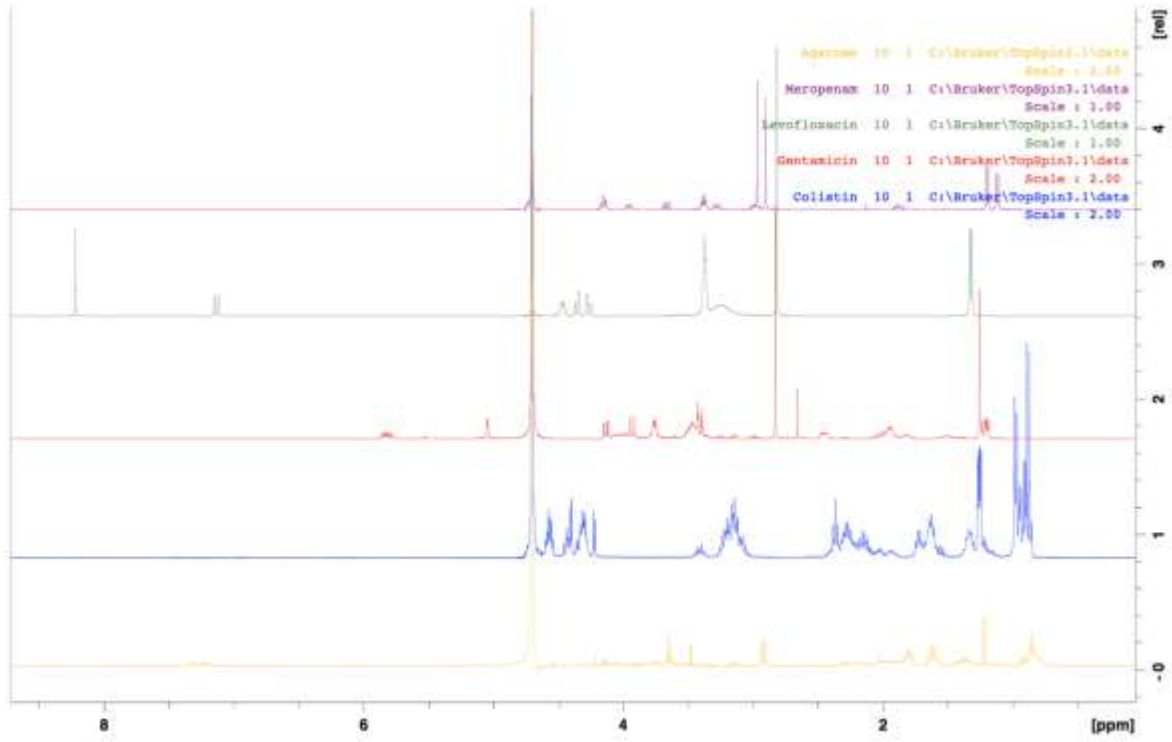


Figure 3.11 One-dimensional (1D) ^{1}H NMR spectra. Image displaying the stack plot representative of the one-dimensional (1D) ^{1}H NMR spectra of the antimicrobial agents: meropenem (purple), levofloxacin (green), gentamicin (red) and the agarose media (yellow) utilised within the study.

Following the successful detection of the independent antimicrobial agents and agarose media using NMR, it was of interest to repeat the analysis to determine whether the principal antimicrobial agent gentamicin could be identified within MH agar samples at varying antibiotic concentrations. This was achieved using the boiling method. The (1D) ^{1}H NMR Spectra below (Figure 3.12) displays three concentrations of gentamicin (0, MIC and x10 MIC), utilised within the ARGP model. When comparing the NMR spectra of gentamicin shown in blue within figure (3.11), to the spectra obtained in figure (3.12), there is a clear reduction in PPM when the antimicrobial agent is added to the agarose containing bacteriological growth media. The result suggested that the detection limits of NMR were not feasible for the identification and quantification of antimicrobial agents within agar samples.

Therefore, it was decided to implement another analytical approach which may provide greater sensitivity.

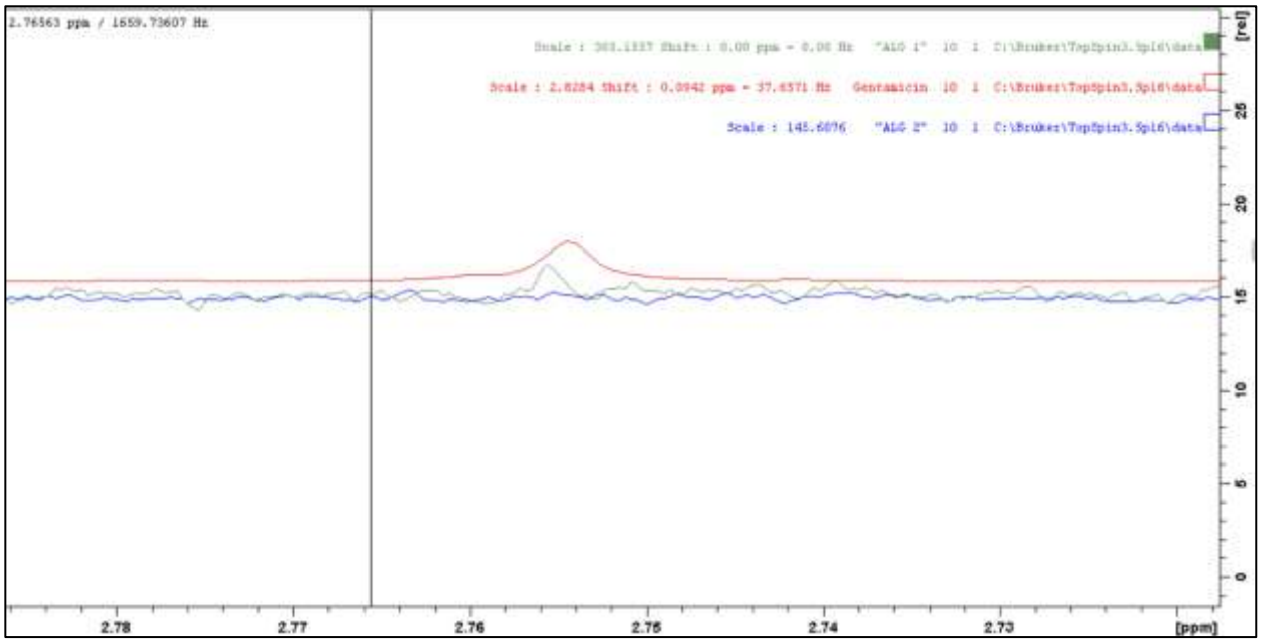


Figure 3.12 Representative 1D ^{1}H NMR spectra of the gentamicin agar extracts. Image displaying the stack plot representative of the one-dimensional (1D) ^{1}H NMR spectra of the gentamicin MH agar extracts. The examined gentamicin agar extract concentrations included: No antibiotic (0mg/L) (green), MIC (2mg/L) (blue) and X10 MIC (20mg/L) (red).

3.3.8.2 Quantification of antimicrobial diffusion and degradation using LC-MS

Further investigation into the analytical approaches for aminoglycoside detection, revealed complications with the antimicrobial agent gentamicin. Gentamicin as a complex consists as four major components, denoted C1, C1a and C2, C2a which can be separated into three distinct groups (1-3) (Figure 3.13) (Heller *et al.*, 2005). The importance of each component was uncovered when LC-MS was employed in the quantification of gentamicin as the four components had three different molecular weights (Table 3.13).

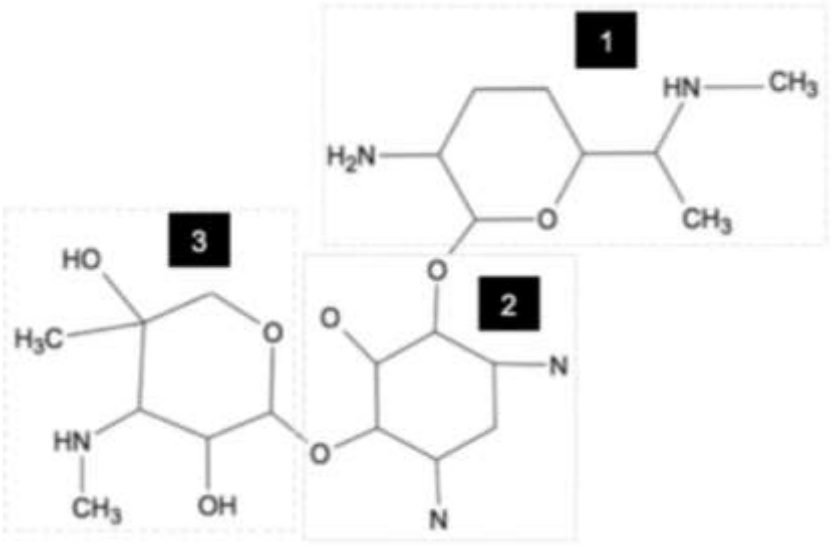


Figure 3.13 2D molecular structure of the gentamicin complex. Figure display the 2D structure of the gentamicin complex. The three sections labelled 1, 2 and 3 represent the three major components C1, C1a and C2.

Table 3.13 Gentamicin components and fragmentation for LC-MS. Table displaying the main mass spectrometric fragmentation pathways of the aminoglycoside gentamicin complex using LC-MS.

Gentamicin	Component	m/Z of MH +Ion
1	C2, C2a	464
2	C1a	450
3	C1	478

The initial LC-MS method was developed primarily to determine whether LC-MS could be utilised as an analytical approach for the detection of antimicrobials within agar extracts of varying antibiotic concentrations. The LC-MS tabular and chromatogram output, suggested the presence of the gentamicin within the agar extracts (Figure 3.14). However, the calculated antimicrobial concentration values

were not as predicted, with the gentamicin concentrations being overestimated for all extracts except 40mg/L (Cal 5 40mg/L) (Table 3.14). Nonetheless, as a valid antimicrobial internal reference standard was not included in the initial analysis and that only one component of the gentamicin complex was considered, the calculated gentamicin concentrations were disregarded.

Table 3.14 LC-MS tabular output for a series of gentamicin agar extractions. Table showing the LC-MS output following gentamicin agar extraction of a series of target antibiotic concentrations (1mg/L, 4mg/L, 10mg/L, 20mg/L and 40mg/L) and the corresponding calculated concentration value based on the LC-MS output.

Sample Name	Sample Type	Area (cps)	RT (min)	Target [Conc]. (mg/L)	Calculated [Conc]. (mg/L)
Cal 1 1mg/l	Standard	1.080e+04	2.09	1.00	N/A
Cal 2 4mg/l	Standard	1.810e+04	2.12	4.00	10.9
Cal 3 10mg/l	Standard	4.400e+04	2.09	10.0	14.4
Cal 4 20mg/l	Standard	2.230e+04	2.07	20.0	26.1
Cal 5 40mg/l	Standard	1.150e+04	2.12	40.0	35.5

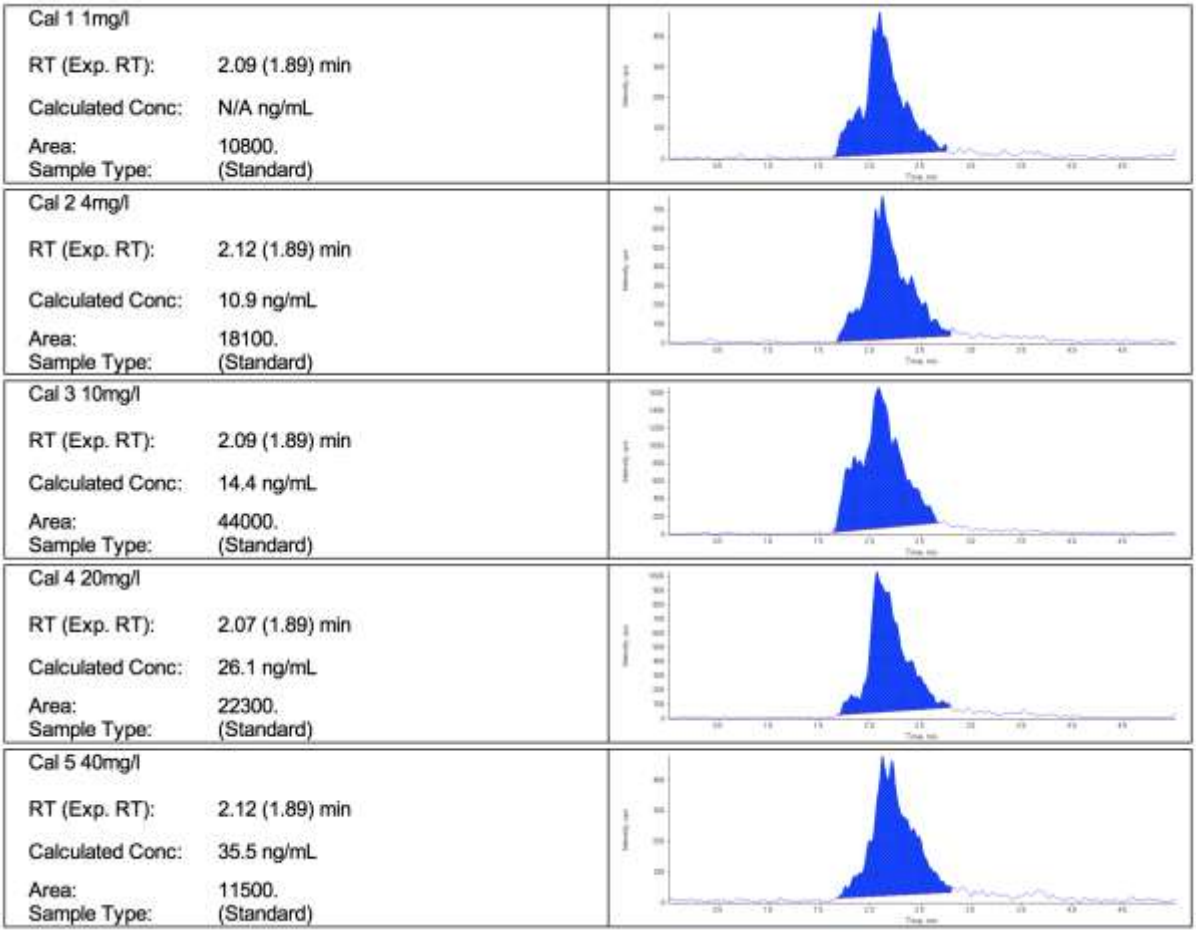


Figure 3.14 The preliminary LC-MS chromatogram output of a series of gentamicin agar extractions. LC-MC chromatogram curves displaying the calculated concentrations of the antimicrobial agent gentamicin from the MH agar extracts.

Following the qualitative assessment of gentamicin using LC-MS, kanamycin was added as an internal reference standard to facilitate the quantification of the gentamicin agar extracts. Kanamycin was selected as the internal reference standard as it is an aminoglycoside antimicrobial agent of similar molecular weight to gentamicin.

The LC-MS quantification of gentamicin by LC-MS was performed on agar extracts of a series of gentamicin concentrations (4mg/L, 8mg/L, 10mg/L, 20mg/L and 40mg/L) spiked with 1mg/L of kanamycin. The LC-MS tabular output for the selected concentrations are discussed independently below, and the LC-MS chromatograms for the agar extractions can be found in the appendix (A.7-A.11).

The LC-MS results for the 4mg/L gentamicin agar extracts (Table 3.15), indicated fairly accurate predictions of retention time. Though, inconsistencies were raised when looking at the area counts and predicted gentamicin concentrations. Like that seen previously, predicted concentrations were significantly higher than the expected 4mg/L, with m/z 464 and 450 concentrations increased by fourfold.

Table 3.15 LC-MS tabular output of the 4mg/L gentamicin extract. Displaying the LC-MS output of the three components of the gentamicin complex following the 4mg/L gentamicin agar extraction.

4mg/L Gentamicin MH Agar Extract			
Compound Name	Gentamicin 1 (464.292/322.300)	Gentamicin 2 (450.317/322.200)	Gentamicin 3 (478.325/322.200)
Expected RT	1.91	1.98	2.15
Actual RT	1.93	1.95	2.12
Area Counts	33200	64100	79000
ITSD Area Counts	9430	9430	9430
Amount (mg/L)	16.4	16.0	12.1

The LC-MS results for the 8mg/L gentamicin agar extracts (Table 3.16), were comparable to the 4mg/L agar extracts in terms of the predicted retention times. However, the ITSD area counts for all three gentamicin components were equal and the expected concentrations were in this case lower than the expected 8mg/L.

Table 3.16 LC-MS tabular output of the 8mg/L gentamicin extract. Displaying the LC-MS output of the three components of the gentamicin complex following the 8mg/L gentamicin agar extraction.

8mg/L Gentamicin Mueller Hinton Agar Extract			
Compound Name	Gent 1 (464.292/322.300)	Gent 2 (450.317/322.200)	Gent 3 (478.325/322.200)
Expected RT	1.86	1.93	2.10
Actual RT	1.96	1.93	2.10
Area Counts	234000	46700	53800
ITSD Area Counts	19700	19700	19700
Amount (mg/L)	5.53	5.58	3.95

The LC-MS results for the 10mg/L gentamicin agar extracts (Table 3.17), were comparable to the 8mg/L agar extracts with equal ITSD area counts for all three gentamicin components and lower than the expected concentrations than 10mg/L.

Table 3.17 LC-MS tabular output of the 10mg/L gentamicin extract. Displaying the LC-MS output of the three components of the gentamicin complex following the 10mg/L gentamicin agar extraction.

10mg/L Gentamicin Mueller Hinton Agar Extract			
Compound Name	Gent 1 (464.292/322.300)	Gent 2 (450.317/322.200)	Gent 3 (478.325/322.200)
Expected RT	1.91	1.98	2.15
Actual RT	2.04	2.08	2.21
Area Counts	17300	37100	48900
ITSD Area Counts	10400	10400	10400
Amount (mg/L)	7.80	8.42	6.83

The LC-MS results for the 20mg/L gentamicin agar extracts (Table 3.18), again had equal ITSD area counts for all three gentamicin components with even lower than the expected concentrations than that seen with the 10mg/L extracts.

Table 3.18 LC-MS tabular output of the 20mg/L gentamicin extract. Displaying the LC-MS output of the three components of the gentamicin complex following the 20mg/L gentamicin agar extraction.

20mg/L Gentamicin Mueller Hinton Agar Extract			
Compound Name	Gent 1 (464.292/322.300)	Gent 2 (450.317/322.200)	Gent 3 (478.325/322.200)
Expected RT	1.97	2.05	2.22
Actual RT	2.01	2.04	2.17
Area Counts	19600	36400	71600
ITSD Area Counts	15700	15700	15700
Amount (mg/L)	5.83	5.47	6.61

The LC-MS results for the 40mg/L gentamicin agar extracts, again had equal ITSD area counts for all three gentamicin components, and surprisingly all three components displayed the lowest predicted concentrations of all tested gentamicin agar extracts.

Table 3.19 LC-MS tabular output of the 40mg/L gentamicin extract. Displaying the LC-MS output of the three components of the gentamicin complex following the 40mg/L gentamicin agar extraction.

40mg/L Gentamicin Mueller Hinton Agar Extract			
Compound Name	Gent 1 (464.292/322.300)	Gent 2 (450.317/322.200)	Gent 3 (478.325/322.200)
Expected RT	1.90	1.97	2.14
Actual RT	2.02	1.99	2.43
Area Counts	24500	48500	65300
ITSD Area Counts	30800	30800	30800
Amount (mg/L)	3.71	3.71	3.07

The results indicated variable responses for m/z 478, 464 and 450 (C1, C2+C2a and C1a), within samples of the same antimicrobial concentration. This strongly suggests inconsistencies within the gentamicin extraction method. Of the three components, it seems m/z 478 would be most appropriate in the detection of gentamicin within the agar extracts.

Looking at the graph (Figure 3.15) of all three components of the gentamicin complex, an inverse relationship is seen between antimicrobial concentration and peak area, again supporting the evidence that the extraction method requires optimisation prior to accurate quantification of gentamicin within the ARGP.

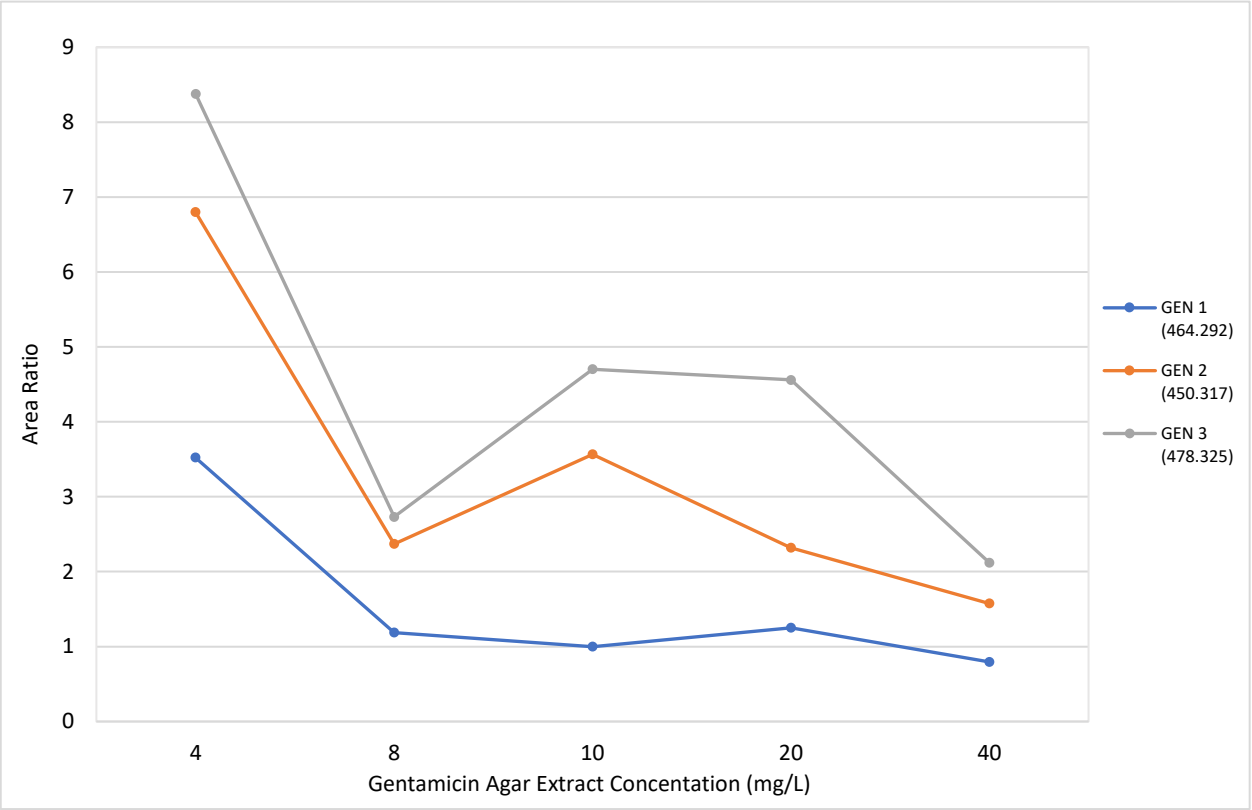


Figure 3.15 LC-MS gentamicin agar extraction quantification. Graph showing the LC-MS output of the five different concentrations of the gentamicin agar extracts for the all three major components of the gentamicin complex: GEN 1 (464.292) (blue), GEN 2 (450.317) (orange) and GEN 3 (478.325) (grey).

To conclude, having assessed all the LC-MS results, there was enough evidence to assume the presence of gentamicin within all MH agar extracts attained from the ARGP experiments. However, until the extraction protocol has been fully optimised, quantification of the antimicrobial concentrations within the agar extracts could not be accurately calculated using LC-MS.

3.3.9 Serial passage: laboratory strain *E. coli* MG1655

To determine the optimal duration of incubation prior to serial transfer an automated 12-hour growth curve was conducted in MH broth. As the mid-late exponential phase of bacterial growth is desirable, the results from the graph below indicate that serial transfer should occur after 4 hours incubation at 37°C (Figure 3.16). It is important to acknowledge that as growth rate of *E. coli* MG1655 is inhibited under antimicrobial selection, the duration of incubation was increased where appropriate.

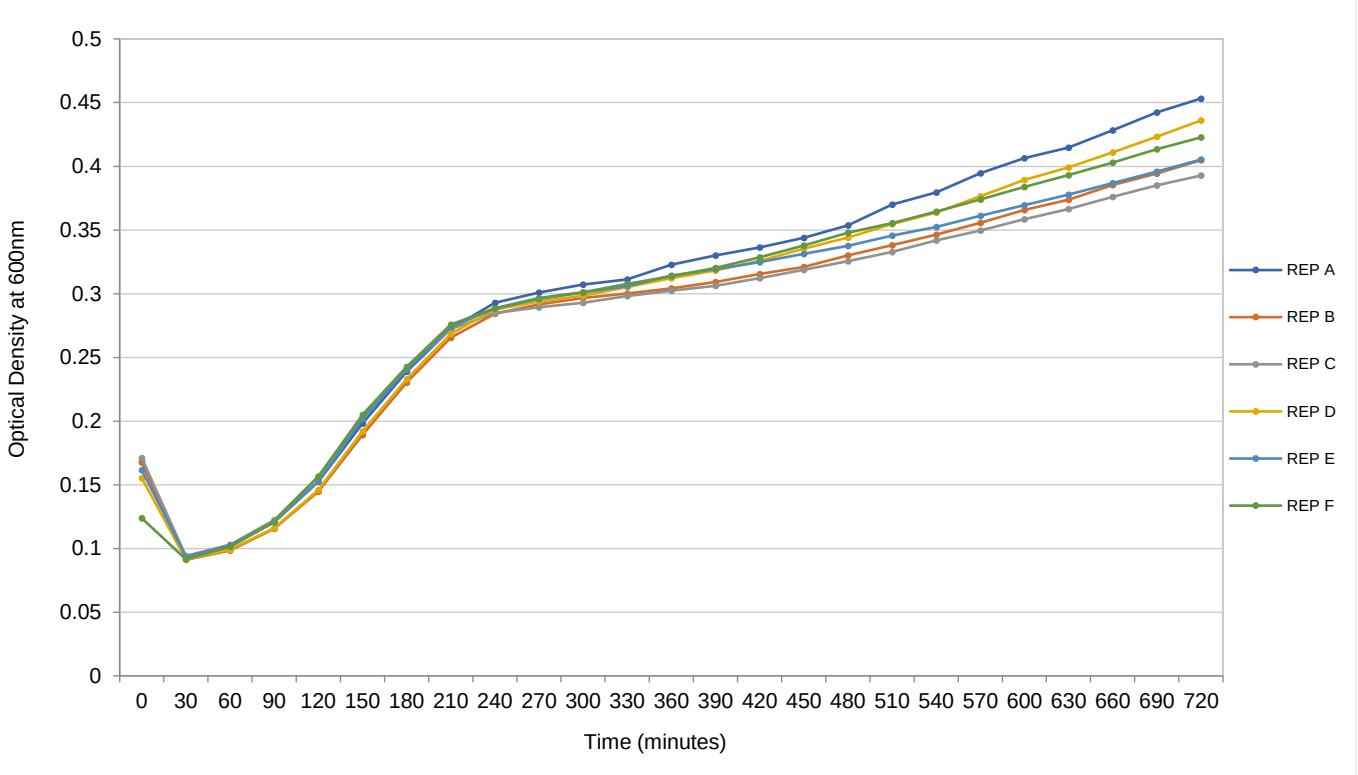


Figure 3.16 Automated growth curve of *E. coli* MG1655 over 12-hours. Graph showing the OD_{600} values of six parallel evolving populations of *E. coli* MG1655 during a 12-hour growth curve using the TECAN automated plate reader to determine the optimal time-period of incubation prior to broth transfer within the serial passage system.

The serial passage results for *E. coli* MG1655 were quantified by both OD measurements and CFU/mL calculations under antimicrobial selection with gentamicin and trimethoprim.

The serial passage experiments were conducted with 3 parallel evolving populations of *E. coli* in triplicate (E1-E9) over a series of increasing antimicrobial concentrations. The results from the graph below displaying the OD measurements of *E. coli* MG1655 under gentamicin selection, revealed similar trends for all 9 parallel evolving populations at most of selected antimicrobial concentrations. What is evident is the fluctuations in OD which occur between transfer (Figure 3.17). The OD values appear to rapidly increase followed by periods of rapid decline, which may reflect the population bottlenecks which arise during serial passage. To comprehend this further bacterial counts were performed and CFU/mL were calculated at each of the antimicrobial concentrations.

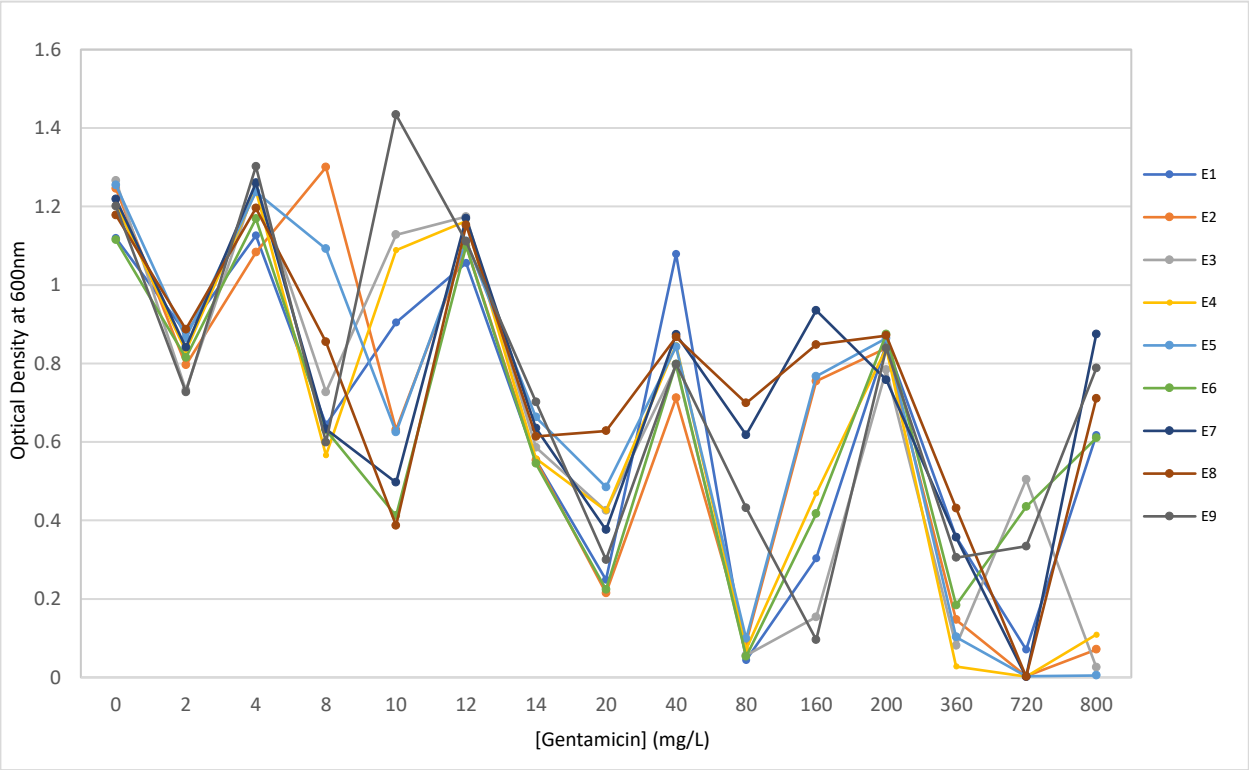


Figure 3.17 Serial passage of *E. coli* MG1655 under gentamicin selection. Graph displaying the evolution of gentamicin resistance for three parallel evolving populations of *E. coli* in triplicate (E1-E9) measuring the OD_{600} over increased concentrations of gentamicin within a serial passage system.

Consistent with the results obtained from the OD measurements, comparable trends are seen with the log10 CFU/mL calculations, with parallel populations displaying similar phenotypic trajectories (Figure 3.18). Repeatedly, periods of rapid bacterial growth are followed by significant log reductions. This is demonstrated at 160mg/L of gentamicin where all parallel evolving strains appear to be non-growing, until serial transfer into a higher antimicrobial concentration where growth rapidly resumes. This could be attributable to mutagenesis events.

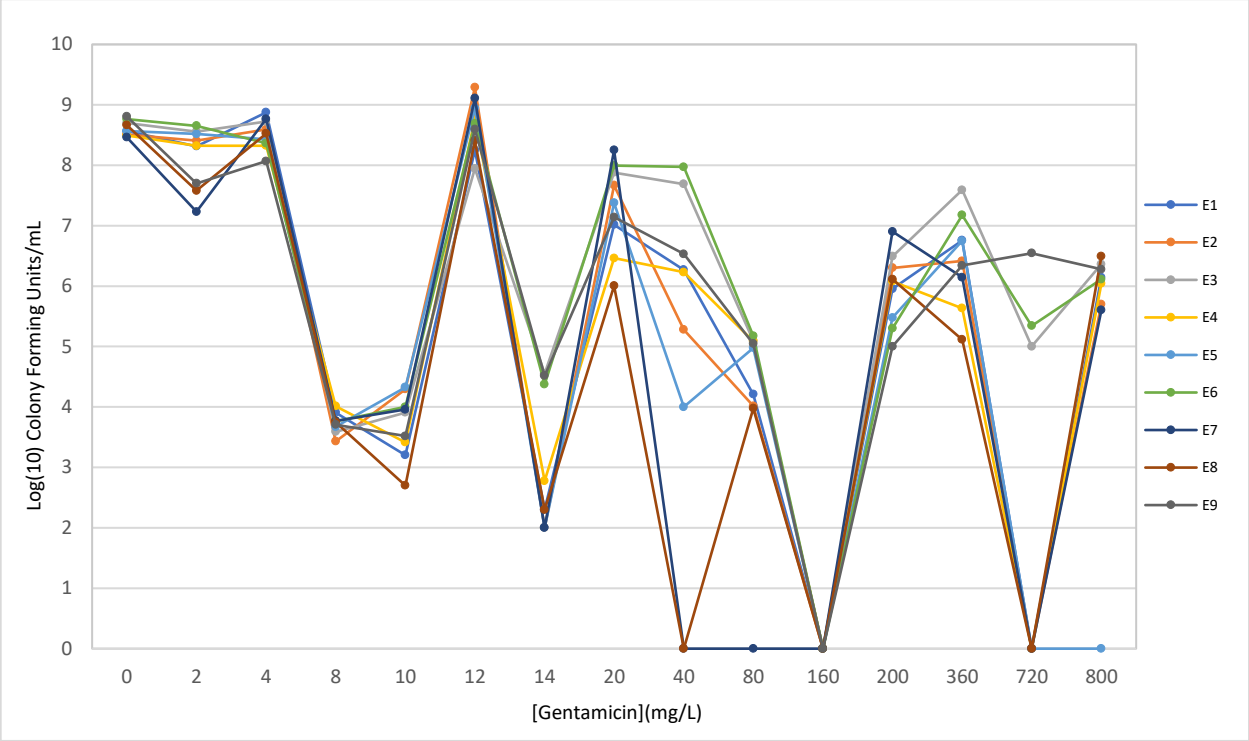


Figure 3.18 Bacterial counts from the serial passage of *E. coli* MG1655 under gentamicin selection. Graph displaying the evolution of gentamicin resistance for three parallel evolving populations of *E. coli* in triplicate (E1-E9) measuring the bacterial colony counts (log10 CFU/mL) over increased concentrations of gentamicin within a serial passage system.

3.3.9.1 Serial passage: *E. coli* MG1655 under trimethoprim selection

Serial passage experiments were repeated with the inclusion of the antimicrobial agent trimethoprim. The OD measurements unlike that for *E. coli* under gentamicin selection, showed inconsistencies between the 9 parallel evolving populations at the lower antimicrobial concentrations, apparent at 10 and 12mg/L (Figure 3.19).

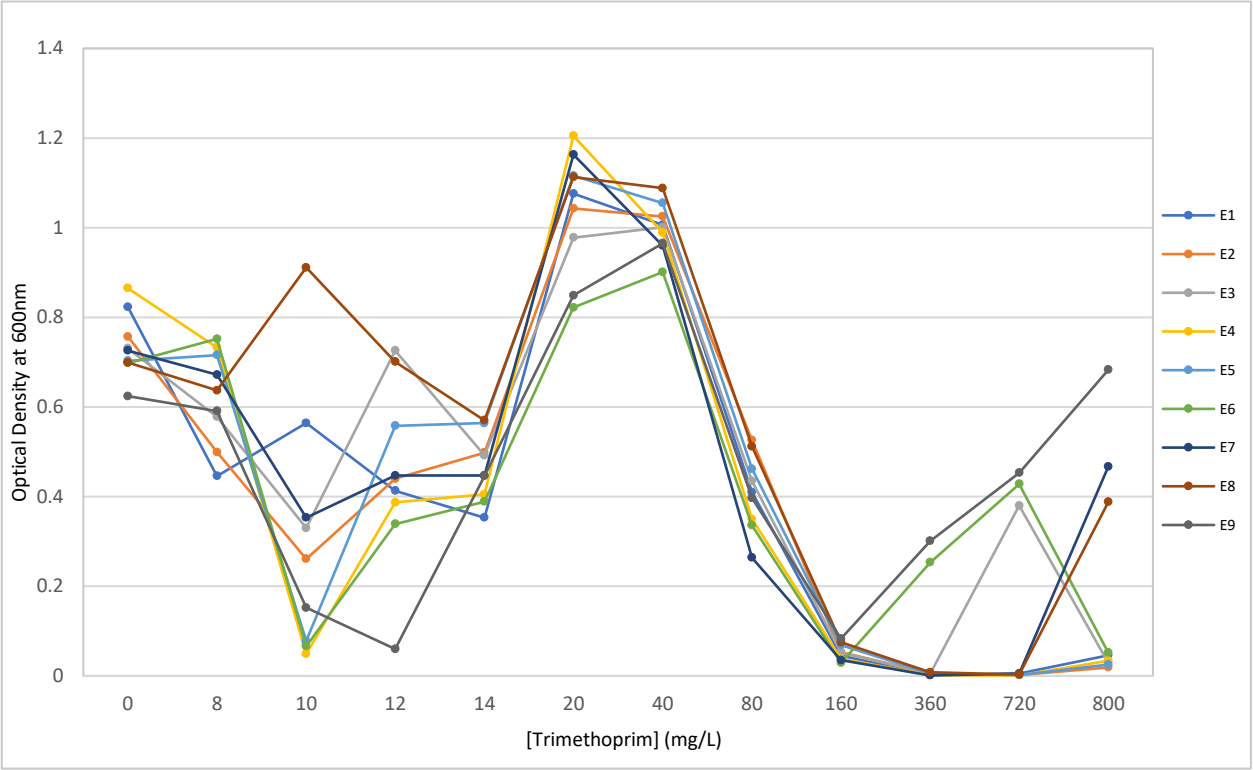


Figure 3.19 Serial passage of *E. coli* MG1655 under trimethoprim selection. Graph displaying the evolution of trimethoprim resistance for three parallel evolving populations of *E. coli* in triplicate (E1-E9) measuring the OD_{600} over increased concentrations of trimethoprim within a serial passage system.

The results from the log10 CFU/mL calculations showed less fluctuations in terms of the periods where rapid bacterial growth followed rapid decline, yet this trend was seen at the two extremes of the selected trimethoprim concentrations, particularly 10-12mg/L and 160-360mg/L (Figure 3.20).

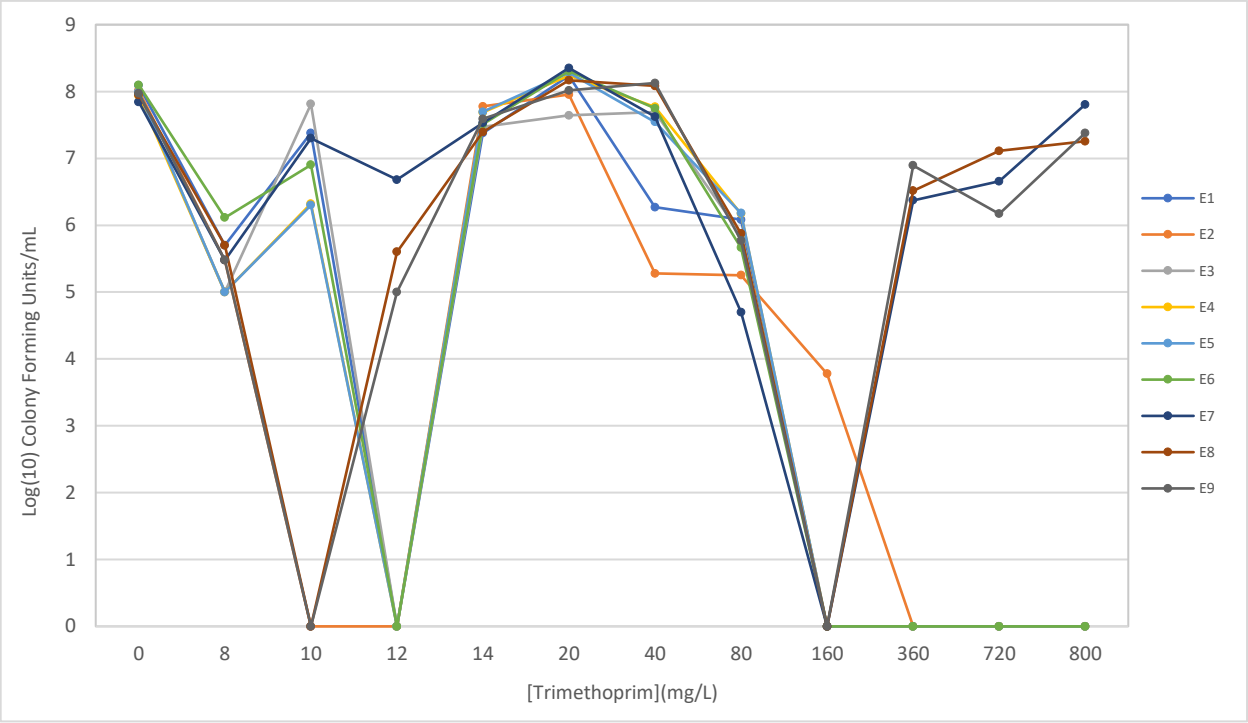


Figure 3.20 Bacterial counts from the serial passage of *E. coli* MG1655 under trimethoprim selection. Graph displaying the evolution of trimethoprim resistance for three parallel evolving populations of *E. coli* in triplicate (E1-E9) measuring the bacterial colony counts (log10 CFU/mL) over increased concentrations of trimethoprim within a serial passage system.

3.3.9.2 Serial passage: uropathogenic clinical strains of *E. coli*

In agreement with the previous antimicrobial susceptibility results for the motile clinical strains of *Escherichia coli* UTI89, CFT073, TPA2743 and TPA4792. Serial passage experiments were not performed using the antibiotic Trimethoprim and experiments with clinical strains were repeated solely under gentamicin selection.

Prior to serial passage experiments, a manual growth curve was conducted on the four new clinical strains to determine the optimal period of incubation prior to transfer (Figure 3.21). As expected, there are clear discrepancies between the OD measurements for the four strains over the 12-hour period. Here when looking at the

averaged OD measurements, the mid-late exponential phase of growth for all four clinical strains appears to occur later than that seen with *E. coli* MG1655 at around the 7-hour time point.

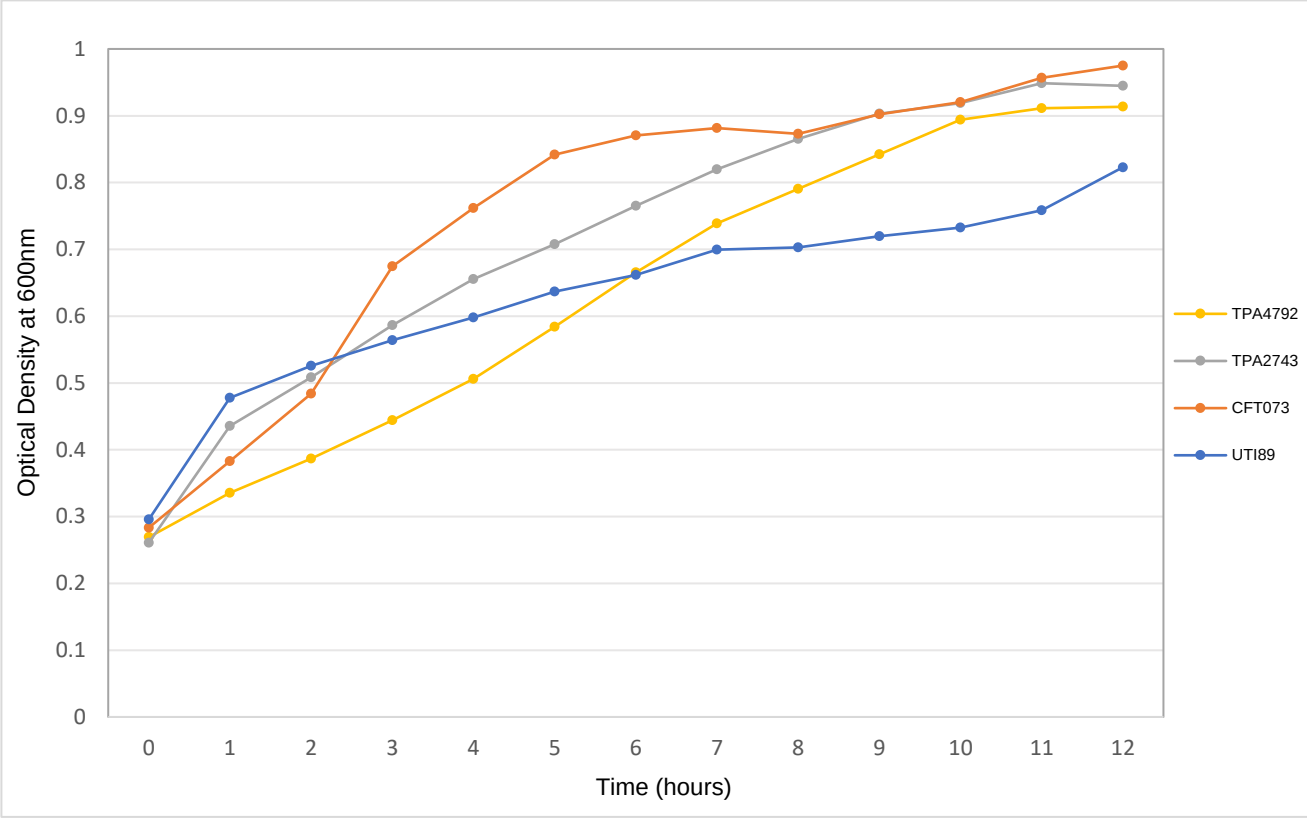


Figure 3.21 Manual growth curve of uropathogenic strains of *E. coli* over 12-hours. Graph showing the averaged OD_{600} values of four uropathogenic strains of *E. coli* (UTI89, CFT073, TPA2743 and TPA4792) over a 12-hour growth curve, to determine the optimal time-period of incubation prior to broth transfer within the serial passage system.

The serial passage experiments for the four clinical strains were conducted with 3 parallel evolving populations per strain (E1-E3). The graph displaying the averaged OD of the four strains showed a rapid reduction upon the first transfer into a broth containing antimicrobial agent (Figure 3.22). This was then followed by bacterial growth recovery and maintenance with fluctuations occurring between OD_{600} of 0.6 and 1 OD units. When comparing these results to that of *E. coli* MG1655 there is a

clear difference in the magnitude of growth fluctuations observed, with the most pronounced reduction here being for the CFT073 strain at 40mg/L gentamicin.

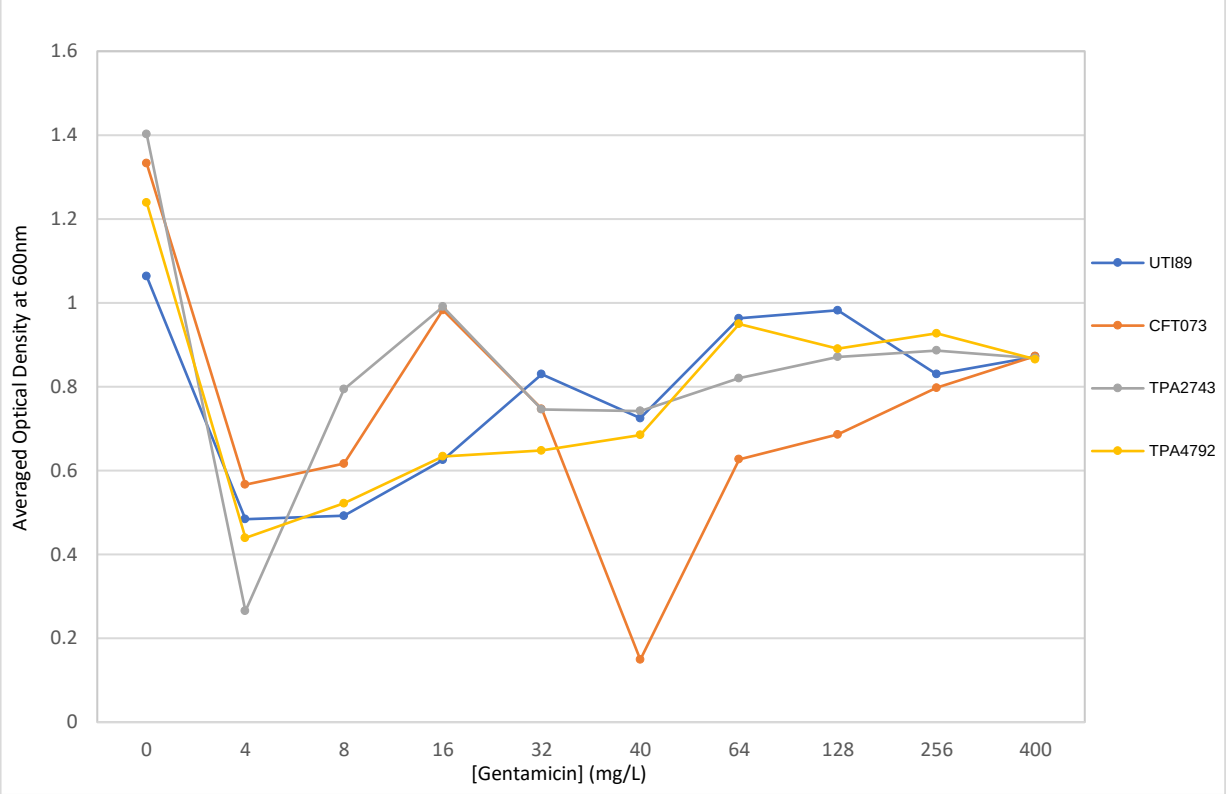


Figure 3.22 Serial passage of uropathogenic *E. coli* strains under gentamicin selection. Graph displaying the evolution of gentamicin resistance for four parallel evolving populations of uropathogenic strains of *E. coli (*UTI89, CFT073, TPA2743 and TPA4792), measuring the averaged OD_{600} over increased concentrations of gentamicin within a serial passage system.

When comparing the serial passage OD results of the clinical strains to that of the log10 CFU/mL calculations, the results are highly comparable (Figure 3.23). Similarly, there is rapid log reduction in growth upon first transfer, this may attributable to bacterial stress responses. The bacterial growth of all strains is then recovered and keep relatively constant, until the rapid reduction in growth of strain CFT073 at 40mg/L. In addition, these results can imply that under gentamicin

selection the four clinical strains of *E. coli* produce similar phenotypic responses, potentially indicating a degree of predictability.

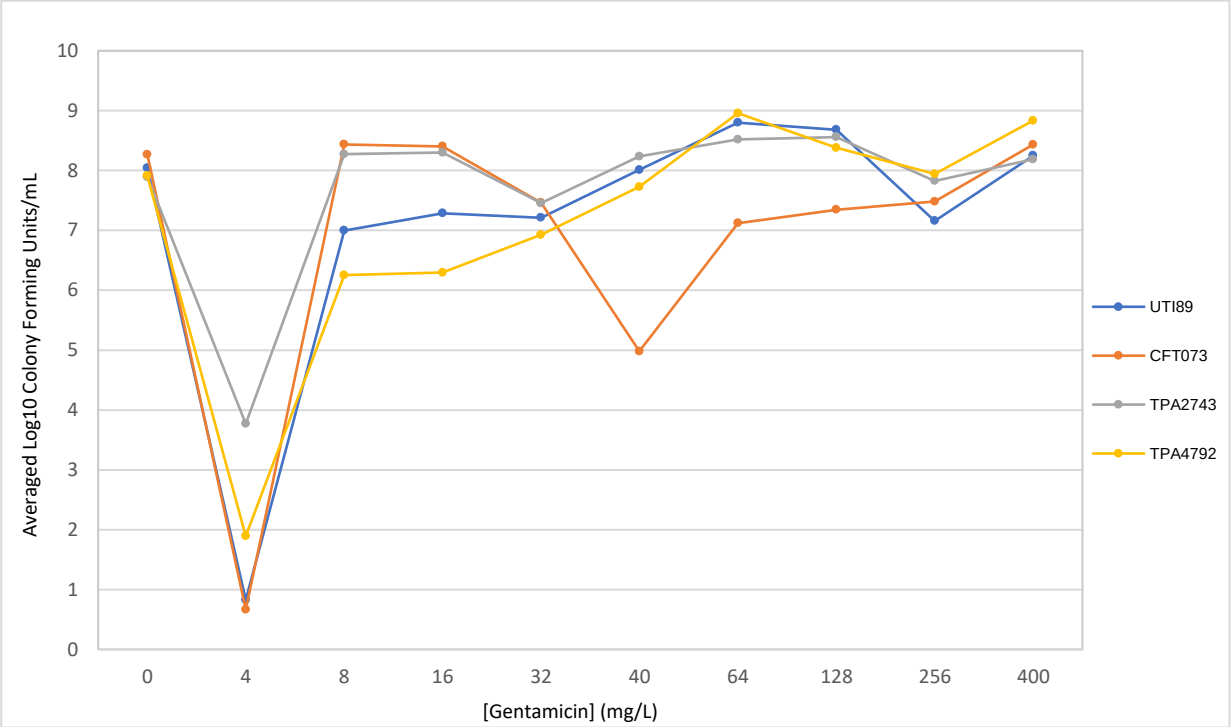


Figure 3.23 Bacterial counts from the serial passage of uropathogenic *E. coli* strains under gentamicin selection. Graph displaying the evolution of gentamicin resistance for four parallel evolving populations of uropathogenic strains of *E. coli* (UTI89, CFT073, TPA2743 and TPA4792) measuring the averaged bacterial colony counts (log10 CFU/mL) over increased concentrations of gentamicin within a serial passage system.

3.4 Discussion

The chapter aims and objectives described in section (3.1) have been met as follows. After extensive provisional method development and optimisation, a finalised tool to explore the evolution of antimicrobial resistance over a spatial temporal antibiotic concentration gradient was established. The application of this experimental evolution tool was directly compared to the traditional experimental evolution approach of serial passage.

3.4.1 Experimental design and optimisation

A major issue experienced with the initial experimental design, was the failed migration of the selected model organism *E. coli.* In regards to bacterial motility, two main processes can occur following the inoculation of bacteria onto agar, one being growth due to nutrient availability and secondly dispersion by cause of chemotaxis, which is dependent on a range of factors (Harshey, 2003). The first of these factors considered, was the inclusion of semi solid media. It is well established that surface motility is dependent on the moisture availability, and that the swarming of bacterial populations is enhanced at agar concentrations of 0.5%-0.7% w/v and compromised at concentrations exceeding 1% w/v (Wolfe and Berg, 1989; Croze *et al.,* 2011). However, despite the alteration in agar consistency there was no significant improvement in the migration of the laboratory strain of *E. coli.*

Ensuing the modifications in agar concentration, nutrient availability within the media was revised through the inclusion of M9 minimal media. The way in which bacterial populations respond to nutrient limitation can fluctuate significantly between

organisms (Martinez, Torello and Kolter, 1999; Toguchi *et al.,* 2000). An initial assumption was made that following utilisation of inadequate nutrients within the zone of inoculation, bacterial populations would migrate towards a nutrient source in response to starvation, as seen with the intracellular signalling and subsequent fruiting body formation in *Myxococcus xanthus* (Shimkets, 1999). However, minimal media failed to support the swarming of *E. coli* within the proposed model (Section 3.3.2.4), and instead swarming appeared to be restricted to the area of inoculation, a behaviour previously described by Harshey and Matsuyama (1994).

Independent of the media condition and the corresponding effects on bacterial surface motility, the force of the antimicrobial selective pressure within the experimental model was considered. Bacterial adaptation can be hindered through the application of strong selective pressures to evolving bacterial populations (Baym *et al.*, 2017). Further to the addition of an intermediate concentration forming the triphasic plate (Section 3.2.2), diphasic antimicrobial slopes were prepared (Section 3.2.2.2), to reduce selective pressures and encourage bacterial adaptation. It was presumed that the gentle diffusion of antimicrobial agents at concentrations ≤MIC along an antimicrobial dose gradient, would replicate the selection of resistant bacteria in natural environments (Gullberg *et al.,* 2011; Harmand *et al.,* 2018). Despite these amendments, the suggested model was unsuccessful in the selection of antibiotic resistant bacteria.

Considering the challenges faced above in the development of a spatial temporal gradient model, the filter tip method was introduced (Section 3.2.3). The

experimental method proposed was subject to scrutiny as the oxygen availability within the system was unknown. *E. coli* is a facultative anaerobe, so although its does not require oxygen for growth, it preferentially utilises oxygen over the reduction of other compounds in its absence (Spiro and Guest, 1991). Oxygen availability has also been closely associated with bacterial motility in *E. coli* (Alder and Templeton, 1967; Douarche *et al.,* 2012). This model effectively facilitated the vertical migration of bacterial populations by chemotaxis through the experimental system (Wang and Ford, 2010). However, concerns over oxygen availability and inconsistencies in the bacterial enumeration from tips presented more questions than answers.

3.4.2 The analysis of bacterial motility

The methods described in (Section 3.2.4) were enhanced through the inclusion of the *E. coli* strain MG1655, frequently employed in experimental evolution under controlled laboratory conditions to address a range of evolutionary questions (Marietou *et al., 2015;* Fong, Joyce and Palsson, 2018; Knöppel *et al.,* 2018). Despite the observed increase in motility of *E. coli* MG1655 when compared to other *E. coli* strains (Wood, 2006), the *E. coli* MG1655 strain failed to migrate across the spatial temporal antibiotic gradient within the proposed model (Section 3.3.5). A key observation following the inclusion of *E. coli* MG1655, was the swarming of bacterial cells within the initial inoculum extending radially within the plate. The outward propelling motion of bacterial cells within the inoculum at the edge of the swarm, appeared to be inhibited at the boundaries of the Triphasic plate (Darnton *et al.,* 2010). Based on these finding the model was altered to a circular format.

The motility of strains was investigated further following the inclusion of uropathogenic motile strains of *E. coli* (UTI89, CFT073, TPA2743 and TPA4792). Bacterial swimming motility was primarily investigated through the preparation of motility agar, by solidifying broth with 0.3-3% w/v bacteriological agar. The motility agar augmented the swimming motility for all *E. coli* strains and was therefore implemented into the ARGP model. The motility of strains was assessed more closely through genomic analysis. Motility in *E. coli* is controlled through a master regulator known as the flhDC operon (Liu and Matsumura, 1994). The motility of *E. coli* strains can vary significantly, depending on the integration of insertion elements within the flhD operon (Blattner *et al.,* 1997). The results from the genomic analysis suggest no insertion elements within the version of *E. coli* MG1655 used within this study. Similar findings were identified by Barker, Pruß and Matsumura (2004), where strains of MG1655 lacking IS elements were poorly motile. The same study suggested that extended incubation of weakly motile strains in motility agar, can reveal motile subpopulations, which should be recommended for future experiments.

3.4.3 Quantification of antimicrobial degradation and diffusion

To validate the ARGP as a tool for exploring the evolution of antibiotic resistance, techniques should be in place to assess the extent of degradation and diffusion of antimicrobial agents (Bonev, Hooper and Parisot, 2008). This is essential as incubation conditions affect the stability of antimicrobial agents during *in vitro* Microbiological testing (Marchbanks, Yost and White, 1987). In addition, the decrease in antimicrobial activity due to degradation can compromise antimicrobial susceptibility results (Lallemand *et al.*, 2016). The concentration gradient present within ARGP, facilitates the diffusion of antimicrobial agents through the agar matrix

(Koch, 1999). The diffusion kinetics of antimicrobial agents within agar matrices is dependent on a range of factors including agar composition and the antimicrobial agent of choice (Kronvall,1983).

The quantification of aminoglycosides by analytical techniques is difficult, owing to their highly polar nature and absence of strong chromophores for chromatographic detection (Lecároz *et al.,* 2006). This was demonstrated with the complications faced in the detection of gentamicin within an agar matrix using NMR. As gentamicin, had been previously quantified in Mueller-Hinton agar by HPLC (Arcelloni *et al.,* 2001), it was decided that liquid chromatography would be combined with mass spectroscopy to increase the sensitivity of the analytical approach (McGlinchey *et al.,* 2008). The detection limits within LC-MS, allowed the detection of gentamicin at all selected concentrations. However, for the accurate quantification of gentamicin, the extraction procedure must be optimised to account for antibiotic loss during extraction which can be a very time-consuming process.

3.4.4 The traditional serial passage approach

Traditional experimental evolution approaches were based on serial passage in well mixed systems (Lenski *et al.,* 1991). Serial passage is well-known for producing population bottlenecks, through the intermittent reduction in bacterial population size during transfer (Wahl, Gerrish and Voivod, 2002). This approach has been criticised for solely permitting the identification of single step resistant mutants, which can survive a given antimicrobial selective pressure (Palmer and Kishony, 2013). This approach has been included in this study for direct comparison with the ARGP, which facilitates the multi-step adaptive evolution of resistance over a spatial concentration

gradient of antibiotics.

The comparative findings, support the accelerated resistance evolution within spatial concentration gradients (Zhang *et al.,* 2011; Baym *et al.,* 2016b). Nevertheless, the serial passage results presented some interesting findings. There were significant fluctuations in the growth of bacterial populations following transfer into increasing concentrations of antibiotics, where periods of drastically reduced growth followed rapid resistance development (section 3.3.9). The phenotypic variants detected which displayed reduced growth and the ability to survive antimicrobial treatment, have been previously defined bacterial persisters (Orman and Brynildsen, 2016). Similar findings were observed by Levin-Reisman (2017), which suggested the incessant survival of sub-populations of bacterial persisters, permitted the establishment of highly resistance bacterial populations. These findings signify the importance of antimicrobial tolerance in resistance evolution.

3.4.5 Conclusion

To conclude, after extensive modification the ARGP facilitates the visualisation of resistance evolution over a spatial gradient of antibiotics, which can be compared to traditional established evolutionary methods. However, the potential applications of this model will only be uncovered following the genomic analysis of the resistant evolved strains.

Chapter 4: Understanding the molecular mechanisms behind the evolution of aminoglycoside resistance in *E. coli* MG1655.

4.1 Introduction

From the development of sequencing methods in 1977 (Sanger, Nicklen and Coulson,1977; Maxam and Gilbert, 1977), to the first sequence of the human genome in 2001 (Venter *et al.,* 2001), whole genome sequencing (WGS) technologies have revolutionised the field of genomics (van Dijk *et al.,* 2014). With the increasing popularity of WGS, advancements have seen significant reductions in costs and WGS at unprecedented speeds with second and third generation sequencing technologies (Ambardar *et al.,* 2016; Pareek, Smoczynski and Tretyn, 2011). The application of WGS technologies within the field of microbiology, has completely transformed the clinical diagnostic paradigm through improved clinical detection, drug susceptibility testing and epidemiological typing (Köser *et al.,* 2012; Reuter *et al.,* 2018; Fournier, Dubourg and Raoult, 2014).

In addition, NGS technologies have facilitated the study of whole genome sequence variation, to provide insights into the adaptive evolution of bacterial populations (Bryant, Chewapreecha and Bentley, 2012). Such approaches have been utilised in the investigation of antimicrobial resistance, through the identification of resistance genes (Yoshida *et al.,* 1990; Wang *et al.,* 2016) and from a surveillance perspective to predict resistance using gene expression profiles (Bradley *et al.*, 2015; Suzuki, Horinouchi and Furusawa, 2014).

Further to the horizontal acquisition of resistant genes, bacteria can evolve resistance through the acquisition mutations (Baquero and Blázquez, 1997; Woodford and Ellington, 2007). This is particularly pertinent in a time where sustained antimicrobial selective pressures are being placed upon bacterial populations (Alonso, Campanario E and Martínez JL, 1999). NGS permits the rapid detection of single nucleotide variants within bacterial populations, and through comparative genomics insights can be provided into molecular mechanisms underlying resistance development (Ramanathan *et al.,* 2017).

It is hypothesised that through comparative analysis SNPs will be identified which are unique to the gentamicin resistance isolate. It is also predicted that the likely locations of these SNPs will be within or associated with the antimicrobial target. The aim of this chapter is to genome sequence the sensitive and resistant isolates and through comparative genomics identify SNPs associated with the acquisition of gentamicin resistance. If gentamicin associated SNPs are identified it is likely these SNPs could be applied as markers for aminoglycoside resistance.

4.2 Materials and methods

4.2.1 Whole genome sequencing at Microbes NG

Genotypic analysis of both the sensitive and confirmed resistant bacterial isolates obtained from the ARGP, was achieved through WGS provided by Microbes NG. Here, sensitivity is defined as the inhibition of bacterial growth by a given concentration of antimicrobial agent. Details of the sequencing protocol and the principles behind illumina sequencing can be found in section (2.8.1).

4.2.2 Whole genome assembly CLC

The bacterial genomes were assembled in CLC genomics workbench 7.7.5. A full description of the assembly protocol is provided in section (2.8.2). The newly assembled genomes were examined and genes interest were retrieved, these included known antimicrobial drug targets as well as genes associated with antimicrobial resistance. The genes of interest for both the sensitive and resistant genomes were exported and analysed using the alignment viewer aliview.

4.2.3 Identification of genes of interest

A preliminary search of the gentamicin antimicrobial drug targets was carried out using the bioinformatic and cheminformatic database DrugBank version 5.1.1 (Wishart *et al.,* 2018), and information regarding the pharmacodynamics and mechanism of action was obtained. Following this a more detailed search of the identified target genes was achieved using the database website EcoGene http://ecogene.org [accessed October 13, 2017] which dedicates itself to the improved functional annotation of the *E. coli* K-12 MG1655 genome and subsequently provided genomic and proteomic details (Zhou and Rudd, 2012). Using the proteomic

information obtained from EcoGene a final analysis was carried out using the online database STRING version 10.5 http://string-db.org [accessed July 30, 2018] to identify proteins forming functional interactions with the antimicrobial targets (Mering *et al.,* 2003). This was achieved simply by supplying the protein name and the organism of interest.

4.2.4 Genomic alignment using Mauve

As previously described (Section 4.2.2) CLC workbench was used in the alignment and assembly of bacterial genomes to identify small-scale evolutionary changes including single nucleotide polymorphisms within known antimicrobial targets. In addition to these local changes, evolving bacterial populations can undergo more large-scale genome wide rearrangements, which were investigated using Mauve available at: http://gel.ahabs.wisc.edu/mauve [accessed August 18, 2018]. Mauve as a comparative genomics platform facilitates the investigation of all scales of evolutionary change which occur within bacterial genomes, through integrating large-scale genomic analysis with multiple sequence alignment using locally collinear blocks (LCBs) (Darling *et al.,* 2004). Details of the Mauve protocol utilised within this study can be found in section (2.8.3).

4.2.4.1 SNP Identification

The SNP export file produced by Mauve contained a list of node ID's, SNP positions and the corresponding coding sequence (CDS) positions for both the sensitive and resistant genomes. Prior to analysis of SNPs, it was necessary to account for the

discrepancies in the annotations of the two genomes specifically the SNP position in relation to the designated scaffold or node ID.

To obtain the nodal information and to ensure that the SNPs observed were located within coding regions of the genomes a python script named get_seq_go was used (Eldridge C, 2017). To execute the script, a text file was made containing a list of the node IDs for the two genomes and it is important to note that if duplicate IDs are present they must be formatted underneath each other within the file. An example of the format used in the node txt. file is presented below (Figure 4.1).

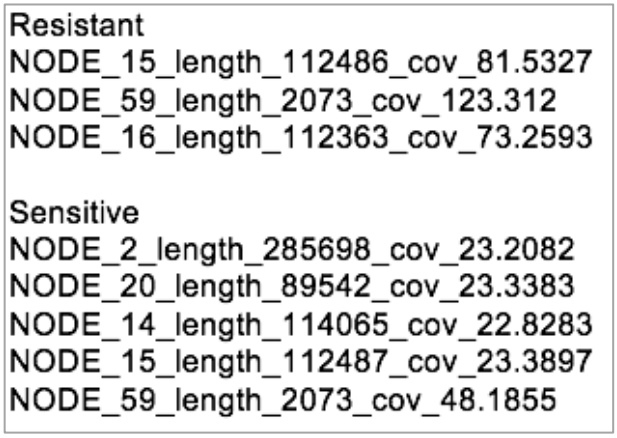


Figure 4.1 The node text file format required for the get_seq_go script. An example of the format required for the execution of the get_seq_go script to obtain the nodal information from the two bacterial genomes. Information regarding the node number, length and coverage must be provided and node IDs must be formatted beneath each other.

Subsequently, all script files must be located and placed within the same folder as the newly created node txt. file or alternatively a path must be written. Once terminal is opened the written command must be replicated per node for all genes indicating the presence of a SNP and their sensitive or resistant counterparts for comparison. Below is an example of the command used for the SNP indicated within node 38 of resistant genome (Figure 4.2).

```
exonerate --model protein2genome –query
Desktop/NODE_14_SEN_clsC/clsC_ProSeq_NODE14_SEN.fasta --target
Desktop/NODE_14_SEN_clsC/NODE_14_SEN.fasta --ryo %qi(%qab
- %qae)\n%qas\n >%ti(%tab - %tae)\n%tas --showtargetgff no --showalignment no --
showvulgar no --geneticcode 11
```

Figure 4.2 The node command executed when running the get_seq_go script. An example of the command utilised when running the get_seq_go script for obtaining node 38 from the gentamicin resistant genome. The command must contain the relevant folder or path, the python script and the .txt file created in figure 4.1. The command was repeated for all nodes of interest indicated with SNPs for both the gentamicin sensitive and resistant *E. coli* MG1655 genomes.

As previously stated, once the nodal information had been obtained using get_seq_go, it was necessary to retrieve the nucleotide sequences of the genes with SNPs from the corresponding nodes. This was achieved through use of the Genbank files to locate the Uniprot IDs (https://www.uniprot.org) and consequently retrieve the protein sequences of the SNP containing genes. The aligned protein sequences were then saved individually in fasta format. The sequence alignment tool Exonerate version 2.2 was downloaded from EMBL-EBI https://www.ebi.ac.uk/about/vertebrate-genomics/software/exonerate [accessed October 30, 2018], to facilitate the mapping of protein coding genes within the genomic sequences contained within each node. To obtain the nucleotide sequences using the exonerate pairwise comparison model 'protein to genome' both the query protein sequence and the target node sequence are required. Below is an example of the command used to obtain the nucleotide sequence of the protein encoding clsC gene which was located within node 14 (Figure 4.3).

```
exonerate --model protein2genome –query
Desktop/NODE_14_SEN_clsC/clsC_ProSeq_NODE14_SEN.fasta --target
Desktop/NODE_14_SEN_clsC/NODE_14_SEN.fasta --ryo %qi(%qab
- %qae)\n%qas\n >%ti(%tab - %tae)\n%tas --showtargetgff no --showalignment no --
showvulgar no --geneticcode 11
```

Figure 4.3 The Exonerate command utilised to obtain the nucleotide sequences of genes within the nodes of interest. An example of Exonerate command utilised to obtain the clsC gene sequence from node 14 of the gentamicin sensitive *E. coli* MG1655 genome. The command must contain the relevant folder or path, the nodal information obtain in figure 4.2, the clsC encoding protein sequence in. fasta format and the bacterial genetic code table 11 must be specified. The command was repeated for all genes with SNPs contained within the formerly obtained nodes.

The final stage in the SNP identification process was to translate the retrieved nucleotide sequences from exonerate, to determine whether the single nucleotide changes observed result in alterations in amino acids. This was conducted within terminal using Transeq using all 6 open reading frames and the bacterial code 11. Below is an example of the command used to translate the nucleotide sequence of the clsC gene (Figure 4.4). The frame was the then selected with the longest open frame and the fewest stop codons and saved in fasta format for comparative analysis in an alignment viewer.

```
transeq -sequence clsC_SEN.fasta -frame 6 -table 11 -outseq clsC_SEN.frames
```

Figure 4.4 The Transeq command utilised to translate the nucleotide sequences. An example of the command used to translate the clsC nucleotide sequence obtained from figure 4.3 to identify whether the SNPs were synonymous or non-synonymous. The command must contain the nucleotide sequence in. fasta format and the maximal number of reading frames (6) and the bacterial translation table (11) must be specified. The command was repeated for all gene sequences identified with SNPs.

4.2.5 Proteome analysis using CD-HIT

All protein annotations from the genbank files were extracted into one proteome file and subject to clustering, using a custom script available at: https://github.com/camilla-eldridge [accessed March 26, 2018]. CD-HIT was used to cluster and compare the sequences from the sensitive and resistant datasets, to reduce redundancy and improve subsequent bioinformatic analysis (Huang *et al.,* 2010). The programme cd-hit-2d which forms part of the CD-HIT package, allows the direct comparison of the two protein sequences facilitating the identification of novel proteins and thus allowing the elimination of contaminated proteins early in the bioinformatic analysis of sequences. Prior to running cd-hit-2d it was necessary to add unique headers and combine the sensitive and resistant proteomes for analysis. The executed commands used are shown below (Figure 4.5).

```
sed -i -e 's/>/>sens_/g' Sensitive_proteome. fasta
sed -i -e 's/>/>res_/g' Resistant_proteome. fasta
grep "sens" combined_sens_res_proteomes. fasta
grep ">res" combined_sens_res_proteomes. fasta
cat Resistant_proteome. fasta Sensitive_proteome. fasta > combined_sens_res_proteomes.fasta
```

Figure 4.5 The command utilised to prepare proteome files for the CD-HIT analysis. The command utilised to add the appropriate headers to the gentamicin sensitive and resistant *E. coli* MG1655 proteome files, before combining the proteomes to a single file for the analysis within CD-HIT.

Once the cd-hit directory was launched the preliminary step involves the identification of novel resistant proteins which are not present in the sensitive proteomes (1) and to identify the novel sensitive proteins which are not found in the resistant proteomes (2)

for the initial analysis clustering was achieved using default parameters with a sequence identity threshold (–c) of 1 for 100% ID (Figure 4.6).

```
1. ./cd-hit-2d -i2 Resistant_proteome.fasta -i Sensitive_proteome.fasta -o Resistant_novel -c 1
2. ./cd-hit-2d -i Resistant_proteome.fasta -i2 Sensitive_proteome.fasta -o Sensitive_novel -c 1
```

Figure 4.6 CD-HIT command for the identification of novel proteins. The command utilised to identify the novel sensitive proteins which are not found within the resistant proteome (1) and the novel resistant proteins not found in the sensitive proteome (2) using cd-hit-2d with a sequence identity threshold (–c) of 1 indicative of 100% ID match.

The outputs from the above commands were then used to detect the sensitive proteins which have not been identified in the resistant novel proteins (3) and secondly to identify the resistant proteins which are not found in the unique sensitive proteins (4) (Figure 4.7). This was achieved by clustering the outputs at 95% (-c = 0.95), to enable the identification of nucleotide polymorphisms and to assist in the removal of potential contaminants.

```
3. ./cd-hit-2d -i2 Sensitive_Novel -i Resistant_Novel -o sens_vs_res -c 0.95
4. ./cd-hit-2d -i2 Resistant_Novel -i Sensitive_Novel -o sens_vs_res_swap -c 0.95
```

Figure 4.7 CD-HIT command for the identification of novel proteins with SNPs. The command utilised to identify the novel sensitive proteins which are not found within the resistant proteome (3) and the novel resistant proteins not found in the sensitive proteome (4) using cd-hit-2d with a sequence identity threshold (–c) of 0.95 indicative of 95% ID match.

The final stage in the proteome analysis using CD-HIT was to upload the identified novel protein variants from the sensitive and resistant datasets into NCBI: Protein BLAST and perform a standard blast search of the non-redundant protein sequence database. For the identification of protein contaminants, the organism was not specified.

4.3 Experimental results

4.3.1 Whole genome sequencing

WGS data obtained from Microbes NG revealed *E. coli* MG1655 resistant strain had a greater number of paired reads both before and after trimming within CLC, when compared to the *E. coli* MG1655 sensitive counterpart (Table 4.1). Despite the variations in coverage of the two sequenced genomes, strains were trimmed by 90.54% and 89.47%, resulting in similar averaged lengths of 168.5 and 166.6 for the sensitive and resistant strains respectively (Figure 4.8).

Table 4.1 Paired sequence read summary of the sensitive and resistant *E. coli* MG1655 genomes. Paired sequence read comparison for *E. coli* MG1655 sensitive and resistant genomes before and after trimming.

***E. coli* MG1655**	**Sensitive**		**Resistant**	
	Before Trim	**After Trim**	**Before Trim**	**After Trim**
Number of Reads	1,460,808	1,322,568	5,345,144	4,782,297

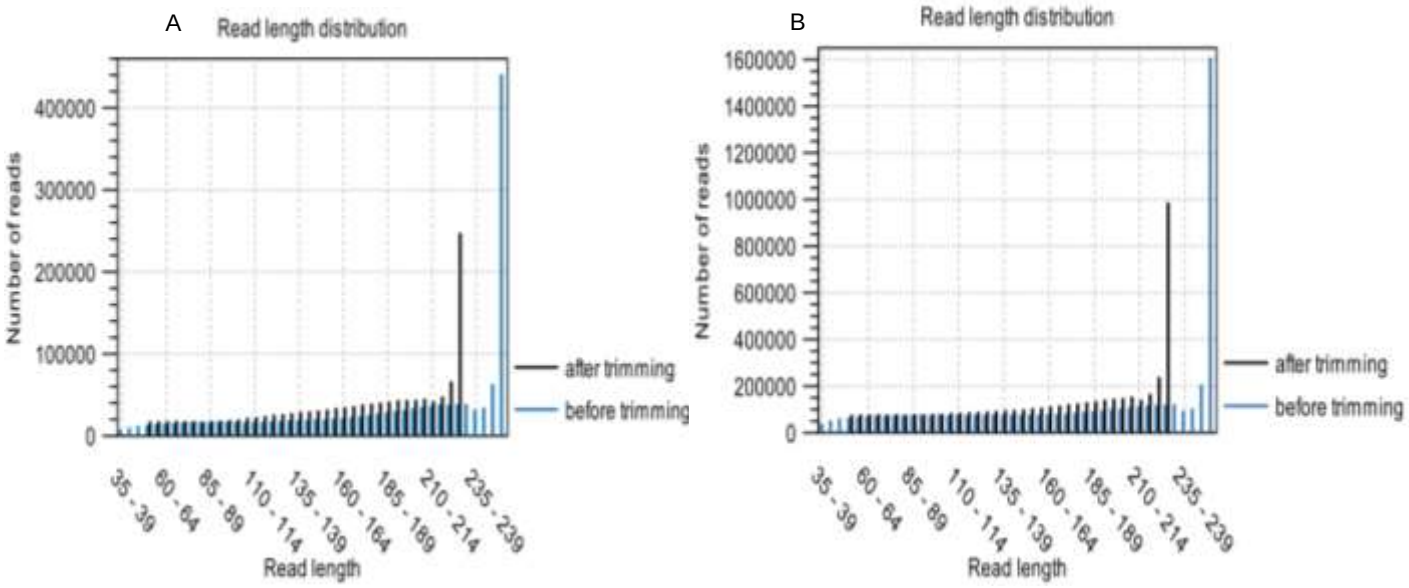


Figure 4.8 Sequence read analysis of the sensitive and resistant *E. coli* MG1655 genomes. Graphical summary of *E. coli* MG1655 sensitive (A) and resistant (B) genome read length before and after trimming.

4.3.2 CLC Sequence quality and trimming

The sequences were then evaluated further by a range of analyses examining the genomes per sequence, per base and by assessing over representation. These analyses were conducted to ultimately improve sequence quality and will be discussed sequentially.

4.3.2.1 Per sequence- length distribution analysis

The length distribution analysis summarises the frequencies of observed sequence lengths, by considering the number of sequences occurring at the specific sequence length and comparing this to the total number of sequences to obtain a percentage. The results show for the two strains of *E. coli* MG1655, similar sequence lengths discussed above (Section 4.3.1), with sequences covering 38-254 base pairs before

being trimmed. Following trimming, no sequences were preserved below 50 and above 229 base pairs for the two genomes respectively. Further, the distribution of most of the sequences for the two genomes post trimming, were above 204 base pairs (Figure 4.9).

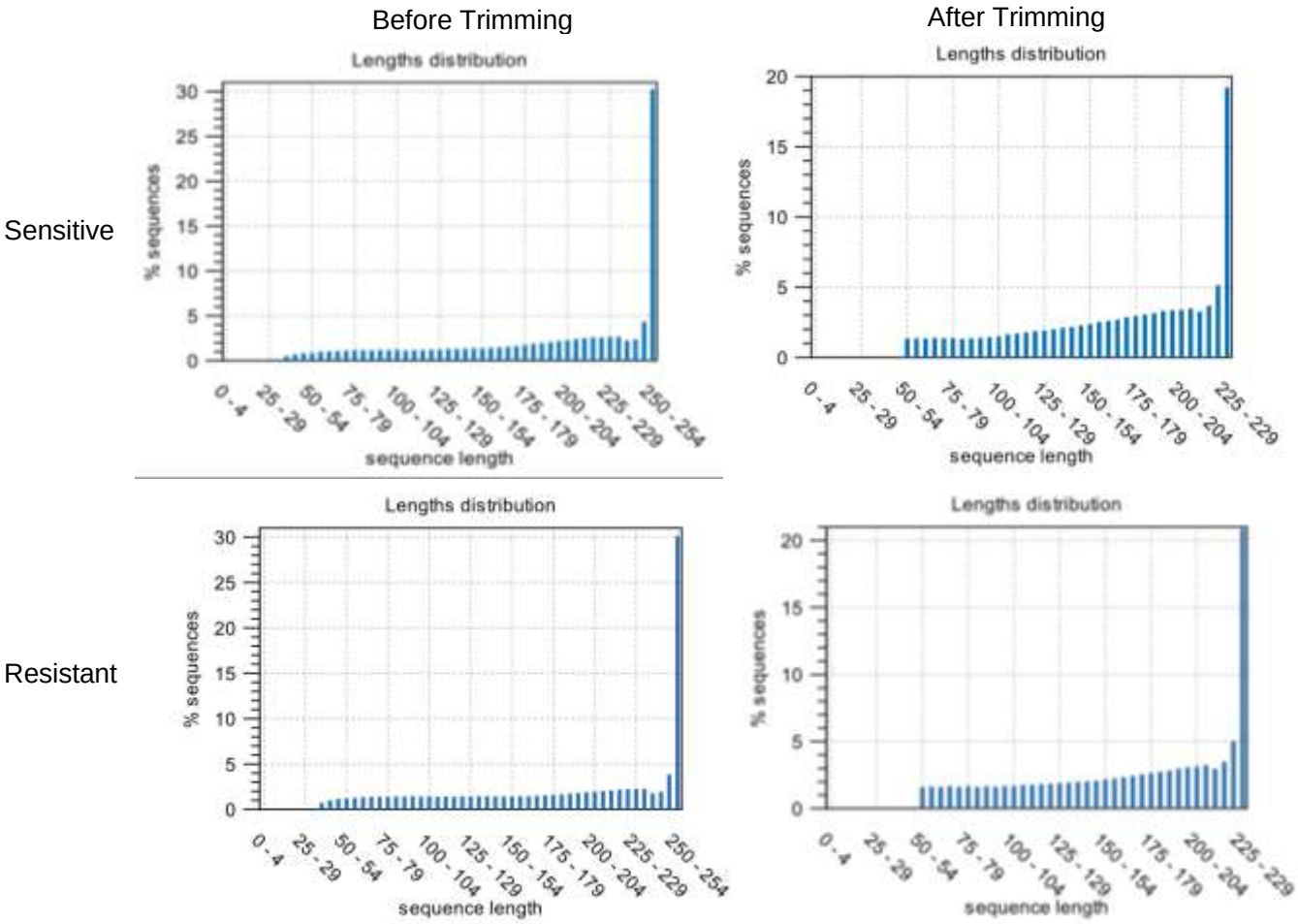


Figure 4.9 Length distribution analysis of the sensitive and resistant *E. coli* MG1655 genomes. The distribution of sequence length of the sensitive and resistant strains of *E. coli* MG1655 produced by illumina sequencing, before and after trimming.

4.3.2.2 Per sequence- GC content analysis

The GC-content analysis summarises the distribution of GC bases compared to all other bases within the genomes, including those that are considered ambiguous. The results show similar GC-content for the two genomes before and after trimming (Figure 4.10). The GC contents of the two sequences ranged from approximately 25% to 70%, with GC-contents peaking at 50%. This value correlates with the accepted GC content of 50.8% for *E. coli* K12.

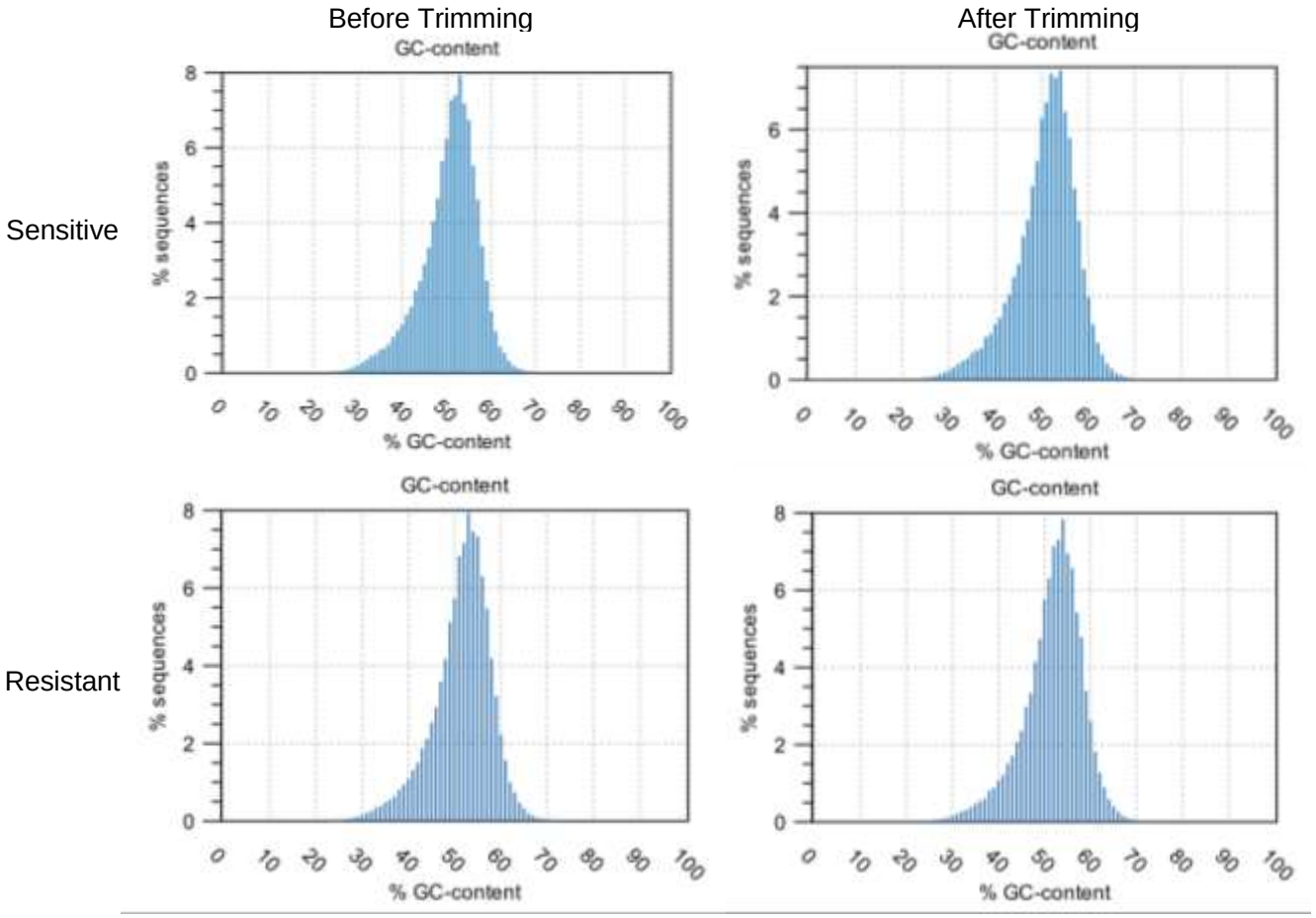


Figure 4.10 GC content analysis of the sensitive and resistant *E. coli* MG1655 genomes. The GC content length for the sensitive and resistant strains of *E. coli* MG1655 produced by illumina sequencing, before and after trimming.

4.3.2.3 Per sequence- quality distribution analysis

The quality distribution analysis summarises the distribution of averaged sequence quality scores using the arithmetic mean of its base qualities. Prior to trimming the phred score for the sensitive and resistant genomes ranged from 20 to 40 and 23 to 40 respectively (Figure 4.11). Post trimming, the sensitive genome had quality scores between 27 to 40, with 87% of sequences scoring between 35 and 40. When comparing this to the resistant genome after trimming had quality scores ranging between 28 to 40, with 91% of sequences scoring between 35 and 40.

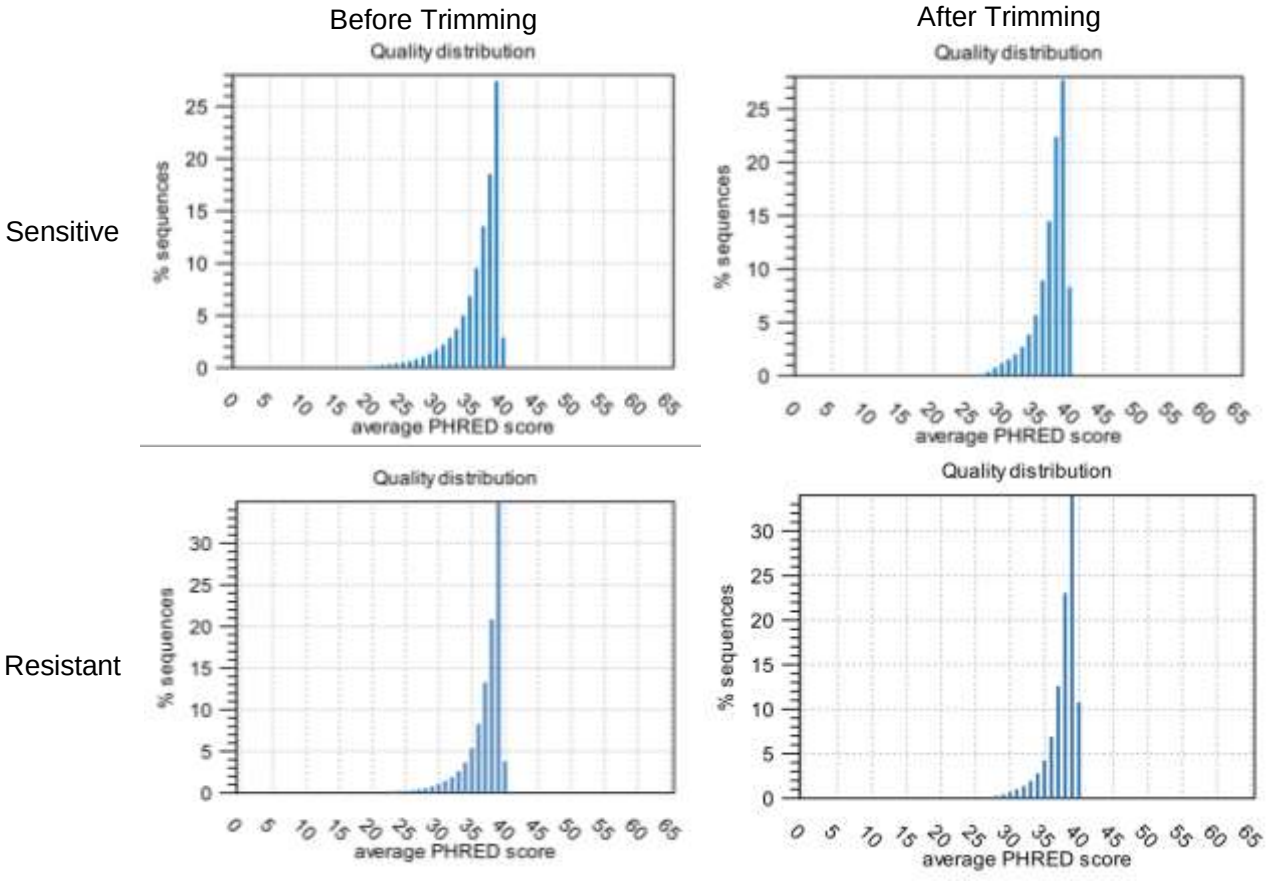


Figure 4.11 Quality distribution analysis of the sensitive and resistant *E. coli* MG1655 genomes. The averaged PHRED score for the sensitive and resistant strains of *E. coli* MG1655 produced by illumina sequencing, before and after trimming.

4.3.2.4 Per base- coverage analysis

The per base coverage analysis summarises the number of sequences which cover the individual base positions within the genomes. For the two genomes, the base coverage is similar pre- and post-trimming with a higher percentage coverage for base positions lower than 100 (Figure 4.12).

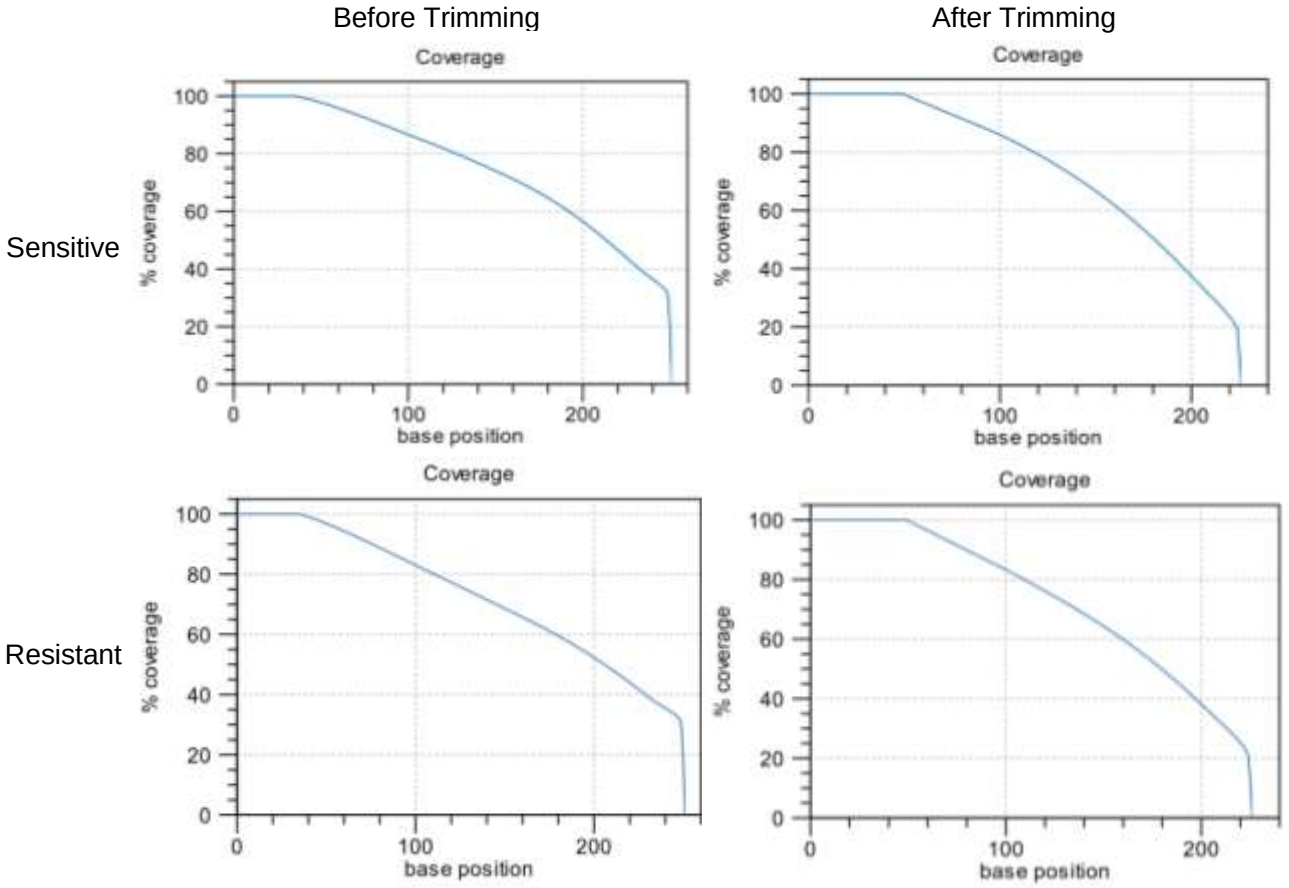


Figure 4.12 Coverage analysis of the sensitive and resistant *E. coli* MG1655 genomes. The percentage coverage per base position for the sensitive and resistant strains of *E. coli* MG1655 produced by illumina sequencing, before and after trimming.

4.3.2.5 Per base- nucleotide contribution analysis

The nucleotide contribution analysis summarises the coverage of all four DNA nucleotides and ambiguous bases at the individual base positions across the genomes. The nucleotide contributions prior to trimming showed great fluctuations ranging from approximately 12-40% for all nucleotides for the first 20 bases, for both the sensitive and resistant genomes (Figure 4.13). Fluctuations were also present at a later base position 240, with a peak in the contribution of the nucleotide adenine and a reduction in nucleotide cytosine within both genomes.

Following trimming, nucleotide contributions for both sequence had levelled however, there were slight differences observed between the two genomes. Within the sensitive genome nucleotide contributions were even at 25% across all base positions. Whereas, within the resistant genome the nucleotide contributions were even at 26% for nucleotides Guanine and cytosine, and even at a slightly lower 24% for thymine and adenine (Figure 4.13).

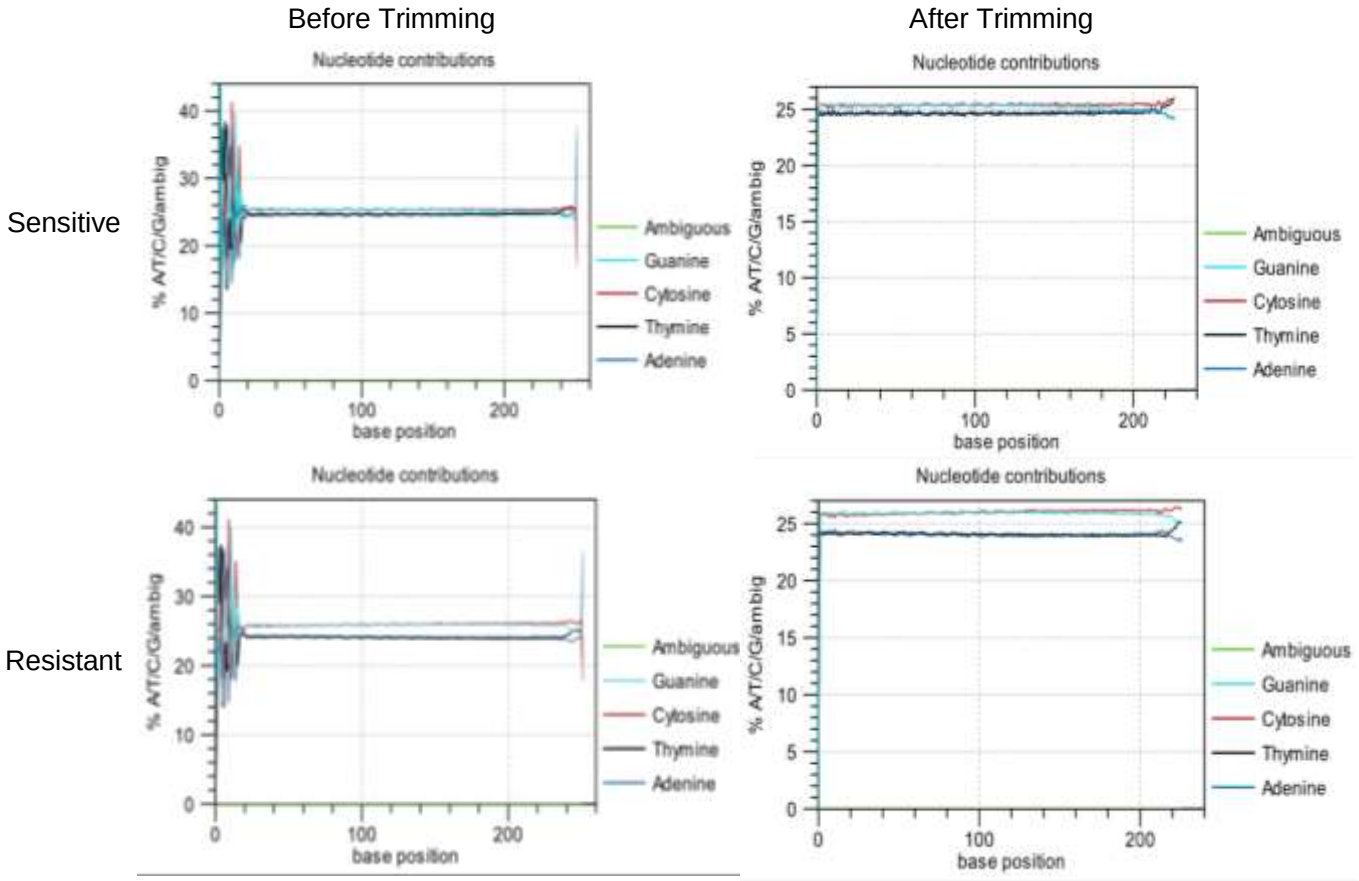


Figure 4.13 Nucleotide contribution analysis of the sensitive and resistant *E. coli* MG1655 genomes. The nucleotide contributions per base position for the sensitive and resistant strains of *E. coli* MG1655 produced by illumina sequencing, before and after trimming.

4.3.2.6 Per base- GC content analysis

The GC content analysis summarises the cumulative coverage of G and C bases, but in this instance per base across the two genomes. As seen above with the nucleotide contribution, prior to trimming fluctuations are present between the two sequences ranging from approximately 30-70% GC contents for the first 20 base positions (Figure 4.14). Following trimming, the GC contents had levelled between 50 and 51% over the 240 base positions for the two genomes respectively. This value again correlates with the accepted 50.8% GC content for *E. coli* K12 (Figure 4.14).

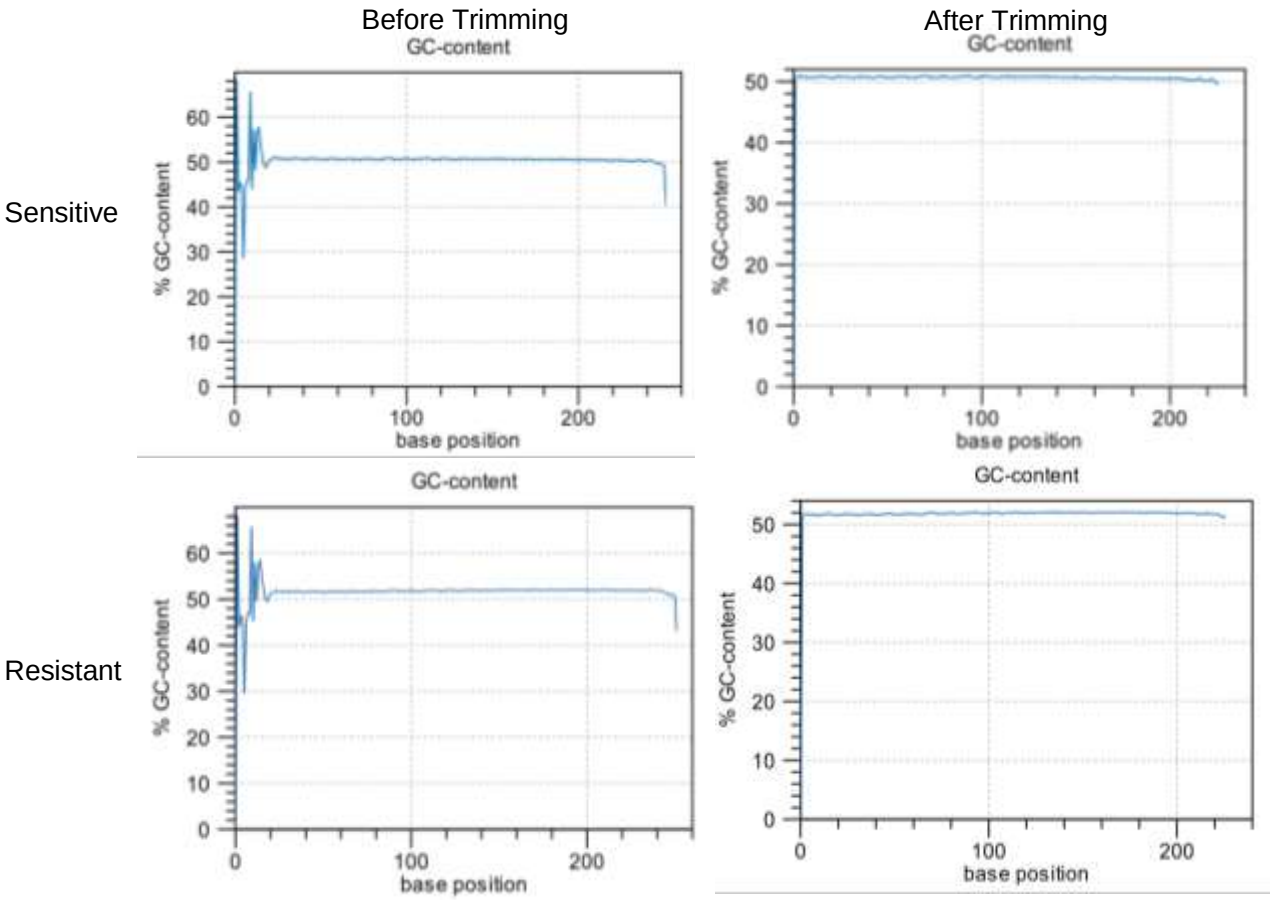


Figure 4.14 GC content per base analysis of the sensitive and resistant *E. coli* MG1655 genomes. The GC content per base position for the sensitive and resistant strains of *E. coli* MG1655 produced by illumina sequencing, before and after trimming.

4.3.2.7 Per base- quality distribution analysis

The quality distribution analysis summarises base quality along the base positions of the two genomes. The phred score quality distributions were very similar for the two genomes, with values ranging from 26-40 and 27-40 for the first 100 base positions for the sensitive and resistant genomes respectively, prior to trimming. Similarly, at base positions exceeding 100, there was a clear reduction in quality distribution for the 5%ile, with phred score values as low as 12 for the two genomes before trimming (Figure 4.15). Following trimming, the overall quality distribution was enhanced with

phred score values of 35-40. Improvements were also evident within the 5%ile where phred scores increased to >20 across all base positions for the two genomes (Figure 4.15).

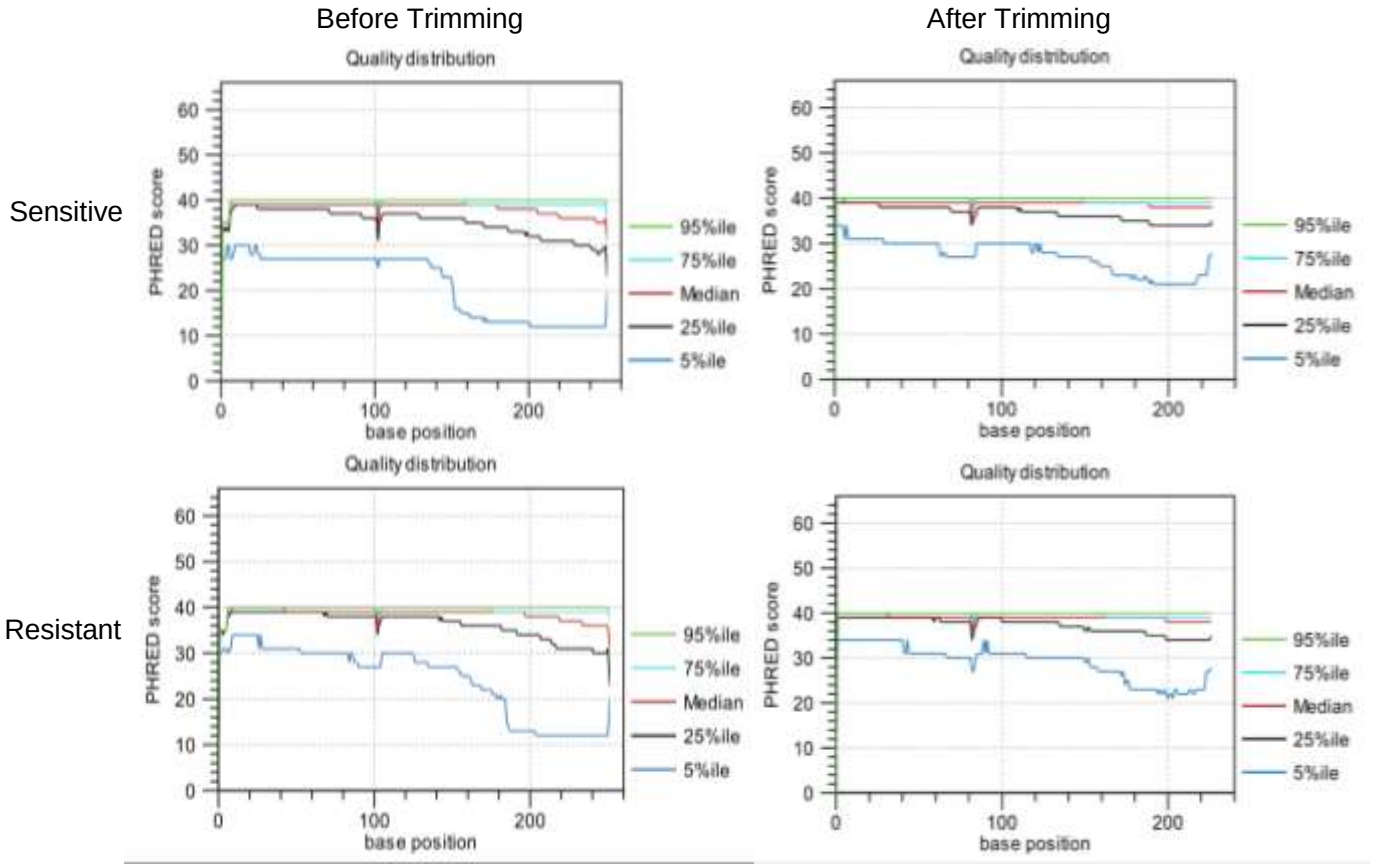


Figure 4.15 Quality distribution per base analysis of the sensitive and resistant *E. coli* MG1655 genomes. The quality distribution analysis per base position for the sensitive and resistant strains of *E. coli* MG1655 produced by illumina sequencing, before and after trimming.

4.3.2.8 Over-representation enriched 5mers analysis

The Enriched 5mers Analysis summarises the ratios of the observed and expected 5mers along base positions across the two genomes. Before trimming, there was substantial fluctuation in the contributions of all 5mers with coverages ranging from 0-0.55 and 0-0.57 for the first 20 base positions for the sensitive and resistant genomes respectively (Figure 4.16). At base positions between 20 to 240, 5mers were equably represented with coverages ranging between 0.2 and 0.3%, with a slight under representation of AAAAA for the two genomes. At base positions exceeding 240, there was an over representation of AAAAA and an under representation of TGGCG for the two genomes.

Following trimming, the overall representation of 5mers had been reduced to between 0.2-0.3% for the two genomes across all base positions. Consistent to the findings of the nucleotide distribution analysis, within the resistant genome there is a distinct separation, in this case of 5mers TTTTT and AAAAA with coverages ranging between 0.22-0.25% and 5mers GCTGG and CCAGC with a coverage of 0.3% (Figure 4.16).

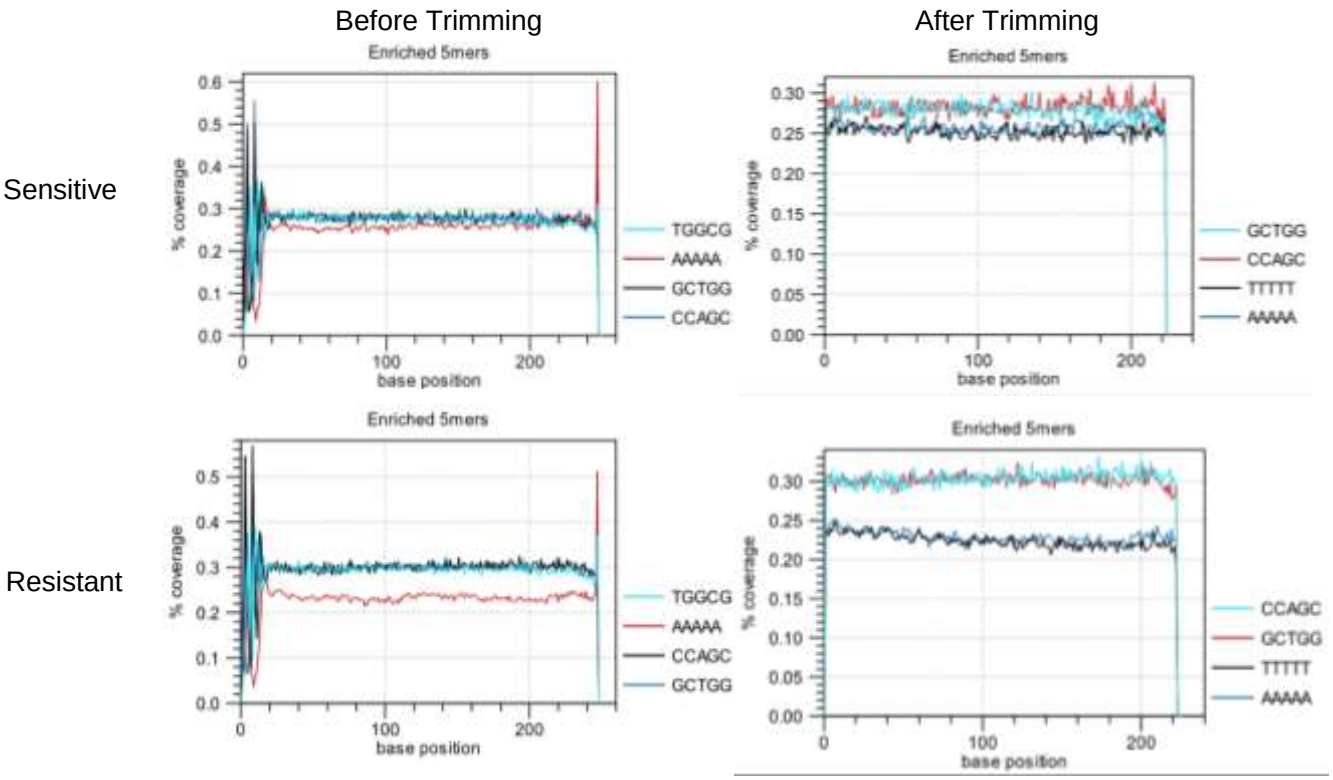


Figure 4.16 Over-representation analysis of the sensitive and resistant *E. coli* MG1655 genomes. The enriched 5mers analysis for the sensitive and resistant strains of *E. coli* MG1655 produced by illumina sequencing, before and after trimming.

4.3.3 Identification of genes associated with aminoglycoside resistance

Having assembled the two bacterial genomes to a high quality. CLC was used to perform a preliminary search of the antimicrobial drug targets. The aminoglycoside gentamicin provided the antimicrobial selective pressure in the sequenced resistant strain of *E. coli* MG1655. Aminoglycosides exert their activity, by binding to specific regions of the 70s-bacterial ribosome, principally through precise interactions with the 16s rRNA at the decoding region of the A-site tRNA, which interrupts ribosomal protein synthesis (Wong *et al.,* 1998). Further structural studies, have revealed gentamicin forms precise interactions within the binding pocket of the major grove of the A-site tRNA and through lateral proximity with the ribosomal subunit protein S12

(Yoshizawa *et al.,* 1988 and Demirci *et al.,* 2013). Supported by the gentamicin drug target information obtained, an initial CLC search was performed on the 16s rRNA genes, of which there are seven variants in *E. coli* and the 30s-ribosomal subunit protein S12, encoded by the *rpsL* gene.

The preliminary CLC search revealed no mutations within the *rpsL* gene, however mutations were found in two of the seven 16s rRNA genes; rna123 and rna154. The precise locations of the single nucleotide changes within the 16s rRNA genes are presented as alignment images obtained from Aliview. There were five mutations identified within the rna123 gene, ranging from base positions 208 to 273 (Figure 4.17). When comparing this to the mutations found within the rna154 gene, there were a total of six mutations, but in two separate locations. Five of the six rna154 mutations were identified at base positions between 79 to 93, and one mutation was found at position 226 (Figure 4.18). Interestingly, the mutation at position 226 was identified in both the rna123 and rna154 genes independently.

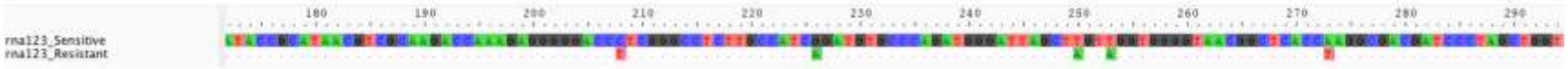


Figure 4.17 The ribosomal rna123 gene alignment. Aliview alignment image displaying the SNPs within the ribonucleotide sequences of the rna123 gene between the sensitive and resistant strain of *E. coli* MG1655. The dots in the alignment are representative of identical bases between the two bacterial genomes.

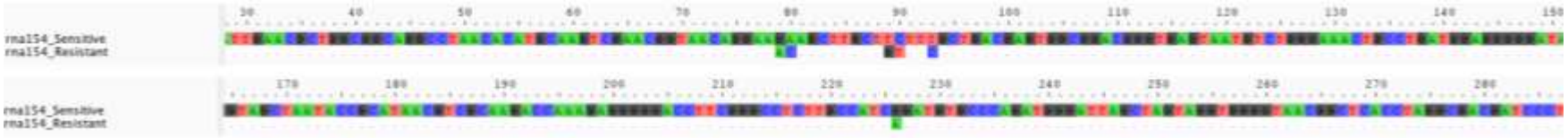


Figure 4.18 The ribosomal rna154 gene alignment. Aliview alignment image displaying the SNPs within the ribonucleotide sequences of the rna154 gene at the two distinct base regions between the sensitive and resistant strain of *E. coli* MG1655. The dots in the alignment are representative of identical bases between the two bacterial genomes.

To broaden the search further, and to identify other genes potentially associated with the acquisition of gentamicin resistance, it was necessary to understand the highly coordinated 30s-ribosomal subunit assembly in greater detail. Particularly, with regards to the interactions of the 16s rRNA with the twenty 30s-ribosomal subunit proteins (Figure 4.19). The assembly can be split into three based on the proteins which associate with the 16s rRNA in the early, mid or late stages of assembly (Holmes and Culver, 2005). Highlighted in green is the gentamicin 30s-ribosomal subunit target protein S12, which associates early in the 30s-ribosomal subunit assembly with the16s rRNA in the body formation. The interaction analysis of S12, revealed key associations with the primary binding proteins S5 and S17. Based on this, both the *rpsE* and the *rpsQ* genes which encode the S5 and S17 protein respectively, were searched within CLC however, no mutations were identified.

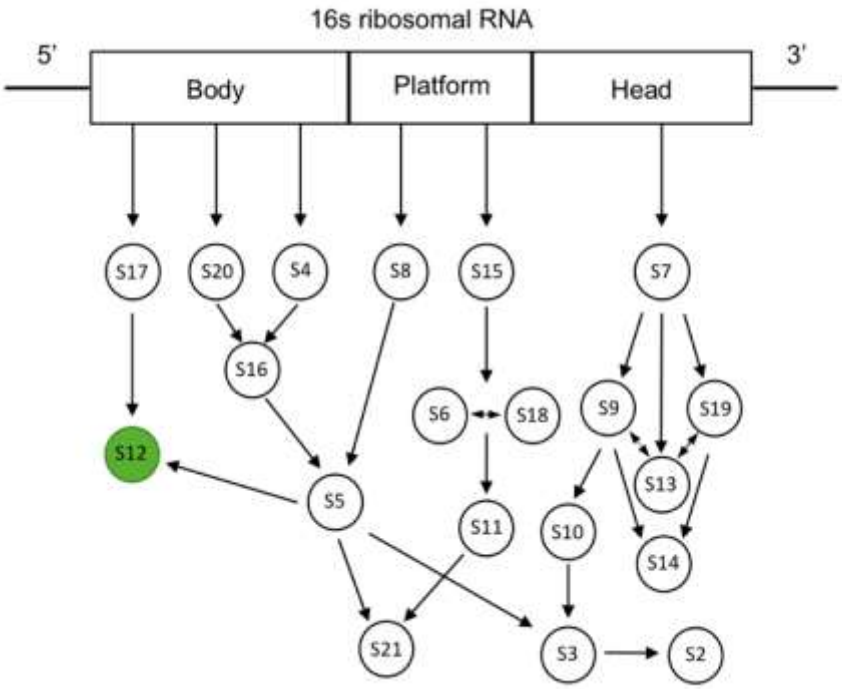


Figure 4.19 Schematic representation of the 30s-ribosomal subunit assembly. Diagram displaying the interactions between the 16s rRNA domains and the 30s ribosomal subunit proteins in the 30s-ribosomal subunit assembly. The arrows represent the hierarchical binding of the r-proteins which can be characterised into early, mid or late binding proteins depending on how early the r-proteins associate with the 16s rRNA domains in the 30s-ribosomal subunit assembly. The ribosomal protein S12 has been highlighted as it is a gentamicin target protein.

Subsequently, STRING was used to predict any additional protein interactions of the *rpsL* gene. The *rpsL* network identified using STRING, contained 11 nodes and 55 edges, with a PPI enrichment p-value of 3.26e-13. All 11 proteins formed structural constituents of the ribosome playing key roles in the biological processes of gene expression and translation (Figure 4.20). A presumptive search of the 11 proteins was conducted in CLC, together with the remaining 8 30s-ribosomal subunit proteins, and again no mutations were identified.

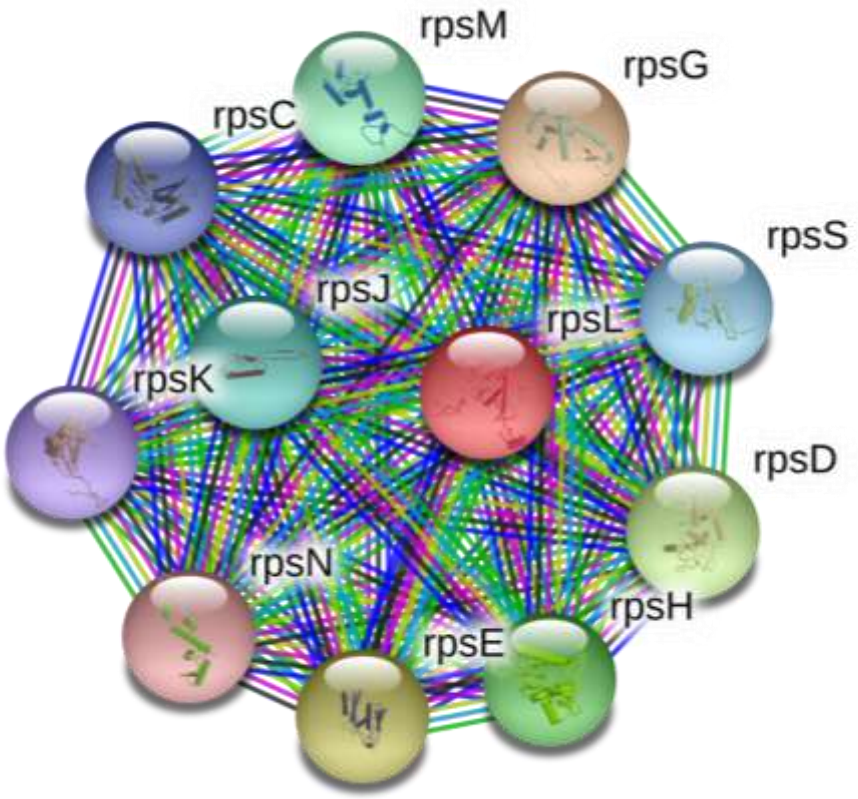


Figure 4.20 STRING protein association network of the *rpsL* gene. Image displaying the predicted interactions of the ribosomal subunit protein S12 encoded by the *rpsL* gene. The coloured nodes represent the first shell of protein interactors with predicted 3D structures available. The edges or lines between the proteins are coloured based on type of interaction: gene neighbourhood (green), gene fusions (red), gene co-occurrence (dark blue). The thickness of the lines relates to the data or evidence available in support of the predicted interaction.

4.3.4 Genome sequence alignment in Mauve

Mauve was then used to Identify any large scale evolutionary events, between the sensitive and resistant genomes. Firstly, the sensitive and resistant genomes were independently aligned to the *E. coli* K-12 MG1655 reference genome.

The initial alignment and movement of contigs of the sensitive strain to the reference genome, revealed a total of four LCBs, with a minimum weight of 1809. The total LCB length was 1173254 weighing 1182222 (Figure 4.21). Further assessment of LCBs revealed regions within the blocks which were not entirely coloured. The alignment was then magnified in order to examine the annotated CDS features, and segments of missing DNA were identified. On closer inspection, most gaps within the alignment were for genes encoding transposases.

When comparing this to the alignment of the resistant strain against the reference genome. There was a total of two LCBs with a minimum weight of 15511. The total LCB length was 4542957 weighing 4584001 (Figure 4.22). The central region of the LCB was connected by a single line indicative of homologous regions the genome. As seen with the sensitive genome gaps were identified within transposase encoding genes.

The final stage after the alignment and movement of contigs to the *E. coli* reference genome, was to align the sensitive and resistant genomes independently. A total of one LCB was identified, with a LCB length of 4574064 weighing 4577513 (Figure 4.23). There was a drastic reduction in the number of genomics gaps identified, with conserved segments free of rearrangements or large scale evolutionary events. To assess the small scale evolutionary events, single nucleotide changes between the two genomes were exported.

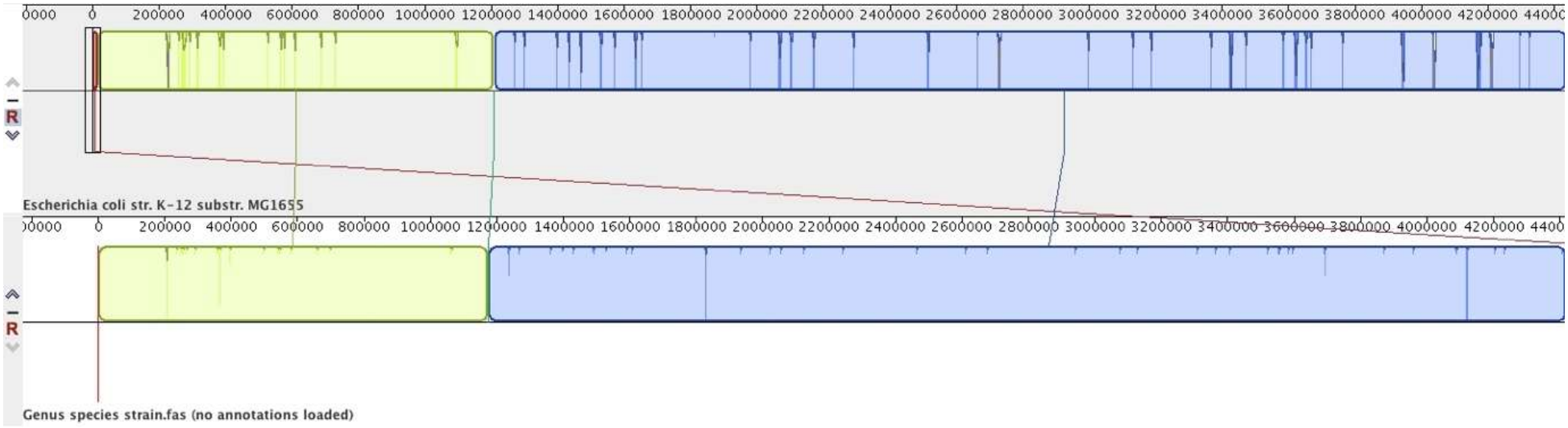


Figure 4.21 Mauve alignment of the gentamicin sensitive *E. coli* MG1655 genome and the *E. coli* K-12 MG1655 reference genome. Alignment image of the Mauve Contig Mover output of the ordered gentamicin sensitive *E. coli* genome relative to the *E. coli* MG1655 K-12 reference genome (top panel). The conserved regions of the genomes are arranged into locally Collinear Blocks(LCBS) indicating regions free of genomic rearrangement. The LCBs lie above the centre line indicating the genomes are in the correct orientation. The connecting lines between the LCBs of the two genomes indicate the aligned segments between genomes. The vertical coloured lines within the LCBs indicates potential gaps or polymorphisms within the conserved genomic regions of the genomes. The red lines signify the contig boundaries in the genome alignment.

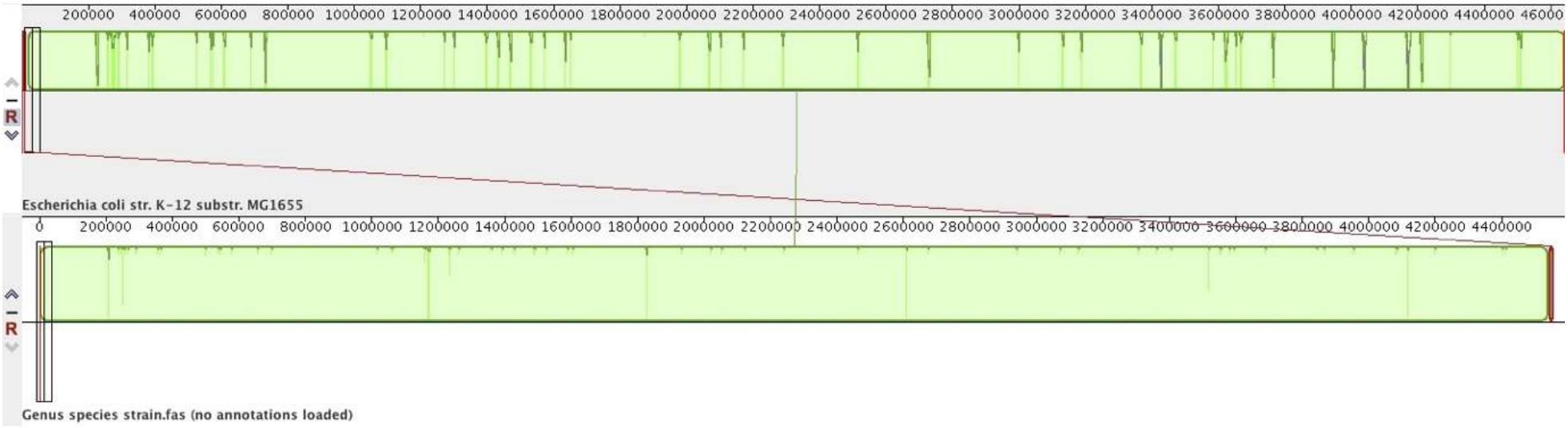


Figure 4.22 Mauve alignment of the gentamicin resistant *E. coli* MG1655 genome and the *E. coli* K-12 MG1655 reference genome. Alignment image of the Mauve Contig Mover output of the ordered gentamicin resistant *E. coli* genome relative to the *E. coli* MG1655 K-12 reference genome (top panel). The conserved regions of the genomes are arranged into locally Collinear Blocks(LCBS) indicating regions free of genomic rearrangement. The LCBs lie above the centre line indicating the genomes are in the correct orientation. The connecting lines between the LCBs of the two genomes indicates the aligned segments between genomes. The vertical coloured lines within the LCBs indicates potential gaps or polymorphisms within the conserved genomic regions of the genomes. The red lines signify the contig boundaries in the genome alignment.

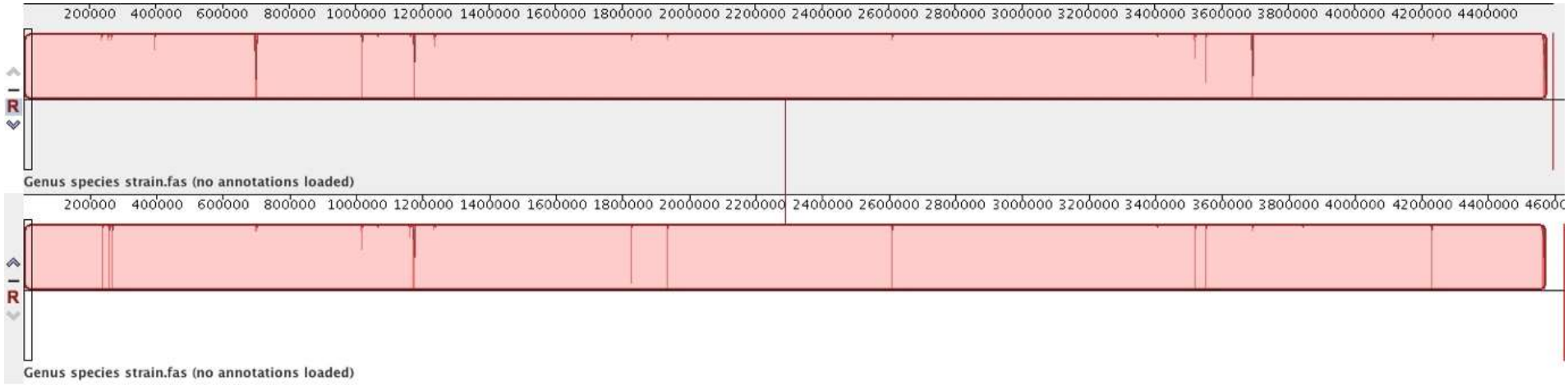


Figure 4.23 The progressive Mauve alignment of the gentamicin sensitive and resistant *E. coli* MG1655 genomes. The pairwise whole genome alignment image of the gentamicin sensitive (top panel) and resistant strains (bottom panel) of *E. coli* MG1655. The single Locally Collinear Block (LCB) between the genomes suggests high genomic conservation. The LCBs are above the genomes centre line and are in forward orientation. The vertical lines within the LCBs indicate regions of the genomes with gaps or polymorphisms. The red lines outside the LCBs indicate the contig boundaries of the genomic alignment.

4.3.4.1 Mauve SNP identification

After the successful alignment of the sensitive and resistant genomes within Mauve. The SNP export function was utilised, to identify the single nucleotide changes observed within the two genomes. A total of 1,166 SNPs were identified, with equal representation of 583 SNPs per genome. Following SNP identification, scripts were used to examine the SNP raw export file, to identify the SNPs which occur within coding regions of the genomes (Appendix table A.1). Confirmatory manual checks of SNP export files were also conducted. A total of 143 SNPs were identified between the sensitive and resistant genomes.

Following the identification of the nodes containing SNPs within the coding regions of the genomes, it was necessary to identify the genes present within each node containing the SNPs for the both the sensitive and resistant genomes (Table 4.2).

Table 4.2 Mauve node SNP summary. Table displaying the nodes containing genes with SNPs and the corresponding CDS positions within the gentamicin sensitive and resistant genomes of *E. coli* MG1655.

Strain	Node	Gene	CDS Position
Sensitive	2	*yjeO*	112915 113217
	20	*ldrD_2*	59566 59673
	57	*spoIIIE*	1114 2310
	15	*fusA*	333 2447
	59	*pinR*	673 1263
	14	*mdoG*	98749 100251
Resistant	16	*mdoC*	98976 100133
	15	*fusA*	333 2447
	58	*spoIIIE*	1307 2089
	59	*pinR*	673 1263

Once the genes containing SNPs had been identified, it was essential to obtain the nucleotide sequences. This was achieved using the script get_seq_go.py, first by obtaining the nodes containing the SNPs and secondly using the CDS positions retrieving the sequences.

A subsequent BLAST search of all sequences, revealed that the *SpoIIIE* gene which encodes a DNA translocase was a contaminant from organism *Bacillus subtilis*. The remaining nucleotide sequences were translated in order to eliminate synonymous mutations, leaving only those non-synonymous mutations which cause alterations in

the amino acid sequence of proteins. Two of the genes identified; the *ldrD_2* small toxic polypeptide encoding gene and the yjeO inner membrane protein encoding gene contained synonymous mutations, resulting in no alterations in amino acid sequence. The *mdoC* and *mdoG* genes were ignored as there were different annotations for the same gene, which might be representative of misannotation and neither genes have an amino acid change. Leaving the *fusA* and *pinR* genes which contained non-synonymous mutations.

The *fusA* gene is 2113 nucleotides in length and encodes the protein Elongation Factor G. The *fusA* gene was identified with one SNP at position 1828, here the nucleotide cytosine (C) was mutated to adenine (A) in the resistant strain (Figure 4.24 A). The given mutation identified resulted in an amino acid change at position 610 from a proline (P) to a threonine (T) residue respectively (Figure 4.24 B).

The *pinR* gene is 591 nucleotides in length, and encodes the protein Serine Recombinase. The *pinR* gene was identified with one SNP at position 8, here the nucleotide Guanine (G) was mutated to adenine (A) in the resistant strain (Figure 4.25 A). This mutation resulted in an amino acid change at position 3 from an arginine (R) to a Glutamine (Q) residue respectively (Figure 4.25 B).

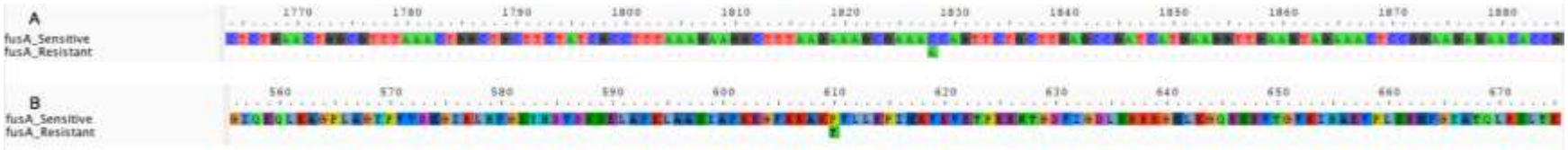


Figure 4.24 The *fusA* gene alignment. Aliview alignment image displaying the non-synonymous SNP within the nucleotide sequence of the *fusA* gene (A) and the corresponding amino acid alteration in the *fusA* protein sequence (B) between the sensitive and resistant strain of *E. coli* MG1655. The dots in the alignment are representative of identical nucleotides/amino acids between the two bacterial genomes.

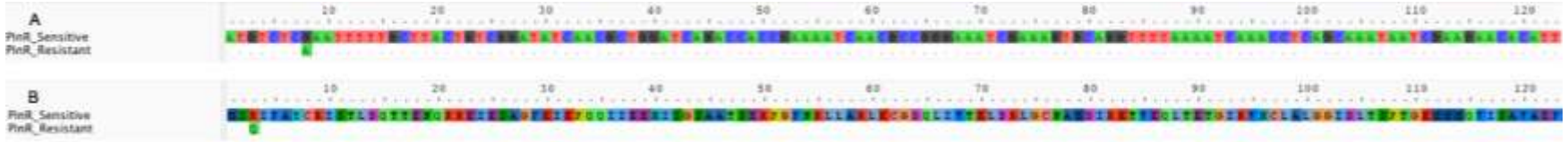


Figure 4.25 The *pinR* gene alignment. Aliview alignment image displaying the non-synonymous SNP within nucleotide sequence of the *pinR* gene (A) and the subsequent amino acid alteration within the *pinR* protein sequence (B) between the sensitive and resistant strain of *E. coli* MG1655. The dots in the alignment are representative of identical nucleotides/amino acids between the two bacterial genomes.

To determine any functional interactions between the mutated *fusA* and *pinR* genes, STRING was utilised. The precise interactions of the proteins containing SNPs could be indicative of key pathways associated with gentamicin resistance.

The *fusA* network contained a total of 11 nodes plus 55 edges with a PPI enrichment p-value of 6.27e-11. All the 10 identified proteins within the *fusA* network formed structural constituents of the ribosome and were involved in the biological processes of gene expression and translation (Figure 4.26). The network comprised of ribosomal proteins for the small 30s and large 50s ribosomal subunits encoded by the *rps* and *rpl* genes respectively. Of significant importance, is the evidence suggesting a functional link between *fusA* and the *rpsL* gene, a known gentamicin target. The two genes have an intergenic distance of 664bp (score 0.535), are both involved in RNA binding and have fundamental roles during translation.

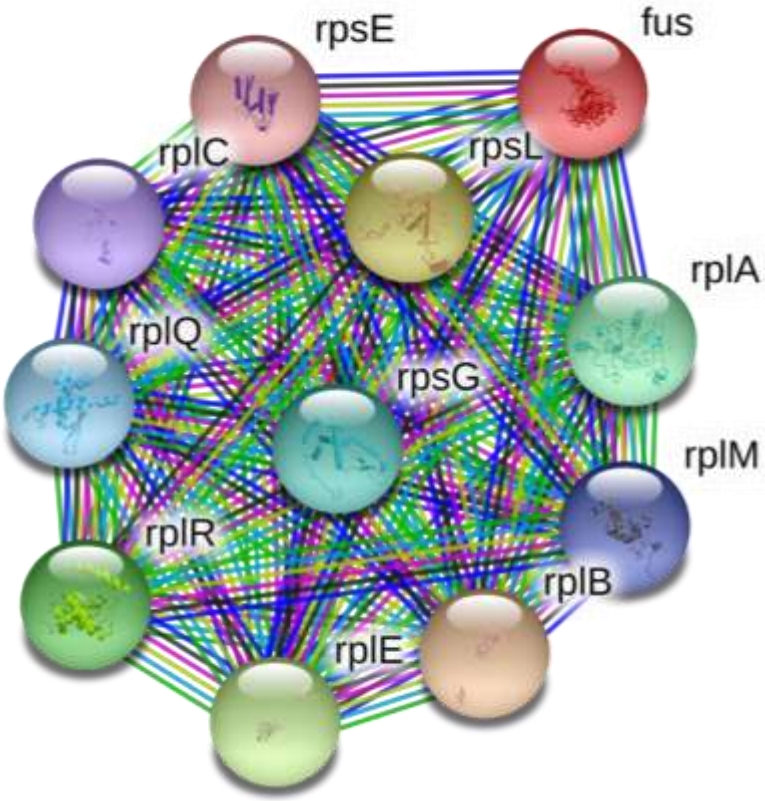


Figure 4.26 STRING protein association network of the *fusA* gene. Image displaying the predicted functional interactions of the ribosomal protein EF-G encoded by the *fusA* gene. The coloured filled nodes indicate the primary protein interactors with known 3D structures. The lines or edges represent the known or predicted interactions and the line thickness correlates to the evidence available in support of the specific interaction.

The *pinR* network contained a total of 8 nodes plus 18 edges with a PPI enrichment p-value of 0.000835. This indicated that the network had significantly more interactions than expected (Figure 4.27). The network identified the *pinR* gene as a Rac Prophage with both recombinase and resolvase activity. The majority of proteins connections within the network comprised of prophages (*pinQ*, *tfaR*, *tfaQ* and *YnaK*) or prophage components (*ynaE* and *rzoR*). An additional experimentally determined interaction of *pinR* was with the *polA* gene a fused DNA polymerase with exonuclease activity (Score 0.460).

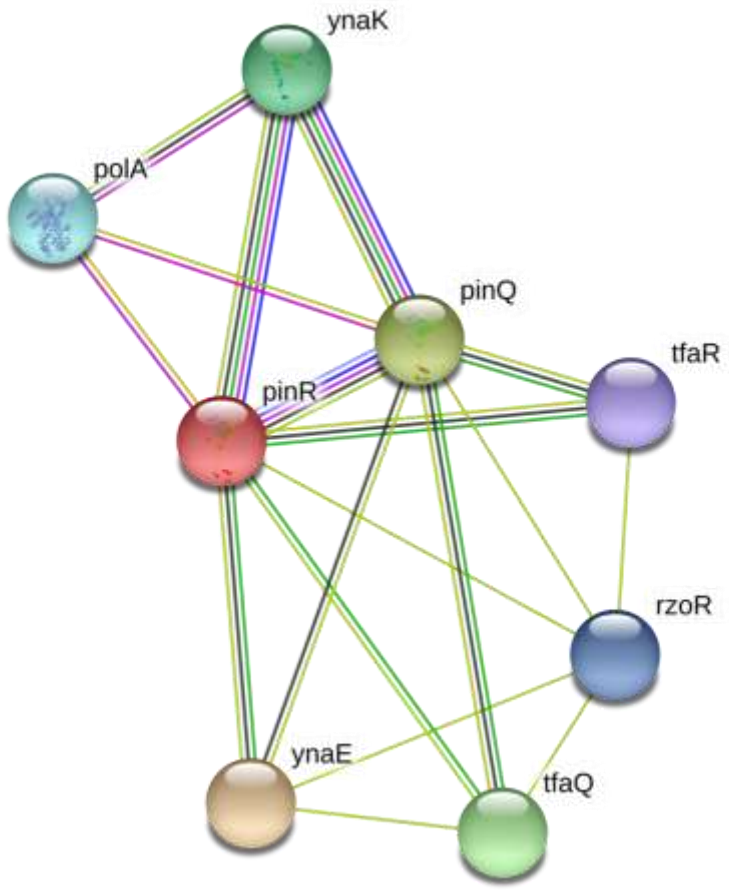


Figure 4.27 STRING protein association network of the *pinR* gene. Image displaying the predicted interactions of the putative serine recombinase *PinR*. Empty coloured nodes represent query proteins and first shell interaction proteins with unknown 3D structures. The lines or edges can be coloured based on other interactions in addition to those which or known or predicted. The yellow/green line represents text mining evidence and the black line suggests evidence of coexpression.

4.3.5 Proteome analysis using CD-HIT

Proteome analysis using CD-Hit was used as a confirmatory test for the identification of mutated proteins within the sensitive and resistant proteomes, but also to analyse the extent of contamination for the improvement of subsequent bioinformatic analyses.

The initial analysis within CD-Hit was conducted to identify any novel proteins which were not 100% identical for the sensitive and resistant proteomes. The analysis revealed 13 novel proteins within the sensitive strain (Table 4.3) and 30 novel proteins within the resistant strain (Table 4.4).

Table 4.3 The non-identical proteins within the proteome of the sensitive strain of *E. coli* MG1655. Summary of the non-identical proteins identified in the proteome of the gentamicin sensitive *E. coli* MG1655 strain.

CD-Hit Non-Identical Proteins Sensitive Strain	
1	sensitive_hypothetical protein
2	sensitive_hypothetical protein
3	sensitive_hypothetical protein
4	sensitive_putative deoxyribonuclease RhsB
5	sensitive_Putative deoxyribonuclease RhsC
6	sensitive_ATP synthase gamma chain
7	sensitive_putative major fimbrial subunit LpfA
8	sensitive_hypothetical protein
9	sensitive_Small toxic polypeptide LdrD
10	sensitive_hypothetical protein
11	sensitive_hypothetical protein
12	sensitive_hypothetical protein
13	sensitive_HTH-type transcriptional regulator ArcR

Table 4.4 The non-identical proteins within the proteome of the resistant strain of *E. coli* MG1655. Summary of the non-identical proteins identified in the proteome of the gentamicin resistant *E. coli* MG1655 strain.

CD-Hit Non-Identical Proteins Resistant Strain	
1	resistant_Putative deoxyribonuclease RhsC
2	resistant_Ethanolamine utilization protein EutM
3	resistant_Putative deoxyribonuclease RhsC
4	resistant_ATP synthase gamma chain
5	resistant_Maltodextrin glucosidase
6	resistant_Elongation factor G
7	resistant_DNA-binding transcriptional repressor YiaJ
8	resistant_Galactoside O-acetyltransferase
9	resistant_Putative deoxyribonuclease RhsC
10	resistant_DNA translocase SpoIIIE
11	resistant_Serine recombinase PinR
12	resistant_hypothetical protein
13	resistant_hypothetical protein
14	resistant_hypothetical protein
15	resistant_hypothetical protein
16	resistant_hypothetical protein
17	resistant_Putative deoxyribonuclease RhsC
18	resistant_hypothetical protein
19	resistant_hypothetical protein
20	resistant_hypothetical protein
21	resistant_hypothetical protein
22	resistant_hypothetical protein
23	resistant_hypothetical protein
24	resistant_PTS system fructose-specific EIIB component
25	resistant_UDP-N-acetylglucosamine 1-carboxyvinyltransferase
26	resistant_hypothetical protein
27	resistant_hypothetical protein
28	resistant_NADP-dependent alcohol dehydrogenase C
29	resistant_hypothetical protein
30	resistant_Lactose operon repressor

The identified proteins were then further clustered at 95% similarity, to identify the sensitive proteins which are not in the identified resistant novel proteins, and the resistant proteins which are not present in the sensitive novel proteins. This allows the identification of altered proteins as a result of nucleotide polymorphisms

assists in the removal of contamination.

Following, the sensitive vs resistant cluster swap at 95%, seven proteins were identified with nucleotide changes between both the sensitive and resistant proteomes (Table 4.5). In addition, 15 proteins were identified exclusively within the sensitive proteome, which could be either novel proteins or contaminants.

Table 4.5 CD-HIT proteome analysis of the *E. coli* MG1655 sensitive vs resistant swap. Summary of the CD-Hit analysis of the gentamicin sensitive proteins clustered at 95% similarity. The black outlined regions of the table indicate clusters of interest with proteins containing nucleotide polymorphisms. Asterisks (*) indicate 100% similarity.

Sensitive Vs Resistant Swap Cluster			
Cluster	**Amino Acids (aa)**	**Predicted Protein**	**% Similarity**
0	105 309	Sens_Putative Res_Putative	* 99.68
1	1032 289	Sens_Putative Res_Putative	* 99.65
2	704 704	Sens_Elongation Res_Elongation	* 99.86
3	477 311	Sens_Putative Res_Putative	* 99.68
4	398 260	Sens_DNA Res_DNA	* 99.23
5	287 258	Sens_ATP Res_ATP	* 95.35
6	262	Sens_Hypothetical	*
7	196 196	Sens_Serine Res_Serine	* 99.49
8	167	Sens_Hypothetical	*
9	165	Sens_Putative	*
10	125	Sens_Hypothetical	*
11	125	Sens_Hypothetical	*
12	125	Sens_Hypothetical	*
13	113	Sens_Hypothetical	*
14	110	Sens_Carboxylestera	*
15	97	Sens_HTH-type	*
16	91	Sens_Hypothetical	*
17	91	Sens_Hypothetical	*
18	91	Sens_Hypothetical	*
19	84	Sens_Hypothetical	*
20	82	Sens_Oxygen-depende	*
21	35	Sens_Small	*

Succeeding, the resistant vs sensitive cluster swap at 95%, two proteins were recognised with nucleotide changes common to both the sensitive and resistant proteomes. Furthermore, 29 proteins were identified within the resistant proteome independently (Table 4.6). This suggested greater contamination within the resistant

proteome.

Table 4.6 CD-HIT proteome analysis of the *E. coli* MG1655 resistant vs sensitive swap. Summary of the CD-Hit analysis of the gentamicin resistant proteins clustered at 95% similarity. The black outlined regions of the table indicate clusters of interest with proteins containing nucleotide polymorphisms. Asterisks (*) indicate 100% similarity.

Resistant Vs Sensitive Cluster			
Cluster	**Amino Acids (aa)**	**Predicted Protein**	**% Similarity**
0	704 704	Res_Elongation Sens_Elongation	* 99.86
1	605	Res_Maltodextrin	*
2	322	Res_Putative	*
3	315	Res_DNA-binding	*
4	311	Res_Putative	*
5	309	Res_Putative	*
6	289	Res_Putative	*
7	260	Res_DNA	*
8	258	Res_ATP	*
9	203	Res_Galactoside	*
10	196 196	Res_Serine Sens_Serine	* 99.49
11	188	Res_Hypothetical	*
12	177	Res_Hypothetical	*
13	148	Res_Hypothetical	*
14	138	Res_Hypothetical	*
15	123	Res_Hypothetical	*
16	113	Res_PTS	*
17	111	Res_Ethanolamine	*
18	109	Res_Hypothetical	*
19	107	Res_UDP-N-acetylglu	*
20	104	Res_Hypothetical	*
21	103	Res_Hypothetical	*
22	91	Res_Lactose	*
23	88	Res_Hypothetical	*
24	86	Res_Hypothetical	*
25	85	Res_Hypothetical	*
26	84	Res_Hypothetical	*
27	80	Res_Hypothetical	*
28	80	Res_Hypothetical	*
29	74	Res_Hypothetical	*
30	63	Res_NADP-dependent	*

The protein sequences of the non-identical proteins for both the sensitive and resistant proteomes (Appendix tables A.2 and A.3) were submitted to BLASTp, to isolate the contaminated proteins. A total of four contaminated proteins were identified within the sensitive proteome from *Shigella* and *staphylococcus* origin (Table 4.7).

Table 4.7 *E. coli* MG1655 sensitive proteome contaminants. Summary of the contaminants identified in the proteome of the gentamicin sensitive *E. coli* MG1655 strain.

Contaminants Sensitive Strain Proteome		
Predicted Protein	**Amino Acid Length**	**Predicted Organism**
Sensitive_Hypothetical Protein (IS1 family transposase)	125	*Shigella dysenteriae*
Sensitive_Hypothetical Protein (IS1 transposase)	125	*Shigella sonnei*
Sensitive_Hypothetical Protein (IS1 ORF2)	125	*Shigella sonnei*
Sensitive_ HTH-type transcriptional regulator ArcR	97	*Staphylococcus epidermidis*

As formerly predicted, there was a greater number of contaminated proteins identified within the resistant proteome. A total of 17 contaminated proteins were classified, from a range of organisms from uncommon sources (Table 4.8). It is therefore probable, that based on the source variability of the predicted proteins that the contaminants were likely introduced during the DNA extraction or genome sequencing process.

Table 4.8 *E. coli* MG1655 resistant proteome contaminants. Summary of the contaminants identified in the proteome of the gentamicin resistant *E. coli* MG1655 strain.

Contaminants Resistant Strain Proteome		
Predicted Protein	**Amino Acid Length**	**Predicted Organism**
DNA Translocase SpoIIIE	260	*Staphylococcus pasteuri*
Resistant_Hypothetical Protein	177	*Staphylococcus*
Resistant_Hypothetical Protein	123	*Bacilli*
Resistant_Hypothetical Protein	103	*Streptococcus pneumoniae*
Resistant_Hypothetical Protein	148	*Bacilli*
Resistant_Hypothetical Protein	85	*Stenotrophomonas maltophilia*
Resistant_Hypothetical Protein	109	*Rhodospirillaceae bacterium*
Resistant_Hypothetical Protein	80	*Ralstonia*
Resistant_Hypothetical Protein	84	*Stenotrophomonas maltophilia*
Resistant_Hypothetical Protein	88	*Stenotrophomonas maltophilia*
Resistant_PTS system fructose-specific EIIB component	113	*Klebsiella pneumoniae*
Resistant_UDP-N-acetylglucosamine 1-carboxyvinyltransferase	107	*Stenotrophomonas maltophillia*
Resistant_Hypothetical Protein	86	*Pseudomonas syringae*
Resistant_Hypothetical Protein	73	*Staphylococcus nepalensis*
Resistant_NADP-dependent alcohol dehydrogenase adhC	62	*Mycobacterium tuberculosis*
Resistant_Hypothetical Protein	104	*Nocardia niwae*
Resistant_Lactose Operon Repressor	91	*Klebsiella pneumoniae*

Proceeding, the removal of contaminated proteins from the sensitive and resistant proteomes. The novel proteins were assessed further to determine their molecular functions. When looking at the novel proteins within the sensitive proteomes, there was an equal representation of proteins with mutations involved in DNA binding and self-proteolysis (Table 4.9).

Table 4.9 Molecular functions of the novel proteins identified in the *E. coli* MG1655 sensitive strain. Summary of the molecular functions of the novel proteins identified in the proteome of the gentamicin sensitive *E. coli* MG1655 strain.

Sensitive Proteins Identified with SNPs		
Sequence ID	**Protein**	**Molecular Function**
KDU50036.1	Putative transposase	DNA Binding Transposase Activity
STM57295.1	IS1 orfA	DNA Binding Transposase Activity
OKV47142.1	Type IV secretion protein Rhs	Self Proteolysis
SQK58704.1	rhsC element core protein RshC	Self Proteolysis
WP_083584225.1	Type IV secretion protein Rhs	Self Proteolysis
EFK13670.1	ATP synthase F1, gamma subunit	ATP production
WP_000554707.1	Tail fiber Protein	Mediated Virion Attachment
KQB23928.1	Toxin Ldr, type I toxin-antitoxin system family protein	Cell Death
STM57295.1	IS1 orfA	DNA Binding Transposase Activity

When comparing this to the novel proteins identified within the resistant proteomes. There was a greater number of proteins identified with unknown functions (Table

4.10). In addition, proteins were identified with diverse biological functions involving ligand binding and sugar metabolism.

Table 4.10 Molecular functions of the novel proteins identified in the *E. coli* MG1655 resistant strain. Summary of the molecular functions of the novel proteins identified in the proteome of the gentamicin resistant *E. coli* MG1655 strain.

Resistant Proteins Identified with SNPs		
Sequence ID	**Proteins**	**Function**
WP_062894316.1	RHS repeat protein	Ligand Binding
WP_001014241.1	Ethanolamine utilization microcompartment protein EutM	Structural Activity Ethanolamine Catabolism
STJ77837.1	RhsA core protein	Self Proteolysis
STM79647.1	Membrane-bound ATP synthase F1 sector gamma-subunit	ATP production
SQR79350.1	Maltodextrin glucosidase	Maltotriose Regulation Hydrolytic Activity
WP_053272490.1	Elongation factor G	GTPase Activity Translation Elongation
WP_001310587.1	DNA-binding transcriptional activator MhpR	DNA Binding Transcription Factor Activity Aromatic Catabolism
WP_001335915.1	Galactoside O-acetyltransferase	Galactoside O-acetyltransferase Activity Lactose Biosynthesis
OZY07411.1	hypothetical protein CIJ89_20860	Unknown
SQK46172.1	YD repeat-containing protein	Self Proteolysis
WP_000078178.1	DNA invertase	DNA binding Recombinase Activity
SRB28838.1	Uncharacterised protein	Unknown
SQN38131.1	Uncharacterised protein	Unknown
WP_112044739.1	RHS repeat	Ligand Binding

When directly assessing the biological functions of the novel sensitive and resistant proteins (Figure 4.28). Proteins were identified within the resistant and sensitive proteins with shared biological functions including ATP production and self-proteolysis. Time permitting, it would be of interest to have analysed the individual proteins identified within the two proteomes further.

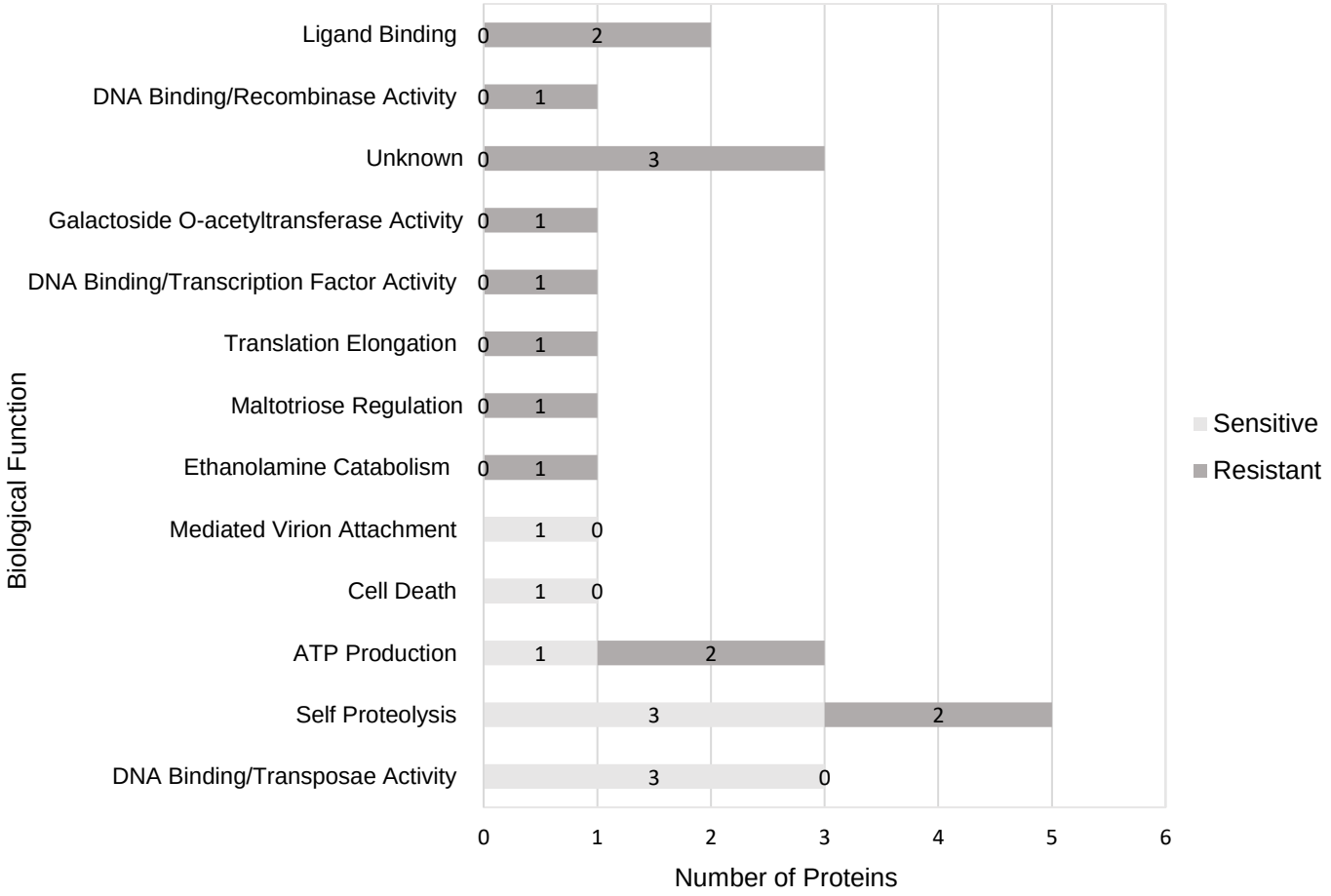


Figure 4.28 Biological functions of the sensitive and resistant proteins of *E. coli* MG1655 containing SNPs. Graph displaying the biological function of the novel proteins identified with mutations in the gentamicin sensitive and resistant strains of *E. coli* MG1655. The numbers on the graph indicate the total number of proteins within the two bacterial strains which contain mutated proteins of a precise biological function.

Conclusively, the evidence from the proteome analysis suggest the Mauve SNP export results were correct, with non-synonymous mutations identified in proteins; Elongation Factor G and Serine Recombinase, encoded by the *fusA* and *pinR* genes respectively (Table 4.11).

Table 4.11 The sensitive and resistant *E. coli* MG1655 proteins containing SNPs. Summary of the protein encoding genes containing non-synonymous SNPs present in the proteomes of both the gentamicin sensitive and resistant strains of *E. coli* MG1655. Asterisks (*) indicate 100% similarity.

Proteins with SNPs within the Sensitive and Resistant Proteomes			
Cluster	**Amino Acids (aa)**	**Predicted Protein**	**% Similarity**
0	704 704	Res_Elongation Sens_Elongation	* 99.86
2	704 704	Sens_Elongation Res_Elongation	* 99.86
10	196 196	Res_Serine Sens_Serine	* 99.49
7	196 196	Sens_Serine Res_Serine	* 99.49

4.4 Discussion

Whole genome sequencing successfully facilitated the identification of SNPs unique to the gentamicin resistant isolate of *E. coli* MG165. The location of SNPs predicted to be related to gentamicin resistance were not disclosed to target genes, instead it emerged resistance conferring mutations could appear in genes more distantly involved in the known molecular action of gentamicin.

4.4.1 Microbial genome sequencing, assembly and annotation

The expansion of NGS technologies has permitted high throughput sequencing and increased data generation, while compromising sequence quality (Utturkar *et al.,* 2014). Demonstrated by the growing number of draft genomes being submitted to the sequence database (Chain *et al.,* 2009). Moreover, it is approximated that sequencing platform errors are accountable for 0.1-1% of bases being miscalled (Fox *et al.,* 2015). Despite the concerns surrounding the accuracy of sequencing technologies, illumina based platforms have predominated within the field due to cost effectiveness and the reduction in the number of missed gene sequences (Mavromatis *et al.,* 2012).

Illumina sequencing of the resistant genome produced the widest coverage with a greater number of paired reads sequenced when compared to the sensitive genome. There was substantial variation in the number of reads from the two genomes; Sensitive: 1,460,808 Resistant: 5,345,144 this was unexpected as the resistant genome is a product of the evolution of its sensitive counterpart. This could suggest a reduced quality of DNA extracted from the sensitive genome or sequencing errors.

The Illumina platform is recognised as a short read sequencing technology, owing to the number of short reads and paired end reads produced following sequencing (Illumina, 2018). The rapid increase in the number of short reads generated from NGS technologies, has meant a greater emphasis has been placed upon assembly techniques for the successful reconstruction and annotation of genomes (Nagarajan and Pop, 2013). Nevertheless, the assembly of short reads is a complex process which is further complicated by the presence of repeat regions of DNA within genomes (Alkan, Sajjadian and Eichler, 2011; Treangen and Salzberg, 2013). Despite such challenges, there are currently no consensus assembly tools available (Alhakami, Mirebrahim and Lonardi, 2017).

The raw data obtained from illumina sequencing at MicrobesNG was analysed using CLC Genomics workbench. An initial pre-processing step was conducted where raw sequence reads were trimmed to discard sequences of partial or poor quality, which can negatively impact the reliability of both genome assembly and SNP detection (Fabbro *et al.,* 2013). Following trimming the ambiguities in sequence read lengths were resolved, with comparable averaged lengths of 168.5 and 166.6 for the sensitive and resistant genomes respectively. A series of quality control steps were introduced at this stage, to minimise errors when independently mapping reads to the reference genomes, a process integral for downstream analyses and accurate variant detection (DePristo *et al.,* 2011).

4.4.2 Gentamicin targets and SNP detection

Following the high-quality genome assembly, known antimicrobial targets were examined for variant detection. The aminoglycoside gentamicin functions through its precise interactions with the 30s-ribosomal subunit, a complex composed of 16s rRNA and 30s ribonucleoproteins (Green and Noller, 1997; Carter *et al.,* 2000). Acquired resistance to aminoglycosides occurs frequently due to mutations within antimicrobial targets, demonstrated by the presence of *rpsL* and 16s rRNA mutations conferring streptomycin resistance in *E. coli* (Toivonen, Boocock and Jacobs, 1999; Springer *et al.,* 2001). The inspection of antimicrobial targets revealed no mutations within the *rpsL* gene encoding the ribosomal protein S12, but mutations were present in the 16s rRNA. Nevertheless, due to the presence of multiple copies of rRNA operons within *E. coli,* it is difficult to suggest that resistance is mediated exclusively by this mechanism (Lambert, 2012). In addition, aminoglycoside resistance conferring mutations have been identified in several of the ribonucleoproteins (DeWilde and Wittmann-Liebold, 1973; Wilcox, Cavey and Pearson, 2001; Kirthi *et al.,* 2006). Despite a broader search directed by STRING of associated proteins, no variation was discovered.

As gentamicin is a broad spectrum antimicrobial agent, it was not feasible to carry out individual CLC searches of known targets and associated proteins, as the likelihood of missing mutations with this approach is extremely high. Therefore, Mauve was employed to identify any major genome rearrangements and for the exportation of SNPs. Using Mauve, 143 SNPs were identified independently within the coding regions of the resistant and sensitive bacterial genomes, through the visual analysis

of the SNP export file which was then verified using a script. Further sequence analysis revealed, the presence of a contaminated gene *SpoIIIE* from the organism *B. subtilis,* encoding a membrane bound DNA translocase which plays an essential role in chromosomal segregation during sporulation (Wu and Errington, 1994; Kaimer, González-Pastor and Graumann, 2009). It is probable that the *B. subtilis* contaminant was introduced during the DNA extraction or sequencing process at MicrobesNG. In addition, the paralogous genes *MdoG* and *MdoC* required for the synthesis of the Osmoregulated Periplasmic Glucan (OPG) backbone were misannotated (Lacroix *et al.,* 1999; Lequette *et al.,* 2004). This occurs commonly in whole genome assembly where paralogous genes, due to their highly similar sequences, fail to be assembled into separate loci (Vinson *et al.,* 2005).

In this study, synonymous SNPs were considered inconsequential, as an assumption was made that these mutations are selectively neutral and result in no alteration in phenotype (Hunt *et al.,* 2009). However, there has been growing emphasis on experimental evolution approaches aimed at exploring the adaptive consequences of synonymous mutations (Zwart *et al.,* 2018). Such experimental evolution studies have been conducted on the *TEM-1* β-lactamase encoding gene, which have linked synonymous mutations with alterations in gene expression, protein folding and fitness (Salverda *et al.,* 2011; Firnberg *et al.,* 2014). This suggests that when investigating the evolutionary dynamics of resistance, all genetic variants should be accounted for.

After discounting the synonymous mutations within the two genomes, a total of two non-synonymous polymorphisms were identified in the *fusA* and *pinR* genes. The

pinR gene encodes a putative DNA invertase/site-specific recombinase, which has been closely associated with the lambdoid prophage RAC in the *E. coli* K-12 genome (Mehta, Casjens and Krishnaswamy, 2004). Serine recombinases are frequently identified as having >30% sequence similarity, across the highly conserved catalytic domains, however the precise DNA inversion reactions are not well characterised (Johnson, 2015). Whereas, the *fusA* gene encodes the protein EF-G essential to ribosomal protein synthesis, providing a well-defined link to the antimicrobial mechanism of action (Agrawal *et al.,* 1998). This is consistent with previous findings, where bacterial populations can survive selective pressures providing the presence of at least one large effect beneficial mutation (Dijk Van *et al.,* 2017).

Previous experimental evolution studies, exploring the evolutionary trajectories of aminoglycoside resistance in *E. coli* have identified mutations in similar positions within the *fusA* gene (Mogre *et al*., 2014, Mogre *et al*., 2017 and Ibacache-Quiroga *et al.*, 2018). Genomic analysis from an *in-vitro* laboratory evolution study investigating the concentration dependent evolution of kanamycin resistance through serial passage, identified multiple kanamycin resistant mutations in the C-terminal end of domain 4 of EF-G (Mogre *et al*., 2014). Significantly, at sublethal concentrations of kanamycin the *fusA* (P610T) was identified, while at higher sublethal concentrations two further *fusA* mutations were discovered; *fusA* (P610L) and *fusA* (A608E) (Mogre *et al*., 2017). In a separate study, using gentamicin as a model to explore the mutational pathways to aminoglycoside resistance in both normo and hypermutable strains of *E. coli*, identified diverse mutations in all gentamicin resistant derivative

including a *fusA* (F593L) substitution in domain 4 of EF-G (Ibacache-Quiroga *et al*., 2018). The findings from the laboratory evolution studies described, indicate that despite the experimental approach utilised, aminoglycoside antibiotics elicit a similar adaptive response within the *fusA* gene encoding *EF-G*. Additionally, these findings emphasise the functional significance of the C-terminal end of domain 4 of EF-G.

4.4.3 Proteome analysis and contaminant identification

The SNPs detected using Mauve were validated using CD-HIT-2D protein sequence analysis, which provides an alternative approach to analysing sequencing data through assigning proteins into clusters based on similarity (Li and Godzik, 2006). CD-HIT-2D facilitated the rapid identification of contaminated proteins and permitted the characterisation of genomic gaps corresponding to previously unidentified protein coding genes (Mavromatis *et al.,* 2012). Interestingly, unlike the consistency seen with the number of SNPs identified between the two genomes using Mauve, there was a greater number of non-identical or contaminated proteins found within the resistant strain than its sensitive counterpart. This increase in the number of divergent proteins present within the resistant dataset, may reflect the differentiating number of sequence reads between the two genomes following illumina sequencing (Olson ND *et al.,* 2012). Following a tblastn search of the non-identical sequences, proteins were detected within the sensitive and resistant strains from a diverse range of species including *Ralstonia*, a common contaminant of ultrapure water systems in NGS technologies (Lawrence, Hatzis and Brash, 2014). The identification and removal of contaminants introduced during the DNA extraction or sequencing process is essential for downstream analysis and the accurate interpretation of results (Sheik *et al.,* 2018). Consistent to the findings from mauve, polymorphisms were confirmed

within the EF-G and Serine Recombinase proteins encoded by *fusA* and *pinR* genes respectively.

4.4.4 Conclusion

In conclusion, of the SNPs identified, those present in the 16s rRNA and *fusA* genes are most likely to be associated with the acquisition of gentamicin resistance. Further analysis will reveal the functional consequences of these SNPs and their potential to be utilised as markers for aminoglycoside resistance.

Chapter 5: Physiochemical analysis of the genes predicted to be involved in the evolution of gentamicin resistance in *E. coli* MG1655.

5.1 Introduction

The evolution of antimicrobial resistance occurs through the process of natural selection, following the emergence of mutations within a bacterial genome. The acquisitions of mutations within a genome is complex and dependent on a range of factors including the genomic stability and mutability of organisms (Martinez and Baquero, 2000; Matic, Taddei and Radman, 2004; Chopra, O'Neil and Miller, 2003). Furthermore, genomic adaptation is known to be induced in response to stress (Galhardo, Hastings and Rosenberg, 2007).

Antimicrobial agents drive resistance evolution in bacteria, however the fixation of resistant conferring mutations is dependent on the site of mutagenesis and whether the mutation confers a selective advantage (Long *et al.,* 2016; Schenk and de Visser, 2013; Bamshad and Wooding, 2003). Thus, the first aim of this chapter is to analyse the mutations identified within chapter 4, and to determine the functionally consequences of these mutations relative to the acquisition of gentamicin resistance.

The rRNA within the 70s ribosomal complex is essential for the functioning of the ribosome, this becomes more evident when the ribosome is targeted by antimicrobial agents which directly interfere with the ribosomal function during protein synthesis (Brodersen *et al.,* 2000). The genetic and functional analysis of the mutations identified within the rRNA is challenging due to the multiple copies of the rRNA genes

within *E. coli*, as the phenotypic expression of these mutations may be altered by the presence of sensitive to resistant ribosome variants (Lee, Holland-Staley and Cunningham, 2001). As rRNA mutations do not encode protein it is fundamental to assess the effect of single nucleotide mutations on the conformational rearrangements of the secondary structures of the mutated rRNA variants. Mutagenesis within the 16s rRNA could lead to alterations in both the free energy and the conformations of the secondary structure motifs, which are key in the maintenance of ribosomal function (Barash and Churkin, 2010).

SNP are the most common form of genetic variation; therefore, it is necessary to determine which SNPs lead to amino acid substitutions in protein encoding genes, which could potentially alter protein function (Clifford *et al.*, 2004; Ng and Henikoff, 2003). The position of the amino acid substitutions is significant, particularly if the non-synonymous mutations are located within conserved proteins or domains which are most likely to be under positive selection (Choi *et al.*, 2012). In addition, the specific amino acid substitutions which arise as a result of mutagenesis are significant due to the variable physical properties of amino acids and the effects this can have on protein function (Kidera *et al.,* 1985). Furthermore, the effects of mutations can often be identified through protein structure analysis, which can often be used to explain the alterations in function (Chasman and Adams, 2001).

As a result of the sequence variation and the potential for wrongly called bases through NGS techniques, it is necessary to validate the SNPs identified which are most likely to be attributable to the acquisition of gentamicin resistance. SNP

validation can be achieved a variety of methods including bi-parental segregating population, but in this chapter the validity of SNPs will be assessed through confirmatory PCR and resequencing (Kumar, Banks and Cloutier, 2012; Li *et al.,* 2009). The SNPs were considered valid based on the recurrence of identical SNPs within the *fusA* and 16s rRNA genes of parallel evolved gentamicin resistant populations of *E. coli* MG1655, using confirmatory PCR and resequencing.

It is hypothesised that there will be physiochemical variations present because of the evolved mutations in genes specific to the function of gentamicin or the associated antimicrobial pathways. Further to this, the mutations identified through WGS will be validated by their presence in other resistant strains isolated from the ARGP.

5.2 Materials and methods

5.2.1 RNA consensus secondary structure XRNA

The consensus secondary structure of the 16s rRNA within *E. coli*, was stipulated using XRNA to provided information regarding the locality of mutations. Further details of the protocol used within the XRNA suite can be found in section (2.9.1).

5.2.2 RNAfold secondary structure analysis

The initial RNA secondary structure predictions and analysis were carried out using RNAfold a stand-alone programme. Details of the RNAfold protocol are described in section (2.9.2).

5.2.2.1 Secondary structure analysis using Vienna

The secondary structure predictions provided by RNAfold did not facilitate the graphical representation of the specific single nucleotide changes which occur following mutagenesis. The ViennaRNA package 2.0 available at: http://www.tbi.univie.ac.at/RNA [accessed July 17, 2018], is an interactive command line tool which consists of a collection of RNA secondary structure programmes, with complementary utilities that assist in the graphical processing of secondary structures (Lorenz *et al.*, 2011). After installation of the Vienna package, the preliminary steps repeated the secondary structure predictions calculated within RNAfold using the commands below. This commands were then repeated for all sensitive and resistance rRNA genes. The output from the command is the MFE structure in dot bracket notation and its calculated free energy. The addition of options (–P) and (–MEA) to the command allows the calculations of partition function (-P) and maximum expected accuracy (MEA) respectively (Figure 5.1).

```
MacBook-Pro:vienna Lucky$ RNAfold < rna154_SEN.seq
MacBook-Pro:vienna Lucky$ RNAfold –p < rna154_SEN.seq
MacBook-Pro:vienna Lucky$ RNAfold –p –MEA < rna154_SEN.seq
```

Figure 5.1 The RNAfold commands utilised in the production of the 16s rRNA secondary structures within Vienna. Displaying the RNAfold command utilised within the ViennaRNA packaged 2.0 to facilitate the centroid and MFE secondary structure predictions.

In addition to using RNAfold within the Vienna package, an additional utility was required to allow the visual representation of secondary structure in the form of a mountain plot using the dp.ps output file from the RNAfold commands above. The mountain plot was previously produced using the RNAfold webserver and was repeated within Vienna using the command below (Figure 5.2).

```
MacBook-Pro:vienna Lucky$ mountain.pl rna154_SEN_dp.ps | xmgrace -pipe
```

Figure 5.2 The command utilised in the production of the 16s rRNA secondary structure mountain plots within Vienna. Displaying the mountain Plot command utilised within ViennaRNA package 2.0.

The results generated from RNAfold within Vienna can be obtained using the RNAfold webserver. However, the final application of Vienna and ultimate reason for its use, requires the .fold files which were generated from the commands shown in figure (5.1). The utility RNAplot was used to allow the graphical representations of the distinct regions containing the ribonucleotide mutations. This was achieved using the (cmark) which circles the base of interest followed by the coordinates of that base, which is then labelled using (label) and finally (omark) is used to stroke the segment by defining a width and a colour according the RGB code (Hofacker *et al.,* 2017). This is then repeated for all the single ribonucleotide changes observed within the rRNA

sequence. See below (Figure 5.3) an example of the command used for the graphical representation of mutagenesis within rna154 resistant strain.

```
MacBook-Pro:vienna Lucky$ RNAplot  --pre "93 cmark 93 -3.5 -0.5 (pos93) Label
93 93 12 GREEN omark 226 cmark 226 -4.0 -0.8 (pos226) Label 226 226 12 RED
omark 90 cmark 90 -3.5 -0.5 (pos90) Label 90 90 12 0 255 255 omark 89 cmark
89 -3.5 -0.3 (pos89) Label 89 89 12 255 230 0 omark 80 cmark 80 80 -3.3 -0.4
(pos80) Label 80 80 12 238 0 238 omark 79 cmark 79 -3.5 -0.5 (pos79) Label 79
```

Figure 5.3 The RNAplot command utilised in the production of the 16s rRNA secondary structures with annotations within Vienna. Displaying the RNAplot command utilised within the ViennaRNA packaged 2.0 to facilitate the graphical representation of ribonucleotide base mutagenesis within the RNA secondary structure.

5.2.3 DNAsp sequence polymorphism analysis

DNAsp version 5.1.0 http://www.ub.edu/dnasp/ [accessed November 28, 2017] was used in the analysis of sequence polymorphisms and divergence across the 16s rRNA genes for both the sensitive and resistant strains. DNAsp was not used in the analysis of diversity for the *fusA* gene, since there was only a single nucleotide polymorphism identified. DNAsp as a tool was used to provide evolutionary insights into the nucleotide diversity of the 16s rRNA gene and to understand the functional importance of genomic regions based on the locality of the polymorphisms observed (Librado and Rozas, 2009). Ribosomal rna sequence polymorphisms were measured using sliding window analysis of nucleotide diversity (π) across the length of the 16s rRNA 123 and 154 genes, with a sliding window of 1.

5.2.4 Phylogenetic analysis using MEGA

To resolve the phylogeny of the mutated 16s rRNA sequences, maximum likelihood trees were constructed using MEGA against the *Escherichia* taxa specified in section (2.9.3).

5.2.5 EMBL SAPS protein structure analysis

The functional properties of the protein with mutations were assessed using EMBL SAPS, to determine any discrepancies in charge distribution across the protein sequence. The protocol for EMBL SAPS is described in section (2.9.4).

5.2.5.1 TreeSAAP selection of amino acid properties

Further physiochemical analysis of protein properties was conducted using TreeSAAP detailed in section (2.9.5).

5.2.5.2 NCBI conserved domains analysis

To understand the protein function in greater detail, NCBI was used to provide a summary of the conserved domains and functional sites of coding nucleotide sequences of the proteins of interest available at: https://www.ncbi.nlm.nih.gov/Structure/cdd/wrpsb.cgi [accessed August 23, 2018]. A Blast search of the CDSEARCH/cdd v3.16 database was conducted using the default search parameter options, for *fusA* and *pinR* respectively.

5.2.6 Homology modelling of protein structure: Swiss Model

The initial protein modelling was carried out using Swiss Model, to provide insights into the homology of mutated proteins between the sensitive and resistant strains of *E. coli*. Details of the Swiss Model methodology are provided in section (2.9.6).

5.2.6.1 Modelling of protein structure: Scigress

SCIGRESS (SCIGRESS, 2013) was used for the molecular modelling of the mutated protein structure further to the predicting the structural effects of mutagenesis. PDB files of both sensitive and resistant strains were imported into Scigress. For simplicity, the backbone of protein structures was changed to solid ribbon. To visualise the mutated residues within the protein structures, the sequences are displayed and observed as a three-letter code and the mutated amino acid residues were presented as ball and cylinders against the ribboned backbones ribbons of EF-G and pinR respectively. Models were comprehensively beatified and duplicate frames for the sensitive and resistant counterparts were exported and directly compared to identify any structural differences due to mutagenesis.

5.2.6.2 Protein structure analysis: PyMOL molecular graphics system

Conclusive structure analysis of the mutated proteins was completed using the molecular graphics system PyMOL (DeLano *et al.,* 2002). Ensuing consistency, all protein PDB files used in this study were rendered using PyMOL with cartoon graphical representations http://www.pymol.org [accessed March 15, 2018]. Using the integrated python feature, protein domains and proteins within large complexes and be selected, highlighted and manipulated using command syntax and atom

selections. In addition, PyMOL was used for the graphical representation of *in sillico* mutagenesis using the PyMOL mutagenesis. The mutated amino acid residues of the resistant proteins were selected using sequence mode and modelled using wizard mutagenesis. The amino acids residues pre- and post-mutagenesis were presented as sticks for visualisation and subsequent comparative analysis of secondary protein structures, supported by the export of the protein molecules as a multi-file.

5.2.7 Molecular analysis based on bioinformatics findings

5.2.7.1 Primer design and DNA extraction

To determine whether the identified mutations were present in additional bacterial strains isolated from both the ARGP and serial passage systems, specific primer sets were designed as per protocol described in section (2.10.1). Details of the primer sets utilised in the PCR reactions are indicated in table (5.1). An alternative approach to DNA extraction was utilised here through thermal lysis, as detailed in section (2.10.2).

Table 5.1 Primers used for the confirmatory PCR amplification of the evolved gentamicin resistance genes. Displaying the sequences and product sizes of the three primer sets: *fusA*, rna123 and rna154 designed using Primer BLAST.

Primer	Sequence (5'-3')	Gene	Expected PCR Product Size
fusA-1	5'-GTGTGGTTAACTCTGGTGAT-3'	*fusA*	1049
fusA-2	5'-CGCTTCATCATACTTCAGGA-3'		
rna123-1	5'-CTAACACATGCAAGTCGAAC-3'	rna123	1432
rna123-2	5'-GAATCACAAAGTGGTAAGCG-3'		
rna154-1	5'-AGAGTTTGATCATGGCTCAG-3'	rna154	964
rna154-2	5'-CGTTGCATCGAATTAAACCA-3'		

5.2.7.2 PCR and agarose gel electrophoresis

PCR reactions were executed on all bacterial DNA extracts, using the specific primers sets for the *fusA* and 16s rRNA genes. Details of the initial gradient PCR reactions and the subsequent thermocycling conditions used for the PCR reactions are available in section (2.10.3). All PCR reactions were analysed through gel electrophoresis using a 2% Top Vision agarose gel, detailed in section (2.10.4).

5.2.7.3 Illumina sequencing at The Natural History Museum facility

PCR products and original primer sets were submitted to the Natural History Museum, London, for sequencing using the Illumina MiSeq for next generation sequencing. The MiSeq platform uses non-targeted shot gun approach with fluorescent dye terminators (Illumina, 2018). Further details of the chemistry behind Illumina sequencing can be found in chapter 3. Sequence files retrieved from the natural history museum were assessed for DNA sequencing quality by base calling from a raw chromatogram using Tracetuner 3.0.6 with the following parameters -Q 3730 and -trim_threshold 20. Sequences were then trimmed based on quality and length ambiguities using a custom script available at: https://github.com/camilla-eldridge [accessed May 18, 2018] before viewing in the alignment viewer aliview.

5.3 Experimental results

5.3.1 RNA consensus secondary structure

The *E. coli* 16s rRNA consensus secondary structure template was exported from XRNA and modified to illustrate the 16s rRNA structural domains and the locations of the identified mutations. The template utilised is the consensus secondary structure model for all seven variants of the *E. coli* 16s rRNA, therefore both the rna123 and rna154 mutations have been included (Figure 5.4).

Based on the results obtained in chapter 4, where the gentamicin ribosomal protein target S12, combines early with the 16s rRNA in the 30s-ribosomal subunit assembly. An assumption could be made that the 16s rRNA mutations will be identified within the first 5' rRNA domain formed.

As predicted, of the four 16s rRNA domains (body, platform, head and penultimate stem) the mutations were identified within the 5' body domain (Figure 5.4 A). Following magnification of the body domain, the rna154 mutations occurred centrally within the structure within helix 6, whereas the rna123 mutations and the shared individual SNP occurred in the lower region of the body domain within helix 6, 7,10 and 11 (Figure 5.4 B). Helix 44 was highlighted within figure (Figure 5.4 A), and its significance will be discussed later in this chapter. The locality of the mutations was further assessed based on the 16s rRNA tertiary interactions within the 30s-ribosomal subunit using PyMOL (Figure 5.5). The rna154 mutations appeared to be situated towards the back of the ribosomal protein S12 (green), in close proximity to the

ribosomal protein S17 within the shoulder the 30s subunit. Whereas, the rna123 mutations were predominantly positioned around domains 1 and 5 of EF-G.

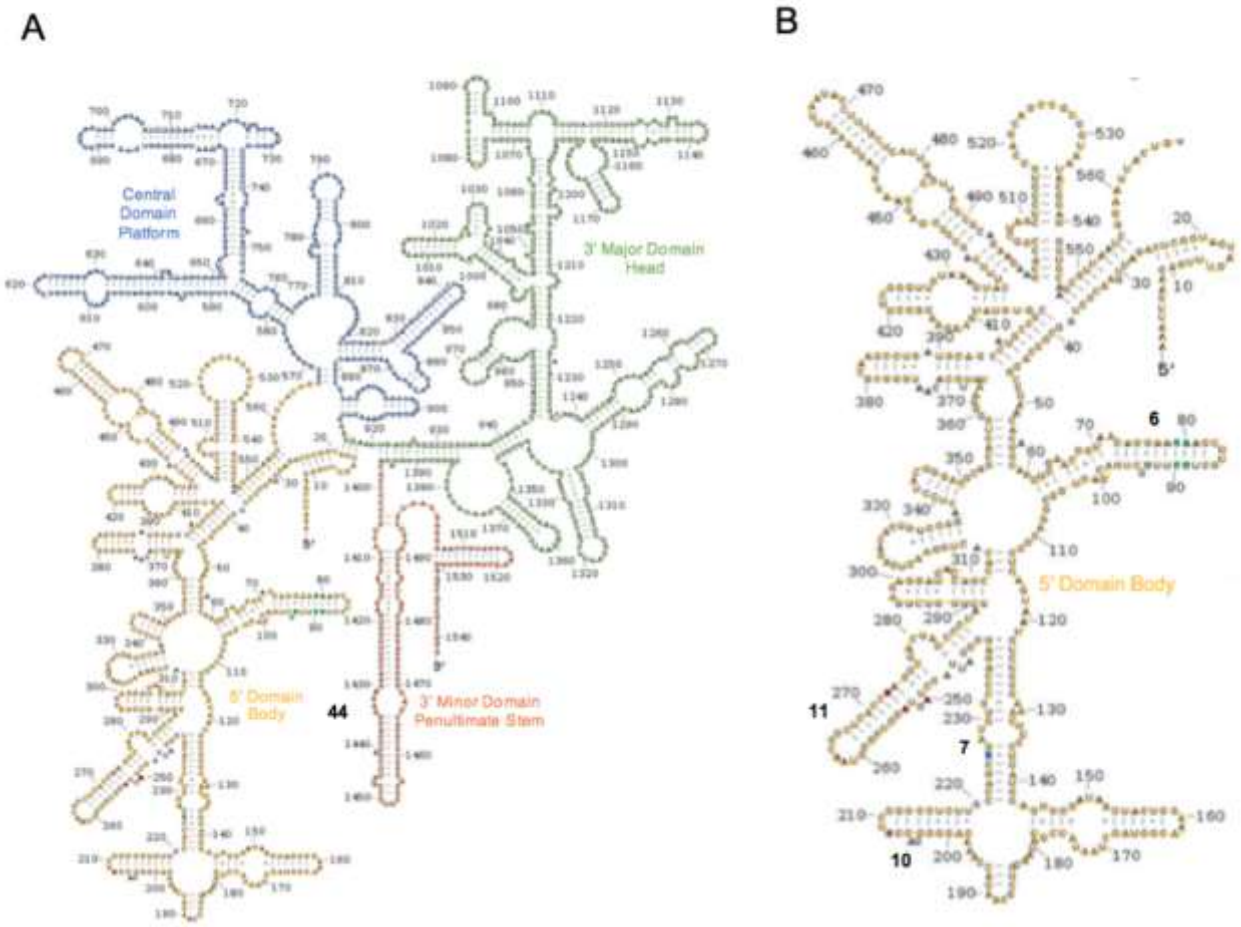


Figure 5.4 XRNA *E. coli* 16s rRNA secondary structure consensus template. (A) The four 16s rRNA domains; body (yellow), platform (blue), head (green) and penultimate stem (orange), in addition the position of helix 44 has been highlighted, for future reference. (B) Magnification of the body domain displaying the rna154 mutations in green and the rna123 mutation in red. The shared mutation at position 226 has been highlighted in blue. The 16s rRNA helices containing the mutations have been numbered in bold.

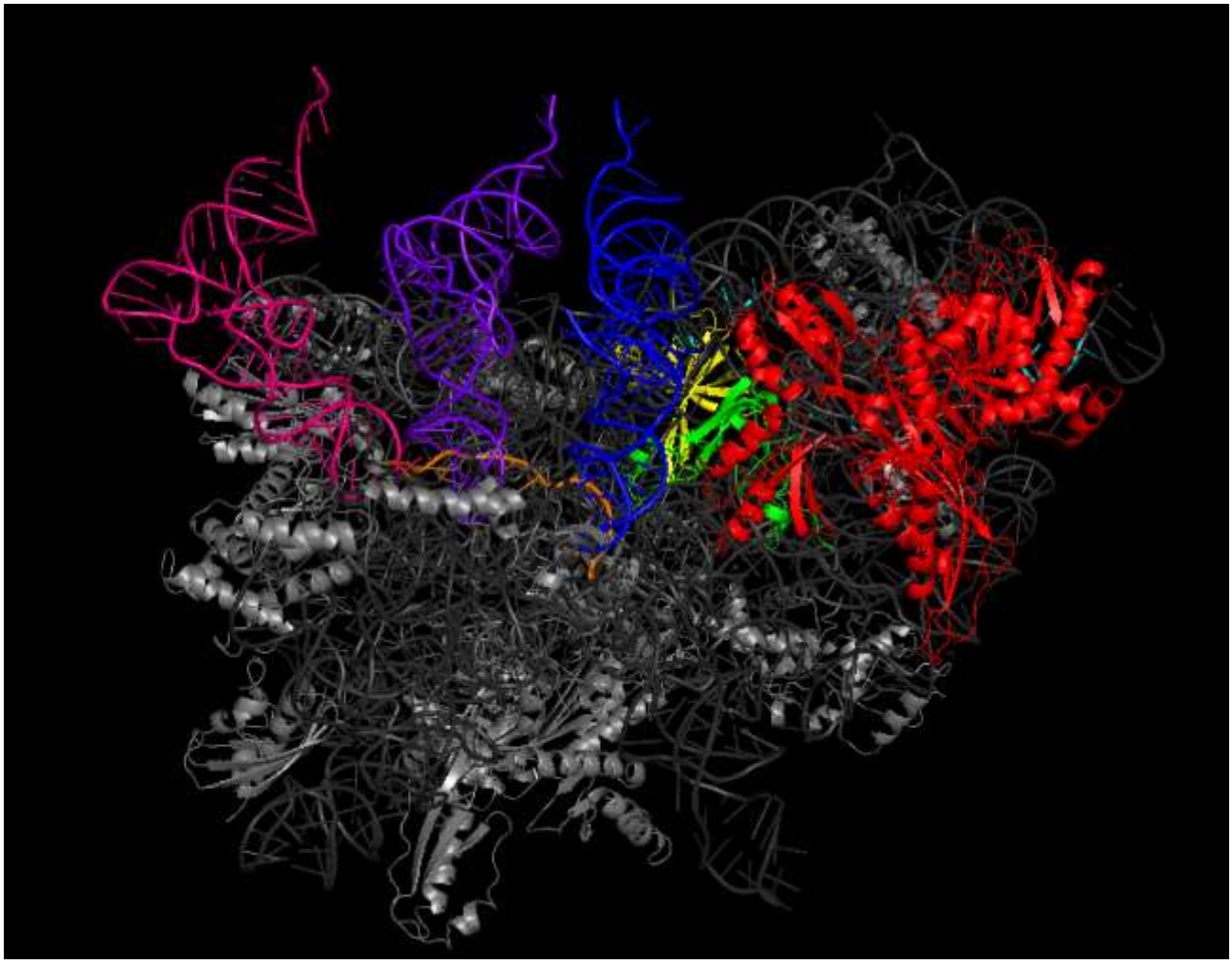


Figure 5.5 The 16s rRNA tertiary Interactions within the 30s-ribosomal subunit. Displaying the tertiary interactions of the 16s rRNA (dark grey) within the 30s-ribosomal subunit (pre-complex). Also included within the image is 30s ribosomal proteins (light grey), the tRNA molecules A (blue), P (purple), E (pink), mRNA (orange), the gentamicin ribosomal binding proteins s12 (green) and s17 (yellow) and compacted EF-G (red). The locations of the 16s rRNA mutations were shown in light blue.

5.3.2 RNA secondary structure predictions using RNAfold

Secondary structures of the individual 16s rRNA variants were obtained from RNAfold. Through the sequence submission, direct comparisons in the structural consequences of mutagenesis can be made between the sensitive and resistant strains of *E. coli*. The RNAfold webserver results can be split into three sections which will be addressed accordingly for both rna123 and rna154.

The results for rna123 based on the minimum free energy (MFE) prediction of the secondary structure, reveals a slight increase in the MFE of the resistant rna123 of -583.40kcal/mol compared to -585.00kcal/mol. In addition, the thermodynamic ensemble prediction results, suggests an increase in the free energy of the ensemble from -609.25kcal/mol in the sensitive strain to -607.70kcal/mol in the resistant counterpart. The ensemble diversity of the two strains was 345.19 and 344.40 respectively, indicating a small change in MFE. When comparing this to the centroid MFE, there was a greater increase in free energy between the two strains, from -474.17kcal/mol in the sensitive strain to -465.77kcal/mol in the resistant strain, suggestive of more apparent alterations in centroid structure.

The final output from RNAfold, is in the form of graphical reports of secondary structure (Figure 5.6). Here secondary structures were coloured specifically, to encode base pair probabilities. When analysing the MFE images, which refers to the secondary structure which contributes minimally to the free energy, there are no visible alterations in structure (Zuker and Stiegler, 1981). When comparing this to the centroid images representative of the central tendencies of a data set, variations are

visible in seven regions of the secondary structure as highlighted using asterisks (Ding, Chan and Lawrence, 2005).

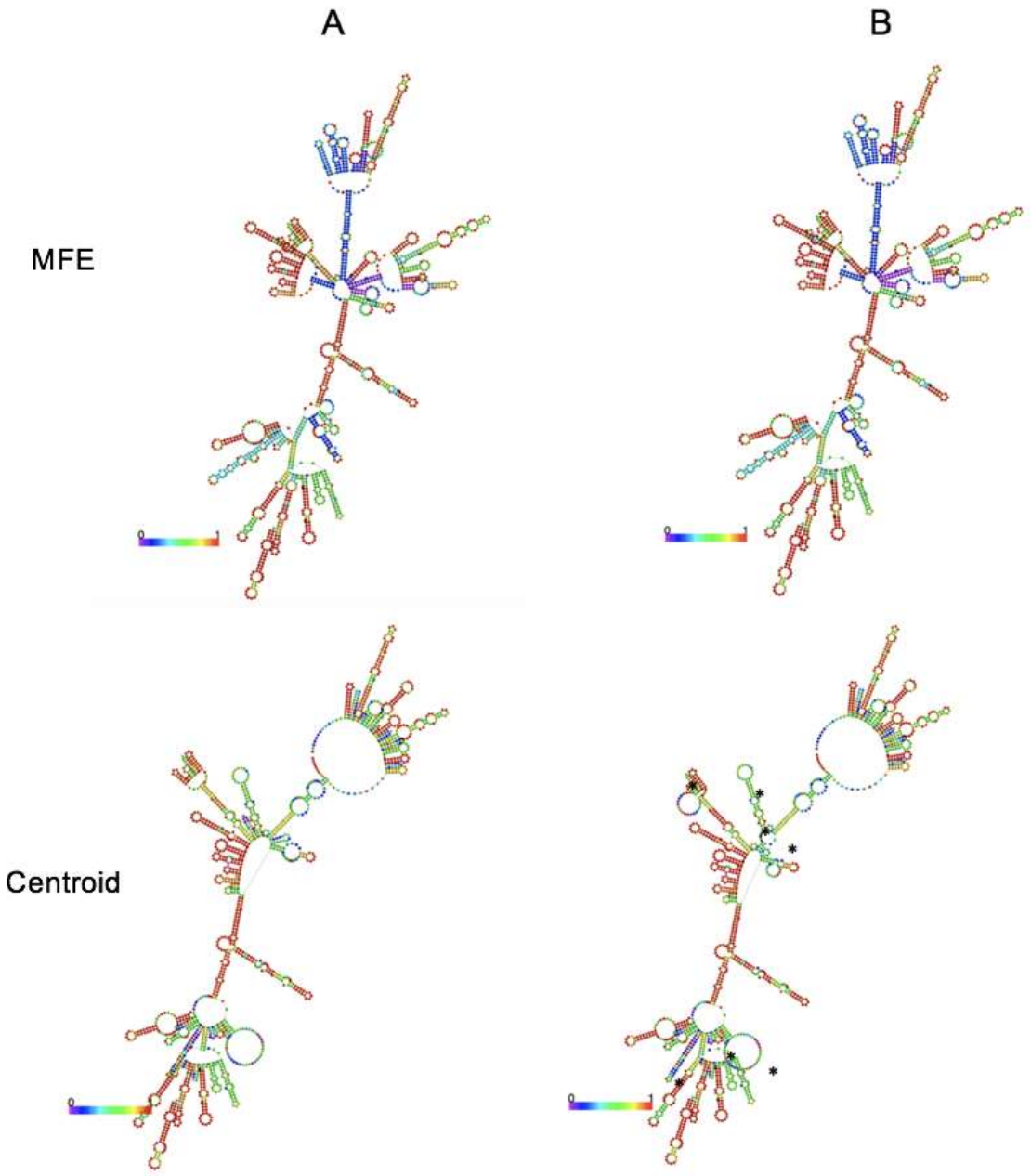


Figure 5.6 Secondary structure analysis of rRNA 123 using RNAfold. Displaying the MFE and Centroid structure images of the gentamicin sensitive (A) and resistant (B) strains of *E. coli* MG1655. The rRNA secondary structures are coloured by base-pairing probabilities scored between 0-1. The structural variation observed between the sensitive and resistant centroid secondary structures have been indicated using an asterisk (*).

When analysing the RNAfold results for rRNA 154, comparable overall secondary structures were observed to that seen with rRNA 123 (Figure 5.7). Looking more closely at the results unlike that seen with rna123, the MFE secondary structure prediction was lower in the resistant strain compared to the sensitive, -583.40kcal/mol and -580.00kcal/mol respectively. The free energy values of the thermodynamic ensemble were similar -607.74kcal/mol for the sensitive strain compared to -607.70kcal/mol for the resistant strain. A more substantial alteration was seen with ensemble diversity with a value of 381.71 for the sensitive strain, whereas the diversity value of the resistant strain was 344.40. As seen previously, there was a greater change when assessing the centroid MFE, here the sensitive strain had a higher MFE of -419.01kcal/mol compared with -465.77kcal/mol for the resistant strain.

When assessing the graphical results, there were no apparent changes in the MFE secondary structure predictions, as seen with rna123 (Figure 5.6). Similarly, apparent changes were observed in the centroid images. Here a total of 8 regions showed modifications within secondary structure, of which an evident change being the collapse of the central loop in the rna154 resistant strain.

An additional graphical report was produced in the form of mountain plots for rna123 and rna154, which can be found in the appendix (Figure A.12). As there is speculation surrounding secondary structure modelling programmes, it was decided as a form of verification to predict the secondary structures of the remaining five 16s rRNA variants. This was based on the principle that given each variant contained identical rna sequences between the sensitive and resistant strains, the structures

therefore should be identical. This proved to be the case, but an interesting finding here was that some of the remaining variants, had dissimilar overall structures to that seen with rna123 and rna154 (Appendix A.13-A.17).

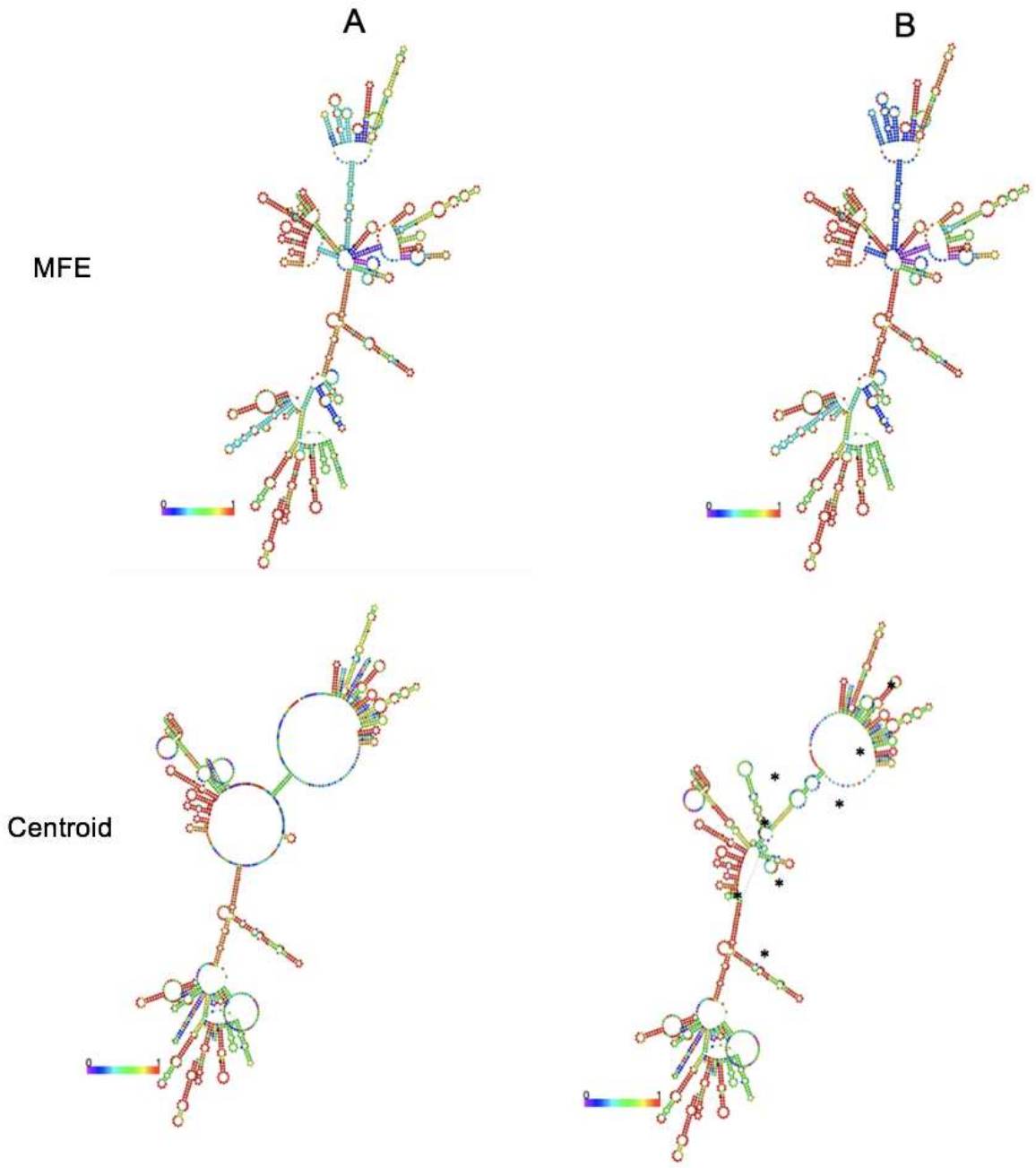


Figure 5.7 Secondary structure analysis of rRNA 154 using RNAfold. Displaying the MFE and Centroid structure images of the gentamicin sensitive (A) and resistant (B) strains of *E. coli* MG1655. The rRNA secondary structures are coloured by base-pairing probabilities scored between 0-1. The structural variation observed between the sensitive and resistant centroid secondary structures have been indicated using an asterisk (*).

5.3.2.1 RNA Secondary Structure analysis using Vienna

The Vienna RNA suite was used to comparatively asses the RNA secondary structures further, through the visualisation of the locality and the precise nucleotide changes which have arisen as a result of mutagenesis. In comparison to the results produced from RNAfold independently, distinct changes were identifiable from the MFE secondary structure images for both rna123 and rna154 post magnification (Figures 5.8 and 5.9).

When looking at the Vienna analysis of rna123, the overall secondary structure appears highly conserved (Figure 5.8). However, following the magnification of the region containing the SNPs, alterations were recognised within two stems of the secondary structure.

Similar results were seen with rna154 with highly comparable MFE structures prior to magnification for the sensitive and resistant strains (Figure 5.9). Following magnification, structural changes were observed predominantly within a single stem of the secondary structure. Intriguingly, the shared mutation at position 226 is found at identical locations within highly conserved secondary structures for the two rRNA variants. These results suggest that the specific mutation identified within two of the seven rRNA variants, may be significant in the acquisition of gentamicin resistance.

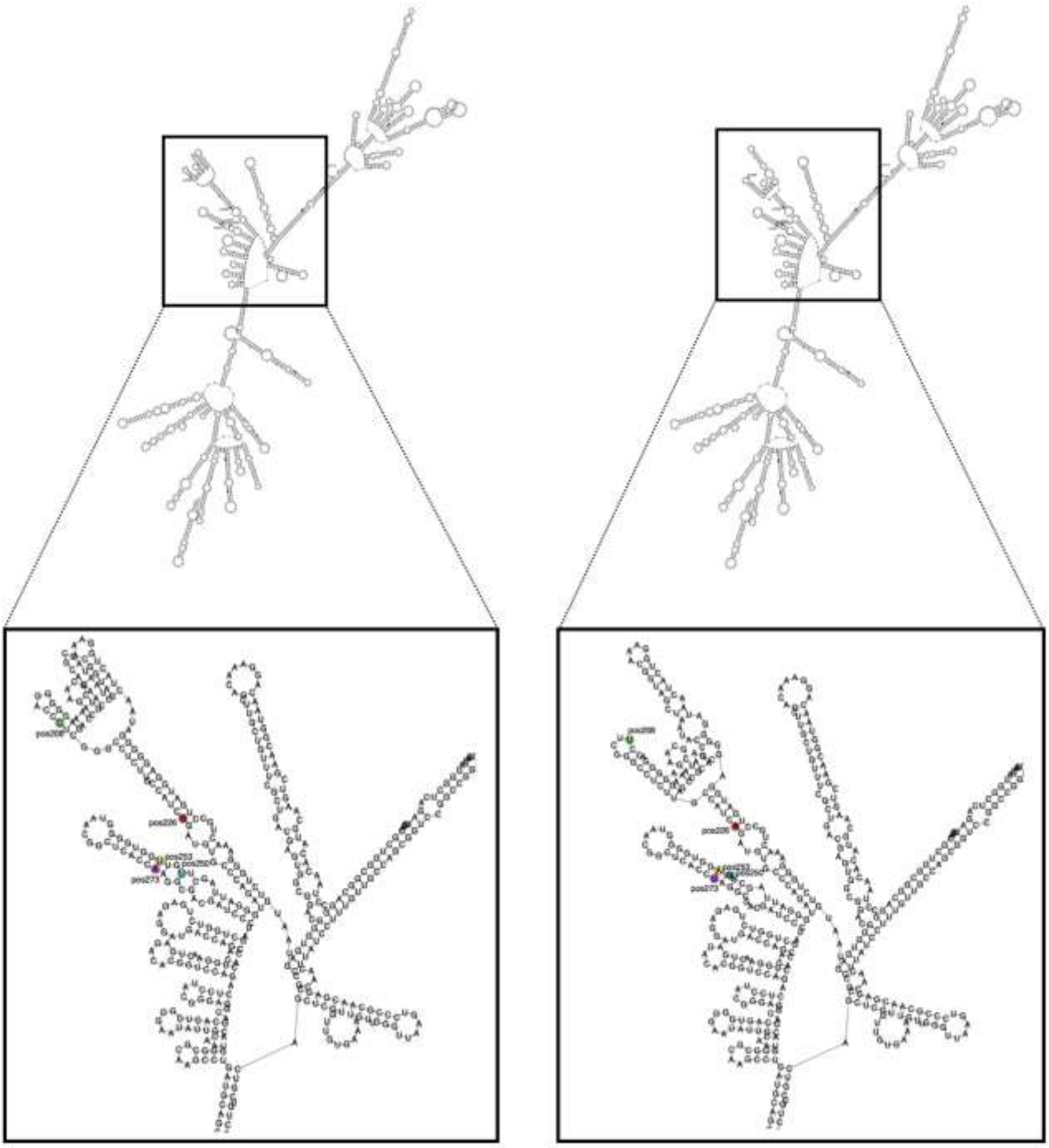


Figure 5.8 RNA secondary structure analysis of rRNA 123 using Vienna. Displaying the MFE secondary structures and magnifications of the rRNA 123 from the sensitive (left) and resistant (right) strains of *E. coli* MG1655. The Vienna package facilitates the identification of the predicted secondary structure alterations between the MFE structures and the annotation of both the mutational position and ribonucleotide base changes. The mutated ribonucleotide bases in the sensitive and resistant strains have been indicated in colour.

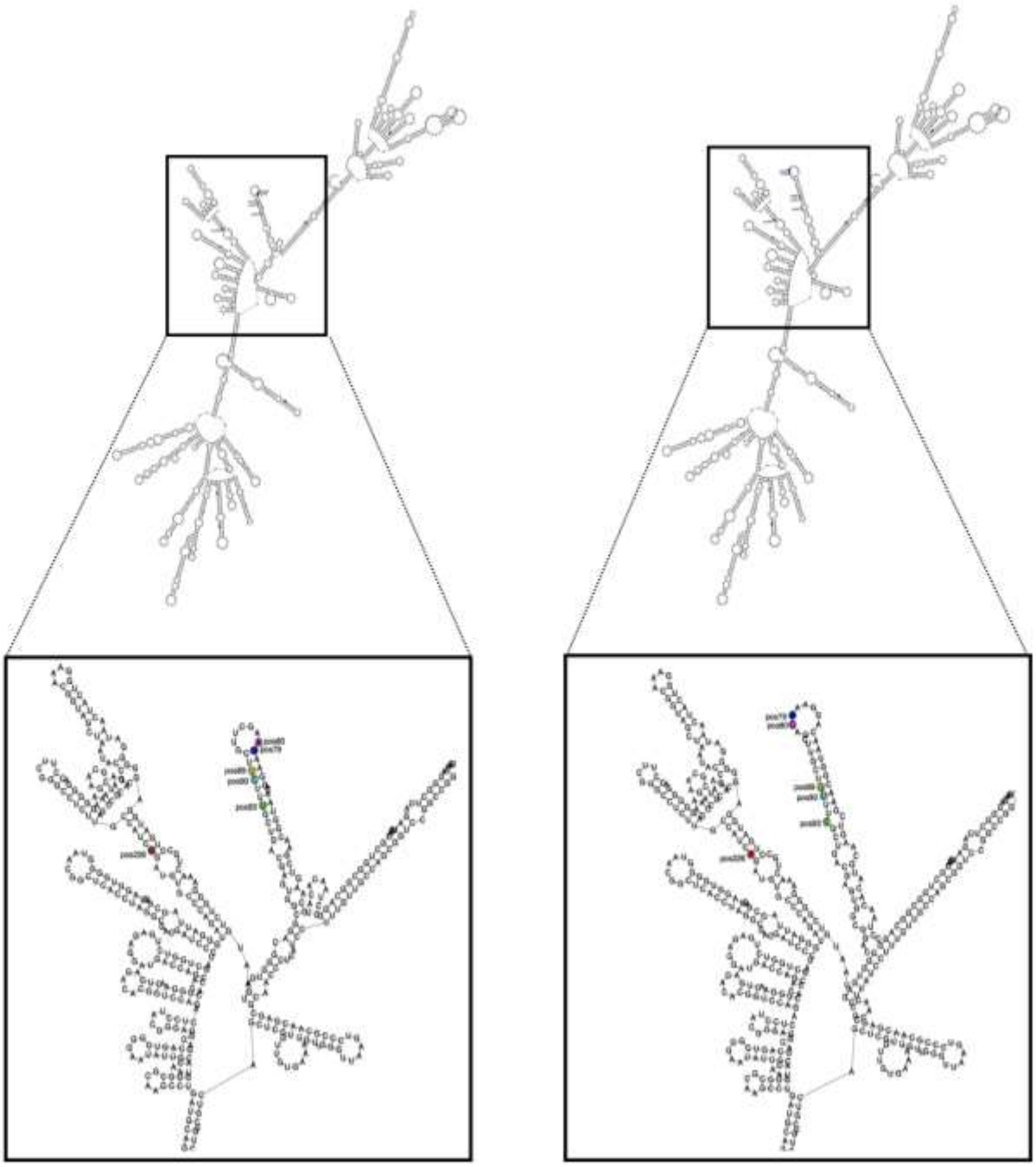


Figure 5.9 RNA secondary structure analysis of rRNA 154 using Vienna. Displaying the MFE secondary structures and magnifications of the rRNA 154 from the sensitive (left) and resistant (right) strains of *E. coli* MG1655. The Vienna package facilitates the identification of the predicted secondary structure alterations between the MFE structures and the annotation of both the mutational position and ribonucleotide base changes. The mutated ribonucleotide bases in the sensitive and resistant strains have been indicated in colour.

5.3.3 RNA sequence polymorphism analysis using DNAsp

Additional rRNA polymorphism analysis was carried out within DNAsp, to assess the nucleotide diversity within the rRNA 123 and 154 genes. The remaining five rRNA were not included in the DNAsp analysis, as they were identical and had nucleotide diversity scores of 0 across the entire sequence length. The nucleotide diversity scores were either 0 or 1 depending on the presence or absence of a single nucleotide polymorphisms within the ribonucleotide sequences. The results suggest a high degree of conservation across the 16s rRNA sequences, with diversity constrained to early segments of the rRNA sequence for the two rRNA variants (Figure 5.10).

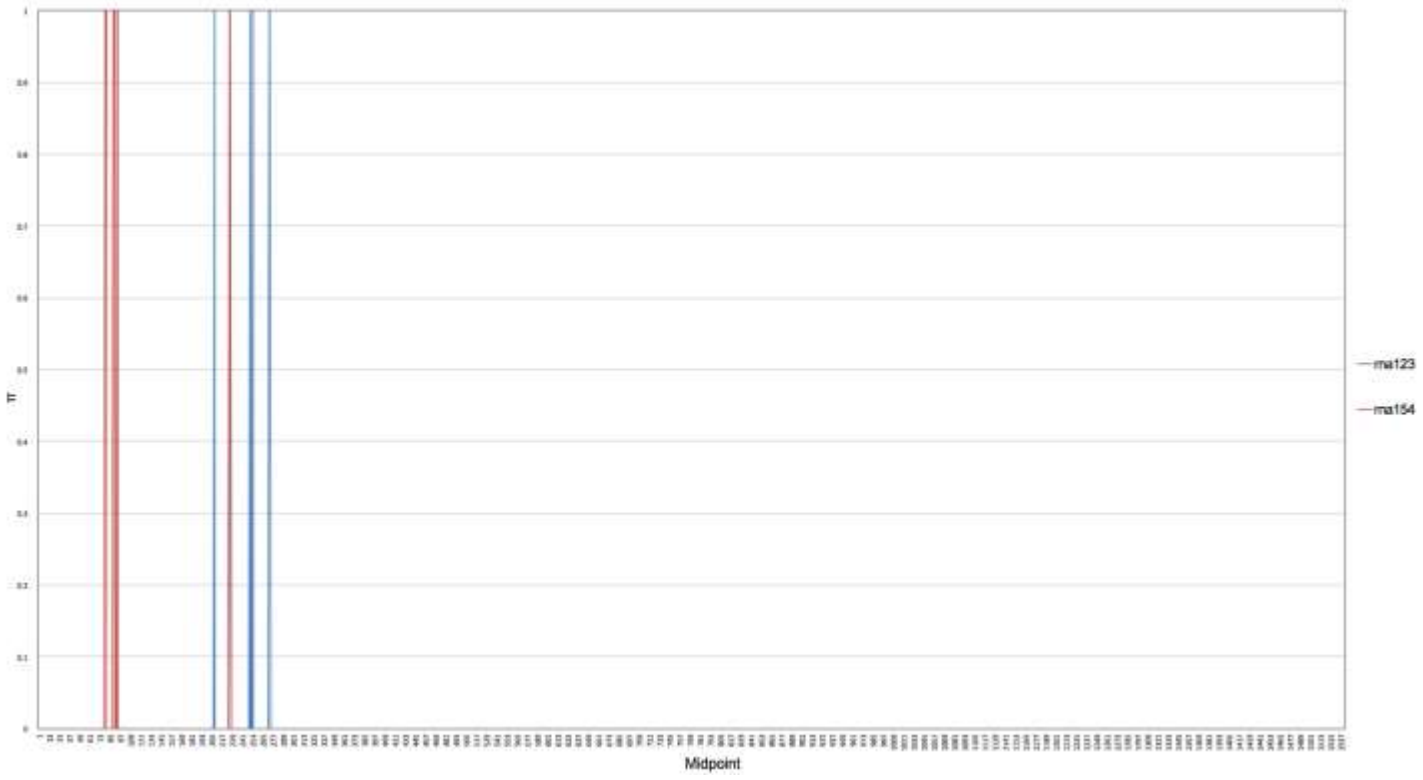


Figure 5.10 Nucleotide diversity of the 16s rRNA 123 and 154 genes. Displaying the mean nucleotide diversity (π) of the 16s rRNA 123 (blue) and 154 (red) genes between the gentamicin sensitive and resistant strain of *E. coli* MG1655 using DNAsp.

5.3.4 Phylogenetic analysis of the rRNA variants

To further assess rRNA conservation and to determine whether the identified rRNA mutations were present within other strains of *E. coli*, the phylogeny of the rna123 and rna154 sensitive and resistant genes were resolved. Firstly, the phylogeny of the rna154 gene was resolved, revealing two clades with the rna154 resistant strain appearing as a lone outgroup within the ML tree (Figure 5.11).

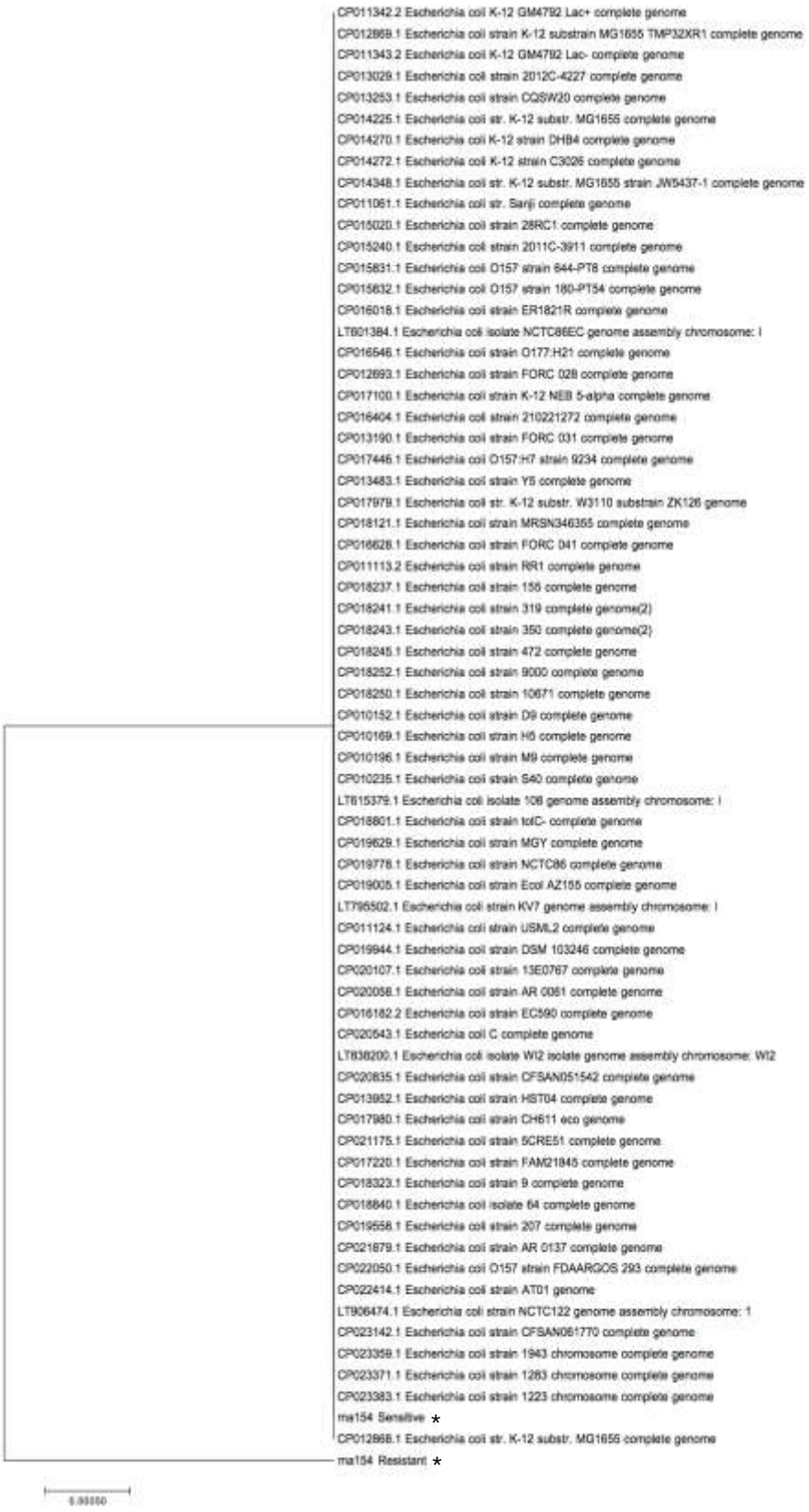


Figure 5.11 Molecular phylogenetic analysis of the 16s rRNA 154 gene. Displaying the resolved phylogeny by maximum likelihood of the rna154 gene within the *Escherichia* taxa using MEGA. The gentamicin sensitive and resistant rna154 isolates have been emphasised using an asterisk.

In comparison, the ML tree revealed similar tree topology for the rRNA 123 gene, with two discrete clades bootstrapped at 99%. There was a clear distinction between the rna123 sensitive and resistant gene sequences, clustering within two separate clades. However, an apparent difference between the rRNA 154 tree, was that the rRNA 123 resistant sequence does not appear as an outgroup, instead the sequence is clustered amongst other *E. coli* strains (Figure 5.12). Upon further inspection within NCBI of the strains which clustered with the resistant rna123 sequence, strains were found to be either genomically modified, or associated with antimicrobial tolerance and/or resistance (Table 5.2). This result indicates that the identified rRNA 123 mutations could be markers associated with the acquisition of antimicrobial resistance.

To conclude, the phylogenetic analysis shows high conservation for the two ML trees, with a single monophyletic unit split into two subgroups with a cladistics distinction of the presence of the resistant rRNA sequence.

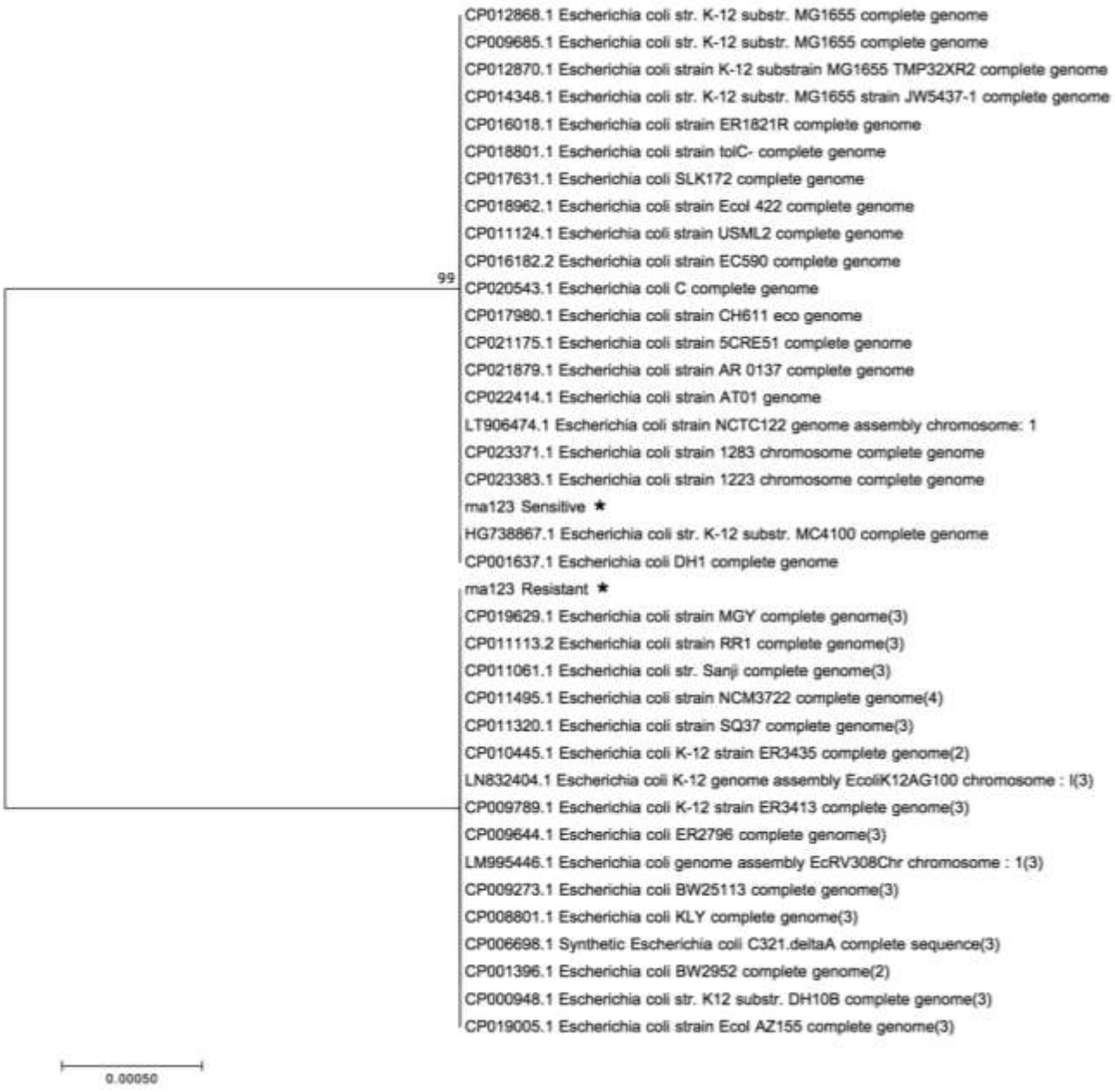


Figure 5.12 Molecular phylogenetic analysis of the 16s rRNA 123 gene.
Displaying the resolved phylogeny by maximum likelihood of the rna123 gene within the *Escherichia* taxa using MEGA. The gentamicin sensitive and resistant rna123 isolates have been emphasised using an asterisk.

Table 5.2 Analysis of the *E. coli* strains which clustered with the gentamicin resistant rna123 isolate. Table displaying the details of the *E. coli* strains which clustered with the rna123 resistant isolate within the phylogenetic analysis.

Strain	Strain ID	Description/Projects
Escherichia coli strain MGY	CP019629.1	Antimicrobial Tolerance/Resistance Evolution
Escherichia coli strain RR1	CP011113.2	Diagnosis and control of Antibiotic Resistant bacteria
Escherichia coli strain Sanji	CP011061.1	Antibiotic Resistant Strain
Escherichia coli strain NCM3722	CP011495.1	Metabolic Deficit Strain
Escherichia coli strain SQ37	CP011320.1	rRNA Operon Deletion Strain
Escherichia coli strain ER3435	CP010455.1	Genomically Recoded Strain
Escherichia coli strain K12AG100	LN832404.1	Sub-lethal Antimicrobial Selection
Escherichia coli strain ER3413	CP009789.1	Methyltransferase Deficient Strain
Escherichia coli strain ER2796	CP009644.1	Methyltransferase Deficient Strain
Escherichia coli strain EcRV308	LM995446.1	Monoisolate WGS
Escherichia coli strain BW25113	CP009273.1	Keio Collection Knockout Mutant (rRNA Deletion)
Escherichia coli strain KLY	CP008801.1	Antimicrobial Tolerance
Escherichia coli strain C321	CP006698.1	Evolution/Genetically Recoded Strain
Escherichia coli strain BW2952	CP001396.1	Strain malG-lacZ fusion/ Mutations/Recombination
Escherichia coli strain DH10B	CP000948.1	Common Laboratory Strain
Escherichia coli strain AZ155	CP019005.1	Global Surveillance of Antimicrobial Resistance

Following the initial assessment of the rRNA mutations which form structural constituents of the ribosome, critical in the processing of genetic material and the production of proteins. The remainder of the results will be focused on the identified mutations within genes encoding proteins, for the *fusA* and *pinR* genes respectively.

5.3.5 EMBL SAPS protein charge analysis

The first assessment of protein property was achieved using EMBL SAPS, to analyse the protein charge distribution following mutagenesis. The *fusA* results suggest that at the amino acid mutation position 610 there was no difference in charge pre-and post-mutagenesis with both the amino acids proline and threonine being of neutral charge (Figure 5.13).

A

```
CHARGE DISTRIBUTIONAL ANALYSIS

     1  00+00000+0 +00000000- 00+0000-+0 00000000+0 0-00-00000 -00-0-0-+0
    61  0000000000 000000+00- 00+0000-00 000-000-0- +00+00-000 0000000000
   121  000-000+00 0+0+00+000 00+0-+0000 00+00000+0 +000000000 00000--000
   181  000-00+0+0 0000-0-000 00-0--000- 00-000-000 000-000-00 --00-+0000
   241  --00-0-0+0 00+0+0000- 0000000000 +0+000000- 000-000000 -00000000-
   301  -0+-000-+0 00---00000 00+000-000 000000+000 000000-000 000+00+-+0
   361  0+0000000+ +--0+-0+00 -0000000+- 0000-000-0 -00000-+0- 00-0000000
   421  -0+0+0-0-+ 000000+00+ --000+000- --00000000 00-000-000 -+0++-000-
   481  0000+00000 +-00+0+00- 0-0+00+000 0+00000000 -0000-0000 0+00-000-0
   541  +000000-00 000-+000-0 0+00000000 00-000+000 0000-0-00- 000+000000
   601  0+-00++0+0 000-000+0- 0-00--000- 000-00+++0 00+00-0-00 00+000-000
   661  0-00000000 +000+0+000 00-00+0--0 000000000- 0+0+
```

B

```
CHARGE DISTRIBUTIONAL ANALYSIS

     1  00+00000+0 +00000000- 00+0000-+0 00000000+0 0-00-00000 -00-0-0-+0
    61  0000000000 000000+00- 00+0000-00 000-000-0- +00+00-000 0000000000
   121  000-000+00 0+0+00+000 00+0-+0000 00+00000+0 +000000000 00000--000
   181  000-00+0+0 0000-0-000 00-0--000- 00-000-000 000-000-00 --00-+0000
   241  --00-0-0+0 00+0+0000- 0000000000 +0+000000- 000-000000 -00000000-
   301  -0+-000-+0 00---00000 00+000-000 000000+000 000000-000 000+00+-+0
   361  0+0000000+ +--0+-0+00 -0000000+- 0000-000-0 -00000-+0- 00-0000000
   421  -0+0+0-0-+ 000000+00+ --000+000- --00000000 00-000-000 -+0++-000-
   481  0000+00000 +-00+0+00- 0-0+00+000 0+00000000 -0000-0000 0+00-000-0
   541  +000000-00 000-+000-0 0+00000000 00-000+000 0000-0-00- 000+000000
   601  0+-00++0+0 000-000+0- 0-00--000- 000-00+++0 00+00-0-00 00+000-000
   661  0-00000000 +000+0+000 00-00+0--0 000000000- 0+0+
```

Figure 5.13 Charge distributional analysis of the *fusA* encoding EF-G protein. Displaying the charge distribution analysis of the EF-G protein sequence between the gentamicin sensitive (A) and resistant (B) strains of *E. coli* MG1655 using EMBL SAPS. The amino acid mutational position is indicated using an asterisk. There is no difference in the charge distribution of the EF-G protein because of the mutation at the amino acid position 610.

When comprising this to the charge distribution of the *pinR*, a change in charge was identified at amino acid position 3. Here the amino acid arginine within the sensitive strain is positively changed and through mutagenesis is replaced with a glutamine residue of neutral charge (Figure 5.14).

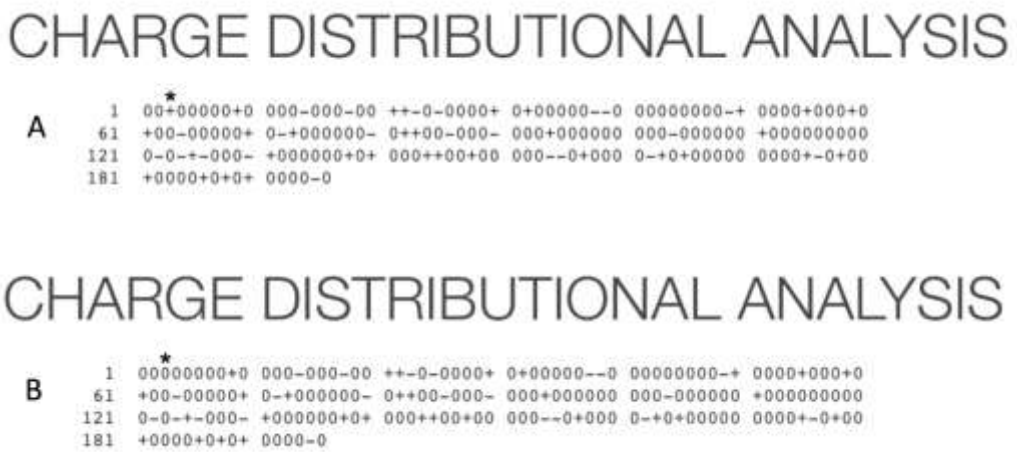


Figure 5.14 Charge distributional analysis of the *pinR* encoding serine recombinase protein. Displaying the charge distribution analysis of the serine recombinase protein sequence between the gentamicin sensitive (A) and resistant (B) strains of *E. coli* MG1655 using EMBL SAPS. The amino acid mutational position is indicated using an asterisk. The *pinR* mutagenesis has resulted in an alteration in the charge distribution at the amino acid position 3.

The SAPS analysis of the two mutated proteins revealed no charge clusters, no high scoring uncharged segments and no unusual spacing within the amino acid sequences.

5.3.5.1 TreeSAAP selection on amino acid properties

In addition to charge distribution the mutated proteins were analysed for alterations in physiochemical properties using TreeSAAP. The TreeSAAP analysis of the *fusA* gene showed significant results within two proteins properties; compressibility and power to be at C-terminus. There was a greater significance in the alteration of protein compressibility post mutagenesis, with a Z score within category 6 of 3.9*** ($p<0.001$) (Figure 5.15). When comparing this to power to be at C-terminus, there was a higher z-score of 4.028*** of equal significance ($p<0.001$). However, this score was obtained within magnitude category 5 (Figure 5.16). These results suggest, the mutated amino acid is located at the C-terminus and causes significant alterations in the protein compressibility.

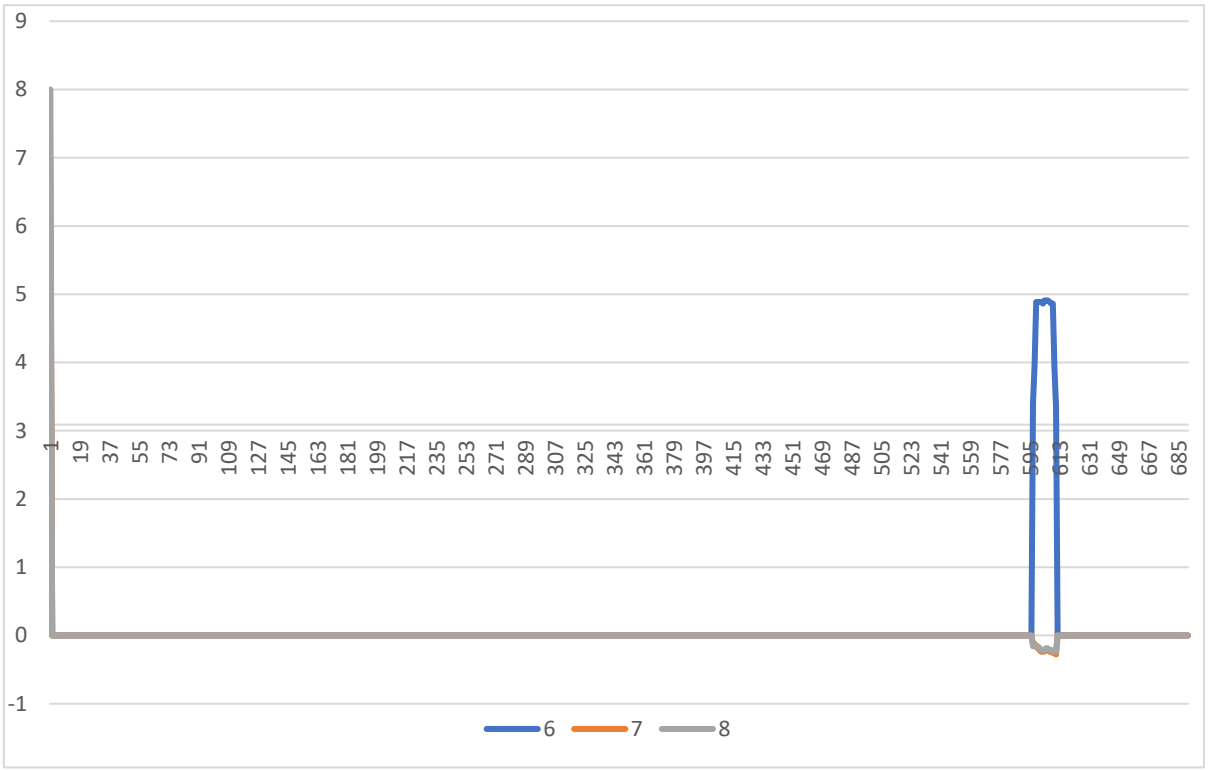


Figure 5.15 Protein compressibility: protein property under positive selection within EF-G. Figure displaying the TreeSAAP analysis of the *fusA* encoding EF-G protein of the drug sensitive and resistant *E. coli* MG1655 strains. The graph displays the frequency of a protein property under selection across the length of the protein sequence There was a significant alteration in protein property of protein compressibility at the mutation position with a Z score of 3.9 ($p<0.001$) within category 6.

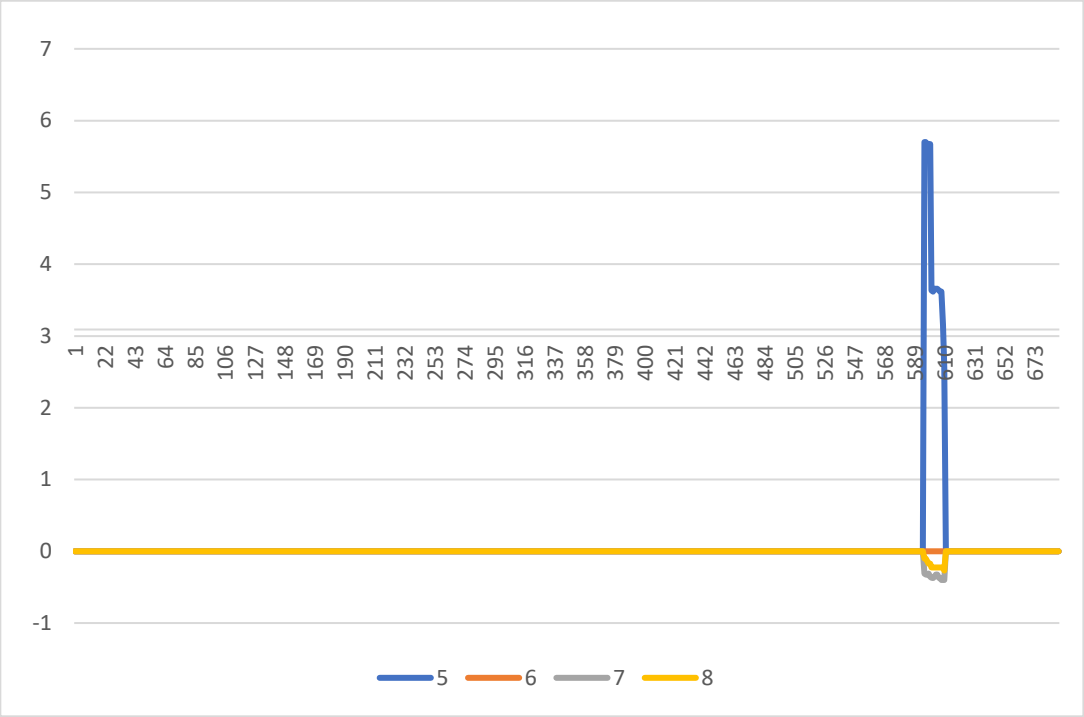


Figure 5.16 The power to be at C-terminus: protein property under positive selection within EF-G. Figure displaying the TreeSAAP analysis of the *fusA* encoding EF-G protein of the drug sensitive and resistant *E. coli* MG1655 strains. The graph displays the frequency of a protein property under selection across the length of the protein sequence. There was a significant alteration in protein property of power to be at C-terminus at the amino acid mutation position with a Z-score of 4.028 ($p<0.001$) within category 5.

The TreeSAAP analysis of the *pinR* gene showed a significant result with one protein property; isoelectric point. Post mutagenesis there was a significant change in the isoelectric point at the mutational position, with a Z score within category 7 of 2.876** ($p<0.01$) (Figure 5.17). This may reflect the alteration in charge identified previously using SAPS (Section 5.3.5). The significance of the observed alteration in protein property within *pinR* was lower than seen for the two alterations in protein property within the *fusA* gene.

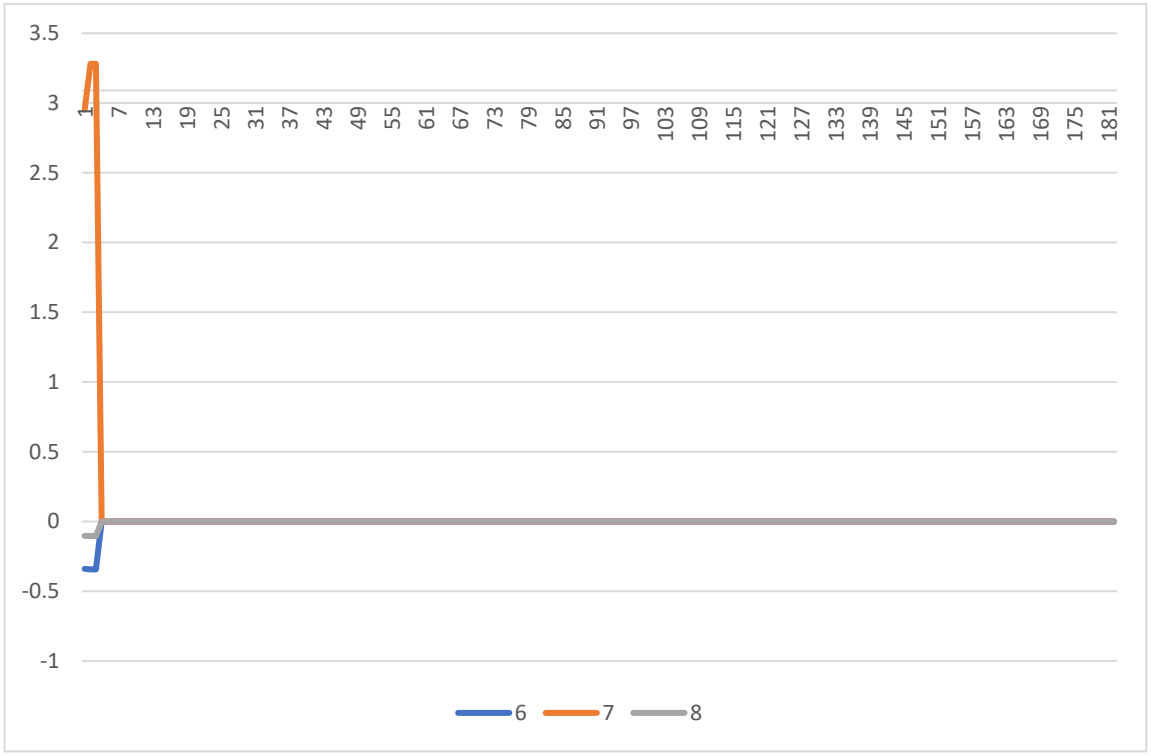


Figure 5.17 Isoelectric point: protein property under positive selection within the *pinR* serine recombinase. Figure displaying the TreeSAAP analysis of *pinR* serine recombinase protein of the drug sensitive and resistant *E. coli* MG1655 strains. The graph displays the frequency of a protein property under selection across the length of the protein sequence. There was a significant alteration in the proteins isoelectric point at the amino acid mutation position with a Z-score of 2.876 (P<0.01) in category 7.

The observed alterations in protein properties are indicative of amino acid adaptation as a result of antimicrobial selection. Additional functional analysis conducted within this chapter will aim to characterise the extent of these adaptations, and their importance in the evolution of gentamicin resistance.

5.3.5.2 Conserved protein domain analysis

The conserved domain analysis of the two mutated proteins was achieved using NCBI. The summary of the conserved domain analysis of the *fusA* sequence revealed 3 superfamilies (Figure 5.18). The GTP-EFU superfamily spans the entire length of the *fusA* sequence (1-2112). The remaining 2 superfamilies include the P-loop_NTPase superfamily and the EF-G_like_IV superfamily. The P-loop_NTPase superfamily is located at the nucleotide interval of 34-867 in the region of conserved functional sites including the GTP binding site. The EF-G_like_IV superfamily appears latter in the *fusA* sequence at an interval of 1495-1827. Together, GTP_EFTU and EF-G_like_IV form fundamental processes within ribosomal protein synthesis. Of particular importance is that the identified *fusA* mutation at the nucleotide interval 1828, occurs at the terminal boundary of the conserved EF-G_like_IV superfamily. All the non-specific hits for the *fusA* nucleotide sequence can similarly be identified in figure (5.18).

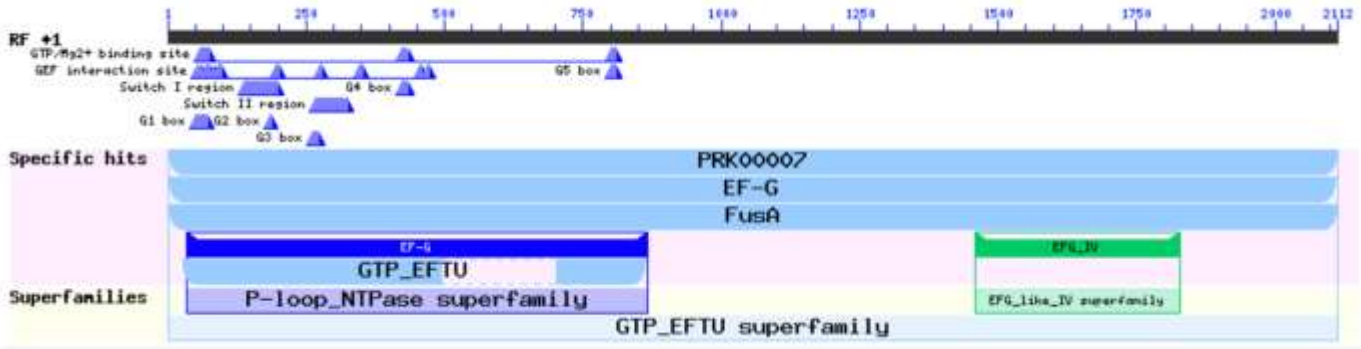


Figure 5.18 Conserved domain analysis of the *fusA* encoding EF-G protein. Displaying the graphical summary from the NCBI conserved domain analysis of the EF-G query sequence. The domains are colour coded per the assigned superfamilies. The functional sites of the protein are mapped to the query sequence and indicated as triangles of the same colour of the domain or superfamily providing the functional activity. The P-loop_NTPase superfamily comprises the functional sites of EF-G. The EF-G like_IV superfamily is of positional significance with regards *fusA* mutation.

In comparison to the conserved domain analysis within the *pinR* nucleotide sequence. In total one superfamily was identified; the Serine Recombinase superfamily, with a specific hit to PinE at an interval of 7-573 covering much of the nucleotide sequence (Figure 5.19). Further specific hits included SR_ResInv and Resolvase at an approximated interval of 10- 440 responsible for site specific recombination. The conserved domain analysis revealed that nucleotide position of the mutation within the *pinR* gene had a specific hit with PinE yet, was in a region containing no conserved functional sites. This result suggested that the identified mutation within *pinR* may be less relevant in the acquisition of gentamicin resistance than the mutation within the functional domain of *fusA*.

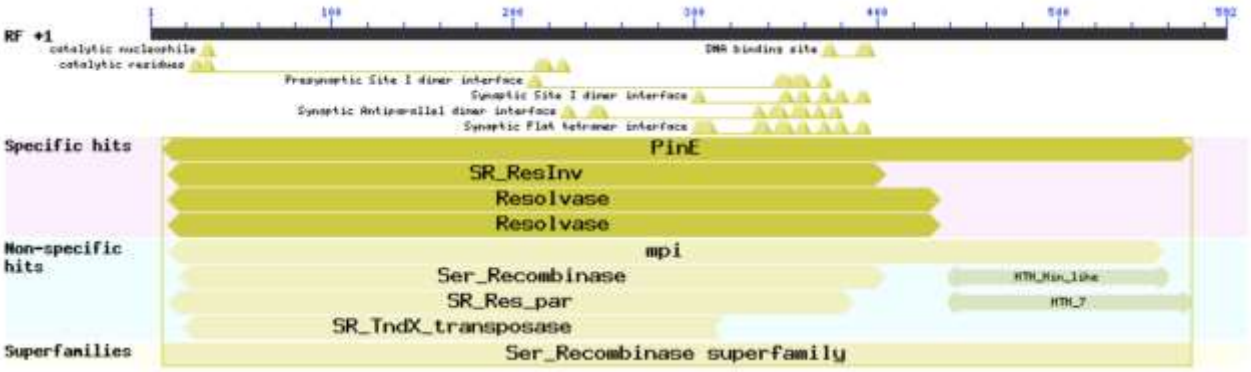


Figure 5.19 Conserved domain analysis of the *pinR* encoding serine recombinase protein. Displaying the graphical summary from the NCBI conserved domain analysis of the serine recombinase query sequence. The domains are colour coded per the assigned superfamilies. The functional sites of the protein are mapped to the query sequence and indicated as triangles of the same colour of the domain or superfamily providing the functional activity. The functional sites of the *pinR* are sequence are annotated within the Ser_Recombinase superfamily.

5.3.6 Protein modelling using Swiss Model

The best available template for the *fusA* gene sequence was the 4v7d.1 model, with a sequence identity of 100% for the sensitive strain (Figure 5.20 A). The resistant gene sequence in comparison had a sequence identity of 99.86% (Figure 5.20 B). Looking comparatively at the two models no obvious changes were present at the three magnifications, excluding the amino acid change at the point of mutagenesis.

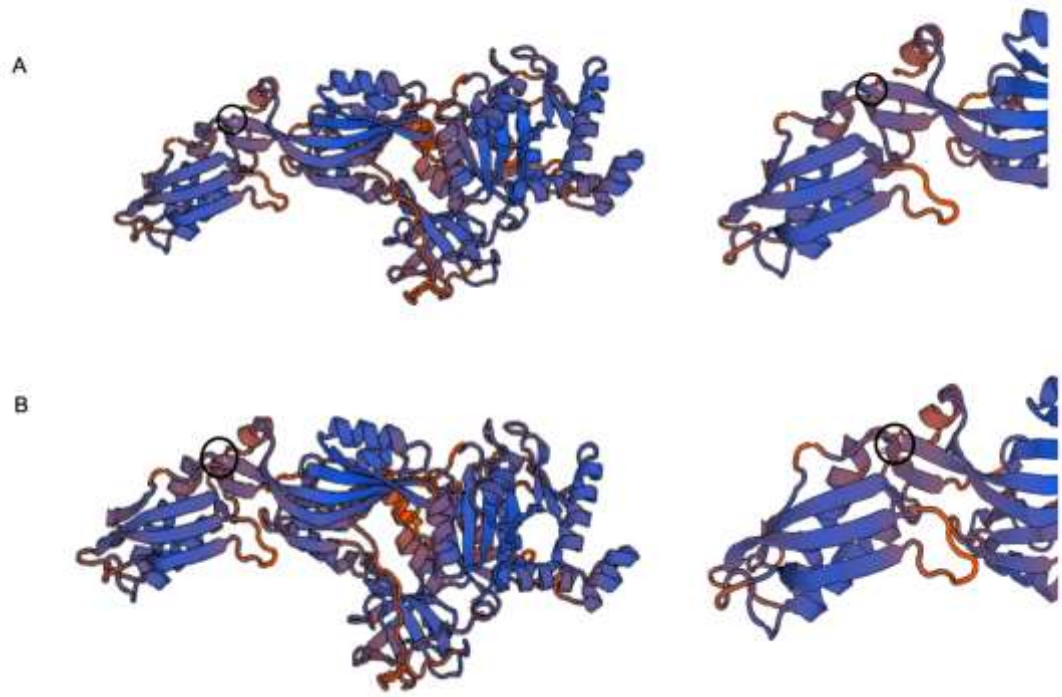


Figure 5.20 Protein homology modelling of EF-G. Displaying the protein models produced within Swiss Model of the EF-G protein within the sensitive (A) and resistant (B) strains of *E. coli* MG1655. The figure shows the overall structure of EF-G and a magnification of the protein model to facilitate the identification of the amino acids at the site of mutagenesis (circled in black).

When comparing this to the homology model for the *pinR* gene sequence. The best available template was the Igdt.1. E model with a sequence identity of 31.32% for both the sensitive and resistant *pinR* sequences, despite the presence of a non-synonymous mutation. Upon closer inspection, all the available model templates for the *pinR* sequence within Swiss Model, did not comprise the first three amino acids and therefore any structural alterations were unidentifiable. There is an increasing number of incomplete crystal structures being deposited to the protein data bank, therefore missing N-terminal strings within a polypeptide chain of a putative protein is not unexpected (Djinovic-Carugo and Carugo, 2015). A single image was included representative of the general structure of the putative serine recombinase for both sequences (Figure 5.21).

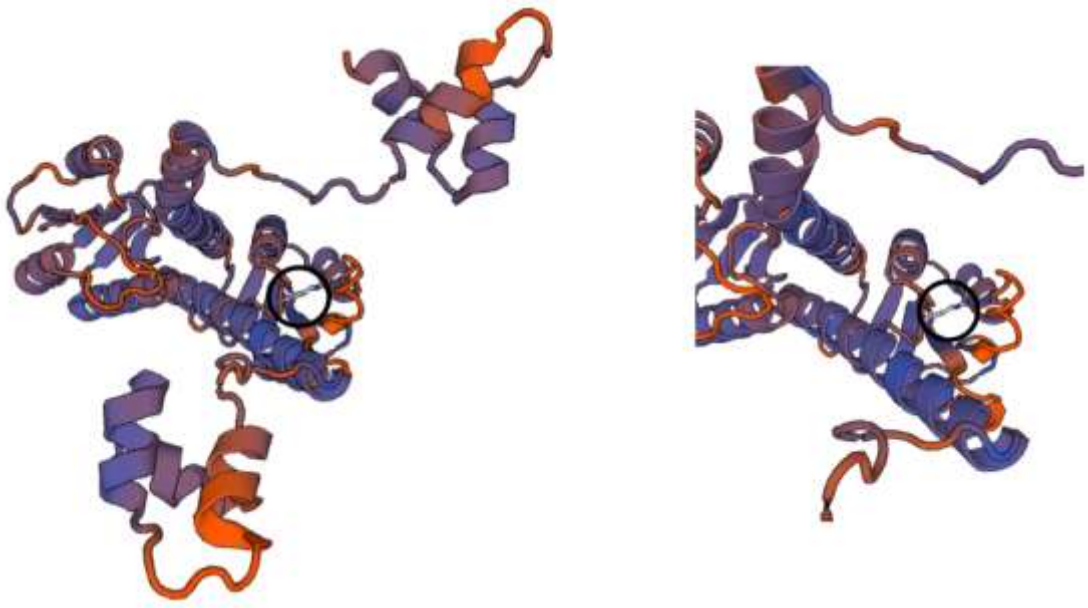


Figure 5.21 Protein homology modelling of Serine Recombinase *pinR*. Displaying the protein model produced within Swiss Model of the Serine Recombinase *pinR* protein within the sensitive strain of *E. coli* MG1655. The figure overall protein structure and a magnification of the protein model to facilitate the identification of the amino acid at the site of mutagenesis (circled in black). The protein model of the resistant *E. coli* MG1655 strain was not included in this image.

This concluded the functional analysis of the *pinR* encoding putative serine recombinase, in the interest of the *fusA* EF-G encoding mutation having greater involvement in the acquisition of gentamicin resistance.

5.3.6.1 Protein modelling using Scigress

The PDB molecules of the *fusA* encoding EF-G protein produced using Swiss Model, were imported into Scigress as chemical sample files (.csf), to geometrically optimise protein structures. The respective protein models were cleaned and beautified using the Valence Shell Electron Pair Repulsion (VSEPR) theory, prior to the visualisation of the mutated amino acid residue (Wong and Currie, 2012).

When comparing the structural features of the geometrically optimised protein structures, there are evident differences between the sensitive (A) and resistant (B) models (Figure 5.22). This was originally identified when looking in the vicinity of the mutated amino acid residue. Within the sensitive model a distinct coil was present adjacent to residue of interest, which was drastically loosened within the resistant model post mutagenesis. Subsequently, a comparative assessment of the overall tertiary structures revealed reduced compactness throughout the entire resistant model, which could reflect the reduction in protein compressibility identified using TreeSAAP.

Figure 5.22 Molecular modelling of EF-G. Displaying a comparison of the protein models of EF-G within the sensitive (A) and resistant (B) strains of *E. coli* MG1655 using Scigress. The protein structures are shown in ribbon representation and the amino acids at the site of mutagenesis are represented using ball-and-stick. The image illustrates the alteration protein compressibility within the two protein structures.

5.3.6.2 Functional analysis of EF-G using PyMOL

To comprehend the reduction in compressibility and alteration in tertiary protein structure, it was necessary to gain further insights into the biological function of EF-G within the 70s-ribosomal complex. The PDB files of EF-G bound to the 70s ribosomal complex were provided by Jinzhong Lin, and visualised within PyMOL. EF-G can be seen to undergo two conformational rearrangements during ribosomal translocation (Figure 5.23). For visual purposes EF-G has been presented both within the 70-s complex and independently. During the Pre–translocation state (A), EF-G is found in a compacted form, with unoccupied A (blue), P (purple) and E (pink) site tRNA's. Post-translocation (B) EF-G undertakes a conformational change to its elongated state occupying the A site tRNA (Lin *et al.,* 2014). The coordinated movement of the peptidyl tRNA from the A to the P site and the resultant movement of the mRNA initiated by EF-G, forms the basis of protein synthesis (Frank and Agrawal, 2000).

The crystal structures of EF-G in complex with the ribosome provided by Lin *et al*. (2014) were utilised for both the protein modelling and molecular docking analysis within this thesis. As despite the availability of previously published cryo-electron microscopy (cryo-EM) and X-ray crystallography structures of ribosome complexes with EF-G (Brilot *et al*., 2013 and Pulk and Cate, 2013), the atomic resolution crystal structures detailed here, in both the pre- and post-translocational states, unveiled a previously unseen transient compact conformation during the pretranslocation complex, which suggested that EF-G exhibits greater conformational flexibility than previously observed (Lin *et al*., 2014). Consequently, to determine the functional significance of *fusA* mutagenesis, structural analysis was conducted on the crystal

structures which provided the most recent insights into the conformational space occupied by EF-G during ribosomal translocation.

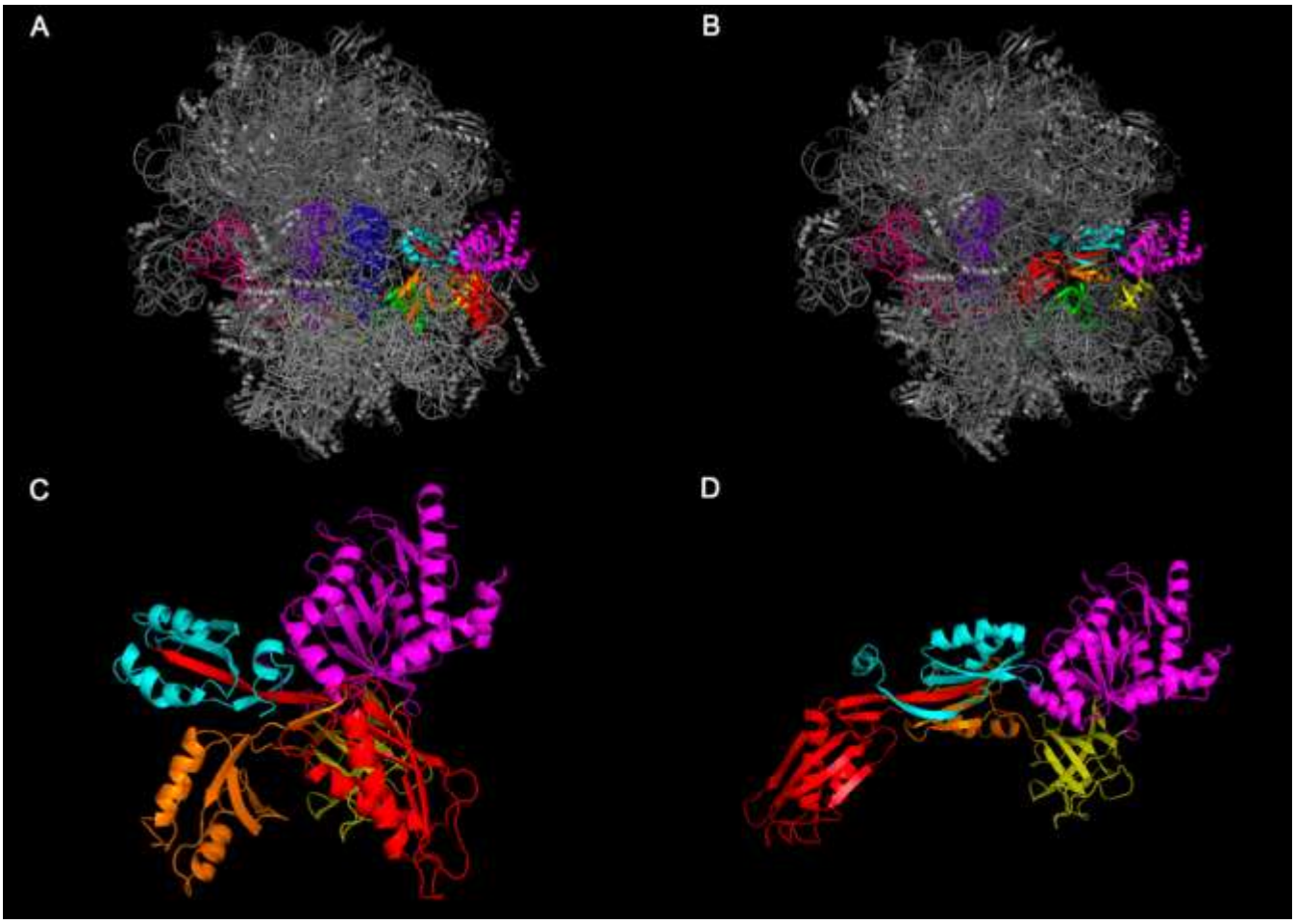


Figure 5.23 Conformational changes of EF-G during ribosomal translocation. Displaying both the pre- (A) and post- (B) 70s-ribosomal complex and the corresponding conformational changes of EF-G from the compact form within the pre-complex (C) to the elongated conformation in the post-complex (D). The 70s-ribosomal subunit was presented in grey with the A-site (blue), P-site (purple) and E-site (pink) TRNA's. The two EF-G conformations were displayed multi-coloured to denote the five individual protein domains.

Above (Figure 5.23), EF-G was presented as five individual segments of differing colours, representative of its five domains. The amino acid boundaries of each domain are shown in figure (5.24). From the functional analysis it is evident that domains 3, 4 and 5 are key to the conformational change of EF-G during ribosomal

translocation. Out of the three domains, domain 4 undergoes the most extensive conformational change with its ‘swivel like’ motion. This domain which was previously acknowledged as a superfamily within the conserved domain analysis. Moreover, the mutation identified is located at the boundary of domains 4 and 5 functioning as a hinge to facilitate the movement of domains.

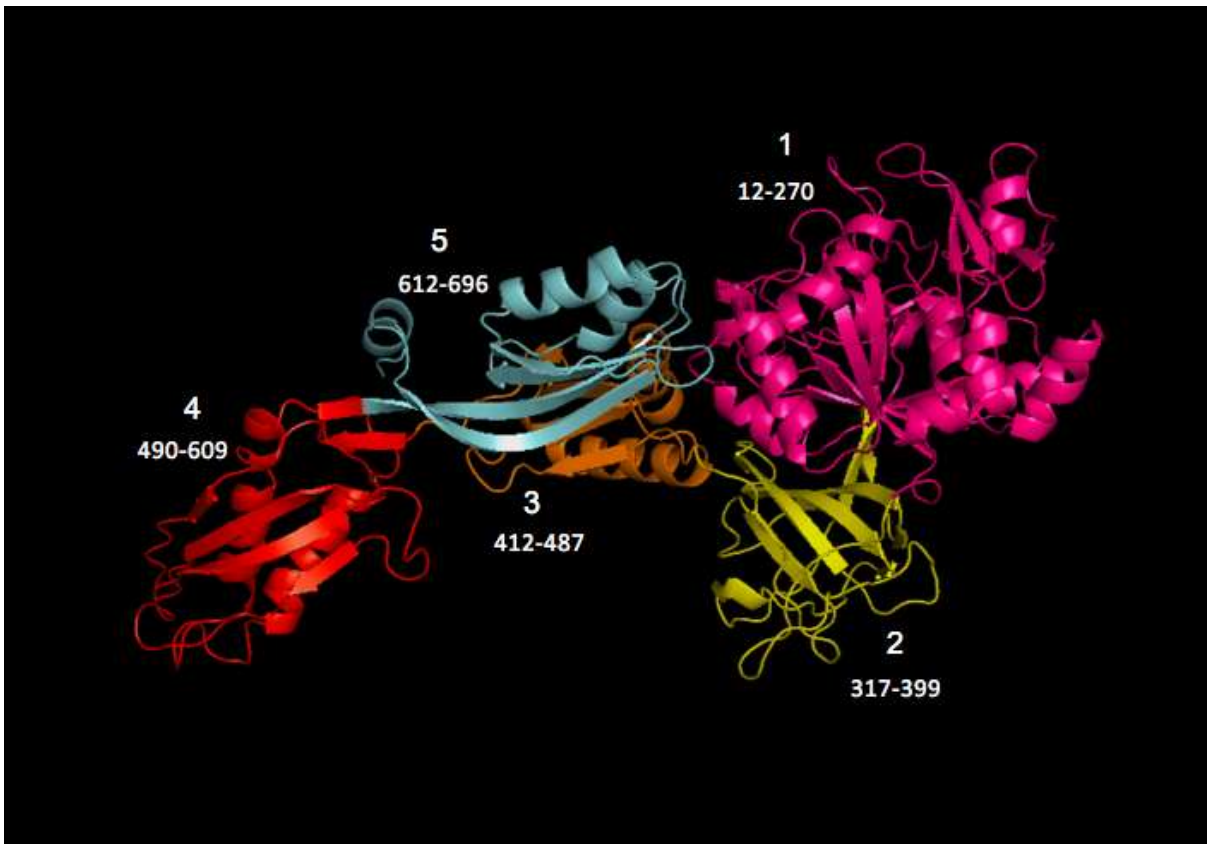


Figure 5.24 The conserved domains of EF-G. Pymol image displaying the five domains of EF-G (domain 1=pink, domain 2=yellow, domain 3=orange, domain 4=red and domain 5=blue) and the corresponding amino acid boundaries of each domain. Each of the five domains form a compact tertiary structure which contributes to the biological function or interaction of EF-G.

The PyMOL mutagenesis function was then used to indicate the position of the mutation, in relation to the domains of EF-G (Figure 5.25). The results here further illustrated the significance of this mutation found at the hinge of domain 4, imperative to the conformational change of this domain during ribosomal translocation. No obvious alterations in the structure of EF-G were observed between the sensitive (A) and resistant (B) strains.

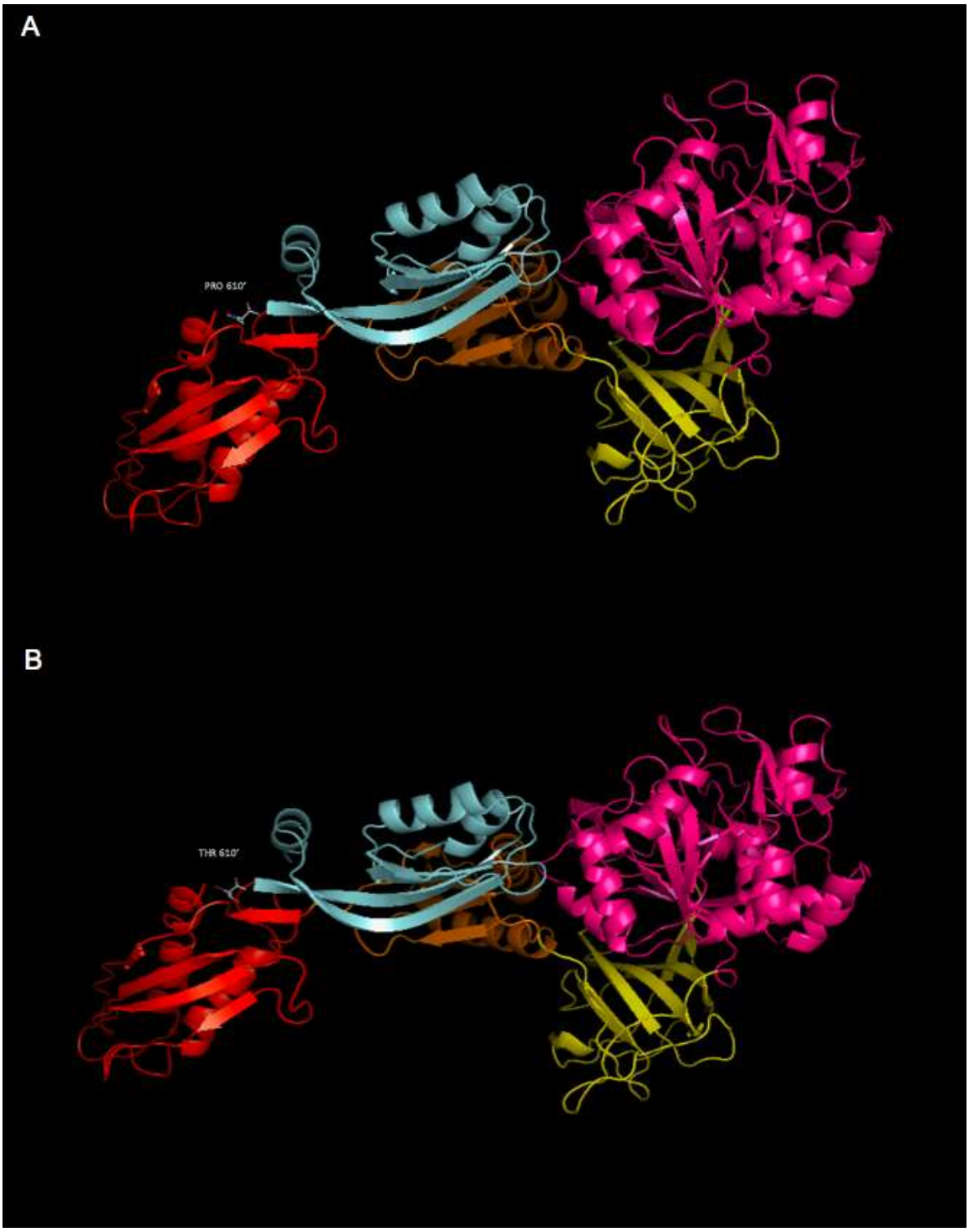


Figure 5.25 PyMOL mutagenesis of EF-G. Pymol image displaying a cartoon representation of the protein structures of the elongated conformation of EF-G within the sensitive (A) and resistant (B) strains of *E. coli* MG1655. The amino acid mutations were illustrated using the ball-stick-representation and the mutational position has been annotated. The five domains of EF-G (domain 1=pink, domain 2=yellow, domain 3=orange, domain 4=red and domain 5=blue) have been highlighted to define the mutational domain. The amino acid mutation at 610 occurs at the boundary of domain 4 and 5 of EF-G.

Further Mutational analysis revealed that the specific amino acid at position 610 pre-mutagenesis, was likely to be essential for the extreme conformational change of domain 4 during translocation (Figure 5.26). Proline residues are referred to as Imino acids with a distinctive ability to associate with the protein backbone in two locations. This contrast to the other amino acids, means proline residues are often associated with polypeptide chains which require tight changes in direction (Barnes, 2003). Therefore, mutagenesis of a proline residue, which owing to its distinctive cyclic structure exhibits unique conformational rigidity, with an amino acid such as threonine, would hinder the structural rigidity at the molecular hinge of domain 4 of EF-G during the conformational transition from a compact to elongated state (Betts and Russell, 2003).

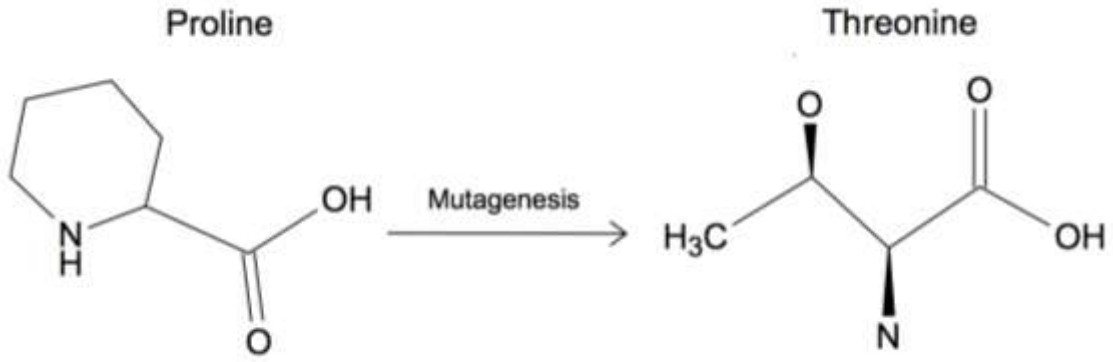


Figure 5.26 The amino acid alteration in the mutagenesis of EF-G. Displaying the 2D molecule sketch of the amino acid change from a proline to a threonine residue during mutagenesis of the *fusA* encoding EF-G. The nitrogen containing ring within the proline residue, is important for conformational changes within protein structures.

5.3.7 PCR amplification and sequencing of genes with mutations

PCR reactions were conducted on the mutated genes most likely to be associated to the acquisition of gentamicin resistance which included both the rRNA genes and the EF-G encoding *fusA* gene.

The initial PCR reactions were carried out on colonies selected at antimicrobial concentrations of 0, MIC and x10 MIC. Colonies were selected at the three concentrations for all primer sets, and were successfully amplified by PCR (Appendix figure A.18) and sequenced at the Natural History Museum illumina sequencing facility.

Successively, sequences were trimmed and aligned prior to analysis. In addition to the sequenced data, comparable sequences of the rRNA and *fusA* genes were added to the alignment to ensure any variations observed were not due to sequencing errors. The analysis of the resequenced rRNA123 genes, displayed conservation across various nucleotide positions within the rRNA sequence. Of interest, the resistant associated mutations located at nucleotide positions 208, 250, 253 and 273, were observed in all resequenced rna123 strains at the two selected gentamicin concentrations (2mg/L and 20mg/L). Further to this, equivalent findings to the phylogenetic analysis of rna123 were observed, with identical polymorphisms being observed in a cluster of *E. coli* strains (Figure 5.27).

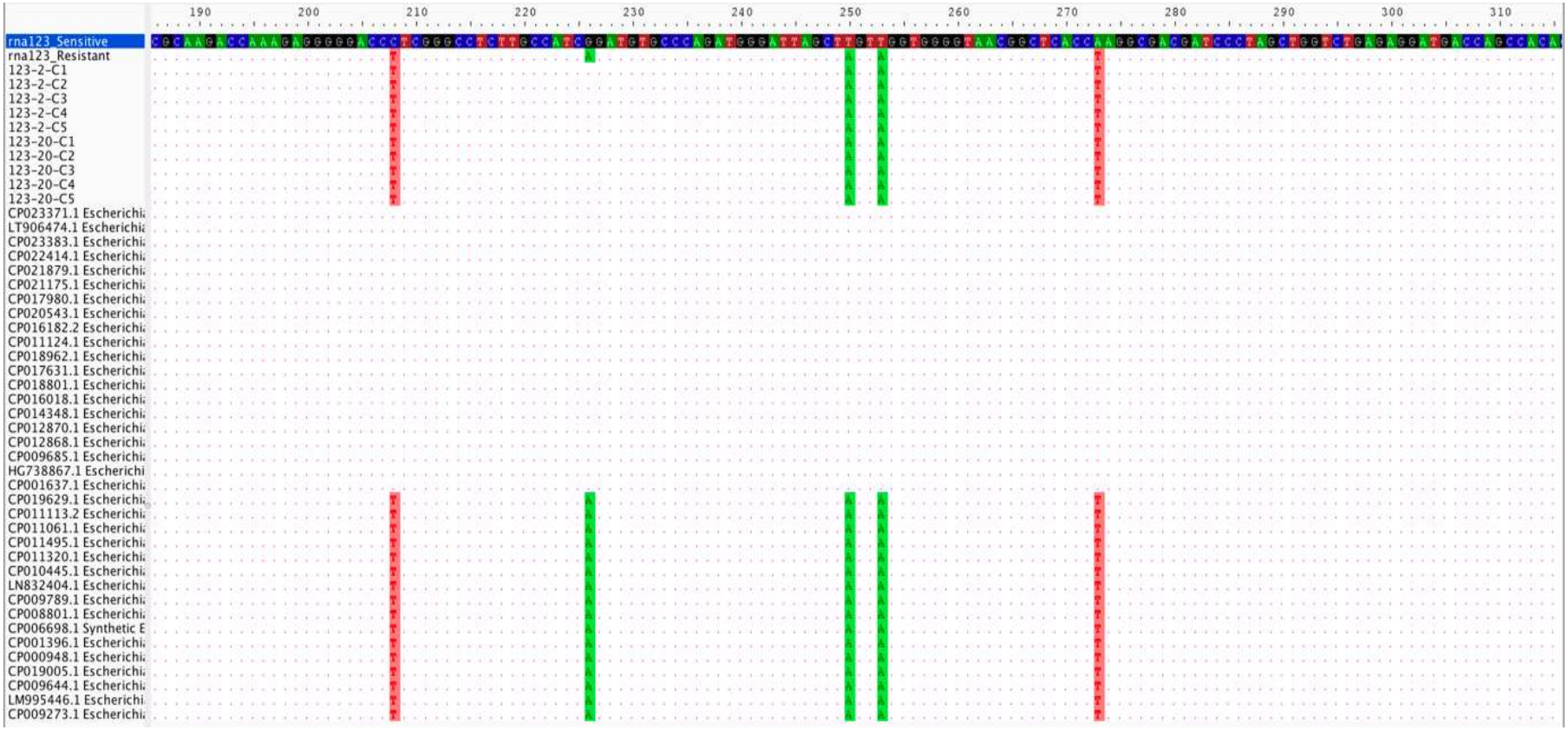


Figure 5.27 The resequenced ribosomal rna123 gene alignment. Aliview alignment displaying the primary rna123 gene sequences of the sensitive and resistant strains of *E. coli* MG1655 strains and the confirmatory rna123 isolates obtained at intermediate antimicrobial concentrations. The alignment includes the rna123 genes utilised in the phylogenetic analysis. The alignment indicates the independent evolution of SNPs in the rna123 resistant isolates selected at the intermediate drug concentrations. There is also clustering observed with strains of E. coli containing identical mutations, supporting the findings of the phylogenetic analysis. The dots in the alignment are representative of identical bases in the bacterial genomes.

In addition, a single point mutation was found at the nucleotide position 1019 for colony 2 selected at 2mg/L gentamicin. Here, a nucleotide substitution from an adenine to a guanine was identified, which was not detected in any of the other strains of *E. coli*, suggestive of an error in sequencing. Moreover, resequencing data revealed substantial rRNA nucleotide discrepancies at positions 1117-1240 (Appendix figure A.19).

The resequencing results for the rRNA 154 genes displayed similarities with the rna123 sequences with high levels of rRNA conservation across the rna154 sequences. The formerly identified mutation in the resistant strain of *E. coli* MG1655 located at position 231, where the nucleotide guanine was substituted for an adenine, was not identified in any of the resequenced resistant strains of *E. coli*. In addition, a unique mutation at position 930 was discovered in colonies 1 and 4 selected at 2mg/L gentamicin, where again a guanine was substituted for adenine in the two resistant strains (Figure 5.28). These results support the findings from the phylogenetic analysis, as no resistant associated mutations were identified in comparable strains of *E. coli*.

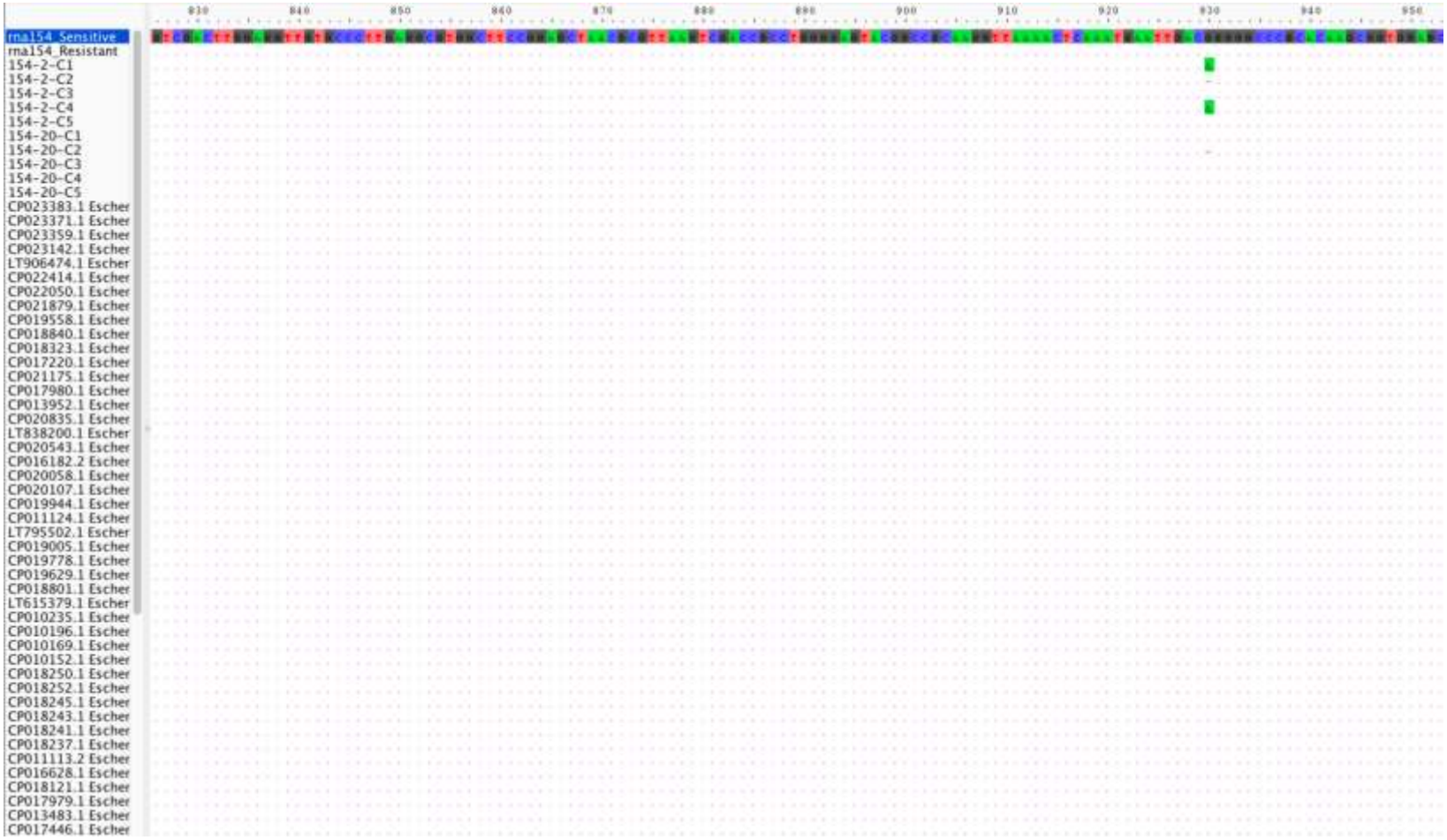


Figure 5.28 The resequenced ribosomal rna154 gene alignment. Aliview alignment displaying the primary rna154 gene sequences of the sensitive and resistant strains of *E. coli* MG1655 strains and the confirmatory rna154 isolates obtained at intermediate antimicrobial concentrations. The alignment includes the rna154 genes utilised in the phylogenetic analysis. The alignment indicates the independent evolution of an SNP in two of the rna54 resequenced isolates at base position 930. There were no mutations identified in the rna154 genes from the phylogenetic analysis. The dots in the alignment are representative of identical bases in the bacterial genomes.

There were nucleotide discrepancies evident in the region of positions 79-98 of the rna154 sequence (Appendix figure A.20). Mutations were formerly identified in this region, in the originally sequenced strain of *E. coli* MG1655. This suggests mutations within this locality despite the gaps within the alignment, may be associated with gentamicin resistance.

In comparison to the resequenced *fusA* isolates, there was a high level of conservation observed, which was maintained across the entire sequence length. An interesting finding here was that the *fusA* mutation at position 1834 was only observed within the 5 colonies selected at the highest antimicrobial concentration (Figure 5.29). As the mutation was not identified within the MIC isolates, this could suggest antimicrobial tolerance at the lower antimicrobial concentration, with the resistant specific SNP only being fixed within the population at the higher 20mg/L concentration.

To determine at what point the *fusA* mutation arises, PCR reactions were completed with colonies selected at intermediate antimicrobial concentrations of 4 and 8 mg/L. All 5 colonies selected at 8mg/L, were successfully amplified using PCR. Though, 2 out of the 5 colonies selected at 4mg/L were unsuccessfully amplified. When viewing the alignments of the intermediated sequences, mutations were identified surrounding the nucleotide position 1834, at positions 1815 and 1860 where in all but one case in both positions a guanine is substituted for an adenine (Figure 5.29). This implied that there are coexisting lineages with a region of mutation surrounding the resistant specific SNP. At position 1860, where guanine was substituted for an adenine, this

lead to no change in amino acid. Whereas, the SNPs located at position 1815 lead to an amino acid change from a glycine to a serine at amino acid position 605 which falls into domain 4. It is likely that the mutation present at 610, is more efficient in function, despite the increasing selective pressure.

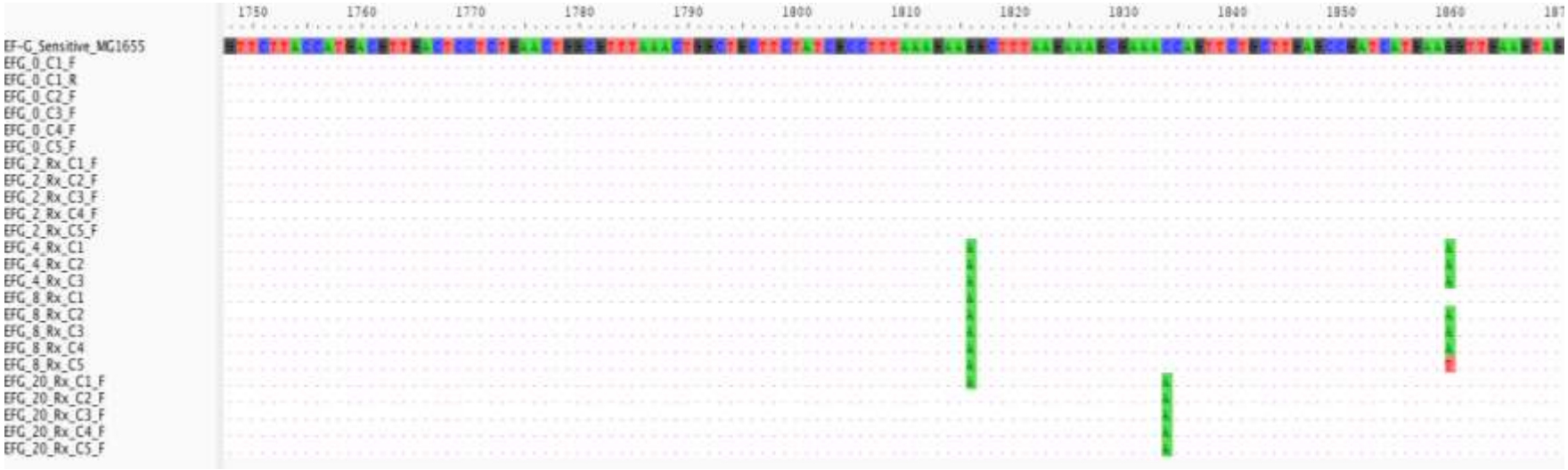


Figure 5.29 The resequenced *fusA* gene alignment. Aliview alignment displaying the primary *fusA* gene sequences of the sensitive and resistant strains of *E. coli* MG1655 strains and the confirmatory *fusA* isolates obtained at intermediate antimicrobial concentrations (4, 8 and 20mg/L gentamicin). The alignment indicates the independent evolution of the *fusA* mutation at the nucleotide position 1834 in the resequenced isolates obtained at 20mg/L gentamicin. There were also mutations identified at position 1815 and 1860 in the resequenced isolates. The dots in the alignment are representative of identical bases in the bacterial genomes.

5.4 Discussion

The research hypothesis described in section (5.1) was supported by the experimental data presented within this chapter as follows. After extensive bioinformatic analysis physiochemical variations were identified in mutated rRNA genes and proteins which related specifically to the functional activity of the aminoglycoside gentamicin. Further to this confirmatory PCR and sequencing revealed the presence of the most significant most likely to be associated with the acquisition of gentamicin resistance in resistant strains, isolated separately within the proposed model and within the traditional serial passage system

5.4.1 Transcriptional regulation and control of the rRNA operon

In bacterial genomes, the rRNA genes are organised into operons, which present typically in several copies within prokaryotic organisms (Pei *et al.*, 2010). The genome of the studied organism *E. coli* K-12 encodes seven copies of the rrn operon named rrnA to rrnH, (Kiss, Sain and Venetianer,1977; Ellwood and Nomura, 1981). It is generally assumed that the presence of multiple copies of the ribosomal RNA (rrn) operon within *E. coli*, is essential for rapid rates of growth (Condon *et al.,* 1992). However, under favourable conditions the maximal growth rates of *E. coli* were maintained following the inactivation of two of the seven rrn operons (Condon *et al.,* 1995). This suggests the multiplicity of the rrn operon, is conserved not only to support maximal growth, but instead to allow bacterial populations to respond to changes in environmental conditions such as resource availability (Davis, Luger and Tai, 1986; Klappenbach, Dunbar and Schmidt, 2000). Therefore, assuming the mutations identified in the two rrn operons are inactivating, the bacterial growth response under selective pressures including antibiotics, could be compromised.

The transcription of rRNA genes is an essential rate-limiting step in ribosomal biosynthesis, and therefore is tightly regulated by several mechanisms to respond to the cellular demands during protein synthesis (Gourse *et al.,* 1996). The initial control of rRNA synthesis is exerted at the level of transcription initiation, by the rrn promoter initiators ATP and GTP and the small molecule effectors initiating nucleoside triphosphate (iNTPs) and nucleotide guanosine5'-diphosphate 3'-diphosphate (ppGpp) which vary significantly in concentration in response to changes in nutritional conditions (Murray and Schneider and Gourse, 2003). The rRNA synthesis is also indirectly regulated through translational feedback from ribosomal protein synthesis, where whilst ribosomal proteins have a high affinity for rRNA to facilitate ribosomes assembly, in the presence of excess rRNA the ribosomal proteins act as repressor and bind to their own mRNA preventing protein synthesis (Burgos *et al.,* 2017).

5.4.2 Secondary structure modelling of the 16s rRNA mutations

Computational modelling permitted the domain mapping of the rrn mutations to the secondary structure of the 16s rRNA, and the comparative analysis of conformational changes in secondary structure post-mutagenesis. Increased knowledge of the interactions of the 16s rRNA domains with ribosomal components, has facilitated the identification of highly conserved regions and structural motifs which are essential to the function of the ribosome (O'Connor, Göringer and Dahlberg, 1992; Vila, Viril-Farley and Tapprich, 1994; Santer *et al.,* 1995; Van Ryk and Dahlberg, 1995). However, an examination of the 16s rRNA mutation database suggested the

identified mutations were not located within domains of high functional importance (Triman, 1996).

Nevertheless, the association of the 16s rRNA with the ribosomal proteins is essential to the sequential assembly of the 30s-ribosomal subunit (Grondek and Culver, 2004). Studies have identified that mutations within conserved regions of the rRNA binding sites, which result in the conformational rearrangements of rRNA, impair ribosomal protein recognition and binding affinity in ribosomal proteins S7 and S8 (Allmang *et al.,* 1994; Dragon, Payant and Brakier-Gingras, 1994). Therefore, according to secondary structure predictions and the tertiary interactions within the 30s ribosomal complex, the altered conformation of the body domain of rRNA 154 proximate to the ribosomal protein S17, could disrupt the initial assembly of the 30s ribosomal subunit. This is significant as the ribosomal protein S12 requires its primary binding protein S17 for its association with the 16s rRNA in the formation of the shoulder of the 30s subunit, which contains the aminoglycoside binding pocket (Jakob *et al.,* 2012).

Unfortunately, the PyMOL mutagenesis function does not operate on ribonucleic acids, and thus the conformational alterations in the tertiary structure of the 16s rRNA within the 30s subunit were unidentified.

5.4.3 16s rRNA mutagenesis and aminoglycoside resistance

In addition to the binding of 16s rRNA with the ribosomal proteins in the assembly of the 30s ribosomal-subunit, there are structural requirements of the 16s rRNA for ligand binding including aminoglycoside antibiotics (Cho, Hamasaki and Rando, 1998). Footprinting studies have revealed a protective interaction between the

antimicrobial streptomycin and concise regions of the 16s rRNA at residues 913 and 915, which when mutated gives rise to phenotypic resistance through the disruption of drug binding (Moazed and Noller, 1987; Leclerc, Melancon and Brakier-Gingras, 1991). In like manner using a single rRNA allelic derivatives of the bacterium *Mycobacterium smegmatis*, hygromycin B resistant mutations were exclusively identified in residues within the helix 44 region of the 16s rRNA, local to the antimicrobial binding site (Pfister, Brodersen and Bottger, 2003). Hence, an assumption could be made that allosteric interactions occurring between the 16s rRNA mutations in helix 6 and the neighbouring helix 44, could disturb the gentamicin binding pocket and lead to phenotypic resistance.

Further studies of the aminoglycoside tobramycin have revealed rRNA aminoglycoside binding, is determined by factors beyond a consensus recognition sequence, with greater emphasis on the folding behaviours and the distinct shapes produced by the 16s rRNA (Wang and Rando, 1995; Tor, Hermann and Westhof, 1998). This suggests that the structural alterations observed within the 16s rRNA could have more profound effects on reducing the affinity of gentamicin to the binding pocket.

Despite the fact previous studies have reported an association between 16s rRNA mutagenesis and resistance development, it is highly unlikely that resistance manifests exclusively due to an alteration within the rRNA target (Kotra, Haddad and Mobashery, 2000). This is predominantly due to the occurrence of multiple copies of the rrn operons which confer resistance in a recessive manner, and thus mutations

are often missed due to the dominance of sensitive rrn genes which contribute to a greater total number of ribosomes (Sigmund and Morgan, 1982). Therefore, it is unrealistic to assume that the acquisition of gentamicin resistance occurs exclusively as a result of the identified mutations within the two 16s rRNA operons. This was further supported through the confirmatory PCR and sequencing of the mutated 16s rRNA genes, which displayed irregularities in terms of the existence and locality of rrn mutations.

5.4.4 Functional analysis of the ribosomal protein EF-G

The functional analysis of the EF-G mutagenesis, holds great significance owing to its critical involvement during protein synthesis, through mediating the translocation of mRNA and tRNA's through the ribosome (Ramrath *et al.,* 2013). Over the last 15 years, substantial insights have been provided into the dynamics of tRNA's during translation (Blanchard *et al.,* 2004; Ermolenko *et al.,* 2007; Agirrezabala *et al.,* 2008; Julian *et al.,* 2008). However, although EF-G is appreciated for its vital catalytic role, the molecular mechanism through which EF-G coordinates ribosomal translocation is not fully understood (Salsai *et al.,* 2014).

The elongation factor EF-G has considerable involvement in two independent phases of ribosomal translocation. The initial phase involves the spontaneous movement of tRNAs, between the ribosomal A- and P- sites relative to the 50s subunit, to form the classical and hybrid pre- translocational conformations (Kim, Puglisi and Chu, 2007; Munro *et al.,* 2010). An important function of the binding of EF-G during this phase is to transiently stabilise the intermediate rotated hybrid state of the pre- translocation

ribosome (Spiegel, Ermolenko and Noller, 2007). During the second phase, EF-G catalyses the movement of mRNA and tRNA combined with the reverse rotation of the 30s-ribosomal subunit (Savelsbergh *et al.,* 2003; Ermolenko and Noller, 2011; Guo and Noller, 2012; Tourigny *et al.,* 2013), which is said to be accelerated through GTP hydrolysis (Rodnina *et al.,* 1997; Holtkamp *et al.,* 2014).

Early structural studies revealed EF-G contains five structurally defined domains, which are fundamental to understanding the molecular mechanisms behind ribosomal translocation (AEvarsson *et al.,* 1994; Czworkowski 1994). EF-G domains 3, 4 and 5 have combined functional significance, through the molecular mimicry of the aminoacyl-tRNA at the ribosomal A-site (Clark *et al.,* 1999). Domain 4 in its elongated state replicates the anticodon arm of the aa-tRNA and deletion of this domain results in diminished translocational activity (Martemyanov and Gudkov,1999). This is compelling evidence to suggest that the mutation identified in the boundary between EF-G domains 4 and 5, could result in altered binding of domain 4 to the ribosomal A site, and subsequently influence ribosomal translocation.

Further structural investigations of the ribosome bound EF-G within the pre- and post translocation states, have indicated that EF-G must undergo substantial structural rearrangement in order to prevent the steric clash of domain 4 at the ribosomal A-site pre-translocation (Gao *et al.,* 2009; Zhou *et al.,* 2012; Brilot *et al.,* 2013). This conformational variability displayed by EF-G in its bound form, is thought to be essential in promoting ribosomal translocation (Munro *et al.,* 2010; Li *et al.,* 2011). Such that, when the intramolecular mobility of EF-G is compromised through the

addition of disulfide crosslinks, EF-G is unable to facilitate ribosomal translocation despite the maintenance of GTPase activity (Peske *et al.,* 2000).

This is particularly meaningful as the non-synonymous mutation at position 610 between the boundary of EF-G domains 4 and 5, leads to an amino acid substitution from a proline to a threonine residue. Proline residues have been recognised for their ability to form molecular hinges, which induce conformational changes within protein structures (Sansom and Weinstein, 2000). Consequently, the replacement of a proline residue at the domain 4 hinge with the amino acid threonine, could significantly effect the intrinsic flexibility of EF-G. This hypothesis was supported through the identification of mutation induced changes in protein compressibility, which is closely associated with modifications in the flexibility and overall functional dynamics of proteins (Gekko *et al.,* 2000; Gekko *et al.,* 2004). In addition, the presence of the EF-G mutation at position 610 in the resequenced resistant isolates, suggests *fusA* mutagenesis is fundamental to the acquisition of gentamicin resistance within the two models.

5.4.5 Conclusion

To conclude, it is therefore tempting to speculate that the combinatory effects of 16s rRNA and *fusA* mutations within the 70s-ribosomal complex, are the intrinsic genotypic adaptations governing gentamicin resistance within the experimental model. Nevertheless, it is currently unknown whether the observed mutations interfere with the antimicrobial binding, and this will be examined independently within chapter 6.

Chapter 6: *In-silico* molecular docking analysis illustrating the adaptive response to gentamicin resistance in *E. coli* MG1655.

6.1 Introduction

Molecular docking aims to predict the precise interactions between a ligand and a receptor using computational approaches (Meng *et al.,* 2011). Such approaches have been fundamental in charactering the behaviours of small molecules and consequently have been applied as tools in molecular biology and computer assisted drug design (Morris and Lim-Wilby, 2008). Molecular docking is routinely used as a screening tool within modern drug design, to explore the drug receptor interactions of novel or synthesised derivatives of compounds (Vijesh *et al.,* 2011).

With the recent increase of antimicrobial resistance, there is an urgent need for approaches to combat this issue. Molecular docking has been utilised to provide insights into the known antimicrobial target mechanisms in the discovery of novel antimicrobial compounds of natural and synthetic origin (Alves *et al.,* 2014; Rajasekaran and Rao, 2015; Gullapelli *et al.,* 2017). In addition, docking can be applied in the study of existing antimicrobial agents to reveal the structural mechanisms of resistance following mutagenesis (Lee, Hong and Kim, 2018; Zhou *et al.,* 2015).

Owing to the application of gentamicin in this study, it was essential to comprehend the docking of antimicrobial agents which target the bacterial ribosome. Docking studies have revealed that mutations within both the ribosomal proteins and rRNA lead to target site perturbations which interfere with the binding of antimicrobial

agents (Long *et al.,* 2009; Klitgaard *et al.,* 2015). Ribosomal binding studies have also shown that in addition to mutations which directly interfere with the drug binding sites, mutations found at distant locations within the ribosome can result in antimicrobial resistance (Blaha *et al.,* 2008).

The structural basis for gentamicin inhibition of bacterial ribosomal protein synthesis in *E. coli* has been previously resolved by Borovinskaya *et al*. (2007), yet this high-resolution structure only exists for EF-G in its elongated conformation of the post-ribosome complex. Further cryo-EM reconstruction of EF-G in the pre-ribosomal complex, revealed small structural rearrangements of EF-G relative to the ribosome during translocation, which is trapped through the addition of the aminoglycoside viomycin (Brilot *et al*., 2013). Nevertheless, as this fails to demonstrate the conformational flexibility of EF-G between its compact and elongated states, the Lin *et al*. (2014) atomic resolution crystal structures were obtained for the *in-silico* molecular docking analysis of *fusA* mutagenesis in the gentamicin resistant strains of *E. coli* MG1655.

The aim of this chapter is therefore to determine whether the binding of gentamicin to the molecular target is interrupted because of the ribosomal mutagenesis of the 16s rRNA genes and EF-G. It is hypothesised that as structural alterations have previously been identified post-mutagenesis, it is likely that the EF-G and rRNA mutations within the 70s-ribosomal complex will disrupt the binding of gentamicin. This would suggest that the identified mutations within this study, are responsible for the *in-vitro* acquisition of gentamicin resistance.

6.2 Materials and methods

6.2.1 Molecular structure analysis

The key components within the docking simulation experiments were sketched using MoleculeSketch, detailed in section (2.11.1).

6.2.2 Preparation of PDB files for docking

The computational ligand docking was achieved using AutoDock suite, a prerequisite of which is a correctly formatted PDB files. Details regarding the source of the PDB files and subsequent modifications for use within Autodock are provided in section (2.11.2).

6.2.3 Preparation of receptor and ligand for docking

The initial stage of molecular docking analysis involved the preparation of ligand and receptor files, which included root detection and computation of charges respectively. Further details of ligand and receptor preparation are specified in sections (2.11.3 and 2.11.4).

6.2.4 Grid box preparation and single docking analysis

To ensure the correct binding of the ligand to the receptor, grid box parameters were specified and modified per docking analysis (Section 2.11.5). Following docking analysis all poses were visualised in PyMOL. Further details are available in section (2.11.6).

6.2.5 Autodock modifications

An issue occurred within Autodock when reading the molecule, as the 70s complex exceeds the maximum number of atoms permitted. To overcome this the retain order was unset by compiling Autogrid and Autodock version 4.2 from source and editing the line in constants.h in the Autodock directory and line autogrid.h within Autogrid to define the max atoms and records to a number exceeding that of the 70s-ribosomal complex. Before compiling, a new directory was made for the new build and ran using autoreconf –i and then built routinely using configure make. In addition to the atom size limit, discrepancies can be observed in molecules containing metal ions, including potassium and magnesium. This error was overcome by adding the following line: atom par K 3.81 0.035 12.000 -0.00110 0.0 0.0 0 -1 -1 1 to the file AD4_parameters.dat file in the Autodock subdirectory, which adds the given charges to the metal ion detailed. To ensure the changes are executed, after the gpf. and dpf. files are created the following line is added to the top of both files respectively: parameter_file AD4_parameters.dat.

6.3 Experimental results

6.3.1 The molecular structure analysis of gentamicin

The structural details of the aminoglycoside gentamicin were discussed previously in chapter 3 (Section 3.9.8.2). Gentamicin has two predominate targets within the 70s-ribosomal complex, which include the 16s rRNA and the 30s-ribosomal subunit protein S12 (Section 4.3.3). Gentamicin has been found to bind to the 16s rRNA in two adjacent positions. Hydrogen bonds are formed between the nitrogen and guanine at position 1405 and between the oxygen and uracil at position 1406. In addition, gentamicin forms a hydrogen bond between the nitrogen and the lysine residue of S12 at position 42. The molecular structure of gentamicin and the corresponding ribonucleotides and the amino acid binding structures can be found below (Figure 6.1).

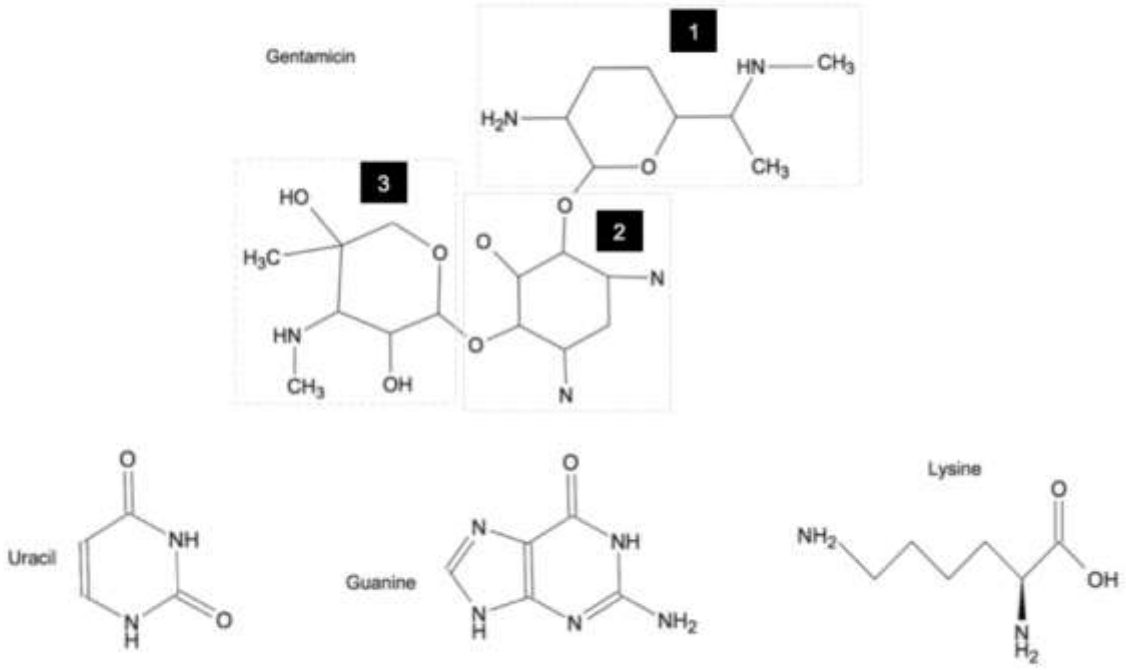


Figure 6.1 Molecular structures of gentamicin and the ribonucleotide and amino acid binding structures. Figure displaying the 2D molecular structure of the gentamicin complex with the three major components labelled 1, 2 and 3. The image also includes the structures of the gentamicin targets including the ribonucleotide bases uracil and guanine of the 16srRNA and the amino acid lysine in the 30s-ribosomal subunit protein S12.

6.3.2 Preparation of the gentamicin-ligand for molecular docking to the 30s-ribosomal subunit protein S12

The primary docking experiments were conducted exclusively on the 30s-ribosomal subunit protein S12. The initial step involved the preparation of the ligand gentamicin. Here the antimicrobial root was detected and a total of 13 rotatable bonds were identified (Figure 6.2), this followed the computation gasteiger charges where a total of 5.0003 charges were added (Figure 6.3).

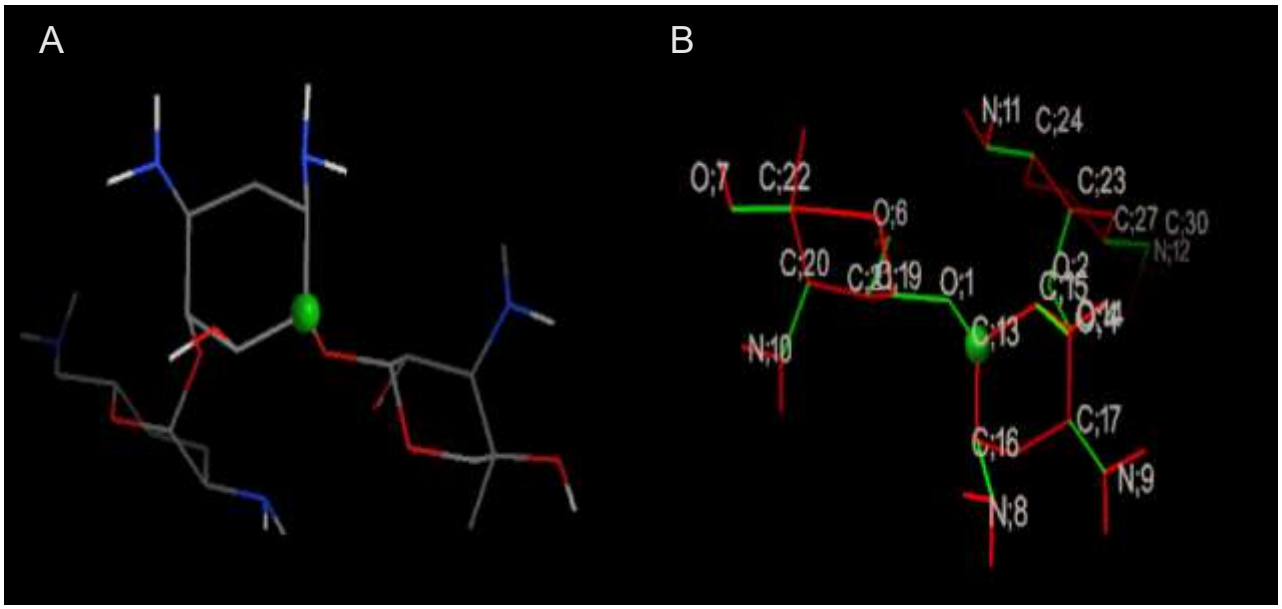


Figure 6.2 Preparation of the gentamicin ligand for docking in Autodock. PyMOL image displaying the root detection of the gentamicin ligand. The root is determined as the fixed portion of the ligand closest to the molecule centre and is shown as a green sphere in both gentamicin structures. (A) Gentamicin structure displaying the root and the chemical composition of the ligand: nitrogen atoms (blue), oxygen atoms (red) and carbon atoms (grey). (B) Gentamicin structure displaying the detected root, the position of each atom and information regarding the rotatable bonds. The rotatable bonds are shown in green and the non-rotatable bonds are shown in red.

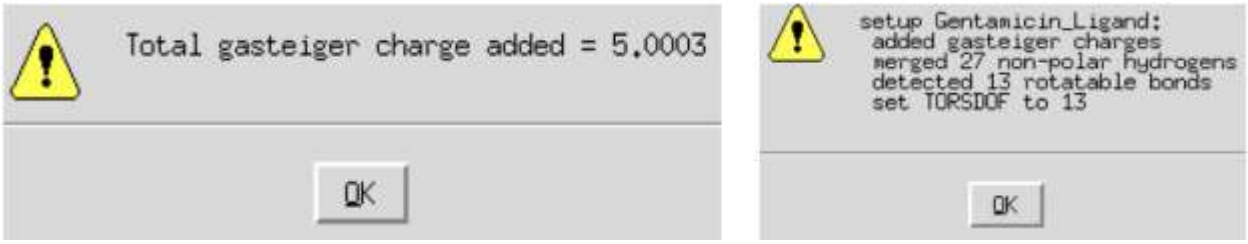


Figure 6.3 The gentamicin ligand preparation docking summary. The autodock output of the final stage in the ligand preparation, displaying the total number of gasteiger charges added the gentamicin ligand (5.003) and the ligand summary which includes the number of detected rotatable bonds (13) and number of merged non-polar hydrogen atoms (27) for ligand docking.

6.3.3 Ligand and grid-box preparation for the molecular docking simulation of gentamicin to the 30s-ribosomal subunit protein S12

Subsequently, the grid box was prepared for the docking of the gentamicin ligand. As previously discussed (Section 6.3.1), gentamicin binds to the lysine residue at position 42 within S12 which was highlighted with yellow asterisks, to assist the grid box optimisation (Figure 6.4). The grid box was positioned directly surrounding the lysine residue and the finalised grid box parameters for docking are shown in table (6.1).

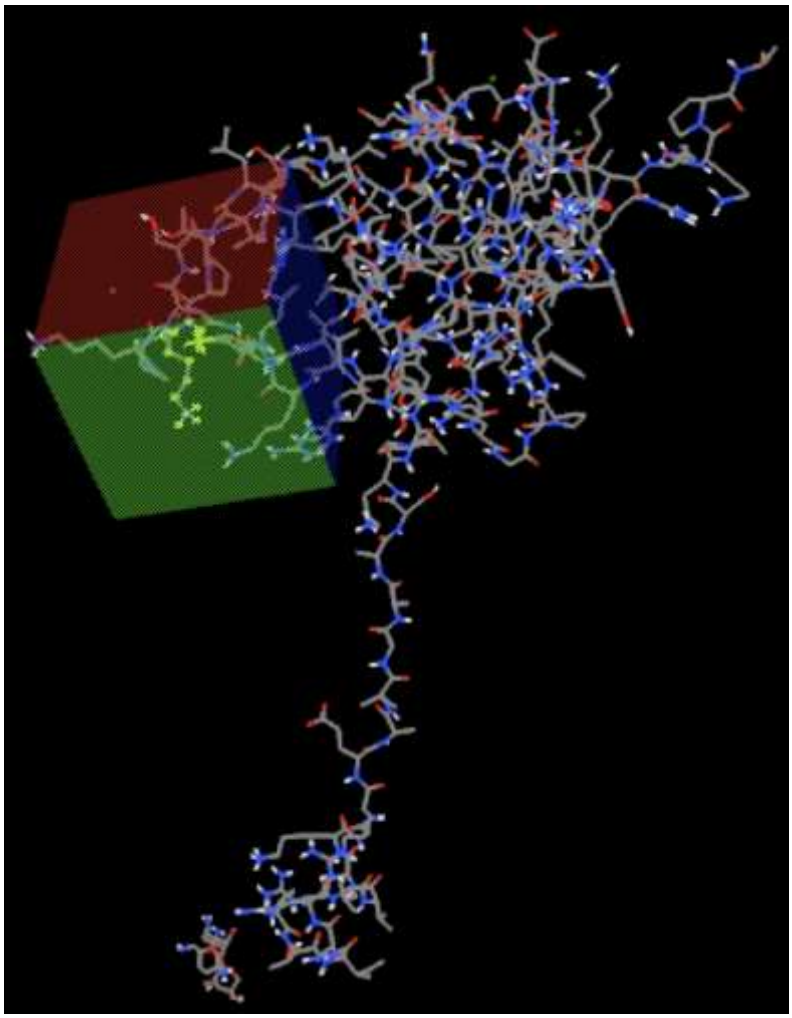


Figure 6.4 The grid box preparation for the 30s-ribsomomal subunit protein S12. PyMOL image displaying the structure of the 30s-ribsomomal subunit protein S12 in stick representation with colours determined by the atoms within the protein structure: carbon atoms (grey), nitrogen atoms (blue) and oxygen atoms (red). The grid box is positioned directly surrounding the lysine residue at position 42 of the ribosomal protein S12 to facilitate the site-specific docking of the ligand to the known binding sites within the protein structure.

Table 6.1 The 30s-ribsomomal subunit protein S12 grid box parameters for the single molecular docking analysis. Displaying the optimised grid box and centre grid box coordinates for the docking simulation of the gentamicin ligand to the 30s-ribsomal subunit protein S12.

Current total grid points per map	6400
Number of points in x-dimension	40
Number of points in y-dimension	40
Number of points in z-dimension	40
Spacing (angstrom)	0.375
Centre grid box	
X centre	19.806
Y centre	11.306
Z centre	-22.389

6.3.3.1 Single docking analysis of gentamicin to the 30s-ribosomal subunit protein S12

The docking analysis of gentamicin to the 30s-ribosomal protein S12 identified five possible antimicrobial conformations, varying in binding energies ranging from 5.48 to 5.535 kcal/mol (Figure 6.5). All predicted gentamicin poses were visualised as sticks, and clear overlays in structures were observed (Figure 6.6).

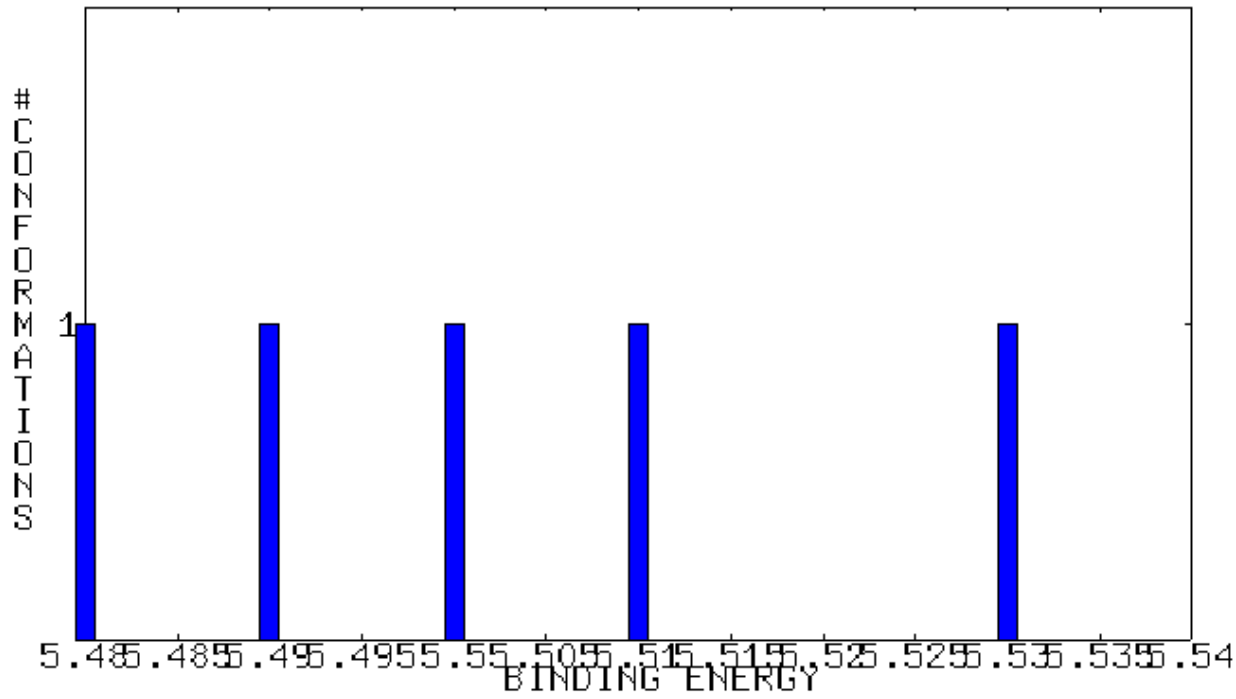


Figure 6.5 Binding energies of the gentamicin conformations obtained from the docking simulation with the 30s-ribosomal subunit protein S12 with Autodock. Graph showing the five predicted gentamicin conformations and the corresponding binding energies.

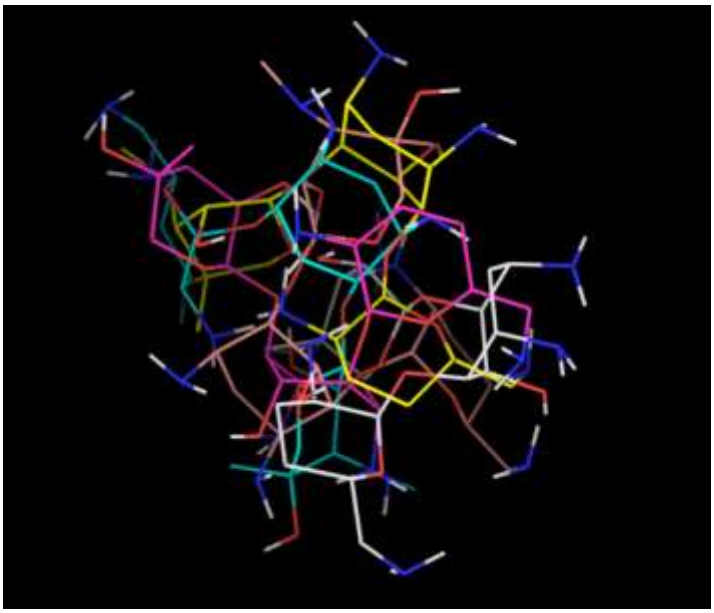


Figure 6.6 Superimposition of the gentamicin binding poses. PyMOL image displaying the structures of the five predicted gentamicin conformations illustrated using the stick model representation. All stick poses are differentiated by the colour of carbon backbone. The remaining elements of the ligand structure are coloured by atom: nitrogen (blue) polar hydrogens (white) and oxygen (red).

Next, all predicted gentamicin conformations were docked to the 30s-ribosomal protein S12, and comparably the overlay in the gentamicin poses was observed as shown (Figure 6.6), with ligand binding appearing in close proximity to the lysine residue at position 42 (Figure 6.7).

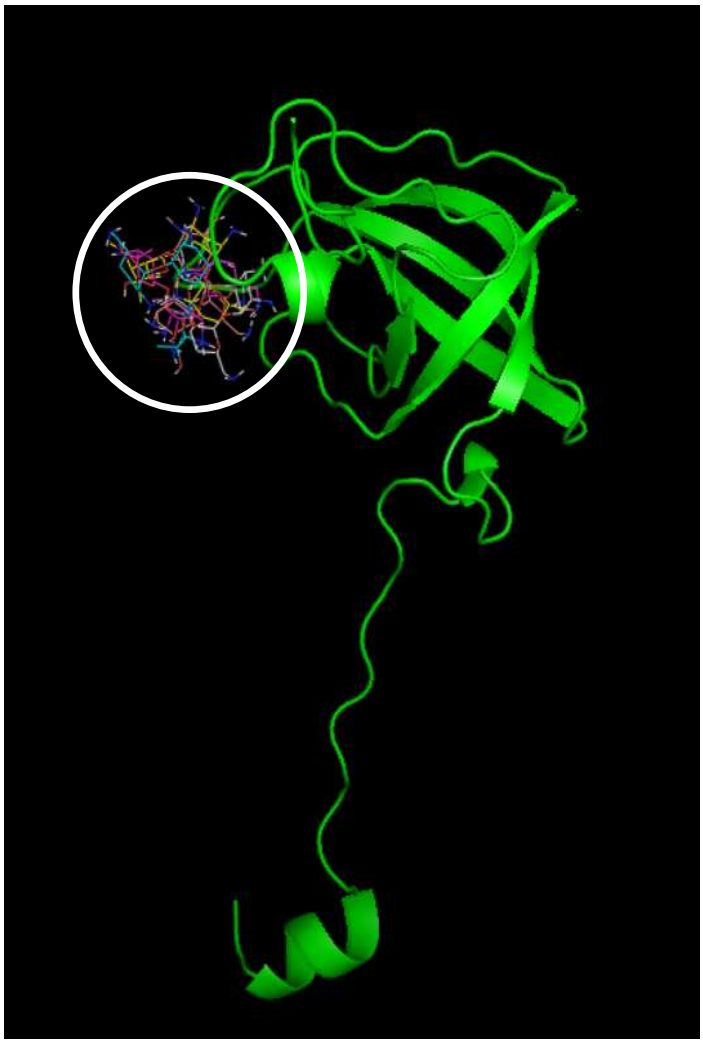


Figure 6.7 Docking simulation of all gentamicin poses to the 30s-ribsomomal subunit protein S12. PyMOL image displaying the single docking analysis image of all predicted gentamicin binding conformations illustrated using the stick model representation (circled in white) docked to the ribosomal subunit protein S12 (green) near the lysine residue at amino acid position 42.

As the binding of molecules is favourable in the lowest energy state, docking was conducted with the gentamicin conformation with a biding energy of 5.48 kcal/mol, using the space filling model (Figure 6.8). Within this model the nitrogen residue can be seen in binding in close proximity to the lysine residue at position 42, supporting the previous docking experiments.

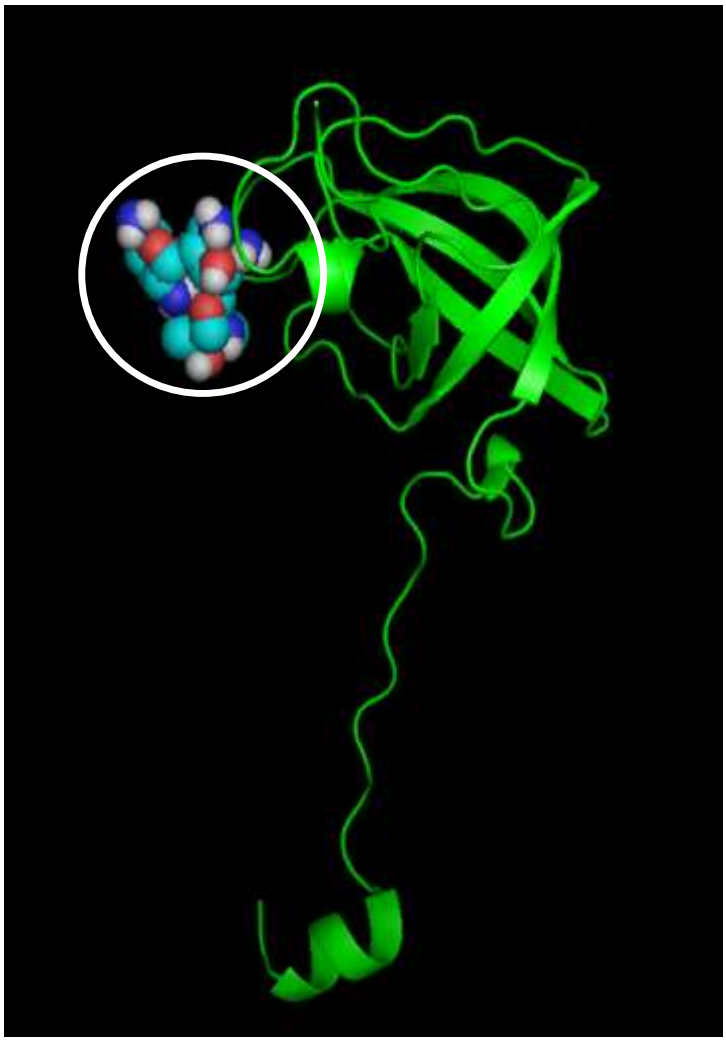


Figure 6.8 Docking simulation of gentamicin to the 30s-ribsomomal subunit protein S12. PyMOL image displaying the single docking analysis image of the lowest gentamicin binding energy conformation represented using the space filling model (circled in white) docked to the ribosomal subunit protein S12 (green) near the lysine residue at amino acid position 42.

To gain a comprehensive understanding of the interactions of the gentamicin ligand within the 30s-ribosomal subunit with both antimicrobial targets (ribosomal protein S12 and the 16s rRNA), docking experiments were repeated with the inclusion of the 70s-ribosomal protein complex.

6.3.4 Gentamicin docking analysis 70s-ribosomal complex

There were extensive modifications made to the PDB files of both the pre- and post 70s-ribosomal complex with EF-G, prior to the successful upload into autodock.

6.3.4.1 Receptor preparation of the 70s-ribosomal complex

A primary stage in the receptor preparation, involved the computation of gasteiger charges. The total number of gasteiger charges added to the 70s-ribosomal complex was -2337.9281 (Figure 6.9).

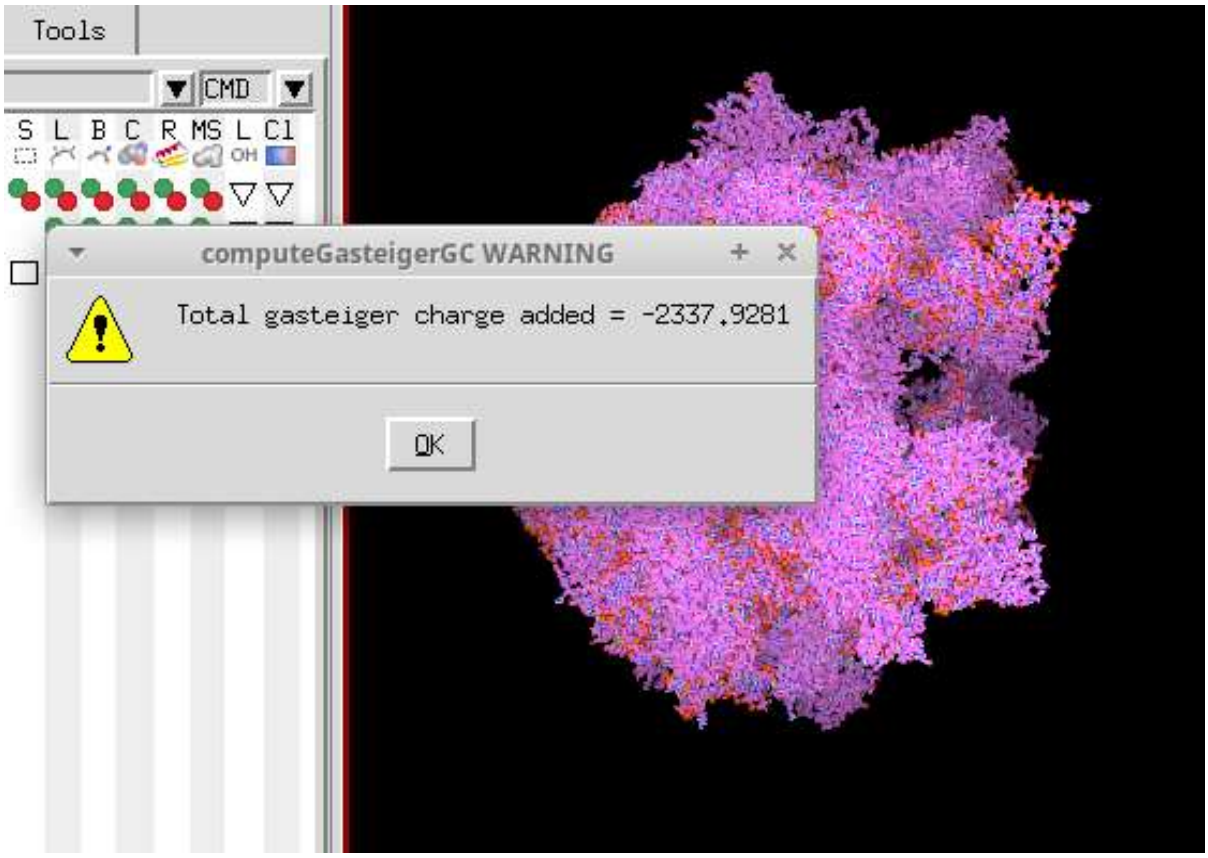


Figure 6.9 Preparation of the 70s-ribosomal complex receptor for docking in Autodock. Displaying the autodock output in the final stage of the 70s-ribosomal complex receptor preparation showing the total number of gasteiger charges added to the receptor macromolecule (-2337.9281).

6.3.4.2 70s pre-ribosomal complex docking analysis pre-mutagenesis

Following receptor preparation, key components for gentamicin docking and ribosomal translocation were highlighted within PyMOL, and included the 30s-ribosomal subunit protein S12, A, P and E site tRNAs, mRNA and EF-G (Figure 6.10).

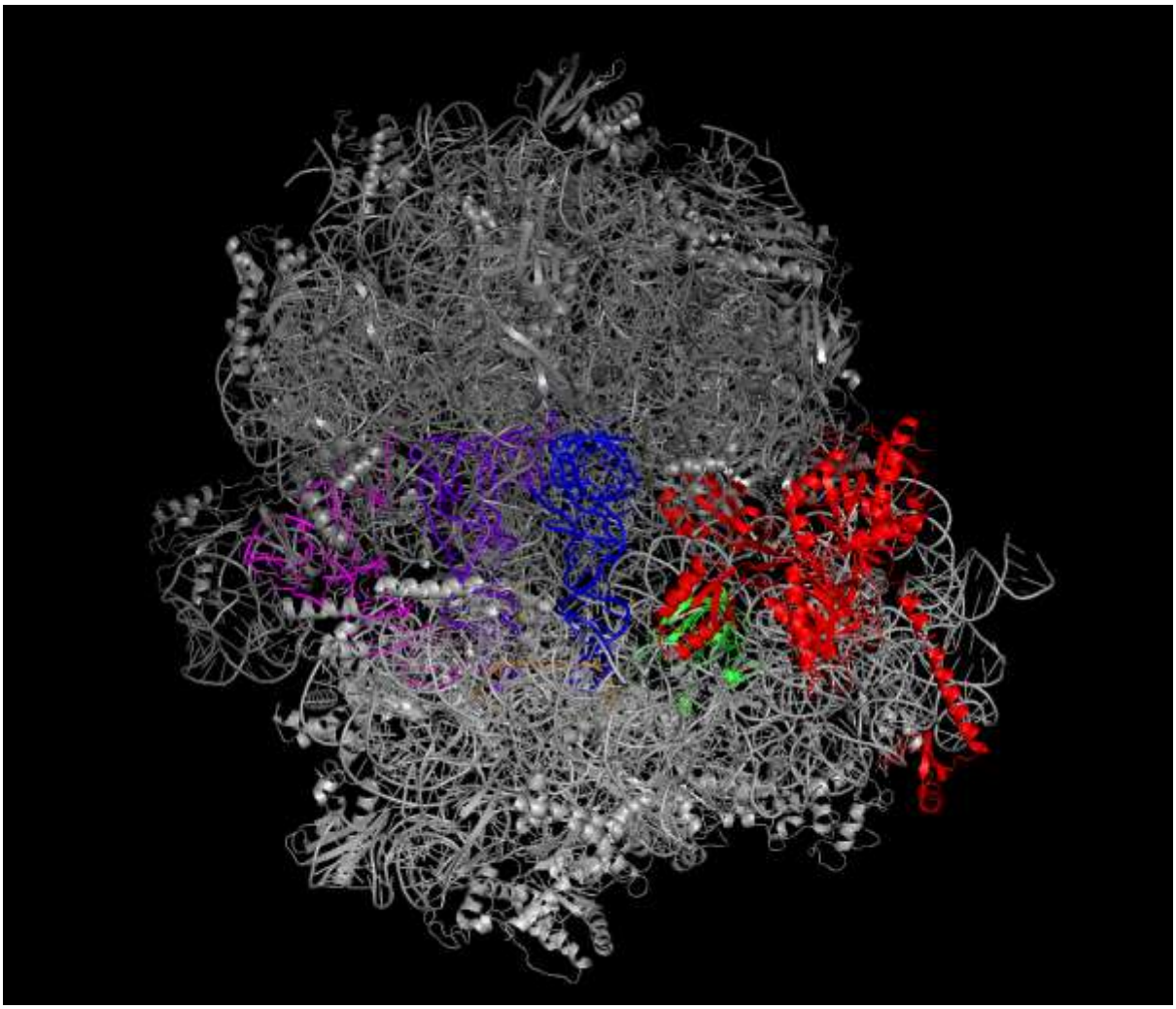


Figure 6.10 The tertiary structure of the 70s pre-ribosomal complex. PyMOL image displaying the 70s pre-ribosomal complex and the key components involved in the ribosomal translocation stage of protein synthesis: EF-G (red), mRNA (orange), A-site tRNA (blue), P-site tRNA (purple) and E-site tRNA (pink). The 30s and 50s ribosomal subunits are shown in light and dark grey respectively and the gentamicin 30s-ribosomal subunit protein target S12 is depicted in green. EF-G (red) is found within its compact conformation within the 70s pre-ribosomal complex.

Successively, the grid box was prepared to encompass both ribosomal targets, comprising the lysine residue at position 42 and the two 16s rRNA positions. The grid box was further expanded to integrate sections of both the A-site tRNA and the mRNA (Figure 6.11). The optimised grid box parameters used in the docking experiments of the 70s-ribosomal complex, are provided in Table (6.2).

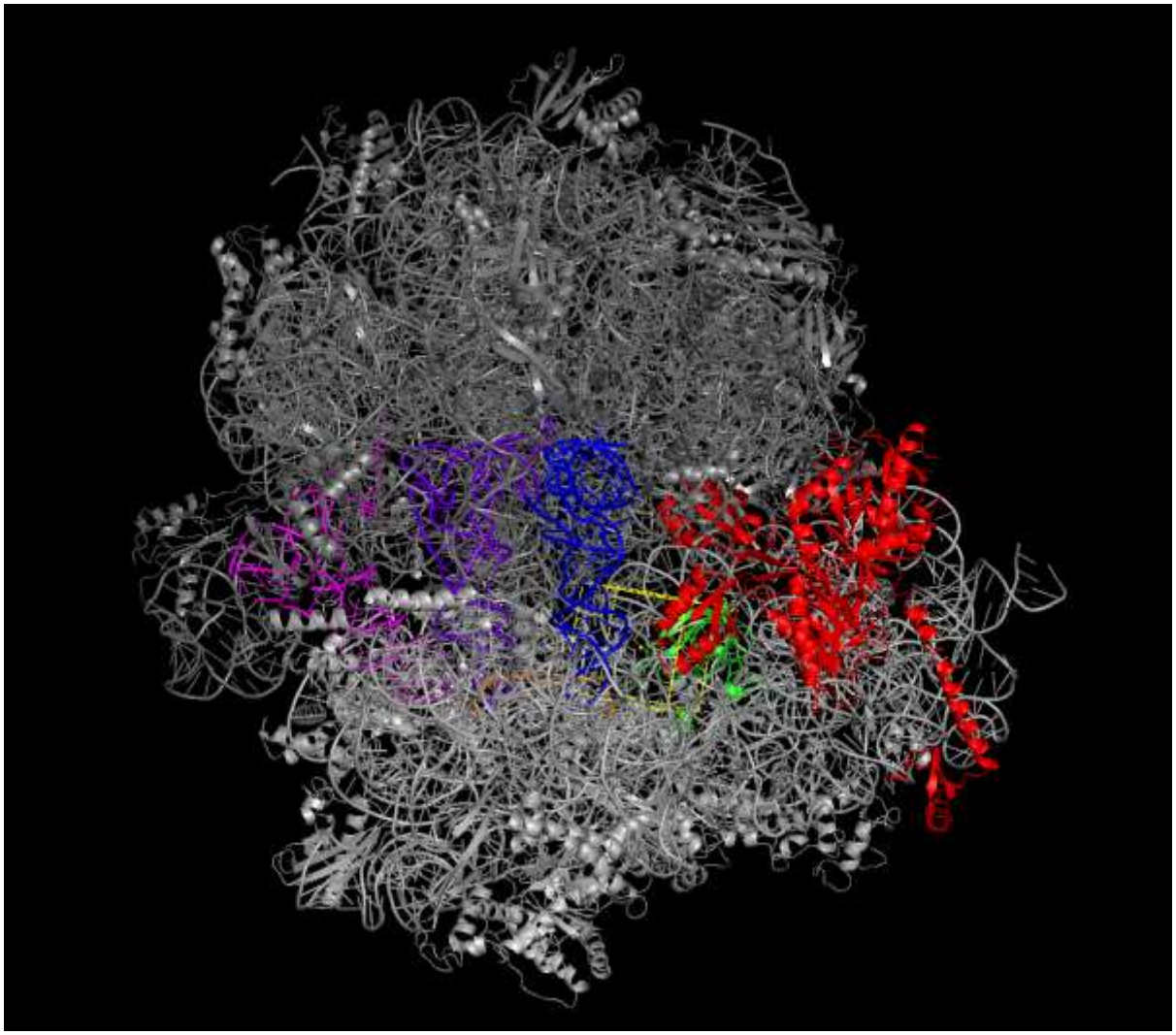


Figure 6.11 The grid box preparation for the 70s pre-ribosomal complex. PyMOL image displaying the cartoon representation of the tertiary structure of the 70s pre-ribosomal complex. The 30s-ribosomal subunit is shown in light grey, 50s-ribosomal subunit in dark grey, EF-G (red), mRNA (orange), A-site tRNA (blue), P-site tRNA (purple), E-site tRNA (pink) and the gentamicin 30s-ribosomal subunit protein target S12 is depicted in green. The grid box (yellow) is positioned proximal to the ribosomal subunit protein S12, A-site tRNA, mRNA and the neighbouring 16s rRNA.

Table 6.2 The 70s pre-ribosomal complex grid box parameters for the molecular docking analysis. Displaying the optimised grid box and centre grid box coordinates for the docking simulation of the gentamicin ligand to the 70s pre-ribosomal complex.

Current total grid points per map	6400
Number of points in x-dimension	60
Number of points in y-dimension	60
Number of points in z-dimension	60
Spacing (angstrom)	0.465
Centre grid box	
X centre	104.64
Y centre	9.769
Z centre	31.18

Docking analysis of gentamicin to the 70s pre-ribosomal complex, revealed a total of 8 gentamicin binding conformations, with binding energies ranging from approximately -108.2 to -105.6 kcal/mol (Figure 6.12).

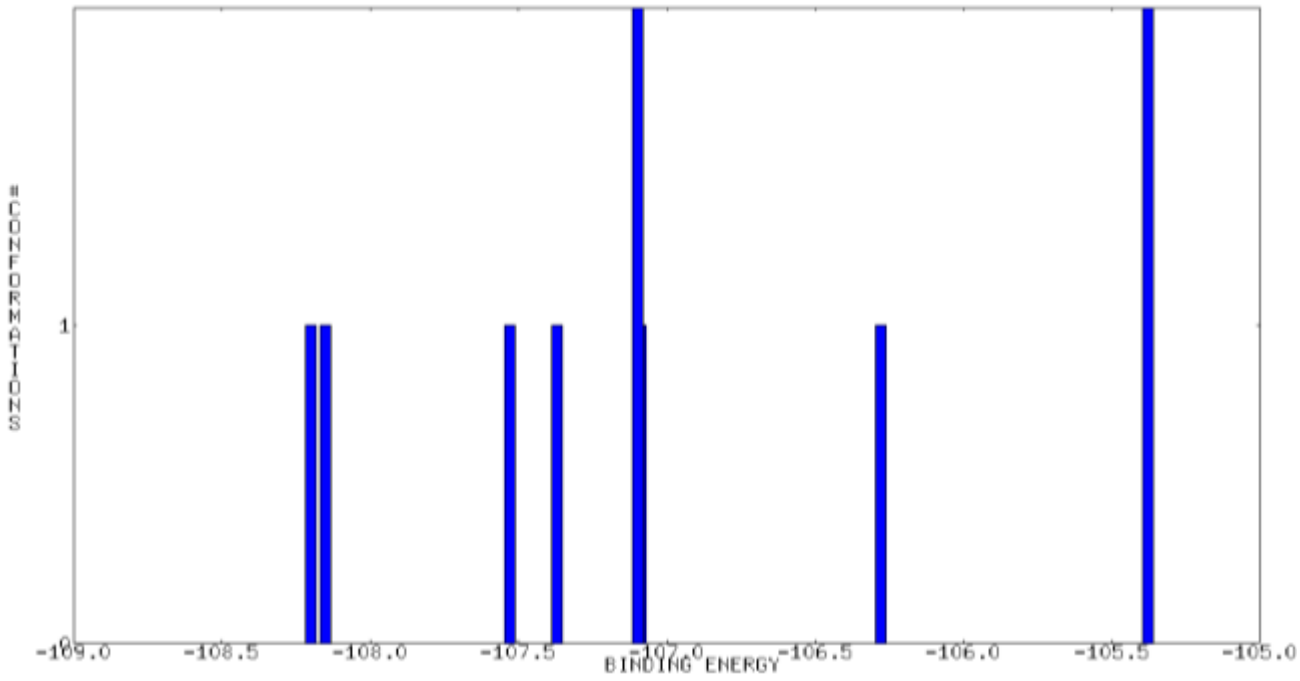


Figure 6.12 Binding energies of the gentamicin conformations obtained from the docking simulation with the 70s pre-ribosomal complex with Autodock. Graph showing the eight predicted gentamicin conformations and the corresponding binding energies.

All the 8 predicted gentamicin binding poses were shown within the 70s pre-ribosomal complex using the space filling model (Figure 6.13). From the figure, all the gentamicin conformations appeared to preferentially bind in closer proximity to the A-site tRNA and mRNA, and at a greater distance from the predicted binding site, the 30s-ribosomal subunit protein S12.

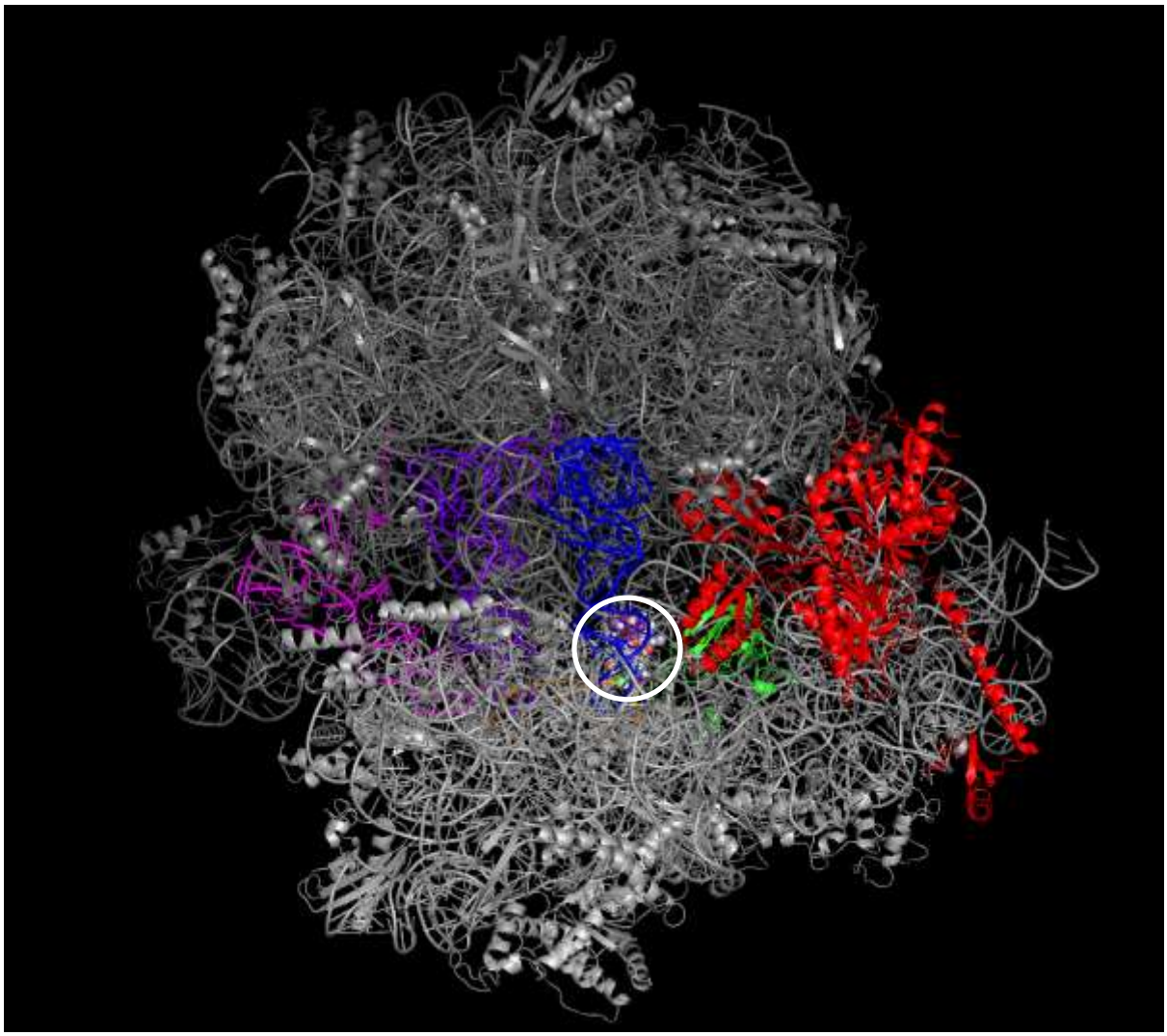


Figure 6.13 Docking simulation of all gentamicin poses within the 70s pre-ribosomal complex. PyMOL image displaying the docking analysis image of all predicted gentamicin binding conformations illustrated using the space filling model representation (circled in white) docked to the 70s pre-ribosomal complex.

Afterwards, the docking results were amended to allow the visualisation of the gentamicin conformation with the lowest binding energy of -108.2 kcal/mol independently within the 70s pre-ribosomal complex (Figure 6.14). The figure below was magnified to determine the precise location of gentamicin within the ribosomal complex, relative to the key components previously described. The gentamicin molecule seemed to be equal distance from the A-site tRNA, mRNA, EF-G domain 4 and the 30s-ribosomal subunit protein S12. It should be noted that the lysine residue at position 42 within the ribosomal protein S12 was highlighted in yellow, as the lowest bound energy conformation of gentamicin appeared in closer vicinity with S12 than previously presumed.

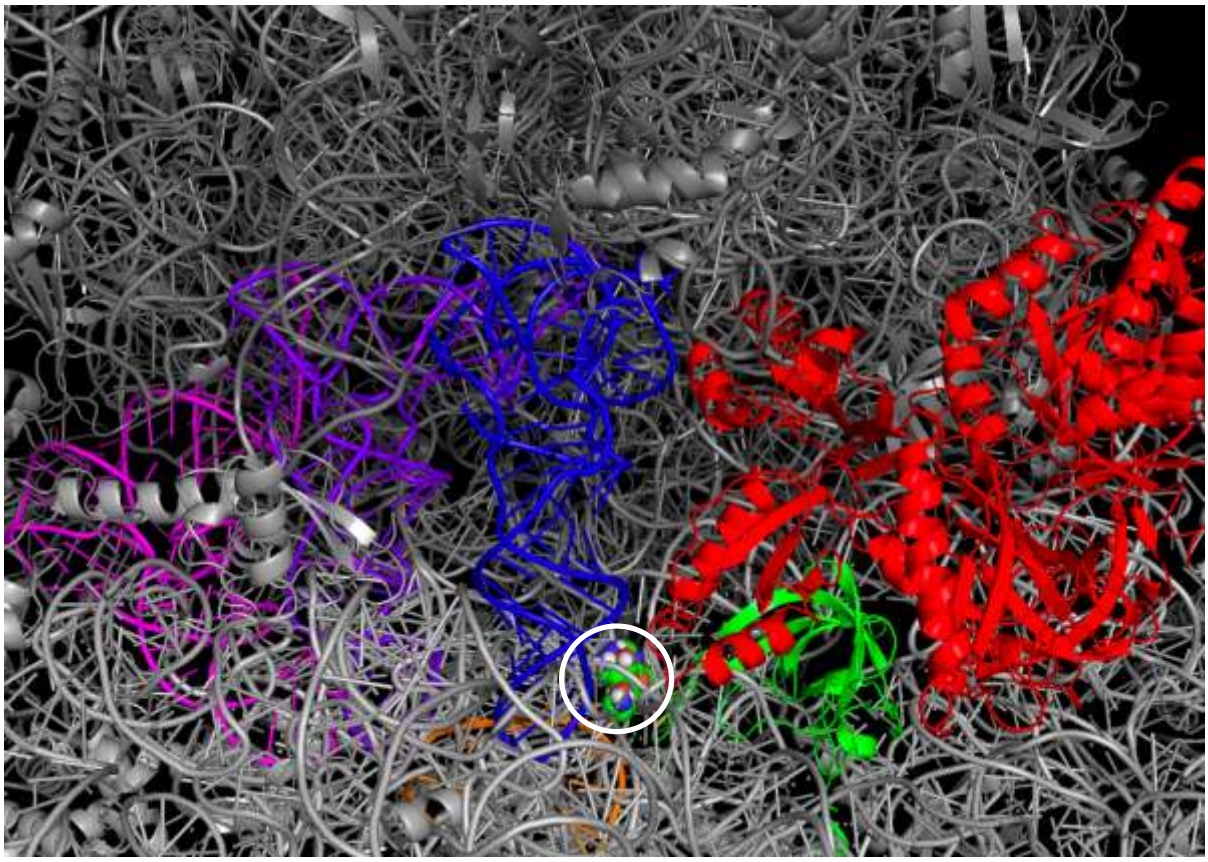


Figure 6.14 Docking simulation of gentamicin to the 70s pre-ribosomal complex. Magnified PyMOL docking analysis image displaying the lowest gentamicin binding energy conformation depicted using a space filling model (circled in white) docked to the 70s pre-ribosomal complex. The gentamicin-ligand is bound to the 70s pre-ribosomal complex near the lysine residue at position 42 of the 30s-ribosomal subunit protein S12 (green), EF-G (red), A-site tRNA (blue) and mRNA (orange).

6.3.4.3 70s post-ribosomal complex docking analysis pre-mutagenesis

The docking analysis was then repeated using identical grid box parameters, but in this case for the 70s post- ribosomal complex, where EF-G undergoes an extensive conformational change with domain 4 occupying the A-site tRNA (Figure 6.15). Again, the 30s-ribosomal subunit protein S12 was highlighted in green, as an indicator of the expected gentamicin target site.

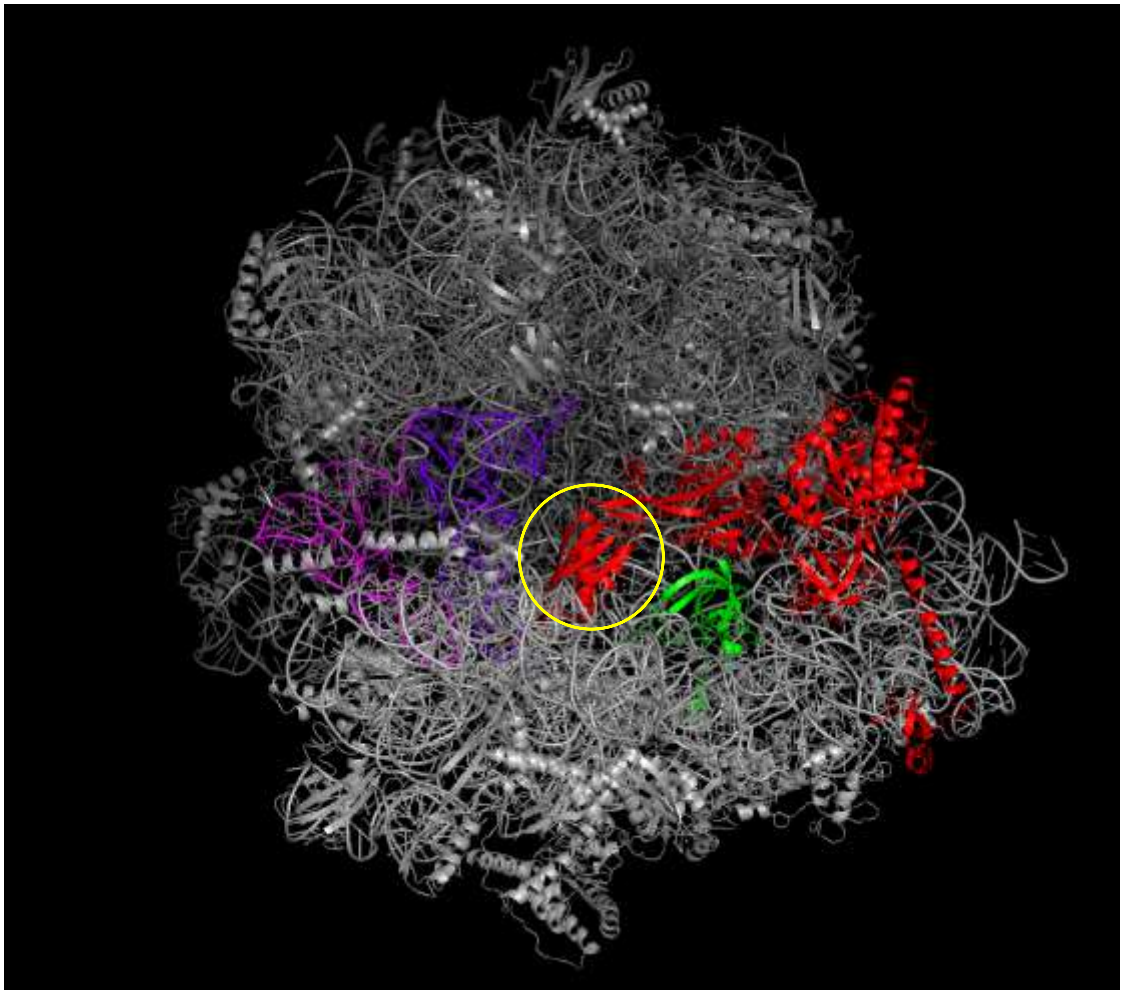


Figure 6.15 The tertiary structure of the 70s post-ribosomal complex. PyMOL image displaying the 70s post-ribosomal complex and the key components involved in the ribosomal translocation stage of protein synthesis: EF-G (red), mRNA (orange), P-site tRNA (purple) and E-site tRNA (pink). The 30s and 50s ribosomal subunits are shown in light and dark grey respectively and the gentamicin 30s-ribosomal subunit protein target S12 is depicted in green. EF-G (red) is found within its elongated conformation within the 70s pre-ribosomal complex. The EF-G domain 4 (circled in yellow) occupies the A-site tRNA within the 70s post-ribosomal complex.

Here, in comparison to the 8 predicted conformations for the 70s pre-ribosomal complex, there were a total of 9 gentamicin conformations predicted with binding energies ranging from -100.0 to -96.50 kcal/mol (Figure 6.16).

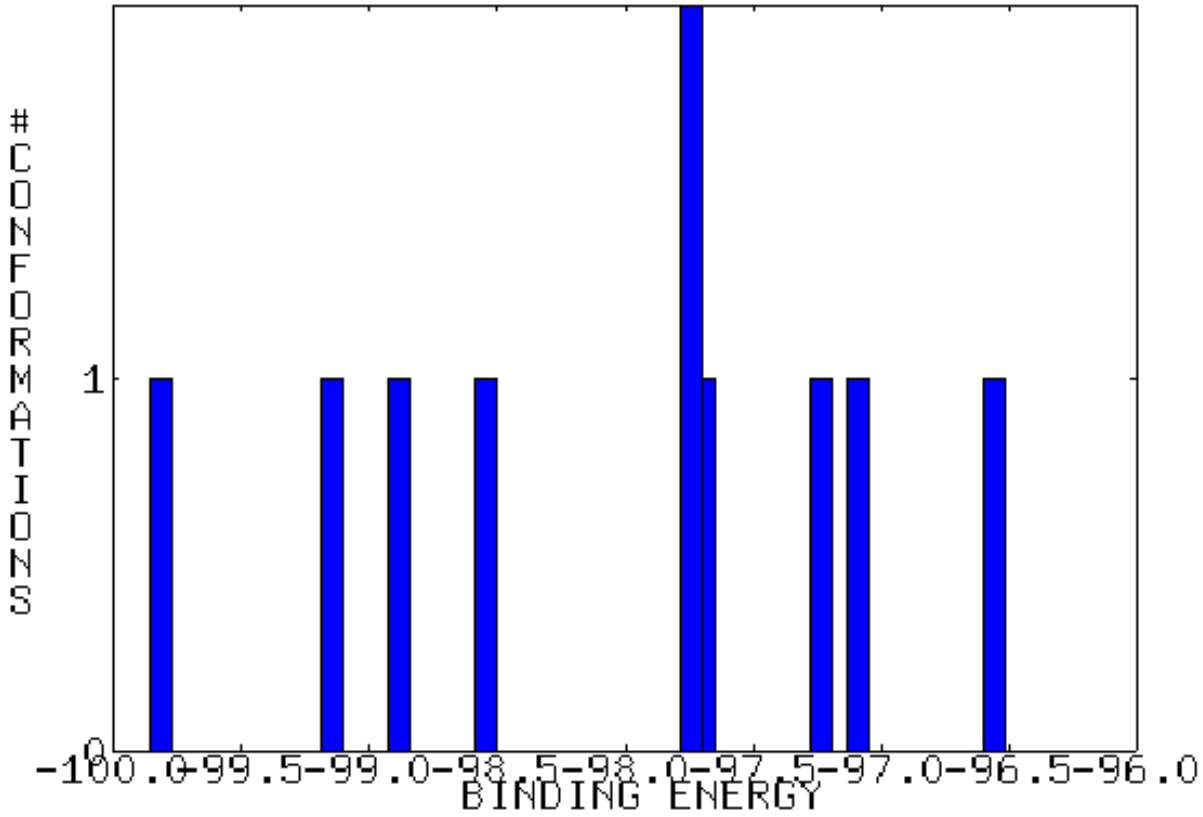


Figure 6.16 Binding energies of the gentamicin conformations obtained from the docking simulation with the 70s post-ribosomal complex with Autodock. Graph showing the nine predicted gentamicin conformations and the corresponding binding energies.

Following the docking predictions, all 9 gentamicin poses were visualised within the 70s post-ribosomal complex (Figure 6.17). The docking analysis revealed discrepancies in the docking position of the antimicrobial gentamicin between the pre- and post- 70s ribosomal complex. Within the post ribosomal complex, all 9 gentamicin poses appeared behind domain 4 of EF-G, and distant from the 30s-ribosomal subunit protein S12.

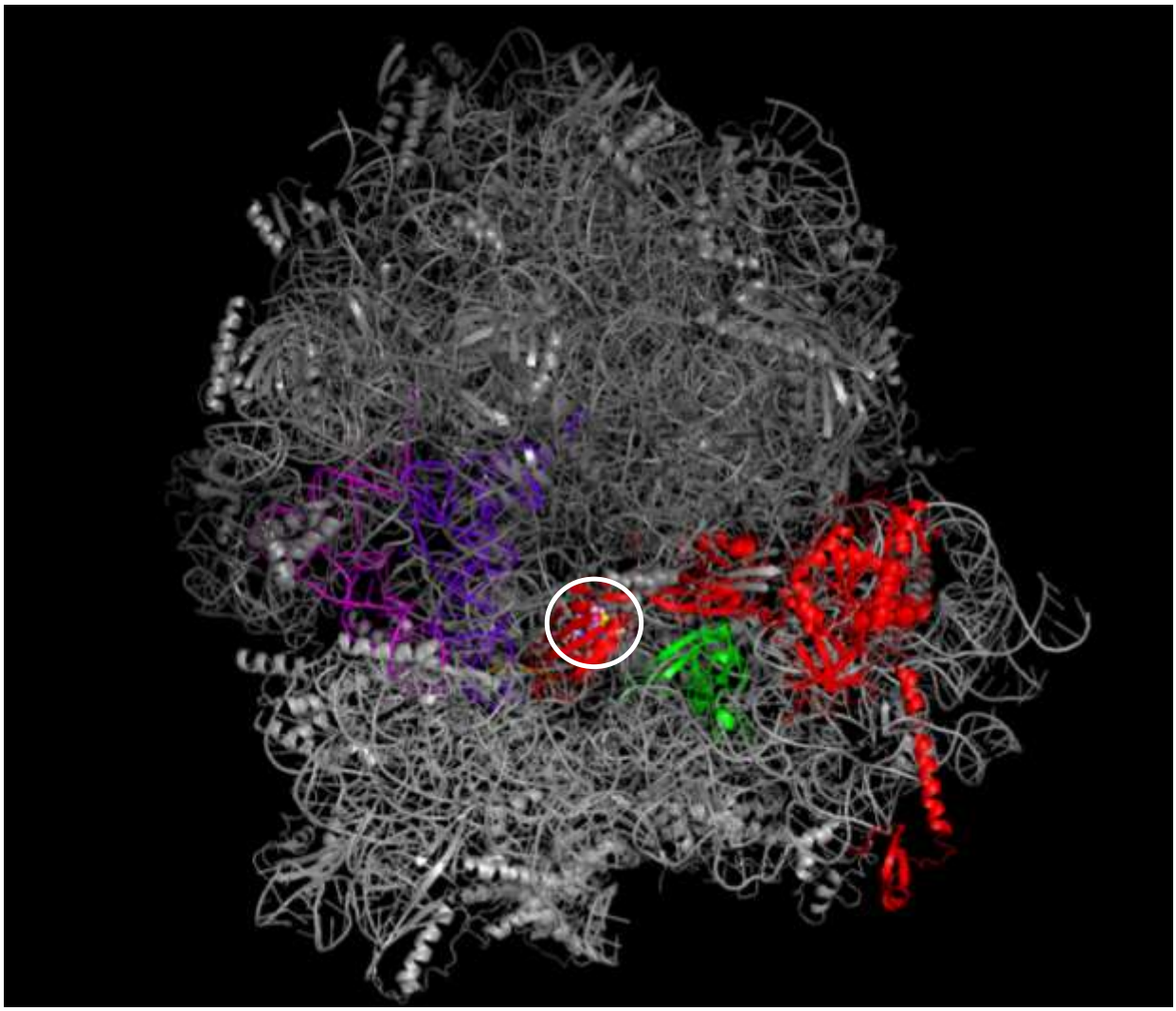


Figure 6.17 Docking simulation of all gentamicin poses within the 70s post-ribosomal complex. PyMOL image displaying the docking analysis image of all predicted gentamicin binding conformations illustrated using the space filling model representation (circled in white) docked to the 70s post-ribosomal complex.

The docking results were then visualised exclusively for the gentamicin conformation with the lowest binding energy of approximately -100.00kcal/mol within the 70s post-ribosomal complex (Figure 6.18). The binding energy for the docking of gentamicin within the post complex was lower than the lowest binding energy of -108.2 kcal/mol within the 70s pre-ribosomal complex. As domain 4 of EF-G is known to occupy the A-site tRNA within the post complex, it can be assumed that as gentamicin appears

behind domain 4 it is within the vicinity of the A-site tRNA. However, its position seemed much higher relative to the 30s-subunit protein of S12 and the mRNA as seen in the 70s pre-complex.

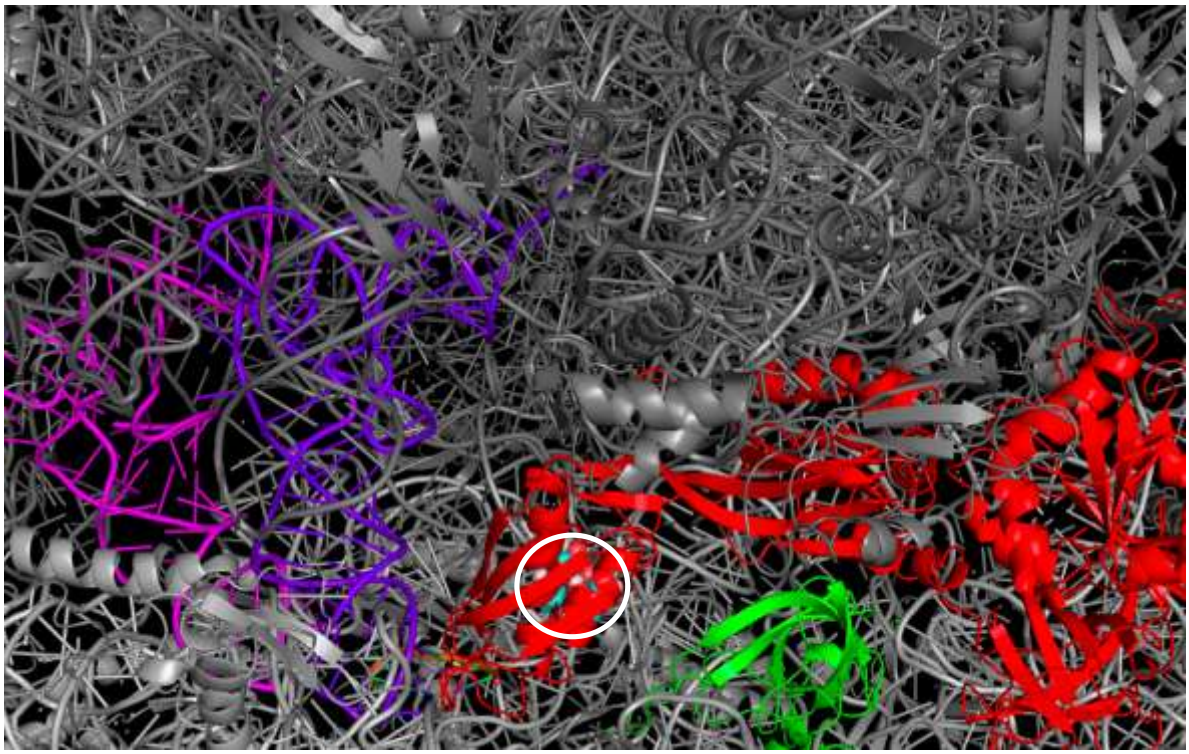


Figure 6.18 Docking simulation of gentamicin to the 70s post-ribosomal complex. Magnified PyMOL docking analysis image displaying the lowest gentamicin binding energy conformation depicted using a space filling model (circled in white) docked to the 70s post-ribosomal complex. The gentamicin-ligand is bound to the 70s post-ribosomal complex behind domain 4 of EF-G (red).

6.3.5 Gentamicin docking analysis post-mutagenesis

Ensuing the docking analysis of gentamicin within the pre- and post 70s ribosomal complex, it was necessary to determine whether the mutagenesis observed within the 16s rRNA and EF-G, affected gentamicin binding. Unfortunately, the 16s rRNA mutations were unable to be assessed using the PyMOL mutagenesis function, however the locality of these mutations were highlighted in yellow for all docking experiments post-mutagenesis.

6.3.5.1 70s pre-ribosomal complex docking analysis post-mutagenesis

Comparable to the docking analysis of gentamicin within the pre-ribosomal complex pre-mutagenesis 8 docking conformations were identified with almost undistinguishable binding energy profiles (Figure 6.19).

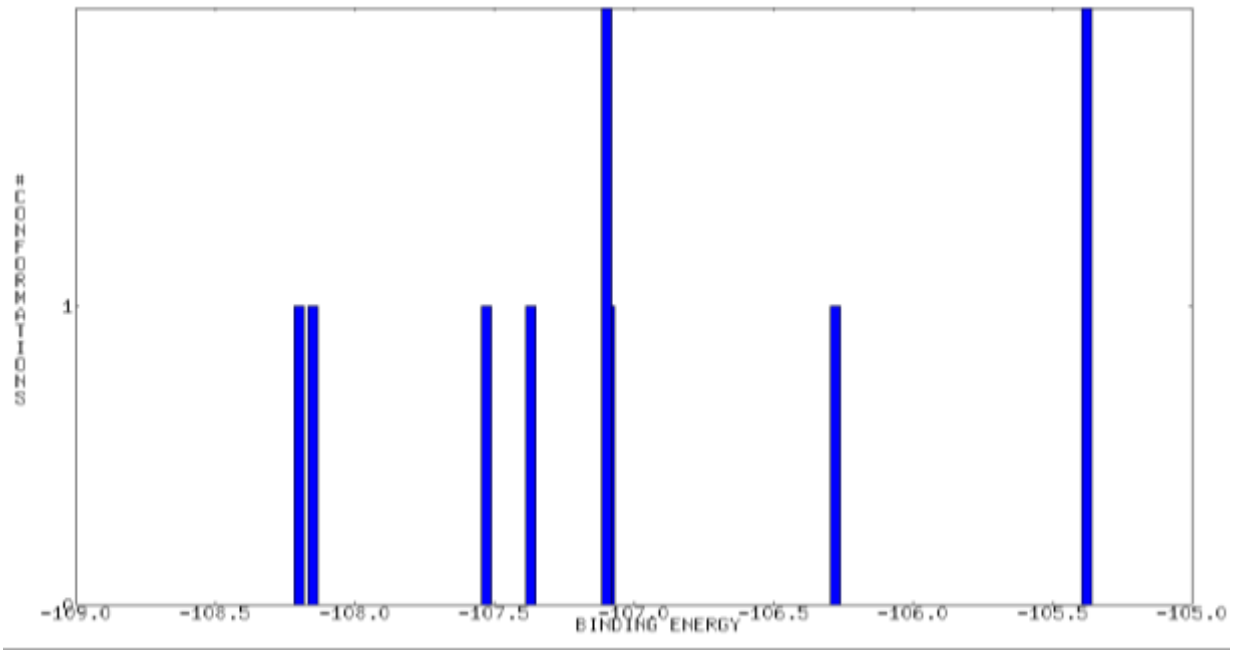


Figure 6.19 Binding energies of the gentamicin conformations obtained from the docking simulation with the 70s pre-ribosomal complex post-mutagenesis with Autodock. Graph showing the eight predicted gentamicin conformations and the corresponding binding energies.

Although there were parallels between the conformational binding energies pre- and post-mutagenesis, such comparisons were not identified when visualising the 8 docking poses within the 70s pre-ribosomal complex post mutagenesis (Figure 6.20). All the 8 gentamicin poses appeared higher than that seen pre-mutagenesis, residing within the A site-tRNA.

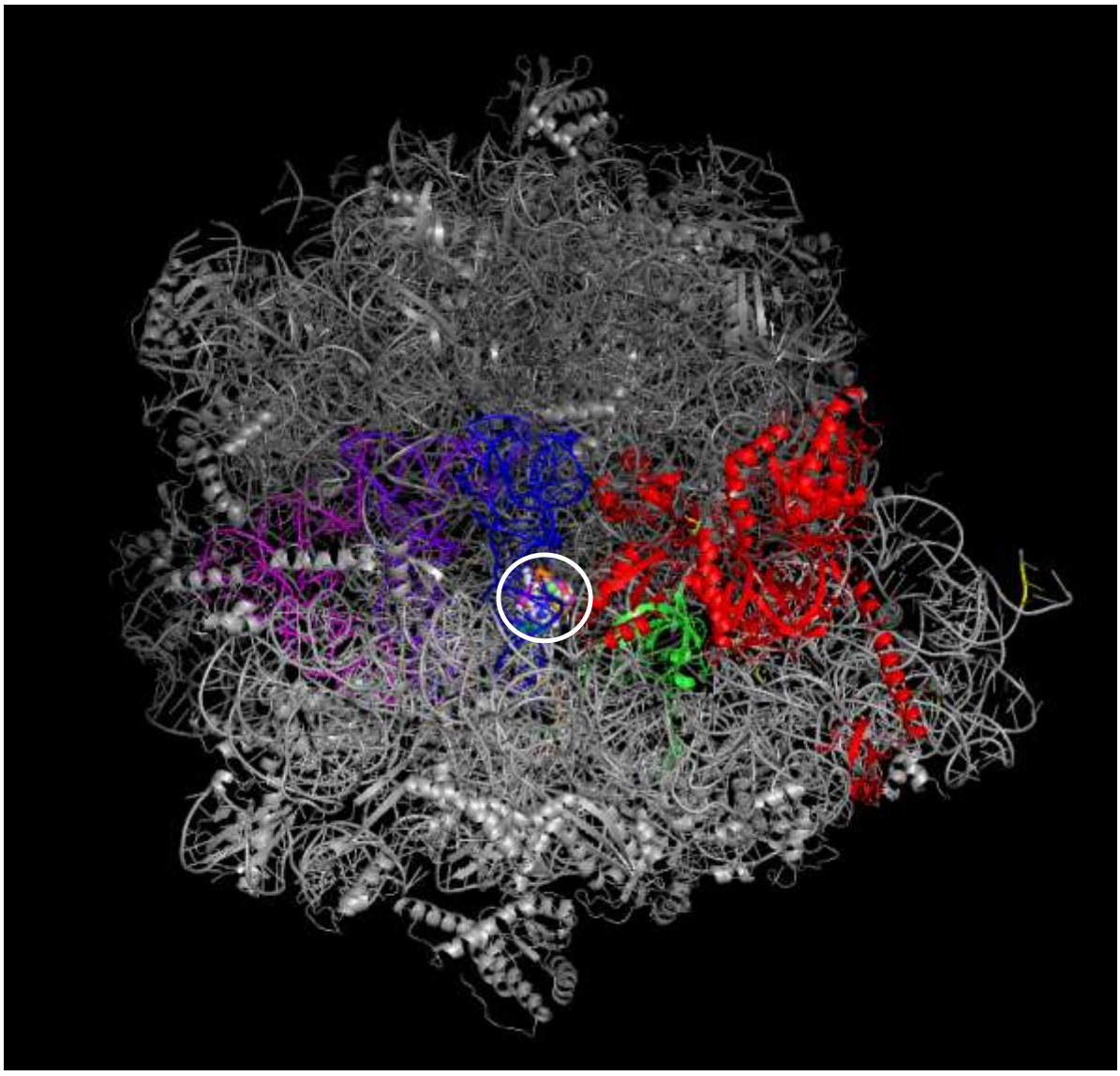


Figure 6.20 Docking simulation of all gentamicin poses within the 70s pre-ribosomal complex post-mutagenesis. PyMOL image displaying the docking analysis image of all predicted gentamicin binding conformations illustrated using the space filling model representation (circled in white) docked to the 70s pre-ribosomal complex post-mutagenesis. The mutations within the 16srRNA of the 30s-ribosomal subunit (light grey) are depicted in yellow and the EF-G (red) mutation is illustrated using the stick model representation of the amino acid alteration at position 610 (yellow).

The docking of gentamicin within the pre- complex post mutagenesis was then visualised singularly for the conformation with the lowest binding energy (-108.20 kcal/mol) (Figure 6.21). This result clearly indicates an alteration in the docking position of gentamicin post-mutagenesis, which appears in a similar location to the docking results for the 70s post-ribosomal complex pre-mutagenesis (Figure 6.17).

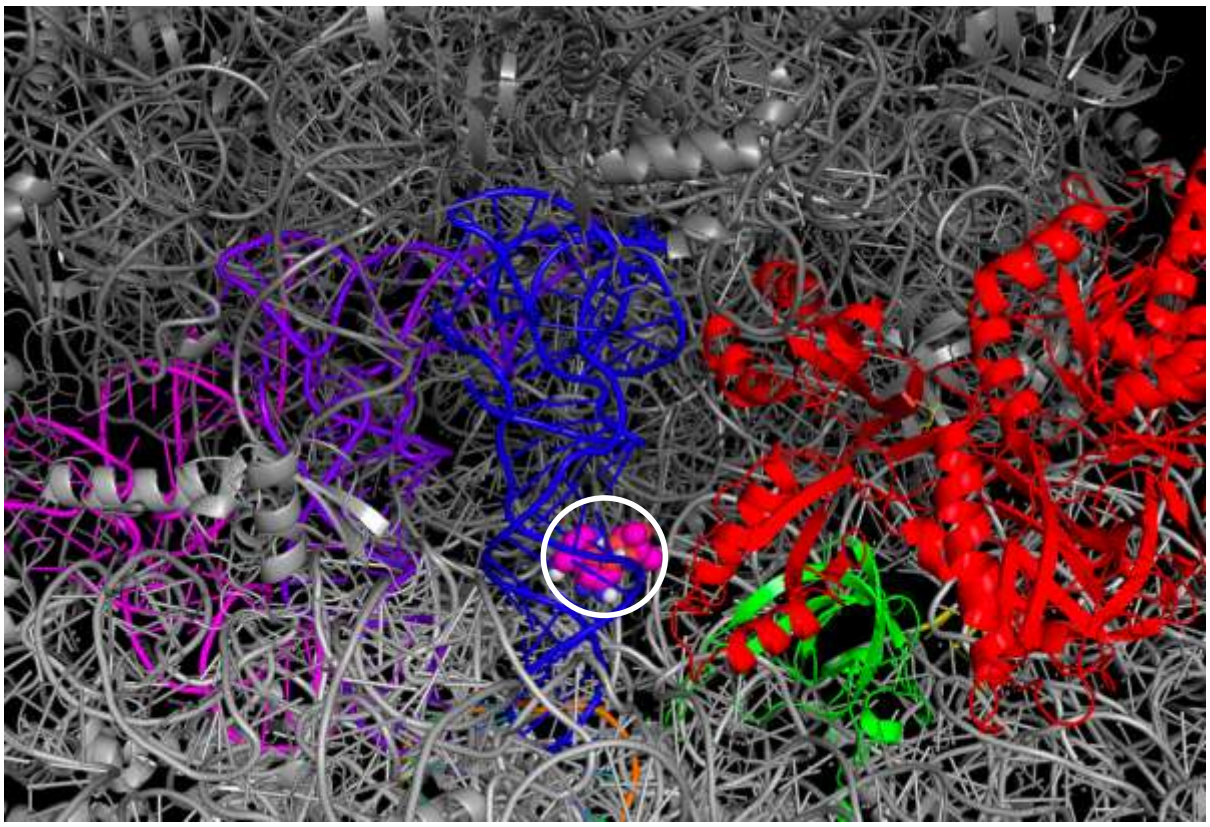


Figure 6.21 Docking simulation of gentamicin bound to the 70s pre-ribosomal complex post-mutagenesis. Magnified PyMOL docking analysis image displaying the lowest gentamicin binding energy conformation depicted using a space filling model (circled in white) docked to the 70s pre-ribosomal complex post-mutagenesis. The gentamicin-ligand is bound higher in the 70s pre-ribosomal complex structure in proximity to the A-site tRNA (blue) and of EF-G (red).

6.3.5.2 70s post-ribosomal complex docking analysis post-mutagenesis

As seen with the predicted binding energies post mutagenesis for the 70s pre-ribosomal complex, corresponding binding energy profiles were observed. A total of 9 conformations were predicted with binding energies ranging from -100.0 to -96.50 kcal/mol (Figure 6.22).

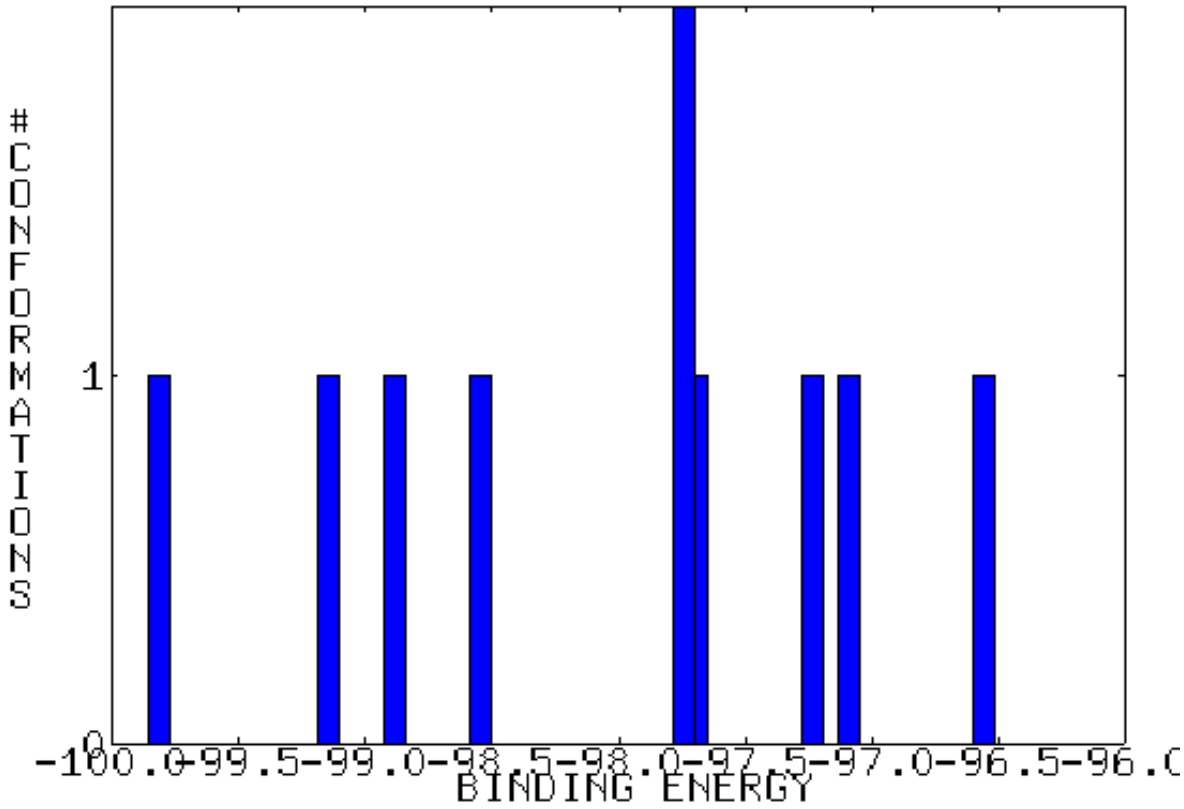


Figure 6.22 Binding energies of the gentamicin conformations obtained from the docking simulation with the 70s post-ribosomal complex post-mutagenesis with Autodock. Graph showing the nine predicted gentamicin conformations and the corresponding binding energies.

This similarity was reiterated when visualising the posed gentamicin docking positions within the 70s post-ribosomal complex post mutagenesis (Figure 6.23). Irrespective of the EF-G mutagenesis, gentamicin binds in precisely the same location within the post ribosomal complex, in the locality of domain 4 of EF-G.

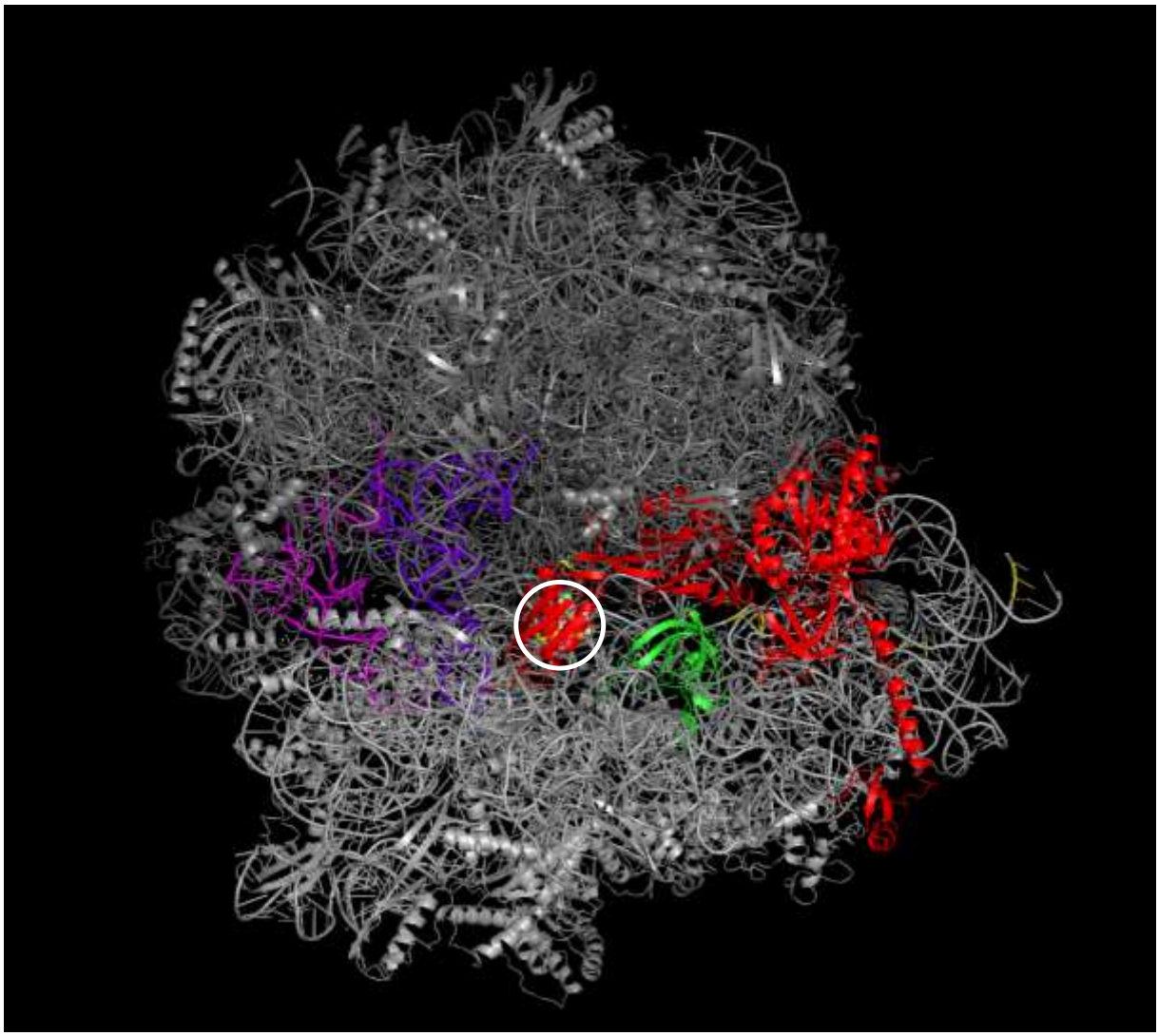


Figure 6.23 Docking simulation of all gentamicin poses within the 70s post-ribosomal complex post-mutagenesis. PyMOL image displaying the docking analysis image of all predicted gentamicin binding conformations illustrated using the space filling model representation (circled in white) docked to the 70s post-ribosomal complex post-mutagenesis. The mutations within the 16s rRNA of the 30s-ribosomal subunit (light grey) are depicted in yellow and the EF-G (red) mutation is illustrated using the stick model representation of the amino acid alteration at position 610 (yellow).

When viewing the specific docking pose of the lowest gentamicin binding conformation (-108.2 kcal/mol) within the 70s post-ribosomal complex post mutagenesis (Figure 6.24), the docking pose is indistinguishable from that observed pre-mutagenesis (Figure 6.20).

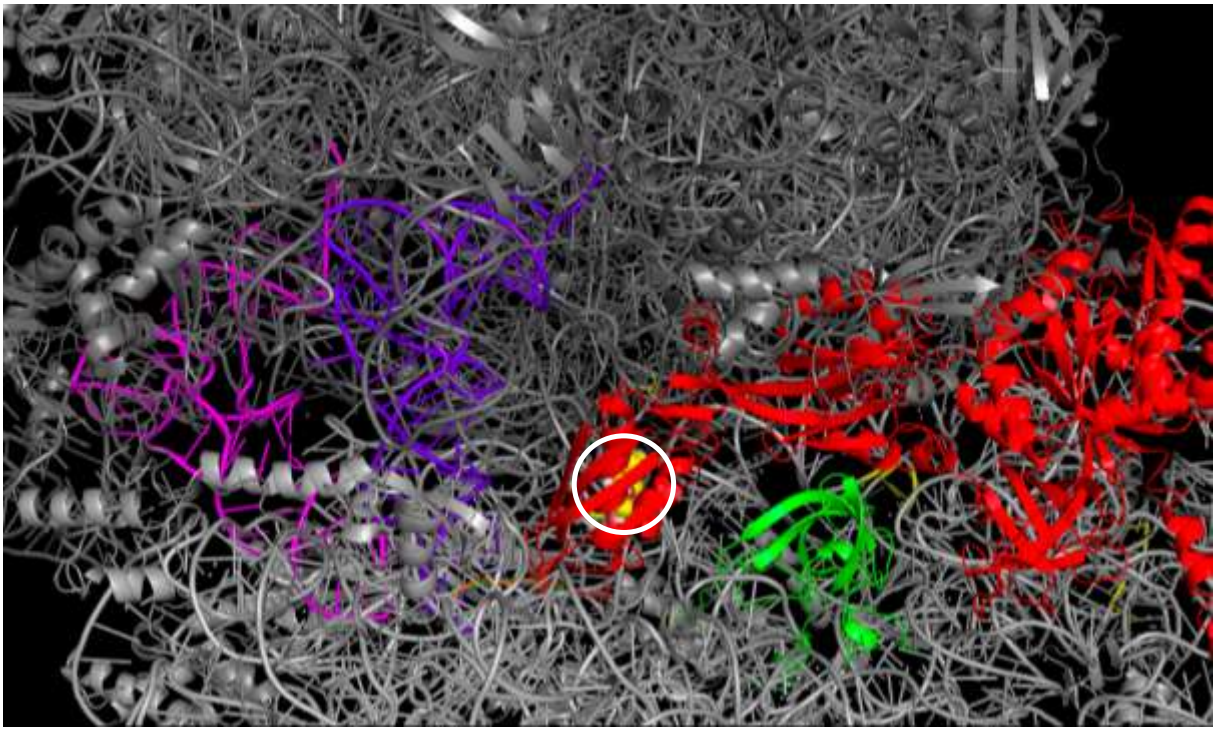


Figure 6.24 Docking simulation of gentamicin bound to the 70s post-ribosomal complex post-mutagenesis. Magnified PyMOL docking analysis image displaying the lowest gentamicin binding energy conformation depicted using a space filling model (circled in white) docked to the 70s post-ribosomal complex post-mutagenesis. The gentamicin-ligand is bound to the 70s post-ribosomal complex behind domain 4 of EF-G (red).

6.3.6 Comparative assessment of gentamicin docking

6.3.6.1 Gentamicin binding analysis 70s pre-ribosomal complex pre- and post-mutagenesis

To conclude the docking analysis, comparative visual assessments were made to determine the impact of the EF-G mutation on ribosomal translocation. There are apparent differences between the docking positions within the pre- ribosomal complex before and after mutagenesis (Figure 6.25).

This result suggests that the mutation identified within EF-G, is culpable for the altered binding of gentamicin post mutagenesis. Functionally, this implies that if gentamicin is unable to bind in close proximity to the mRNA and the base of the A-site tRNA, it may no longer be able to interrupt the wobble base formation and consequentially lead to a disruption of protein synthesis. This would ultimately result in gentamicin resistance.

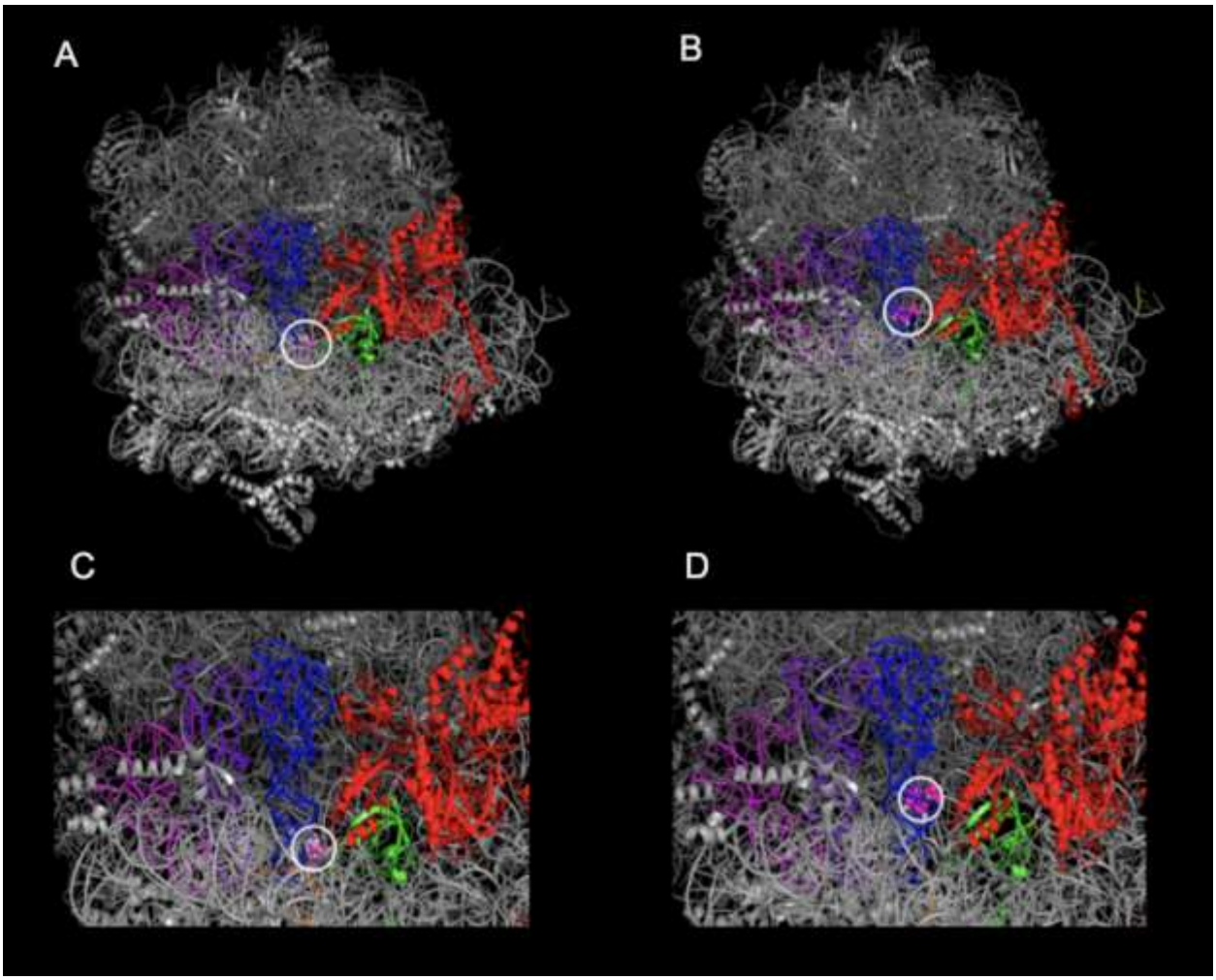


Figure 6.25 Docking simulation comparison of gentamicin bound to the 70s pre-ribosomal complex pre- and post-mutagenesis. PyMOL docking analysis image displaying the lowest gentamicin binding energy conformation depicted using a space filling model (circled in white) docked to the 70s pre-ribosomal complex pre-mutagenesis (A) and post-mutagenesis (B). The magnified docking images of gentamicin-ligand bound to the 70s pre-ribosomal complex pre-mutagenesis (C) and post-mutagenesis (D) demonstrate altered drug binding post-mutagenesis shown by the position of gentamicin in relation to the 30s-ribosomal subunit protein S12 (green), EF-G (red), A-site tRNA (blue) and mRNA (orange).

6.3.6.2 Gentamicin binding analysis 70s post-ribosomal complex pre- and post-mutagenesis

When visually comparing the docking positions within the post ribosomal complex before and after mutagenesis, no differences were identified (Figure 6.26). This result suggests that for the antimicrobial gentamicin to be effective in its activity, it must dock to the 70s-ribosome within the pre- ribosomal complex. If gentamicin fails to dock within the pre-state, regardless of the presence of the EF-G mutation it will not successfully bind in a position which leads to the interruption of protein synthesis.

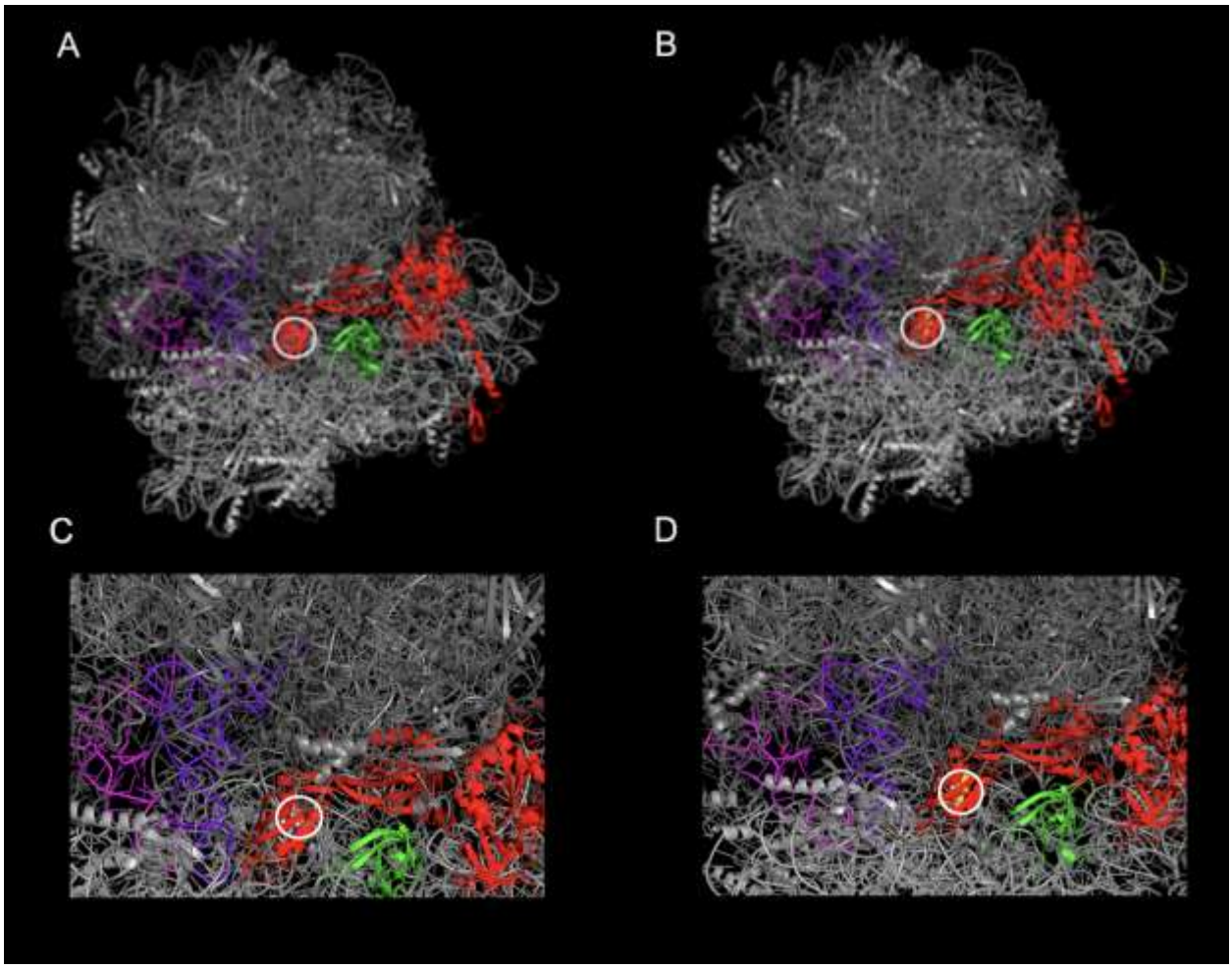


Figure 6.26 Docking simulation comparison of gentamicin bound to the 70s post-ribosomal complex pre- and post-mutagenesis. PyMOL docking analysis image displaying the lowest gentamicin binding energy conformation depicted using a space filling model (circled in white) docked to the 70s post-ribosomal complex pre-mutagenesis (A) and post-mutagenesis (B). The magnified docking images of gentamicin-ligand bound to the 70s pre-ribosomal complex pre-mutagenesis (C) and post-mutagenesis (D) suggests no alterations in drug binding post-mutagenesis indicated by location of gentamicin behind domain 4 of EF-G in all four images.

6.4 Discussion

The aim of this chapter described in section (6.1), was achieved through the identification of altered gentamicin binding to the 70s-ribosomal complex in its pre-translocation state, following *fusA* mutagenesis. The conclusions from the molecular docking analysis suggest that the mutation within EF-G, contributes significantly to the acquisition of gentamicin resistance within the proposed model.

The sequential multistep process of ribosomal translocation, was described extensively within chapter 4. Although EF-G independent translocation has been observed in-vitro (Southworth, Brunelle and Green, 2002), it was revealed that the GTPase EF-G, was an essential requirement in the rapid response to the cellular demands during translation elongation (Voorhees and Ramakrishnan, 2013). Major advancements in understanding the complex mechanisms in which EF-G catalyses ribosomal translocation, have been provided through the addition of small molecules which inhibit the translocation stage of protein synthesis (Peske *et al.,* 2004).

The macrocyclic thiopeptide, thiostrepton is an antimicrobial agent which inhibits translation through directly interfering with the function of EF-G (Bowen *et al.,* 2004). Interestingly, truncation of EF-G domains 4 and 5 permitted ribosomal translocation in the presence of thiostrepton (Walter *et al.,* 2011). Therefore, it is plausible to suggest mutagenesis at the boundary of domains 4 and 5, could permit EF-G to induce translocation despite the presence of the aminoglycoside gentamicin.

6.4.1 Aminoglycoside inhibition of ribosomal protein synthesis

Aminoglycoside antibiotics disrupt translation by inducing pleotropic effects on the ribosome in its pre-translocation complex, through diminishing the accuracy of tRNA selection at the decoding site and through the inhibition EF-G dependent translocation (Davies and Davies, 1967; Fredrick and Noller, 2003; Peske *et al.,* 2004). This confirms why no alterations were observed in the binding of gentamicin to the post-translocation complex ensuing mutagenesis, as aminoglycoside activity is dependent on drug binding to the ribosome in its pre- translocation state.

The predominant binding site of aminoglycoside antibiotics is helix 44 of the 16s rRNA major groove, where precise interactions are made with the decoding bases A1492 and A1493 (Ogle, Carter and Ramakrishnan, 2003). The perturbations which arise following aminoglycoside binding to the decoding region, are thought to stabilise the classical state of the pre-translocation complex and increase the affinity of tRNAs to the A-site (Vaiana and Sanbonmatsu, 2009). Consequently, as gentamicin operates through classical state stabilisation, and post- mutagenesis is unable to bind to the decoding region of the pre-translocation complex, it is highly unlikely that ribosomal translocation will be inhibited (Feldman *et al.,* 2010).

Although the structural basis of aminoglycoside induced miscoding is well understood, high resolution crystallography and single-molecule fluorescence resonance energy transfer (smFRET) techniques have been applied to understand the molecular mechanisms behind the highly dynamic process of translation elongation inhibition (Tsai *et al.,* 2013). Such studies have revealed secondary binding sites of aminoglycoside antibiotics including neomycin, which preferentially

binds to helix 69 of the 23s rRNA stabilising the hybrid state (P/E) position and consequently blocks the selection and translocation of aa-tRNA (Wang *et al.,* 2012).

6.4.2 The 30s-ribosomal subunit protein S12 and EF-G

An independent docking analysis was conducted with the ribosomal protein S12 owing to its involvement in the secondary binding interactions of gentamicin. The ribosomal protein S12 has been previously acknowledged for its functional role in the regulation of the pre-translocation state during aa-tRNA selection (Yates, 1979; Cukras *et al.,* 2003). Further studies have suggested that upon codon recognition the ribosomal protein S12 is involved in the signalling between the 30s-ribosomal subunit and elongation factor-Tu (EF-Tu) (Gregory, Carr and Dahlberg, 2009). Functional insights into the interactions between the ribosomal protein S12 and EF-G have been provided by biochemical studies of the antibiotics dityromycin and GE82832, which through binding to S12 inhibit the EF-G dependent ribosomal translocation (Brandi *et al.,* 2006; Buckley *et al.,* 2014).

The independent docking analysis, confirmed an interaction between gentamicin and the upper region of the ribosomal protein S12. This is significant as preceding the rotation of the 30s-subunit, the ribosomal protein S12 interacts with domain 3 of EF-G and blocks the conformational change from the compact to an elongated state, which becomes entirely inhibited in the presence of dityromycin (Lin *et al.,* 2016).

Therefore, based on the atomic resolution crystal structures of EF-G in complex with the 70s-ribosomal complex in the pre-translocation state, and through use of the PyMOL mutagenesis function, it is evident that the *fusA* mutation at the amino acid position 610 significantly effects the binding of gentamicin to the decoding region of

helix 44 near the ribosomal protein S12. This strongly indicates that the *fusA* mutagenesis, is the primary mechanism preventing the inhibition of ribosomal translocation by gentamicin within the resistant isolates of *E. coli* MG1655 obtained from the ARGP *in vitro*.

Chapter 7: Phylogeny and evolution of EF-G: a phylogenetic perspective

7.1 Introduction

In the previous chapters, WGS and subsequent bioinformatic analysis has suggested that the observed *fusA* mutation is strongly associated with the acquisition of gentamicin resistance. This SNP has not only been shown to alter the structure and physiochemical properties of the *fusA* encoding EF-G protein, but docking analysis has revealed the unsuccessful binding of gentamicin to the 70s-ribosome post mutagenesis in its pre-translocation state. Together, these results imply that if the *fusA* mutation was identified within other bacterial species under gentamicin selection, populations may resist the lethal action of aminoglycoside antibiotics. A mutation identified in a functional gene such as *fusA* could be used as a marker of resistance or assist in antimicrobial resistance surveillance. Phylogenetics, will be applied within this chapter to unravel the evolutionary information stored within the identified genes of interest, to gain better insights into the patterns and processes of molecular evolution (Soltis and Soltis, 2003).

Phylogenetics has facilitated the study of patterns in the evolutionary conservation of genes across organisms (Bergmiller, Ackerman and Silander, 2012). This is significant, as orthologs of genes which are functionally essential and highly conserved, should remain essential across organisms of different bacterial taxa (Jordan *et al.,* 2002). Relating specifically to selection, if orthologous genes are well conserved between organisms, this is indicative of purifying selection acting strongly on essential genes, manifesting in slow rates of evolution to maintain function (Silander and Ackerman, 2009). The evolutionary conservation of an elongation

factor gene '*tuf*' which encodes EF-Tu has been previously described (Isabel *et al.,* 2008). The EF-Tu gene is highly conserved, and has been used to resolve the phylogeny of the *Enterobacteriaceae* family, with better discrimination in bacterial classification than the 16s rRNA (Paradis *et al.,* 2005).

In addition to evolutionary conservation, phylogenetics enables the study of how genes evolve. Evolution occurs when the genetic composition of a population changes over a period time, and such variations can occur through a range of processes including natural selection, mutation and genetic drift (Lenski and Wiser, 2009). Adaptive evolution arises through the selection of mutations which are associated with enhancements of the fitness or survivability, in response to environmental conditions (Charlesworth, Barton and Charlesworth, 2017). Evolutionary adaptation is common amongst bacterial populations due to the frequent exposure to variable environments (Bennett and Lenski, 2007). However, the way in which organisms evolve in response to their environment can vary significantly. Stochastic processes such as mutation and drift lead to species divergence, whereas natural selection promotes convergence through similar adaptations in distantly related organisms (Deatherage *et al.,* 2017).

The adaptive evolution of bacterial populations has largely been explored through experimental evolution (Dettman *et al.,* 2012). Microorganisms are applicable for experimental evolution studies due to their rapid generation times and genetic adaptation when placed in varying environments (Elena and Lenski, 2003). The combination of laboratory evolution with comparative genomics, allows the direct observation of adaptive evolution within a laboratory timescales (Herring *et al.,* 2006).

The highly controlled nature of experimental evolution studies facilitates the direct application of selective pressures to evolving populations to determine the underlying adaptive mutations leading to evolutionary change (Lang and Desai, 2014). Nevertheless, the short evolutionary timescales and regulation of selective pressures within experimental evolution studies, forces the selection of beneficial mutations, which may not present in natural environments (Hegreness and Kishony, 2007).

Phylogenetic perspectives provide the basis for comparative genomics. Advancements in DNA sequencing technologies have facilitated an increase in the number of comparative genomic studies aimed at detecting sequence diversity amongst bacterial populations. Despite the success of comparative genomics, the identification of variants alone does not provide enough information regarding bacterial evolution, pathogenicity and epidemiology (Alland *et al.,* 2003). Evolutionary events leave distinct patterns within bacterial populations which can be characterised through phylogenetic analysis (Guttman and Dykhuizen, 1994). Combined phylogenetics and comparative genomics reveals evolutionary relationships between bacterial strains, and allows the clustering of nucleotide polymorphisms of greater biological significance (Filliol *et al.,* 2005).

Combined phylogenetic analysis and comparative genomics has been applied in epidemiological studies of disease outbreaks, the classification of emerging bacterial pathogens and the identification of disease specific markers (Brenner *et al.,* 2017; Kuroda *et al.,* 2010; Gagneux and Small, 2007). Of significance, is the identification of phylogenetic markers associated with the acquisition of antimicrobial resistance (Feuerriegel, Köser and Niemann, 2014). This has been decisive in the emergence

of resistance mutations among organisms of clinical significance such as carbapenemase producing *Enterobacter* spp. (Chavda *et al.,* 2016).

The initial aim of this chapter is to determine the level of gene conservation of *fusA* through the reconstruction of evolutionary relationships among the *Escherichia* taxa and *Enterobacteriaceae* family. Secondly, to establish how the *fusA* sequence evolves, explicitly whether the identified SNP is an adaptive response to the antimicrobial selective pressure, and thus convergence may have occurred within distantly related species within *Enterobacteriaceae* family. Finally, to discover whether the SNP identified, is found within the wild amongst clinically significant bacterial pathogens which pose a threat in the antimicrobial resistance crisis.

It is hypothesised that the identified SNP within the *fusA* gene will remain unique to the experimentally evolved strain of *E. coli* MG1655 after screening distantly related strains within the *Enterobacteriaceae* family and strains of clinical significance.

7.2 Materials and methods

7.2.1 Phylogenetic analysis of *fusA* using PhyML

PHYML software version 3.1 was used for estimating maximum likelihood phylogenies, described previously in section (2.12). Phylogenies of the *fusA* gene were resolved for the *Enterobacteriaceae* family. The optimal model for each of the data sets was determined using JModelTest2 (Section 2.12.1). Conclusively, PhyML was executed under the parameters outlined in section (2.12.2), for the construction of the phylogenetic trees. The PHYML command line strings used to construct the *fusA* ML trees, are shown below (Figure 7.1). All phylogenetic trees were visualised and annotated using iTOL.

1. MacBook-Pro:~ Lucky$ PhyML-3.1_macOS-MountainLion -b 1000 -m 011023 -i Desktop/*Enterobacteriaceae*_Tree.phylip -d nt

2. MacBook-Pro:~ Lucky$ PhyML-3.1_macOS-MountainLion -b 1000 -m 011023 -i Desktop/Enterobacteriaceae_Distinct_Clade_Tree.phylp -d nt

Figure 7.1 The PhyML commands executed in the *fusA* ML tree construction. Image displaying the PhyML commands utilised to construct the ML trees of the *fusA* gene within the *Enterobacteriaceae* family (1) and within distinct clades of the *Enterobacteriaceae* family (2). The command contains the PhyML execution followed by a list of parameters: number of bootstraps (-b), model (-m), alignment file (-i) and the data type (-d). The commands utilised the same model (011023), 1000 bootstraps and the data type was nucleotide (nt).

7.2.2 SMART structure analysis of EF-G

To identify any structural variants within the *fusA* encoding EF-G protein within the *Enterobacteriaceae* family, a Simple Modular Architecture Research Tool (SMART) was used available at: http://smart.embl-heidelberg.de [accessed October 4, 2018], to provide detailed annotations regarding the protein architecture and functionally important domains.

7.2.3 Phylogenetic analysis of the 16s rRNA

Escherichia coli ATCC 11775T strain 16s rRNA gene was used as the consensus gene sequence for the resultant 16s species tree for the Enterobacteriaceae family. A BLAST search was conducted as previously described in section (2.12) and downloaded from the NCBI database. A total of 19 Enterobacteriaceae 16s rRNA genes along with the corresponding gene variants, which for most strains there was a total of seven 16s rRNA variants downloaded. Ribonucleotide sequences were reverse complemented were necessary. Sequence files were then imported into EMBL: MUSCLE and progressively aligned.

RNA alignment files were visualised using Aliview to assess the degree of gaps and sequence length discrepancies. TrimAl version 1.2 available at http://trimal.cgenomics.org [accessed August 29, 2018]. was used to remove all positions in the alignment with gaps in more than 10% of the sequences and prints the 60% best positions using the command below (Figure 7.2) and files were outputted in phylip format for direct manipulation in PhyML.

```
MacBook-Pro:~ Lucky$./trimal -in /Users/Lucky/Desktop/16s_New_3.fasta -out
/Users/Lucky/Desktop/16s_Trimmed.phylip -gt 0.9 -cons 60 –phylip
```

Figure 7.2 The TrimAl command executed for the removal of poorly aligned regions of the 16s rRNA sequences. Image displaying the TrimAl command utilised to remove regions of the 16s rRNA sequences with gaps or sequence read discrepancies. The command contains the TrimAl execution followed by a series of parameters: input file (-in) in fasta format, output file (-out) in phylip format, gap threshold (-gt) and the minimum threshold of positions conserved within the original alignment (-cons). The command utilised specifies a gap threshold of 0.9 with a 60% sequence conservation threshold.

Alongside this, the alignment file (fasta format) was imported directly into JModelTest2 to determine the best fit substitution model. Following the likelihood calculations, the model selection uncertainty was calculated for the three implemented information criteria models (AIC, BIC and decision theory (DT)), as per specifications described in section (2.12.1). The rRNA alignment (phylip format) was used to construct a ML tree using PhyML with a bootstrap of 1000 using the model 012342 as illustrated in the command below (Figure 7.3). The 16s rRNA species tree was visualised using iTOL.

```
MacBook-Pro:~ Lucky$ PhyML-3.1_macOS-MountainLion -b 1000 -m 012342 -i
Desktop/16s_Trimmed.phylip -d nt.
```

Figure 7.3 The PhyML command executed in the 16s rRNA ML tree construction. Image displaying the PhyML command utilised to construct the ML tree of the 16s rRNA gene within the *Enterobacteriaceae* family. The command contains the PhyML execution followed by a list of parameters: number of bootstraps (-b), model (-m), alignment file (-i) and the data type (-d). The command specifies 1000 bootstraps, the 012342 model, the 16s-trimmed alignment file and the data type was nucleotide (nt).

7.2.4 Protein conservation analysis of EF-G

Distinct Enterobacteriaceae sequences used for the construction of the *fusA* nucleotide tree, were used to assess the level of protein conservation for the *fusA* encoding protein Elongation Factor G. The *fusA* nucleic acid sequences were translated using EMBOSS Transeq available at: https://www.ebi.ac.uk/Tools/st/emboss_transeq/ [accessed August 20, 2018]. The selected parameters included the selection of the bacterial code 11 and all six frames. The translated sequences for the correct frame, were then downloaded and inputted into PRALINE, where in addition to a multi sequence alignment information can be obtained regarding the homology and amino acid conservation per residue of the given protein sequences. PRALINE was run under standard conditions, aside from the deselection of structural features which was not required.

7.2.5 Bioinformatic analysis of EF-G: *Enterobacteriaceae* database

The phylogeny of the *E. coli* MG1655 *fusA* gene was resolved by conducting a nucleotide BLAST search of the NCBI database repository as described in section (2.12). A further local alignment search was conducted at Public Health England (PHE), of the National Collection of Type Cultures (NCTC) culture collection in collaboration with Dr Katie Hopkins and Professor Neil Woodford. The *fusA* nucleotide sequence of the resistant strain of *E. coli* MG1655, was submitted to PHE and screened by the bioinformatician Dr Michel Doumith against a collection of 840 carbapenemase-producing and 958 ESBL-producing *E. coli* genomes.

7.3 Experimental results

7.3.1 Best fit model selection for the phylogenetic analysis of *fusA*

Preceding the phylogenetic *fusA* nucleotide tree construction within PhyML, best fit models were selected based upon the likelihood (–InL) scores for the AIC, BIC and DT model criterion. The preferred model for the data sets used in the phylogenetic analysis of *fusA* amongst the *Enterobacteriaceae* family was 011023. When looking more closely at the predicted scores generated using JModelTest2, all three of the selection criterion predicted models, 011020 and 011023 as best fitting for the two data sets (Tables 7.1 and 7.2).

Table 7.1 JModelTest2 model selection summary for the *fusA* nucleotide sequence data within *Enterobacteriaceae* family. Displaying the selection of best fit model using the AIC, BIC and DT selection criterion. K indicates the number of parameters. Based on the JModel results summary, the 011023 model was selected as the best fit model. The best model was defined as the model with the lowest criterion value (-InL) indicated in bold.

Model Selection Criterion	Model	Partition	-InL	K	AIC/BIC/DT
AIC[1]	011023+I+G+F	011023	**5779.887**	44	11647.775
AIC[2]	011020+I+G	011020	5786.407	40	11652.813
BIC[1]	011020+I+G	011020	5786.407	40	11879.029
BIC[2]	011023+I+G	011023	5783.029	41	1879.928
DT[1]	011020+I+G	011020	5786.407	40	0.008
DT[2]	011023+I+G	011023	5783.029	41	0.008

Table 7.2 JModelTest2 model selection summary for the *fusA* nucleotide sequence data within distinct clades of the *Enterobacteriaceae* family. Displaying the selection of best fit model using the AIC, BIC and DT selection criterion. K indicates the number of parameters. Based on the JModel results summary, the 011023 model was selected as the best fit model. The best model was defined as the model with the lowest criterion value (-InL) indicated in bold.

Model Selection Criterion	Model	Partition	-InL	K	AIC/BIC/DT
AIC[1]	011023+I+G+F	011023	**5517.405**	38	11110.809
AIC[2]	011020+I+G	011020	5524.851	34	11117.701
BIC[1]	011020+I+G	011020	5524.851	34	11309.984
BIC[2]	011023+I+G	011023	5521.026	35	11309.990
DT[1]	011020+I+G	011020	5524.851	34	11309.984
DT[2]	011023+I+G	011023	5521.026	35	11309.990

7.3.2 Phylogenetic analysis of *fusA* within the *Escherichia* genus

The ML tree constructed for the analysis of *fusA* amongst the *Escherichia genus* has not been included here as all strains were monophyletic, and the slight topology evident was due to the inclusion of *fusA* sequences of the *Shigella* genus (Appendix A.21). Further haplotype analysis of *fusA* within the species *E. coli (sensus stricto)*, revealed a total of two haplotypes, the resistant and the susceptible.

7.3.3 Phylogenetic analysis of *fusA* amongst the *Enterobacteriaceae* family

The resolved phylogeny of the *fusA* gene within the *Enterobacteriaceae* family demonstrated a high level of conservation amongst the *Escherichia* genus. Although the conservation of *fusA* within the *Escherichia* genus, is distributed amongst strains of the *Shigella* genus. The sensitive and resistant *fusA* sequences of the laboratory strains of *Escherichia coli* MG1655, have been highlighted in yellow and the observed distances between the two strains was likely reflective of the uncovered nucleotide polymorphism. There is slight topology evident within the ML tree, shown as an

outgroup comprising members of the *Enterobacteriaceae* family which do not fall into the *Escherichia*/*Shigella* genus. The tree topology is therefore suggestive of speciation or related to the taxonomic identification of bacteria (Figure 7.4 A).

It is essential to identify whether the topological differences observed, are reflective of any structural alterations within the functional domains of EF-G. SMART analysis of the *fusA* encoding EF-G protein between distantly related species within the *Enterobacteriaceae* family*,* revealed identical EF-G protein structures with conserved functional domains (Figure 7.4 B). The conservation of the EF-G domains despite the observed tree topology, indicates the biological significance of EF-G and further suggests that the tree topology is related to speciation.

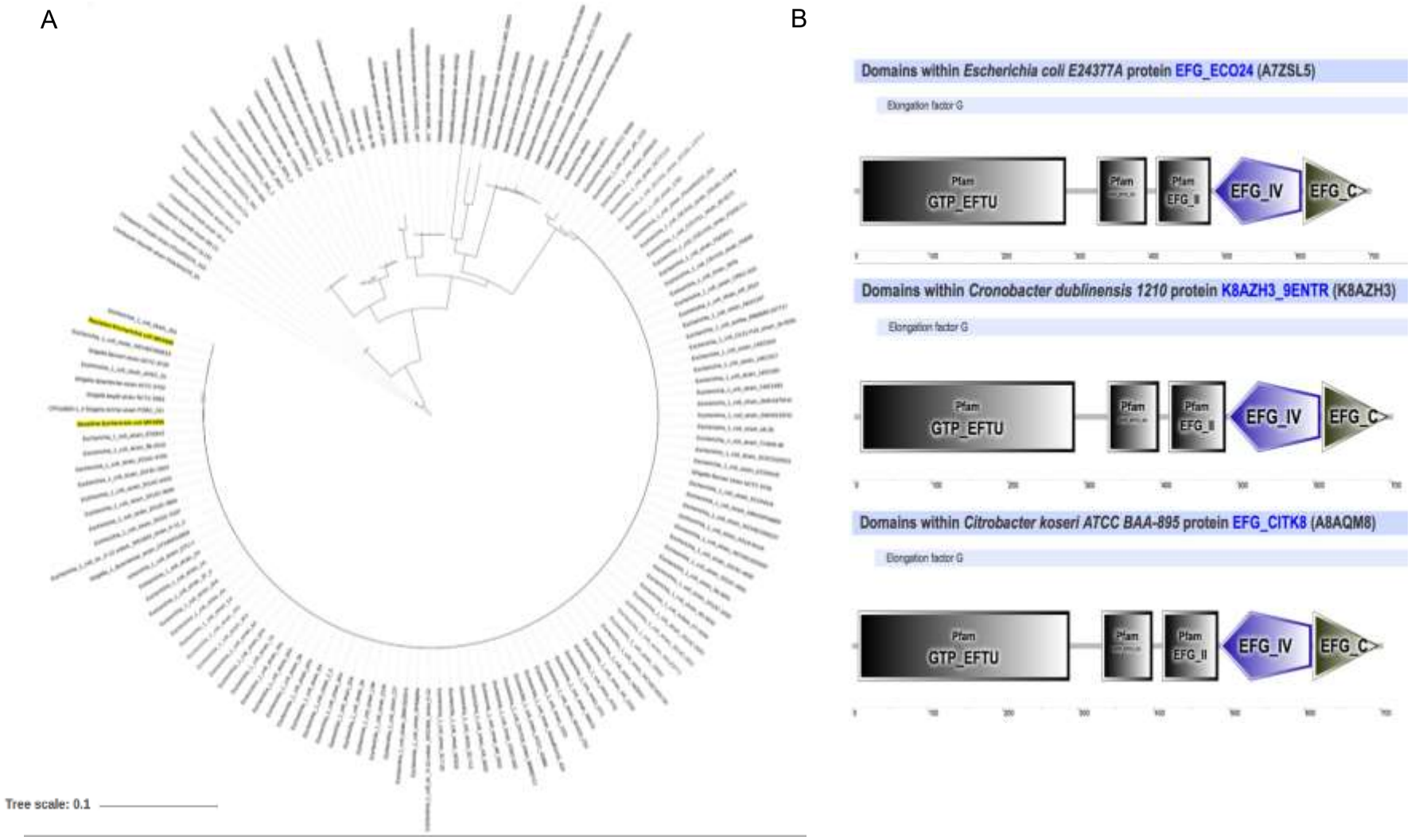


Figure 7.4 Phylogenetic analysis of the *fusA* gene and EF-G within the *Enterobacteriaceae* family. (A) Displaying the ML tree for the *fusA* gene within species of the *Enterobacteriaceae* family constructed using PhyML. The best fit model was calculated for the data set (011023 model) and trees were bootstrapped 1000 times. The ML tree was visualised in iTOL using the radial display mode. The sensitive and resistant strains of *E. coli* MG1655 were highlighted in yellow. (B) SMART Analysis of EF-G. Displaying the functional domains within the EF-G protein sequence. SMART analysis was conducted on distantly related species within the *Enterobacteriaceae* family (*E. coli*, *Cronobacter dublinensis* and *Citrobacter Koseri*).

7.3.3.1 Phylogenetic analysis of *fusA* amongst distinct clades of the *Enterobacteriaceae* family

To determine whether the evolution of the *fusA* gene suggests convergent evolution or speciation, a ML tree was constructed for the most distantly related species within the *Enterobacteriaceae* family and information was obtained regarding the strain source and locality (Figure 7.5). If convergent evolution of the *fusA* gene had occurred, it would be expected that strains would cluster based on source irrespective of the species. However, strains within the *Enterobacteriaceae* family which are classified as pathogenic are found disseminated throughout the tree and are largely clustered based on genus. However, supporting the previous findings again the *fusA* sequence is highly conserved amongst the *Escherichia* genus.

7.3.4 Haplotype analysis of *fusA* within the *Enterobacteriaceae* family

In addition to the phylogenetic analysis of *fusA* within the *Enterobacteriaceae* family, an examination of *fusA* haplotypes was completed. A total of 147 *fusA* sequences were examined, and 38 haplotypes were identified. Of the 38 haplotypes, 37 haplotypes were species specific with one isolated haplotype specific to the *fusA* mutation in the resistant strain of *E. coli* MG1655 (Appendix Figure A.22). Looking more closely at the 37 species specific haplotypes there was significant variation within 340 independent nucleotide bases of the 2112 nucleotides of the *fusA* sequence (Appendix Figure A.23).

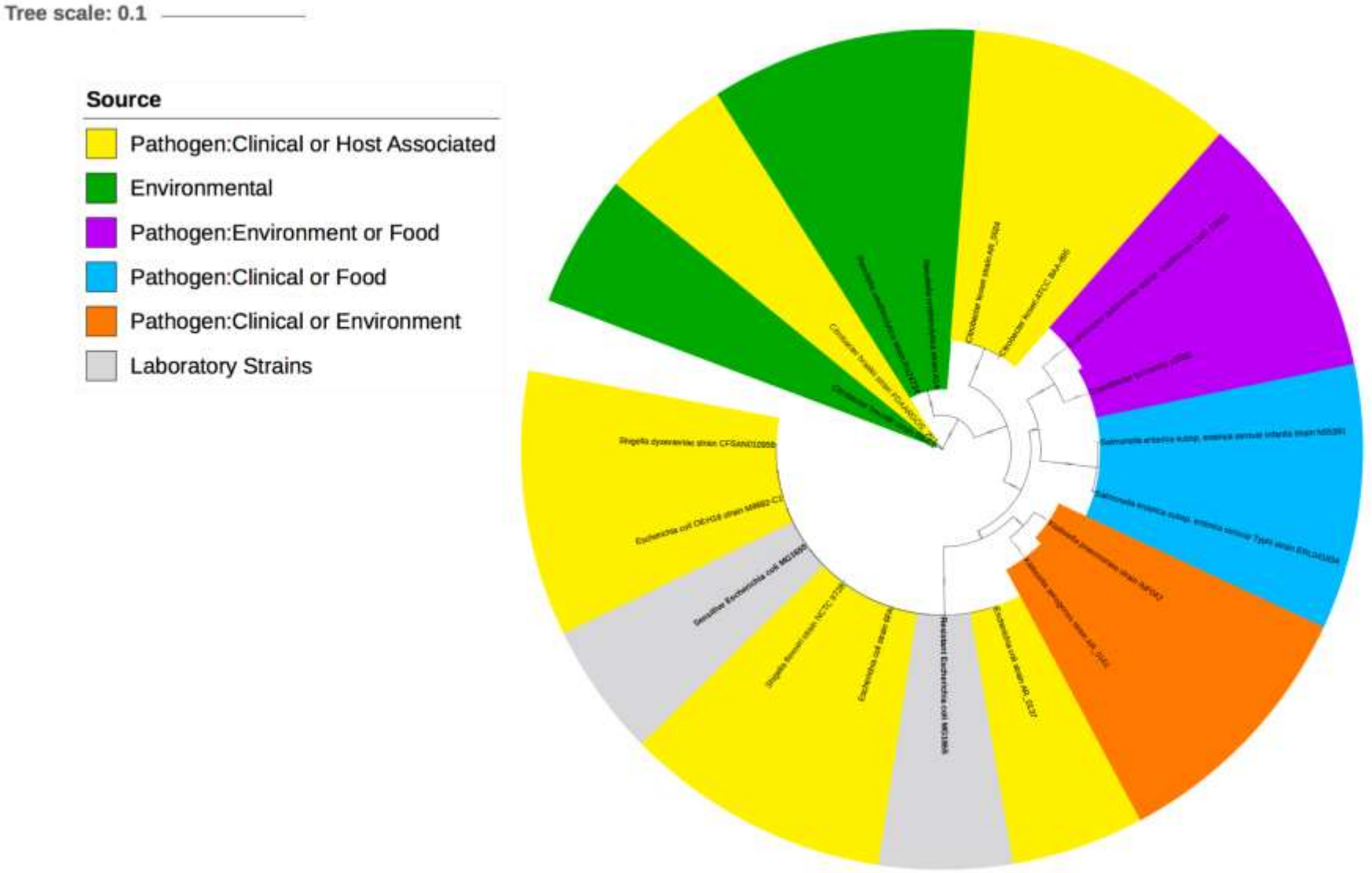


Figure 7.5 Phylogenetic analysis of the *fusA* gene within distantly related species of the *Enterobacteriaceae* family. Displaying the ML tree for the *fusA* gene within distinct species of the *Enterobacteriaceae* family using PhyML. The best fit model was calculated for the data set (011023 model) and trees were bootstrapped 1000 times. The ML tree was visualised in iTOL using the radial display mode. The *Enterobacteriaceae* species within the ML were labelled at the branch tips and coloured based on the strain source/locality: clinical (yellow), environmental (green), pathogen environment/food (purple), pathogen clinical/food (blue), pathogen clinical/environment(orange) and laboratory (grey). The sensitive and resistant laboratory strains of *E. coli* MG1655 were highlighted in bold.

7.3.5 Phylogenetic analysis of the 16s rRNA genes amongst distinct clades of the *Enterobacteriaceae* family

All the evidence gathered formerly suggests the evolution of *fusA* is directly related to speciation. To reaffirm the findings, a phylogenetic tree was constructed based on the 16s rRNA sequences of the most distantly related species within the *Enterobacteriaceae* family. The 16s rRNA sequences are often used in the study of bacterial phylogeny and taxonomy, owing to the high conservation and slow evolution of the rRNA genes plus their high prevalence amongst bacterial species (Janda and Abbott, 2007). Members of the *Enterobacteriaceae* family were coloured based on species and numbered based on the rRNA gene variant (Figure 7.6). The tree suggests that the 16s rRNA sequences are clustered into phylogenetically groups based on species. As previously described, there are deviations between the *Escherichia* and *Shigella* genus, shown by the intermittent strain highlighted in grey between the *Escherichia* strains coloured in red. Owing to the comparable clustering of sequences shown in figures (7.5 and 7.6), it can be interpreted that the evolution of *fusA* is species specific.

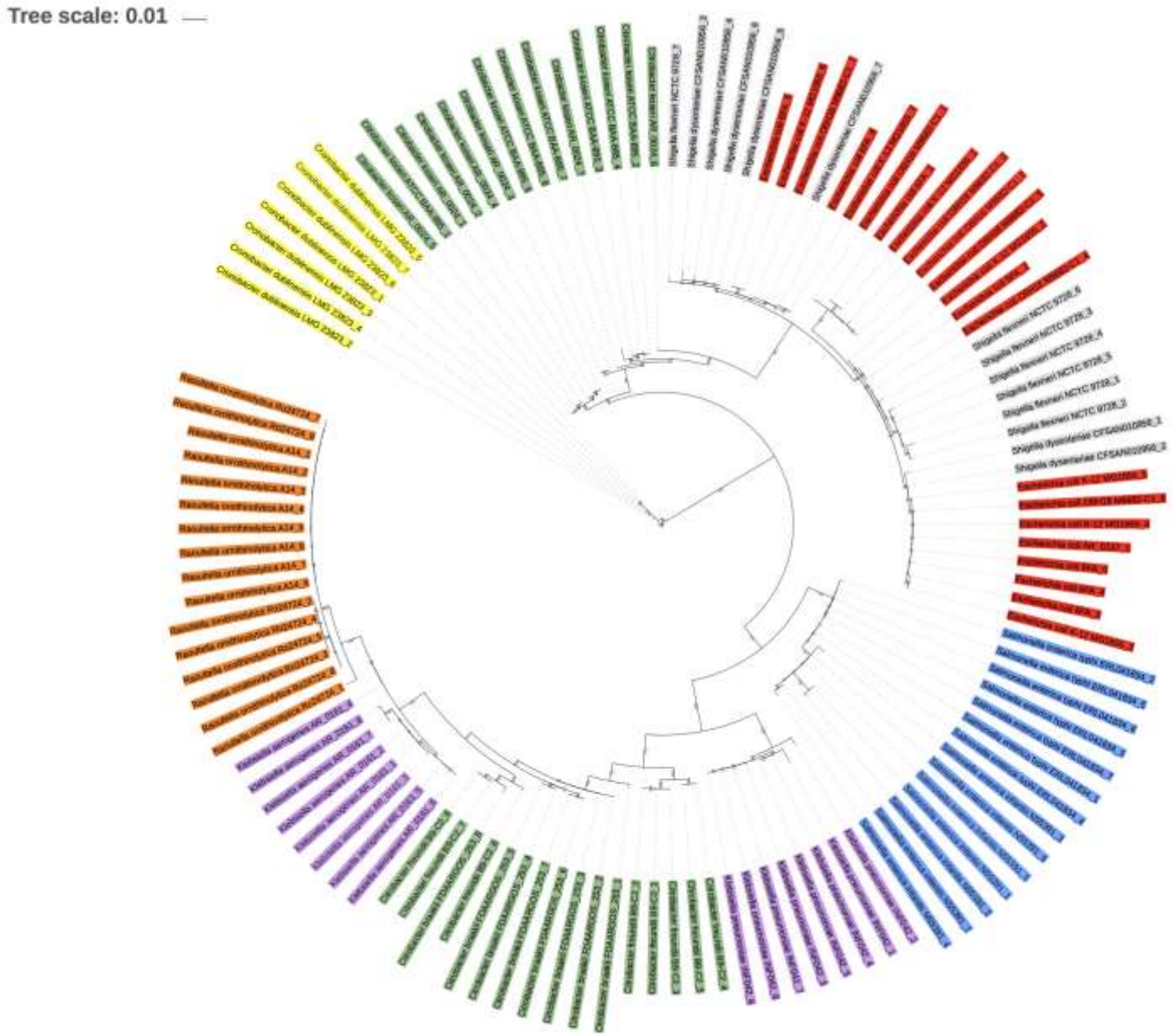


Figure 7.6 Phylogenetic analysis of the 16s rRNA genes within the *Enterobacteriaceae* family. Displaying the ML tree for the 16s rRNA genes within the *Enterobacteriaceae* family using PhyML. The best fit model was calculated for the data set (012342 model) and trees were bootstrapped 1000 times. The ML tree was visualised in iTOL using the radial display mode. The *Enterobacteriaceae* strains were coloured by species: *Cronobacter* (yellow), *Citrobacter* (green), *Klebsiella* (purple), *Escherichia* (red), *Shigella* (grey) and *Raoultella* (orange). The *Enterobacteriaceae* strains were labelled with numbers indicating the 16s rRNA gene variant.

7.3.6 EF-G protein conservation analysis PRALINE

The PRALINE protein conservation analysis of EF-G exposed a high degree of protein conservation across the entire protein sequence, with 4 being the lowest protein conservation score reported at a single amino acid position (Figure 7.7). There was complete conservation achieved for the first 150 amino acids, reflective of the conserved domain analysis (Section 7.3.3). Variations in amino acid conservation scores were detected between positions 150-250 and 400-500, with the lowest conservation score described at position 178 which falls within domain 1 of EF-G. When analysing the amino acid conservation at the point of mutagenesis (610), a conservation score of 9 was predicted with the only amino acid polymorphism being that of the evolved resistant strain of *E. coli* MG1655.

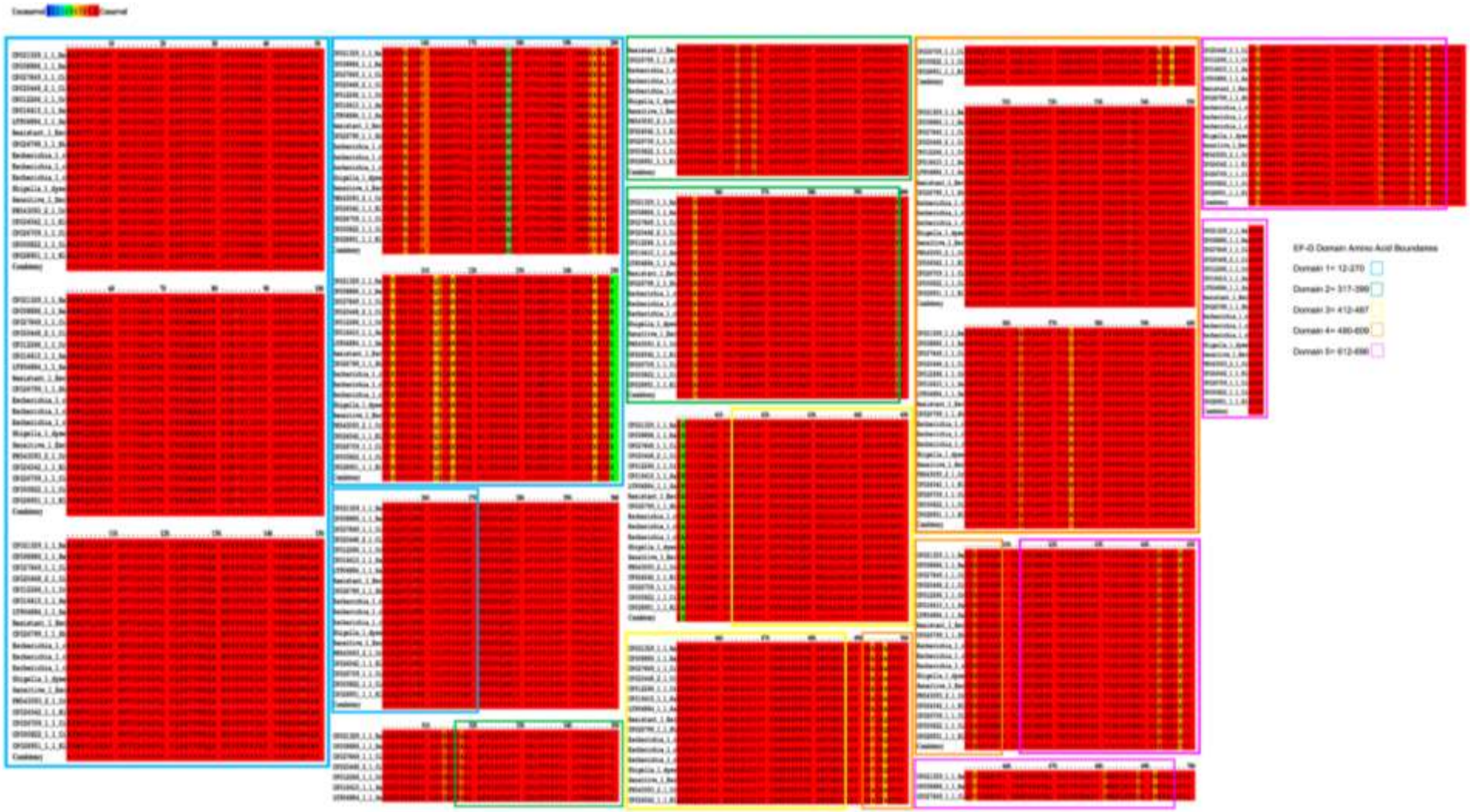


Figure 7.7 PRALINE protein conservation analysis of EF-G within the *Enterobacteriaceae* family. Displaying the multiple protein sequence alignment of EF-G within the *Enterobacteriaceae* family. The amino acid within the *fusA* protein sequences are coloured based on position conservation. The colour scheme display conservation levels per the colour key (top left), where highly conserved amino acids are shown in red and non-conserved amino acids are depicted in blue. The alignment has been sectioned into coloured blocks to highlight the amino acid boundaries of the EF-G protein domains: domain 1(blue), domain 2 (green), domain 3 (yellow), domain 4 (orange) and domain 5 (pink).

7.3.7 Bioinformatic analysis of the *fusA* mutation at PHE

The *fusA* resistant nucleotide sequence was screened at PHE against a collection of carbapenemase and ESBL producing *E. coli*. This is relevant as mutations in *fusA*, a common cause of fusidic acid resistance, have been associated with multiple phenotypic characteristics including reduced growth rates, increased sensitivity to oxidative stress and hyper-susceptibility to various classes of antibiotics including B-lactams (Macvanin *et al*., 2004 and Macvanin and Hughes, 2005). The search revealed that of the 840 carbapenemase-producing and 958 ESBL-producing *E. coli* genomes, there were no mutations identified at position 610 that could be associated with gentamicin resistance.

7.4 Discussion

The aims of this chapter described in section (7.1), were achieved as follows. The *fusA* gene was found to be highly conserved amongst the *Escherichia* taxa and *Enterobacteriaceae* family. Examination of the phylogenetic relationships of the *fusA* gene between distinct members of the *Enterobacteriaceae* revealed signatures of divergent evolution, which suggested a potential role for the *fusA* gene in bacterial phylogeny. Conclusively, as hypothesised the SNP within the *fusA* gene was unsuccessfully identified within the clinically significant bacterial pathogens.

The elongation factor EF-G, is an important bacterial protein of the translational GTPases (trGTPases), which comprise the essential translation factors characteristically conserved across all domains of life (Atkinson, 2015). The main prokaryotic elongation factors are EF-Tu and EF-G, whilst independently involved

within two distinct phases of ribosomal translation, their mutual GTPase activity and structural similarities indicate they arose from a common ancestor (Laursen and Lawrence, 1978). The highly conserved EF-Tu, acts as an evolutionary marker in bacterial phylogeny, and has provided phenotypic details conventionally missed using 16s rRNA (Kamla, Henrich and Hadding, 1996). More recently, the EF-G encoding *fusA* gene has been employed in Multilocus Sequence Typing (MLST), in the discrimination of diverse isolates within the *Cronobacter* bacterial genus (Joseph *et al.,* 2012).

Therefore, it is then unsurprising that the phylogenetic analysis of the *fusA* gene revealed monophyly within the *Escherichia* genus, owing to the highly conserved nature of the *fusA* gene and its application in the study of bacterial taxonomy. The observed overlay between the *E. coli* and *Shigella* species within the *Enterobacteriaceae* family phylogeny, further questions whether these closely related organisms should be reclassified as the same species (Chattaway *et al.,* 2017). Furthermore, the highly conserved *fusA* gene performed similarly to the 16s rRNA genes in the classification of bacterial species within the *Enterobacteriaceae* family, which suggested that the *fusA* gene evolves through evolutionary divergence not convergence (Merhej *et al.,* 2009; Wang and Huang, 2018).

In addition to the phylogenetic analysis of the *fusA* gene, functional genomics was applied to provide a link between the protein sequence conservation and the function of EF-G (Benner *et al.,* 2000). The protein conservation analysis uncovered

highly conserved regions within the EF-G protein sequence consistent across the *Enterobacteriaceae* family. This is likely to reflect the functional constraints on the divergent evolution of the elongation factor proteins, related to their essential roles during protein synthesis (Gaucher, Miyamoto and Benner, 2001). Therefore, it can be assumed that the amino acid variability detected within the non-conserved regions of EF-G, are species specific and do not lead to any significant alterations in protein structure or function (Chelvanayagam *et al.,* 1997). Granting this the resistant mutation identified, was located within a functionally conserved region of EF-G, and thus any alteration is expected to drastically alter the protein function and the phenotype of the organism (Marini, Thomas and Rine, 2010).

Despite all the evidence suggesting the mutation identified within the universally conserved elongation factor was of both functional and structural significance, the EF-G mutation was restricted to the experimentally evolved strain of *E. coli* MG1655. The given outcome does not undermine the experimental findings but instead outlines the possible limitations of adaptive laboratory evolution. Here, within the proposed model bacterial populations are subject to an artificial highly controlled environment and driven to resistance through the application of an increasing antimicrobial selective pressure. Although antimicrobial concentration gradients have been employed to reflect the heterogeneity within natural environments, it is unlikely that in nature bacterial populations would be exposed to the elevated concentrations utilised within the model (Baquero and Negri, 2005; Ghaddar, Hashemidahaj and Findlay, 2018). Likewise, the evolutionary model does not represent all the selective pressures facing

bacterial populations in nature and clinical settings including; fluctuating environmental conditions, bacterial competition and the immune system involvement (Reznick and Ghalambor, 2005; Dragosits and Mattanovich, 2013).

It is also important to acknowledge the mutational dynamics within bacterial populations, where beneficial mutations will be selected for, except when coexisting populations are present which harbour a substantial fitness advantage (Conrad, Lewis and Palsson, 2011). Consequently, it is feasible to suggest that the *fusA* mutation identified could have been eliminated from the population before reaching a detectable limit. The fitness costs associated with the acquisition of gentamicin resistance will be discussed in detail within the next chapter.

Chapter 8: The fitness costs associated with the acquisition of gentamicin resistance in *E. coli* MG1655

8.1 Introduction

The emergence of antimicrobial resistance within bacterial populations is contingent upon several factors including: mutation rates, antibiotic administration patterns and the associated fitness costs (Andersson and Levin, 1999; Björkman and Andersson, 2000; Lipsitch and Levin 1997). In general, it is commonly accepted that resistance conferring mutations impose biological costs on bacterial fitness in the absence of antimicrobial selective pressures (Levin *et al.,*1997). This is particularly relevant where mutations result in drug target alterations, in cellular components key to biological functions such as protein biosynthesis within the bacterial ribosome (Springer *et al.,*2001; Paulander, Maisnier-Patin and Andersson, 2009).

Consequently, this resulted in strategies to reduce antimicrobial therapy following resistance development, as an approach to combat antimicrobial resistance (Hall, 2004). Founded upon the principal that sensitive genotypes would outcompete the 'less fit' resistance mutants within a bacterial population, leading to the reversion to sensitivity (Lenski, 1998; Johnsen *et al.,*2009; Melnyk, Wong and Kassen, 2014). This provided incentives for an alternative approach to drug design, through the identification of antimicrobial targets which had the most detrimental effects on bacterial fitness (Andersson and Levin, 1999; Andersson, 2006).

However, while antimicrobial suppression has the potential to revert resistance, this cannot explain the persistence of resistant bacteria in the absence of antibiotics

(Andersson, 2003). There is growing evidence to suggest that bacterial adaptation successfully facilitates the maintenance of resistance and alleviation of the associated fitness costs through three key mechanisms (Andersson and Hughes, 2010). Compensatory adaptation can counteract the fitness costs associated with resistance, through mutations at secondary sites which restore bacterial fitness (Levin, Perrot and Walker, 2000; Kugelberg *et al.,*2005; Besier *et al.,*2005). In addition, resistance mutations may arise within bacterial populations which encounter no fitness costs in the absence of selective pressures (Ramadhan and Hegedus, 2005). The final mechanism is genetic co-selection, where genetic linkages are formed between resistance markers, enhancing host fitness (Enne *et al.,*2004; Andersson, 2006).

The assessment of bacterial fitness is relatively complex, as contributory elements such as in host competition, transmission and clearance have proven difficult to measure (Andersson and Levin, 1999). Subsequently, bacterial fitness has been experimentally determined through the comparative assessment of bacterial growth, virulence and competitiveness (Schrag, Perrot and Levin, 1997; Björkman, Hughes and Andersson,1998; Björkman *et al.,*1999).

This chapter aims to experimentally determine the fitness costs associated with the acquisition of gentamicin resistance, through bacterial growth and virulence assays. Further bacterial assessments will include biofilm formation, dose response and morphological analysis. It is hypothesised that there will be fitness cost associated with resistance development, which could result in reductions in both bacterial growth and virulence. Nevertheless, an assumption can be made that subsequent evolution of the resistant population could have compensated for any observed fitness costs.

8.2 Materials and methods

8.2.1 Manual growth curves

The initial assessment of bacterial fitness was conducted through a direct comparison of bacterial growth using a manual growth curve over a 12-hour period, to obtain quantifiable bacterial colony counts. Details of the manual growth curve protocol can be found in section (2.13.1).

8.2.2 Automated growth curves

In addition to the manual growth curve described section (2.13.1), an automated growth curve was conducted using the TECAN Infinite M200 Pro. Here bacterial cultures were prepared as stated in section (2.3), and standardised to an OD_{600} of 0.08 OD units within falcon tubes. Subsequently, 50µL of each bacterial culture in triplicate was added to wells of a 96-well microtiter plate containing 150µL MH broth, to achieve a starting OD_{600} of 0.002 OD units. Negative controls containing 200µL of MH broth were included. Plates were sealed and placed in the plate TECAN plate reader under the conditions specified in table (8.1).

Table 8.1 TECAN Infinite M200 Pro plate reader conditions. Displaying the plate reader conditions utilised in the automated analysis of bacterial growth.

Temperature	37ºC
Absorbance	600nm
Shaking	5 seconds (prior to OD_{600})
Kinetic cycles	12
Duration of each cycle	30 minutes

8.2.3 Dose response analysis

In addition to the analysis of fitness by assessing the growth of sensitive and resistant strains of *E. coli,* grown within identical conditions in MH broth. It was of interest to see how bacterial growth deviates in the presence of the antimicrobial agents at two given does; the MIC and x10 the MIC. The dose response analysis was conducted by manual and automated methods as stated above in sections (8.2.1 and 8.2.2). The only modification made here was the addition of antimicrobial agents at the selected concentrations to the MH broth prior to inoculation. Positive controls were included for the comparative analysis of growth in the absence of the antimicrobial selective pressures.

8.2.4 *Galleria mellonella* virulence model

Furthermore, the bacterial fitness of the sensitive and resistant strains of *E. coli* MG1655 were analysed by assessing bacterial virulence using the *G. mellonella* toxicity model. The details of the protocol used in *Galleria* toxicity model are provided in section (2.13.2).

8.2.5 Scanning electron microscopy

The final bacterial comparison of the sensitive and resistant strains of *E. coli*, was a morphological assessment using SEM. Details of the SEM protocol are described previously in section (2.13.3).

8.3 Experimental results

Several assays were conducted to determine the costs associated with the acquisition of gentamicin resistance, through a comparative assessment of fitness of sensitive and resistant strains of *E. coli* MG1655. For convenience, the sensitive strain of *E. coli* MG1655 was referred to as the WT, and the resistant strains were denoted MIC and x10 MIC, reflecting the selected antimicrobial concentration. The statistical analysis and graphical visualisation of data was completed using the software programme RStudio.

8.3.1 Automated bacterial growth analysis

Following the confirmation of phenotypic resistance and the identification of mutations related specifically to the aminoglycoside mechanism of action, it was of interest to determine any alterations in the growth of resistant strains when placed in a medium in the absence of the antimicrobial selective pressure. Considering the colony variation amongst bacterial populations, growth analysis was conducted on five individual colonies from the WT, MIC and x10 MIC strains in triplicate within a 96-well plate format. To assist in the graphical interpretation of results, the triplicates of the five individual colonies from each strain were presented in the same colour.

Prior to the comparative analysis of the sensitive and resistant strains, the growth of each strain will be discussed independently. The growth analysis of *E. coli* MG1665 WT, revealed similar growth trajectories for all 15 cultures over a period of 12 hours. The greatest fluctuations between colonies appeared between 3-12 hours, with replicate two of colony five attaining an OD_{600} value of 0.2166 OD units after 12 hours

of growth, compared to the remaining WT cultures achieving OD_{600} values closer to 0.3 OD units. Looking more broadly at the individual colonies, all colony triplicates followed similar growth trajectories except for C5 where disparities were observed (Figure 8.1). However, there was no significant difference in the growth of five colonies from the WT strain of *E. coli* MG1655 after 12 hours of growth (P=0.4279). This indicated reduced inter-colony variation between the WT population.

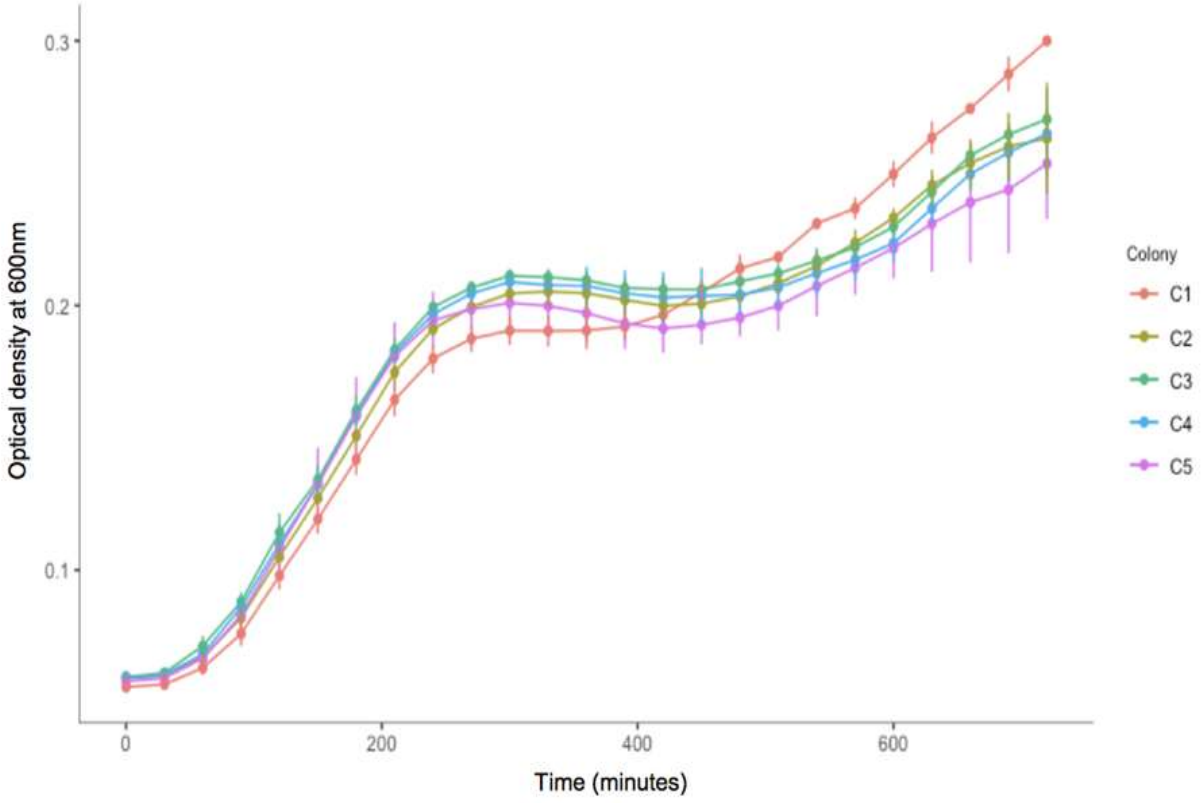


Figure 8.1 Automated growth analysis of *E. coli* MG1655 WT strain. Graph showing the growth curve of the *E. coli* MG1655 WT strain in MH broth at 37ºC. All five colonies (C1-C5) were grown in triplicate in 96-well plates and the OD_{600} values were measured every 30minutes for 12 hours in the Infinite M200 PRO plate reader. The colony triplicates were presented in colour: C1 (red), C2 (Khaki green), C3 (green), C4 (blue) and C5 (magenta). The triplicate values per colony were then averaged and the standard error was calculated. The error bars represent standard deviation. There was no significant difference in the growth rates between all replicates of the *E. coli* MG1655 WT strain (P=0.4279).

For the *E. coli* MG1655 resistant strain selected at the MIC of gentamicin. Again, similar phenotypic trajectories were observed after 12 hours of growth (Figure 8.2). However, variations were identified relatively early in the growth of replicate two of colony two and for all three replicates of colony four. Further deviations in growth were observed at the end of the growth curve for replicates one and two of colony five. In comparison to the growth of the WT population, the gradient of the initial phase of growth over the first 300 minutes, was less steep for the MIC cultures. In addition, the maximum growth for the MIC cultures did not exceed 0.25 OD units. There was no significant difference in the growth of the five individual colonies of the MIC strain of *E. coli* MG1655 after 12 hours of growth (P= 0.4456). However, although statistically insignificant, in comparison to the WT strain there was greater inter-colony variation observed with the MIC cultures.

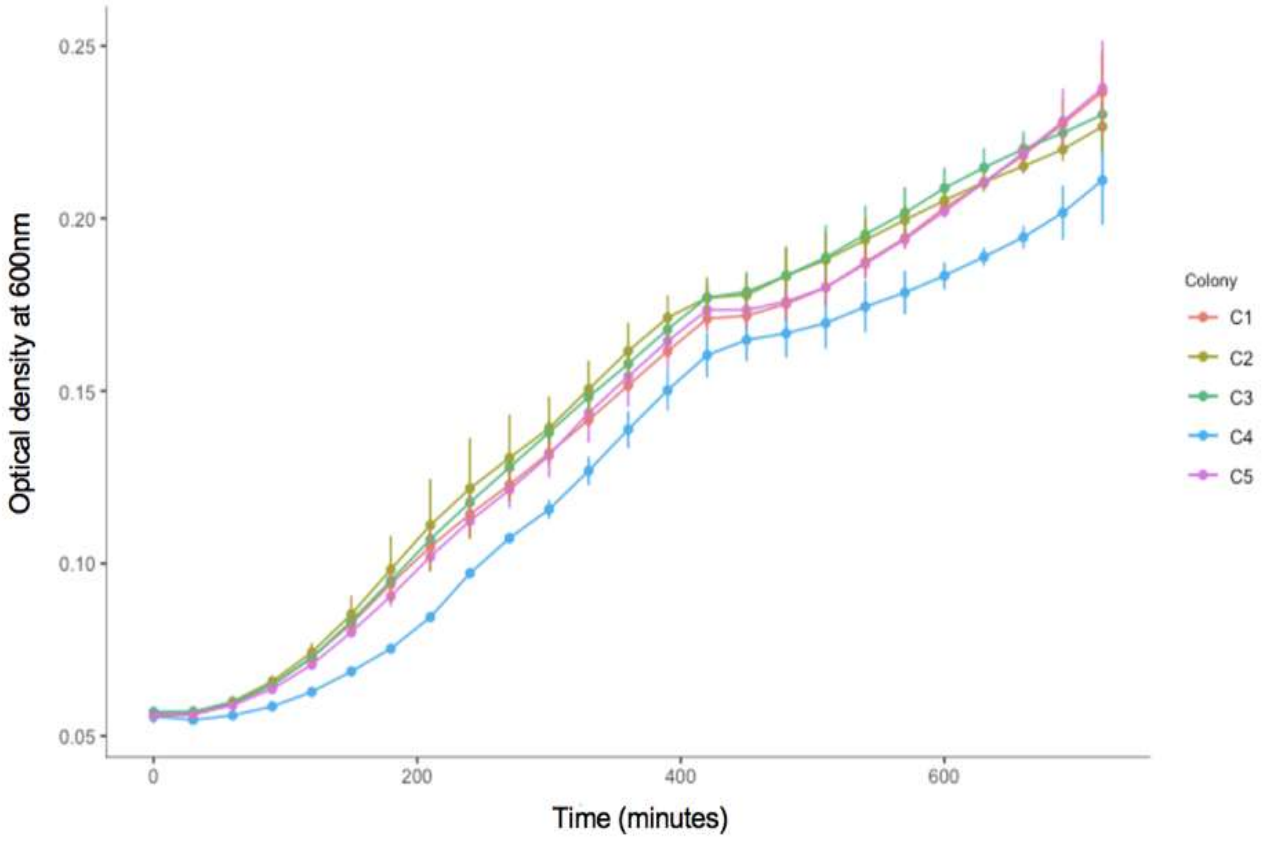


Figure 8.2 Automated growth analysis of *E. coli* MG1655 MIC strain. Graph showing the growth curve of the *E. coli* MG1655 MIC strain in MH broth at 37ºC. All five colonies (C1-C5) were grown in triplicate in 96-well plates and the OD_{600} values were measured every 30 minutes for 12 hours in the Infinite M200 PRO plate reader. The colony triplicates were presented in colour: C1 (red), C2 (Khaki green), C3 (green), C4 (blue) and C5 (magenta). The triplicate values per colony were averaged and the standard error was calculated. The error bars represent standard deviation. There was no significant difference in the growth rates between all replicates of the *E. coli* MG1655 MIC strain (P=0.4456).

For the *E. coli* MG1655 resistant strain selected at x10 the MIC of gentamicin. The overall trend of growth was similar for all 15 resistant cultures (Figure 8.3). Direct comparisons could be drawn between the growth of the two resistant strains, as fluctuations appeared in a subset of colonies or replicates which diverged from the overall trend of growth. This was observed within the x10 MIC resistant strain for all replicates of colony five between 90-480 minutes of growth and between individual replicates of colonies three and four after 60-420 minutes of growth.

However, dissimilar to the MIC resistant strain, the maximum OD_{600} value achieved for the x10 MIC cultures was less than that of the WT and MIC strains, with the final OD_{600} values ranging from 0.15-0.2 OD units after 12 hours of growth. There was no significant difference in the growth of the x10 MIC strain of *E. coli* MG1655 after 12 hours of growth ($P=0.4781$). Again, although statistically insignificant, in comparison to the WT and MIC strains there was greater inter-colony variation observed with the x10 MIC cultures. These results proposed that as the level of resistance increases, there is greater variation amongst bacterial colonies.

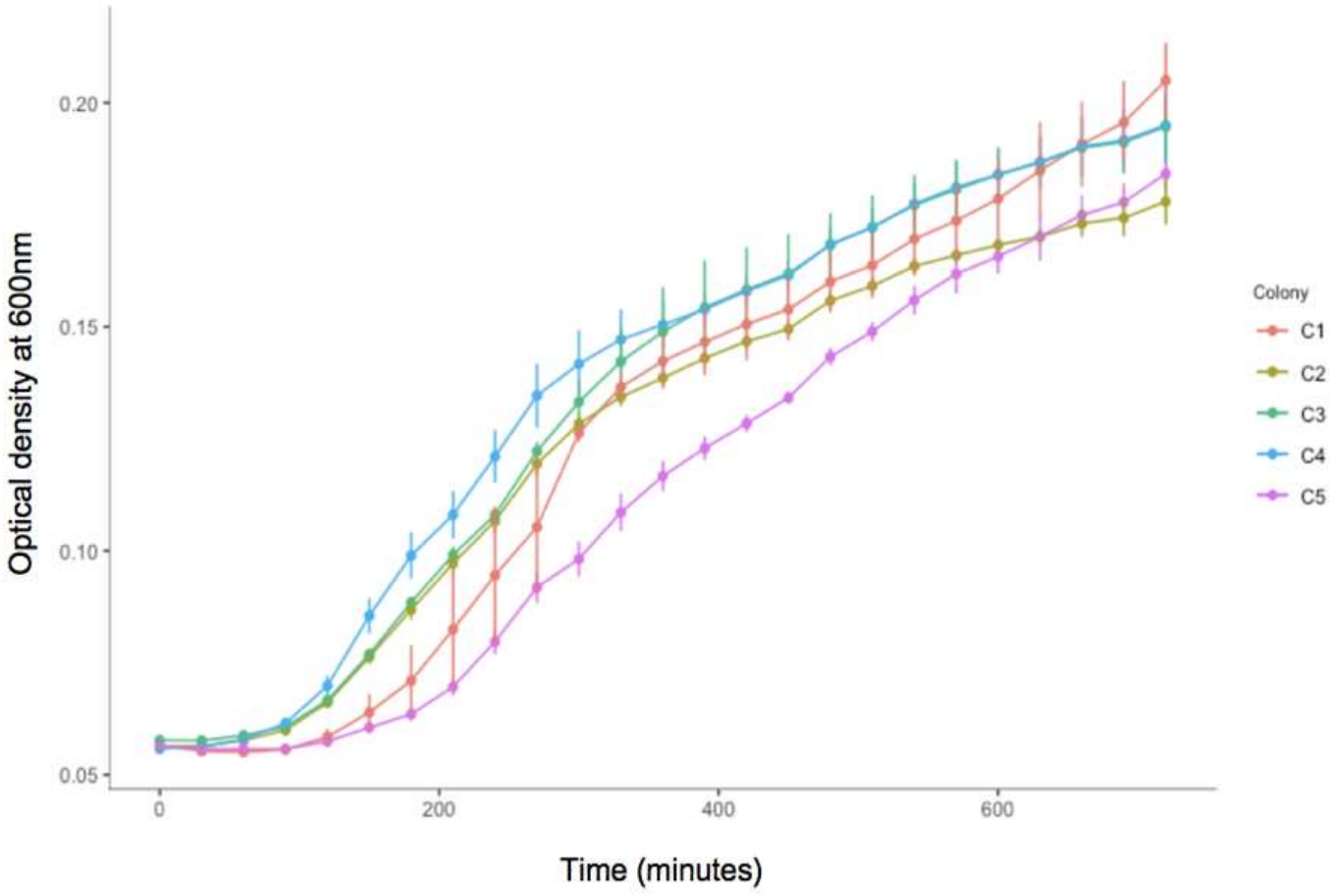


Figure 8.3 Automated growth analysis of *E. coli* MG1655 x10 MIC strain. Graph showing the growth curve of the *E. coli* MG1655 x10 MIC strain in MH broth at 37ºC. All five colonies (C1-C5) were grown in triplicate in 96-well plates and the OD_{600} values were measured every 30 minutes for 12 hours in the Infinite M200 PRO plate reader. The colony triplicates were presented in colour: C1 (red), C2 (Khaki green), C3 (green), C4 (blue) and C5 (magenta). The triplicate values per colony were averaged and the standard error was calculated. The error bars represent standard deviation. There was no significant difference in the growth rates between all replicates of the *E. coli* MG1655 MIC strain (P=0.4781).

Successively, the OD_{600} values of the sensitive and resistant strains were averaged to comparatively asses the growth of the WT, MIC and x10MIC strains over 12 hours (Figure 8.4). Looking more briefly at the overall trend of growth between the strains, it was evident that the growth rate was higher for the WT in comparison to the two resistant strains, which shared more closely related growth profiles. The greatest disparities in growth amid the sensitive and resistant strains were observed over the initial 400 minutes. After which, the growth of the resistant strains started to diverge, with the MIC strain increasing more rapidly in growth, than its highly resistant counterpart. These preliminary growth results, suggest compromised growth following the acquisition of resistance.

There were significant differences in growth observed between the sensitive and resistant strains over the 12-hour growth curve (P=1.1e-09). The greatest significant difference in growth was observed when comparing the growth of the
WT with the two resistant strains independently. The bacterial growth of the WT strain differed most significantly to the x10MIC strain (P=3.094e-1) followed by the MIC strain (P=1.609e-05). Unsurprisingly, there was no significant difference between the growth of the two resistant strains (P=0.06019).

The shaded ribbons surrounding the lines on the graph (Figure 8.4), represent the standard error of the triplicate OD_{600} values from each strain. The error values were relatively small for all three strains, yet the degree of error in the x10MIC strain appeared larger than that of the WT and MIC.

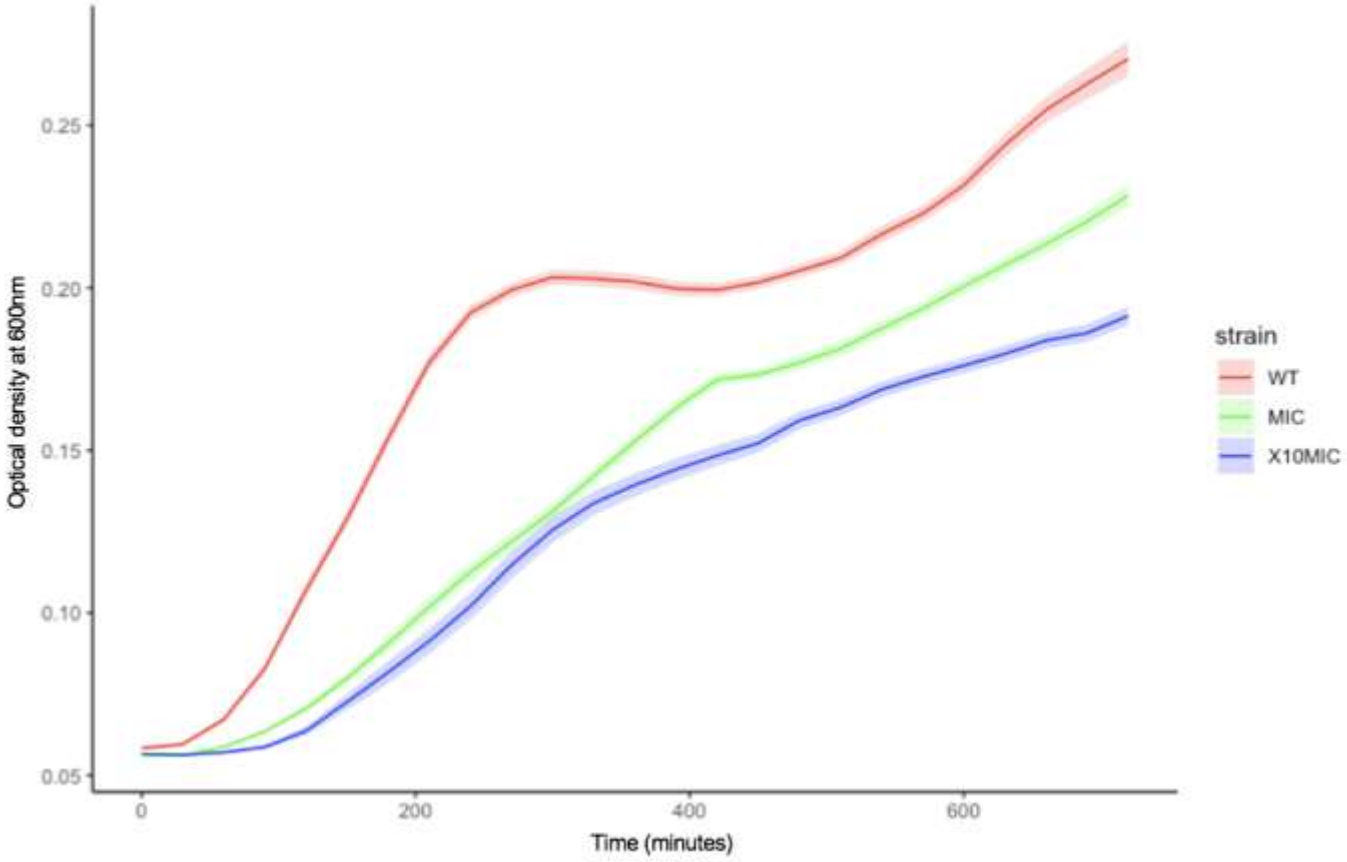


Figure 8.4 Automated growth analysis of gentamicin sensitive and resistant strains of *E. coli* MG1655. Graph showing the growth curve of the *E. coli* MG1655 WT, MIC and x10 MIC strains in MH broth at 37ºC. Five colonies (C1-C5) per strain were grown in triplicate in 96-well plates and the OD_{600} values were measured every 30 minutes for 12 hours in the Infinite M200 PRO plate reader. The colony triplicates per strain were averaged and the standard error was calculated. The error ribbons represent the standard error of the mean. There was a significant difference in the growth rates of the sensitive and resistant strains of E. coli MG1655 (P= 1.1e-09).

Prior to the statistical analysis, the data was assessed to determine the correct statistical approach. The normality tests for the statistical analysis revealed the automated growth curve data did not follow normal distribution, illustrated via histograms and the QQ plots. The histogram graphics suggested highly variable distribution between with the WT and resistant strains (Figure 8.5 A). This was validated using QQ plots, where the quantiles for each strain did not form a straight line (Figure 8.5 B). As normality was desired, the data was transformed using various functions in RStudio (Figure 8.5 C), however this was unsuccessful. Therefore, a non-parametric approach to an annova was conducted in the form of a Kruskal–Wallis test, to determine significance.

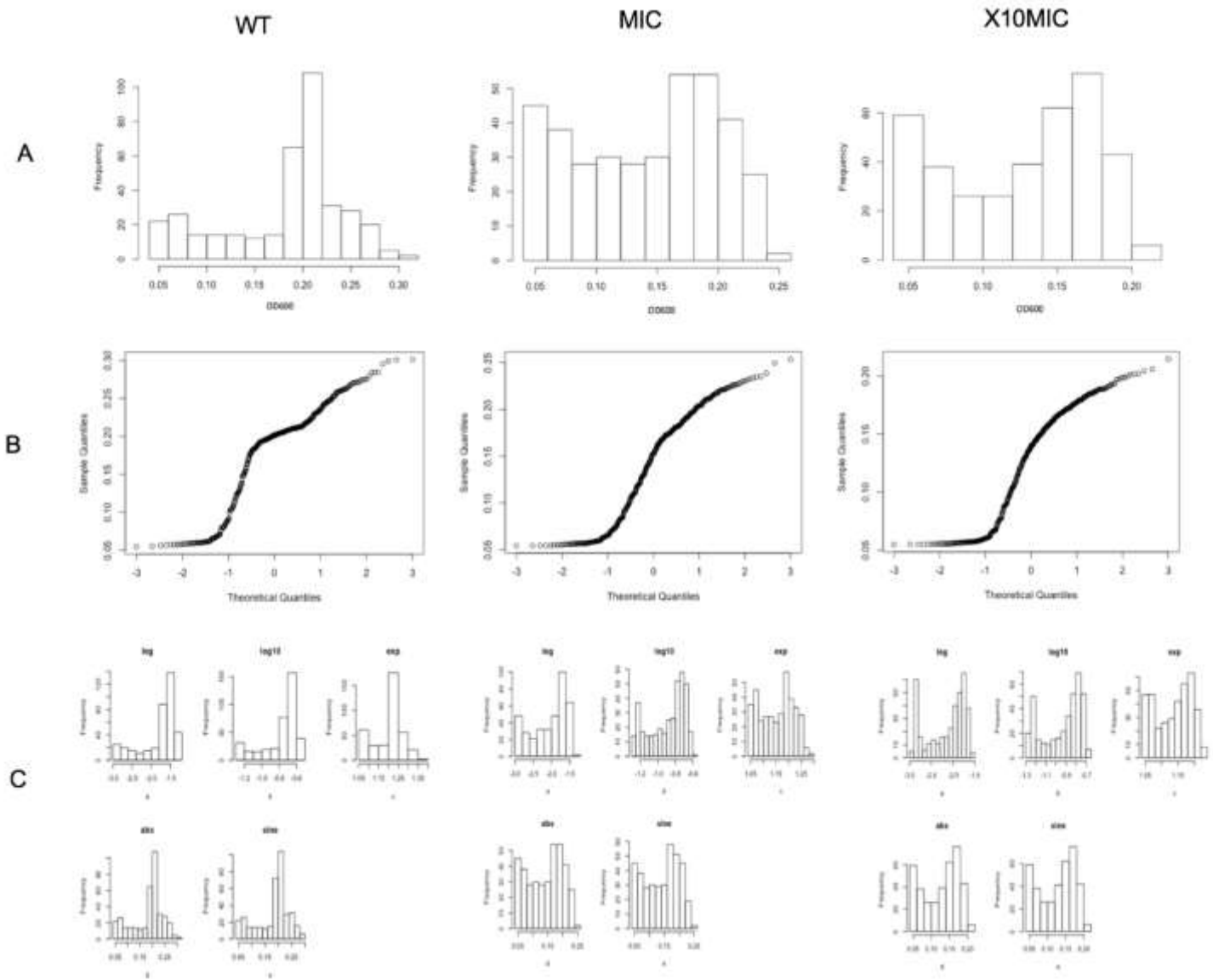


Figure 8.5 Normality tests for the statistical analysis of the automated bacterial growth curves. (A) Histograms of WT, MIC and x10MIC, displaying the frequency of optical density values for the automated growth curve data. (B) QQ plots of WT, MIC and x10MIC, displaying the theoretical quantiles of the optical density values for the automated growth curve data. (C) Transformation plots (log, log10, exp, abs and sine) of the *E. coli* MG1655 WT MIC and x10 MIC automated growth curve data.

8.3.1.1 Automated exponential phase growth analysis

Following the growth curve analysis of the sensitive and resistant strains over a duration of 12 hours, it was of interest to further analyse the bacterial growth in relation to the specific phases of growth. The exponential phase of growth was selected, as it is the phase in which bacterial cells are most rapidly dividing.

The exponential growth analysis of the WT strain, revealed highly similar phenotypic trajectories for all five colonies (Figure 8.6). The exponential phase of the WT strain, was initiated following 60 minutes of growth in MH broth, which followed a rapid increase in OD_{600} values to between 0.188 and 0.208 OD units after 270 minutes for all selected colonies.

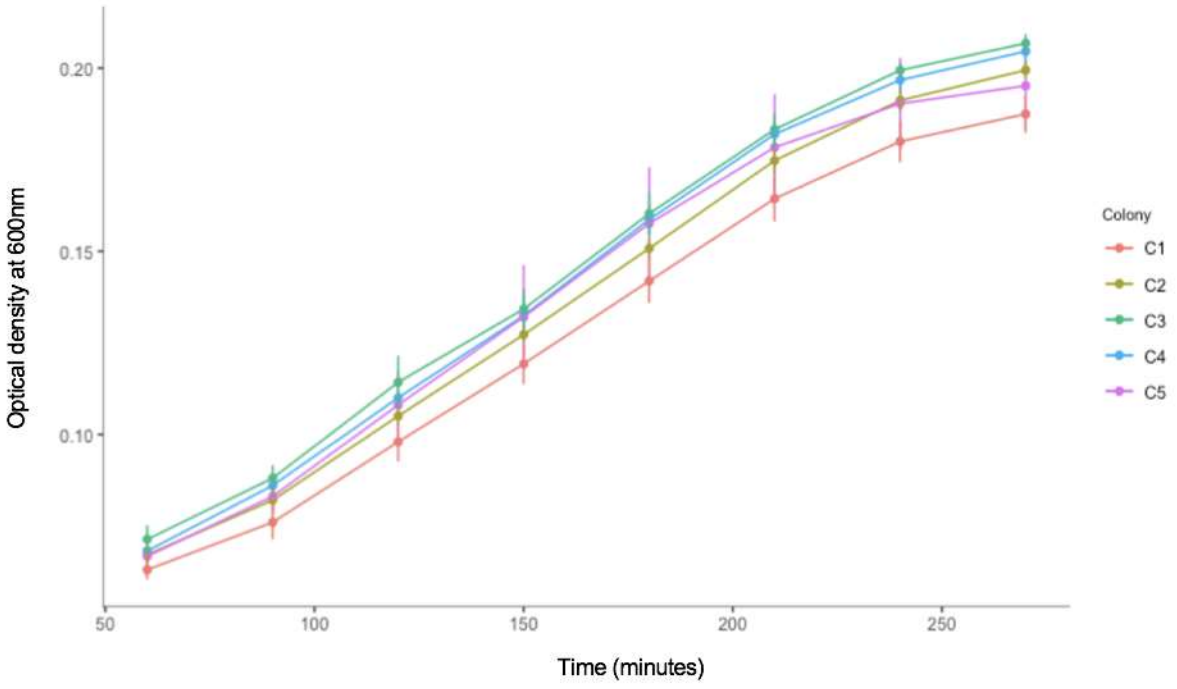


Figure 8.6 Automated exponential growth analysis of *E. coli* MG1655 WT strain. Graphs showing the exponential growth curve of the *E. coli* MG1655 WT strain in MH broth at 37ºC. All five colonies (C1-C5) were grown in triplicate in 96-well plates and the OD_{600} values were measured in the Infinite M200 PRO plate reader. The triplicate values per colony were averaged and the standard error was calculated. The error bars represent standard deviation. There was no significant difference in the exponential growth rates between WT colonies of *E. coli* MG1655.

In comparison to the WT strain, the exponential phase of growth of the MIC strain, followed very similar trajectories in all but one colony (Figure 8.7). The slope of the is less steep than that seen with the WT cultures, and this is further demonstrated by the significantly lower OD_{600} values attained after 270 minutes of growth, ranging from 0.106 to 0.145 OD units for all five colonies. The substantial difference in exponential growth identified within colony C4, inferred greater colony variation of resistant cultures as described previously.

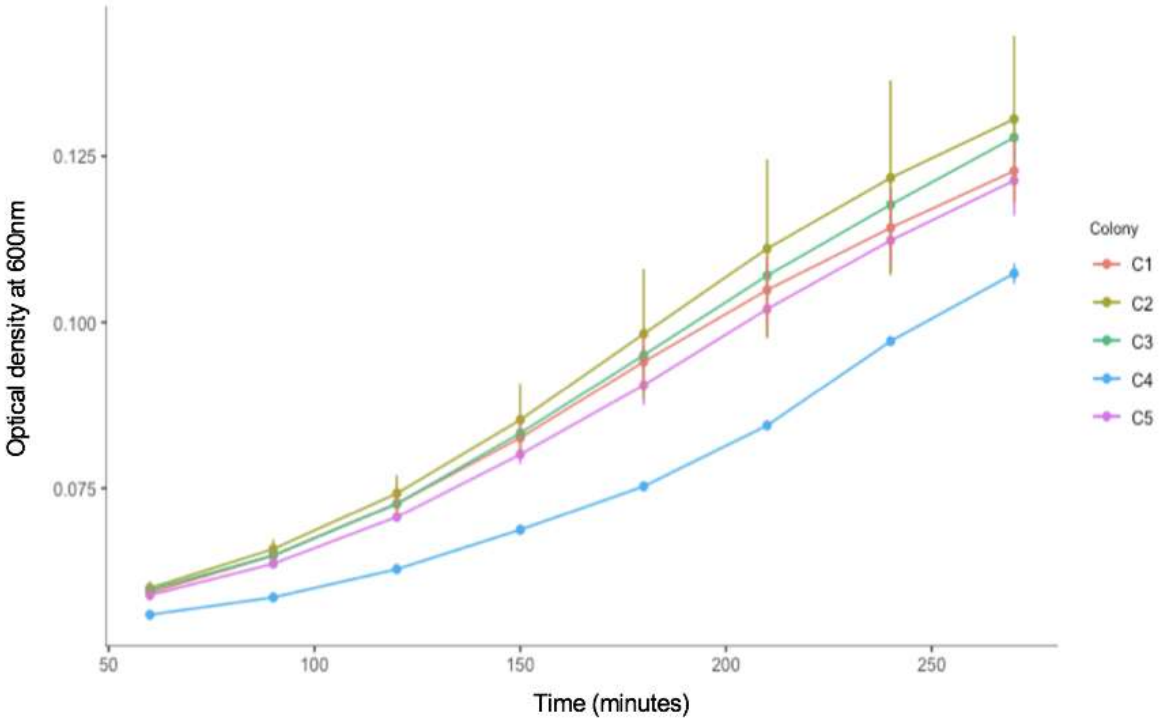


Figure 8.7 Automated exponential growth analysis of *E. coli* MG1655 MIC strain. Graphs showing the exponential growth curve of the *E. coli* MG1655 MIC strain in MH broth at 37ºC. All five colonies (C1-C5) were grown in triplicate in 96-well plates and the OD_{600} values were measured in the Infinite M200 PRO plate reader. The triplicate values per colony were averaged and the standard error was calculated. The error bars represent standard deviation.

The fluctuations observed between colonies was further demonstrated when looking at the exponential growth of the x10MIC resistant strain (Figure 8.8). Similarities can be drawn when looking at the overall trend in colony growth however, the inter-colony variation was more pronounced, excluding colonies C2 and C3. The exponential growth analysis further supported the findings of an extended lag phase within the X10 MIC resistant cultures, where the most pronounced increase in OD_{600} was observed following 120 minutes. Equivalent, OD_{600} values were achieved after 270 minutes of growth for the two resistant strains, which ranged between 0.081 to 0.142 OD units.

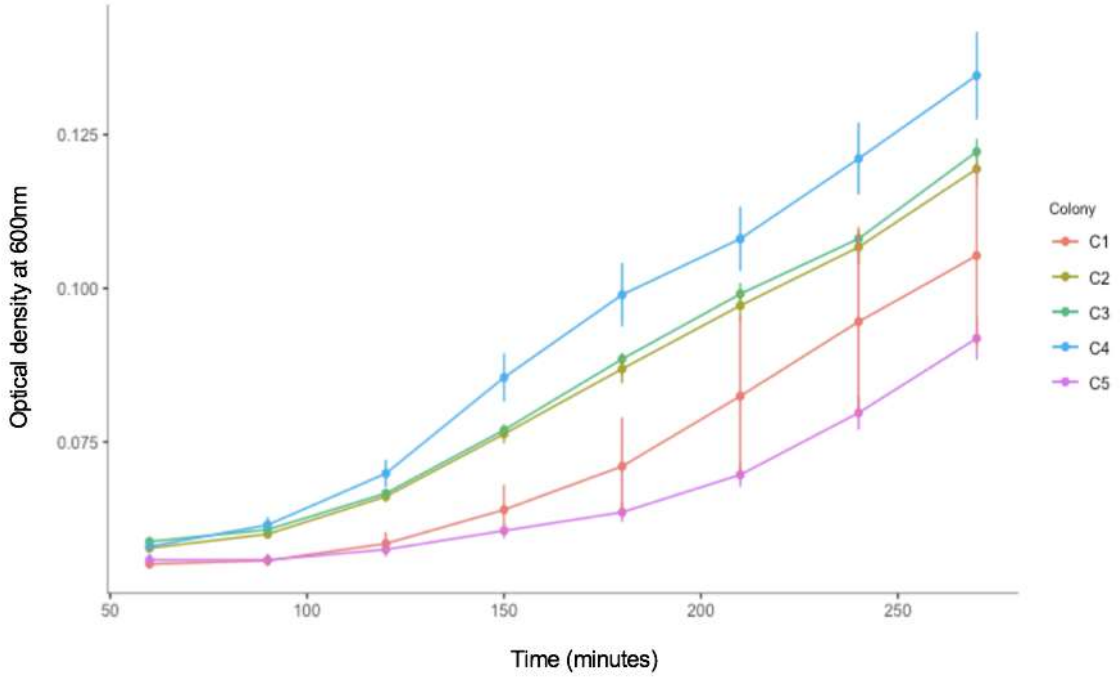


Figure 8.8 Automated exponential growth analysis of *E. coli* MG1655 x10 MIC strain. Graphs showing the exponential growth curve of the *E. coli* MG1655 x10 MIC strain in MH broth at 37ºC. All five colonies (C1-C5) were grown in triplicate in 96-well plates and the OD_{600} values were measured in the Infinite M200 PRO plate reader. The triplicate values per colony were averaged and the standard error was calculated. The error bars represent standard deviation.

In addition to independently assessing the growth of the five individual colonies per strain, the OD_{600} values were averaged to directly compare the exponential growth of WT with the MIC and x10MIC resistant strains (Figure 8.9). There were pronounced differences in the exponential growth of the WT strain, compared to the growth between 60-270 minutes for the resistant strains. This was most evident at 270 minutes where the OD_{600} achieved for WT strain was 0.199 OD units compared to OD_{600} value of 0.122 and 0.115 OD units for the MIC and x10 MIC resistant strains respectively. As described previously (Section 8.3.1), the two resistant strains displayed similar profiles over the defined growth phase. Nevertheless, the MIC growth was marginally faster than the x10MIC strain, suggesting an indirect relationship between the level of resistance and the proficiency of bacterial growth.

The statistical analysis revealed there was again a significant difference in the bacterial growth of all three strains over the exponential phase of growth ($P=1.667e-05$). Comparably, the greatest difference in exponential growth was observed between the WT and x10 MIC strain ($P= 2.368e-05$) and there was no significant difference between the exponential growth of the two gentamicin resistant strains ($P=0.2655$).

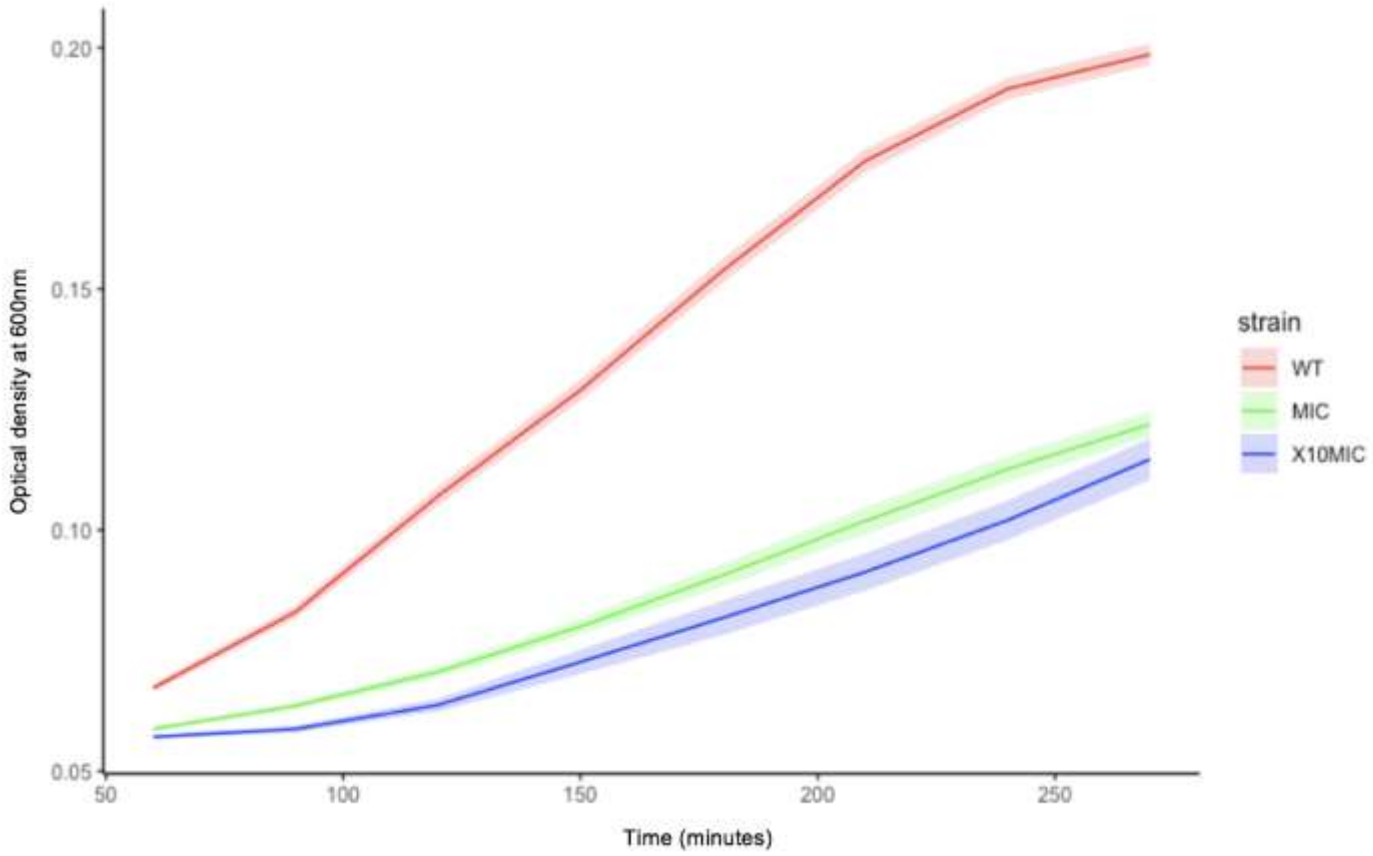


Figure 8.9 Automated exponential growth analysis of gentamicin sensitive and resistant strains of *E. coli* MG1655. Graph showing the exponential growth curve of the *E. coli* MG1655 WT, MIC and x10 MIC strains in MH broth at 37ºC. Five colonies (C1-C5) per strain were grown in triplicate in 96-well plates and the OD_{600} values were measured using the Infinite M200 PRO plate reader. The colony triplicates per strain were averaged and the standard error was calculated. The error ribbons represent the standard error of the mean. There was a significant difference in the growth rates of the sensitive and resistant strains of E. coli MG1655 (P= 1.677e-05).

8.3.2 Manual assessment of bacterial growth

To determine the reproducibility of the automated growth curve findings, the experiment was repeated manually with the inclusions of 5 individual colonies from each strain, as opposed to the 15 cultures per strain tested in the automated 96-well plate format. The growth of strains will be discussed independently, prior to the comparative assessments of growth between strains. As seen previously, the five colonies of *E. coli* MG1665 WT revealed similar phenotypic trajectories of growth (Figure 8.10). However, there appeared to be an extended lag period over the first 2.5 hours of growth, which was not evident in the automated method. Another clear discrepancy between the growth of the WT cultures, was the significantly higher maximal OD_{600} achieved following 12 hours of growth within the manual method. There was no significant difference in the growth of the WT strain of *E. coli* MG1655 after 12 hours of growth (P=0.3068).

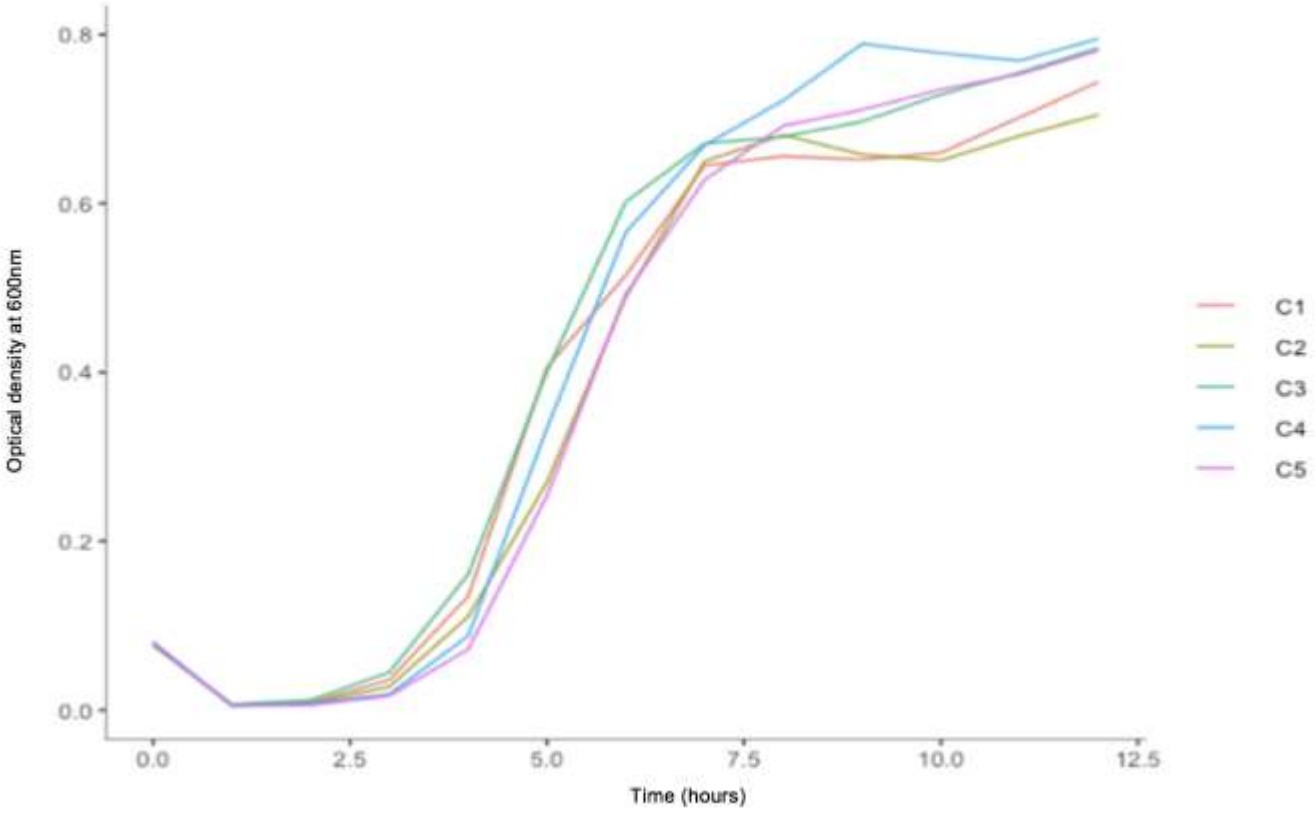


Figure 8.10 Manual growth analysis of *E. coli* MG1655 WT strain. Displaying the growth curve of the *E. coli* MG1655 WT strain in MH broth at 37ºC. All five colonies (C1-C5) were grown within conical flasks and the OD_{600} values were measured every hour for 12 hours manually using a spectrophotometer. There was no significant difference in the growth rates between all colonies of the *E. coli* MG1655 WT strain (P=0.3068).

When comparing this to the resistant strain selected at the MIC of gentamicin, similar phenotypic trajectories were observed for all MIC cultures, except for C2 which displayed a rapid increase in OD_{600} ensuing three hours of growth (Figure 8.11). Additionally, a primary lag phase of growth over the first 2.5 hours was detected, which parallels that seen with the WT cultures. A clear distinction in the growth of the MIC compared to the WT cultures, is reflected in the maximum OD_{600} achieved after 12 hours of growth, which was 0.448 and 0.795 OD units respectively. Overall, there was no significant difference in the growth of the five MIC cultures (0.2495).

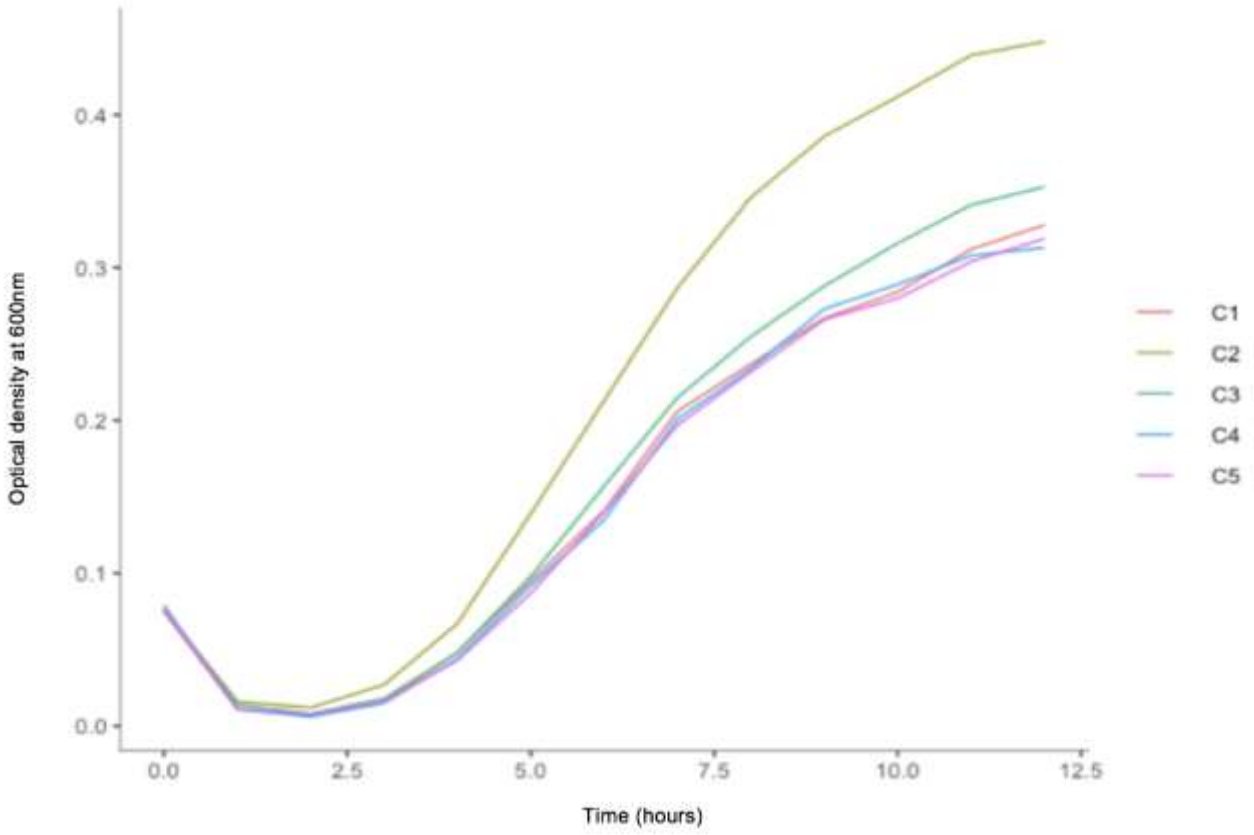


Figure 8.11 Manual growth analysis of *E. coli* MG1655 MIC strain. Displaying the growth curve of the *E. coli* MG1655 WT strain in MH broth at 37ºC. All five colonies (C1-C5) were grown within conical flasks and the OD_{600} values were measured every hour for 12 hours manually using a spectrophotometer. There was no significant difference in the growth rates between all colonies of the *E. coli* MG1655 MIC strain (P=0.2495).

Similar trends in growth were observed for the *E. coli* MG1655 resistant strain selected at x10 the MIC of gentamicin (Figure 8.12). The growth of all five cultures was highly comparable, with no major colony discrepancies as seen above with the MIC strain (Figure 8.11). This resemblance was further justified following statistical analysis, where there was found to be no significant difference in the growth of the x10 MIC resistant strain of *E. coli* MG1655 after 12 hours of growth (P=0.0782). There appeared to be a trend emerging, whereby the maximum OD_{600} achieved following 12 hours of growth, is lowest for the more highly resistant strain of *E. coli* MG1655.

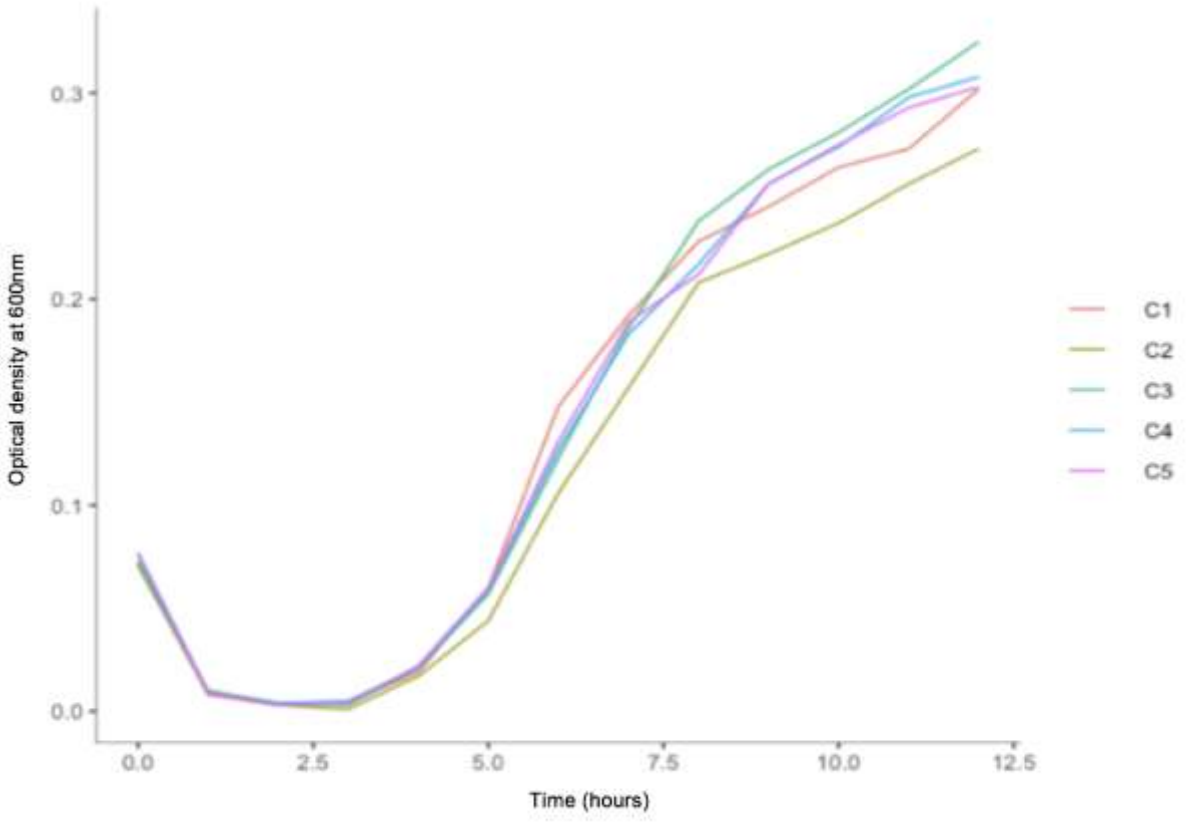


Figure 8.12 Manual growth analysis of *E. coli* MG1655 x10MIC strain. Displaying the growth curve of the *E. coli* MG1655 x10MIC strain in MH broth at 37ºC. All five colonies (C1-C5) were grown within conical flasks and the OD_{600} values were measured every hour for 12 hours manually using a spectrophotometer. There was no significant difference in the growth rates between all colonies of the *E. coli* MG1655 X10MIC strain (P=0.0782).

To assess the growth of the sensitive and resistant strains collectively, the OD_{600} values were averaged per time point (Figure 8.13). When comparing the manual and automated growth curves the general trends of the graphs are analogous, however variations were identifiable in terms of the OD_{600} values achieved, and the steepness of the gradient obtained within the manual method for the WT strain. There was a significant difference in the bacterial growth between the sensitive and resistant strains succeeding the lag phase (P=5.044e-06). However, there was no significant

difference in the growth between the two resistant strains (P=0.07526). Again, as seen with the automated bacterial analysis, the greatest difference in growth was observed between the WT and x10MIC resistant strain of *E. coli* MG1655 (P=6.849e-06).

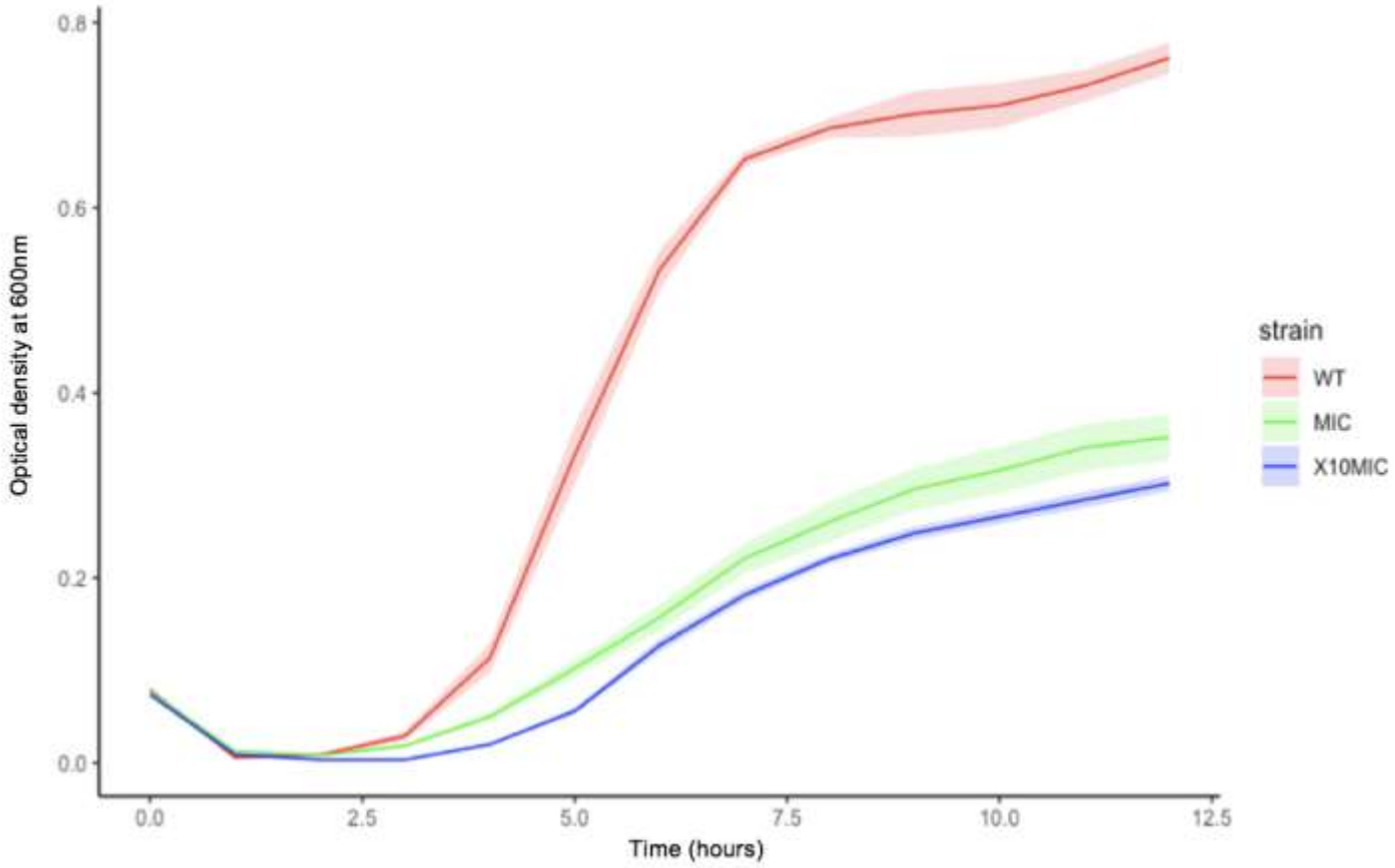


Figure 8.13 Manual growth analysis of gentamicin sensitive and resistant strains of *E. coli* MG1655. Graph showing the growth curve of the *E. coli* MG1655 WT, MIC and x10 MIC strains in MH broth at 37ºC. All five colonies (C1-C5) per strain were grown within conical flasks and the OD_{600} values were measured every hour for 12 hours manually using a spectrophotometer. The OD_{600} values obtained from each colony per strain were averaged and the standard error was calculated. The error ribbons represent the standard error of the mean. There was a significant difference in the growth rates of the sensitive and resistant strains of *E. coli* MG1655 (P= 5.044e-06).

As above, prior to the statistical analysis, it was necessary to determine the correct statistical approach based on the manual growth curve data set. Comparable to the

previous data set, the normality tests established that the data did not follow normal distribution, even following transformation. The graphical normality and transformation plots can be found in the appendix (Figure A.24). Consequently, the non-parametric *Kruskal–Wallis test was applied.*

8.3.2.1 Manual exponential phase growth analysis

As with the automated growth curve, the exponential phase of growth was examined more thoroughly. The duration of the exponential phase was determined for the WT strain before a comparative analysis of growth within the selected period of the two resistant strains. The exponential growth of the WT strain was initiated after 3 hours of growth and progressed for a further five hours, during which the OD_{600} values increased progressively from a minimum of 0.017 OD units to a maximum of 0.722 OD units. The independent exponential growth analysis plots for the *E. coli* MG1655 WT, MIC and x10 MIC strains can be found in the appendix (Figure A.25-A.27).

To directly compare the exponential growth of the sensitive and resistant strains collectively, the OD_{600} values for each colony were averaged (Figure 8.14). There was a significant difference in the exponential growth of the sensitive and resistant strains of *E. coli* MG1655 (P=2.852e-05). The greatest difference in exponential growth of the three strains was seen after 8 hours, with an OD_{600} value of 0.686 OD units achieved for the WT compared to 0.261 and 0.221 OD units for the MIC and x10MIC resistant strains respectively. Correspondingly, the greatest difference in exponential growth was observed between the WT and x10MIC strain (P=3.705e-05) and there was no significant difference amongst the two resistant strains (P=0.1353).

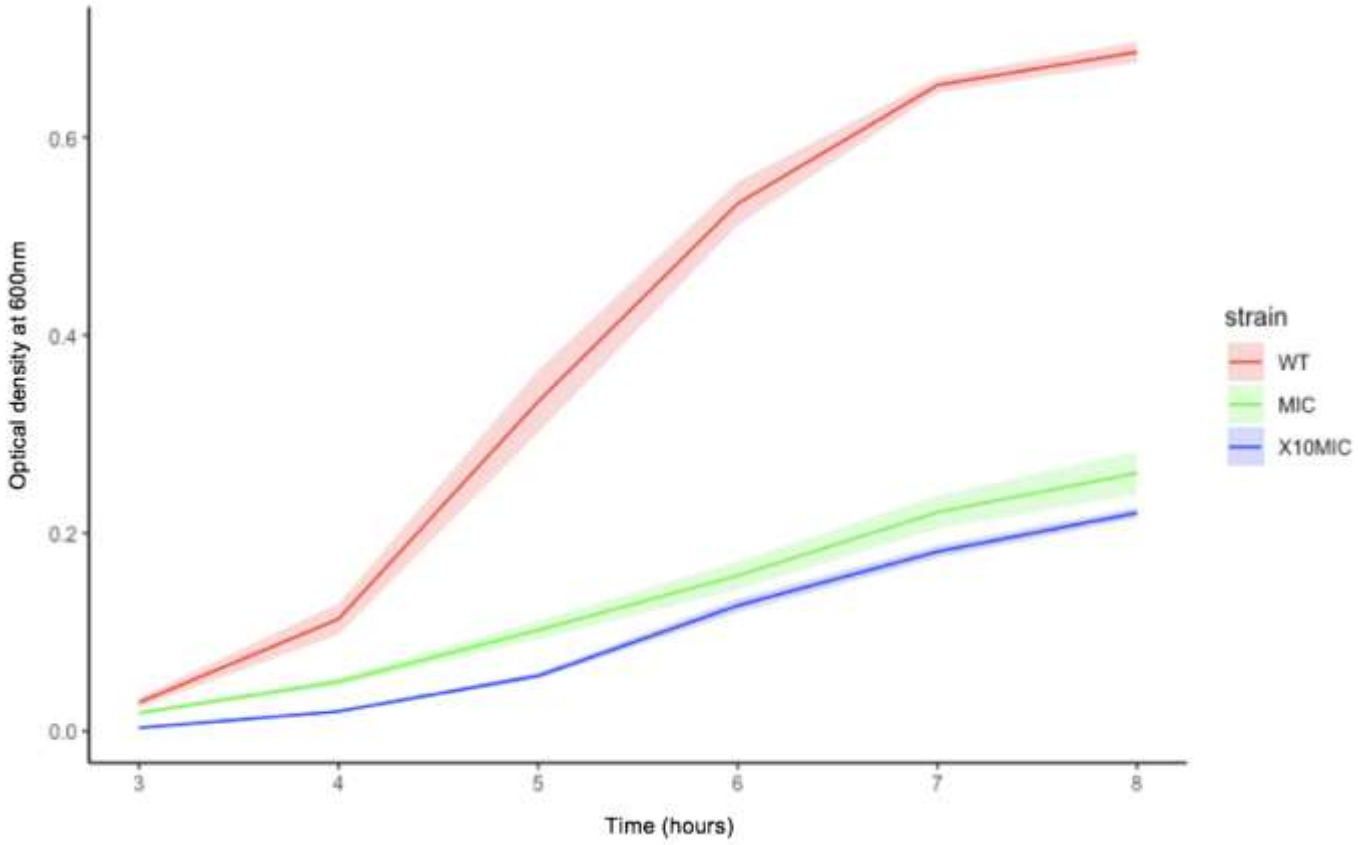


Figure 8.14 Manual exponential growth analysis of gentamicin sensitive and resistant strains of *E. coli* MG1655. Graph showing the exponential growth curve of the *E. coli* MG1655 WT, MIC and x10 MIC strains in MH broth at 37ºC. All five colonies (C1-C5) per strain were grown within conical flasks and the OD_{600} values were measured manually using a spectrophotometer. The OD_{600} values obtained from each colony per strain were averaged and the standard error was calculated. The error ribbons represent the standard error of the mean. There was a significant difference in the growth rates of the sensitive and resistant strains of *E. coli* MG1655 (P=2.852e-05).

8.3.3 Manual bacterial growth analysis bacterial enumeration

To compare the OD_{600} values obtained from both the automated and manual turbidity methods, viable cell counts of the sensitive and resistant strains were performed and the logarithm of the number of CFU/mL was calculated. Again, the viable cell counts of the sensitive and resistant strains will be discussed independently before a direct comparative analysis.

Following the enumeration of bacterial counts, the WT growth curve did not follow a smooth trajectory as seen with the two turbidity methods, however the growth of the five independent WT cultures proceeded in a comparable trend (Figure 8.15). This was supported through the statistical analysis of the WT cultures, where there was no significant difference observed in the log10 CFU/mL counts of the WT strain over the duration of the bacterial growth curve (P= 0.2573). The *E. coli* WT strain cultures appeared to reach the stationary phase of growth after approximately 6 hours of growth in MH broth. The maximum bacterial count obtained within the growth analysis was 8.27875 log10 CFU/mL which was enumerated following 690 minutes of growth.

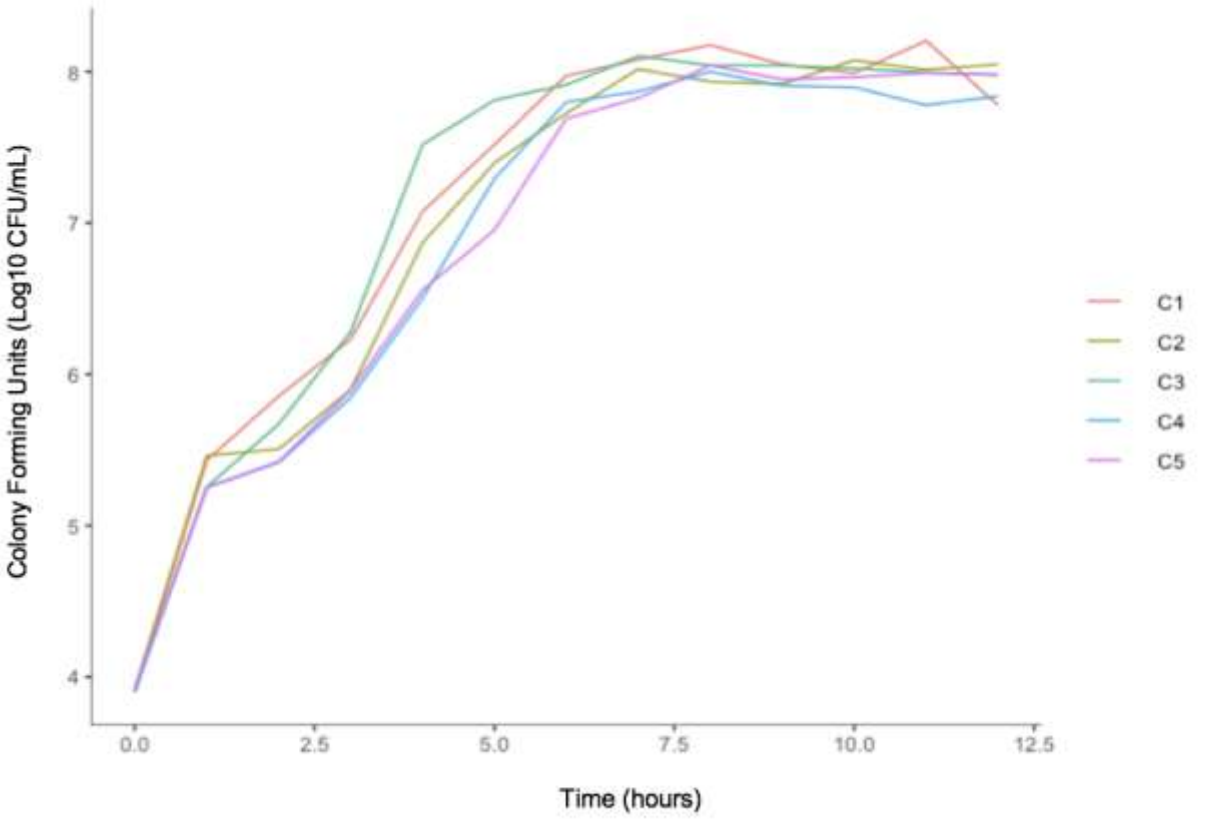


Figure 8.15 Manual growth analysis of *E. coli* MG1655 WT strain by bacterial enumeration. Displaying the growth curve of the *E. coli* MG1655 WT strain in MH broth at 37ºC. All five colonies (C1-C5) were grown within conical flasks and samples were taken every hour for 12 hours. The bacterial samples were used to perform serial dilutions which were plated on MH plates and incubated O/N at 37ºC. Following incubation, the log10 CFU/mL was determined. There was no significant difference in the bacterial counts between all colonies of the *E. coli* MG1655 WT strain (P=0.2573).

Comparable growth profiles were seen for the MIC cultures, with similar overall trajectories for all five cultures (Figure 8.16). The stationary phase of growth emerged approximately at the same time point as the WT, however the initial gradient during the exponential phase of growth was steeper for the WT cultures. Additionally, the maximum bacterial count obtained within the growth analysis of the MIC cultures was 7.77085 log10 CFU/mL. There was no significant difference detected in the

bacterial counts of the MIC cultures over the 12-hour duration of the growth curve (P=0.08545). Though the p-value implied borderline significance.

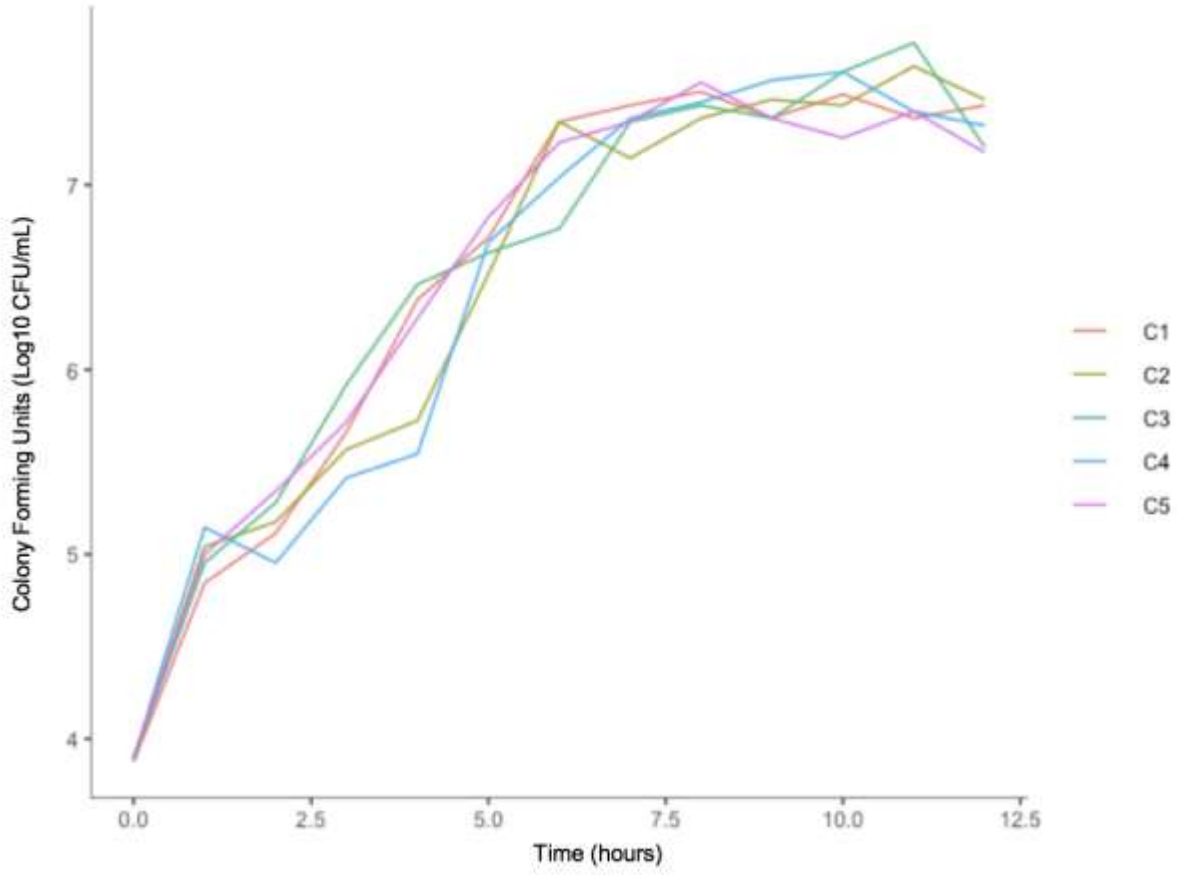


Figure 8.16 Manual growth analysis of *E. coli* MG1655 MIC strain by bacterial enumeration. Displaying the growth curve of the *E. coli* MG1655 MIC strain in MH broth at 37ºC. All five colonies (C1-C5) were grown within conical flasks and samples were taken every hour for 12 hours. The bacterial samples were used to perform serial dilutions which were plated on MH plates and incubated O/N at 37ºC. Following incubation, the log10 CFU/mL was determined. There was no significant difference in the bacterial counts between all colonies of the *E. coli* MG1655 MIC strain (P=0.08545).

The manual growth curve of the x10 MIC strain (Figure 8.17), displayed similar trajectories with no significant difference in the bacterial counts of the five independent cultures during the 12-hour growth analysis (P=0.09198). However, again the obtained p-value suggested borderline significance. In contrast to the rapid

onset of bacterial growth over the first 2 hours observed for the WT and MIC strains, the x10 MIC cultures appeared to have an extended lag phase where bacterial counts remained low for the first 2.5 hours. Here, unlike the turbidity results, the maximal bacterial count obtained for the x10MIC cultures was comparable to the MIC strain, with a bacterial count of 7.57978 log10 CFU/mL after 9 hours of growth.

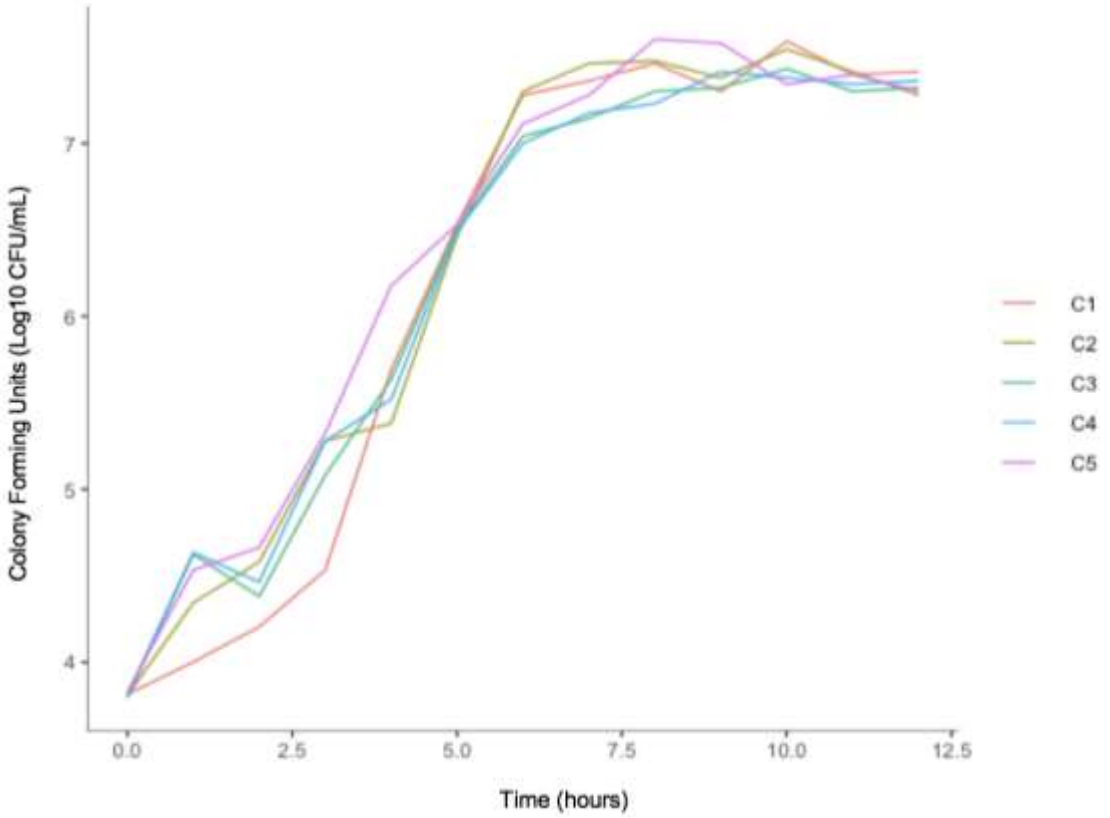


Figure 8.17 Manual growth analysis of *E. coli* MG1655 X10MIC strain by bacterial enumeration. Displaying the growth curve of the *E. coli* MG1655 X10MIC strain in MH broth at 37ºC. All five colonies (C1-C5) were grown within conical flasks and samples were taken every hour for 12 hours. The bacterial samples were used to perform serial dilutions which were plated on MH plates and incubated O/N at 37ºC. Following incubation, the log10 CFU/mL was determined. There was no significant difference in the bacterial counts between all colonies of the *E. coli* MG1655 X10MIC strain (P=0.09198).

When analysing the averaged log10 CFU/mL growth collectively for the sensitive and resistant strains (Figure 8.18). There were similarities in the growth over the first hours for the WT and MIC strains, which was dissimilar to the extended two-hour lag phase recognised with the x10 MIC cultures. As previously discussed all strains appeared to plateau after approximately 6 hours of growth, with almost identical stationary phase profiles for the two resistant strains of *E. coli* MG1655. Despite this over the 12-hour growth analysis there was a significant difference in the bacterial counts for the sensitive and resistant strains (P=5.02e-06). The greatest variation in bacterial counts was observed independently between all the WT and X10MIC resistant strain (P=1.177e-05). There was no significant difference in the bacterial counts of the two resistant strains (P=0.2519) consistent with the OD_{600} findings. These results indicated compensated bacterial growth due to the acquisition of gentamicin resistance.

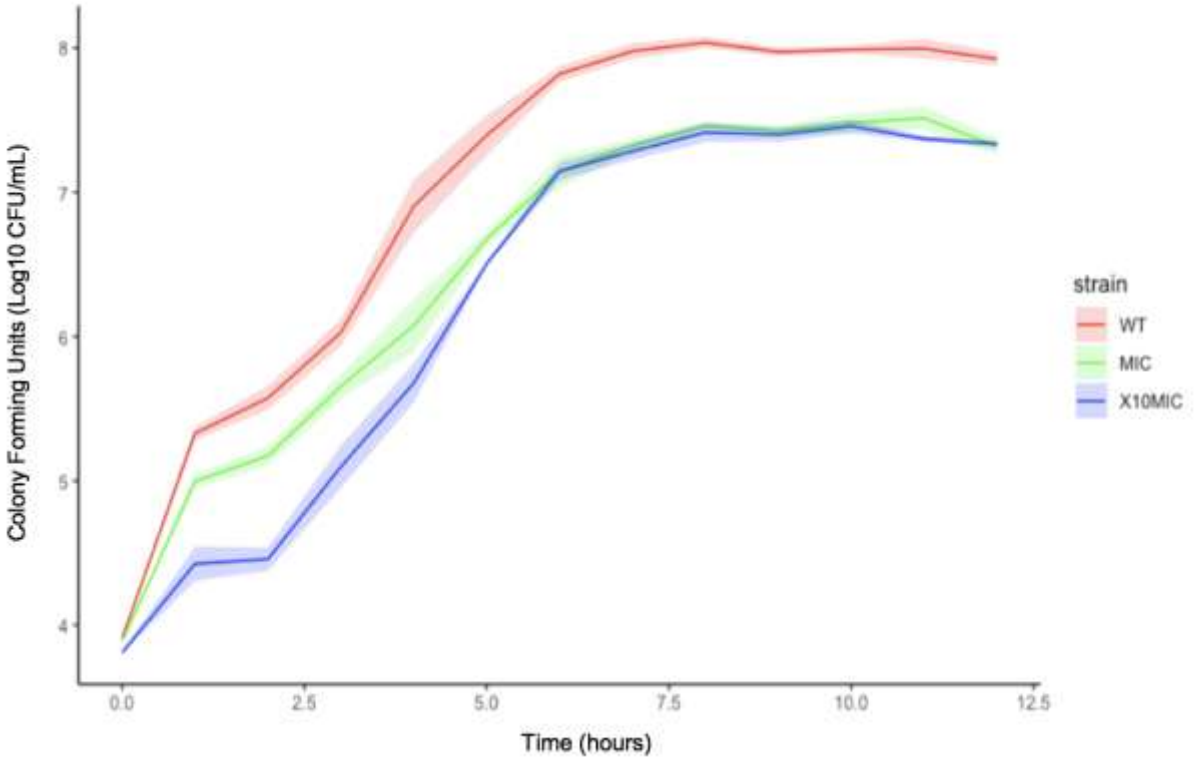


Figure 8.18 Manual growth analysis of gentamicin sensitive and resistant strains of *E. coli* MG1655 by bacterial enumeration. Graph showing the growth curve of the *E. coli* MG1655 WT, MIC and x10 MIC strains in MH broth at 37ºC. The five colonies (C1-C5) per strain were grown within conical flasks and samples were taken every hour for 12 hours. The bacterial samples were used to perform serial dilutions which were plated on MH plates and incubated O/N at 37ºC. Following incubation, the log10 CFU/mL was determined. The bacterial counts (log10 CFU/mL) per colony for each strain were averaged and the standard error was calculated. The error ribbons represent the standard error of the mean. There was a significant difference in the growth rates of the sensitive and resistant strains of *E. coli* MG1655 (P=5.02e-06).

Comparable to the previous data sets, the normality tests for the statistical analysis revealed the bacterial enumeration data did not follow normal distribution even following transformation. The graphical normality and transformation plots can be found in the appendix (Figure A.28). Thus, the non-parametric *Kruskal–Wallis* test was applied to determine significance.

8.3.3.1 Manual exponential growth analysis bacterial enumeration

As described previously with the turbidity methods, a supplementary analysis was performed on the viable bacterial counts of the sensitive and resistant strains during the exponential phase of growth. The bacterial enumeration of the *E. coli* MG1655 WT strain, revealed an earlier initiation of exponential growth when compared to the turbidity methods. The exponential phase followed one hour of growth and extended for a further five. The independent exponential growth analysis plots for the *E. coli* MG1655 WT, MIC and x10 MIC strains can be found in the appendix (Figure A.29-A.31).

To summarise and directly compare the exponential growth of the three strains, the averaged bacterial counts were determined. This revealed clear distinctions in the exponential growth of the WT strain compared the two resistant strains (Figure 8.19). The bacterial counts achieved by the WT strain during the exponential growth extended from 5.3305 log10 CFU/mL at beginning of the phase to 7.8202 log10 CFU/mL after six hours of growth. The bacterial counts obtained by the MIC strain were lower than that of the WT at the start of the exponential phase 4.9974 log10 CFU/mL and this was maintained throughout the duration of the phase. Following this trend, the x10MIC strain had an even further reduced bacterial count after 1 hour of 4.4261 log10 CFU/mL, which was sustained for a further hour. Despite this after three hours of exponential, the x10MIC strain rapidly increased in bacterial number to achieve almost identical averaged log10 CFU/mL values to that of the MIC, which were 7.1440 log10 CFU/mL and 7.1470 log10 CFU/mL for the MIC and x10MIC strains respectively.

There were significant differences in the exponential growth counts between the sensitive and resistant strains of *E. coli* (P=0.001568). Comparable to the previous bacterial growth analysis findings, the greatest difference in exponential counts was observed between the WT and X10MIC strain (P=0.0009513). There was no significant difference in the exponential counts between the two resistant strains (P=0.01836).

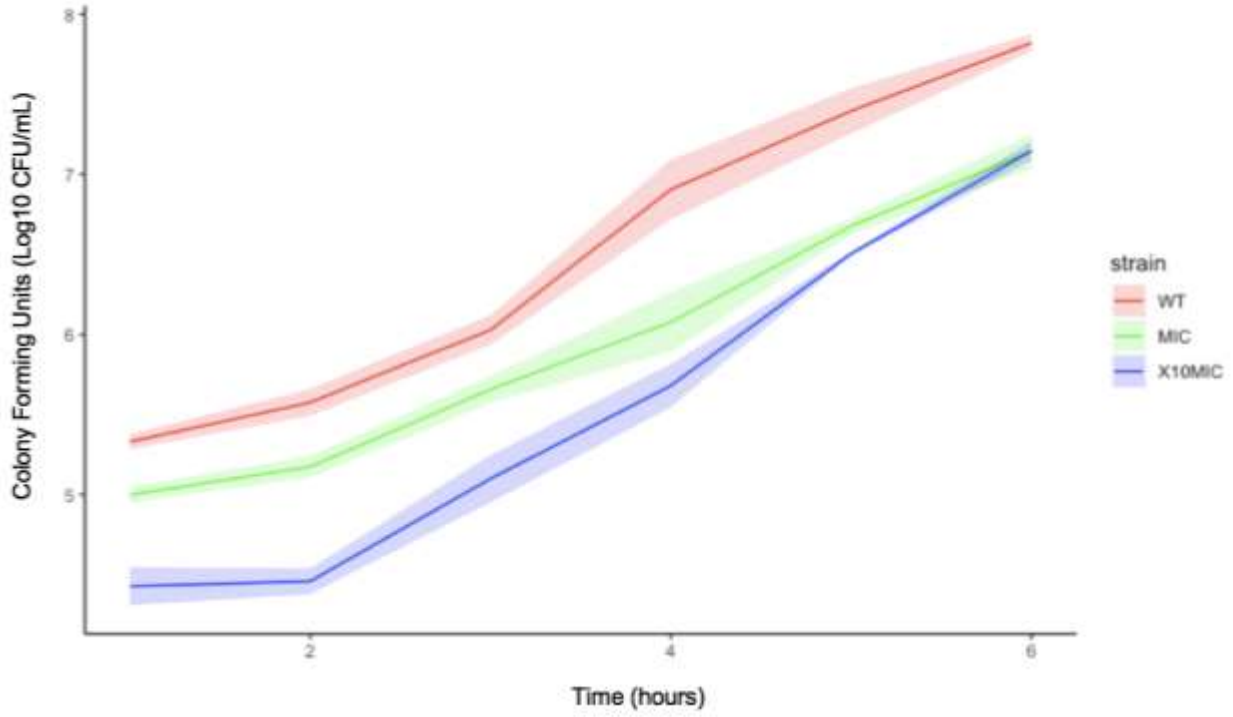


Figure 8.19 Manual exponential growth analysis of gentamicin sensitive and resistant strains of *E. coli* MG1655 by bacterial enumeration. Graph showing the exponential growth curve of the *E. coli* MG1655 WT, MIC and x10 MIC strains in MH broth at 37ºC. The five colonies (C1-C5) per strain were grown within conical flasks and samples were taken every hour for 12 hours. The bacterial samples were used to perform serial dilutions which were plated on MH plates and incubated O/N at 37ºC. Following incubation, the log10 CFU/mL was determined. The bacterial counts (log10 CFU/mL) per colony for each strain were averaged and the standard error was calculated. The error ribbons represent the standard error of the mean. There was a significant difference in the growth rates of the sensitive and resistant strains of *E. coli* MG1655 (P=0.001568).

8.3.4 Bacterial dose response analysis

A comparative dose response analysis was then conducted on the resistant strains in the presence of the selective concentrations of the antimicrobial agents. This was then directly compared to the growth of the WT population in a non-selective environment. Owing to the previously determined accuracy of OD_{600} measurements, it was decided that the dose response analysis will be conducted manually using OD_{600} measurements. As with the growth curve analysis, the response of each strain will be discussed individually, prior to a comparative analysis.

As seen with the growth curve analysis, the five colonies of *E. coli* MG1665 WT revealed similar phenotypic trajectories of growth (Figure 8.20). Comparable to the manual growth curve turbidity analysis, a lag phase was observed over the first 200 minutes and the maximal OD_{600} value attained ranged between 0.6 and 0.8 OD units. The discrepancies between the five cultures emerged following the exponential phase of growth, nevertheless there was no significant difference in the growth of the WT strain of *E. coli* MG1655 after 12 hours of growth (P=0.08049).

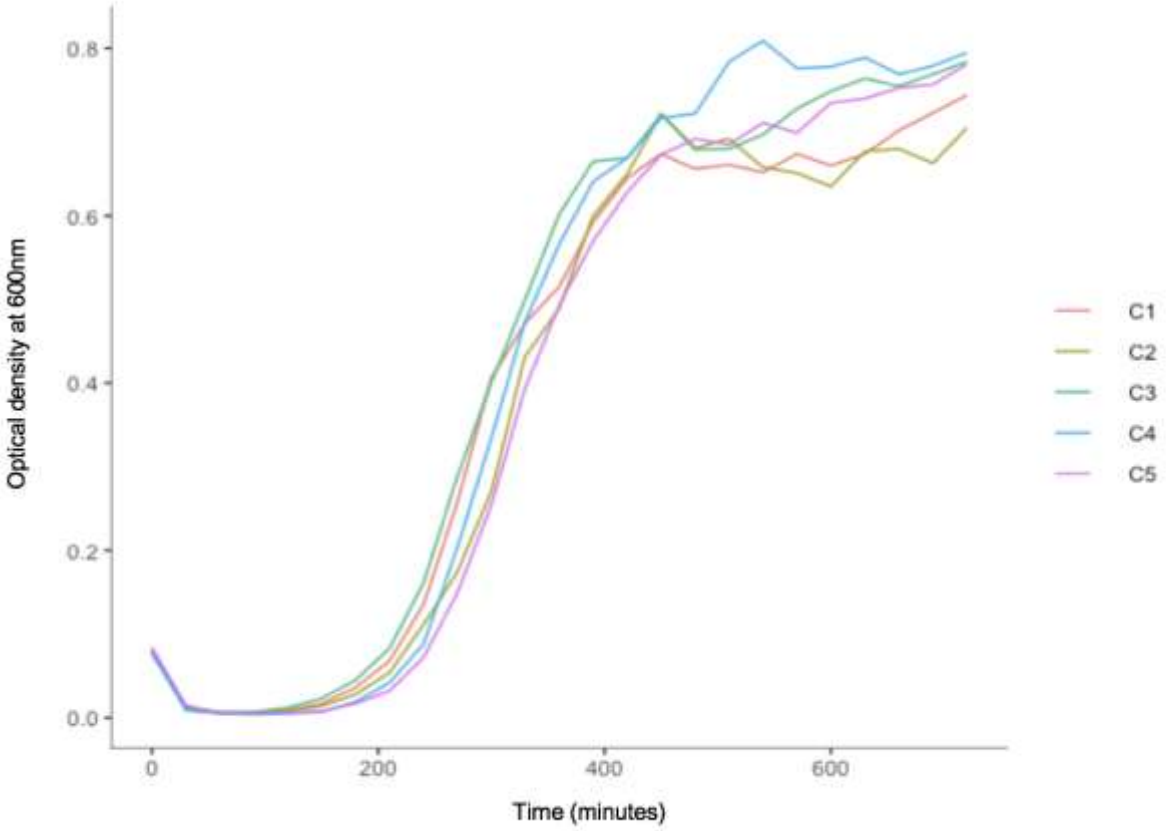


Figure 8.20 Manual dose response analysis of *E. coli* MG1655 WT strain. Displaying the growth curve of the *E. coli* MG1655 WT strain in MH broth at 37ºC. All five colonies (C1-C5) were grown within conical flasks and the OD_{600} values were measured every hour for 12 hours manually using a spectrophotometer. There was no significant difference in the growth rates between all colonies of the *E. coli* MG1655 WT strain (P=0.08049).

In comparison to the growth curve analysis of the resistant strain selected at the MIC of gentamicin in a non-selective environment, there appeared to be an extended lag phase of growth for four of the five MIC cultures lasting approximately 300 minutes (Figure 8.21). Succeeding this prolonged lag phase, there was a rapid growth response for all five cultures, surpassing the optimal OD_{600} of the MIC strain in the absence of the antimicrobial selective pressures. This result suggests, providing there is a genotypic adaptation within the MIC cultures, there is increased growth in the presence of antimicrobial agents. There was a significant difference in the growth of

the MIC cultures in the presence of the antimicrobial selective pressure (2mg/L gentamicin) over 12 hours (P=0.007877).

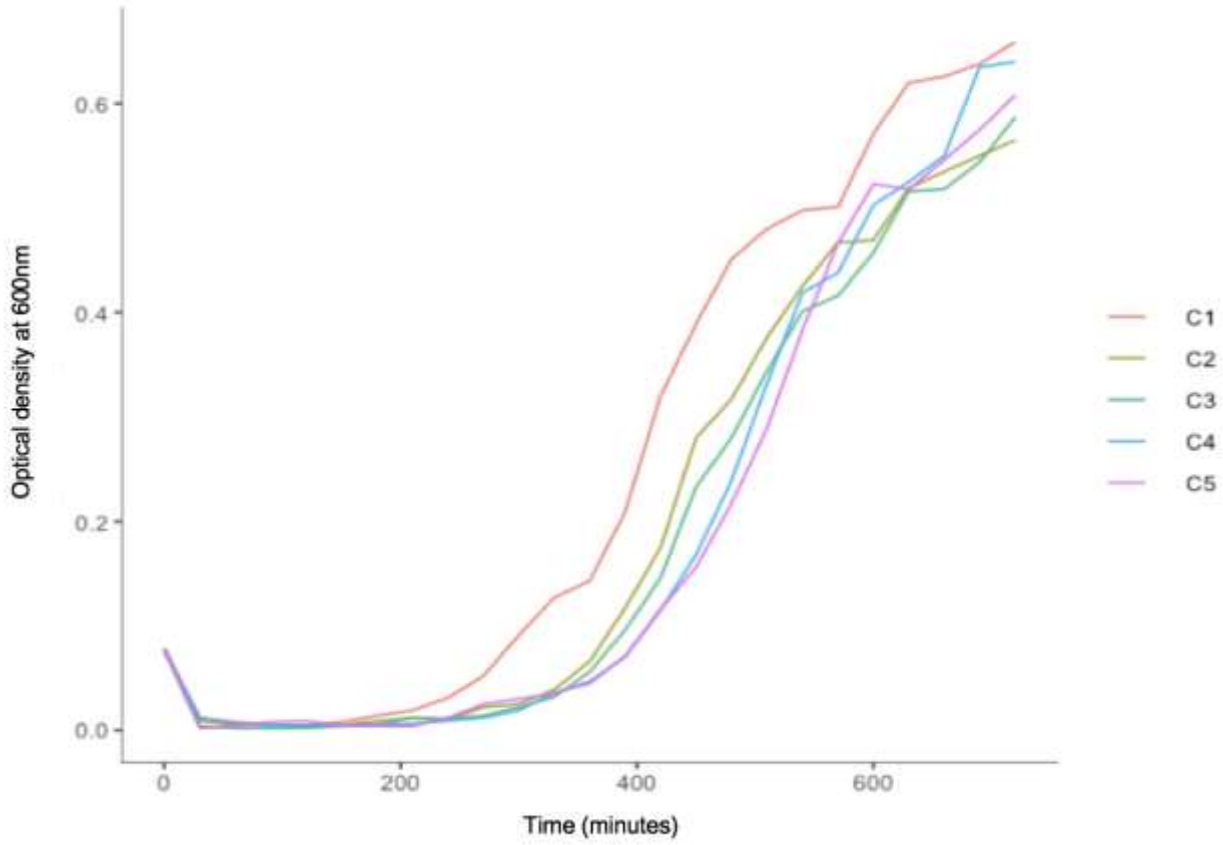


Figure 8.21 Manual dose response analysis of *E. coli* MG1655 MIC strain. Displaying the growth curve of the *E. coli* MG1655 MIC strain in MH broth supplemented with the MIC of gentamicin (2mg/L) at 37ºC. All five colonies (C1-C5) were grown within conical flasks and the OD_{600} values were measured every hour for 12 hours manually using a spectrophotometer. There was no significant difference in the growth rates between all colonies of the *E. coli* MG1655 MIC strain (P=0.007877).

Conversely, this was not observed for the x10 MIC strain (Figure 8.22). Following, the inoculation into a medium containing the high dose antimicrobial agent, there was a rapid reduction in the OD_{600} values which was not recovered within the 12-hour growth analysis. The data here revealed rapid bacterial decline upon the initial introduction to media containing x10MIC of gentamicin. This suggested that selection strength is an

important parameter which affects the phenotypic diversity of evolving bacterial populations, such that under strong selective pressures bacterial survival takes immediate priority over the presence of resistance-conferring mutations (Oz *et al*., 2014). This trend in dose response was seen for all five cultures of the X10 MIC strain and therefore unsurprisingly, there was no statistical significance difference in the dose response of the x10 MIC strain (P>0.9). Further, this suggested that for an accelerated response to occur in the presence of an antimicrobial agent, the bacterial populations must adapt to the selective pressure, and failing this the bacterial response is inhibited.

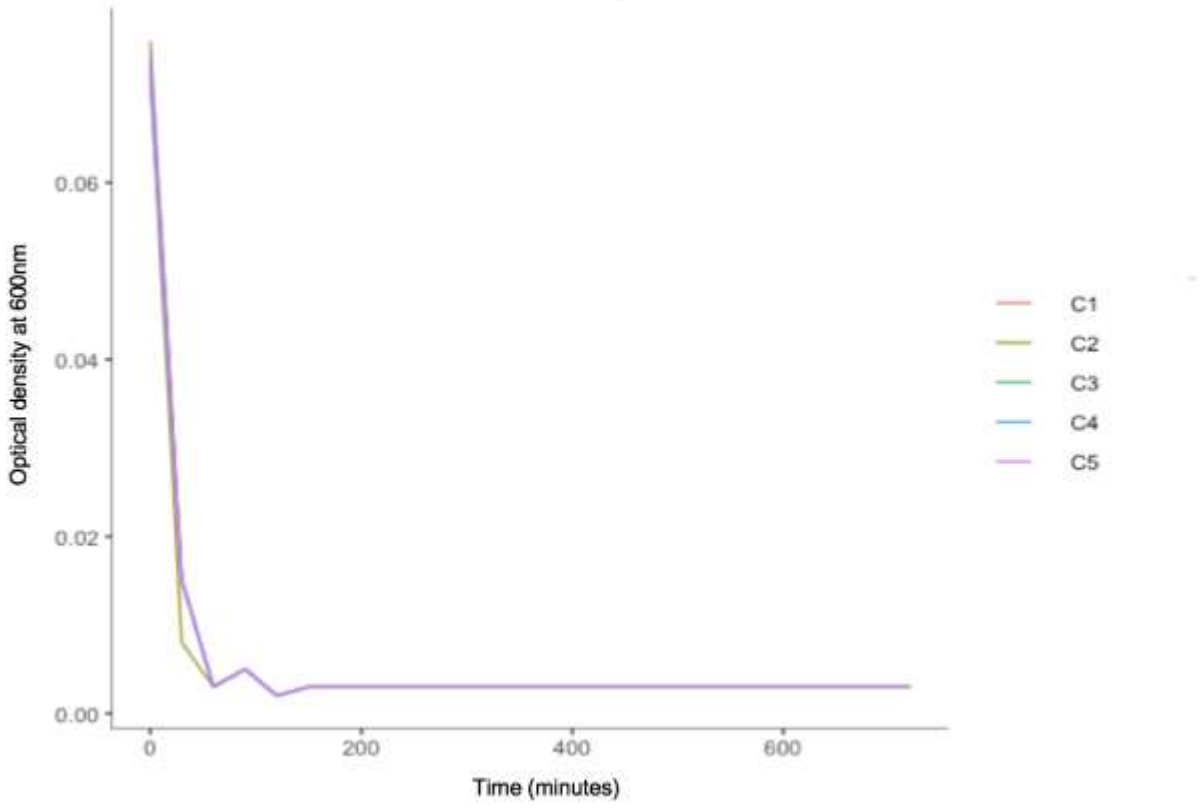


Figure 8.22 Manual dose response analysis of *E. coli* MG1655 X10MIC strain. Displaying the growth curve of the *E. coli* MG1655 X10MIC strain in MH broth supplemented with X10 the MIC of gentamicin (20mg/L) at 37ºC. All five colonies (C1-C5) were grown within conical flasks and the OD_{600} values were measured every hour for 12 hours manually using a spectrophotometer. There was no significant difference in the growth rates between all colonies of the *E. coli* MG1655 X10MIC strain.

When examining the averaged dose response of the three strains directly, the extended lag phase displayed within the resistant cultures is accentuated (Figure 8.23). There were clear discrepancies between gradients of the WT and MIC strains, where there appears to be a right shift in the positive slope of the MIC strain. This result also highlighted how rapidly the MIC cultures responded once adapted to the antimicrobial selective pressure. Here, there was a significant difference in the bacterial growth in response to presence of antimicrobial agents between the sensitive and resistant strains of *E. coli* MG1655 (P=2.2e-16).

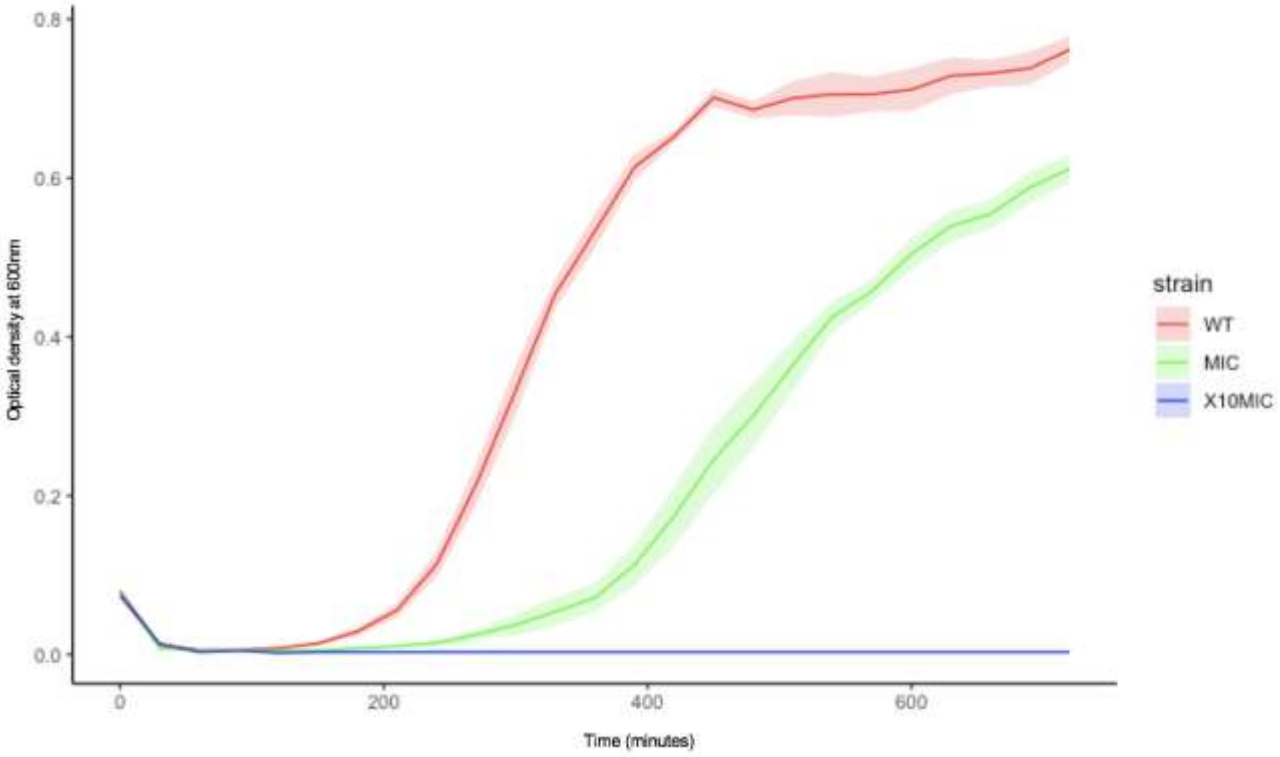


Figure 8.23 Manual dose response analysis of gentamicin sensitive and resistant strains of *E. coli* MG1655. Graph showing the growth curve of the *E. coli* MG1655 WT, MIC and x10 MIC strains in MH broth supplemented with 0, 2 and 20mg/L respectively at 37ºC. All five colonies (C1-C5) per strain were grown within conical flasks and the OD_{600} values were measured manually using a spectrophotometer. The OD_{600} values obtained from each colony per strain were averaged and the standard error was calculated. The error ribbons represent the standard error of the mean. There was a significant difference in the growth rates of the sensitive and resistant strains of *E. coli* MG1655 (P= 2.2e-16).

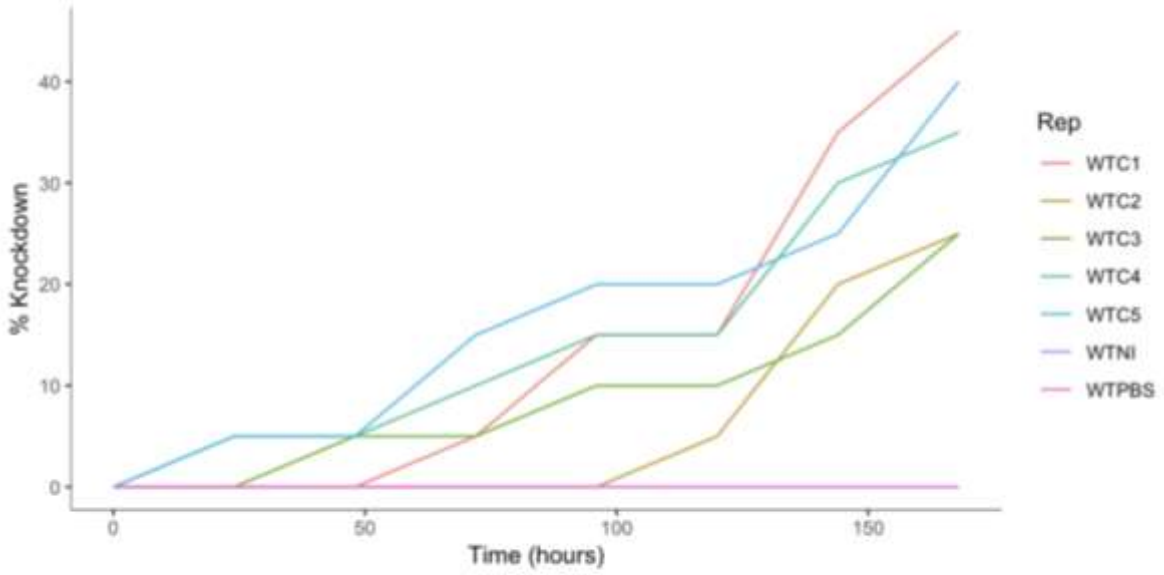
Similarly, to the previous bacterial growth data, the dose response data set was not normally distributed and therefore the non-parametric *Kruskal–Wallis test was used to determine significance.* The graphical normality and transformation plots can be found in the appendix (Figure A.32).

8.3.5 Bacterial virulence analysis

The microbial virulence using the *G. mellonella* toxicity model was assessed over seven days. Twenty *G. mellonella* were inoculated with five cultures taken from the sensitive and resistant strains of *E. coli* MG1655 cultivated in a non-selective environment, to determine the fitness costs associated with resistance development. The larvae were assessed daily using four key observations: survival, activity, cocoon formation and melanisation. As with the bacterial growth analysis, the statistics and graphical visualisation of data was completed using RStudio using the non-parametric *Kruskal–Wallis test*. The graphical normality and transformation plots for the virulence data can be found in the appendix (A.33).

The virulence of the sensitive and resistant strains will be discussed independently, prior to a comparative analysis. To assist in the interpretation of survival plots, a supplementary table has been included within the figure displaying the daily accumulated *G. mellonella* post-infection. Furthermore, an image has been added of the *G. mellonella* following the 7-day experimentation period.

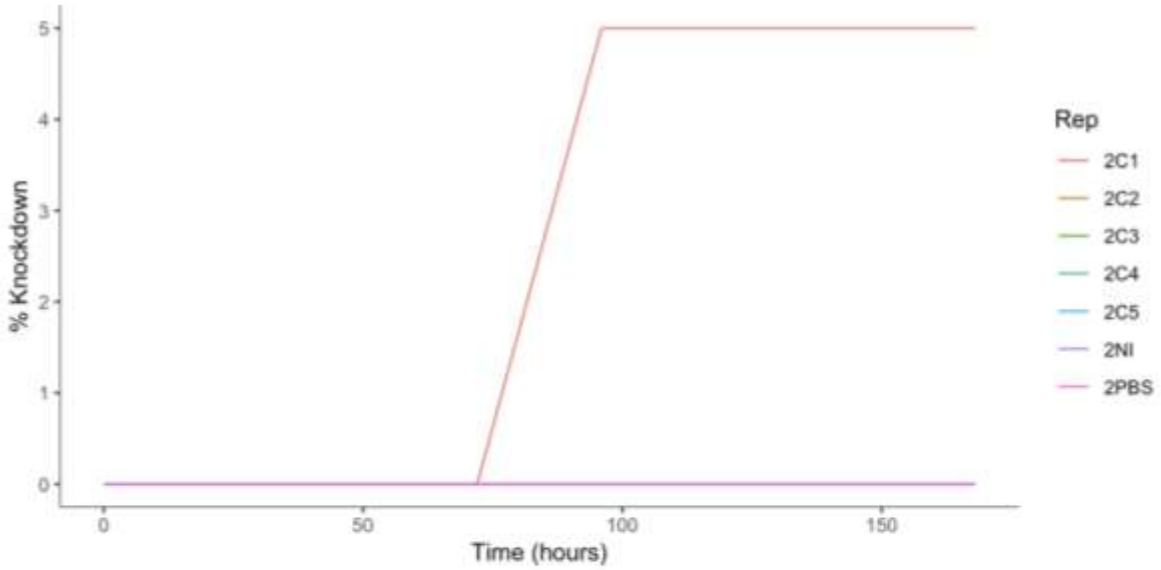
Five bacterial WT cultures were each individually inoculated into twenty *G. mellonella* larvae at time point zero and the percentage knockdown was monitored daily over a period of 7 days (Figure 8.24). Following inoculation, the initial death of larvae was observed after 1 day for WT cultures C4 and C5. The survival deteriorated thereafter for larvae inoculated with all five WT cultures. After day seven, the % knockdown of larvae inoculated with all five bacterial WT cultures was between 25% and 45%. Although there was significant variability in the virulence profiles of the five WT cultures, the overall trend depicts increased virulence after one-day post-inoculation. There was 100% survival for the non-injected and PBS controls after 7 days.



Gentamicin Sensitive *E. coli* (WT)	Accumulative *Galleria mellonella* deaths over time post-infection						
	WT colonies					Controls	
Time (hours)	C1	C2	C3	C4	C5	PBS	NI
0	0	0	0	0	0	0	0
24	0	0	0	1	1	0	0
48	0	0	1	1	1	0	0
72	1	0	1	2	2	0	0
96	3	0	2	4	4	0	0
120	3	1	2	4	4	0	0
144	7	4	3	5	5	0	0
168	9	5	5	5	8	0	0

Figure 8.24 Survival analysis of *G. mellonella* after *E. coli* MG1655 WT infection. A total of twenty *G. mellonella* larvae were injected with five independent culture of the *E. coli* MG1655 WT (WT C1-C5) and incubated at 37ºC for a period of seven days, and mortality (% Knockdown) was monitored daily. The two experimental controls correspond to the non-infected *G. mellonella* (WTNI) and larvae injected with sterile PBS (WT PBS). The accumulative *G. mellonella* deaths over time post-infection were represented within a tabular format, to stipulate the individual larvae deaths per WT colony which was difficult to interpret from the survival curve.

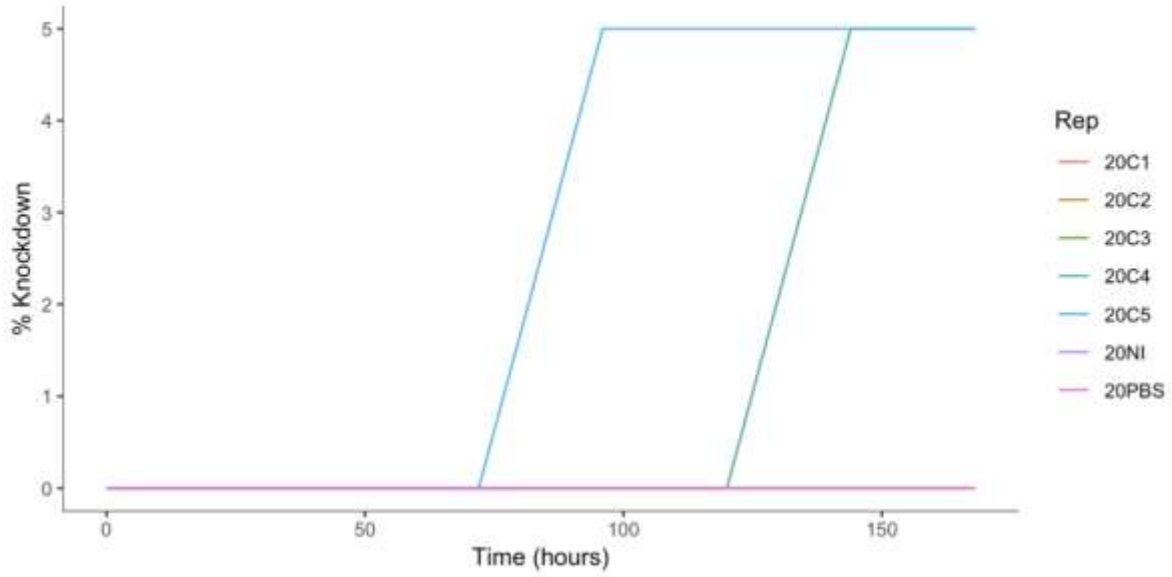
This was then compared to the virulence profiles of the larvae inoculated with the five bacterial resistant cultures selected at the MIC of gentamicin (Figure 8.25). The initial knockdown of a single larvae in MIC cultures 2C1 and 2C3, was observed 96 hours post-inoculation. This 5% knockdown of larvae was only apparent in *galleria* inoculated with two of the five MIC cultures, and the remaining larvae inoculated with the three remaining MIC bacterial cultures exhibited 100% survival. The 5% reduction in survival, is difficult to identify within the survival plot owing to the identical virulence profiles of MIC cultures 2C1 and 2C3, though evident within the table (Figure 8.25). There was 100% survival for the non-injected and PBS controls after 7 days. This result suggested a reduced virulence of resistant cultures.



Gentamicin Resistant *E. coli* (MIC)	Accumulative *Galleria mellonella* deaths over time post-infection						
	MIC colonies					Controls	
Time (hours)	C1	C2	C3	C4	C5	PBS	NI
0	0	0	0	0	0	0	0
24	0	0	0	0	0	0	0
48	0	0	0	0	0	0	0
72	0	0	0	0	0	0	0
96	1	0	1	0	0	0	0
120	1	0	1	0	0	0	0
144	1	0	1	0	0	0	0
168	1	0	1	0	0	0	0

Figure 8.25 Survival analysis of *G. mellonella* after *E. coli* MG1655 MIC infection. A total of twenty *G. mellonella* larvae were injected with five independent cultures of the *E. coli* MG1655 MIC (2mg/L) strain (2 C1-C5) and incubated at 37ºC for a period of seven days, and mortality (% Knockdown) was monitored daily. The two experimental controls correspond to the non-infected *G. mellonella* (2 NI) and larvae injected with sterile PBS (2 PBS). The accumulative *G. mellonella* deaths over time post-infection were represented within a tabular format, to stipulate the individual larvae deaths per MIC colony which was difficult to interpret from the survival curve.

Similar virulence profiles were observed for larvae inoculated with the five bacterial resistant cultures selected at the x10 MIC of gentamicin (Figure 8.26). As with the larvae inoculated with the MIC cultures, the initial knockdown of a single larvae was observed 96 hours post-inoculation for the x10 MIC culture 20C5. The later reduction in *galleria* survival occurred six days post-infection for the x10 MIC cultures 20C1 and 20C4. There was 100% survival for the non-injected and PBS controls after 7 days. These results correspond to the virulence profiles of the MIC strains and further suggest reduced virulence of resistant cultures.



Gentamicin Resistant *E. coli* (X10 MIC)	Accumulative *Galleria mellonella* deaths over time post-infection						
	X10 MIC colonies					Controls	
Time (hours)	C1	C2	C3	C4	C5	PBS	NI
0	0	0	0	0	0	0	0
24	0	0	0	0	0	0	0
48	0	0	0	0	0	0	0
72	0	0	0	0	0	0	0
96	0	0	0	0	0	0	0
120	0	0	0	0	0	0	0
144	1	0	0	1	1	0	0
168	1	0	0	1	1	0	0

Figure 8.26 Survival analysis of *G. mellonella* after *E. coli* MG1655 X10MIC infection. A total of twenty *G. mellonella* larvae were injected with five independent cultures of the *E. coli* MG1655 X10MIC (20mg/L) strain (20 C1-C5) and incubated at 37ºC for a period of seven days, and mortality (% Knockdown) was monitored daily. The two experimental controls correspond to the non-infected *G. mellonella* (20 NI) and larvae injected with sterile PBS (20 PBS). The accumulative *G. mellonella* deaths over time post-infection were represented within a tabular format, to stipulate the individual larvae deaths per X10MIC colony which was difficult to interpret from the survival curve.

To provide a direct comparison of the virulence profiles of the sensitive and resistant strains, the percentage knockdown values were averaged for all five bacterial cultures per strain (Figure 8.27). The survival graph accentuated the substantial differences in virulence of the sensitive and resistant strains of *E. coli* MG1655. The two resistant strains displayed highly similar virulence profiles. Following one-day post-infection, there was a significant difference in the virulence of the WT compared to both gentamicin resistant strains (P= 3.928e-09). Unsurprisingly, there was no significant difference between the virulence of the MIC and X10 MIC resistant cultures (P>0.999). The bacterial virulence results indicated a direct relationship between *galleria* survival and time post inoculation and more significantly a reduction in the virulence of resistant strains. This suggested a fitness cost in virulence associated with the acquisition of gentamicin resistance.

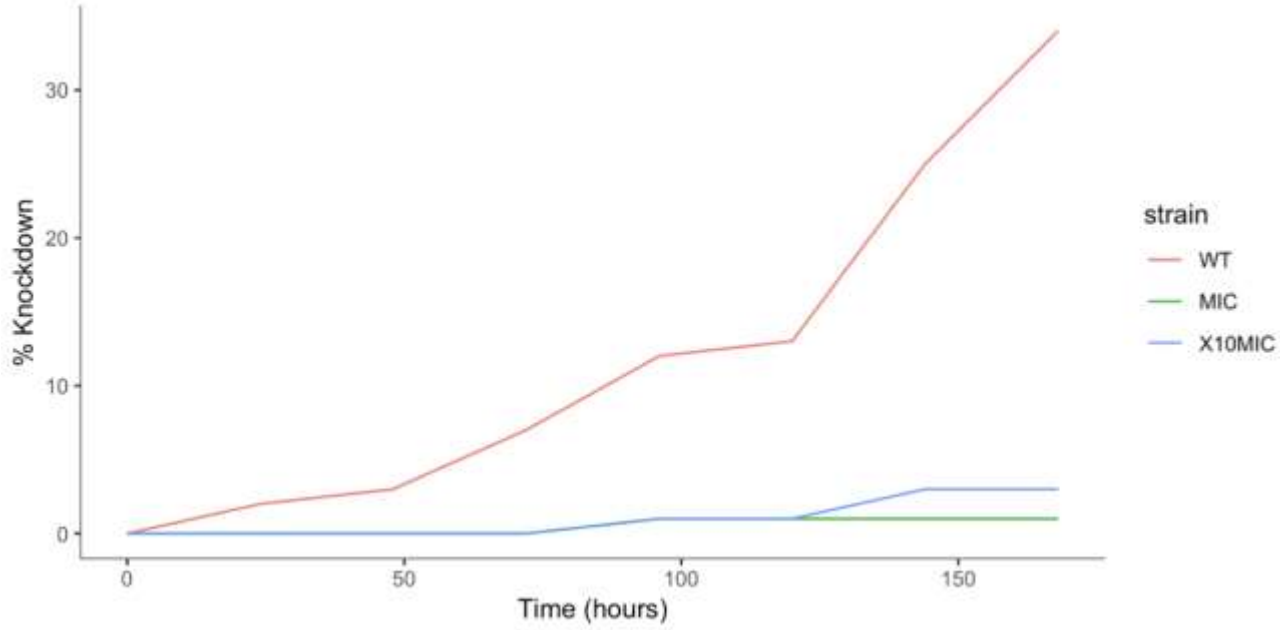


Figure 8.27 Survival analysis of *G. mellonella* after infection with sensitive and resistant strains of *E. coli* MG1655. A total of twenty *G. mellonella* larvae were injected with five independent cultures of *E. coli* MG1655 WT, MIC and X10MIC strains and incubated at 37ºC for a period of seven days, and mortality was monitored daily. The survival graph displays the averaged *G. mellonella* deaths over time post-infection. There were statistically significant differences in the virulence of the *E. coli* MG1655 WT when compared to the two resistant strains (P=3.928e-09).

As indicated previously an image has been included showing all *galleria* larvae seven-days post-inoculation (Figure 8.28). Unfortunately, the lighting within the microbiology laboratory had compensated the image quality, making it difficult to visualise the larvae inoculated with cultures C4 and C5 for all three bacterial strains.

Taking into consideration the four key observations, and the knowledge that melanisation typically begins with distinct black spots on the larvae, which progresses to the larvae turning black when they die. It was evident that there were detrimental consequences on the physical state of the *galleria* seven days post WT infection, as seen with the increased number of black larvae when compared to the larvae groups infected with the resistant cultures.

Interestingly, it was concluded that all four-observations should be considered when assessing the physical conditions of the *G. mellonella* larvae. This was more evident when evaluating the virulence of the resistant strains than their sensitive WT counterpart. Looking at the graph in figure (8.25), there was 100% survival observed for all larvae inoculated with the MIC cultures 2C2 and 2C3. Nevertheless, when looking at the image below some of the *galleria* within these groups appeared to be melanised. Despite this, a physical assessment identified that the larvae were either reactive post-stimulation or had partial cocoon formation. An additional observation not included in this image, was the visualisation of increased silk production within the larvae injected with PBS in the experimental control group.

Figure 8.28 *G. mellonella* infection with gentamicin sensitive and resistant strains of *E. coli* MG1655. Examination of *G. mellonella* larvae seven days post-infection with sensitive and resistant strains of *E. coli* MG1655. The image displays the twenty *galleria* larvae injected with five independent cultures (C1-C5) of the three bacterial strains: WT, MIC (2mg/L GEN Rx) and X10 MIC (20mg/L GEN RX). The *galleria* larvae were considered dead when melanised and non-responsive upon stimulation. Healthy larvae were typically cream in colour.

8.3.6 SEM bacterial morphology analysis

In addition to an assessment of fitness, a preliminary study was conducted to examine the morphophysiological changes in the bacterial cells induced by antimicrobial therapy. Through high resolution SEM valuable information was obtained regarding the surface morphology, shape and spatial organisation of sensitive and antibiotic treated *E. coli* MG1655.

8.3.6.1 SEM morphological analysis of *E. coli* MG1655 WT Strain

The sensitive WT strain of *E. coli* MG1655, was prepared for SEM in MH medium. At a low magnification, the distribution of WT bacterial cells appeared uniform across the SEM micrograph (Figure 8.29 A). At a higher magnification, the length of the WT bacterial cells was determined to be approximately 2µm, and most WT bacterial cells displayed smooth intact surfaces with little surface topology (Figure 8.29 B). In addition, ruptured WT bacterial cells were detected infrequently across the SEM film, indicated by the black arrows in figure (8.29 B). This disruption of WT bacterial cells could have occurred during the process of sample preparation and fixation.

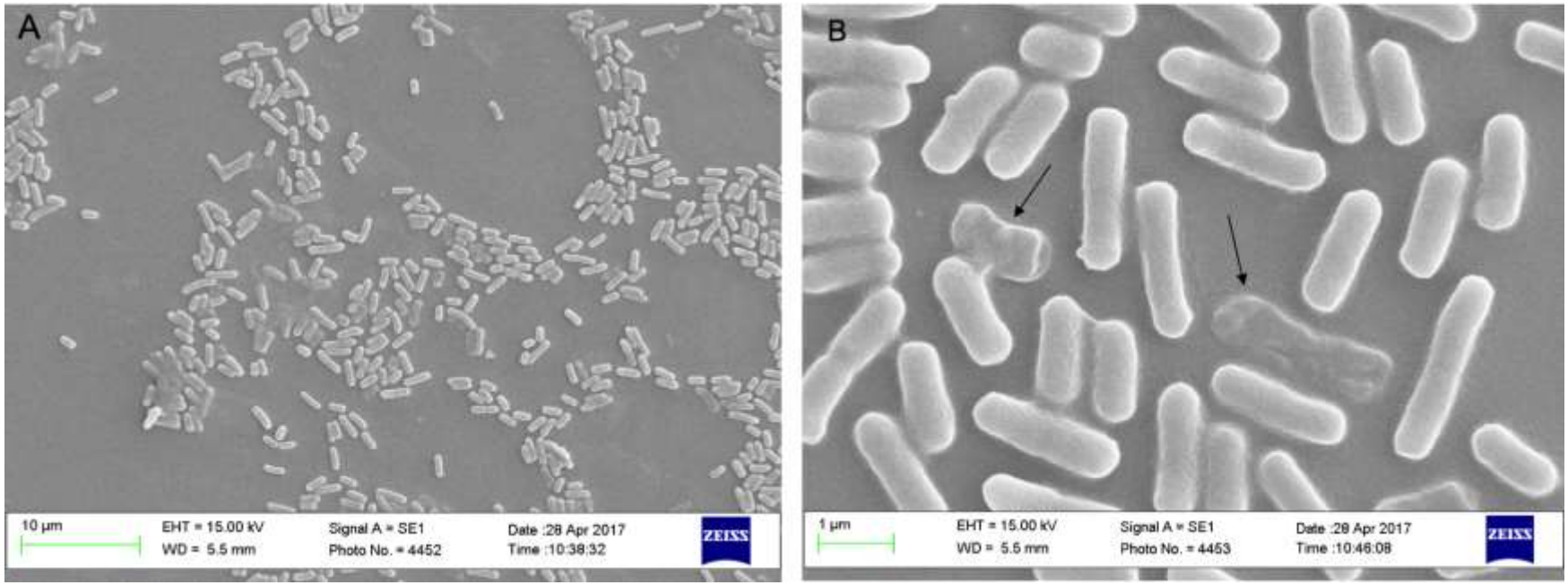


Figure 8.29 SEM of *E. coli* MG1655 WT strain. SEM micrographs of the *E. coli* MG1655 WT strain at low magnification (A) and high magnification (B). The black arrows denote ruptured WT bacterial cells. The size bar was 10μm and 1μm for the two respective magnifications.

8.3.6.2 SEM morphological analysis of the gentamicin resistant *E. coli* MG1655 strain

The resistant strain of *E. coli* MG1655 was selected at 2mg/L of gentamicin, which was the predetermined MIC value. The resistant bacterial cells were prepared for SEM analogously to the sensitive cells in MH medium. At the lowest magnification, the resistant bacterial cells displayed significant changes in their spatial distribution. The bacterial cells appeared to aggregate, and this clustering was consistent across the entire SEM film (Figure 8.30 A). In addition, a matrix like substance appeared concealing some of the resistant bacterial cells, notably reducing the bacterial surface topology as indicated by the black arrow in figure (8.30 A).

At a higher magnification, the average length of the resistant bacterial cells was significantly reduced to approximately 1µm (Figure 8.30 B). At this magnification, there was a distinct increase in the bacterial compactness of resistant cells in comparison to the sensitive WT in figure (8.29 B). Comparable to that seen at a lower magnification, there appears to be a layer over the resistant bacterial cells reducing the surface topology. In addition, although relatively infrequent, cell surface protrusions were identified on several resistant bacterial cells on the SEM film, as indicted by the black arrow (Figure 8.30 B), discussed further in the discussion.

The preliminary results derived from SEM, indicated a reduced ability of the gentamicin resistant strain of *E. coli* MG1655 to grow to maximum lengths, which supported the findings from the bacterial fitness assessments, where reduced growth was observed within the resistant strains (Sections 8.3.1-8.3.3).

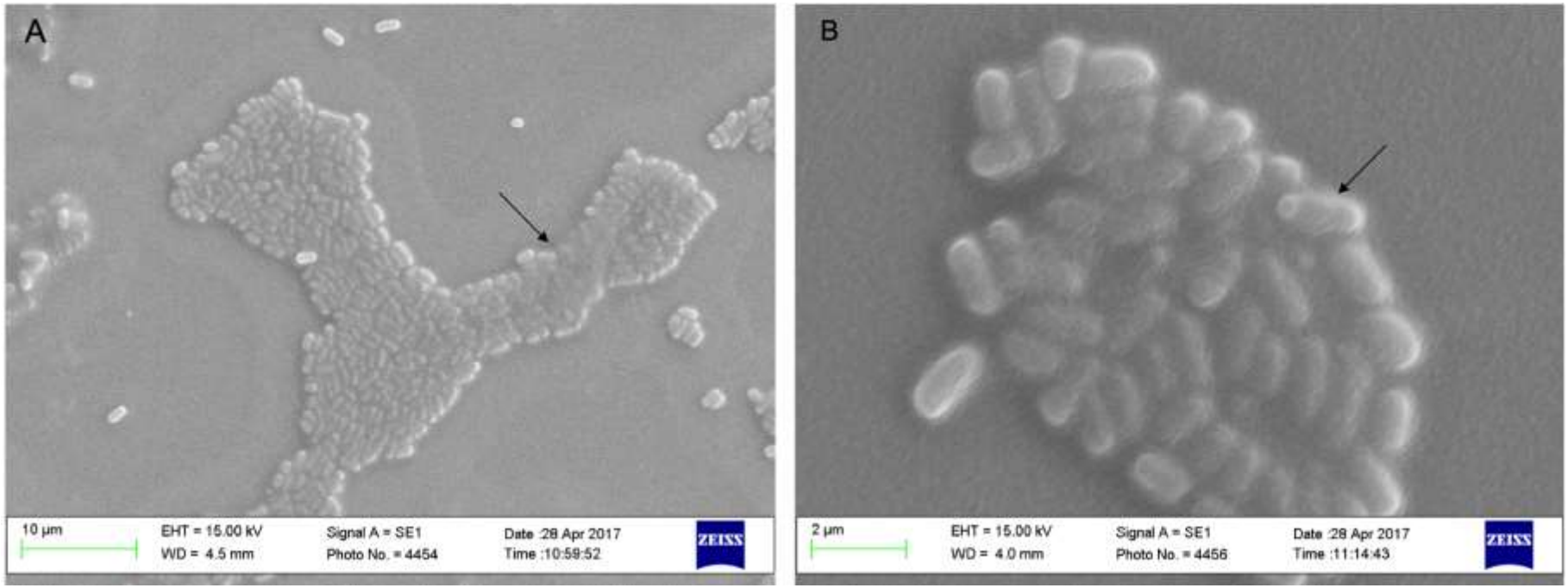


Figure 8.30 SEM of the gentamicin resistant *E. coli* MG1655 MIC strain. SEM micrographs of the *E. coli* MG1655 MIC strain at low magnification (A) and high magnification (B). The black arrows denote the matrix coating of bacterial cell and bacterial cell protrusions respectively. The size bar was 10μm and 2μm for the two individual magnifications.

8.4 Discussion

8.4.1 Fitness costs of antimicrobial resistance

The aim of this chapter described in section (8.1), was completed as follows. The fitness costs associated with the acquisition of gentamicin resistance were determined through the experimental assessment of both bacterial growth and virulence. As hypothesised there was a substantial reduction in fitness observed within the resistant isolates of *E. coli* MG1655. A further morphological assessment of the resistant isolates using SEM revealed alterations in the size and spatial organisation of resistant colonies. In accordance with the WGS data, compensatory evolution had not occurred within the resistant isolates to alleviate the fitness cost of antibiotic resistance.

The acquisition of drug resistant mutations within bacterial populations is frequently associated with biological costs which alter the fitness of organisms (Spratt, 1996). Predominantly, as resistant conferring mutations target fundamental biological processes within the bacterial cell, which in the absence of antibiotics can impose substantial costs (Lenski, 1998). Therefore, it is unsurprising that the mutations identified in both the rrn and *fusA* genes confer fitness costs due to their essential involvement in ribosomal protein synthesis. In addition, it is likely that the fitness costs associated with the acquisition of gentamicin resistance were accentuated, as no secondary site compensatory mutations were identified during experimental evolution to alleviate such costs (Björkman *et al.,* 2000).

The substantial reduction in fitness within the experimentally evolved resistant strains of *E. coli* MG1655, reaffirmed the findings from the phylogenetic analysis where the resistant mutations discovered in the laboratory, were not identified within natural environments. Since natural selection acts to eliminate resistant mutants with elevated fitness costs, it might be presumed that under natural conditions the evolved resistant mutants, impaired in both growth and virulence would be outcompeted by mutants which do not confer these substantial fitness costs (Andersson and Hughes, 2011). Furthermore, the selection of aminoglycoside resistant mutants, relies heavily upon the fitness costs associated with the nature of the resistance mechanisms (Beceiro, Tomás and Bou, 2013). As demonstrated by the acquisition of aminoglycoside resistance by aminoglycoside modifying enzymes, such as Rmt methyltransferases, which confer no fitness costs (Gutierrez *et al.,*2012). Hence, it could be predicted that the acquisition of aminoglycoside resistance through ribosomal mutagenesis of the *fusA* and rrn genes would not be detected in a clinical setting, when alternative mechanisms exist which confer no fitness costs.

8.4.2 The fitness costs associated with chromosomal mutations

The fitness costs associated with antimicrobial resistance have been found to be dependent upon the precise mutations which mediate resistance to a given antimicrobial agent (Bottger, Pletschette and Andersson, 2005). A well-established model used to explore the relationship between chromosomal mutations and bacterial fitness, is through mutagenesis of the *rpsL* gene, which is renowned to confer streptomycin resistance (Nair *et al.,*1993). In streptomycin, resistant strains of *E. coli*, the replacement of the lysine residue at position 42, with the amino acids arginine,

asparagine and threonine, caused apparent discrepancies in bacterial fitness due to the restrictiveness of the mutations (Funatsu and Wittmann, 1972). This is extremely pertinent as resistant strains with restrictive ribosomal mutations have been shown to exhibit slower rates of translation compared to strains with non-restrictive mutants (Sander *et al.,*2002). Therefore, an assumption can be made that the reduced growth displayed by the gentamicin resistant strains of *E. coli* MG1655, may be enhanced by the replacement of a proline residue with the restrictive amino acid threonine within the ribosomal protein EF-G, which has been correlated with slower growing, less fit bacterial strains (Chiou and Jones, 1995).

8.4.3 The fitness costs associated with rRNA mutations

Ribosomal RNA mutations in antibiotic resistant bacteria, are frequently associated with structural alterations within the 70s-ribosomal complex and reduced translation efficiency, therefore rRNA mutations would be predicted to confer substantial fitness costs (Binet and Maurelli, 2005). However, the extent in which rRNA mutations contribute to bacterial fitness was found to vary significantly depending on the degree of conservation at the mutation position, and the level of resistance conferred (Shcherbakov *et al.,*2010). Several studies have reported that rrn mutations which confer high levels of antibiotic resistance, have little or no effect on bacterial fitness within the mycobacterial species (Sander *et al.,*2002; Shcherbakov *et al.,*2010; Nessar *et al.,*2011). In addition, a frequently detected laboratory rrn mutation linked to streptomycin resistance, exhibits high fitness costs and has yet to be identified within clinical strains (Springer *et al.,*2001). Therefore, owing to the variability in fitness costs of resistant rrn mutations, it is extremely difficult to make

any definitive conclusions regarding the contribution of these mutations to the observed reduction in bacterial fitness of the experimentally evolved gentamicin resistant strains.

8.4.4 Bacterial virulence and antimicrobial resistance

As previously described the larvae of *G. mellonella* are being used increasingly in the study of bacterial pathogenicity and virulence (Ramarao, Neilsen-Leroux and Lereclus, 2012; Cook and McArthur, 2013; Li *et al.,*2018). To date, very few studies report using *G. mellonella* to explore the fitness costs in bacterial virulence associated with antimicrobial resistance. Nevertheless, mice models have formerly been implemented to study the fitness costs in bacterial virulence of antibiotic resistant *Salmonella typhimurium,* where resistant mutants were avirulent in mice (Bjorkman, Hughes and Andersson, 1998). A comparative study using mice and *G. mellonella* as models to study the virulence of *Pseudomonas aeruginosa* mutants disclosed highly comparable virulence findings (Jander, Rahme and Ausubel, 2000; Norville *et al.,*2014).

Therefore, as non-human primate models are not used in research at Kingston University, *Galleria* was utilised as an alternative model within this study. The results from the virulence analysis were comparable to the findings of Bjorkman *et al*. (1998), as antibiotic resistant strains of bacteria exhibited substantial fitness costs in bacterial virulence. In addition, the observed reduction in bacterial virulence, provides an explanation for the persistence of resistant bacteria in animals long after the termination of antimicrobial treatment (Smith, 1975).

8.4.5 Morphological assessments using SEM

Scanning electron microscopy has been used to explore the detrimental effects on the bacterial cell envelope induced by antimicrobial agents (Hartmann *et al.,*2010). SEM analysis revealed the aminoglycoside gentamicin used within this study, severely reduces the integrity of the *P. aeruginosa* cell envelope, which manifest as holes due to the displacement of critical metal ions in the membrane of bacterial cells (Martin and Beveridge, 1986). This is the first report, of SEM being utilised to visualise the structural morphology of resistant strains in the absence of antimicrobial selective pressures. The SEM analysis exposed a reduction in the size of resistant colonies of *E. coli* MG1665, which supports the observation of reduced growth of resistant strains. Despite these preliminary observations, the preparation of bacterial samples for SEM generally requires a dehydration phase which is associated with the collapse and shrinking of bacterial cells (Golding *et al.,*2016). Therefore, before any conclusions can be made, SEM must be repeated with an increased number of replicates.

8.4.6 Conclusion

The observed fitness costs associated with the acquisition of gentamicin resistance, are most likely due to the restrictive mutation in the *fusA* gene. The contribution of the 16s rRNA mutations in the reduction of bacterial fitness, is yet to be elucidated. As the resistant mutations are costly to the growth and virulence of *E. coli* MG1655, it would be expected that the gentamicin resistant strains would be outcompeted in the natural environment.

Chapter 9: General discussion and conclusions

The aims of the work outlined within this thesis were developed to produce an experimental tool to explore the evolution of antimicrobial resistance *in vitro*. Detailed bioinformatic and phylogenetic analysis, was used to determine the significance of the genotypic findings and finally the fitness of the resistant evolved strains were determined. In the following chapter, the individual aims outlined within each chapter will be addressed and any future work will be stipulated.

9.1 The ARGP experimental evolution model

Chapter 3, describes the extensive method development of the adaptive laboratory evolution tool, the ARGP, which facilitates the direct visualisation of resistance evolution across an antibiotic spatial concentration gradient. The application of spatial antibiotic concentration gradients in the study of microbial evolution has been previously described by Baym *et al*. (2016) and more recently by Ghaddar *et al*. (2018). The direct visualisation of the evolutionary progress of bacterial populations within the ARGP, enables researchers to determine the exact point in which resistance emerges across the antimicrobial gradient. Through subsequent mutational analysis of resistant isolates selected from the ARGP, using the indispensable tool of WGS, the underlying genotypic mutational pathways to resistance could be predicted (Baym *et al.,* 2017; Spicknall *et al.,* 2013).

The ARGP is a well-established model to monitor the evolution of antimicrobial resistance *in vitro.* The simplicity of the ARGP, suggests scope for its application within an educational setting, where it can be installed in teaching laboratories to

demonstrate the rapid rates of resistance evolution amongst bacterial species (Krist and Showsh, 2007; Kawecki *et al.,* 2012).

The ARGP holds potential for other areas of antimicrobial resistance research. In modern medicine, combinational therapy has been highly successful, however its effect on the evolution of antibiotic resistance is unclear (Hegreness *et al.,* 2008; Torella, Chait and Kishony, 2010). It has been suggested that synergistic combinations which are preferred clinically, may favour resistance evolution while antagonistic interactions can suppress the evolution of resistance (Michel *et al.,* 2008; Yeh *et al.,* 2010). There is also opportunity for investigating the effects of sub-lethal antibiotic concentrations on resistance evolution within the ARGP (Andersson and Hughes, 2012; Hughes and Andersson, 2012).

An interesting prospective application of the ARGP is within diagnostic experimental evolution during complex cases, where bacterial stains are removed from patients, evolved in real-time against a range of antimicrobial agents, and the information generated can be used to advise clinical decisions (Wood and Read, 2015). If successful, in incidences of chronic bacterial infections, treatment regimes could be selected based on evolutionary risk management (Palmer and Kishony, 2013). It is hoped that the ARGP will assist therapeutic strategies, in informing the design of novel antimicrobials and manipulating current antimicrobial agents to restrict or reverse the evolution of antimicrobial resistance (Baym, Stone and Kishony, 2016a). Finally, the ARGP could be applied to other areas of applied microbiology, including

the evolution of biocide resistance which is essential from an infection control perspective (Webber *et al.,* 2015).

9.2 Comparative genomic analysis through WGS

The WGS analysis described within chapter 4 revealed that the SNPs identified within the *fusA* and 16s rRNA genes, are the genotypic adaptations most likely associated with the acquisition of the gentamicin resistance. These findings in support of Toprak *et al*. (2011), suggest WGS should be routinely applied within laboratory evolution experiments, due to the discovery of resistance conferring mutations which would have been previously missed by the conventional sequencing of primary antimicrobial targets (Köser *et al.,* 2014; Baym *et al.,* 2015). These observations suggest adaptation to the aminoglycoside gentamicin, is not exclusive to mutations of the 16s rRNA target, and can occur more broadly through mutations in genes such as *fusA*, encoding proteins directly involved in the underlying biological process of ribosomal protein synthesis (Rodnina *et al.,* 1997). The genomic conclusions reached within this chapter, provided the grounds for subsequent investigations into the functional consequences of these adaptive mutations.

9.3 Bioinformatic analysis: the functional consequences of mutagenesis

Chapter 5 describes the identification of structural and physiochemical variations within the mutated 16s rRNA and *fusA* genes, which related specifically to the functional activity of gentamicin. The mutagenesis of the *fusA* gene, resulted in an amino substitution within a highly conserved region of EF-G between the boundary of domains 4 and 5, which plays an essential role in the conformational changes required during the ribosomal translocation stage of protein synthesis (Munro *et al.,*

2010). The replacement of a flexible proline residue at the molecular hinge of domain four, produced apparent alterations in the protein compressibility and tertiary structure of EF-G (Sansom and Weinstein, 2000). Furthermore, the identification of the *fusA* mutation in resequenced isolates obtained from the ARGP and serial passage system, suggested repeatability in the emergence of a resistant genotype in independently evolving populations of *E. coli* MG1655 conferring resistance to gentamicin (Weinreich *et al.,* 2006; Toprak *et al.,* 2013; Furusawa, Horinouchi and Maeda, 2018).

The functional analysis of the 16s rRNA mutations, was complicated by the presence of several copies of the rRNA operon in *E. coli* MG1655 (Kiss, Sain and Venetianer, 1977). Therefore, despite the observed structural alterations in the secondary structures of rna123 and rna154, it was impossible to determine whether gentamicin resistance was mediated by the mutations within the 16s rRNA genes (Springer *et al.,* 2001).

9.4 Molecular docking analysis of gentamicin

The data presented in chapter 6, employed atomic resolution crystal structures of EF-G bound to the 70s-ribosomal complex, to examine whether the binding of gentamicin was disrupted through ribosomal mutagenesis of the *fusA* gene (Lin *et al.,* 2015). Significantly, the *in silico* molecular docking analysis revealed that gentamicin was unable to bind to the decoding region of the 70s-ribosomal complex in its pre-translocational state, following the ribosomal mutagenesis of EF-G. Together with the findings from previous chapters, this undoubtedly suggested that *fusA* mutagenesis is the primary mechanism driving the evolution of gentamicin resistance *in vitro*.

9.5 Phylogenetic analysis of the *fusA* gene

Chapter 7 describes the phylogenetic analysis of the *fusA* gene. Reconstruction of the *fusA* phylogeny revealed evolutionary conservation across the *Escherichia* taxa and *Enterobacteriaceae* family. The phylogenetic analysis of the *fusA* mutant distribution within the *Enterobacteriaceae* family, showed no evidence of convergent evolution, as the mutant appeared independently and was not distributed across multiple branches of the phylogenetic tree (Hazbon *et al.,* 2008). The protein conservation analysis of EF-G, exposed positional conservation of the proline residue at codon 610 across the *Enterobacteriaceae* family, indicative of functional and structural significance (Ashkenazy *et al.,* 2010). Chapter 7 was concluded, by the unsuccessful identification of the experimentally evolved *fusA* mutant genotype within the PHE collection of carbapenemase and ESBL producing strains of *E. coli*.

The *fusA* mutant genotype has so far only been identified *in vitro*, which suggests that the genotype may be unable to thrive in natural and clinical settings owing to the presence of increased selective pressures, which were not incorporated within the experimental evolution model. Thus, it is likely that the resistant mutant may be eliminated from the population before reaching detectable limits. The work presented within this chapter, paved way for the investigation into the fitness costs associated with the mutagenesis of the *fusA* gene.

9.6 The bacterial fitness costs of gentamicin resistance

Chapter 8 describes the fitness costs associated with the acquisition of gentamicin resistance. The bacterial fitness investigations confirmed biological costs associated with gentamicin resistance development, which were reflected in reduced growth and virulence of the experimentally evolved resistant strain of *E. coli* MG1655. Interestingly, it has been reported that restrictive ribosomal mutations have detrimental impacts on bacterial growth, colonisation and pathogenesis (Bohman *et al.,* 1984; Chiou and Jones, 1995; Sander *et al.,* 2002; Chen *et al.,* 2013). Therefore, the fitness costs inferred by the resistant strain of *E. coli* MG1655, could be directly linked to the restrictive substitution of the amino acid proline within EF-G. The large fitness costs associated with the *fusA* resistant genotype, suggest bacterial populations carrying this mutation are likely to be outcompeted within the natural environment (Melnyk, Wong and Kassen, 2014).

9.7 Future work

Future work will include the study of a wide variety of antimicrobial agents with differing mechanisms and spectrums of activity. In addition, a more diverse range of microorganisms will be incorporated within the ARGP model, including clinically significant pathogens which appear on the antimicrobial resistance threat list (Ventola, 2015).

To demonstrate the importance of the single amino acid alteration in EF-G in the generation of gentamicin resistance in *E. coli*, site directed mutagenesis will be applied to generate *fusA* 610T mutants which will be introduced by plasmids into

gentamicin-susceptible strains of *E. coli* (Heermann, Zeppenfeld and Jung, 2008). Subsequent expression analysis of the *fusA* mutants in the recombinant strains of *E. coli*, will determine whether the 610T mutation in the *fusA* gene results in gentamicin resistance (Besier *et al.,* 2003). In addition, gene-inactivation techniques will be employed to determine whether rrn deletion strains of rna123 and rna154, can maintain a gentamicin resistant phenotype in *E. coli* MG1655 (Asai *et al.,*1999).
To further enhance the computational docking, it would be of interest to determine the crystal structures of EF-G bound to the ribosome in complex with gentamicin (Gagnon *et al*., 2012; Lin *et al.,* 2015). Supplementary to the structural investigations, smFRET could provide insights into the conformational dynamics during ribosomal translocation following aminoglycoside binding and *fusA* mutagenesis (Ha, 2001; Feldman *et al*., 2010; Petrov *et al.,* 2012; Shebl *et al.,* 2016).

To augment the bacterial fitness analysis and to further explore the evolutionary fate of the *fusA* resistant genotype, *in vitro* competition assays will be conducted to determine the relative fitness of the resistant and sensitive strains of *E. coli* MG1655 (Wiser and Lenski, 2015).

9.8 Concluding remarks

The ARGP experiments are conducted in a highly controlled laboratory environment, with minimal selective pressures acting on the evolving bacterial populations, which facilitates the identification of the core adaptive mutations governing resistance development *in vitro*. The aim of this thesis have been met as described above, however unfortunately due to time constraints mathematical modelling was not

performed. The findings presented here provide a grounded platform for the in depth analysis of the evolution of antimicrobial resistance. The research approach outlined within this thesis, will be utilised in a future CARA funded research project at Kingston University for the continued exploration of the evolutionary pathways to antimicrobial resistance in *Klebsiella*, *Pseudomonas* and *Mycoplasma* spp. under various environmental stress conditions.

Appendix

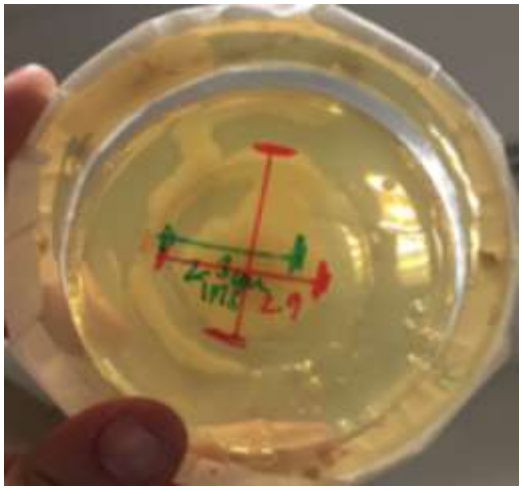


Figure A.1 The triphasic circle plate experiment: *E. coli* under trimethoprim selection. Image of the preliminary triphasic circle plate experiment within a petri dish displaying *E. coli* under trimethoprim selection following three days incubation at 37°C. The image clearly demonstrates the agar dehydration between the two outer antimicrobial phases. The lines drawn on the triphasic plates indicate the daily bacterial migration.

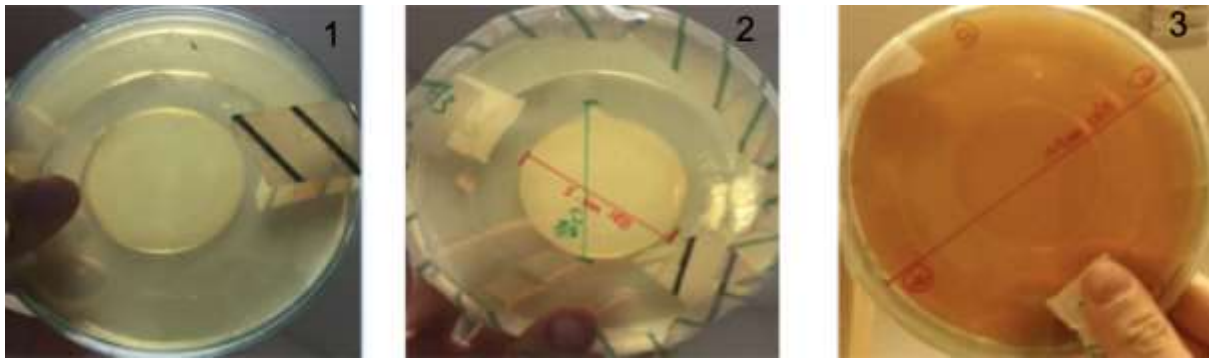


Figure A.2 The triphasic circle plate experiment: *P. mirabilis* under trimethoprim selection. The enlarged triphasic circle plate experiment within glass petri dishes. Displaying the daily images of *P. mirabilis* under trimethoprim selection over a period of three days (1-3) incubation at 37°C. The image demonstrates the rapid resistance evolution in the motile organism *P. mirabilis.* This was apparent by the decolourisation in the MH media after 3 days incubation (3). The lines drawn on the triphasic plates indicate the daily bacterial migration.

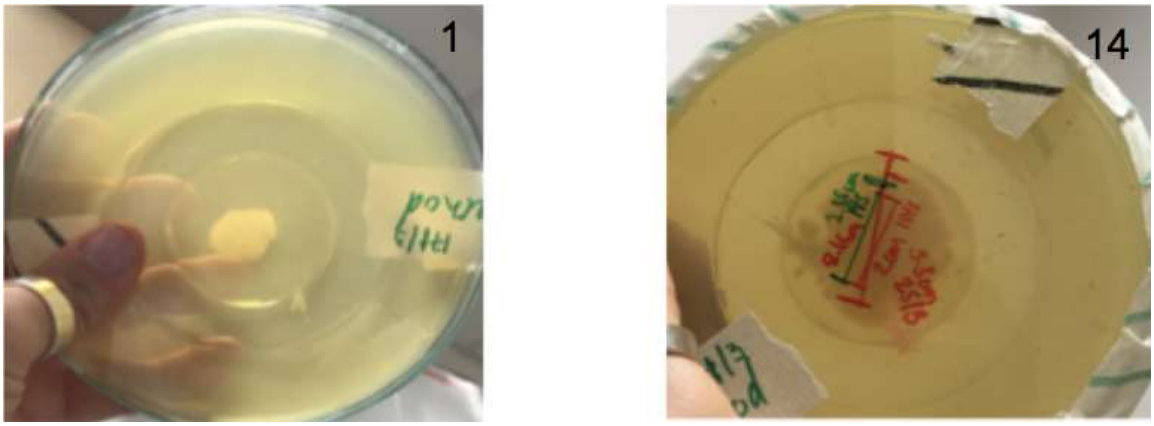


Figure A.3 The triphasic circle plate experiment: *E. coli* under trimethoprim selection over 2 weeks. The enlarged triphasic circle plate experiment within glass petri dishes. Displaying *E. coli* under trimethoprim selection following one day (1) and 14 days (14) incubation at 37°C. The image demonstrates the slow migration of *E. coli* after two weeks incubation. The growth of E. coli also appears sporadic within the initial phase contain no antimicrobial agent (14). The lines drawn on the triphasic plates indicate the daily bacterial migration. The decolourisation of the MH media was evident after 14 days incubation.

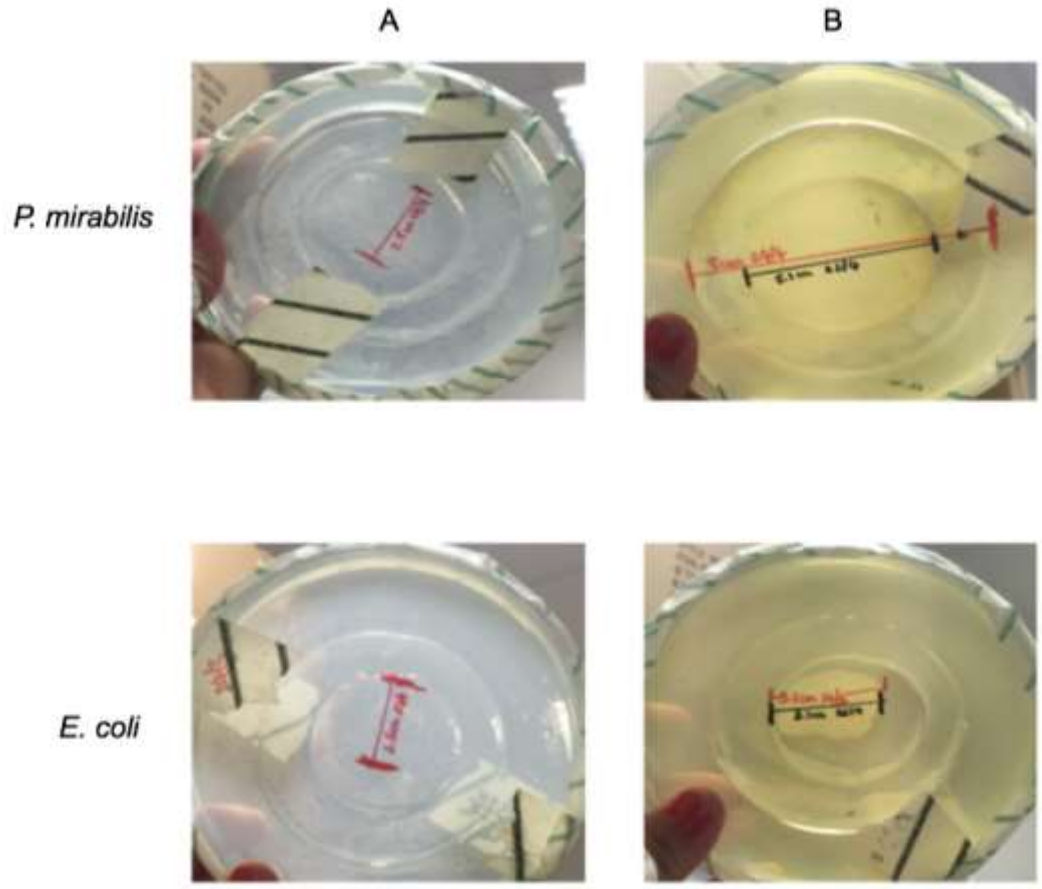


Figure A.4 The triphasic circle plate experiment: *E. coli* and *P. mirabilis* under trimethoprim selection in MH and M9 minimal media. The image displays a comparative analysis of the triphasic circle plate experiment for *E. coli* and *P. mirabilis* under trimethoprim selection following two days incubation at 37°C within two different growth medium M9 minimal media (A) and MH (B). The lines drawn on the triphasic plates indicate the daily bacterial migration. The M9 minimal media did not enable the growth of the two organisms. The MH triphasic circle plates facilitated the growth of the two organisms, but resistance evolution proceeded at faster rates for the motile organism *P. mirabilis*.

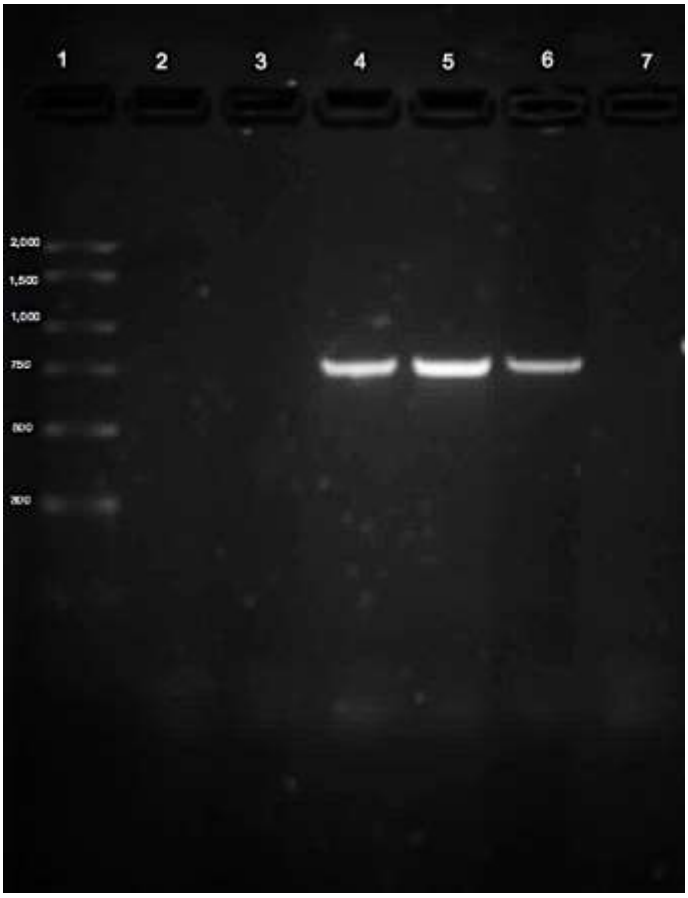


Figure A.5 Gel electrophoresis image of the genomic analysis of bacterial motility. The gel electrophoresis image displaying the PCR reactions of the five strains of *E. coli*: MG1655, UTI89, CFT073, TPA2743 and TPA4792 (wells 2-6). The gel electrophoresis image indicated apparent flhDC expression in *E. coli* strains: CFT073, TPA2743 and TPA4792.

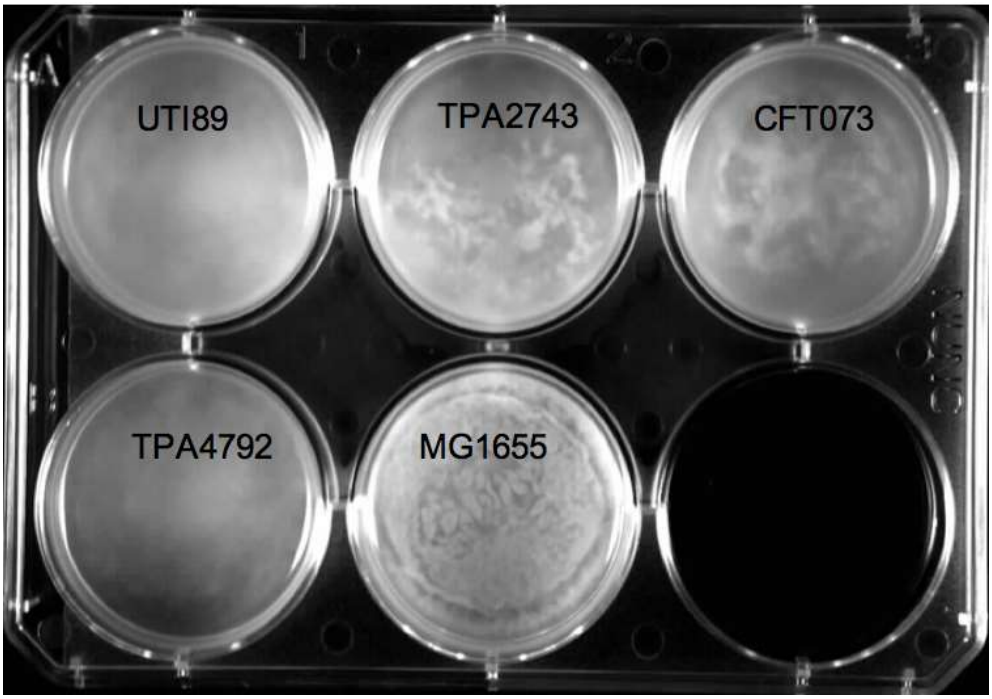


Figure A.6 The 0.3% 6-well plate motility assay. The phenotypic motility assay of the five *E. coli* strains (UTI89, TPA2743, CFT073, TPA4792 and MG1655) using a motility layer formed by MH broth solidified with 3% bacteriological agar [wt/vol]. Following O/N incubation at 37°C the bacterial motility can be observed on the surface of the motility layer within the each well.

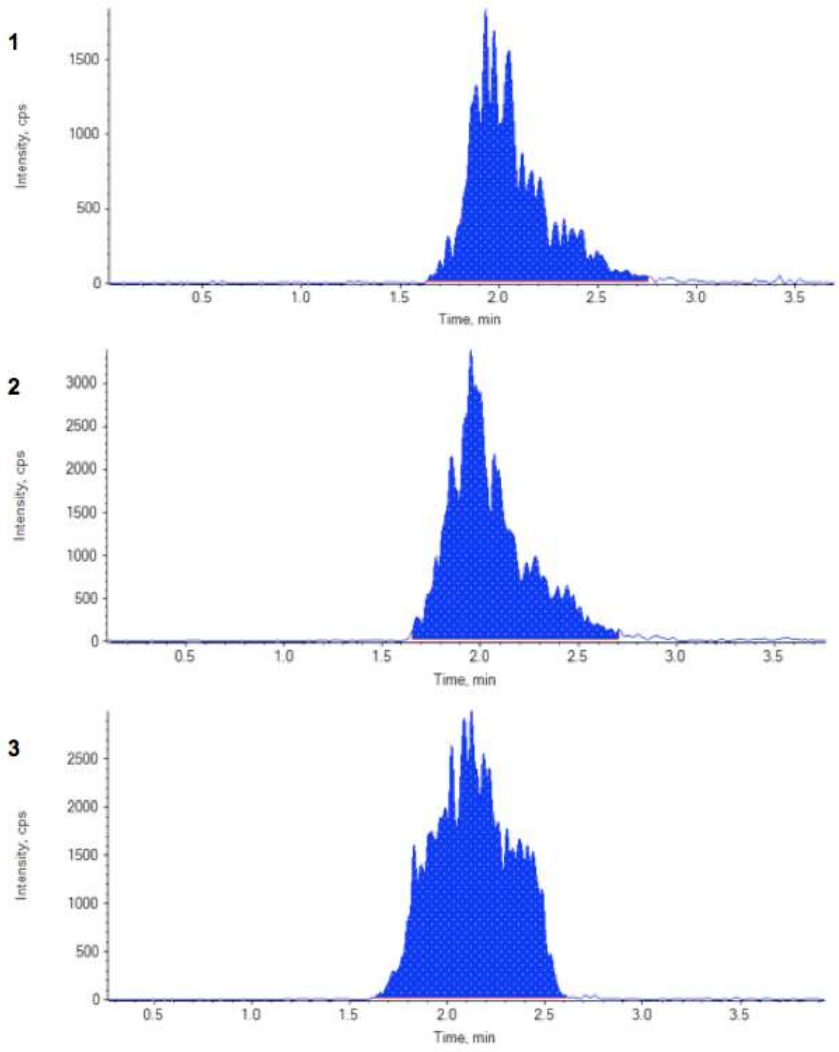


Figure A.7 LC-MS chromatogram 4mg/L gentamicin agar extraction. Image displaying the LC-MS chromatogram output of the 4mg/L Gentamicin agar extract for all three Gentamicin components: 1 (464.292/322.300), 2 (450.317/322.200) and 3 (478.325/322.200) presenting LC-MS intensity over time.

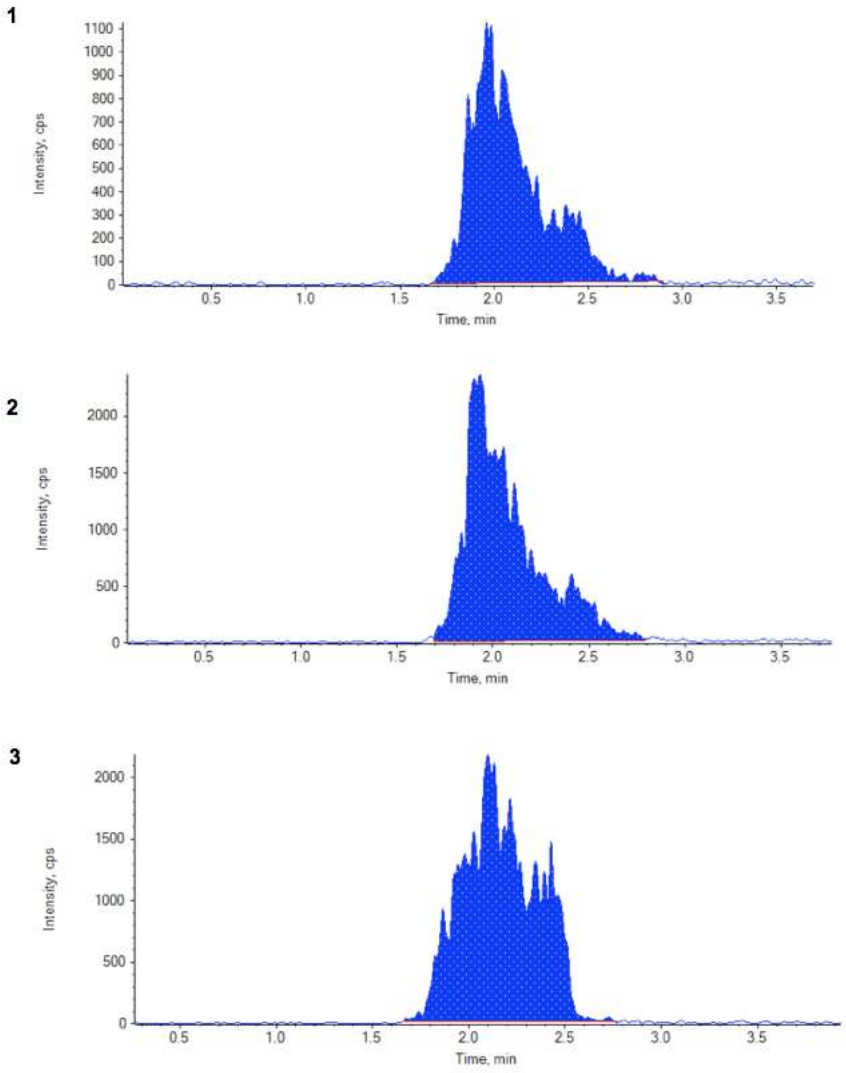


Figure A.8 LC-MS chromatogram 8mg/L gentamicin agar extraction. Image displaying the LC-MS chromatogram output of the 8mg/L Gentamicin agar extract for all three Gentamicin components: 1 (464.292/322.300), 2 (450.317/322.200) and 3 (478.325/322.200) presenting LC-MS intensity over time.

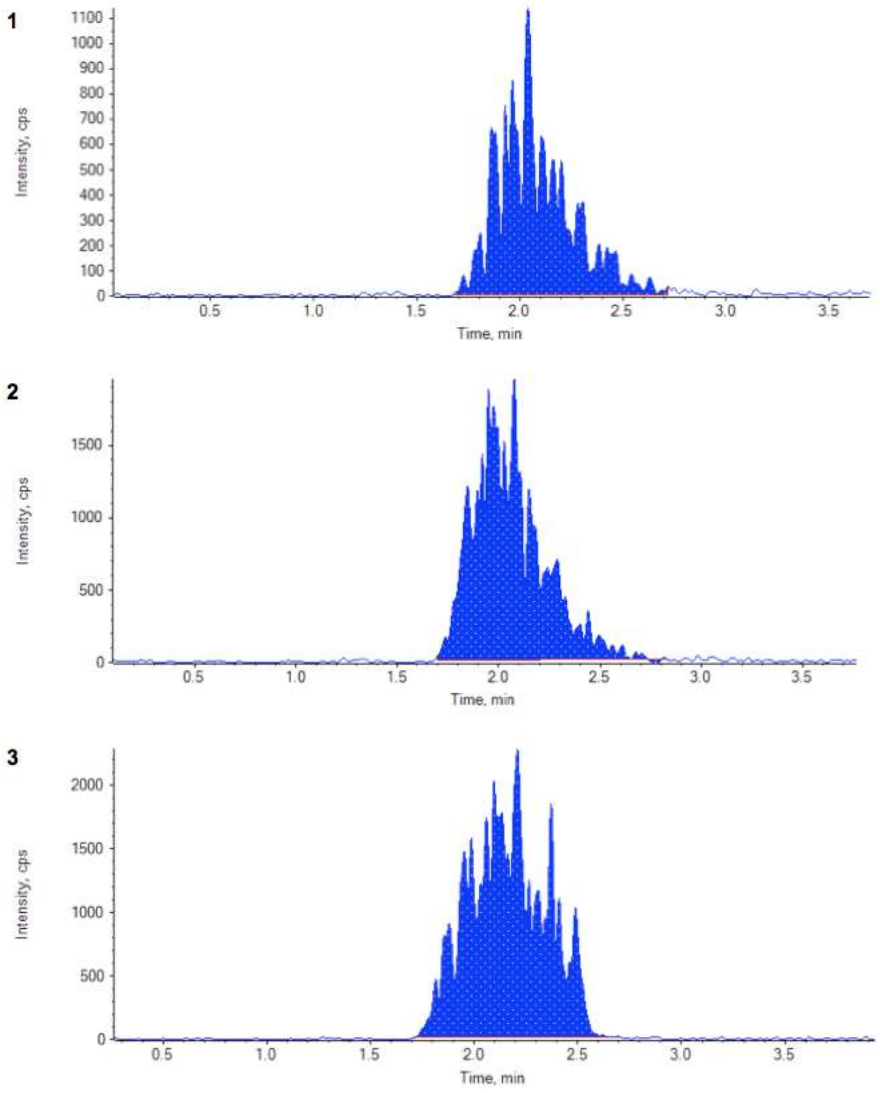


Figure A.9 LC-MS chromatogram 10mg/L gentamicin agar extraction. Image displaying the LC-MS chromatogram output of the 10mg/L Gentamicin agar extract for all three Gentamicin components: 1 (464.292/322.300), 2 (450.317/322.200) and 3 (478.325/322.200) presenting LC-MS intensity over time.

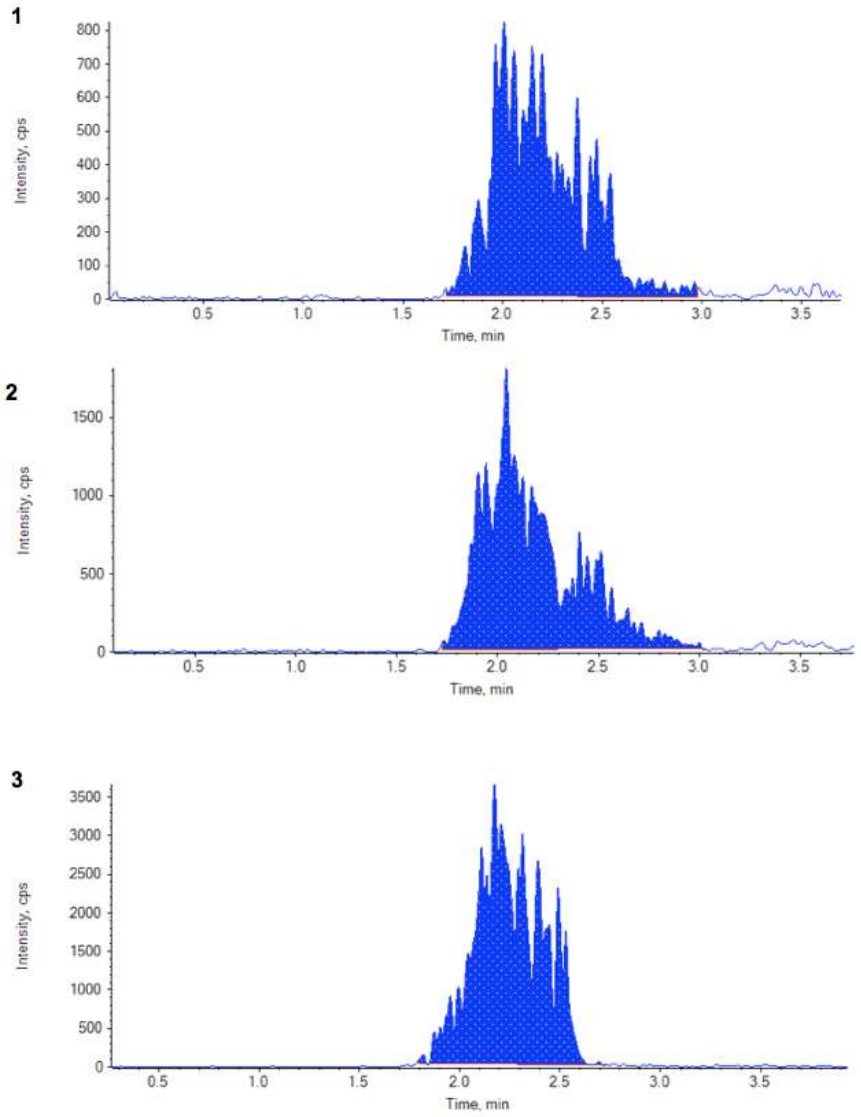


Figure A.10 LC-MS chromatogram 20mg/L gentamicin agar extraction. Image displaying the LC-MS chromatogram output of the 20mg/L Gentamicin agar extract for all three Gentamicin components: 1 (464.292/322.300), 2 (450.317/322.200) and 3 (478.325/322.200) presenting LC-MS intensity over time.

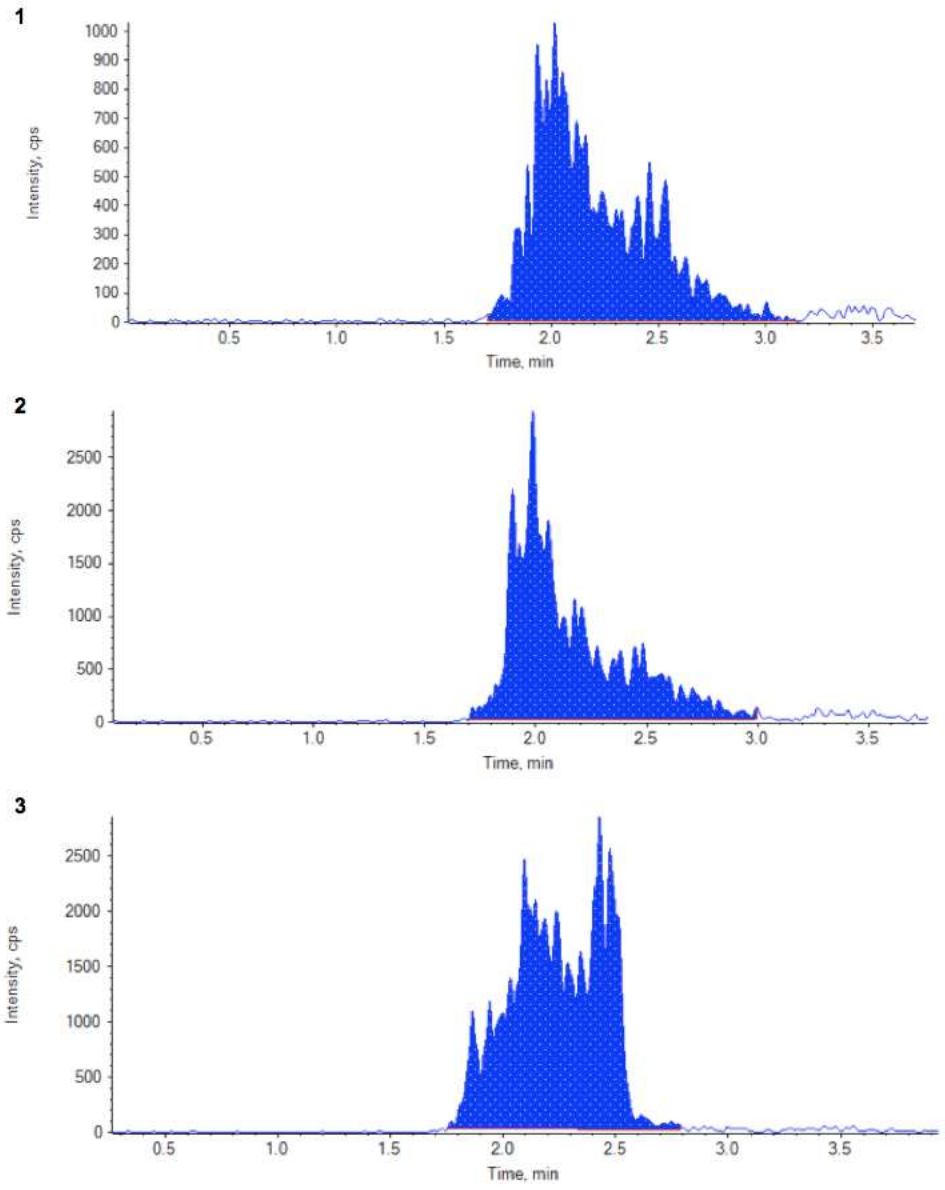


Figure A.11 LC-MS chromatogram 40mg/L gentamicin agar extraction. Image displaying the LC-MS chromatogram output of the 40mg/L Gentamicin agar extract for all three Gentamicin components: 1 (464.292/322.300), 2 (450.317/322.200) and 3 (478.325/322.200) presenting LC-MS intensity over time.

Table A.1 MAUVE SNP export. Table displaying the nodal information of all SNPs within the coding regions of the sensitive and resistant genomes obtained using Mauve SNP export.

SNP pattern	Sensitive Contig	Sensitive Position Contig	Sensitive Genome Wide Position	Resistant Contig	Resistant Position in Contig	Resistant Genome Wide Position
ac	NODE_14	1746	1062759	NODE_16	1746	1060305
tc	NODE_14	1786	1062799	NODE_16	1786	1060345
tc	NODE_14	99500	1160513	NODE_16	99500	1158059
ga	NODE_14	112605	1173618	NODE_62	222	1171144
ct	NODE_14	112609	1173622	NODE_62	226	1171148
ga	NODE_14	112611	1173624	NODE_62	228	1171150
ac	NODE_14	112613	1173626	NODE_62	230	1171152
tg	NODE_14	112614	1173627	NODE_62	231	1171153
tc	NODE_14	112615	1173628	NODE_62	232	1171154
at	NODE_14	112619	1173632	NODE_62	236	1171158
ga	NODE_14	112620	1173633	NODE_62	237	1171159
ct	NODE_14	112621	1173634	NODE_62	238	1171160
ca	NODE_14	112623	1173636	NODE_62	240	1171162
ca	NODE_14	112624	1173637	NODE_62	241	1171163
gt	NODE_14	112625	1173638	NODE_62	242	1171164
tc	NODE_14	112628	1173641	NODE_62	244	1171166
ta	NODE_14	112629	1173642	NODE_62	245	1171167
ag	NODE_14	112630	1173643	NODE_62	246	1171168
ac	NODE_14	112634	1173647	NODE_62	250	1171172

ca	NODE_14	112636	1173649	NODE_62	252	1171174
tg	NODE_14	112638	1173651	NODE_62	254	1171176
tg	NODE_14	112639	1173652	NODE_62	255	1171177
tg	NODE_14	112640	1173653	NODE_62	256	1171178
ag	NODE_14	112641	1173654	NODE_62	257	1171179
ag	NODE_14	112642	1173655	NODE_62	258	1171180
ag	NODE_14	112643	1173656	NODE_62	259	1171181
ct	NODE_14	112645	1173658	NODE_62	261	1171183
tc	NODE_14	112649	1173662	NODE_62	265	1171187
ac	NODE_14	112652	1173665	NODE_62	268	1171190
ac	NODE_14	112655	1173668	NODE_62	271	1171193
ct	NODE_14	112657	1173670	NODE_62	273	1171195
ac	NODE_14	112658	1173671	NODE_62	274	1171196
gt	NODE_14	112661	1173674	NODE_62	277	1171199
gt	NODE_14	112662	1173675	NODE_62	278	1171200
gt	NODE_14	112665	1173678	NODE_62	280	1171202
ga	NODE_14	112666	1173679	NODE_62	281	1171203
ta	NODE_14	112669	1173682	NODE_62	284	1171206
ac	NODE_14	112670	1173683	NODE_62	285	1171207
ga	NODE_14	112686	1173699	NODE_62	299	1171221
cg	NODE_14	112698	1173711	NODE_62	303	1171225
ag	NODE_14	112702	1173715	NODE_62	307	1171229
at	NODE_14	112704	1173717	NODE_62	309	1171231

ta	NODE_14	112705	1173718	NODE_62	310	1171232
gt	NODE_14	112707	1173720	NODE_62	312	1171234
ca	NODE_14	112708	1173721	NODE_62	313	1171235
ac	NODE_14	112709	1173722	NODE_62	314	1171236
ta	NODE_14	112711	1173724	NODE_62	316	1171238
ag	NODE_14	112717	1173730	NODE_62	322	1171244
tg	NODE_14	112718	1173731	NODE_62	323	1171245
ct	NODE_14	112721	1173734	NODE_62	326	1171248
tc	NODE_14	112723	1173736	NODE_62	328	1171250
gt	NODE_14	112724	1173737	NODE_62	329	1171251
ct	NODE_14	112729	1173742	NODE_62	336	1171258
gt	NODE_14	112731	1173744	NODE_62	338	1171260
ct	NODE_14	112732	1173745	NODE_62	339	1171261
ca	NODE_14	112733	1173746	NODE_62	340	1171262
ta	NODE_14	112734	1173747	NODE_62	341	1171263
ca	NODE_14	112739	1173752	NODE_62	346	1171268
at	NODE_14	112744	1173757	NODE_62	350	1171272
ga	NODE_14	112745	1173758	NODE_62	351	1171273
ct	NODE_14	112746	1173759	NODE_62	352	1171274
ac	NODE_14	112748	1173761	NODE_62	354	1171276
ag	NODE_14	112750	1173763	NODE_62	356	1171278
ag	NODE_14	112753	1173766	NODE_62	359	1171281
ac	NODE_14	112754	1173767	NODE_62	360	1171282

at	NODE_14	112755	1173768	NODE_62	361	1171283
ga	NODE_14	112758	1173771	NODE_62	364	1171286
ta	NODE_14	112759	1173772	NODE_62	365	1171287
gt	NODE_14	112760	1173773	NODE_62	366	1171288
at	NODE_14	112766	1173779	NODE_62	377	1171299
gt	NODE_14	112769	1173782	NODE_62	380	1171302
ct	NODE_14	112770	1173783	NODE_62	381	1171303
at	NODE_14	112772	1173785	NODE_62	383	1171305
cg	NODE_14	112773	1173786	NODE_62	384	1171306
ag	NODE_14	112774	1173787	NODE_62	385	1171307
at	NODE_14	112776	1173789	NODE_62	387	1171309
at	NODE_14	112780	1173793	NODE_62	391	1171313
gt	NODE_14	112782	1173795	NODE_62	393	1171315
gt	NODE_14	112787	1173800	NODE_62	402	1171324
tc	NODE_14	112788	1173801	NODE_62	403	1171325
tc	NODE_14	112789	1173802	NODE_62	404	1171326
ag	NODE_14	112790	1173803	NODE_62	405	1171327
ga	NODE_14	112793	1173806	NODE_62	408	1171330
tc	NODE_14	112795	1173808	NODE_62	413	1171335
ta	NODE_14	112798	1173811	NODE_62	416	1171338
ag	NODE_14	112800	1173813	NODE_62	418	1171340
ca	NODE_14	112801	1173814	NODE_62	419	1171341
ta	NODE_14	112802	1173815	NODE_62	420	1171342

ct	NODE_14	112804	1173817	NODE_62	422	1171344
ga	NODE_14	112805	1173818	NODE_62	423	1171345
tc	NODE_14	112807	1173820	NODE_62	425	1171347
tc	NODE_14	112810	1173823	NODE_62	428	1171350
ga	NODE_14	112811	1173824	NODE_62	429	1171351
ag	NODE_14	112812	1173825	NODE_62	430	1171352
gt	NODE_14	112818	1173831	NODE_62	436	1171358
ac	NODE_14	112821	1173834	NODE_62	439	1171361
ga	NODE_20	59648	1234726	NODE_20	59759	1232510
tg	NODE_20	59719	1234797	NODE_20	59830	1232581
tc	NODE_20	59848	1234926	NODE_20	59959	1232710
tg	NODE_20	59849	1234927	NODE_20	59960	1232711
ta	NODE_20	59872	1234950	NODE_20	59983	1232734
ag	NODE_20	59912	1234990	NODE_20	60023	1232774
ag	NODE_20	59919	1234997	NODE_20	60030	1232781
ct	NODE_20	59928	1235006	NODE_20	60039	1232790
ga	NODE_20	59944	1235022	NODE_20	60055	1232806
ga	NODE_20	59950	1235028	NODE_20	60061	1232812
tg	NODE_20	59972	1235050	NODE_20	60083	1232834
at	NODE_20	60010	1235088	NODE_20	60121	1232872
ag	NODE_20	60446	1235524	NODE_20	60557	1233308
gt	NODE_15	620	3406523	NODE_15	620	3404619
tg	NODE_2	113054	4232928	NODE_2	113254	4229015
gc	NODE_55	91	4562477	NODE_55	91	4558564

gc	NODE_55	1277	4563663	NODE_55	1277	4559750
ga	NODE_55	1278	4563664	NODE_55	1278	4559751
ct	NODE_55	1285	4563671	NODE_55	1285	4559758
ct	NODE_55	1285	4563671	NODE_55	1285	4559758
gt	NODE_57	31	4565444	NODE_58	224	4561724
at	NODE_57	90	4565503	NODE_58	283	4561783
at	NODE_57	1891	4567304	NODE_58	2084	4563584
tc	NODE_57	1893	4567306	NODE_58	2086	4563586
tc	NODE_57	1894	4567307	NODE_58	2087	4563587
ac	NODE_57	1898	4567311	NODE_58	2091	4563591
ta	NODE_57	1907	4567320	NODE_58	2100	4563600
tc	NODE_57	1909	4567322	NODE_58	2102	4563602
ta	NODE_57	1910	4567323	NODE_58	2103	4563603
ga	NODE_59	680	4568505	NODE_59	680	4564289
tg	NODE_59	975	4568800	NODE_59	975	4564584
tg	NODE_61	1	4569899	NODE_60	1	4565683
cg	NODE_61	2	4569900	NODE_60	2	4565684
ct	NODE_61	3	4569901	NODE_60	3	4565685
ac	NODE_61	8	4569906	NODE_60	8	4565690
gc	NODE_61	9	4569907	NODE_60	9	4565691
cg	NODE_61	10	4569908	NODE_60	10	4565692
ga	NODE_61	21	4569919	NODE_60	15	4565697
ga	NODE_61	23	4569921	NODE_60	17	4565699

ta	NODE_61	25	4569923	NODE_60	19	4565701
tc	NODE_61	30	4569928	NODE_60	25	4565707
ag	NODE_61	31	4569929	NODE_60	26	4565708
tc	NODE_61	33	4569931	NODE_60	28	4565710
cg	NODE_61	38	4569936	NODE_60	33	4565715
tc	NODE_61	39	4569937	NODE_60	34	4565716
tc	NODE_61	44	4569942	NODE_60	39	4565721
cg	NODE_70	319	4580401	NODE_63	1375	4569052
ca	NODE_70	345	4580427	NODE_63	1401	4569078

Table A.2 CD-Hit export non-identical proteins in the sensitive strain. The CD-Hit analysis export displaying the protein sequences of the non-identical proteins within the sensitive strain of *E. coli* MG1655.

CD-Hit Non-Identical Proteins Sensitive Strain
>sensitive_hypothetical protein MDEQWGYVGAKSRQRWLFYAYDRLRKTVVAHVFGERTMATLGRLMSLLSPFDV VIWMTDGWPLYESRLKGKLHVISKRYTQRIERYNLNLRQHLARLGRKSLSFSKSV ELHDKVIGHYLNIKHYQ
>sensitive_hypothetical protein MASVSISCPSCSATDGVVRNGKSTAGHQRYLCSHCRKTWQLQFTYTASQPGTH QKIIDMAMNGVGCRATARIMGVGLNTIFRHLKNSGRSR
> sensitive_hypothetical protein MWPDNRIARDAHYLYRYDRHGRLTEKTDLIPEGVIRTDDERTHRYHYDSQHRLVH YTRTQYAEPLVESRYLYDPLGRRVAKRVWRRERDLTGWMSLSRKPQVTWYGWD GDRLTTIQNDRTRIQTIYQPGSFTPLIRVETATGELAKTQRRSLADTLQQSGGEDG GSVVFPPVLVQMLDRLESEILADRVSEESRRWLASCGLTVEQMQNQMDPVYTPA RKIHLYHCDHRGLPLALVSTEGATEWCAEYDEWGNLLNEENPHQLQQLIRLPGQ QYDEESGLYYNRHRYYDPLQGRYITQDPIGLKGGWNFYQYPLNPVQYIDSMGLA SKYGHLNNGGYGARPNKPPTPDPSKLPDIAKQLRLPYPIDQASSAPNVFKTFFRAL SPYDYTLYCRKWVKPNLTCTPQDDSQYPGMDTKTASDYLPQTNWPTTQLPPGYT CAEPYLFPDINKPDGPATAGIDDLGEILAKMKQRTSRGIRK
>sensitive_putative deoxyribonuclease RhsB MVAHRHTGRPEICYRYDSDGRVTEQLNPAGLSYTYQYEKDRITITDSLNRREVLHT QGEGGLKRVVKKEHADGSVTQSQFDAVGRLRAQTDAAGRTTEYSPDVVTGLITRI TTPDGRASAFYYNHHSQLTSATGPDGLEIRREYDEWGRLIQETAPDGDITRYRYD NPHSDLPCATEDATGSRKTMTWSRYGQLLSFTDCSGYVTRYDHDRFGQVTAVH REEGLSQYRAYDSRGQLIAVKDTQGHETRYEYNAAGDLTTVIAPDGSRNGTQYD AWGKAICTTQGGLTRSMEYDAAGRVIRLTSENGSHTTFRYDVLDRLIQETGFDGR TQRYHHDLTGKLIRSEDEGLVTHWHYDEADRLTHRTVNGETAERWQYDERGWL TDISHISEGHRVTVHYGYDSKGRLASEHLTVHHPQTNELLWQHETRHAYNAQGLA NRCIPDSLPAVEWLTYGSGWLSGMKLGDTPLVEYTRDRLHRETLRSFGRYELTTA YTPAGQLQSQHLNSLLSDRDYTWNDNGELIRISSPRQTRSYSYSTTGRLTGVHTT AANLDIRIPYTTDPAGNRLPDPELHPDSALSMWPDNRIARDAHYLYRYDRHGRLT EKTDLIPEGVIRTDDERTHRYHYDSQHRLVHYTRTQYAEPLVESRYLYDPLGRRV AKRVWRRERDLTGWMSLSRKPQVTWYGWDGDRLTTIQNDRTRIQTIYQPGSFTP LIRVETATGELAKTQRRSLADTLQQSGGEDGGSVVFPPVLVQMLDRLESEILADR VSEESRRWLASCGLTVAQMQSQMDPVYTPARKIHLYHCDHRGLPLALISTEGTTA WYAEYDEWGNLLNEENPHQLQQLIRLPGQQYDEESGLYYNRHRYYDPLQGRYIT QDPIGLKGGWNFYQYPLNPISNIDPLGLETLKCIKPLHSMGGTGERSGPDIWGNPF YHQYLCVPDGKGDYTCGGQDQRGESKGDGLWGPGKASNDTKEAAGRCDLVET DNSCVENCLKGKFKEVRPRYSVLPDIFTPINLGLFKNCQDWSNDSLETCKMKCSG NNIGRFIRFVFTGVM

>sensitive_Putative deoxyribonuclease RhsC
MVAHRHTGRPEIRYRYDSDGRVTEQLNPAGLSYTYQYEKDRITITDSLDRREVLHT
QGEAGLKRVVKKEHADGSVTQSQFDAVGRLRAQTDAAGRTTEYSPDVVTGLITRI
TTPDGRASAFYYNHHNQLTSATGPDGLELRREYDELGRLIQETAPDGDITRYRYD
NPHSDLPCATEDATGSRKTMTWSRYGQLLSFTDCSGYVTRYDHDRFGQMTAVH
REEGLSQYRAYDSRGQLIAVKDTQGHETRYEYNIAGDLTAVIAPDGSRNGTQYDA
WGKAVRTTQGGLTRSMEYDAAGRVIRLTSENGSHTTFRYDVLDRLIQETGFDGRT
QRYHHDLTGKLIRSEDEGLVTHWHYDEADRLTHRTVKGETAERWQYDERGWLT
DISHISEGHRVAVHYRYDEKGRLTGERQTVHHPQTEALLWQHETRHAYNAQGLA
NRCIPDSLPAVEWLTYGSGYLAGMKLGDTPLVEYTRDRLHRETLRSFGRYELTTA
YTPAGQLQSQHLNSLLSDRDYTWNDNGELIRISSPRQTRSYSYSTTGRLTGVHTT
AANLDIRIPYATDPAGNRLPDPELHPDSTLSMWPDNRIARDAHYLYRYDRHGRLT
EKTDLIPEGVIRTDDERTHRYHYDSQHRLVHYTRTQYEEPLVESRYLYDPLGRRV
AKRVWRRERDLTGWMSLSRKPQVTWYGWDGDRLTTIQNDRTRIQTIYQPGSFTP
LIRVETATGELAKTQRRSLADALQQSGGEDGGSVVFPPVLVQMLDRLESEILADR
VSEESRRWLASCGLTVEQMQNQMDPVYTPARKIHLYHCDHRGLPLALISKEGTTE
WCAEYDEWGNLLNEENPHQLQQLIRLPGQQYDEESGLYYNRHRYYDPLQGRYIT
QDPIGLKGGWNFYQYPLNPVTNTDPLGLEVFPRPFPLPIPWPKSPAQQQADDNA
AKALTKWWNDTASQRIFDSLILNNPGLALDITMIASRGNVADTGITDRVNDIINDRF
WSDGKKPDRCDVLQELIDCGDISAKDAKSTQKAWNCRHSRQSNDKKR
>sensitive_ATP synthase gamma chain
MAGAKEIRSKIASVQNTQKITKAMEMVAASKMRKSQDRMAASRPYAETMRKVIGH
LAHGNLEYKHPYLEDRDVKRVGYLVVSTDRGLCGGLNINLFKKLLAEMKTWTDKG
VQCDLAMIGSKGVSFFNSVGGNVVAQVTGMGDNPSLSELIGPVKVMLQAYDEGR
LDKLYIVSNKFINTMSQVPTISQLLPLPASDDDDLKHKSWDYLYEPDPKALLDTLLR
RYVESQVYQGVVENLASEQAARMVAMKAATDNGGSLIKELQLVYNKARQASITQ
ELTEIVSGAAAV
>sensitive_putative major fimbrial subunit LpfA
MFALAGNKWNTTLPGGNMQFQGVIIAETCRIEAGDKQMTVNMGQISSNRFHAVG
EDSAPVPFVIHLRECSTVVSERVGVAFHGVADGKNPDVLSVGEGPGIATNIGVALF
DDEGNLVPINRPPANWKRLYSGSTSLHFIAKYRATGRRVTGGIANAQAWFSLTYQ
>sensitive_hypothetical protein
MHRIDTKTAQKDKFGAGKNGFTRGNPQTGTPATDLDDDYFDMLQEELCSVVEAS
GASLEKGRHDQLLTALRALLLSRKNPFGDIKSDGTVQTALENLGLGEGSALPVGV
PVPWPSATPPTGWLKCNGAAFSAEEYPELAKAYPTNKLPDLRGEFIRGWDDGRG
IDTGRSILSIQGYATEDHAHGLPSRSTIVTDATINFYFDEIWVNSGTDIIKRGNTNDA
GLPAPDYGTFKTYKQSVDGLGAAASETRPRNIAFNYIVRAA
>sensitive_Small toxic polypeptide LdrD
MTLAQFAMTFWHDLAAPILAGIITAAIVSWWRNRK
>sensitive_hypothetical protein
MASVSISCPSCSATDGVVRNGKSTAGHQRYLCSHCRKTWQLQFTYTASQPGTH
QKIIDMAMNGVGCRATARIMGVGLNTIFRHLKNSGRSR
>sensitive_hypothetical protein
MDEQWGYVGAKSRQRWLFYAYDRLRKTVVAHVFGERTMATLGRLMSLLSPFDV
VIWMTDGWPLYESRLKGKLHVISKRYTQRIERHNLNLRQHLARLGRKSLSFSKSV
EQHDKVIGHYLNIKHYQ

>sensitive_hypothetical protein MDEQWGYVGAKSRQRWLFYAYDSLRKTVVAHVFGERTMATLGRLMSLLSPFDV VIWMTDGWPLYESRLKGKLHVISKRYTQRIERHNLNLRQHLARLGRKSLSFSKSV ELHDKVIGHYLNIKHYQ
>sensitive_HTH-type transcriptional regulator ArcR MKLFSLLSETQCQHIEYNMALTSKLAKERVTKILRYLCQTVGYDHDEFYEIKHFMTI QLLSDMAGISRETTSHIINELREEKILFKNSKNWLVSKDL

Table A.3 CD-Hit export non-identical proteins in the resistant strain. The CD-Hit analysis export displaying the protein sequences of the non-identical proteins within the resistant strain of *E. coli* MG1655.

CD-Hit Non-Identical Proteins Resistant Strain
>resistant_Putative deoxyribonuclease RhsC MVFPPVLVQMLDRLESEILADRVSEESRRWLASCGLTVEQMQNQMDPVYTPARK IHLYHCDHRGLPLALVSTEGATEWCAEYDEWGNLLNEENPHQLQQLIRLPGQQY DEESGLYYNRHRYYDPLQGRYITQDPIGLKGGWNFYQYPLNPVQYIDSMGLASKY GHLNNGGYGARPNKPPTPDPSKLPDIAKQLRLPYPIDQASSAPNVFKTFFRALSPY DYTLYCRKWVKPNLTCTPQDDSQYPGMDTKTASDYLPQTNWPTTQLPPGYTCAE PYLFPDINKPDGPATAGIDDLGEILAKMKQRTSRGIRK
>resistant_Ethanolamine utilization protein EutM MNGPVLDPDKEESTMEALGMIETRGLVALIEASDAMVKAARVKLVGVKQIGGGLC TAMVRGDVAACKAATDAGAAAAQRIGELVSVHVIPRPHGDLEEVFPIGLKGDSSN L
>resistant_Putative deoxyribonuclease RhsC MVFPPVLVQMLDRLESEILADRVSEESRRWLASCGLTVEQMQNQMDPVYTPARK IHLYHCDHRGLPLALISKEGTTEWCAEYDEWGNLLNEENPHQLQQLIRLPGQQYD EESGLYYNRHRYYDPLQGRYITQDPIGLKGGWNFYQYPLNPVTNTDPLGLEVFPR PFPLPIPWPKSPAQQQADDNAAKALTKWWNDTASQRIFDSLILNNPGLALDITMIA SRGNVADTGITDRVNDIINDRFWSDGKKPDRCDVLQELIDCGDISAKDAKSTQKA WNCRHSRQSNDKKR
>resistant_ATP synthase gamma chain MAGAKEIRSKIASVQNTQKITKAMEMVAASKMRKSQDRMAASRPYAETMRKVIGH LAHGNLEYKHPYLEDRDVKRVGYLVVSTDRGLCGGLNINLFKKLLAEMKTWTDKG VQCDLAMIGSKGVSFFNSVGGNVVAQVTGMGDNPSLSELIGPVKVMLQAYDEGR LDKLYIVSNKFINTMSQVPTISQLLPLPASDDDDLKHKSWDYLYEPDPKALLDTLLR RYVESQVYQGVVENLASEQAARYGGDESRDRQWRQPD
>resistant_Maltodextrin glucosidase MMLNAWHLPVPPFVKQSKDQLLITLWLTGEDPPQRIMLRTEHDNEEMSVPMHKQ RSQPQPGVTAWRAAIDLSSGQPRRRYSFKLLWHDRQRWFTPQGFSRMPPARLE QFAVDVPDIGPQWAADQIFYQIFPDRFARSLPREAEQDHVYYHHAAGQEIILRDW DEPVTAQAGGSTFYGGDLDGISEKLPYLKKLGVTALYLNPVFKAPSVHKYDTEDY RHVDPQFGGDGALLRLRHNTQQLGMRLVLDGVFNHSGDSHAWFDRHNRGTGG ACHNPESPWRDWYSFSDDGTALDWLGYASLPKLDYQSESLVNEIYRGEDSIVRH WLKAPWNMDGWRLDVVHMLGEAGGARNNMQHVAGITEAAKETQPEAYIVGEHF GDARQWLQADVEDAAMNYRGFTFPLWGFLANTDISYDPQQIDAQTCMAWMDNY RAGLSHQQQLRMFNQLDSHDTARFKTLLGRDIARLPLAVVWLFTWPGVPCIYYGD EVGLDGKNDPFCRKPFPWQVEKQDTALFALYQRMIALRKKSQALRHGGCQVLYA EDNVVVFVRVLNQQRVLVAINRGEACEVVLPASPFLNAVQWQCKEGHGQLTDGIL ALPAISATVWMN

>resistant_Elongation factor G
MARTTPIARYRNIGISAHIDAGKTTTTERILFYTGVNHKIGEVHDGAATMDWMEQE
QERGITITSAATTAFWSGMAKQYEPHRINIIDTPGHVDFTIEVERSMRVLDGAVMV
YCAVGGVQPQSETVWRQANKYKVPRIAFVNKMDRMGANFLKVVNQIKTRLGANP
VPLQLAIGAEEHFTGVVDLVKMKAINWNDADQGVTFEYEDIPADMVELANEWHQN
LIESAAEASEELMEKYLGGEELTEAEIKGALRQRVLNNEIILVTCGSAFKNKGVQAM
LDAVIDYLPSPVDVPAINGILDDGKDTPAERHASDDEPFSALAFKIATDPFVGNLTF
FRVYSGVVNSGDTVLNSVKAARERFGRIVQMHANKREEIKEVRAGDIAAAIGLKDV
TTGDTLCDPDAPIILERMEFPEPVISIAVEPKTKADQEKMGLALGRLAKEDPSFRV
WTDEESNQTIIAGMGELHLDIIVDRMKREFNVEANVGKPQVAYRETIRQKVTDVEG
KHAKQSGGRGQYGHVVIDMYPLEPGSNPKGYEFINDIKGGVIPGEYIPAVDKGIQE
QLKAGPLAGYPVVDMGIRLHFGSYHDVDSSELAFKLAASIAFKEGFKKAKTVLLEPI
MKVEVETPEENTGDVIGDLSRRRGMLKGQESEVTGVKIHAEVPLSEMFGYATQLR
SLTKGRASYTMEFLKYDEAPSNVAQAVIEARGK

>resistant_DNA-binding transcriptional repressor YiaJ
MIFYCALSIGRVFSATIKTCPNVHQVHHVVLTIEMSINMQNNEQTEYKTVRGLTRGL
MLLNMLNKLDGGASVGLLAELSGLHRTTVRRLLETLQEEGYVRRSPSDDSFRLTIK
VRQLSEGFRDEQWISALAAPLLGDLLREVVWPTDVSTLDVDAMVVRETTHRFSRL
SFHRAMVGRRLPLLKTASGLTWLAFCPEQDRKELIEMLASRPGDDYQLAREPLKL
EAILARARKEGYGQNYRGWDQEEKIASIAVPLRSEQRVIGCLNLVYMASAMTIEQA
AEKHLPALQRVAKQIEEGVESQAILVAGRRSGMHLR

>resistant_Galactoside O-acetyltransferase
MNMPMTERIRAGKLFTDMCEGLPEKRLRGKTLMYEFNHSHPSEVEKRESLIKEMF
ATVGENAWVEPPVYFSYGSNIHIGRNFYANFNLTIVDDYTVTIGDNVLIAPNVTLSV
TGHPVHHELRKNGEMYSFPITIGNNVWIGSHVVINPGVTIGDNSVIGAGSIVTKDIP
PNVVAAGVPCRVIREINDRDKHYYFKDYKVESSV

>resistant_hypothetical protein
MLYFSGLGLSVSDSANPVHHYGHVQGGYSVPLIITASDITSHQPVSRKISARHFAGI
FQWMTGICTENIPPFNPLTDEDN

>resistant_Putative deoxyribonuclease RhsC
MVFPPVLVQMLDRLESEILADRVSEESRRWLASCGLTVEQMQNQMDPVYTPARK
IHLYHCDHRGLPLALISTEGATAWCAEYDEWGNLLNEENPHQLQQLIRLPGQQYD
EESGLYYNRHRYYDPLQGRYITQDPIGLKGGWNLYGYQLNPISDIDPLGLSMWED
AKSGACTNGLCGTLSAMIGPDKFDSIDSTAYDALNKINSQSICEDKEFAGLICKDNS
GRYFSTAPNRGERKGSYPFNSPCPNGTEKVSAYHTHGADSHGEYWDEIFSGKDE
KIVKSKDNNIKSFYLGTPSGNFKAIDNHGKEITNRKGLPNVCRVHGNM

>resistant_DNA translocase SpoIIIE
MELSKTPHMLISGLTGSGKSYSMYYLMYSLILKGHEVFVIDRKQVLTKFGTVIGND
HVADENPNESEQIFELIERVNNIMLKRQKILKNDDRFKKDIEAGFQNANWNNICLVI
DELGALTQDLAMMKKAERDRFYSALGNIAMKGRNTGVSLMISLQQANAQSFNGT
GIRDQLSFKMVLGNSDRQTREVLFSSQDISDVKLKPGQAFYTKADTRNKPGFLFM
PTFDFELTIPNLEKLIELQNRDKRSFMPRKFIEVEKNP

>resistant_Serine recombinase PinR
MSQIFAYCRISTLDQTTENQRREIESAGFKIKPQQIIEEHISGSAATSERPGFNRLLA
RLKCGDQLIVTKLDRLGCNAMDIRKTVEQLTETGIRVHCLALGGIDLTSPTGKMMM
QVISAVAEFERDLLLERTHSGIVRARGAGKRFGRPPVLNEEQKQVVFERIKSGVSI
SAIAREFKTSRQTILRAKAKLQTPDI

>resistant_hypothetical protein MLPEQCFRLRVMHRLTLTGQPAFLIIPVMHRHPVPFADMADVCQPATFVILPPLCG FTLHGAVREAVCFVIVPVGDQTLILAADKFAGQVVVITLCAAVKAGFLYQPVKYIVT EGGVAAVFTGQADDPSGSIVFHTAR
>resistant_hypothetical protein MEEYYSIKDVANICNKNKSSVSRKLTNLCFEIMNDDFDMHFKKQKGYNNIEQFFFN EYAVKYIISLFYKDLDYNIIKDMPLNQVLKKINTTKINTKLSNIEILIDLVSNPNSDTIDIL NTISNIKSDFNKLNDEINLLKSLENEHKKDLAHMDFWIDKQNEEIDFLKNEILKRLKK D
>resistant_hypothetical protein MSFRLSDNQLKNMLNATHSEIYELMNSQCNDISHFDLEVISNIELPIDYRKNLSLLF ELLKFEKSLIEIYLKLVNVYSQNDDLFNYDLKNNRRELKGQLEKIDLRISELSKGYIQ FLEYENGV
>resistant_hypothetical protein MCEENYEDNIYESEKFINIENVDDIKKEVIIFFIQLGKYFENEQNEAFKNLSDENKMN IFDVLNVIYPNIKKLFDNVNNITVCLLYTSPSPRDRTRSRMPSSA
>resistant_hypothetical protein MLPEQFIRLRMMHRQVFTGEAAFTVITIVHRHPVPFADMADVCQPAAFVILPPLCG FAIHGAVREAVCFVIVPVGDQPLILAADKFAGQVVVITLCAAVKAGFLYQPVEYIVT EGGVAAVFTGQADDPSGSIVFHTARQTALRGTDGFSPRIVLCPVSAAVRGNDGG QVTGGVVFIPRFMALRVFHGN
>resistant_Putative deoxyribonuclease RhsC MVFPPVLVQMLDRLESEILADRVSEESRRWLASCGLTVAQMQSQMDPVYTPARK IHLYHCDHRGLPLALISTEGTTAWYAEYDEWGNLLNEENPHQLQQLIRLPGQQYD EESGLYYNRHRYYDPLQGRYITQDPIGLKGGWNFYQYPLNPISNIDPLGLETLKCI KPLHSMGGTGERSGPDIWGNPFYHQYLCVPDGKGDYTCGGQDQRGESKGDGL WGPGKASNDTKEAAGRCDLVETDNSCVENCLKGKFKEVRPRYSVLPDIFTPINLG LFKNCQDWSNDSLETCKMKCSGNNIGRFIRFVFTGVM
>resistant_hypothetical protein MPGLKGLDNFVKFDVKSFLEKKELVFISCEKNYEKKLDEEGRKVPDINKPRGLKFE INIDKDDTIYSIYDFNSKERKEEKGSNFRKSITVWINNPKLEPEKFDNFAPGDKIFLK GVDEVETMQMTSHLIIVADNIVKVEKNEQQLPKQ
>resistant_hypothetical protein MLQRPADAPPEMQTQPMASVLRDLDIYSVRLEALRGHFWGCMDRCKALALLERP ARSPEVNTHAGREYDLWVWMPQLSVGTLVEP
>resistant_hypothetical protein MRAHQHRDITRLHRAPTQHRFARTRLDQGLVDRSDAGLGRCLTRGIGTPRLVTPA AQHAQGECLGGVAVVQIVLIGAHATGTHPLELDVRLEEGVLTPVVVQRLKRAQH
>resistant_hypothetical protein MSFEVLLASFGLDGNHGLARLGAVVHALDVGGTTVPEAGGFVAILAGARKRLADD DALLTEIGSVLDSLYAHFSGSRKPS
>resistant_hypothetical protein MLLTQPFFKQFTTKPIADARSQTEVITCLSADSRKAVDVMVDKALAAGASEPQPAR DYGFMYQRGFQDLDGHLWEIAHMDGEPG
>resistant_hypothetical protein MPSSRFFAGAYPCCGWRGARTNLLLQVSLDAEQRAPAIVQLGTLLQWQLEPLDY LEQVLAGVPEPVLLDRVLQILASDVPDGARCGVP

>resistant_PTS system fructose-specific EIIB component MSLNLAKPAGVTLDIATVEQAAEKLNSGKLDHKKVMVVLGNTAETLALVEKVPGIS AINYGGLPQKEGARQFGKAIYLTEEEIAHSRALKEKGIRLEMRQVPAHSAELLNDQ L
>resistant_UDP-N-acetylglucosamine 1-carboxyvinyltransferase MALNCVADGVGVINETIFENRFMHVNELLRLGADIQVEGHTAIVRGSEHLSGAPVM ATDLRASASLILAGLMASGETTIDRIYHLDRGYENIEEKLSSLGATIRRVP
>resistant_hypothetical protein MAVTERQQFGDFFQAEAQRLGAADEAQPAQVLLTEDAVAPGRACRRGQQLLAF VVAHRIDRHAGLACQLTNVHGGFLRMLPLIVKP
>resistant_hypothetical protein MNIESIEMISITIIFVILISLATFIIDKMNWWGGVKKTLFIMWNIVCAIFILLGTLVSLIFIV FTAMLFASGD
>resistant_NADP-dependent alcohol dehydrogenase C MRRSLAGSNIGGIAETQEMLNFCAEHGVTPEIELIEPDYINDAYERVLASDVRYRFV IDISL
>resistant_hypothetical protein MHLVDGQAERVAVQVLAGIGEAVGQVDEVRDGFLQHLEVGNPRDLAQRLCCHP HGTVELADGGQPRFQAPVEAFVLEEEARLFNLDQRESLTIHALADPEEGVV
>resistant_Lactose operon repressor MPGEKSVIGYDDTYESSFFYPALSTVSLDLDLQGKEAVRRILASTSGAPHTSSILPA RLVIRHSSGARIEQGKDLQAIAEQLRAIAHRLAP

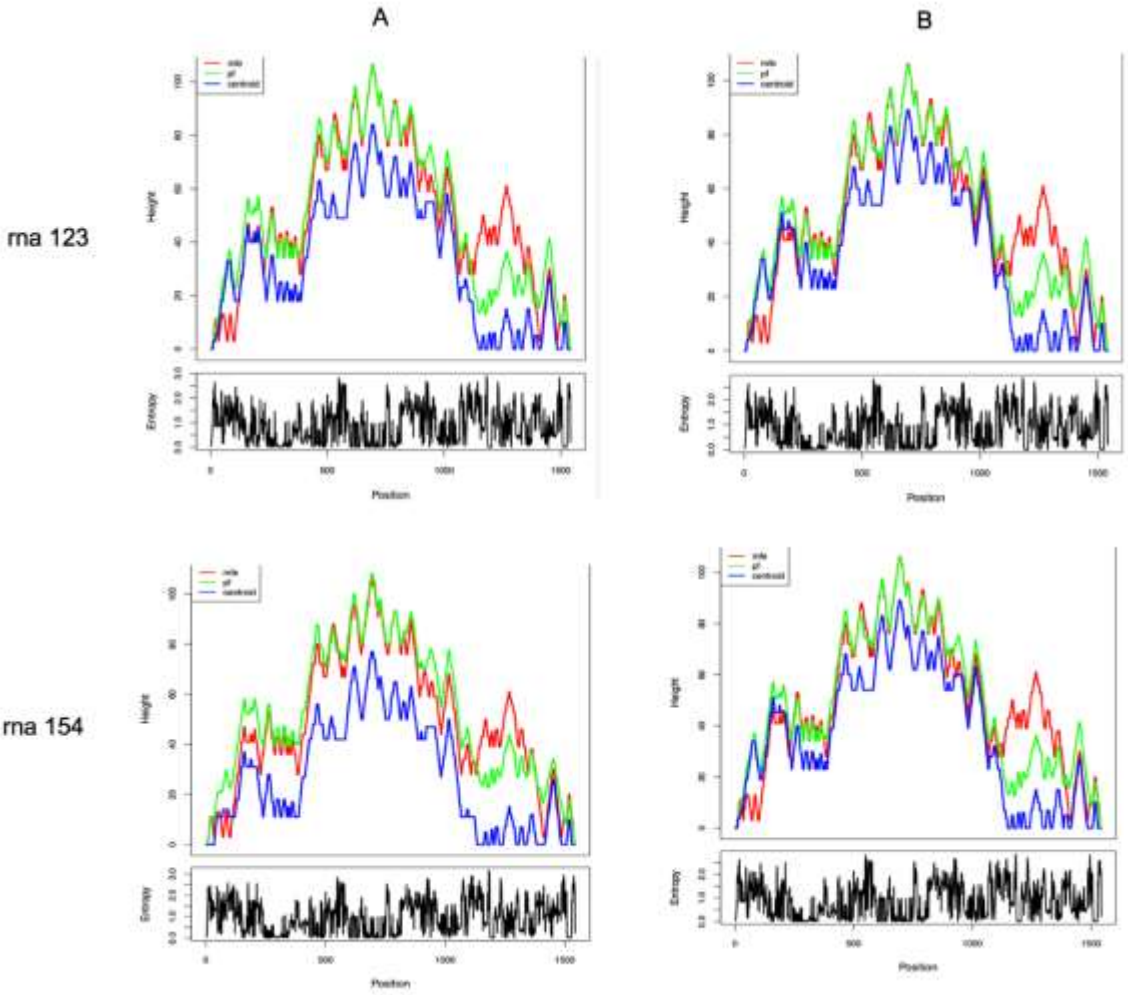


Figure A.12 Mountain plot representation of the rRNA secondary structure analysis using RNA fold. Displaying the base pair probabilities based on the genomic position for the sensitive (A) and resistant (B) 16s rRNA genes rna 123and 154 secondary structures predicted using RNA fold. There were apparent differences in the centroid secondary structure predictions between the sensitive and resistant strains of *E. coli* MG1655.

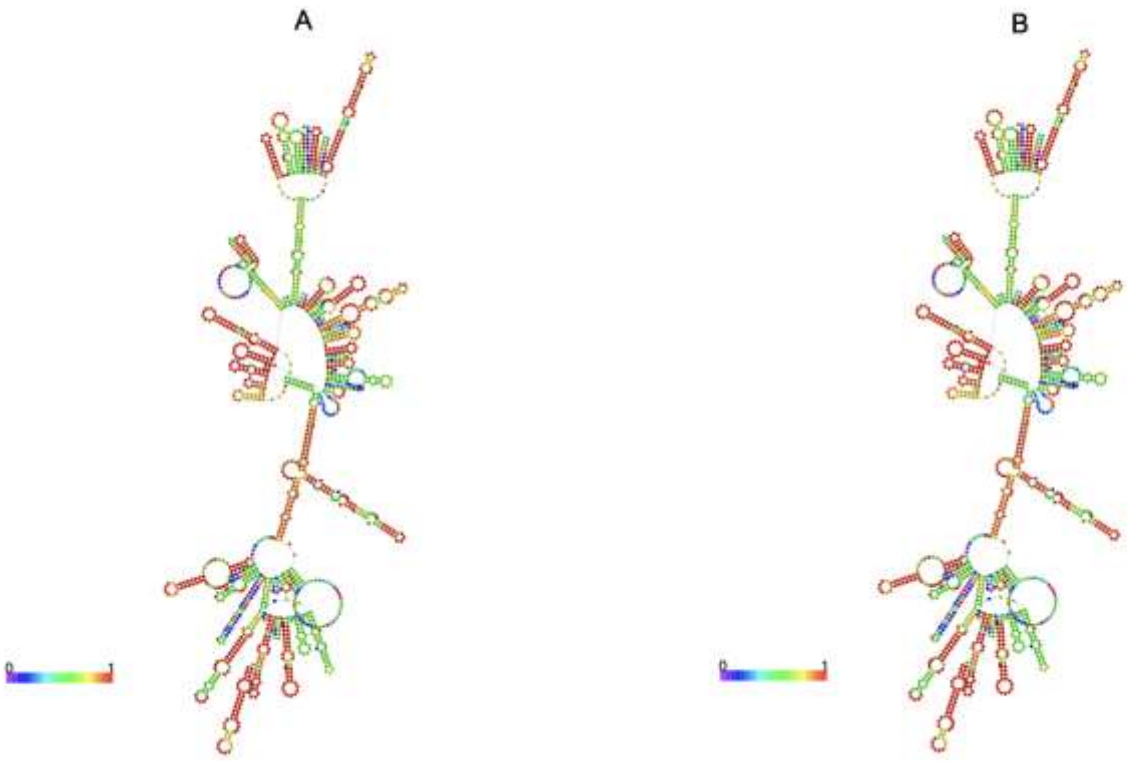


Figure A.13 Secondary structure analysis of rRNA 86 using RNA fold. Displaying the centroid structure image of the rRNA 86 gene in gentamicin sensitive (A) and resistant (B) strains of *E. coli* MG1655.There was no apparent difference in the centroid secondary structure predictions between the sensitive and resistant strains.

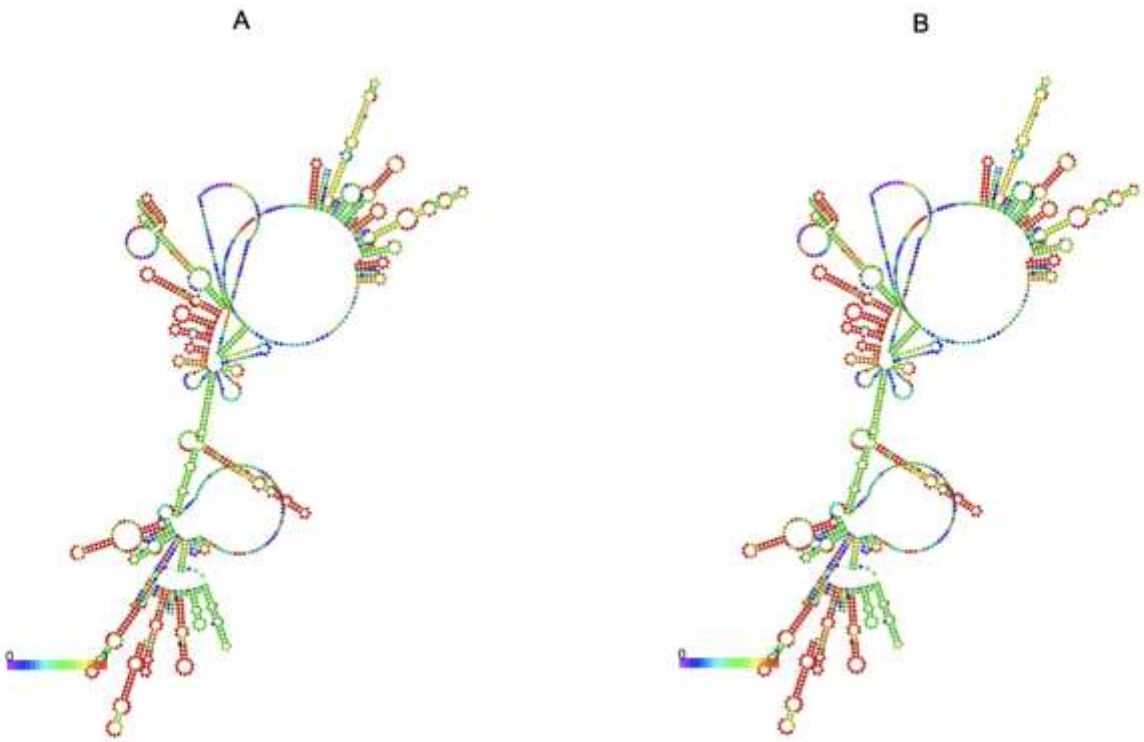


Figure A.14 Secondary structure analysis of rRNA 3 using RNA fold. Displaying the centroid structure image of the rRNA 3 gene in gentamicin sensitive (A) and resistant (B) strains of *E. coli* MG1655. There was no apparent difference in the centroid secondary structure predictions between the sensitive and resistant strains.

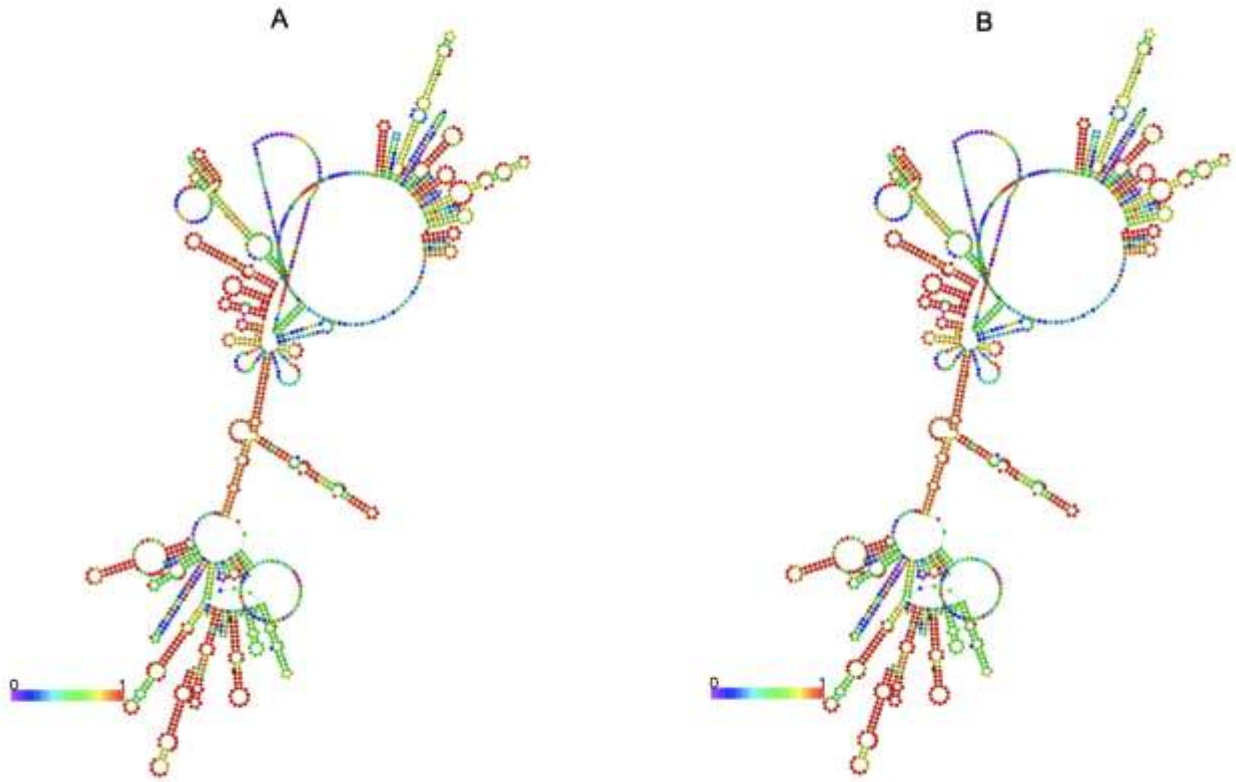


Figure A.15 Secondary structure analysis of rRNA 134 using RNA fold. Displaying the centroid structure image of the rRNA 134 gene in gentamicin sensitive (A) and resistant (B) strains of *E. coli* MG1655. There was no apparent difference in the centroid secondary structure predictions between the sensitive and resistant strains.

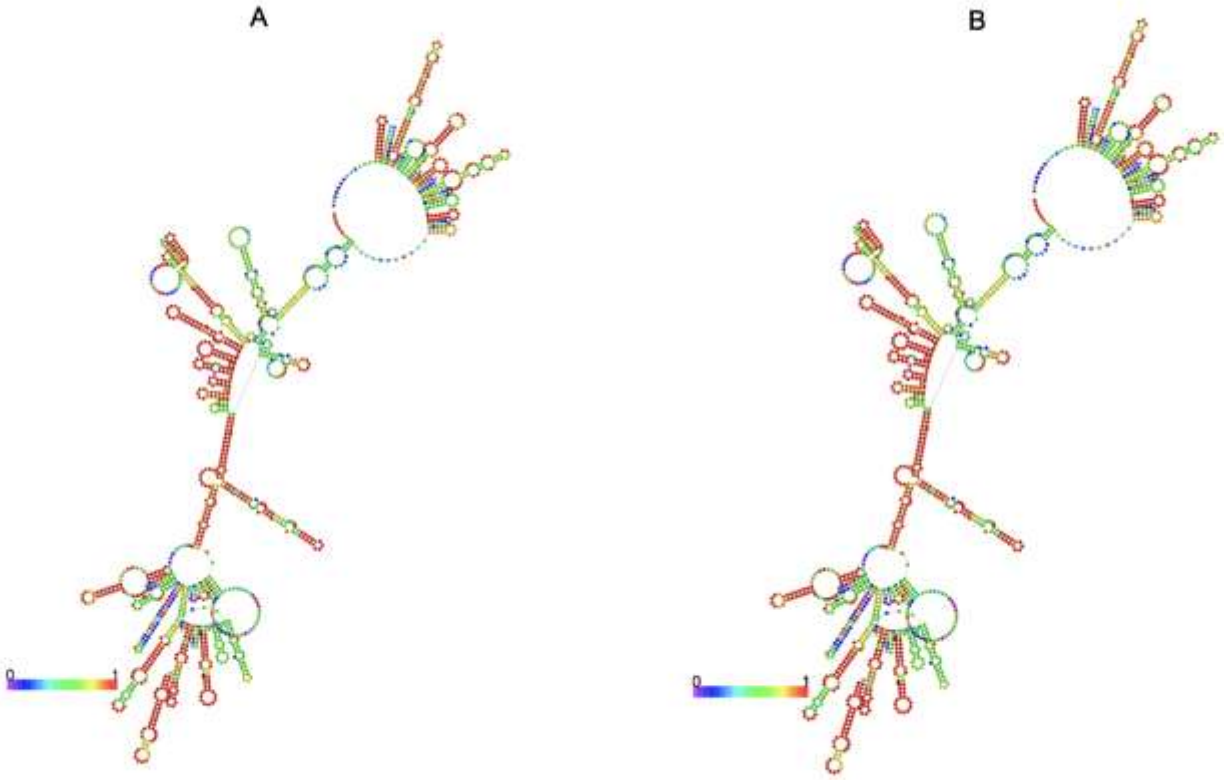


Figure A.16 Secondary structure analysis of rRNA 146 using RNA fold. Displaying the centroid structure image of the rRNA 146 gene in gentamicin sensitive (A) and resistant (B) strains of *E. coli* MG1655. There was no apparent difference in the centroid secondary structure predictions between the sensitive and resistant strains.

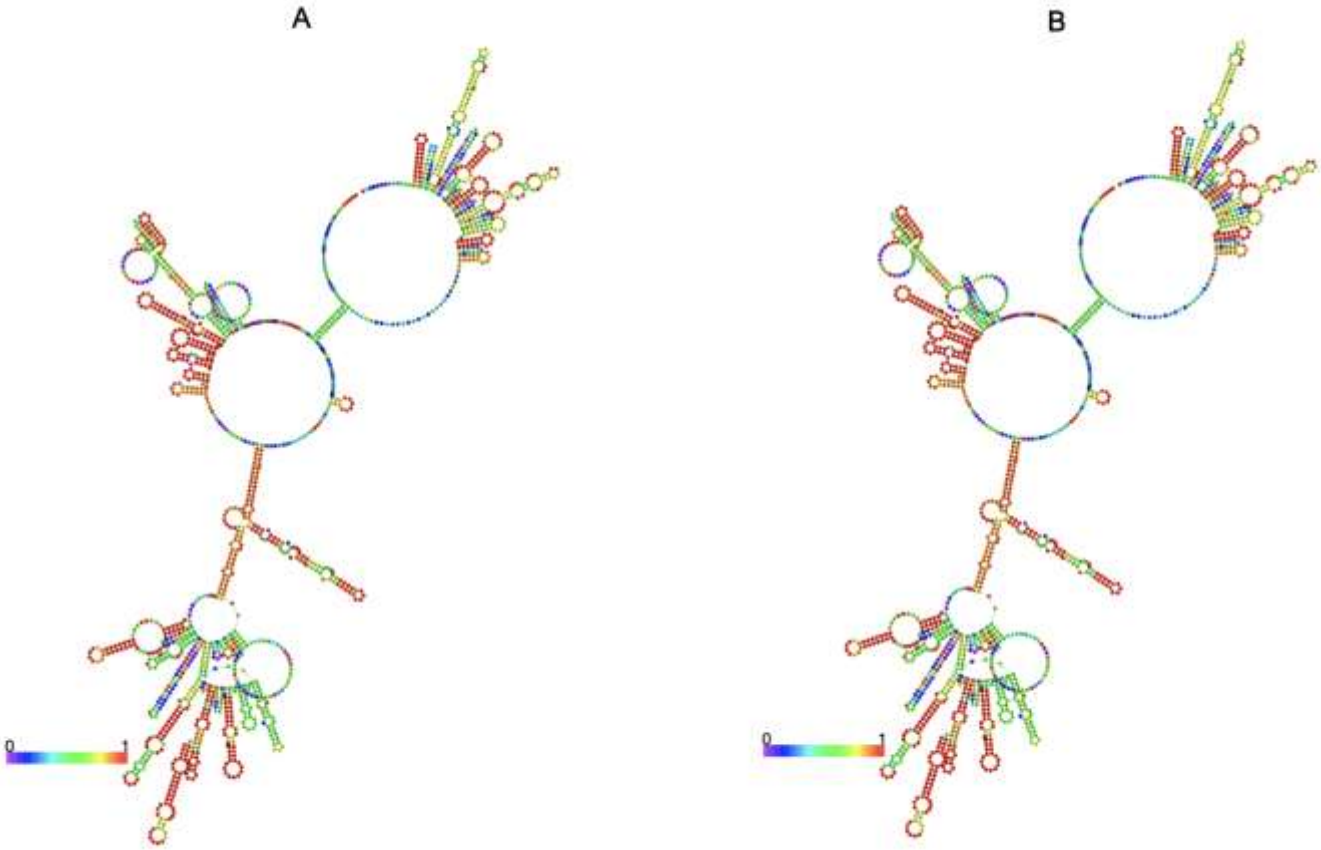


Figure A.17 Secondary structure analysis of rRNA 156 using RNA fold. Displaying the centroid structure image of the rRNA 156 gene in gentamicin sensitive (A) and resistant (B) strains of *E. coli* MG1655. There was no apparent difference in the centroid secondary structure predictions between the sensitive and resistant strains.

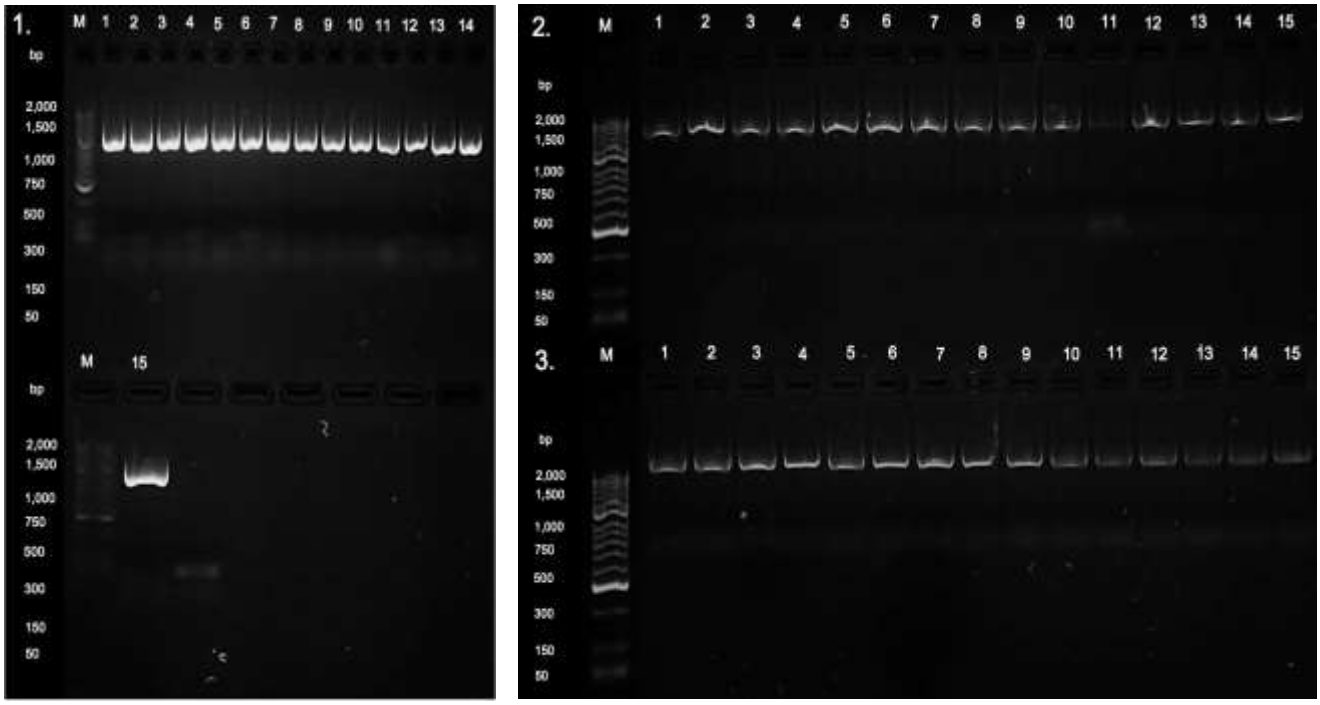


Figure A.18 Gel electrophoresis image of the genomic analysis of genes with mutations. 2% w/v Top Vision Agarose gel electrophoresis image displaying the EF-G (1.), rRNA 123 (2.) and rRNA 154 (3.) PCR reactions. Five colonies were selected at three antibiotic concentrations (0, MIC, X10MIC) for each of the mutated genes (lanes 1-15 respectively). 2Kb DNA ladder (lane M) was used as standard. All colonies were successfully amplified for the three genes selected at the three antimicrobial concentrations.

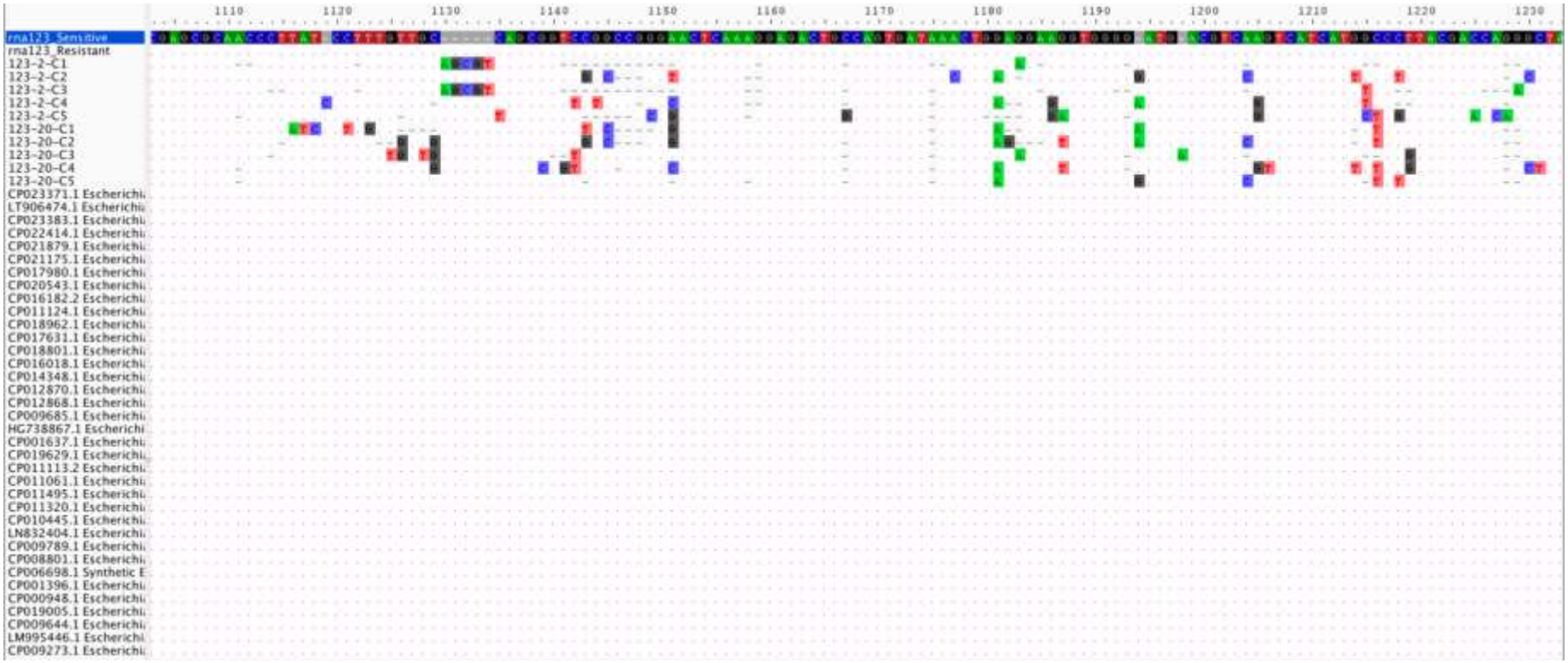


Figure A.19 Resequenced rRNA 123 gene alignment. Aliview alignment image of the resequenced rRNA 123 genes, together with strains of *E. coli* used in the phylogenetic analysis of rna123, displaying the identified variants between the resequenced strain.

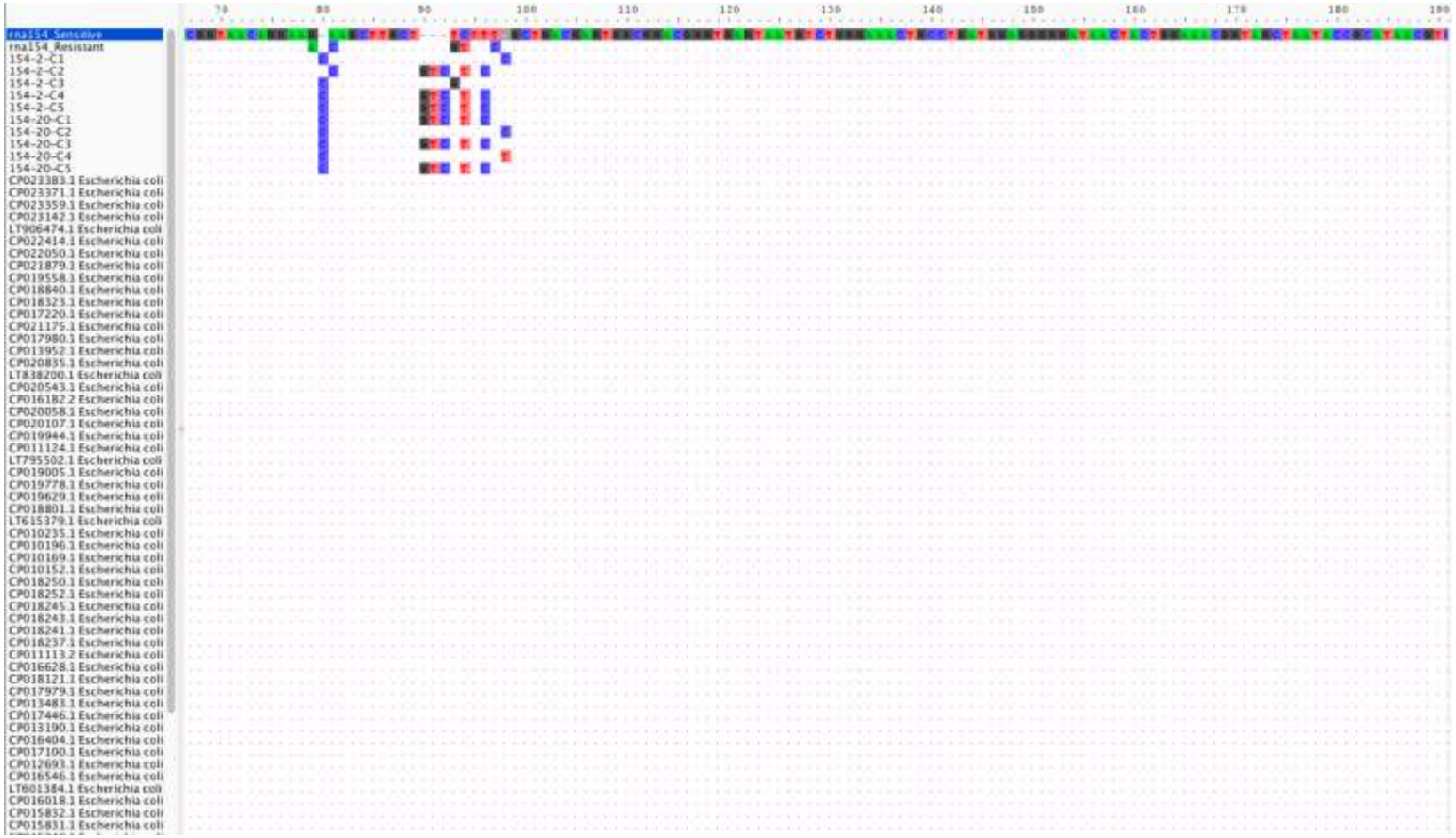


Figure A.20 Resequenced rRNA 154 gene alignment. Aliview alignment image Alignment of the resequenced rRNA154 genes, together with strains of *E. coli* used in the phylogenetic analysis of rna154, displaying the identified variants between the resequenced strains.

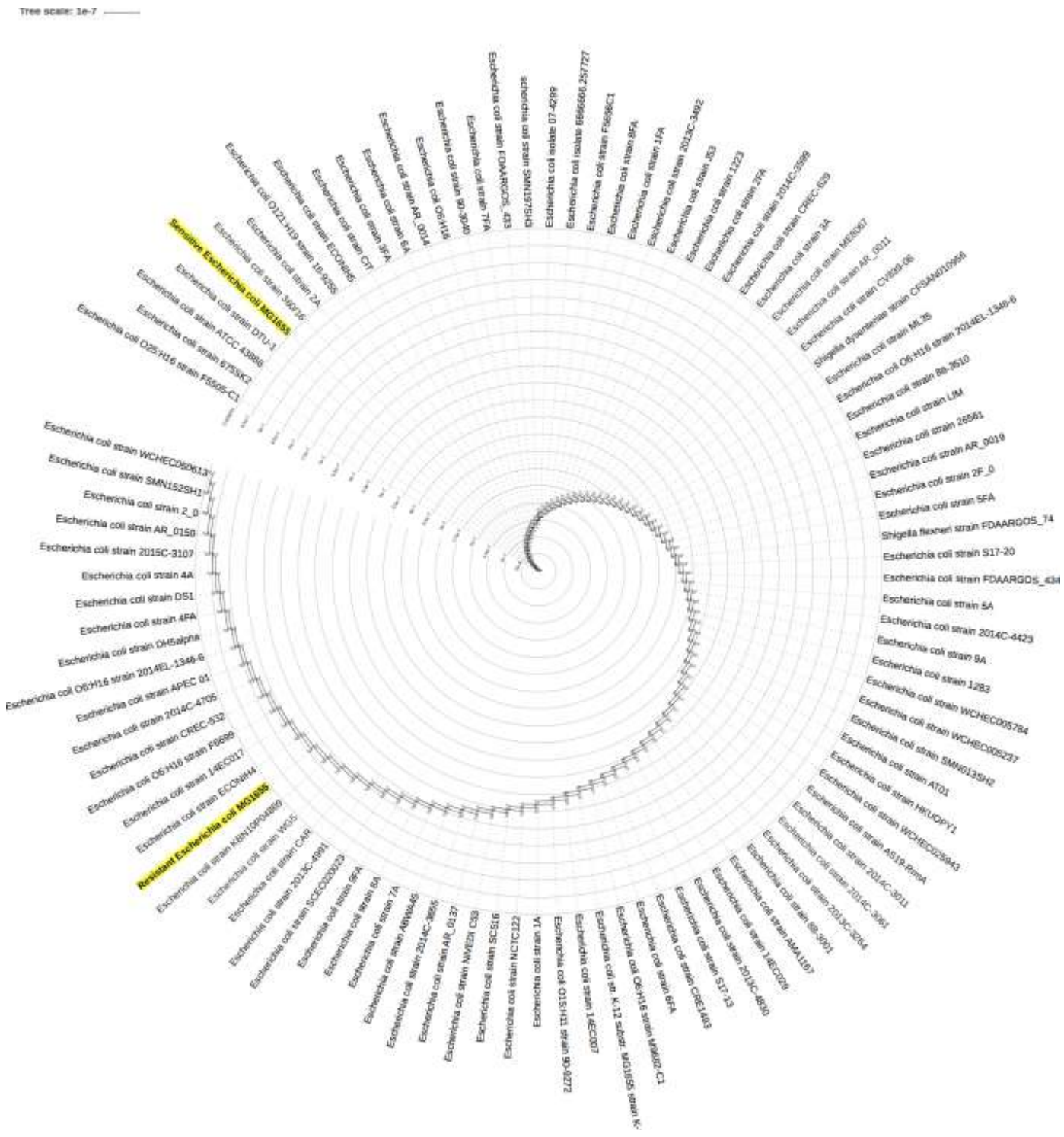


Figure A.21 Phylogenetic analysis of *fusA* within the *Escherichia* genus. iTOL image displaying the ML tree of the *fusA* gene within the *Escherichia* genus constructed using PhyML with a bootstrap of 1000. The E. coli MG1655 sensitive and resistant strains have been highlighted in yellow.

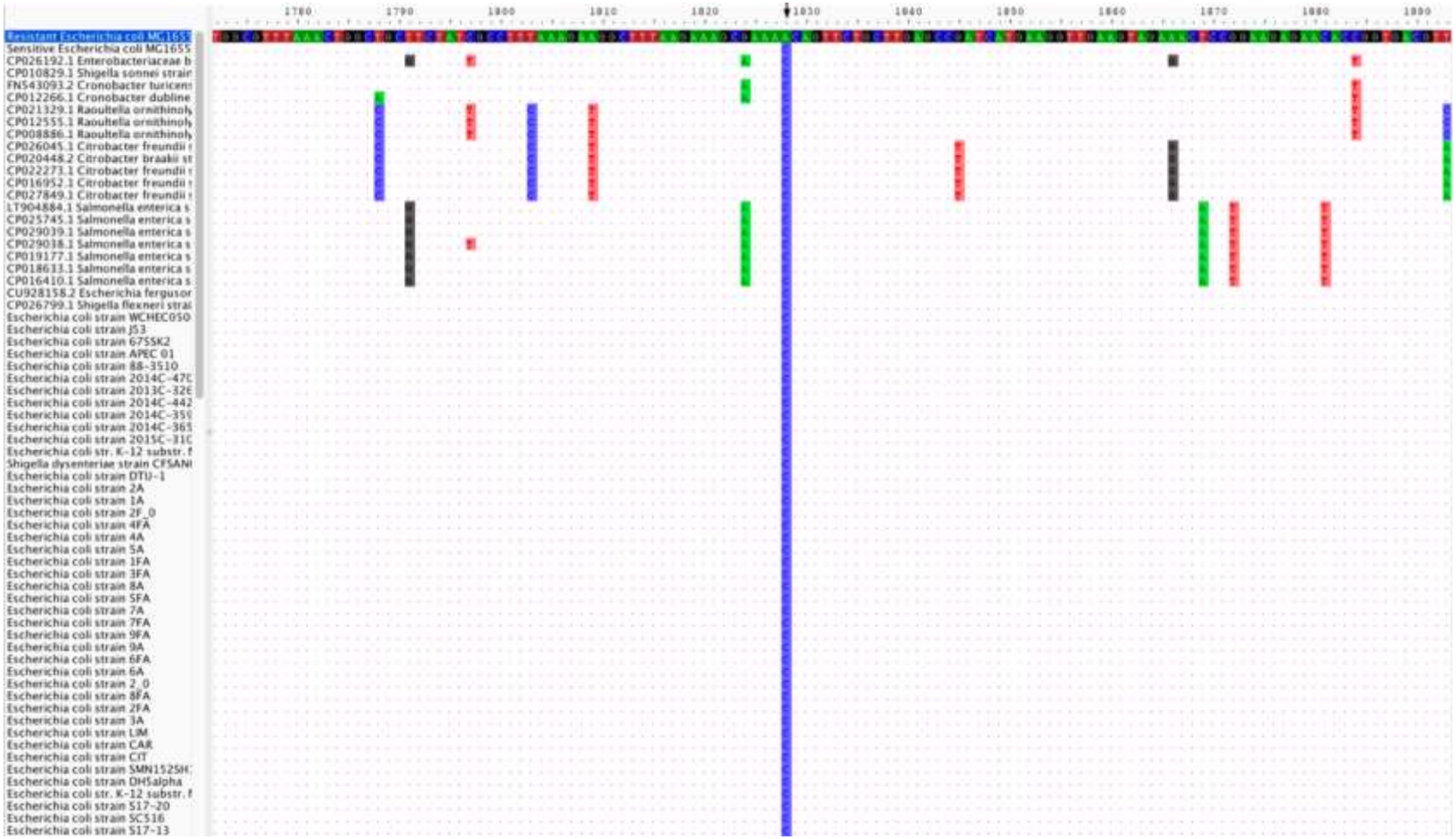


Figure A.22 Haplotype analysis of the *fusA* gene within the *Enterobacteriaceae* family. Aliview alignment image displaying the single haplotype within the *Enterobacteriaceae* family containing the *fusA* mutation depicted using the small black arrow.

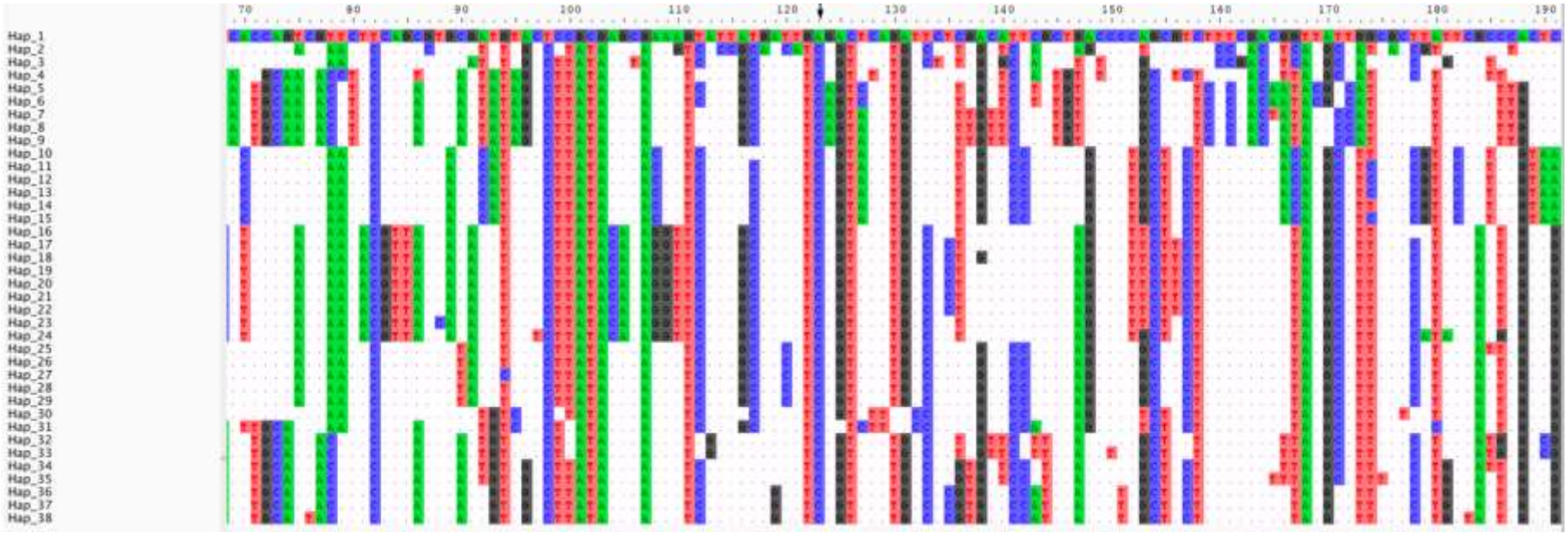


Figure A.23 Haplotype variation analysis of the *fusA* gene within the *Enterobacteriaceae* family. Aliview alignment image displaying the 38 haplotypes of *fusA* and the nucleotide variation within 340 nucleotides of the *fusA* gene sequence.

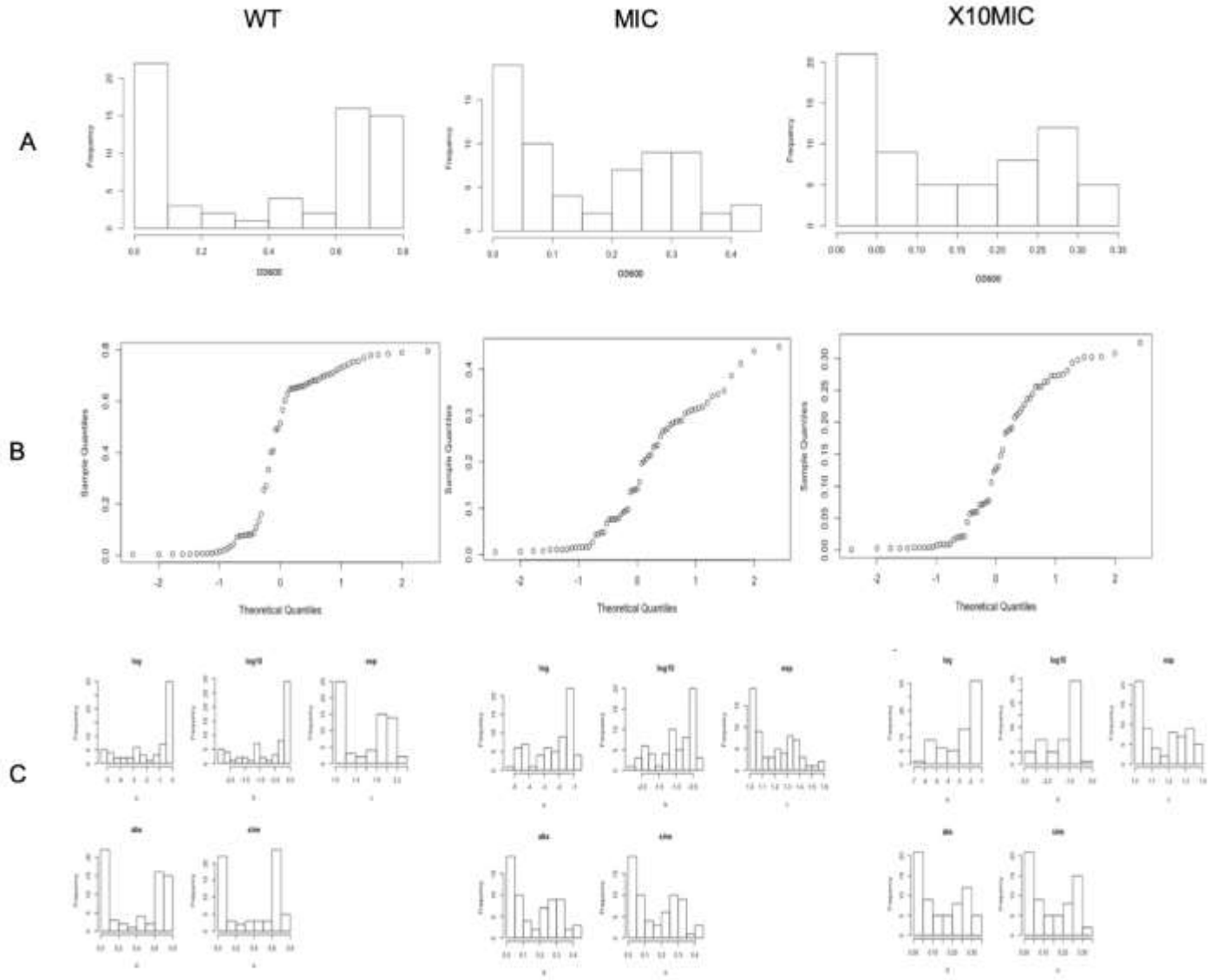


Figure A.24 Normality tests for the statistical analysis of the manual bacterial growth OD curves. (A) Histograms of WT, MIC and x10MIC, displaying the frequency of optical density values for the manual growth curve data. (B) QQ plots of WT, MIC and x10MIC, displaying the theoretical quantiles of the optical density values for the manual growth curve data. (C) Transformation plots (log, log10, exp, abs and sine) of the *E. coli* MG1655 WT MIC and x10 MIC manual growth curve data.

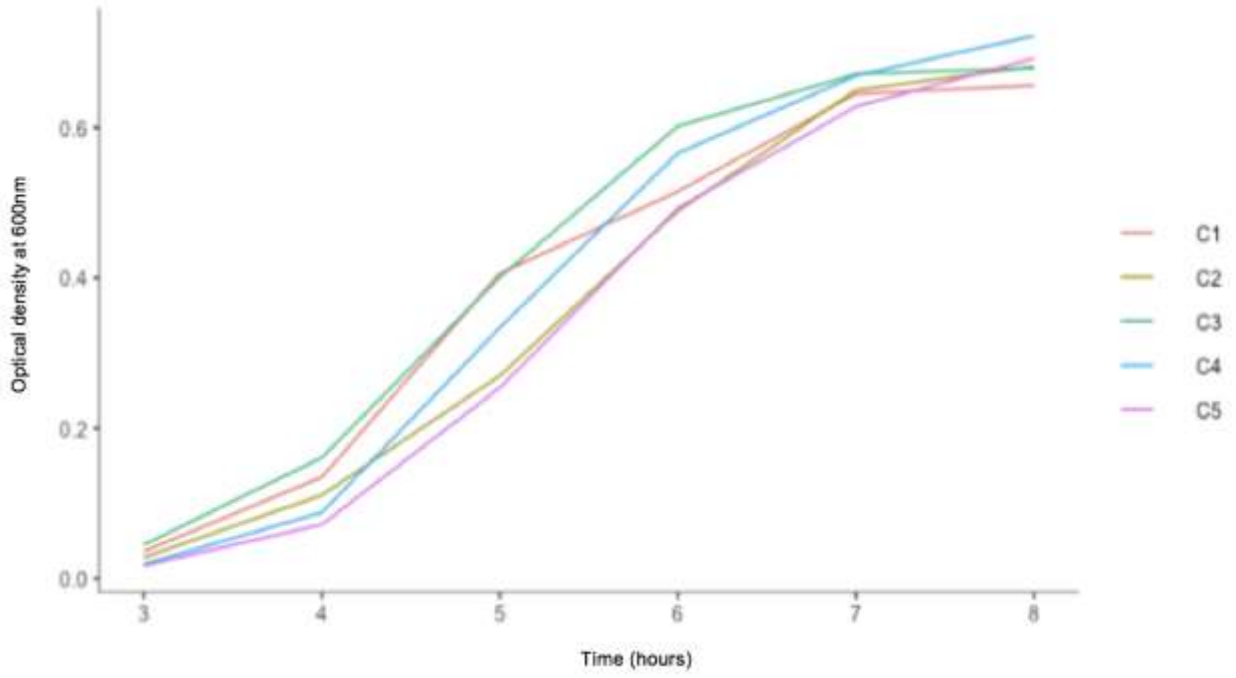


Figure A.25 Manual exponential growth analysis of *E. coli* MG1655 WT strain. Displaying the exponential growth curve of the *E. coli* MG1655 WT strain in MH broth at 37ºC. All five colonies (C1-C5) were grown within conical flasks and the OD_{600} values were measured using a spectrophotometer.

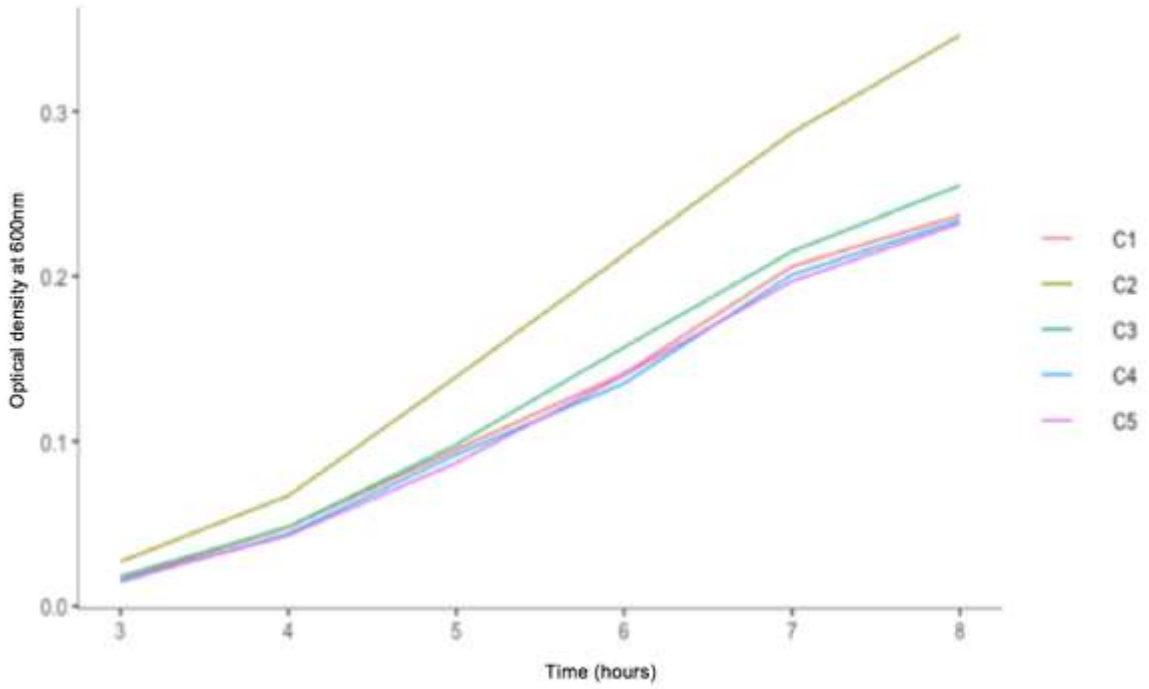


Figure A.26 Manual exponential growth analysis of *E. coli* MG1655 MIC strain. Displaying the exponential growth curve of the *E. coli* MG1655 MIC strain in MH broth at 37ºC. All five colonies (C1-C5) were grown within conical flasks and the OD_{600} values were measured using a spectrophotometer.

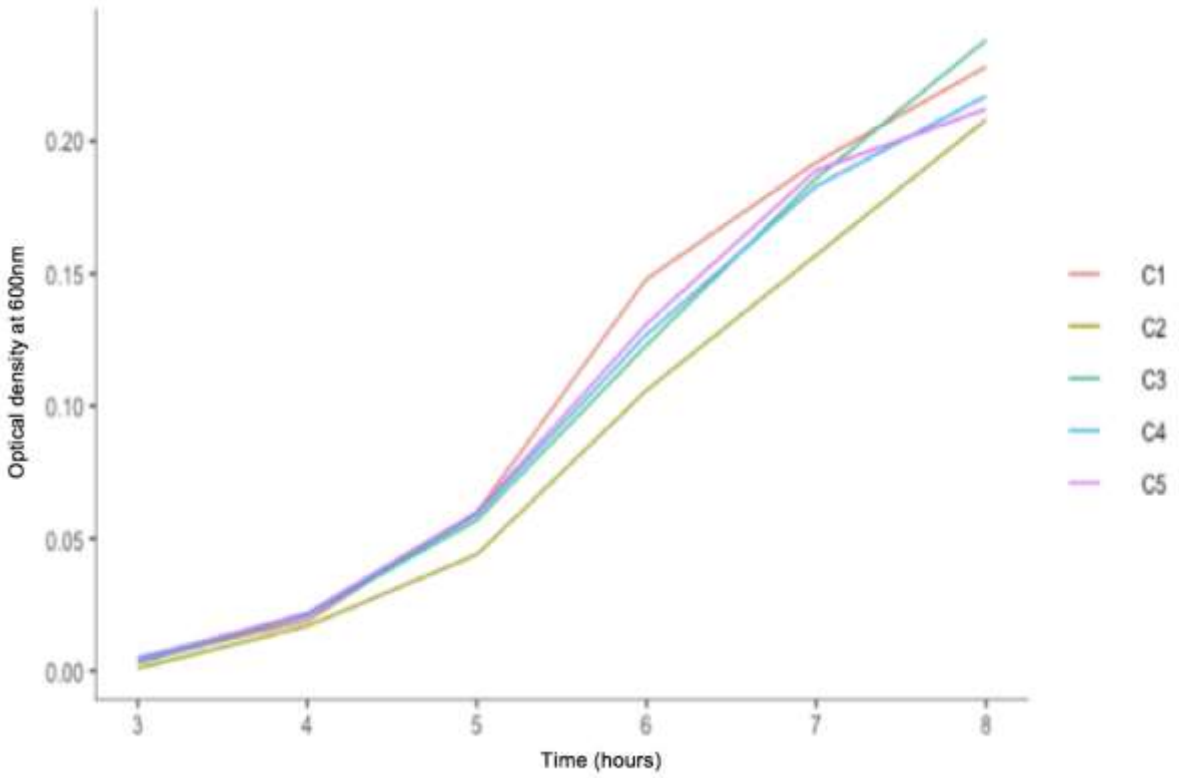


Figure A.27 Manual exponential growth analysis of *E. coli* MG1655 x10MIC strain. Displaying the exponential growth curve of the *E. coli* MG1655 x10MIC strain in MH broth at 37ºC. All five colonies (C1-C5) were grown within conical flasks and the OD_{600} values were measured using a spectrophotometer.

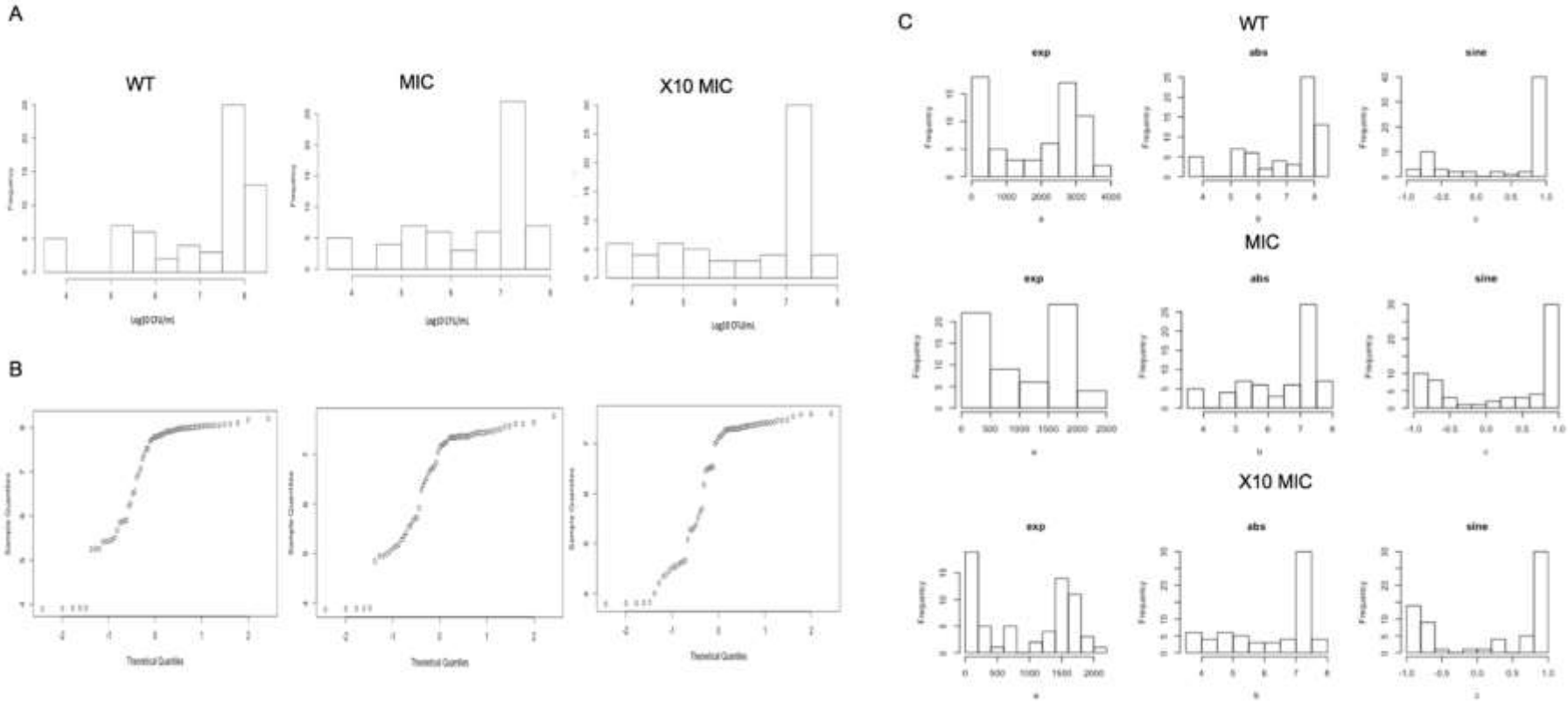


Figure A.28 Normality tests for the statistical analysis of the manual bacterial growth log10 CFU/mL curves. (A) Histograms of WT, MIC and x10MIC, displaying the frequency of log10 CFU/mL values for the manual growth curve data. (B) QQ plots of WT, MIC and x10MIC, displaying the theoretical quantiles of the log10 CFU/mL values for the manual growth curve data. (C) Transformation plots (exp, abs and sine) of the *E. coli* MG1655 WT MIC and x10 MIC manual growth curve data.

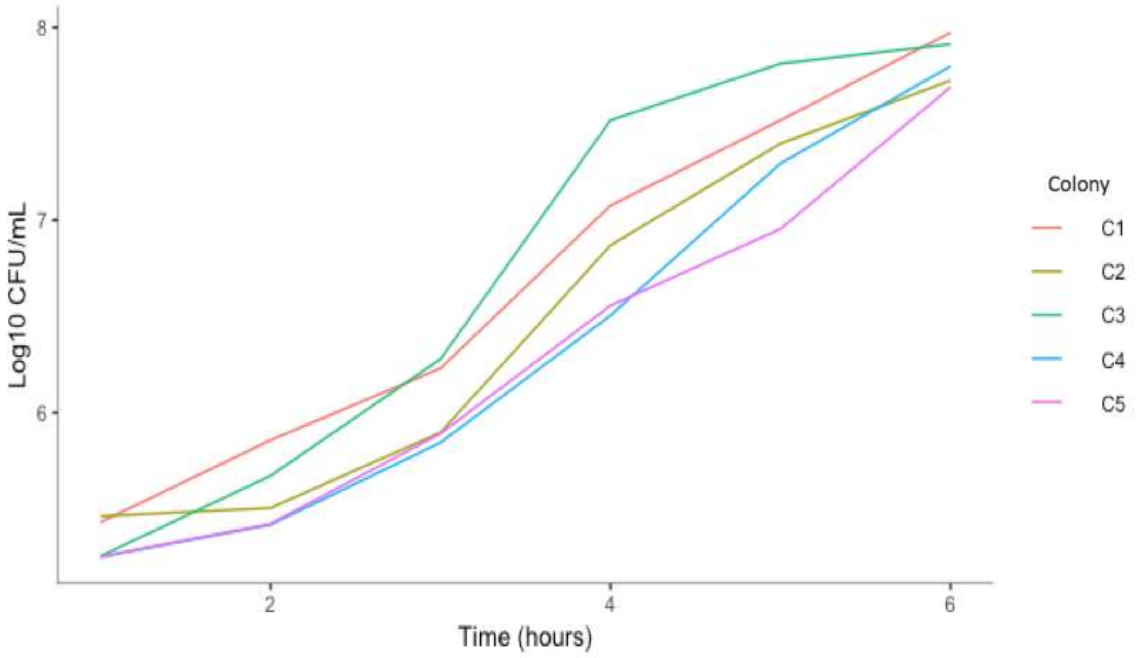


Figure A.29 Manual exponential growth analysis of *E. coli* MG1655 WT strain by bacterial enumeration. Displaying the exponential growth curve of the *E. coli* MG1655 WT strain in MH broth at 37ºC. All five colonies (C1-C5) were grown within conical flasks and samples were taken every hour for 12 hours. The bacterial samples were used to perform serial dilutions which were plated on MH plates and incubated O/N at 37ºC. Following incubation, the log10 CFU/mL was determined.

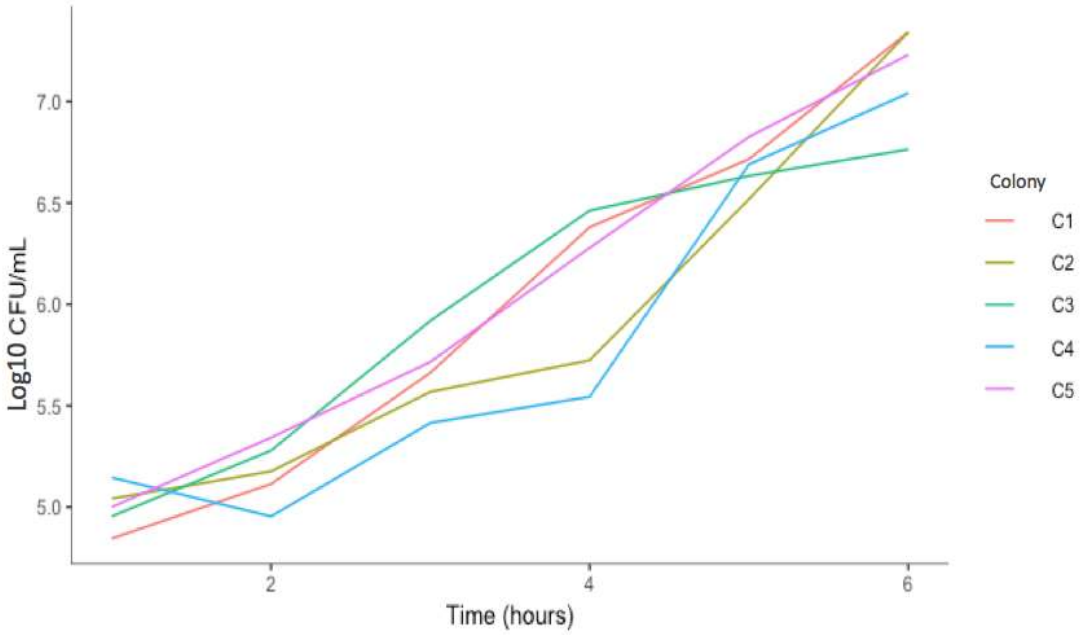


Figure A.30 Manual exponential growth analysis of *E. coli* MG1655 MIC strain by bacterial enumeration. Displaying the exponential growth curve of the *E. coli* MG1655 MIC strain in MH broth at 37ºC. All five colonies (C1-C5) were grown within conical flasks and samples were taken every hour for 12 hours. The bacterial samples were used to perform serial dilutions which were plated on MH plates and incubated O/N at 37ºC. Following incubation, the log10 CFU/mL was determined.

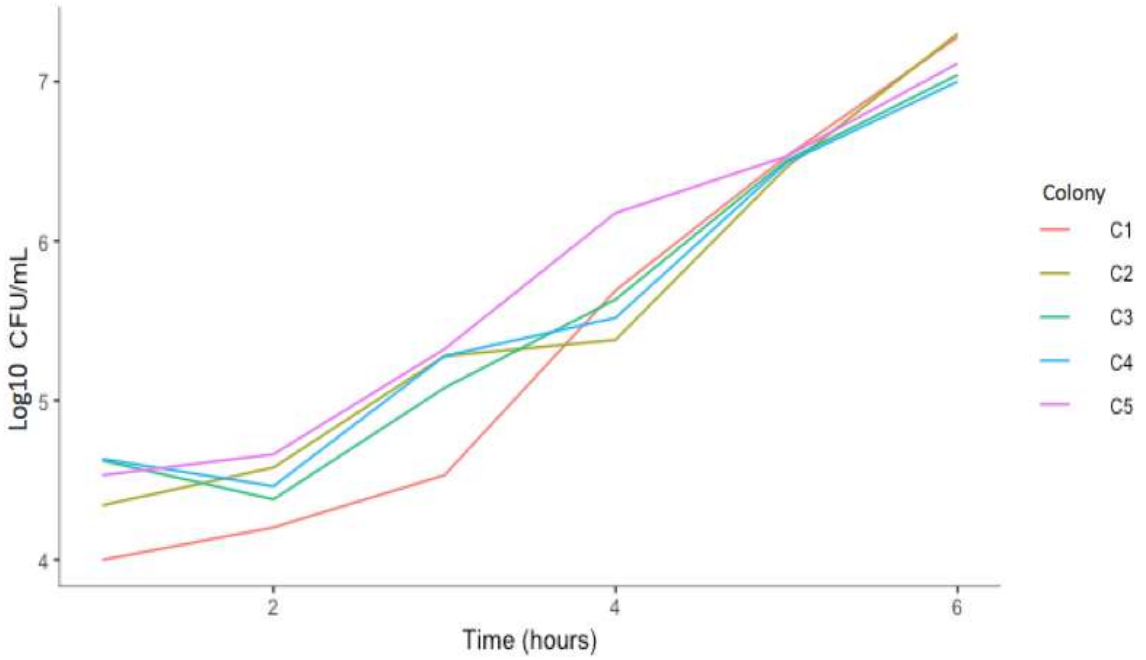


Figure A.31 Manual exponential growth analysis of *E. coli* MG1655 X10MIC strain by bacterial enumeration. Displaying the exponential growth curve of the *E. coli* MG1655 X10MIC strain in MH broth at 37ºC. All five colonies (C1-C5) were grown within conical flasks and samples were taken every hour for 12 hours. The bacterial samples were used to perform serial dilutions which were plated on MH plates and incubated O/N at 37ºC. Following incubation, the log10 CFU/mL was determined.

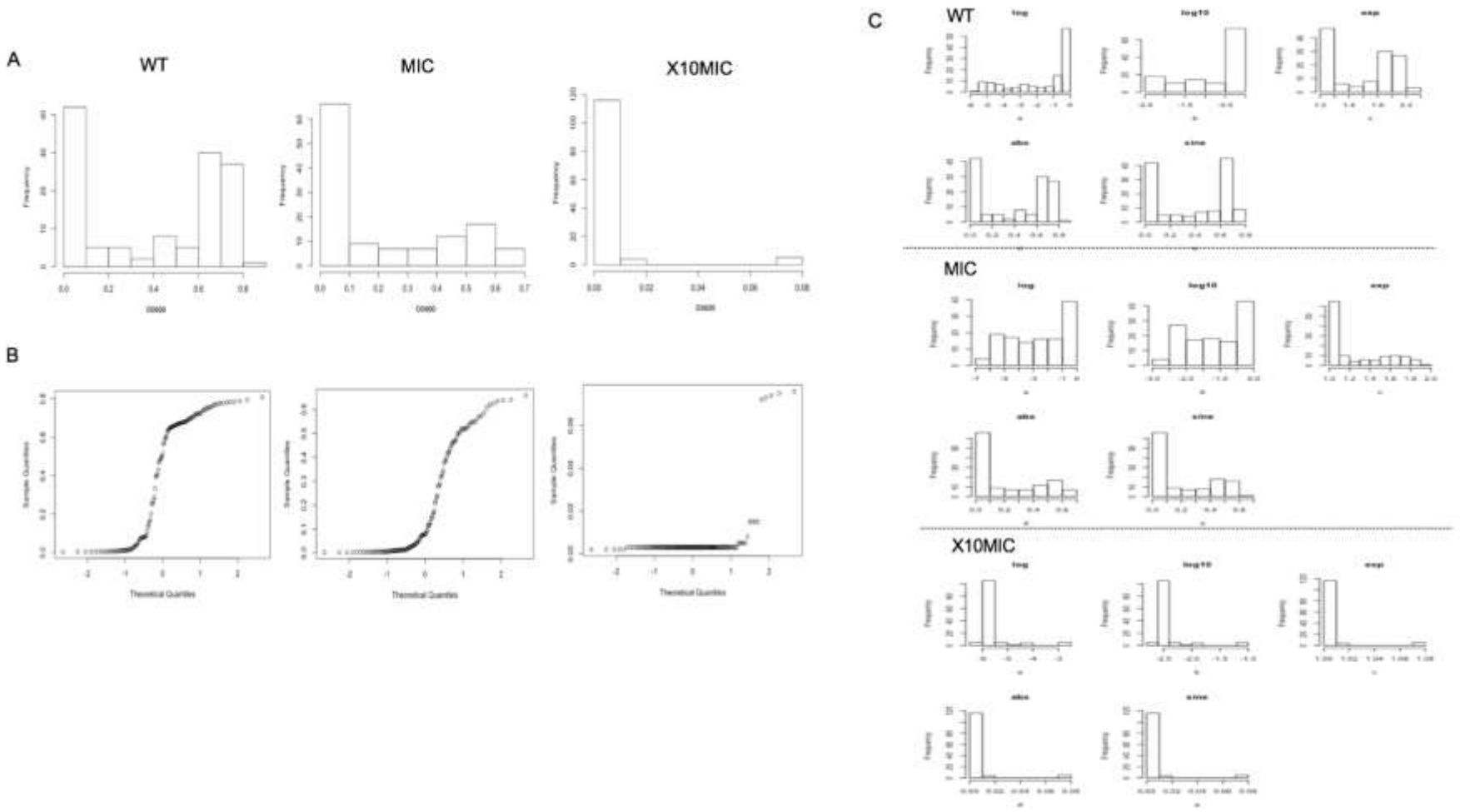


Figure A.32 Normality tests for the statistical analysis of the dose response OD curves. (A) Histograms of WT, MIC and x10MIC, displaying the frequency of optical density values for the dose response curve data. (B) QQ plots of WT, MIC and x10MIC, displaying the theoretical quantiles of the optical density values for the dose response data. (C) Transformation plots (log, log10, exp, abs and sine) of the *E. coli* MG1655 WT MIC and x10 MIC dose response data.

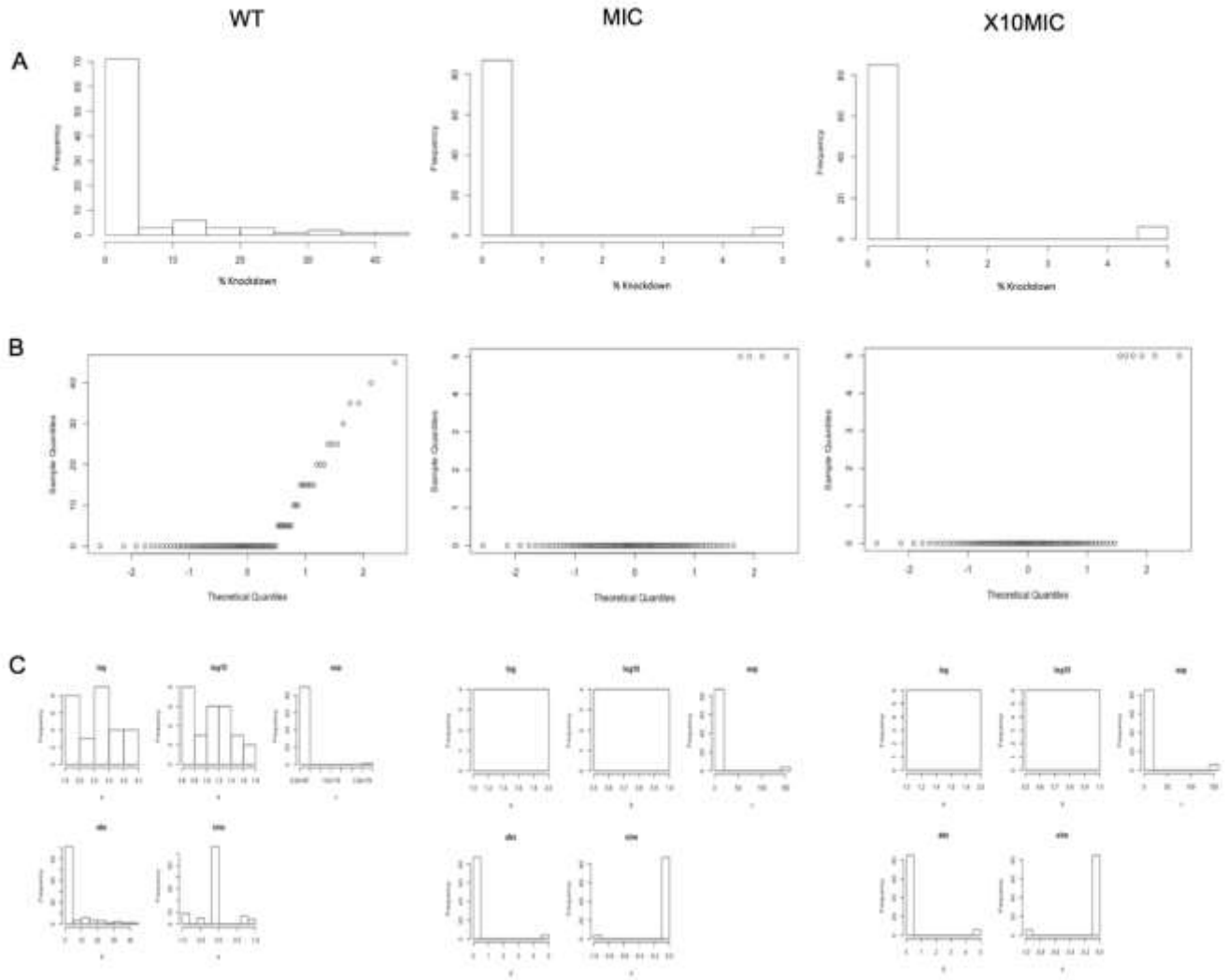


Figure A.33 Normality tests for the statistical analysis of the bacterial virulence. (A) Histograms of WT, MIC and x10MIC, displaying the frequency of %survival values for the virulence data. (B) QQ plots of WT, MIC and x10MIC, displaying the theoretical quantiles of the %survival values for the virulence data. (C) Transformation plots (log, log10, exp, abs and sine) of the *E. coli* MG1655 WT MIC and x10 MIC virulence data.

References

Abraham EP and Chain E. (1940) An enzyme from bacteria able to destroy penicillin. *Nature* 146: 837. doi: 10.1038/146837a0

Adler J and Templeton B. (1967) The effect of environmental conditions on the motility of *Escherichia coli*. *Journal of General Microbiology*. doi: 10.1099/00221287-46-2-175

Aggen JB, Armstrong ES, Goldblum AA, Dozzo P, Linsell MS, Gliedt MJ and Moser HE. (2010) Synthesis and spectrum of the neoglycoside ACHN-490. *Antimicrobial Agents and Chemotherapy*. doi: 10.1128/AAC.00572-10

Agirrezabala X, Lei J, Brunelle JL, Ortiz-Meoz RF, Green R and Frank, J. (2008) Visualization of the hybrid state of tRNA binding promoted by spontaneous ratcheting of the ribosome. *Molecular Cell*. doi: 10.1016/j.molcel.2008.10.001

Agrawal RK, Penczek P, Grassucci RA, and Frank, J. (1998) Visualization of elongation factor G on the *Escherichia coli* 70S ribosome: the mechanism of translocation. *Proceedings of the National Academy of Sciences of the United States of America* 95 (11): 6134–6138. doi: 10.1073/PNAS.95.11.6134

Albert TJ, Dailidiene D, Dailide G, Norton JE, Kalia A, Richmond TA, Molla M, Singh J, Green RD and Berg DE. (2005) Mutation discovery in bacterial genomes: metronidazole resistance in *Helicobacter pylori*. *Nature Methods 2* 951. doi: 10.1038/nmeth805

Alhakami H, Mirebrahim, H and Lonardi S. (2017) A comparative evaluation of genome assembly reconciliation tools. *Genome Biology 18*(93). doi: 10.1186/s13059-017-1213-3

Alkan C, Sajjadian S and Eichler EE. (2011) Limitations of next-generation genome sequence assembly. *Nature Methods 8*(1): 61–65.

Alland D, Whittam TS, Murray MB, Cave MD, Hazbon MH, Dix K, Kokoris M, Duesterhoeft A, Eisen JA, Fraser CM and Fleischmann RD. (2003) Modeling bacterial evolution with comparative-genome-based marker systems: application to *Mycobacterium tuberculosis* evolution and pathogenesis. *Journal of Bacteriology*. doi: 10.1128/JB.185.11.3392-3399.2003

Allmang C, Mougel M, Westhof E, Ehresmann B, and Ehresmann C. (1994) Role of conserved nucleotides in building the 16S rRNA binding site of *E. coli* ribosomal protein S8. *Nucleic Acids Research* (22).

Alonso A, Campanario E and Martínez JL. (1999) Emergence of multidrug-resistant mutants is increased under antibiotic selective pressure in *Pseudomonas aeruginosa*. *Microbiology*. doi: 10.1099/00221287-145-10-2857

Alves MJ, Froufe HJ, Costa AF, Santos AF, Oliveira LG, Osório SR, Abreu RM, Pintado M and Ferreira IC. (2014) Docking studies in target proteins involved in antibacterial action mechanisms: Extending the knowledge on standard antibiotics to antimicrobial mushroom compounds. *Molecules*. doi: 10.3390/molecules19021672

Ambardar S, Gupta R, Trakroo D, Lal R and Vakhlu J. (2016) High Throughput Sequencing: An Overview of Sequencing Chemistry. *Indian Journal of Microbiology*. doi: 10.1007/s12088-016-0606-4

Aminov RI. (2010) A brief history of the antibiotic era: Lessons learned and challenges for the future. *Frontiers in Microbiology*. doi: 10.3389/fmicb.2010.00134

Andersson DI and Levin BR. (1999) The biological cost of antibiotic resistance. *Current Opinion in Microbiology* 2: 489–493.

Andersson DI. (2003) Persistence of antibiotic resistant bacteria. *Current Opinion in Microbiology 6*(5): 452–456. doi: doi: 10.1016/j.mib.2003.09.001

Andersson DI. (2006) The biological cost of mutational antibiotic resistance: any practical conclusions? *Current Opinion in Microbiology*. doi: 10.1016/j.mib.2006.07.002

Andersson DI and Hughes D. (2010) Antibiotic resistance and its cost: is it possible to reverse resistance? *Nature Reviews Microbiology* 8: 260–271.

Andersson DI and Hughes D. (2011) Persistence of antibiotic resistance in bacterial populations. *FEMS Microbiology Reviews.* doi: 10.1111/j.1574-6976.2011.00289.x

Andersson DI and Hughes D. (2012) Evolution of antibiotic resistance at non-lethal drug concentrations. *Drug Resistance Updates*. doi: 10.1016/j.drup.2012.03.005

Ankomah P and Levin BR. (2014) Exploring the collaboration between antibiotics and the immune response in the treatment of acute, self-limiting infections. *Proceedings of the National Academy of Sciences*. doi: 10.1073/pnas.1400352111

Arcelloni C, Comuzzi B, Vaiani R and Paroni R. (2001) Quantification of gentamicin in mueller-hinton agar by high-performance liquid chromatography. *Journal of Chromatography B: Biomedical Sciences and Applications*. doi: 10.1016/S0378-4347(00)00460-6

Asai T, Zaporojets D, Squires C and Squires CL. (1999) An *Escherichia coli* strain with all chromosomal rRNA operons inactivated: Complete exchange of rRNA genes between bacteria. *Proceedings of the National Academy of Sciences*. doi: 10.1073/pnas.96.5.1971

Asai T, Ciara C, Condon C, Voulgaris J, Zaporojets D, Shen B, Al-Omar M, Squires C and Squires CL. (1999) Construction and initial characterization of *Escherichia coli* strains with few or no intact chromosomal rRNA operons. *Journal of Bacteriology* (181).

Ashkenazy H, Erez E, Martz E, Pupko T and Ben-Tal N. (2010) ConSurf 2010: calculating evolutionary conservation in sequence and structure of proteins and nucleic acids. *Nucleic Acids Research*. doi: 10.1093/nar/gkq399

Asuncio A, Martinez A, Martinez M, Torello S and Kolter R. (1999) Sliding motility in *Mycobacteria*. *Journal of Bacteriology* (181).

Atkinson CC. (2015) The evolutionary and functional diversity of classical and lesser-known cytoplasmic and organellar translational GTPases across the tree of life. *BMC Genomics*. doi: 10.1186/s12864-015-1289-7

Aubry-Damon H, Soussy CJ and Courvalin P. (1998) Characterization of mutations in the rpoB gene that confer rifampin resistance in *Staphylococcus aureus*. *Antimicrobial Agents and Chemotherapy*. doi: 10.1128/aac.49.3.903-907.2005

AVMA. (2008) One Health Initiative Task Force: Final Report. Available at: https://www.avma.org/KB/Resources/Reports/Documents/onehealth_final.pdf

AEvarsson A, Brazhnikov E, Garberl M, Zheltonosova J, Chirgadze Y, Al-Karadaghi S, Svensson LA and Liljas A. (1994) Three-dimensional structure of the ribosomal translocase: elongation factor G from *Thermus thermophilus*. *The EMBO Journal* (13).

Baldauf SL. (2003) Phylogeny for the faint of heart: a tutorial. *Trends in Genetics* 19(6): 345–351. doi: 10.1016/S0168-9525(03)00112-4

Bamshad M and Wooding SP. (2003). Signatures of natural selection in the human genome. *Nature Reviews Genetics*. doi: 10.1038/nrg999

Baquero F and Blázquez J. (1997) Evolution of antibiotic resistance. *Trends in Ecology and Evolution*. doi: 10.1016/S0169-5347(97)01223-8

Barash D and Churkin A. (2010) Mutational analysis in RNAs: comparing programs for RNA deleterious mutation prediction. *Briefings in Bioinformatics 12*(2): 104–114. doi: 10.1093/bib/bbq059

Barnes MR and Gray IC. (2003) Amino acid properties and consequences of substitutions. Bioinformatics for Geneticists 289-316. doi: 10. 1002/0470867302

Barker CS, Prüß BM and Matsumura P. (2004) Increased motility of *Escherichia coli* by insertion sequence element integration into the regulatory region of the flhD operon. *Journal of Bacteriology*. doi: 10.1128/JB.186.22.7529-7537.2004

Barrick JE, Yu DS, Yoon SH, Jeong H, Oh TK, Schneider D, Lenski RE and Kim JF. (2009) Genome evolution and adaptation in a long-term experiment with *Escherichia coli*. *Nature* 461(7268). doi: 10.1038/nature08480

Barrick JE and Lenski RE. (2013) Genome dynamics during experimental evolution. *Nature Reviews Genetics*. doi: 10.1038/nrg3564

Bartholomew JW and Mittwer T. (1952) The Gram Stain. *Bacteriology Reviews* 16(1): 1-29.

Baym M, Stone LK, and Kishony R. (2016a) Multidrug evolutionary strategies to reverse antibiotic resistance. *Science*. doi: 10.1126/science.aad3292

Baym M, Lieberman TD, Kelsic ED, Chait R, Gross R, Yelin I and Kishony R. (2016b) Spatiotemporal microbial evolution on antibiotic landscapes. *Science* 353(6304): 1147–1151. doi: 10.1126/science.aag0822

Baquero F and Negri MC. (2005) Challenges: selective compartments for resistant microorganisms in antibiotic gradients. *Bioessays 19*(8). doi: 10.1002/bies.950190814

Beceiro A, Tomás M and Bou G. (2013) Antimicrobial resistance and virulence: a successful or deleterious association in the bacterial world? *Clinical Microbiology Reviews*. doi: 10.1128/CMR.00059-12

Begg EJ and Barclay ML. (1995) Aminoglycosides--50 years on. *British Journal of Clinical Pharmacology*. doi: 10.1051/0004-6361/200913605

Benner SA, Chamberlin SG, Liberles DA, Govindarajan S and Knecht L. (2000) Functional inferences from reconstructed evolutionary biology involving rectified databases - An evolutionarily grounded approach to functional genomics. *Research in Microbiology*. doi: 10.1016/S0923-2508(00)00123-6

Bennett AF and Lenski RE. (1997) Phenotypic and evolutionary adaptation of a model bacterial system to stressful thermal environments. *Environmental Stress, Adaptation and Evolution*. 135–154.

Benveniste R and Davies J. (1973) Structure-activity relationships among the aminoglycoside antibiotics: role of hydroxyl and amino groups. *Antimicrobial Agents and Chemotherapy*. doi: 10.1128/AAC.4.4.402

Bergmiller T, Ackermann M and Silander OK. (2012). Patterns of evolutionary conservation of essential genes correlate with their compensability. *PLoS Genetics*. doi: 10.1371/journal.pgen.1002803

Besier S, Ludwig A, Brade V and Wichelhaus TA. (2003) Molecular analysis of fusidic acid resistance in *Staphylococcus aureus*. *Molecular Microbiology*. doi: 10.1046/j.1365-2958.2003.03307.x

Besier S, Ludwig A, Brade V and Wichelhaus TA. (2005) Compensatory adaptation to the loss of biological fitness associated with acquisition of fusidic acid resistance in *Staphylococcus aureus*. *Antimicrobial Agents and Chemotherapy*. doi: 10.1128/AAC.49.4.1426-1431.2005

Betts MJ and Russell RB. (2003) Amino acid properties and consequences of substitutions. In Barnes MR and Gray IC (Eds), *Bioinformatics for genetics.* Wiley.

Binet R and Maurelli AT. (2005) Fitness cost due to mutations in the 16S rRNA associated with spectinomycin resistance in *Chlamydia psittaci* 6BC. *Antimicrobial Agents and Chemotherapy*. doi: 10.1128/AAC.49.11.4455-4464.2005

Binet R and Maurelli AT. (2005) Frequency of spontaneous mutations that confer antibiotic resistance in *Chlamydia* spp. *Antimicrobial Agents and Chemotherapy*. doi: 10.1128/AAC.49.7.2865-2873.2005

Björkman J, Hughes D and Andersson DI. (1998) Virulence of antibiotic-resistant *Salmonella typhimurium*. *Proceedings of the National Academy of Sciences* 95(7): 3949-53.

Björkman J, Nagaev I, Berg OG, Hughes D and Andersson DI. (2000) Effects of environment on compensatory mutations to ameliorate costs of antibiotic resistance. *Science* 287(5457):1479-82. doi: 10.1126/science.287.5457.1479

Björkman J and Andersson DI. (2000) The cost of antibiotic resistance from a bacterial perspective. *Drug Resistance Updates*. doi: 10.1054/drup.2000.0147

Björkman J, Samuelsson P, Andersson DI and Hughes D. (1999) Novel ribosomal mutations affecting translational accuracy, antibiotic resistance and virulence of *Salmonella typhimurium*. *Molecular Microbiology*. doi: 10.1046/j.1365-2958.1999.01142.x

Blaha G, Gürel G, Schroeder SJ, Moore PB and Steitz TA. (2008) Mutations outside the anisomycin-binding site can make ribosomes drug-resistant. *Journal of Molecular Biology*. doi: 10.1016/j.jmb.2008.03.075

Blakely GW. (2015) Mechanisms of horizontal gene transfer and DNA recombination. In Tang YW, Sussman M, Liu D, Poxton I and Schwartzman J (Eds), *Molecular Medical Microbiology (Second Edition)*. Boston: Academic Press. 291–302. doi: 10.1016/B978-0-12-397169-2.00015-9

Blanchaert B, Huang S, Wach K, Adams E and Schepdael AV. (2017) Assay Development for Aminoglycosides by HPLC with Direct UV Detection. *Journal of Chromatographic Science* 55(3): 197-204.

Blattner FR, Plunkett G, Bloch CA, Perna NT, Burland V, Riley M, Collado-Vides J, Glasner JD, Rode CK, Mayhew GF, Gregor J, Davis NW, Kirkpatrick HA, Goeden

MA, Rose DJ, Mau B and Shao Y. (1997) The Complete Genome Sequence of Escherichia coli K-12. *Science*, *277*(5331): 1453–1462. doi: 10.1126/science.277.5331.1453

Blount, Z. D, Borland, C. Z, and Lenski, R. E. (2008). Historical contingency and the evolution of a key innovation in an experimental population of *Escherichia coli*. *Proceedings of the National Academy of Sciences*. doi: 10.1073/pnas.0803151105

Blount ZD. (2015) The unexhausted potential of *E. coli*. *Elife*. doi: 10.7554/eLife.05826.001

Bogdanovich T, Adams-Haduch JM, Tian GB, Nguyen MH, Kwak EJ, Muto CA, and Doi Y. (2011) Colistin-resistant, *Klebsiella pneumoniae* Carbapenemase (KPC)-producing *Klebsiella pneumoniae* belonging to the international epidemic clone ST258. *Clinical Infectious Diseases*. doi: 10.1093/cid/cir401

Bohman K, Ruusala T, Jelenc PC and Kurland CG. (1984) Kinetic impairment of restrictive streptomycin-resistant ribosomes. *MGG Molecular and General Genetics*. doi: 10.1007/BF00328706

Bonev B, Hooper J and Parisot, J. (2008) Principles of assessing bacterial susceptibility to antibiotics using the agar diffusion method. *Journal of Antimicrobial Chemotherapy*. doi: 10.1093/jac/dkn090

Bonhoeffer S, Lipsitch M and Levin BR. (1997) Evaluating treatment protocols to prevent antibiotic resistance. *Medical Sciences* (94). Retrieved from www.pnas.org.

Bonten MJM, Austin DJ and Lipsitch M. (2001) Understanding the spread of antibiotic resistant pathogens in hospitals: mathematical models as tools for control. *Clinical Infectious Diseases*. doi: 10.1086/323761

Borovinskaya MA, Pai RD, Zhang W, Schuwirth BS, Holton JM, Hirokawa G, Kaji H, Kaji A and Cate JHD. (2007) Structural basis for aminoglycoside inhibition of bacterial ribosome recycling. *Nature Structural and Molecular Biology* 14(8). doi: 10.1038/nsmb1271

Bottger EC, Pletschette M and Andersson DI. (2005) Drug resistance and fitness in *Mycobacterium tuberculosis* infection. *Journal of Infectious Diseases* 191(5): 823-4.

Bowen WS, van Dyke N, Murgola EJ, Lodmell JS and Hill WE. (2005) Interaction of thiostrepton and elongation factor-G with the ribosomal protein L11-binding domain. *Journal of Biological Chemistry*. doi: 10.1074/jbc.M407008200

Bradley P, Gordon NC, Walker TM, Dunn L, Heys S, Huang B, Earle S, Pankhurst LJ, Anson L, de Cesare M, Piazza P, Votintseva AA, Golubchik T, Wilson DJ, Wyllie DH, Diel R, Niemann S, Feuerriegel S, Kohl TA, Ismail N, Omar SV, Smith GE, Buck D, McVean, Walker SA, Peto TEA, Crook DW and Iqbal Z. (2015) Rapid antibiotic-resistance predictions from genome sequence data for *Staphylococcus aureus* and *Mycobacterium tuberculosis*. *Nature Communications*. doi: 10.1038/ncomms10063

Brandi, L, Fabbretti, A, Di Stefano, M, Lazzarini, A, Abbondi, M, and Gualerzi, C. O. (2006). Characterization of GE82832, a peptide inhibitor of translocation interacting with bacterial 30S ribosomal subunits. *RNA*. doi: 10.1261/rna.61206

Brdicka R. (1995) The human genome—chromosome 17. *Casopis Lekaru Ceskych* 134(17):555-7. doi: 10.1371/journal.pone.0128036

Brilot AF, Korostelev AA, Ermolenko DN and Grigorieff N. (2013) Structure of the ribosome with elongation factor G trapped in the pretranslocation state. *Proceedings of the National Academy of Sciences*. doi: 10.1073/pnas.1311423110

Brodersen DE, Clemons WM Jr, Carter AP, Morgan-Warren RJ, Wimberly BT and Ramakrishnan V. (2000) The structural basis for the action of the antibiotics tetracycline, pactamycin, and hygromycin B on the 30S Ribosomal Subunit. *Cell* 103(7): 1143–1154. doi: 10.1016/S0092-8674(00)00216-6

Brown D. (2015) Antibiotic resistance breakers: can repurposed drugs fill the antibiotic discovery void? *Nature Reviews Drug Discovery* 14(12): 821-32. doi: 10.1038/nrd4675

Bryant JCC and Bentley SD. (2012) Developing insights into the mechanisms of evolution of bacterial pathogens from whole-genome sequences europe PMC funders group. *Future Microbiol* 7(11): 1283–1296. doi: 10.2217/fmb.12.108

Bulkley D, Brandi L, Polikanov YS, Fabbretti A, O'Connor M, Gualerzi CO and Steitz TA. (2014) The antibiotics dityromycin and GE82832 bind protein S12 and block EF-G-Catalyzed translocation. *Cell Reports*. doi: 10.1016/j.celrep.2013.12.024

Burgos HL, O'Connor K, Sanchez-Vazquez P and Gourse RL. (2017) Roles of transcriptional and translational control mechanisms in regulation of ribosomal protein synthesis in Escherichia coli. *Journal of Bacteriology*. doi: 10.1128/JB.00407-17

Calderwood SA, Wennersten C, Moellering RC, Kunz LJ and Krogstad DJ. (1977) Resistance to six aminoglycosidic aminocyclitol antibiotics among enterococci: prevalence, evolution, and relationship to synergism with penicillin. *Antimicrobial Agents and Chemotherapy*. doi: 10.1128/AAC.12.3.401

Cantón R and Coque TM. (2006) The CTX-M β-lactamase pandemic. *Current Opinion in Microbiology*. doi: 10.1016/j.mib.2006.08.011

Carter AP, Clemons WM, Brodersen DE, Morgan-Warren RJ, Wimberly BT and Ramakrishnan V. (2000) Functional insights from the structure of the 30S ribosomal subunit and its interactions with antibiotics. *Nature*, *407*(6802): 340-8. doi: 10.1038/35030019

Cassini A, Högberg LD, Plachouras D, Quattrocchi A, Hoxha A, Simonsen GS, Colomb-cotinat M, Kretzschmar ME, Devleesschauwer B, Cecchini M, Ouakrim DA, Oliveira TC, Struelens MJ, Suetens C and Monnet DL. (2018) Attributable deaths and disability-adjusted life-years caused by infections with antibiotic-resistant bacteria in the EU and the european economic area in 2015: a population-level modelling analysis. *The Lancet Infectious Diseases*. doi: 10.1016/S1473-3099(18)30605-4

CDC. (2013) Antibiotic Resistance Threats in the United States. Available at: https://www.cdc.gov/drugresistance/pdf/ar-threats-2013-508.pdf

Chain PSG *et al*. (2009) Genome project standards in a new era of sequencing. *Science 9*(326). doi:10.1126/science.1180614

Charlesworth D, Barton NH and Charlesworth B. (2017) The sources of adaptive variation. *Proceedings of the Royal Society B: Biological Sciences* 284(1855). doi: 10.1098/rspb.2016.2864

Chasman D and Adams MR. (2001) Predicting the functional consequences of non-synonymous single nucleotide polymorphisms: structure-based assessment of amino acid variation. *Journal of Molecular Biology* 307(2): 683–706. doi: 10.1006/jmbi.2001.4510

Chattaway MA, Schaefer U, Tewolde R, Dallman TJ and Jenkins C. (2017) Identification of *Escherichia coli* and *Shigella* species from whole-genome sequences. *Journal of Clinical Microbiology*. doi: 10.1128/JCM.01790-16

Chavda KD, Chen L, Fouts DE, Sutton G, Brinkac L, Jenkins SG, Bonomo RA, Adams MD and Kreiswirth, BN. (2016) Comprehensive genome analysis of carbapenemase-producing Enterobacter spp.: New insights into phylogeny, population structure, and resistance mechanisms. *MBio*. doi: 10.1128/mBio.02093-16

Chelvanayagam G, Eggenschwiler A, Knecht L, Gonnet GH and Benner SA. (1997) An analysis of simultaneous variation in protein structures. *Protein Engineering, Design and Selection* 10(4): 307–16.

Chen C, Blumentritt CA, Curtis MM, Sperandio V, Torres AG, and Dudley EG. (2013) Restrictive streptomycin resistance mutations decrease the formation of attaching and effacing lesions in *Escherichia coli* O157:H7 strains. *Antimicrobial Agents and Chemotherapy*. doi: 10.1128/AAC.00709-13

Chiou CS and Jones AL. (1995) Molecular analysis of high-level streptomycin resistance in *Erwinia amylovora*. *Phytopathology* 85(3). doi: 10. 1094/Phyto-85-324

Cho J, Hamasaki K and Rando RR. (1998) The binding site of a specific aminoglycoside binding RNA molecule. *Biochemistry* 37(14): 4985–4992. doi: 10.1021/bi972757h

Choi Y, Sims GE, Murphy S, Miller JR, Chan AP (2012). Predicting the functional effect of amino acid substitutions and indels. *PLOS ONE*, *7*(10). doi: 10.1371/journal.pone.0046688

Chopra I, O'Neill AJ and Miller K. (2003) The role of mutators in the emergence of antibiotic-resistant bacteria. *Drug Resistance Updates*. doi: 10.1016/S1368-7646(03)00041-4

Clark BFC, Thirup S, Kjeldgaard M and Nyborg J. (1999) Structural information for explaining the molecular mechanism of protein biosynthesis. *FEBS Letters*. doi: 10.1016/S0014-5793(99)00562-1

Clements A, Young JC, Constantinou N and Frankel G. (2012) Infection strategies of enteric pathogenic *Escherichia coli*. *Gut Microbes*. doi: 10.4161/gmic.19182

Comas I, Borrell S, Roetzer A, Rose G, Malla B, Kato-Maeda M, Galagan S and Gagneux S. (2012) Whole-genome sequencing of rifampicin-resistant *Mycobacterium tuberculosis* strains identifies compensatory mutations in RNA polymerase genes. *Nature Genetics*. doi: 10.1038/ng.1038

Condon C, Liveris D, Squires C, Schwartz I and Squires, C. L. (1995) rRNA operon multiplicity in *Escherichia coli* and the physiological implications of rrn inactivation. *Journal of Bacteriology*. doi: 10.1128/jb.177.14.4152-4156.1995

Condon C, Philips J, Fu ZY, Squires C and Squires CL. (1992) Comparison of the expression of the seven ribosomal RNA operons in Escherichia coli. *The EMBO Journal* (11): 4175-85.

Conrad TM, Lewis NE and Palsson BO. (2011) Microbial laboratory evolution in the era of genome-scale science. *Molecular Systems Biology* 7(509). doi: 10.1038/msb.2011.42

Consortium REX. (2007) Structure of the scientific community modelling the evolution of resistance. *PLOS ONE* 2(12): 1–9. doi: 10.1371/journal.pone.0001275

Conway T, Krogfelt KA and Cohen PS. (2004) The life of commensal *Escherichia co*li in the mammalian intestine. *EcoSal Plus*. doi:10.1128/ecosalplus.8.3.1.2

Conway T and Cohen PS. (2015) Commensal and pathogenic *Escherichia coli* metabolism in the gut. *Metabolism and Bacterial Pathogenesis*. doi: 10.1128/microbiolspec.MBP-0006-2014

Cook SM and McArthur JD. (2013) Developing *Galleria mellonella* as a model host for human pathogens. *Virulence*. doi: 10.4161/viru.25240

Cox G and Wright GD. (2013) Intrinsic antibiotic resistance: Mechanisms, origins, challenges and solutions. *International Journal of Medical Microbiology*. doi: 10.1016/j.ijmm.2013.02.009

Croucher NJ, Harris SR, Fraser C, Quail MA, Burton J, Van Der Linden M, McGee L, von Gottberg A, Song JH, Ko KS, Pichon B, Baker S, Parry CM Lambersten LM, Shahinas D, Pillai DR, Michell TJ, Dougan G, Tomasz A, Klugman KP, Parkhill J, Hanage WP and Bentley SD. (2011) Rapid pneumococcal evolution in response to clinical interventions. *Science*. doi: 10.1126/science.1198545

Croxen MA, Law RJ, Scholz R, Keeney KM, Wlodarska M and Finlay BB. (2013) Recent advances in understanding enteric pathogenic *Escherichia coli*. *Clinical Microbiology Reviews*. doi: 10.1128/CMR.00022-13

Croxen MA and Finlay BB. (2009) Molecular mechanisms of *Escherichia coli* pathogenicity. *Nature Reviews Microbiology* 8: 26-38. doi: 10.1038/nrmicro2265

Croze OA, Ferguson GP, Cates ME and Poon WCK. (2011) Migration of chemotactic bacteria in soft agar: role of gel concentration. *Biophysical Journal*. doi: 10.1016/j.bpj.2011.06.023

Johnson CN, Briles DE, Benjamin WH. Jr, Hollingshead SK and Waites KB. (2005) Relative fitness of fluoroquinolone-resistant *Streptococcus pneumoniae*. *Emerging Infectious Diseases* 11(6): 814-820.

Cukras AR, Southworth DR, Brunelle JL, Culver GM and Green R. (2003) Ribosomal proteins S12 and S13 function as control elements for translocation of the mRNA:tRNA complex. *Molecular Cell* 12(2): 321-8.

Czworkowski J, Wang J, Steitz TA and Moore PB. (1994) The crystal structure of elongation factor G complexed with GDP, at 2.7 A resolution. *The EMBO Journal* 13(16): 3661-3668.

D'Agata EMC, Dupont-Rouzeyrol M, Magal P, Olivier P and Ruan S. (2009) The impact of different antibiotic regimens on the emergence of antimicrobial-resistant bacteria. *PLOS ONE* 3(12): 1–10. doi: 10.1371/journal.pone.0004036

D'Agata EMC, Magal P, Olivier D, Ruan S and Webb GF. (2007) Modeling antibiotic resistance in hospitals: The impact of minimizing treatment duration. *Journal of Theoretical Biology* 249(3):487-499. doi: 10.1016/j.jtbi.2007.08.011

Darmon E and Leach DRF. (2014) Bacterial genome instability. *Microbiology and Molecular Biology Reviews*. doi: 10.1128/MMBR.00035-13

Darwin CR. (1859) On the origin of species by means of natural selection. Available at: http://www.vliz.be/docs/Zeecijfers/Origin_of_Species.pdf

Davies JE. (1997) Origins, acquisition and dissemination of antibiotic resistance determinants. *Ciba Foundation Symposium* 207: 15–27.

Davies J. (2006) Where have all the antibiotics gone? *Canadian Journal of Infectious Diseases and Medical Microbiology* 17(5): 287-290.

Davies J and Davies D. (2010) Origins and evolution of antibiotic resistance. *Microbiology and Molecular Biology Reviews*. doi: 10.1128/MMBR.00016-10

Davies J and Davis BD. (1968) Misreading of ribonucleic acid code words induced by aminoglycoside antibiotics the effect of drug concentration. *The Journal of Biological Chemistry* 243 (12): 3312-6.

Davis BD, Luger SM and Tai PC. (1986) Role of ribosome degradation in the death of starved *Escherichia coli* cells. *Journal of Bacteriology*. doi: 10.1128/jb.166.2.439-445.1986

D'Costa VM, McGrann KM, Hughes DW and Wright GD. (2006) Sampling the antibiotic resistome. *Science 311*(5759): 374-7. doi: 10.1126/science.1120800

Deatherage DE, Kepner JL, Bennett AF, Lenski RE and Barrick JE. (2017) Specificity of genome evolution in experimental populations of *Escherichia coli* evolved at different temperatures. *Proceedings of the National Academy of Sciences*. doi: 10.1073/pnas.1616132114

Del Fabbro C, Scalabrin S, Morgante M and Giorgi FM. (2013) An extensive evaluation of read trimming effects on Illumina NGS data analysis. *PLOS ONE* 8(12). doi: 10.1371/journal.pone.0085024

Demerec M. (1945) Production of *Staphylococcus* strains resistant to various concentrations of penicillin. *Proceedings of the National Academy of Sciences* 31(1):16-24.

Demirci H, Murphy F, Murphy E, Gregory ST, Dahlberg AE and Jogl G. (2013) A structural basis for streptomycin-induced misreading of the genetic code. *Nature Communications*. doi: 10.1038/ncomms2346

Denamur E and Matic I. (2006) Evolution of mutation rates in bacteria. *Molecular Microbiology*. doi: 10.1111/j.1365-2958.2006.05150.x

DePristo MA, Banks E, Poplin R, Garimella KV, Maguire JR, Hartl C, Philippakis AA, del Angel G, Rivas MA, Hanna M, McKenna A, Fennell TJ Kernytsky AM, Sivachenko AY, Cibulskis K, Gabriel SB, Altshuler D and Daly MJ. (2011) A framework for variation discovery and genotyping using next-generation DNA sequencing data. *Nature Genetics*. doi: 10.1038/ng.806

Julian P, Konevega AL, Scheres SHW, Lazaro M, Gil D, Wintermeyer W, Rodnina MV and Valle M. (2008) Structure of ratcheted ribosomes with tRNAs in hybrid states. *Proceedings of the National Academy of Sciences* 105(44): 16924-27.

Darriba D, Taboada GL, Doallo R and Posada D. (2012) jModelTest 2: more models, new heuristics and parallel computing. *Nature Methods* 9(8), 772.

Dettman JR, Rodrigue N, Melnyk AH, Wong A, Bailey SF and Kassen R. (2012) Evolutionary insight from whole-genome sequencing of experimentally evolved microbes. *Molecular Ecology*. doi: 10.1111/j.1365-294X.2012.05484.x

DeWilde M and Wittmann-Liebold B. (1973) Localization of the amino-acid exchange in protein S5 from an *Escherichia coli* mutant resistant to spectinomycin. *Molecular and General Genetics* 127(3): 273–276.

Didelot X and Falush D. (2007) Inference of bacterial microevolution using multilocus sequence data. *Genetics*. doi: 10.1534/genetics.106.063305

Ding Y, Chi YC and Lawrence CE. (2005) RNA secondary structure prediction by centroids in a Boltzmann weighted ensemble. *RNA*. doi: 10.1261/rna.2500605

Djinovic-Carugo K and Carugo O. (2015) Missing strings of residues in protein crystal structures. *Intrinsically Disordered Proteins*. doi: 10.1080/21690707.2015.1095697

Doi Y, Wachino JI and Arakawa Y. (2016) Aminoglycoside resistance: the emergence of acquired 16S ribosomal RNA methyltransferases. *Infectious Disease Clinics of North America*. doi: 10.1016/j.idc.2016.02.011

Domingues S, da Silva GJ and Nielsen KM. (2012) Integrons vehicles and pathways for horizontal dissemination in bacteria. *Mobile Genetic Elements*. doi: 10.1016/B978-0-12-374984-0.00804-4

Dong Y, Zhao X, Domagala J and Drlica K. (1999) Effect of fluoroquinolone concentration on selection of resistant mutants of *Mycobacterium bovis* BCG and *Staphylococcus aureus*. *Antimicrobial Agents and Chemotherapy* 43(7): 1756-8.

Douarche C, Buguin A, Salman H, and Libchaber A. (2009). *E. Coli* and oxygen: a motility transition. *Physical Review Letters* 102(19). doi: 10.1103/PhysRevLett.102.198101

Dragon F, Payant C and Brakier-Gingras L. (1994) Mutational and structural analysis of the RNA binding site for *Escherichia coli* ribosomal protein S7. *Journal of Molecular Biology 244*(1): 74–85. doi: 10.1006/jmbi.1994.1705

Drake JW, Charlesworth B, Charlesworth D and Crow JF. (1998) Rates of spontaneous mutation. *Genetics* 148(4): 1667-86.

Dragosits M and Mattanovich D. (2013) Adaptive laboratory evolution - principles and applications for biotechnology. *Microbial Cell Factories*. doi: 10.1186/1475-2859-12-64

Drlica K. (2003) The mutant selection window and antimicrobial resistance. *Journal of Antimicrobial Chemotherapy*. doi: 10.1093/jac/dkg269

Drlica K, Malik M, Kerns RJ and Zhao X. (2008) Quinolone-mediated bacterial death. *Antimicrobial Agents and Chemotherapy*. doi: 10.1128/AAC.01617-06

Duriez P, Clermont O, Bonacorsi S, Bingen E, Chaventre A, Elion J, Picard B and Denamur E. (2001) Commensal *Escherichia coli* isolates are phylogenetically distributed among geographically distinct human populations. *Microbiology* 147(6): 1671-6.

Dykhuizen DE. (1990) Experimental Studies of Natural Selection in Bacteria. *Annual Review of Ecology and Systematics* 21(1). doi: 10.1146/annurev.es.21.110190.002105

Dykhuizen DE and Hartl DL. (1983) Selection in chemostats. *Microbiology Reviews* 47(2) 150-168.

Eagle H, Fleischman R and Levy M. (1952) Development of increased bacterial resistance to antibiotics I. Continuous spectrum of resistance to penicillin, chloramphenicol and streptomycin. *Journal of Bacteriology* 63(5): 623-38.

Elena SF and Lenski RE. (2003) Evolution experiments with microorganisms: the dynamics and genetic bases of adaptation. *Nature Reviews Genetics* 4: 457-469. doi: 10.1038/nrg1088

Ellwood M and Nomura M. (1982) Chromosomal locations of the genes for rRNA in *Escherichia coli* K-12. *Journal of Bacteriology* 149(2): 458-468.

Engberg J, Aarestrup FM, Taylor DE, Gerner-Smidt P and Nachamkin I. (2001) Quinolone and macrolide resistance in *Campylobacter jejuni* and *C. coli*: resistance mechanisms and trends in human isolates. *Emerging Infectious Diseases*. doi: 10.3201/eid0701.010104

Enne VI, Bennett PM, Livermore DM and Hall LMC. (2004). Enhancement of host fitness by the sul2-coding plasmid p9123 in the absence of selective pressure. *Journal of Antimicrobial Chemotherapy*. doi: 10.1093/jac/dkh217

Ermolenko DN and Noller HF. (2011) MRNA translocation occurs during the second step of ribosomal intersubunit rotation. *Nature Structural and Molecular Biology*. doi: 10.1038/nsmb.2011

Ermolenko DN, Majumdar ZK, Hickerson RP, Spiegel PC, Clegg RM and Noller, HF. (2007) Observation of Intersubunit Movement of the Ribosome in Solution Using FRET. *Journal of Molecular Biology* 370(3): 530–540. doi: 10.1016/j.jmb.2007.04.042

Escherich T. (1988) The intestinal bacteria of the neonate and breast-fed infant:1884. *Reviews of Infectious Diseases* 10: 1220–25.

Exner M, Bhattacharya S and Christiansen B. (2017) Antibiotic resistance: what is so special about multidrug-resistant Gram-negative bacteria*? GMS Hygiene and Infection Control*. doi: 10.3205/dgkh000290

Falagas ME and Kasiakou SK. (2005) Colistin: the revival of polymyxins for the management of multidrug-resistant Gram-negative bacterial infections. *The Pediatric Infectious Disease Journal*. doi: 10.1097/01.inf.0000174577.97635.7b

Feldman MB, Terry DS, Altman RB and Blanchard SC. (2010) Aminoglycoside activity observed on single pre-translocation ribosome complexes. *Nature Chemical Biology*. doi: 10.1038/nchembio.274

Felmingham D. (2002) The need for antimicrobial resistance surveillance. *The Journal of Antimicrobial Chemotherapy* 50(1): 1-7.

Feuerriegel S, Köser CU and Niemann S. (2014) Phylogenetic polymorphisms in antibiotic resistance genes of the *mycobacterium tuberculosis* complex. *Journal of Antimicrobial Chemotherapy*. doi: 10.1093/jac/dkt535

Filliol I, Motiwala AS, Cavatore M, Qi W, Hazbón MH, del Valle MB, Fyfe J, Garcia-Garcia L, Rastogi N, Sola C, Zozio T, Guerrero MI, Leon CI, Crabtree J, Angiuoli S, Eisenach KD, Durmaz R, Joloba ML, Rendon A, Siguentes-Osornio J, Ponce de Leon A, Cave MD, Fleischmann R, Whittam TS and Alland D. (2006) Global phylogeny of *Mycobacterium tuberculosis* based on single nucleotide polymorphism (SNP) analysis: insights into tuberculosis evolution, phylogenetic accuracy of other DNA fingerprinting systems, and recommendations for a minimal standard SNP set. *Journal of Bacteriology*. doi: 10.1128/JB.188.2.759-772.2006

Firnberg E, Labonte JW, Gray JJ and Ostermeier M. (2014) A comprehensive, high-resolution map of a gene's fitness landscape. *Molecular Biology and Evolution*, *31*(6): 1581–1592. doi: 10.1093/molbev/msu081

Fong SS, Joyce AR and Palsson BØ. (2005) Parallel adaptive evolution cultures of *Escherichia coli* lead to convergent growth phenotypes with different gene expression states. *Genome Research*. doi: 10.1101/gr.3832305

Foster PL. (2000). Adaptive mutation: implications for evolution. *Bioessays*. 22(12): 1067-1074.

Foster PL, Lee H, Popodi E, Townes JP and Tang H. (2015) Determinants of spontaneous mutation in the bacterium *Escherichia coli* as revealed by whole-genome sequencing. *Proceedings of the National Academy of Sciences*. doi: 10.1073/pnas.1512136112

Foster TJ. (2017) Antibiotic resistance in Staphylococcus aureus. Current status and future prospects. *FEMS Microbiology Reviews* 41(3): 430-449.

Fournier PE, Dubourg G and Raoult D. (2014). Clinical detection and characterization of bacterial pathogens in the genomics era. *Genome Medicine*. doi: 10.1186/s13073-014-0114-2

Fox EJ, Reid-Bayliss KS, Emond MJ and Loeb LA. (2014) Accuracy of next generation sequencing platforms. *Journal of Next Generation Sequencing and Applications*. doi: 10.4172/jngsa.1000106

Frank J and Agrawal RK. (2000) A ratchet-like inter-subunit reorganization of the ribosome during translocation. *Nature* 406(6793):318-22. doi: 10.1038/35018597

Fredrick K and Noller HF. (2003) Catalysis of ribosomal translocation by sparsomycin. *Science* 300(5622): 1159–62. doi: 10.1126/science.1084571

Friedman L, Alder JD and Silverman JA. (2006) Genetic changes that correlate with reduced susceptibility to daptomycin in *Staphylococcus aureus*. *Antimicrobial Agents and Chemotherapy*. doi: 10.1128/AAC.00039-06

Funatsu G and Wittmann HG. (1972) Ribosomal proteins: XXXIII. Location of amino-acid replacements in protein S12 isolated from *Escherichia coli* mutants resistant to streptomycin. *Journal of Molecular Biology* 68(3): 547–550. doi: doi: 10.1016/0022-2836(72)90108-8

Furusawa C. (2017). Analysis of drug resistance using experimental evolution. *The Pharmaceutical Society of Japan* 137(4): 373–376. doi: 10.1248/yakushi.16-00235-1

Furusawa C, Horinouchi T and Maeda T. (2018) Toward prediction and control of antibiotic-resistance evolution. *Current Opinion in Biotechnology*. doi: 10.1016/j.copbio.2018.01.026

G20 Germany. (2017) Berlin Declaration of the G20 Health Ministers Together Today for a Healthy Tomorrow. Available at: https://www.bundesgesundheitsministerium.de/fileadmin/Dateien/3_Downloads/G/G20-Gesundheitsministertreffen/G20_Health_Ministers_Declaration_engl.pdf

Gagnon MG, Seetharaman SV, Bulkley D and Steitz TA. (2012) Structural basis for the rescue of stalled ribosomes: Structure of YaeJ bound to the ribosome. *Science*. doi: 10.1126/science.1217443

Galhardo RS, Hastings PJ and Rosenberg SM. (2007) Mutation as a stress response and the regulation of evolvability. *Critical Reviews in Biochemistry and Molecular Biology*. doi: 10.1080/10409230701648502

Gao YG, Selmer M, Dunham CM, Weixlbaumer A, Kelley AC and Ramakrishnan V. (2009) The structure of the ribosome with elongation factor G trapped in the post translocational state. *Science*. doi: 10.1126/science.1179709

Garneau-Tsodikova S and Labby KJ. (2016) Mechanisms of resistance to aminoglycoside antibiotics: Overview and perspectives. MedChemComm. doi: 10.1039/c5md00344j

Gaucher EA, Miyamoto MM and Benner SA. (2000) Function-structure analysis of proteins using covarion-based evolutionary approaches: Elongation factors. *Proceedings of the National Academy of Sciences* 98(2): 548-552.

Gekko K, Kamiyama T, Ohmae E and Katayanagi K. (2000) Single amino acid substitutions in flexible loops can induce large compressibility changes in dihydrofolate reductase. *The Journal of Biochemistry* 128(1): 21-7.

Gekko K, Obu N, Li J and Lee JC. (2004) A linear correlation between the energetics of allosteric communication and protein flexibility in the *Escherichia coli* cyclic AMP receptor protein revealed by mutation-induced changes in compressibility and amide hydrogen−deuterium exchange. *Biochemistry* 43(13): 3844–3852. doi: 10.1021/bi036271e

Geli P, Laxminarayan R, Dunne M and Smith DL. (2012) “one-size-fits-all”? optimizing treatment duration for bacterial infections. *PLOS ONE*. doi: 10.1371/journal.pone.0029838

Germovsek E, Barker CI and Sharland M. (2017) What do I need to know about aminoglycoside antibiotics? *Archives of Disease in Childhood - Education and Practice* 102(2): 89–93. doi: 10.1136/archdischild-2015-309069

Ghaddar N, Hashemidahaj M and Findlay BL. (2018) Access to high-impact mutations constrains the evolution of antibiotic resistance in soft agar. *Scientific Reports*. doi: 10.1038/s41598-018-34911-9

Gillon D. (1984) Stability of bidimensional stellar systems for axisymmetric modes. *Astrophysics and Space Science*. doi: 10.1093/aje/kwt017

Giraud A, Matic I, Tenaillon O, Clara A, Radman M, Fons M and Taddei F. (2001) Costs and benefits of high mutation rates: adaptive evolution of bacteria in the mouse gut. *Science* 291(5513): 2606–2608. doi: 10.1126/science.1056421

Golding CG, Lamboo LL, Beniac DR and Booth TF. (2016) The scanning electron microscope in microbiology and diagnosis of infectious disease. *Scientific Reports*. doi: 10.1038/srep26516

Goodarzi H, Hottes AK and Tavazoie S. (2009) Global discovery of adaptive mutations. *Nature Methods*. doi: 10.1038/nmeth.1352

Gourse RL, Gaal T, Bartlett MS, Appleman JA and Ross W. (1996) rRNA transcription and growth rate–dependent regulation of ribosome synthesis in *Escherichia coli*. *Annual Review of Microbiology* 50(1): 645–677. doi: 10.1146/annurev.micro.50.1.645

Gram HC. (1884) Über die isolierte Färbung der Schizomyceten in Schnitt- und Trockenpräparaten. *Fortschritte Der Medizin* 2: 185–189.

Gregory ST, Carr JF and Dahlberg AE. (2009) A signal relay between ribosomal protein S12 and elongation factor EF-Tu during decoding of mRNA. *RNA*. doi: 10.1261/rna.1355709

Gresham D and Hong J. (2015) The functional basis of adaptive evolution in chemostats. *FEMS Microbiology Reviews*. doi: 10.1111/1574-6976.12082

Grondek JF and Culver GM. (2004) Assembly of the 30S ribosomal subunit: positioning ribosomal protein S13 in the S7 assembly branch. *RNA*. doi: 10.1261/rna.7130504

Grundmann H and Hellriegel B. (2006) Mathematical modelling: a tool for hospital infection control. *The Lancet Infectious Diseases* 6(1): 39–45. doi: 10.1016/S1473-3099(05)70325-X

Guindon S and Gascuel O. (2003) A simple, fast and accurate method to estimate large phylogenies by maximum-likelihood. *Systematic Biology* 52: 696-704.

Guindon S, Dufayard JF, Lefort V, Anisimova M, Hordijk W and Gascuel O. (2010) New Algorithms and Methods to Estimate Maximum-Likelihood Phylogenies: Assessing the Performance of PhyML 3.0. *Systematic Biology* 59(3): 307-321.

Gullapelli K, Brahmeshwari G, Ravichander M and Kusuma U. (2017) Synthesis, antibacterial and molecular docking studies of new benzimidazole derivatives. *Egyptian Journal of Basic and Applied Sciences*. doi: 10.1016/j.ejbas.2017.09.002

Gullberg E, Cao S, Berg OG, Ilbäck C, Sandegren L, Hughes D and Andersson DI. (2011) Selection of resistant bacteria at very low antibiotic concentrations. *PLoS Pathogens*. doi: 10.1371/journal.ppat.1002158

Guo Z and Noller HF. (2012) Rotation of the head of the 30S ribosomal subunit during mRNA translocation. *Proceedings of the National Academy of Sciences*. doi: 10.1073/pnas.1218999109

Gutierrez B, Escudero JA, Millan AS, Hidalgo L, Carrilero L, Ovejero CM, Santa-Lopez A, Thomas-Lopez A and Gonzalez-Zorn B. (2012) Fitness cost and interference of Arm/Rmt aminoglycoside resistance with the RsmF housekeeping methyltransferases. *Antimicrobial Agents and Chemotherapy*. doi: 10.1128/AAC.06066-11

Guttman DS and Dykhuizen DE. (1994) Detecting selective sweeps in naturally occurring *Escherichia coli*. *Genetics* 138(4): 993-1003.

Ha T. (2001) Single-molecule fluorescence resonance energy transfer. *Methods*. doi: 10.1006/meth.2001.1217

Hall BG. (2004) Predicting the evolution of antibiotic resistance genes. *Nature Reviews Microbiology* 2: 430–435.

Hancock RE, Raffle VJ and Nicas TI. (1981). Involvement of the outer membrane in gentamicin and streptomycin uptake and killing in *Pseudomonas aeruginosa*. *Antimicrobial Agents and Chemotherapy*. doi: 10.1128/AAC.19.5.777

Harmand N, Gallet R, Martin G and Lenormand T. (2018) Evolution of bacteria specialization along an antibiotic dose gradient. doi: 10.1002/evl3.52

Harris SR, Feil EJ, Holden MT, Quail MA, Nickerson EK, Chantratita N, Gardete S, Tavares A, Day N, Lindsay JA, Edgeworth JD, de Lencastre H, Parkhill J, Peacock SJ and Bentley SD. (2010) Evolution of MRSA during hospital transmission and intercontinental spread. *Science*. doi: 10.1126/science.1182395

Harrison E, Wood JA, Dytham C, Pitchford JW, Truman J, Spiers A, Paterson S and Brockhurst MA. (2015) Bacteriophages limit the existence conditions for conjugative plasmids. *MBio*. doi: 10.1128/mBio.00586-15

Harshey RM. (2003) Bacterial motility on a Surface: many ways to a common goal. *Annual Review of Microbiology*. doi: 10.1146/annurev.micro.57.030502.091014

Harshey RM and Matsuyamat T. (1994) Dimorphic transition in *Escherichia coli* and *Salmonella typhimurium:* surface-induced differentiation into hyperflagellate swarmer cells. *Developmental Biology* 91(18): 8631-8635.

Hartmann M, Berditsch M, Hawecker J, Ardakani MF, Gerthsen D and Ulrich AS. (2010) Damage of the bacterial cell envelope by antimicrobial peptides gramicidin S and PGLa as revealed by transmission and scanning electron microscopy. *Antimicrobial Agents and Chemotherapy*. doi: 10.1128/AAC.00124-10

Hawkey PM and Jones AM. (2009) The changing epidemiology of resistance. *Journal of Antimicrobial Chemotherapy*. doi: 10.1093/jac/dkp256

Hazbón MH, Motiwala AS, Cavatore M, Brimacombe M, Whittam TS and Alland D. (2008) Convergent evolutionary analysis identifies significant mutations in drug resistance targets of *Mycobacterium tuberculosis*. *Antimicrobial Agents and Chemotherapy*. doi: 10.1128/AAC.00309-08

Heermann R, Zeppenfeld T and Jung K. (2008) Simple generation of site-directed point mutations in the Escherichia coli chromosome using Red®/ET® recombination. *Microbial Cell Factories*. doi: 10.1186/1475-2859-7-14

Hegreness M, Shoresh N, Damian D, Hartl D and Kishony R. (2008) Accelerated evolution of resistance in multidrug environments. *Proceedings of the National Academy of Sciences*. doi: 10.1073/pnas.0805965105

Hegreness M and Kishony R. (2007) Analysis of genetic systems using experimental evolution and whole-genome sequencing. *Genome Biology*. doi: 10.1186/gb-2007-8-1-201

Herring CD, Raghunathan A, Honisch C, Patel T, Applebee MK, Joyce AR, Albert TJ, Blattner FR, van den Boom D, Cantor CR and Palsson BØ. (2006) Comparative genome sequencing of *Escherichia coli* allows observation of bacterial evolution on a laboratory timescale. *Nature Genetics 38*: 1406. Retrieved from doi: 10.1038/ng1906

Holberger LE and Hayes CS. (2009) Ribosomal protein S12 and aminoglycoside antibiotics modulate A-site mRNA cleavage and transfer-messenger RNA activity in *Escherichia coli*. *Journal of Biological Chemistry*. doi: 10.1074/jbc.M109.062745

Holmes KL and Culver GM. (2005) Analysis of conformational changes in 16S rRNA during the course of 30S subunit assembly. *Journal of Molecular Biology*. doi: 10.1016/j.jmb.2005.09.056

Hollis A and Ahmed Z. (2013) Preserving antibiotics rationally. *New England Journal of Medicine*. doi: 10.1056/NEJMp1312654

Holtkamp W, Cunha CE, Peske F, Konevega AL, Wintermeyer W and Rodnina MV. (2014) GTP hydrolysis by EF-G synchronizes tRNA movement on small and large ribosomal subunits. *EMBO Journal*. doi: 10.1002/embj.201387465

Hughes D and Andersson DI. (2012) Selection of resistance at lethal and non-lethal antibiotic concentrations. *Current Opinion in Microbiology*. doi: 10.1016/j.mib.2012.07.005

Huijben S, Sim DG, Nelson WA and Read AF. (2011) The fitness of drug-resistant malaria parasites in a rodent model: multiplicity of infection. *Journal of Evolutionary Biology*. doi: 10.1111/j.1420-9101.2011.02369.x

Hunt R, Sauna ZE, Ambudkar SV, Gottesman MM and Kimchi-Sarfaty C. (2009) Silent (synonymous) SNPs: should we care about them? *Methods in Molecular Biology* 578: 23–39. doi: 10.1007/978-1-60327-411-1_2

Huse SM, Huber JA, Morrison HG, Sogin ML and Welch DM. (2007) Accuracy and quality of massively parallel DNA pyrosequencing. *Genome Biology* 8(7): 143. doi: 10.1186/gb-2007-8-7-r143

Iannini PB, Ehret J and Eickhoff TC. (1976) Effects of ampicillin amikacin and ampicillin rifampin on *enterococci*. *Antimicrobial Agents and Chemotherapy*. doi: 10.1128/AAC.9.3.448

Ibacache-Quiroga C, Oliveros JC, Couce A and Blázquez J. (2018) Parallel Evolution of High-Level Aminoglycoside Resistance in *Escherichia coli* Under Low and High Mutation Supply Rates. *Frontiers in Microbiology* 9: 427. doi: 10.3389/fmicb.2018.00427

Isabel S, Leblanc É, Boissinot M, Boudreau DK, Grondin M, Picard FJ, Martel EA, Parham NJ, Chain PS, Bader DE, Mulvey MR, Bryden L, Roy PH, Ouellette M and Bergeron MG. (2008). Divergence among genes encoding the elongation factor Tu of *Yersinia* species. *Journal of Bacteriology*. doi: 10.1128/JB.01067-08

Jahn LJ, Munck C, Ellabaan MMH and Sommer MOA. (2017) Adaptive laboratory evolution of antibiotic resistance using different selection regimes lead to similar phenotypes and genotypes. *Frontiers in Microbiology* 8: 816. doi: 10.3389/fmicb.2017.00816

Janda JM and Abbott SL. (2007) 16S rRNA gene sequencing for bacterial identification in the diagnostic laboratory: Pluses, perils, and pitfalls. *Journal of Clinical Microbiology*. doi: 10.1128/JCM.01228-07

Jander G, Rahme LG and Ausubel FM. (2000) Positive correlation between virulence of *Pseudomonas aeruginosa* mutants in mice and insects. *Journal of Bacteriology* 182(13): 3843-5.

Jansen G, Barbosa C and Schulenburg H. (2013) Experimental evolution as an efficient tool to dissect adaptive paths to antibiotic resistance. *Drug Resistance Updates* 16(6): 96–107. doi: 10.1016/j.drup.2014.02.002

Johnsen PJ, Townsend JP, Bøhn T, Simonsen GS, Sundsfjord A, Nielsen KM. (2009) Factors affecting the reversal of antimicrobial-drug resistance. *The Lancet Infectious Diseases* 9(6): 357–64.

Johnson JR, Gajewski A, Lesse AJ and Russo TA. (2003) Extraintestinal pathogenic *Escherichia coli* as a cause of invasive nonurinary infections. *Journal of Clinical Microbiology.* doi: 10.1128/JCM.41.12.5798-5802.2003

Johnson JR and Russo TA. (2002) Extraintestinal pathogenic *Escherichia coli*: "The other bad *E. coli*." *The Journal of Laboratory and Clinical Medicine* 139(3): 155–162. doi: 10.1067/mlc.2002.121550

Johnson RC. (2015) Site-specific DNA inversion by serine recombinases. *Microbiology Spectrum* 3(3): 1–36. doi: 10.1128/microbiolspec.MDNA3-0047-2014

Johnston C, Martin B, Fichant G, Polard P and Claverys JP. (2014) Bacterial transformation: distribution, shared mechanisms and divergent control. *Nature Reviews Microbiology* 12(3):181-96. doi: 10.1038/nrmicro3199

Jordan IK, Rogozin IB, Wolf YI and Koonin EV. (2002) Essential genes are more evolutionarily conserved than are nonessential genes in bacteria. *Genome Research*. doi: 10.1101/gr.87702

Joseph S, Sonbol H, Hariri S, Desai P, McClelland M and Forsythe SJ. (2012) Diversity of the *Cronobacter* genus as revealed by multilocus sequence typing. *Journal of Clinical Microbiology*. doi: 10.1128/JCM.00905-12

Kabsch W. (1993) Automatic processing of rotation diffraction data from crystals of initially unknown symmetry and cell constants. *Journal of Applied Crystallography* 40, 658-674.

Kaimer C, González-Pastor JE and Graumann PL. (2009) SpoIIIE and a novel type of DNA translocase, SftA, couple chromosome segregation with cell division in *Bacillus subtilis*. *Molecular Microbiology*. doi: 10.1111/j.1365-2958.2009.06894.x

Kamla V, Henrich B and Hadding U. (1996) Phylogeny based on elongation factor Tu reflects the phenotypic features of mycoplasmas better than that based on 16S rRNA. *Gene*. doi: 10.1016/0378-1119(95)00884-5

Kaper JB, Nataro JP and Mobley HLT. (2004) Pathogenic *Escherichia coli*. *Nature Reviews Microbiology* 2: 123-140. doi: 10.1038/nrmicro818

Karki R, Pandya D, Elston RC and Ferlini C. (2015) Defining "mutation" and "polymorphism" in the era of personal genomics. *BMC Medical Genomics*. doi: 10.1186/s12920-015-0115-z

Kawecki TJ, Lenski RE, Ebert D, Hollis B, Olivieri I and Whitlock MC. (2012) Experimental evolution. *Trends in Ecology and Evolution*. doi: 10.1016/j.tree.2012.06.001

Kennemann L, Didelot X, Aebischer T, Kuhn S, Drescher B, Droege M, Reinhardt R, Correa P, Meyer TF, Josenhans C, Falush D and Suerbaum S. (2011) *Helicobacter pylori* genome evolution during human infection. *Proceedings of the National Academy of Sciences*. doi: 10.1073/pnas.1018444108

Kidera A, Konishi Y, Oka M, Ooi T and Scheraga HA. (1985) Statistical analysis of the physical properties of the 20 naturally occurring amino acids. *Journal of Protein Chemistry* 4(1): 23–55.

Kim HD, Puglisi JD and Chu S. (2007) Fluctuations of transfer RNAs between classical and hybrid states. *Biophysical Journal*. doi: 10.1529/biophysj.107.109884

Kim S, Lieberman TD and Kishony R. (2014) Alternating antibiotic treatments constrain evolutionary paths to multidrug resistance. *Proceedings of the National Academy of Sciences*. doi: 10.1073/pnas.1409800111

Kirthi N, Roy-Chaudhuri B, Kelley T and Culver GM. (2006) A novel single amino acid change in small subunit ribosomal protein S5 has profound effects on translational fidelity. *RNA 12*(12): 2080–2091.

Kiss A, Sain B and Venetianer P. (1977) The number of rRNA genes in *Escherichia coli*. *FEBS Letters*. doi: 10.1016/0014-5793(77)80354-2

Kitchel B, Rasheed JK, Patel JB, Srinivasan A, Navon-Venezia S, Carmeli Y, Brolund A and Giske CG. (2009) Molecular epidemiology of KPC-producing *Klebsiella pneumoniae* isolates in the United States: Clonal expansion of multilocus sequence type 258. *Antimicrobial Agents and Chemotherapy*. doi: 10.1128/AAC.00126-09

Klappenbach JA, Dunbar JM and Schmidt TM. (2000) rRNA operon copy number reflects ecological strategies of bacteria. *Applied and Environmental Microbiology* 66(4): 1328-33.

Klitgaard RN, Ntokou E, Nørgaard K, Biltoft D, Hansen LH, Trædholm NM, Kongsted J and Vester B. (2015) Mutations in the bacterial ribosomal protein l3 and their association with antibiotic resistance. *Antimicrobial Agents and Chemotherapy*. doi: 10.1128/AAC.00179-15

Knöppel A, Knopp M, Albrecht LM, Lundin E, Lustig U, Näsvall J and Andersson DI. (2018) Genetic adaptation to growth under laboratory conditions in *Escherichia coli* and *Salmonella enterica*. *Frontiers in Microbiology*. doi: 10.3389/fmicb.2018.00756

Koch AL. (1999) Diffusion through agar blocks of finite dimensions: a theoretical analysis of three systems of practical significance in microbiology. *Microbiology*. doi: 10.1099/13500872-145-3-643

Kohanski MA, Dwyer DJ and Collins JJ. (2010) How antibiotics kill bacteria: from targets to networks. *Nature Reviews Microbiology*. doi: 10.1038/nrmicro2333

Köser CU, Ellington MJ, Cartwright EJP, Gillespie SH, Brown NM, Farrington M, Holden MTG, Dougan G, Bentley SD, Parkhill J and Peacock SJ. (2012) Routine use of microbial whole genome sequencing in diagnostic and public health microbiology. *PLoS Pathogens*. doi: 10.1371/journal.ppat.1002824

Köser CU, Ellington MJ and Peacock SJ. (2014) Whole-genome sequencing to control antimicrobial resistance. *Trends in Genetics*. doi: 10.1016/j.tig.2014.07.003

Kotra LP, Haddad J and Mobashery S. (2000) Aminoglycosides: perspectives on mechanisms of action and resistance and strategies to counter resistance. *Antimicrobial Agents and Chemotherapy* 44(12): 3249-56. doi: 10.1128/AAC.44.12.3249-3256.2000

Kouyos RD, Metcalf CJE, Birger R, Klein EY, zur Wiesch PA, Ankomah P, Arinaminpathy N, Bogich TL, Bonhoeffer S, Brower C, Chi-Johnston G, Cohen T, Day T, Greenhouse B, Huijben S, Metlay J, Mideo N, Pollitt LC, Read AF, Smith DL, Standley C, Wale N and Grenfell B. (2014) The path of least resistance: Aggressive or moderate treatment? *Proceedings of the Royal Society B: Biological Sciences*. doi: 10.1098/rspb.2014.0566

Krause KM, Serio AW, Kane TR and Connolly LE. (2016) Aminoglycosides: an overview. *Cold Spring Harbor Perspectives in Medicine*. doi: 10.1101/cshperspect.a027029

Krist AC and Showsh SA. (2007) Experimental evolution of antibiotic resistance in bacteria. *The American Biology Teacher* 69(2): 94–98. doi: 10.1662/0002-7685

Kronvall G. (1983) Single-strain regression analysis for determination of interpretive breakpoints for cefoperazone disk diffusion susceptibility testing. *Journal of Clinical Microbiology* 17(6): 975-980.

Kugelberg E, Löfmark S, Wretlind B and Andersson DI. (2005) Reduction of the fitness burden of quinolone resistance in *Pseudomonas aeruginosa*. *Journal of Antimicrobial Chemotherapy*. doi: 10.1093/jac/dkh505

Kumana CR and Yuen KY. (1994) Parenteral aminoglycoside therapy: selection, administration and monitoring. *Drugs*. doi: 10.2165/00003495-199447060-00004

Kumar S, Banks TW and Cloutier S. (2012) SNP discovery through next-generation sequencing and its applications. *International Journal of Plant Genomics*. doi: 10.1155/2012/831460

Lacroix JM, Lanfroy E, Cogez V, Lequette Y, Bohin A and Bohin JP. (1999) The mdoC gene of *Escherichia coli* encodes a membrane protein that is required for succinylation of osmoregulated periplasmic glucans. *Journal of Bacteriology 181*(12): 3626–3631.

Laehnemann D, Peña-Miller R, Rosenstiel P, Beardmore R, Jansen G and Schulenburg H. (2014) Genomics of rapid adaptation to antibiotics: convergent evolution and scalable sequence amplification. *Genome Biology and Evolution*. doi: 10.1093/gbe/evu106

Lagier JC, Hugon P, Khelaifia S, Fournier PE, La Scola B and Raoult D. (2015) The rebirth of culture in microbiology through the example of culturomics to study human gut microbiota. *Clinical Microbiology Reviews*. doi: 10.1128/CMR.00014-14

Lallemand EA, Lacroix MZ, Toutain PL, Boullier S, Ferran AA and Bousquet-Melou A. (2016) *In vitro* degradation of antimicrobials during use of broth microdilution method can increase the measured minimal inhibitory and minimal bactericidal concentrations. *Frontiers in Microbiology*. doi: 10.3389/fmicb.2016.02051

Lambert T. (2012) Antibiotics that affect the ribosome. *Reviews Science Technology* 31(1): 57–64.

Lang GI and Desai MM. (2014) The spectrum of adaptive mutations in experimental evolution. *Genomics*. doi: 10.1016/j.ygeno.2014.09.011

Laurence M, Hatzis C and Brash DE. (2014) Common contaminants in next-generation sequencing that hinder discovery of low-abundance microbes. *PLOS ONE* 9(5). doi: 10.1371/journal.pone.0097876

Laursen RA and Duffy L. (1978) The evolution of elongation factors Tu and G by gene duplication. *FEBS Letters*. doi: 10.1016/0014-5793(78)80753-4

Lecároz C, Campanero MA, Gamazo C and Blanco-Prieto MJ. (2006) Determination of gentamicin in different matrices by a new sensitive high-performance liquid chromatography-mass spectrometric method. *Journal of Antimicrobial Chemotherapy*. doi: 10.1093/jac/dkl258

Leclerc D, Melancon P and Brakier-gingras L. (1991). Mutations in the 915 region of *Escherichia coli* 16S ribosomal RNA reduce the binding of streptomycin to the ribosome. *Nucleic Acids Research*. doi: 10.1093/nar/19.14.3973

Lederberg J and Tatum El. (1946) Gene recombination in *Escherichia coli*. *Nature* 158(4016): 558.

Lederberg J and Lederberg EM. (1951) Replica plating and indirect selection of bacterial mutants. *Journal of bacteriology* 63(3): 399-406.

Lee A, Hong S and Kim D. (2018) KRDS: a web server for evaluating drug resistance mutations in kinases by molecular docking. *Journal of Cheminformatics*. doi: 10.1186/s13321-018-0274-y

Lee H, Popodi E, Tang H and Foster PL. (2012) Rate and molecular spectrum of spontaneous mutations in the bacterium *Escherichia coli* as determined by whole-genome sequencing. *Proceedings of the National Academy of Sciences*. doi: 10.1073/pnas.1210309109

Lee HH, Molla MN, Cantor CR and Collins JJ. (2010) Bacterial charity work leads to population-wide resistance. *Nature*. doi: 10.1038/nature09354

Lee K, Holland-Staley CA and Cunningham PR. (2001) Genetic approaches to studying protein synthesis: effects of mutations at 516 and A535 in *Escherichia coli* 16S rRNA. *American Society for Nutritional Sciences* 131(11): 2994-3004.

Lenski RE. (1998) Bacterial evolution and the cost of antibiotic resistance. *International Microbiology* 1(4): 265–270.

Lenski RE. (2010) Phenotypic and genomic evolution during a 20,000-generation experiment with the bacterium *Escherichia coli*. *Plant Breeding Reviews*. doi: 10.1002/9780470650288.ch8

Lenski RE, Rose MR, Simpson SC and Tadler SC. (1991) Long-term experimental evolution in *Escherichia coli*. I. Adaptation and divergence during 2,000 generations. *The American Naturalist* 138(6): 1315–1341. doi: 10.1086/285289

Lerminiaux NA and Cameron ADS. (2018) Horizontal transfer of antibiotic resistance genes in clinical environments. *Canadian Journal of Microbiology*. doi: 10.1139/cjm-2018-0275

Levin BR, Lipsitch M, Perrot V, Schrag S, Antia R, Simonsen L, Walker NM and Stewart FM. (1997) The population genetics of antibiotic resistance. *Clinical Infectious Diseases* 24(1): 9–16.

Levin BR, Perrot V and Walker N. (2000) Compensatory mutations, antibiotic resistance and the population genetics of adaptive evolution in bacteria. *Genetics*. doi: 10.1534/genetics.110.124628

Levin BR. (2001) Minimizing potential resistance: a population dynamics view. *Clinical Infectious Diseases*. doi: 10.1086/321843

Levin-Reisman I, Ronin I, Gefen O, Braniss I, Shoresh N and Balaban NQ. (2017) Antibiotic tolerance facilitates the evolution of resistance. *Science*. doi: 10.1126/science.aaj2191

Lequette Y, Ödberg-Ferragut C, Bohin JP and Lacroix JM (2004). Identification of mdoD, an mdoG paralog which encodes a twin-arginine-dependent periplasmic protein that controls osmoregulated periplasmic glucan backbone structures. *Journal of Bacteriology* 186(12): 3695–3702.

Levy SB and Marshall B. (2004) Antibacterial resistance worldwide: causes, challenges and responses. *Nature Medicine*. doi: 10.1038/nm1145

Li R, Li Y, Fang X, Yang H, Wang J, Kristiansen K and Wang J. (2009) SNP detection for massively parallel whole-genome resequencing. *Genome Research*. doi: 10.1101/gr.088013.108

Li T, Chen J, Chang D, Fang X, Wang J, Guo Y, Su L, Xu G, Wang Y, Chen Z and Liu C. (2013) Draft Genome Sequence of *Escherichia coli* Strain LCT-EC59. *Genome Announcements*. doi: 10.1128/genomeA.00049-13

Li W, Trabuco LG, Schulten K and Frank J. (2011) Molecular dynamics of EF-G during translocation. *Proteins: Structure, Function and Bioinformatics*. doi: 10.1002/prot.22976

Lieberman TD, Michel JB, Aingaran M, Potter-Bynoe G, Roux D, Davis MR Jr, Skurnik N, Leiby N, LiPuma JJ, Goldberg JB, McAdam AJ, Priebe GP and Kishony R. (2011) Parallel bacterial evolution within multiple patients identifies candidate pathogenicity genes. *Nature Genetics*. doi: 10.1038/ng.997

Liévin-Le Moal V and Servin AL. (2006) The front line of enteric host defense against unwelcome intrusion of harmful microorganisms: Mucins, antimicrobial peptides, and Microbiota. *Clinical Microbiology Reviews*. doi: 10.1128/CMR.19.2.315-337.2006

Lin J, Gagnon MG, Bulkley D and Steitz TA. (2015) Conformational changes of elongation factor g on the ribosome during tRNA translocation. *Cell*. doi: 10.1016/j.cell.2014.11.049

Lipsitch M and Levin BR. (1997) The Population Dynamics of Antimicrobial Chemotherapy. *Antimicrobial Agents and Chemotherapy* 41(2): 363-373.

Liu A, Tran L, Becket E, Lee K, Chinn L, Park E, Tran K and Miller JH. (2010) Antibiotic sensitivity profiles determined with an *Escherichia coli* gene knockout collection: Generating an antibiotic bar code. *Antimicrobial Agents and Chemotherapy*. doi: 10.1128/AAC.00906-09

Liu X and Matsumura P. (1994) The FlhD/FlhC complex, a transcriptional activator of the *Escherichia coli* flagellar class II operons. *Journal of Bacteriology* 176(23): 7345-51.

Livermore DM. (2011) Discovery research: the scientific challenge of finding new antibiotics. *Journal of Antimicrobial Chemotherapy*. doi: 10.1093/jac/dkr262

Livermore DM and Woodford N. (2006) The beta-lactamase threat in *Enterobacteriaceae*, *Pseudomonas* and *Acinetobacter*. *Trends in Microbiology* 14(9): 413–420. doi: 10.1016/j.tim.2006.07.008

Long H, Miller SF, Strauss C, Zhao C, Cheng L, Ye Z, Griffin K, Te R, Lee H, Chen CC and Lynch M. (2016) Antibiotic treatment enhances the genome-wide mutation rate of target cells. *Proceedings of the National Academy of Sciences*. doi: 10.1073/pnas.1601208113

Long KS, Poehlsgaard J, Hansen LH, Hobbie, SN, Böttger EC and Vester B. (2009) Single 23S rRNA mutations at the ribosomal peptidyl transferase centre confer resistance to valnemulin and other antibiotics in *Mycobacterium smegmatis* by perturbation of the drug binding pocket. *Molecular Microbiology*. doi: 10.1111/j.1365-2958.2009.06596.x

Lowy FD. (2003) Antimicrobial resistance: the example of *Staphylococcus aureus*. *Journal of Clinical Investigation*. doi: 10.1172/JCI18535

Lozovsky ER, Chookajorn T, Brown KM, Imwong M, Shaw PJ, Kamchonwongpaisan S, Neafsey DE, Weinreich DM and Hartl DL. (2009) Stepwise acquisition of pyrimethamine resistance in the malaria parasite. *Proceedings of the National Academy of Sciences*. doi: 10.1073/pnas.0905922106

Lukačišinová M and Bollenbach T. (2017) Toward a quantitative understanding of antibiotic resistance evolution. *Current Opinion in Biotechnology* 46: 90–97. doi: 10.1016/j.copbio.2017.02.013

Macvanin M, Ballagi A and Hughes D. (2004) Fusidic Acid-Resistant Mutants of *Salmonella enterica* Serovar Typhimurium Have Low Levels of Heme and a Reduced Rate of Respiration and Are Sensitive to Oxidative Stress. *Antimicrobial Agents and Chemotherapy* 48(10): 3877-3883.

Macvanin M and Hughes D. (2005) Hyper-susceptibility of a fusidic acid-resistant mutant of *Salmonella* to different classes of antibiotics. *FEMS Microbiology Letters* 247: 215-220.

Marchbanks CR, Yost RL and White RL. (1987) Cefotaxime stability during *in vitro* microbiological testing. *Antimicrobial Agents and Chemotherapy*. doi: 10.1128/AAC.31.9.1375

Mariam DH, Mengistu Y, Hoffner SE and Andersson DI. (2004) Effect of rpoB mutations conferring rifampin resistance on fitness of *Mycobacterium tuberculosis*. *Antimicrobial Agents and Chemotherapy*. doi: 10.1128/AAC.48.4.1289-1294.2004

Martinez JL. (2009) The role of natural environments in the evolution of resistance traits in pathogenic bacteria. *Proceedings of the Royal Society Biology: Biological Sciences* 276(1667): 2521–2530. doi: 10.1098/rspb.2009.0320

Marietou A, Nguyen ATT, Allen EE and Bartlett DH. (2015) Adaptive laboratory evolution of *Escherichia coli* K-12 MG1655 for growth at high hydrostatic pressure. *Frontiers in Microbiology*. doi: 10.3389/fmicb.2014.00749

Marini NJ, Thomas PD and Rine J. (2010) The use of orthologous sequences to predict the impact of amino acid substitutions on protein function. *PLoS Genetics*. doi: 10.1371/journal.pgen.1000968

Martemyanov KA and Gudkov AT. (2000) Domain III of elongation factor G from *Thermus thermophilus* is essential for induction of GTP hydrolysis on the ribosome. *Journal of Biological Chemistry*. doi: 10.1074/jbc.M002656200

Martin NL and Beveridge TJ. (1986) Gentamicin interaction with *Pseudomonas aeruginosa* cell envelope. *Antimicrobial Agents and Chemotherapy*. doi: 10.1128/AAC.29.6.1079

Martinez JL and Baquero F. (2000) Mutation Frequencies and Antibiotic Resistance. *Antimicrobial Agents and Chemotherapy* 44(7): 1771-7.

Matic I, Taddei F and Radman M. (2004) Survival versus maintenance of genetic stability: a conflict of priorities during stress. *Research in Microbiology*. doi: 10.1016/j.resmic.2004.01.010

Mavromatis K, Land ML, Brettin TS, Quest DJ, Copeland A, Clum A, … and Kyrpides NC. (2012) The fast changing landscape of sequencing technologies and their impact on microbial genome assemblies and annotation. *PLOS ONE*. doi: 10.1371/journal.pone.0048837

Maxam AM and Gilbert W. (1977) A new method for sequencing DNA. *Proceedings of the National Academy of Sciences* 74(2): 560-4.

McGlinchey TA, Rafter PA, Regan F and McMahon GP. (2008) A review of analytical methods for the determination of aminoglycoside and macrolide residues in food matrices. *Analytica Chimica Acta*. doi: 10.1016/j.aca.2008.05.054

Mehta P, Casjens S and Krishnaswamy S. (2004) Analysis of the lambdoid prophage element e14 in the *E. coli* K-12 genome. *BMC Microbiology* 4: 4. doi: 10.1186/1471-2180-4-4

Melnyk AH, Wong A and Kassen R. (2015) The fitness costs of antibiotic resistance mutations. *Evolutionary Applications*. doi: 10.1111/eva.12196

Meng XY, Zhang HX, Mezei M and Cui M. (2012) Molecular Docking: A powerful approach for structure-based drug discovery. *Current Computer Aided Drug Design* 7(2): 146-157.

Merhej V, Royer-Carenzi M, Pontarotti P and Raoult D. (2009) Massive comparative genomic analysis reveals convergent evolution of specialized bacteria. *Biology Direct*. doi: 10.1186/1745-6150-4-13

Metcalf CJE, Birger RB, Funk S, Kouyos RD, Lloyd-Smith JO and Jansen VAA. (2015) Five challenges in evolution and infectious diseases. *Epidemics*. doi: 10.1016/j.epidem.2014.12.003

Michel JB, Yeh PJ, Chait R, Moellering RC and Kishony R. (2008) Drug interactions modulate the potential for evolution of resistance. *Proceedings of the National Academy of Sciences*. doi: 10.1073/pnas.0800944105

Miller GH, Doukas PH and Seydel JK. (1972) Sulfonamide structure-activity relation in a cell-free system. Correlation of inhibition of folate synthesis with antibacterial activity and physicochemical parameters. *Journal of Medicinal Chemistry 15*(7): 700–706. doi: 10.1021/jm00277a002

Mingeot-Leclercq MP, Glupczynski Y and Tulkens PM. (1999) Aminoglycosides: activity and resistance. *Antimicrobial Agents and Chemotherapy*. doi: 10.1128/AAC.43.4.727

Moazed D and Noller HF. (1987) Interaction of antibiotics with functional sites in 16S ribosomal RNA. *Nature* 327(6121): 389-94. doi: 10.1038/327389a0

Mogre A, Sengupta T, Veetil RT, Ravi P and Seshasayee ASN. (2014) Genomic Analysis Reveals Distinct Concentration-Dependent Evolutionary Trajectories for Antibiotic Resistance in *Escherichia coli*. *DNA Research* 21, 711-726. doi: 10.1093/dnares/dsu032

Mogre A, Sengupta T, Veetil RT, Ravi P and Seshasayee ASN. (2017) Modulation of Global Transcriptional Regulatory Networks as a Strategy for Increasing Kanamycin Resistance of the Translational Elongation Factor-G Mutants in *Escherichia coli*. *Genes Genomes Genetics* 7(12): 3955-3966. doi: 10.1534/g3.117.300284

Morris GM and Lim-Wilby M. (2008) Molecular Docking. *Methods Molecular Biology* 443: 365-82. doi: https:doi.org/10.1007/978-1-59745-177-2_19

Morris GM, Huey R, Lindstrom W, Sanner MF, Belew RK, Goodsell DS and Olson AJ. (2009) Autodock4 and AutoDockTools4: automated docking with selective receptor flexibility. *Journal of Computational Chemistry* 16: 2785-91.

Munro JB, Altman RB, Tung CS, Sanbonmatsu KY and Blanchard SC. (2010) A fast dynamic mode of the EF-G-bound ribosome. *EMBO Journal*. doi: 10.1038/emboj.2009.384

Munro JB, Wasserman MR, Altman RB, Wang L and Blanchard SC. (2010) Correlated conformational events in EF-G and the ribosome regulate translocation. *Nature Structural and Molecular Biology*. doi: 10.1038/nsmb.1925

Murray HD, Schneider DA and Gourse RL. (2003) Control of rRNA expression by small molecules is dynamic and nonredundant. *Molecular Cell*. doi: 10.1016/S1097-2765(03)00266-1

Mwangi MM, Wu SW, Zhou Y, Sieradzki K, de Lencastre H, Richardson P, Bruce D, Rubin E, Myers E, Siggia ED and Tomasz A. (2007) Tracking the *in vivo* evolution of multidrug resistance in *Staphylococcus aureus* by whole-genome sequencing. *Proceedings of the National Academy of Sciences*. doi: 10.1073/pnas.0609839104

Nair J, Rouse DA, Bai GH and Morris SL. (1993) The rpsL gene and streptomycin resistance in single and multiple drug-resistant strains of *Mycobacterium tuberculosis*. *Molecular Microbiology* 10(3): 521–527. doi: 10.1111/j.1365-2958.1993.tb00924.x

Nessar R, Reyrat JM, Murray A and Gicquel B. (2011) Genetic analysis of new 16s rRNA mutations conferring aminoglycoside resistance in *Mycobacterium abscessus*. *Journal of Antimicrobial Chemotherapy*. doi: 10.1093/jac/dkr209

Ng PC and Henikoff S. (2003) SIFT: Predicting amino acid changes that affect protein function. *Nucleic Acids Research* 31(13): 3812–3814. doi: 10.1093/nar/gkg509

Nikaido H. (1994) Prevention of drug access to bacterial targets: permeability barriers and active efflux. *Science* 264(5157): 382-8. doi: 10.1126/science.8153625

Nikaido H. (1996) Multidrug efflux pumps of gram-negative bacteria. *Journal of Bacteriology*. doi: 10.1128/jb.178.20.5853-5859.1996

Nonaka L, Connell SR and Taylor DE. (2005) 16S rRNA mutations that confer tetracycline resistance in *Helicobacter pylori* decrease drug binding in *Escherichia coli* ribosomes. *Journal of Bacteriology*. doi: 10.1128/JB.187.11.3708-3712.2005

Norville IH, Hartley MG, Martinez E, Cantet F, Bonazzi M and Atkins TP. (2014) *Galleria mellonella* as an alternative model of *Coxiella burnetii* infection. *Microbiology*. doi: 10.1099/mic.0.077230-0

O'connor M, Göringer HU and Dahiberg AE. (1992) A ribosomal ambiguity mutation in the 530 loop of *E. coli* 16S rRNA. *Nucleic Acids Research*. doi: 10.1093/nar/20.16.4221

Ogle JM, Carter AP and Ramakrishnan V. (2003) Insights into the decoding mechanism from recent ribosome structures. *Trends Biochemical Sciences* 28(5): 259–266. doi: 10.1016/S0968-0004(03)00066-5

Olson ND, Zook JM, Morrow JB and Lin NJ. (2017) Challenging a bioinformatic tool's ability to detect microbial contaminants using *in silico* whole genome sequencing data. *PeerJ*. doi: 10.7717/peerj.3729

O'Neill AJ, Huovinen T, Fishwick CWG and Chopra I. (2006) Molecular genetic and structural modeling studies of *Staphylococcus aureus* RNA polymerase and the fitness of rifampin resistance genotypes in relation to clinical prevalence. *Antimicrobial Agents and Chemotherapy*. doi: 10.1128/AAC.50.1.298-309.2006

O'Neill J. (2016) Tackling Drug-Resistant Infections Globally: final report and recommendations. *Review on Antimicrobial Resistance*. Available at: https://amr-review.org/sites/default/files/160525_Final%20paper_with%20cover.pdf

Opatowski L, Guillemot D, Boëlle PY and Temime L. (2011) Contribution of mathematical modeling to the fight against bacterial antibiotic resistance. *Current Opinion in Infectious Diseases* 24(3). doi: 10.1097/QCO.0b013e3283462362

Orman MA and Brynildsen MP. (2016) Persister formation in *Escherichia coli* can be inhibited by treatment with nitric oxide. *Free Radical Biology and Medicine*. doi: 10.1016/j.freeradbiomed.2016.02.003

Ouderkirk JP, Nord JA, Turett GS and Kislak JW. (2003) Polymyxin B nephrotoxicity and efficacy against nosocomial infections caused by multiresistant gram-negative bacteria. *Antimicrobial Agents and Chemotherapy*. doi: 10.1128/AAC.47.8.2659-2662.2003

Oz T, Guvenek A, Yildiz S, Karaboga E, Tamer YT, Mumcuyan N, Ozan VB, Senturk GH, Cokol M, Yeh P and Toprak E. (2014) Strength of Selection Pressure us an Important Parameter Contributing to the Complexity of Antibiotic Resistance Evolution. *Molecular Biology and Evolution* 31(9): 2387-240.

Page RDM and Holmes EC (1998) Molecular Evolution: A Phylogenetic Approach (3rd edition). Wiley-Blackwell. Oxford, London.

Palmer AC and Kishony R. (2013) Understanding, predicting and manipulating the genotypic evolution of antibiotic resistance. *Nature Reviews Genetics*. doi: 10.1038/nrg3351

Palmer AC, Toprak E, Baym M, Kim S, Veres A, Bershtein S and Kishony R. (2015) Delayed commitment to evolutionary fate in antibiotic resistance fitness landscapes. *Nature Communications*. doi: 10.1038/ncomms8385

Pankey GA and Sabath LD. (2004) Clinical relevance of bacteriostatic versus bactericidal mechanisms of action in the treatment of Gram-positive bacterial infections. *Clinical Infectious Diseases*. doi: 10.1086/381972

Paradis S, Boissinot M, Paquette N, Bélanger SD, Martel EA, Boudreau DK, Picard FJ, Ouellette M, Roy PH and Bergeron MG. (2005) Phylogeny of the *Enterobacteriaceae* based on genes encoding elongation factor Tu and F-ATPase β-subunit. *International Journal of Systematic and Evolutionary Microbiology*. doi: 10.1099/ijs.0.63539-0

Pareek CS, Smoczynski R and Tretyn A. (2011) Sequencing technologies and genome sequencing. *Journal of Applied Genetics*. doi: 10.1007/s13353-011-0057-x

Pauly HM, Larson BE, Coatney GA, Button KD, DeCamp CE, Fajardo RS, Haut RC, Haut Donahue TL. (2015) Assessment of cortical and trabecular bone changes in two models of post-traumatic osteoarthritis. *Journal of Orthopaedic Research*. doi: 10.1016/S1473-3099(18)30475-4

Payne DJ, Gwynn MN, Holmes DJ and Pompliano DL. (2006) Drugs for bad bugs: confronting the challenges of antibacterial discovery. *Nature Reviews Drug Discovery* 6(1): 29-40. doi: 10.1038/nrd2201

Pearson T, Busch JD, Ravel J, Read TD, Rhoton SD, U'Ren JM, Simonson TS, Kachur SM, Leadem RR, Cardon ML, van ERT MN, Huynh LY, Fraser CM and Keim P. (2004) Phylogenetic discovery bias in *Bacillus anthracis* using single-nucleotide polymorphisms from whole-genome sequencing. *Proceedings of the National Academy of Sciences*. doi: 10.1073/pnas.0403844101

Pei AY, Oberdorf WE, Nossa CW, Agarwal A, Chokshi P, Gerz EA, Zhida J, Lee P, Yang L, Poles M, Brown SM, Sotero S, DeSantis T, Brodie E, Nelson K and Pei Z. (2010) Diversity of 16S rRNA genes within individual prokaryotic genomes. *Applied and Environmental Microbiology*. doi: 10.1128/AEM.02953-09

Pendleton JN, Gorman SP and Gilmore B. (2013) Clinical relevance of the ESKAPE pathogens. *Expert Review of Anti-Infective Therapy* 11(3): 297–308. doi: 10.1586/eri.13.12

Peske F, Matassova NB, Savelsbergh A, Rodnina MV and Wintermeyer W. (2000) Conformationally restricted elongation factor G retains GTPase activity but is inactive in translocation on the ribosome. *Molecular Cell*. doi: 10.1016/S1097-2765(00)00049-6

Petrov A, Chen J, O'Leary S, Tsai A and Puglisi JD. (2012) Single-molecule analysis of translational dynamics. *Cold Spring Harbor Perspectives in Biology*. doi: 10.1101/cshperspect.a011551

Pfister P, Risch M, Brodersen DE and Böttger EC. (2003) Role of 16S rRNA helix 44 in ribosomal resistance to hygromycin B. *Antimicrobial Agents and Chemotherapy*. doi: 10.1128/AAC.47.5.1496-1502.2003

Pitout JDD. (2012) Extraintestinal pathogenic *Escherichia coli*: a combination of virulence with antibiotic resistance. *Frontiers in Microbiology*. doi: 10.3389/fmicb.2012.00009

Pitout JDD and Laupland KB. (2008) Extended-spectrum beta-lactamase-producing *Enterobacteriaceae*: an emerging public-health concern. *The Lancet Infectious Diseases* 8(3): 159–166. doi: 10.1016/S1473-3099(08)70041-0

Pulk A and Cate JHD. (2013) Control of ribosomal subunit rotation by elongation factor G. Science 340(6140): 1235970. doi: 10.1126/science.1235970

Rajasekaran S and Rao GK. (2015) Synthesis, anti-microbial and molecular docking studies of some 2,3-disubsttuted quinazolinone analogs. *Journal of Computational Methods in Molecular Design* 5(4): 11–15.

Rajendhran J and Gunasekaran P. (2011) Microbial phylogeny and diversity: small subunit ribosomal RNA sequence analysis and beyond. *Microbiological Research*. doi: 10.1016/j.micres.2010.02.003

Ramadhan AA and Hegedus E. (2005) Survivability of vancomycin resistant enterococci and fitness cost of vancomycin resistance acquisition. *Journal of Clinical Pathology*. doi: 10.1136/jcp.2004.024091

Ramanathan B, Jindal HM, Le CF, Gudimella R, Anwar A, Razali R, Poole-Johnson J, Manikam R and Sekaran SD. (2017) Next generation sequencing reveals the antibiotic resistant variants in the genome of *Pseudomonas aeruginosa*. *PLOS ONE* 12(8). doi: 10.1371/journal.pone.0182524

Ramarao N, Nielsen-Leroux C and Lereclus D. (2012) The insect *Galleria mellonella* as a powerful infection model to investigate bacterial pathogenesis. *Journal of Visualized Experiments*. doi: 10.3791/4392

Ramirez MS and Tolmasky ME. (2010) Aminoglycoside modifying enzymes. *Drug Resistance Updates*. doi: 10.1016/j.drup.2010.08.003

Ramrath DJF, Lancaster L, Sprink T, Mielke T, Loerke J, Noller HF and Spahn CM. (2013) Visualization of two transfer RNAs trapped in transit during elongation factor G-mediated translocation. *Proceedings of the National Academy of Sciences*. doi: 10.1073/pnas.1320387110

Reuter S, Ellington MJ, Cartwright EJP, Köser CU, Török ME, Gouliouris T, Harris SR, Brown NM, Holden MT, Quail M, Parkhill J, Smith GP, Bentley SD and Peacock SJ. (2013) Rapid bacterial whole-genome sequencing to enhance diagnostic and public health microbiology. *JAMA Internal Medicine*. doi: 10.1001/jamainternmed.2013.7734

Reznick DN and Ghalambor CK (2005). Selection in nature: experimental manipulations of natural populations. *Integrative and Comparative Biology*. doi: 10.1093/icb/45.3.456

Riley M, Abe T, Arnaud MB, Berlyn MKB, Blattner FR, Chaudhuri RR, Glasner JD, Horiuchi T, Keseler IM, Kosuge T, Mori H, Perna NT, Plunkett G, and Rudd KE, Serres MH, Thomas GH, Thomson NR, Wishart D and Wanner BL. (2006) *Escherichia coli* K-12: a cooperatively developed annotation snapshot—2005. *Nucleic Acids Research*. doi: 10.1093/nar/gkj405

Rissman AI, Mau B, Biehl BS, Darling AE, Glasner JD and Perna NT. (2009) Reordering contigs of draft genomes using the Mauve Aligner. *Bioinformatics*. doi: 10.1093/bioinformatics/btp356

Robinson TP, Bu DP, Carrique-Mas J, Fèvre EM, Gilbert M, Grace D, Hay SI, Jiwakanon J, Kakkar M, Kariuki S, Laxminarayan R, Lubroth J, Magnusson U, Ngoc TP, Boeckel TP and Woolhouse MEJ. (2016). Antibiotic resistance is the quintessential one health issue. *Transactions of The Royal Society of Tropical Medicine and Hygiene*. doi: 10.1093/trstmh/trw048

Rodnina MV, Savelsbergh A, Katunin VI and Wintermeyer W. (1997) Hydrolysis of GTP by elongation factor G drives tRNA movement on the ribosome. *Nature* 385(6611): 37-41. doi: 10.1038/385037a0

Rosche WA and Foster PL. (2000) Determining mutation rates in bacterial populations. *Methods*. doi: 10.1006/meth.1999.0901

Rossolini GM, Arena F, Pecile P and Pollini S. (2014) Update on the antibiotic resistance crisis. *Current Opinion in Pharmacology*. doi: 10.1016/j.coph.2014.09.006

Russo TA and Johnson JR. (2003) Medical and economic impact of extraintestinal infections due to *Escherichia coli*: focus on an increasingly important endemic problem. *Microbes and Infection*. doi: 10.1016/S1286-4579(03)00049-2

Rybak MJ and McGrath BJ. (1996) Combination antimicrobial therapy for bacterial infections. *Drugs* 52(3): 390–405. doi: 10.2165/00003495-199652030-00005

Salsi E, Farah E, Dann J and Ermolenko DN. (2014) Following movement of domain IV of elongation factor G during ribosomal translocation. *Proceedings of the National Academy of Sciences*. doi: 10.1073/pnas.1410873111

Salverda ML, Dellus E, Gorter FA, Debets AJ, van der Oost J, Hoekstra RF, Tawfik DS and de Visser JA. (2011) Initial mutations direct alternative pathways of protein evolution. *PLoS Genetics* 7(3). doi: 10.1371/journal.pgen.1001321

Samuelsen Ø, Naseer U, Tofteland S, Skutlaberg DH, Onken A, Hjetland R, Sundsfjord A and Giske CG. (2009) Emergence of clonally related *Klebsiella pneumoniae* isolates of sequence type 258 producing plasmid-mediated KPC carbapenemase in Norway and Sweden. *Journal of Antimicrobial Chemotherapy*. doi: 10.1093/jac/dkp018

Sander P, Springer B, Prammananan T, Sturmfels A, Kappler M, Pletschette M and Böttger EC. (2002) Fitness cost of chromosomal drug resistance-conferring mutations. *Antimicrobial Agents and Chemotherapy*. doi: 10.1128/AAC.46.5.1204-1211.2002

Sanger F, Nicklen S and Coulson AR. (1977) DNA sequencing with chain-terminating inhibitors. *Proceedings of the National Academy of Sciences*. doi: 10.1073/pnas.74.12.5463

Sansom MSP and Weinstein H. (2000) Hinges, swivels and switches: the role of prolines in signalling via transmembrane alpha-helices. *Trends in Pharmacological Sciences* 21(11):445-51.

Santajit S and Indrawattana N. (2016) Mechanisms of antimicrobial resistance in ESKAPE Pathogens. *BioMed Research International*. doi: 10.1155/2016/2475067

Santer UV, Cekleniak J, Kansil S, Santer M, O'Connor M and Dahlberg A. (1995) A mutation at the universally conserved position 529 in *Escherichia coli* 16S rRNA creates a functional but highly error prone ribosome. *RNA* 1(1): 89–94.

Sato K, Hamada M, Asai K and Mituyama T. (2009) CentroidFold: A web server for RNA secondary structure prediction. *Nucleic Acids Research*. doi: 10.1093/nar/gkp367

Savelsbergh A, Katunin VI, Mohr D, Peske F, Rodnina MV and Wintermeyer W. (2003) An elongation factor G-induced ribosome rearrangement precedes tRNA-mRNA translocation. *Molecular Cell* 11(6): 1517-1523.

Schatz A, Bugle E and Waksman SA. (1944) Streptomycin, a substance exhibiting antibiotic activity against Gram-Positive and Gram-negative bacteria. *Experimental Biology and Medicine* 55(1): 66–69.

Schenk MF and de Visser JAGM. (2013) Predicting the evolution of antibiotic resistance. *BMC Biology*. doi: 10.1186/1741-7007-11-14

Schluenzen F, Tocilj A, Zarivach R, Harms J, Gluehmann M, Janell D, Bashan A, Bartels H, Agmon I, Franceschi F and Yonath A. (2000) Structure of functionally activated small ribosomal subunit at 3.3 Å resolution. *Cell*. doi: 10.1016/S0092-8674(00)00084-2

Schrag SJ, Perrot V and Levin BR. (1997) Adaptation to the fitness costs of antibiotic resistance in *Escherichia coli*. *Proceedings of the Royal Society B: Biological Sciences* 264(1386): 1287-1291.

Schroeder RE, Waldsich C and Wank H. (2000) Modulation of RNA function by aminoglycoside antibiotics. *The EMBO Journal* 19(1): 1-9.

Schwaber MJ, Navon-Venezia S, Kaye KS, Ben-Ami R, Schwartz D and Carmeli Y. (2006) Clinical and economic impact of bacteremia with extended-spectrum-β-lactamase-producing *Enterobacteriaceae*. *Antimicrobial Agents and Chemotherapy*. doi: 10.1128/AAC.50.4.1257-1262.2006

Shcherbakov D, Akbergenov R, Matt T, Sander P, Andersson DI and Böttger EC. (2010) Directed mutagenesis of *Mycobacterium smegmatis* 16S rRNA to reconstruct the *in vivo* evolution of aminoglycoside resistance in *Mycobacterium tuberculosis*. *Molecular Microbiology*. doi: 10.1111/j.1365-2958.2010.07218.x

Shebl B, Menke DE, Pennella M, Poudyal RR, Burke DH and Cornish PV. (2016) Preparation of ribosomes for smFRET studies: a simplified approach. *Archives of Biochemistry and Biophysics* 603: 118–130. doi: doi: 10.1016/j.abb.2016.05.010

Sheik CS, Reese BK, Twing KI, Sylvan JB, Grim SL, Schrenk MO, Sogin ML and Colwell FS. (2018) Identification and removal of contaminant sequences from ribosomal gene databases: lessons from the census of deep life. *Frontiers in Microbiology* 9: 840. doi: 10.3389/fmicb.2018.00840

Shimkets LJ. (1999) Intercellular Signaling During Fruiting-Body Development of *Myxococcus xanthus*. *Annual Review of Microbiology*. doi: 10.1146/annurev.micro.53.1.525

Sigmund CD and Morgan EA. (1982) Erythromycin resistance due to a mutation in a ribosomal RNA operon of *Escherichia coli*. *Proceedings of the National Academy of Sciences* 79(18): 5602-6.

Silander OK and Ackermann M. (2009) The constancy of gene conservation across divergent bacterial orders. *BMC Research Notes*. doi: 10.1186/1756-0500-2-2

Silhavy TJ, Kahne D and Walker S. (2010) The bacterial cell envelope. *Cold Spring Harbor Perspectives in Biology*. doi: 10.1101/cshperspect.a000414

Silver LL. (2011) Challenges of antimicrobial discovery revisited. *Clinical Microbiology Reviews*. doi: 10.1111/j.1749-6632.2010.05828.x

Singh PK. (2017) One health approach to tackle antimicrobial resistance in South East Asia. *British Medical Journal* 358. doi: 10.1136/bmj.j3625

Sjölund M, Tano E, Blaser MJ, Andersson DI and Engstrand L. (2005) Persistence of resistant *Staphylococcus epidermidis* after single course of clarithromycin. *Emerging Infectious Diseases*. 11(9): 1389-1393.

Smith HW. (1975) Persistence of tetracycline resistance in pig *E. coli*. *Nature* 258: 628. doi: 10.1038/258628a0

Soltis DE. (2003) The role of phylogenetics in comparative genetics. *Plant Physiology*. doi: 10.1104/pp.103.022509

Southworth DR, Brunelle JL and Green R. (2002) EFG-independent translocation of the mRNA:tRNA complex is promoted by modification of the ribosome with thiol specific reagents. *Journal of Molecular Biology*. doi: 10.1016/S0022-2836(02)01196-8

Spiegel PC, Ermolenko DN and Noller HF. (2007) Elongation factor G stabilizes the hybrid-state conformation of the 70S ribosome. *RNA*. doi: 10.1261/rna.601507

Spiro S and Guest JR. (1991) Adaptive responses to oxygen limitation in *Escherichia coli*. *Trends in Biochemical Sciences* 16(8): 310–4.

Spratt BG. (1994) Resistance to antibiotics mediated by target alterations. *Science* 264(5157): 388. doi: 10.1126/science.8153626

Spratt BG. (1996) Antibiotic resistance: counting the cost. *Current Biology* 6: 1219–1221.

Springer B, Kidan YG, Prammananan T, Ellrott K, Böttger EC and Sander P. (2001) Mechanisms of streptomycin resistance: selection of mutations in the 16S rRNA gene conferring resistance. *Antimicrobial Agents and Chemotherapy*. doi: 10.1128/AAC.45.10.2877-2884.2001

Sprouffske K and Wagner A. (2016) Growthcurver: An R package for obtaining interpretable metrics from microbial growth curves. *BMC Bioinformatics*. doi: 10.1186/s12859-016-1016-7

Srinivasan R, Karaoz U, Volegova M, MacKichan J, Kato-Maeda M, Miller S, Nadarajan R, Brodie EL and Lynch SV. (2015) Use of 16S rRNA gene for identification of a broad range of clinically relevant bacterial pathogens. *PLOS ONE*. doi: 10.1371/journal.pone.0117617

Stein A and Raoult D. (2002) Colistin: An Antimicrobial for the 21st Century? *Clinical Infectious Diseases* 35(7): 901-2.

Stucki D and Gagneux S. (2013) Single nucleotide polymorphisms in *Mycobacterium tuberculosis* and the need for a curated database. *Tuberculosis*. doi: 10.1016/j.tube.2012.11.002

Sun D. (2018) Pull in and push out: mechanisms of horizontal gene transfer in bacteria. *Frontiers in Microbiology*. doi: 10.3389/fmicb.2018.02154

Suzuki S, Horinouchi T and Furusawa C. (2014). Prediction of antibiotic resistance by gene expression profiles. *Nature Communications*. doi: 10.1038/ncomms6792

Swenson JM, Wallace RJ, Silcox VA and Thornsberry C. (1985) Antimicrobial susceptibility of five subgroups of *Mycobacterium fortuitum* and *Mycobacterium chelonae*. *Antimicrobial Agents and Chemotherapy*. doi: 10.1128/AAC.28.6.807

Temime L, Hejblum G, Setbon M and Valleron AJ. (2008) The rising impact of mathematical modelling in epidemiology: antibiotic resistance research as a case study. *Epidemiology and Infection*. doi: 10.1017/S0950268807009442

Tenaillon O, Skurnik D, Picard B and Denamur E. (2010) The population genetics of commensal *Escherichia coli*. *Nature Reviews Microbiology* 8: 207. doi: 10.1038/nrmicro2298

Tenover FC. (2006) Mechanisms of antimicrobial resistance in bacteria. *American Journal of Infection Control* 34(5). doi: 10.1016/j.ajic.2006.05.219

Thermes C. (2014) Ten years of next-generation sequencing technology. *Trends in Genetics: TIG*. doi: 10.1016/j.tig.2014.07.001

Toguchi A, Siano M, Burkart M and Harshey RM. (2000) Genetics of swarming motility in *Salmonella enterica* Serovar *Typhimurium*: critical role for lipopolysaccharide. *Journal of Bacteriology* 182(22): 6308-21.

Toivonen JM, Boocock MR and Jacobs HT. (1999) Modelling in *Escherichia coli* of mutations in mitoribosomal protein S12: novel mutant phenotypes of rpsL. *Molecular Microbiology* 31(6): 1735–1746.

Tomasz A. (1979) The mechanism of the irreversible antimicrobial effects of penicillins: how the beta-lactam antibiotics kill and lyse bacteria. *Annual Reviews Microbiology*. 33: 133-37.

Toprak E, Veres A, Michel JB, Chait R, Hartl DL and Kishony R. (2012) Evolutionary paths to antibiotic resistance under dynamically sustained drug selection. *Nature Genetics*. doi: 10.1038/ng.1034

Toprak E, Veres A, Yildiz S, Pedraza JM, Chait R, Paulsson J and Kishony R. (2013) Building a morbidostat: an automated continuous-culture device for studying bacterial drug resistance under dynamically sustained drug inhibition. *Nature Protocols*. doi: 10.1038/nprot.2013.021

Tor Y, Hermann T and Westhof E. (1998) Deciphering RNA recognition: aminoglycoside binding to the hammerhead ribozyme. *Chemistry and Biology*. doi: 10.1016/S1074-5521(98)90286-1

Torella JP, Chait R and Kishony R. (2010) Optimal drug synergy in antimicrobial treatments. *PLoS Computational Biology*. doi: 10.1371/journal.pcbi.1000796

Tourigny DS, Fernández IS, Kelley AC and Ramakrishnan V. (2013) Elongation factor G bound to the ribosome in an intermediate state of translocation. *Science*. doi: 10.1126/science.1235490

Treangen TJ and Salzberg SL. (2013) Repetitive DNA and next-generation sequencing: computational challenges and solutions. *Nature Reviews Genetics* 13(1): 36–46.

Triman KL. (1996) *The 16S ribosomal RNA mutation database (16SMDB)*. *Nucleic Acids Research* 24(1): 166-8.

Tsai A, Uemura S, Johansson M, Puglisi EV, Marshall RA, Aitken CE, Korlach J, Ehrenberg M and Puglisi JD. (2013) The impact of aminoglycosides on the dynamics of translation elongation. *Cell Reports*. doi: 10.1016/j.celrep.2013.01.027

Tzouvelekis LS, Markogiannakis A, Psichogiou M, Tassios PT and Daikos GL. (2012) Carbapenemases in *Klebsiella pneumoniae* and other *Enterobacteriaceae*: an evolving crisis of global dimensions. *Clinical Microbiology Reviews*. doi: 10.1128/CMR.05035-11

Utturkar SM, Klingeman DM, Land ML, Schadt CW, Doktycz MJ, Pelletier DA and Brown SD. (2014) Evaluation and validation of *de novo* and hybrid assembly techniques to derive high-quality genome sequences. *Bioinformatics* 30(19): 2709–2716.

Vaara M (1992). Agents that increase the permeability of the outer membrane. *Microbiological Reviews* 56(3): 395-411.

Vaiana AC and Sanbonmatsu KY. (2009) Stochastic gating and drug-ribosome interactions. *Journal of Molecular Biology*. doi: 10.1016/j.jmb.2008.12.035

Vakulenko SB and Mobashery S. (2003) Versatility of aminoglycosides and prospects for their future. *Clinical Microbiology Reviews*. doi: 10.1128/CMR.16.3.430-450.2003

van Dijk T, Hwang S, Krug J, de Visser JAGM and Zwart MP. (2017) Mutation supply and the repeatability of selection for antibiotic resistance. *Physical Biology 14*(5). doi: 10. 1088/1478-3975/aa7f36

van Niel CB. (1944) The culture, general physiology, morphology and classification of the non-sulfur purple and brown bacteria. *Bacteriological Reviews* 8(1): 1-118.

van Ryk DI and Dahlberg AE. (1995) Structural changes in the 530 loop of *Escherichia coli* 16S rRNA in mutants with impaired translational fidelity. *Nucleic Acids Research*. doi: 10.1093/nar/23.17.3563

Venter JC *et al*. (2001) The sequence of the human genome. *Sciences* 291(5507): 1304-51.

Ventola CL. (2015) The antibiotic resistance crisis. *Pharmacy and Therapeutics* 40(4): 277-283.

Vestergaard M, Paulander W, Leng B, Nielsen JB, Westh HT and Ingmer H. (2016) Novel pathways for ameliorating the fitness cost of gentamicin resistant small colony variants. *Frontiers in Microbiology*. doi: 10.3389/fmicb.2016.01866

Větrovský T and Baldrian P. (2013) The variability of the 16S rRNA gene in bacterial genomes and its consequences for bacterial community analyses. *PLOS ONE*. doi: 10.1371/journal.pone.0057923

Vijesh AM, Isloor AM, Telkar S, Arulmoli T and Fun HK. (2013) Molecular docking studies of some new imidazole derivatives for antimicrobial properties. *Arabian Journal of Chemistry*. doi: 10.1016/j.arabjc.2011.10.007

Vila A, Viril-Farley J and Tappricht WE. (1994) Pseudoknot in the central domain of small subunit ribosomal RNA is essential for translation. *Proceeding of the National Academy of Sciences USA* 91(23): 11148-11152.

Vinson JP, Jaffe DB, O'Neill K, Karlsson EK, Stange-Thomann N, Anderson S, Mesirov JP, Satoh N, Satou Y, Nusbaum C, Birren B, Galagan JE, Lander ES. (2005) Assembly of polymorphic genomes: algorithms and application to Ciona savignyi. *Genome Research* 15(8):1127–35.

Vogwill T and Maclean RC. (2015) The genetic basis of the fitness costs of antimicrobial resistance: a meta-analysis approach. *Evolutionary Applications*. doi: 10.1111/eva.12202

von Wintersdorff CJH, Penders J, van Niekerk JM, Mills ND, Majumder S, van Alphen LB, Savelkoul PHM and Savelkoul PHM. (2016) Dissemination of antimicrobial resistance in microbial ecosystems through horizontal gene transfer. *Frontiers in Microbiology*. doi: 10.3389/fmicb.2016.00173

Voorhees RM and Ramakrishnan V. (2013) Structural basis of the translational elongation cycle. *Annual Review of Biochemistry* 82: 203–236.

Wahl LM, Gerrish PJ and Saika-Voivod I. (2002) Evaluating the impact of population bottlenecks in experimental evolution. *Genetics*. doi: 10.1002/ijc.23980

Walsh C. (2003) Antibiotics: actions, origins, resistance. *American Society for Microbiology*, Washington.

Walter JD, Hunter M, Cobb M, Traeger G and Spiegel PC. (2012) Thiostrepton inhibits stable 70S ribosome binding and ribosome-dependent GTPase activation of elongation factor G and elongation factor 4. *Nucleic Acids Research*. doi: 10.1093/nar/gkr623

Wang G, Wilson TJM, Jiang Q and Taylor DE. (2001) Spontaneous mutations that confer antibiotic resistance in *Helicobacter pylori*. *Antimicrobial Agents and Chemotherapy*. doi: 10.1128/AAC.45.3.727-733.2001

Wang H, Wang J, Yu P, Ge P, Jiang Y, Xu R, Chen R and Liu X. (2017) Identification of antibiotic resistance genes in the multidrug-resistant *Acinetobacter baumannii* strain, MDR-SHH02, using whole-genome sequencing. *International Journal of Molecular Medicine* 39(2): 364-372. doi: 10.3892/ijmm.2016.2844

Wang L, Pulk A, Wasserman MR, Feldman MB, Altman RB, Doudna Cate JH and Blanchard SC. (2012) Allosteric control of the ribosome by small-molecule antibiotics. *Nature Structural and Molecular Biology*. doi: 10.1038/nsmb.2360

Wang M and Ford RM. (2009) Transverse bacterial migration induced by chemotaxis in a packed column with structured physical heterogeneity. *Environmental Science and Technology* 43(15): 5921-7.

Wang Y and Huang JM. (2018) Assessment of molecular markers for classification of bacterial phyla using topological dissimilarity of phylogenetic trees. *Journal of Phylogenetics and Evolutionary Biology*. doi: 10.4172/2329-9002.1000204

Wang Y and Rando RR. (1995) Specific binding of aminoglycoside antibiotics to RNA. *Chemistry and Biology*. doi: 10.1016/1074-5521(95)90047-0

Webber MA, Whitehead RN, Mount M, Loman NJ, Pallen MJ and Piddock LJV. (2015) Parallel evolutionary pathways to antibiotic resistance selected by biocide exposure. *Journal of Antimicrobial Chemotherapy*. doi: 10.1093/jac/dkv109

Weinreich DM, Delaney NF, DePristo MA and Hartl DL. (2006) Darwinian evolution can follow only very few mutational paths to fitter proteins. *Science* 312(5770): 111-4. doi: 10.1126/science.1123539

Weinstein MJ, Luedemann GM, Oden EM, Wagman GH, Rosselet JP, Marquez JA, Coniglio CT, Charney W, Herzog HL and Black J. (1963) Gentamicin, a new antibiotic complex from *Micromonospora*. *Journal of Medicinal Chemistry* 6(4): 463–464. doi: 10.1021/jm00340a034

Weizhong Li and Adam Godzik. (2006) Cd-hit: a fast program for clustering and comparing large sets of protein or nucleotide sequences. *Bioinformatics* 22(13): 1658–1659.

WHO. (2014) Antimicrobial Resistance: Global Report on Surveillance. Available at: https://apps.who.int/iris/bitstream/handle/10665/112642/9789241564748_eng.pdf

WHO. (2018) Tackling Antimicrobial Resistance (AMR) Together. Working Paper 1.0: Multisectoral coordination. Available at: https://www.who.int/antimicrobial-resistance/publications/Tackling-AMR-multisectoral-coordination-june2018.pdf

Wilcox SK, Cavey GS and Pearson JD. (2001) Single ribosomal protein mutations in antibiotic-resistant bacteria analysed by mass spectrometry. *Antimicrobial Agents and Chemotherapy* 45(11): 3046–3055. doi: 10.1128/AAC.45.11.3046–3055.2001

Wiser MJ and Lenski RE. (2015) A comparison of methods to measure fitness in *Escherichia coli*. *PLOS ONE*. doi: 10.1371/journal.pone.0126210

Woerther PL, Burdet C, Chachaty E and Andremont A. (2013) Trends in human faecal carriage of extended-spectrum β-lactamases in the community: toward the globalisation of CTX-M. *Clinical Microbiology Reviews*. doi: 10.1128/CMR.00023-13

Wolfe AJ and Berg HC. (1989) Migration of bacteria in semisolid agar. *Proceedings of the National Academy of Sciences*. doi: 10.1073/pnas.86.18.6973

Wong C and Currie J. (2001) Teaching with SCIGRESS exercises on molecular modeling in chemistry. Pacific University, Forest Grove, Oregon.

Wood TK, González Barrios AF, Herzberg M and Lee J. (2006) Motility influences biofilm architecture in *Escherichia coli*. *Applied Microbiology and Biotechnology*. doi: 10.1007/s00253-005-0263-8

Woodcock DJ, Krusche P, Strachan NJC, Forbes KJ, Cohan FM, Méric G and Sheppard SK. (2017) Genomic plasticity and rapid host switching can promote the evolution of generalism: A case study in the zoonotic pathogen *Campylobacter*. *Scientific Reports*. doi: 10.1038/s41598-017-09483-9

Woodford N and Ellington MJ. (2007) The emergence of antibiotic resistance by mutation. *Clinical Microbiology and Infection*. doi: 10.1111/j.1469-0691.2006.01492.x

Woods RJ and Read AF. (2015) Clinical management of resistance evolution in a bacterial infection: a case study. *Evolution, Medicine and Public Health*. doi: 10.1093/emph/eov025

Worthington RJ and Melander C. (2013) Combination approaches to combat multidrug-resistant bacteria. *Trends in Biotechnology*. doi: 10.1016/j.tibtech.2012.12.006

Wright GD. (2011) Molecular mechanisms of antibiotic resistance. *Chemical Communications* 47(14): 4055–4061. doi: 10.1039/C0CC05111J

Wu LJ and Errington J. (1994) *Bacillus subtilis* SpoIIIE protein required for DNA segregation during asymmetric cell division. *Science* 264(5158): 572–575.

Yang B, Wang Y and Qian PY. (2016) Sensitivity and correlation of hypervariable regions in 16S rRNA genes in phylogenetic analysis. *BMC Bioinformatics*. doi: 10.1186/s12859-016-0992-y

Yang L, Jelsbak L, Marvig RL, Damkiaer S, Workman CT, Rau MH, Hansen SK, Folkesson A, Johansen HK, Ciofu O, Høiby N, Sommer MO and Molin S. (2011) Evolutionary dynamics of bacteria in a human host environment. *Proceedings of the National Academy of Sciences*. doi: 10.1073/pnas.1018249108

Yang Z. (2005) The power of phylogenetic comparison in revealing protein function. *Proceedings of the National Academy of Sciences*. doi: 10.1073/pnas.0500371102

Yang Z and Nielsen R. (1998) Synonymous and nonsynonymous rate variation in nuclear genes of mammals. *Journal of Molecular Evolution* 46(4): 409–418. doi: 10.1007/PL00006320

Yates JL. (1979) Role of ribosomal protein S12 in discrimination of aminoacyl-tRNA. *Journal of Biological Chemistry*. doi: 10.1016/j.neucom.2011.09.039

Ye J, Coulouris G, Zaretskaya I, Cutcutache I, Rozen S and Madden T. (2012) Primer-BLAST: A tool to design target-specific primers for polymerase chain reaction. *BMC Bioinformatics*. 13:134.

Yeh PJ, Hegreness MJ, Aiden AP and Kishony R. (2009) Drug interactions and the evolution of antibiotic resistance. *Nature Reviews Microbiology*. doi: 10.1038/nrmicro2133

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

Yoshizawa S, Fourmy D and Puglisi JD. (1998) Structural origins of gentamicin antibiotic action. *EMBO Journal*. doi: 10.1093/emboj/17.22.6437

Zhang Q, Lambert G, Liao D, Kim H, Robin K, Tung CK, Pourmand N and Austin RH. (2011) Acceleration of emergence of bacterial antibiotic resistance in connected microenvironments. *Science* 333(6050): 1764-1767.

Zhou C, Niu H, Yu H, Zhou L and Wang Z. (2015) Effects of two novel amino acid substitutions on the penicillin binding properties of the PBP5 C-terminal from *Enterococcus faecium*. *Molecular Medicine Reports*. doi: 10.3892/mmr.2015.4057

Zhou J, Lancaster L, Trakhanov S and Noller HF. (2012) Crystal structure of release factor RF3 trapped in the GTP state on a rotated conformation of the ribosome. *RNA*. doi: 10.1261/rna.031187.111

Zhou J and Rudd KE. (2013) EcoGene 3.0. *Nucleic Acids Research*. doi: 10.1093/nar/gks1235

Zuker M and Stiegler P. (1981) Optimal computer folding of large RNA sequences using thermodynamics and auxiliary information. *Nucleic Acids Research*. doi: 10.1093/nar/9.1.133

Zwart MP, Schenk MF, Hwang S, Koopmanschap B, de Lange N, van de Pol L, Nga TT, Szendro IG, Krug J and de Visser JAGM. (2018) Unravelling the causes of adaptive benefits of synonymous mutations in TEM-1 β-lactamase. *Heredity* 121(5): 406–421. doi: 10.1038/s41437-018-0104-z