

Developing a new efficient CRISPR-Cas9 base Homology Directed Repair oriented gene editing plasmid for Streptomyces

Kate Murray

Oxford Brookes University

Health and Life Sciences Department

September 2024

Supervisors: Dr Hee-Jeon Hong, Dr Barbara Jennings and Professor Susan Brookes

This thesis is submitted in partial fulfilment of the requirements of the award of Master.

Contents

Acknowledgements.....	4
Abbreviations.....	5
Abstract	7
1. Introduction.....	9
1.1. Streptomyces.....	9
1.2. Antibiotics and the resistance crisis.....	11
1.3. Glycopeptide antibiotics	14
1.4. Glycopeptide resistance studies in <i>S. coelicolor</i>	18
1.5. The CRISPR-Cas gene editing system	22
1.5.1. Discovery of CRISPR-Cas	22
1.5.2. CRISPR-Cas9.....	24
1.5.3. CRISPR-Cas9 based gene editing systems developed in Streptomyces.....	28
1.6. Aims and objectives	35
2. Materials and Methods.....	37
2.1. Bacterial strains, plasmids and oligos for PCR.....	37
2.2. Growth media	39
2.2.1. Soya flour mannitol agar (SFM)	39
2.2.2. Luria-Bertani broth (LB).....	39
2.2.3. Luria-Bertani agar (LA).....	39
2.2.4. 2x Yeast extract tryptone medium (2x YT).....	39
2.3. Buffers and solutions.....	40
2.3.1. Transformation and storage solution (TSS).....	40
2.3.2. Super optimal media with catabolic repressor for transformation (SOC).....	40
2.3.3. Solutions for small-scale plasmid extraction from <i>E. coli</i>	41
2.3.3.1. Sodium chloride Tris-EDTA (STE)	41
2.3.3.2. Solution 1	41
2.3.3.3. Solution 2	41
2.3.3.4. Solution 3	41

2.3.3.5. TE buffer	42
2.3.3.6. Others	42
2.4. Bacterial culture and storages	42
2.4.1. <i>E. coli</i> strains	42
2.4.2. Preparation of chemically competent <i>E. coli</i> cells	42
2.4.3. <i>S. coelicolor</i> cells	43
2.5. Microbial DNA cloning	43
2.5.1. Small-scale plasmid DNA extraction from <i>E. coli</i>	43
2.5.2. Restriction endonuclease digestion	44
2.5.3. Gel electrophoresis	44
2.5.4. DNA extraction and purification from Agarose gel	45
2.5.5. Ligation	45
2.5.6. Transformation of competent <i>E. coli</i> cells	45
2.5.7. PCR	45
2.6. Other microbiological methods and materials	46
2.6.1. Conjugation between <i>E. coli</i> and <i>S. coelicolor</i>	46
2.6.2. Miscellaneous reagents, enzymes and kits used for this study	47
3. Results	48
3.1. Configuration of pCRISPR-Cas9-KM	48
3.1.1. Sequence analysis and confirmation of pCRISPR-Cas9 and pCRISPR-cBEST plasmids by restrictions	50
3.1.2. Construction of pCRISPR-Cas9-KM	50
3.1.3. Confirmation of pCRISPR-Cas9-KM plasmid	53
3.1.4. Re-configuration of pCRISPR-Cas9-KM	57
3.2. Analysis of the inability to create pCRISPR-Cas9-KM	60
3.2.1. Colony PCR of pCRISPR-Cas9-KM following plasmid re-configuration	60
3.2.2. Restriction enzyme analysis of subsequent pCRISPR-Cas9-KM*	60
3.2.3. Sequencing analysis of reconfigured pCRISPR-Cas9-KM* plasmid indicates unexpected rearrangements	60

4. Discussion	66
4.1. Inability to produce desired pCRISPR-Cas9-KM plasmid due to unexpected rearrangement of plasmid DNA.....	66
4.2. Critical review of methods	68
4.3. Trial of Gibson Assembly to ligate 1 Kb up- and down-stream region of <i>vanK</i> into <i>Stu</i> I site of pCRISPR-Cas9-KM* using NEBuilder HiFi DNA Assembly	70
4.4. Conclusions and future perspectives.....	73
4.4.1. Functional genomic data concerning teicoplanin resistance clusters.....	73
4.4.2. Implications and utilisations of an optimised CRISPR-Cas9 based plasmid	74
References	76

Acknowledgements

I would like to thank the many people who have made the completion of this project and thesis possible through their continued support and aid. Firstly, I want to thank Dr Hee-Jeon Hong who has been an excellent teacher and provided immense support throughout this project. She has provided the technical expertise and knowledge that I required, and nurtured my desire to pursue a research career. Secondly, I would like to thank all the technical support staff in the Health and Life Sciences department within the Sinclair building who always provided support and assistance when needed. I would also like to thank Prof. Colin Smith and Dr. Giselda Bucca from the University of Brighton who kindly gifted the pCRISPR-Cas9 plasmid utilised in the present study. Again, the completion of this thesis could not have been completed without the endless research carried out developing the CRISPR-Cas editing plasmids and understanding glycopeptide resistance mechanisms, which were used as reference material in the writing of this thesis. Lastly I would like to thank my family and my wonderful partner for their continued support, without which, the completion of this project and thesis would not have been completed.

Abbreviations

AMR: Antimicrobial Resistance

cAMP: Cyclic Adenosine Monophosphate

CAP: Catabolic activator protein

Cas: CRISPR-associated system

cBEST: pCRISPR-cBEST

CRISPR: Clustered regularly interspaced short palindromic repeats

(p)CRISPR-BEST: CRISPR Base Editing System

crRNA: CRISPR RNA

DI: De-ionised

DMSO: Dimethyl Sulfoxide

DNA: Deoxyribonucleic acid

DSB: Double-stranded break

EDTA: Ethylenediaminetetraacetic acid

HDR: Homology-directed repair

IPTG: Isopropyl-beta-D-thiogalactopyranoside

Kb: Kilo-base

KM: pCRISPR-Cas9-KM

LA: Luria-Bertani Agar

LB: Luria-Bertani Broth

MIC: Minimum Inhibitory Concentration

MMR: Mismatch repair

MRSA: Methicillin-resistant *Staphylococcus aureus*

NAG: N-acetyl glucosamine

NAM: N-acetyl muramic acid

NER: Nucleotide excision repair

NHEJ: Non-Homologous End Joining

NTP: Nucleoside triphosphates

PAM: Protospacer Adjacent Motif

PAP: Phosphatidic acid phosphatase

PCR: Polymerase chain reaction

PEG: Polyethylene glycol

PG: Peptidoglycan

RE: Restriction Enzyme

RNA: Ribonucleic acid

RPM: Rotations per Minute

SDS: Sodium Dodecyl-Sulphate

SFM: Soya Flour Mannitol

sgRNA: Single-guide RNA

SOC: Super Optimal media with Catabolic repressor

ssDNA: Single-stranded DNA

STE: Sodium chloride Tris-EDTA

TE: Tris-EDTA

tracrRNA: Trans-Activating CRISPR RNA

TCS: Two-component system

TSS: Transformation and Storage solution

UDP: Uridine diphosphate

UV: Ultraviolet

VRE: Vancomycin-resistant *Enterococci*

VRSA: Vancomycin-resistant *Staphylococcus aureus*

Abstract

The CRISPR-Cas9 editing mechanism revolutionized genetic engineering, providing a precise and versatile tool for editing the genomes of virtually any organism. *Streptomyces* are soil-dwelling bacteria with a complex life cycle and a reputation for producing a wide array of bioactive compounds, including many clinically important antibiotics. Despite their significance, genetic manipulation in *Streptomyces* has traditionally been challenging. Their complex and large genomes, slow growth, and the presence of multiple, sometimes redundant, biosynthetic gene clusters made traditional methods of mutagenesis laborious and time-consuming. The implementation of CRISPR-Cas9 in *Streptomyces* therefore has a profound impact on both basic research and industrial applications. In particular, it is important for the development of new antibiotics at a time when rising antibiotic resistance is growing into a global health crisis.

This study sought to develop a new efficient CRISPR-Cas9 gene editing plasmid that works based on the homology directed repair system in *Streptomyces*. The fundamental goal of the study is to characterise the phenotype of a deletion mutant in the gene cluster, SCO6357-53, by using the newly developed CRISPR-Cas9 plasmid. Previous study has shown that SCO6357-53 is a suspected teicoplanin tolerance/resistance cluster in *Streptomyces coelicolor*. A natural product glycopeptide, teicoplanin, is an important alternative front-line therapy to vancomycin for the treatment of Gram-positive nosocomial infections but, unlike vancomycin, its mechanism of resistance is not clearly understood. While this study sought to characterise this cluster, an inability to create the new CRISPR-Cas9 based editing plasmid prevented any genetic manipulations to be carried out. Extensive plasmid DNA analysis using restriction enzyme digestion, PCR and sequencing was conducted. An initial inability to produce the desired plasmid led to a reconfiguring of the DNA assembly, however this second method also failed to produce the desired resulting plasmid. Whole plasmid DNA sequencing analysis revealed that a range of different unexpected rearrangements of DNA occurred, with either the loss of DNA segments which are functionally vital or a gain of DNA due to an unexpected insertion of external DNA fragments in various loci in the reconfigured plasmid. Although additional unwanted DNA fragments were present, one plasmid sample seemed to preserve all functionally vital regions for the HDR oriented targeted gene editing, so further cloning based on Gibson assembly was undertaken. However, unfortunately this also failed to obtain a desired product. Thus, an inability to construct the required HDR oriented gene editing plasmid, pCRISPR-Cas9KM, meant that further investigation into the teicoplanin resistance cluster could not be carried out.

In light of this, further advancements into generating this actinomycete-codon optimised genome editing tool needs to be carried out. By investigating the knockout of this teicoplanin resistance-conferring cluster, further insights into the associated resistance mechanisms can be determined. Gaining insights into these resistance mechanisms will provide a fundamental keystone in tackling the rising Anti-Microbial Resistance crisis. Through understanding these mechanisms, current antibiotics can be developed in a way that circumvent the arising bacterial resistance mechanisms, thus helping to alleviate the increasingly prevalent AMR crisis.

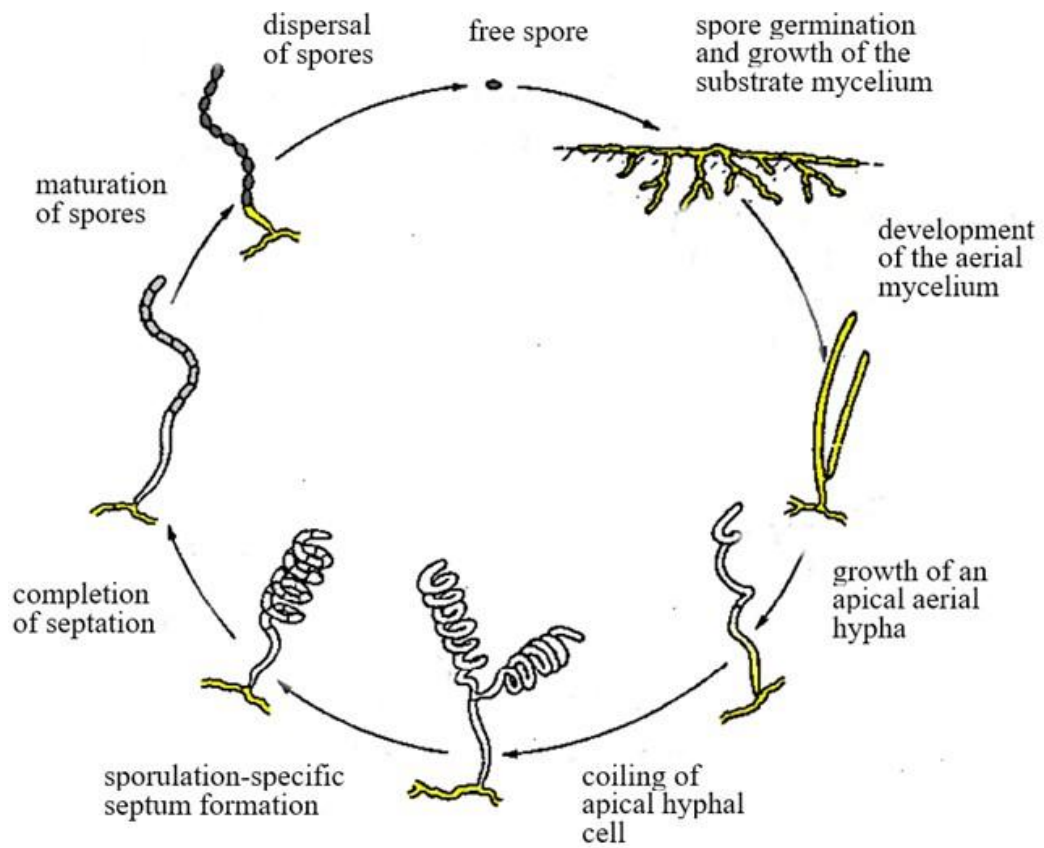
1. Introduction

1.1. Streptomyces

Actinomycetes are heterogeneous bacteria found in the Actinomycetales order and are Gram-positive filamentous bacteria. They are characterised by their high G+C content (over 70%) and their morphological complexity (Hu *et al.*, 2018). Actinomycetes form a distinct group of bacteria due to their unique life cycle: actinomycetes form mycelium and hyphae during development, which initially caused them to be mistaken for fungi, but have since been found to be distinct (they are smaller than fungi) (Pepper *et al.*, 2015). These bacteria predominantly found in soil and decaying organic matter, play a crucial role in decomposition and nutrient cycling. Streptomyces is the most dominant group in the order Actinomycetales and is well-known for its prolific production of secondary metabolites. In particular, many Streptomyces species are antibiotic producers (responsible for over 75% of current antimicrobial agents) that contain innate antibiotic resistance, which make them a genus of significant interest in the molecular microbiology of antibiotic resistance mechanisms.

The complex development of Streptomyces species results in the production of a wide range of bioactive compounds and secondary metabolites, including antibiotics (Fig. 1A). During development, Streptomyces form complex multicellular sporulating structures which differentiate into distinct “tissues”. Branching filamentous vegetative growth produces aerial hyphae that bear long chains of reproductive spores. During maturation, the aerial hyphae undergo a series of complex changes and coiling to form the sporulation-specific septa. Spore maturation is then completed by the thickening of the spore walls and deposition of spore pigment. For example, the model actinomycete, *Streptomyces coelicolor* colonies will change colour from white to grey as spores mature and this stage of maturation is what triggers the antibiotic biosynthesis gene expression (Fig. 1B) (Kieser *et al.*, 2000). When mature, the branching mycelium are found in the lower parts of bacterial growth, and the upper surface is covered in aerial hyphae, bearing chains of spores. These morphologically distinct hyphae are characterised by their dense, white, fuzzy appearance (Wildermuth, 1970). Morphological differentiation requires the presence of developmental genes *bld* and *whi*. When *bld* genes are blocked in early development and grow as substrate mycelium, therefore do not produce aerial mycelium, they have a shiny “bald” appearance. When *whi* genes are blocked in the later stages of development, aerial hyphae are produced that cannot complete the developmental process. These hyphae cannot form mature spore chains, thus do not exhibit the grey colour, and as such, have a white appearance (Chater, 1972).

A



B



Figure 1. (A) A diagram illustrates a complex life cycle of *Streptomyces* species (Kieser *et al.*, (2000). (B) Images of *S. coelicolor* model actinomycete grown on SFM agar medium.

S. coelicolor is the model organism of this study. Its full genome sequences were annotated and published for the first time in 2002 and was the first throughout the Actinomycete phylum (Bentley *et al.*, 2002). The genome consists of an 8.67 Mb linear chromosome, within which there are 7,825 predicted genes. The genome encodes ~20 clusters that are known or predicted to confer secondary metabolites. During normal *S. coelicolor* growth, four antibiotic compounds are expressed. Two of these compounds are pigmented: actinorhodin (an aromatic polyketide) presents with a blue pigment and undecylprodigiosin (a tripyrrole) with a red pigment. Methylenomycin, a cyclopentanone derivative and calcium-dependent antibiotic (CDA), a cyclic lipopeptide are also produced (Chater and Bibb, 1997). While many *Streptomyces* strains including *S. coelicolor* do not affect humans, some bacteria within the same Actinomycetales order do, e.g. pathogenic *Mycobacterium tuberculosis* and *Mycobacterium leprae*. Genomic-level comparison between the harmless model actinomycete, *S. coelicolor*, and these pathogens allows for insights into the pathogenic mechanisms to be drawn, and how they differ from their saprophytic relatives (Bentley *et al.*, 2002). In addition to this, the analysis of the *S. coelicolor* genome is a crucial tool in tackling the rising antimicrobial resistance (AMR) crisis. Analysis of their innate antibiotic resistance genes allows for a greater understanding of how resistance mechanisms arise, and how to develop antimicrobial agents to circumvent these mechanisms.

1.2. Antibiotics and the resistance crisis

Antibiotics are chemical compounds that have bacteriostatic or bactericidal properties. Such compounds that occur naturally are utilised by organisms in order to out-compete other microbial or fungal species in order to survive. Through the advancements of human knowledge and technology, these compounds have been enhanced to treat microbial infections and can be produced naturally, semi-synthetically or synthetically (Manyi-Loh *et al.*, 2018). To overcome antimicrobial degradation and to survive exposure to these bactericidal compounds, bacteria have developed resistance mechanisms, or they express their innate resistome. Bacteria that produce antibiotic compounds will often contain intrinsic antibiotic resistance mechanisms. Such bacteria are often Gram-negative, as their outer membrane acts as a natural permeability barrier against a wide range of antibacterial compounds. Other inherently resistant bacteria will contain an intrinsic resistome that naturally lowers the bacterial level of susceptibility to antibiotics. Most producer strains have the ability to confer resistance as their genomes contain cryptic intrinsic resistance gene(s) often closely associated with the synthetic gene(s). Such intrinsic resistomes explain the emergence of resistance mechanisms, and how acquired resistance may spread across bacterial species

particularly in pathogens. Antibiotic resistance genes can be horizontally transferred between bacterial species via transposable plasmids. The lack of an outer membrane makes Gram-positive bacteria inherently less resistant to antibiotics, hence why a large variety of resistance mechanisms have arisen in Gram-positive bacteria.

The mass production of natural product penicillin during World War II revolutionised medicine. Its production was relatively easy and proved to be clinically effective. Penicillin significantly reduced not only the number of deaths from a wide range of bacterial infectious diseases, such as pneumonia, gonorrhoea, rheumatic fever and sepsis, but also the occurrence of various other nosocomial infections (Chain *et al.*, 1940). The introduction of penicillin sparked a surge of antibiotic discovery and production, known as the 'Golden Era' which peaked in the mid-1950s. Four significant sources of antibiotic compounds were distinguished: actinomycete natural products, other bacterial natural products, fungal natural products and synthetic antibiotics (Fig. 2). Filamentous actinomycetes are the largest source of antimicrobial compounds, as they account for almost 70% of all known natural products (Hutchings, Truman and Wilkinson, 2019). However, it did not take long for AMR to occur and render many newly discovered antibiotics redundant. Even at the time of penicillin's discovery, Fleming himself predicted the development of resistance, as his studies had already created resistant microbes, thus he speculated it was only a matter of time before resistance occurred in the body. Resistance against penicillin was first reported by Abraham and Chain in 1940 (Abraham and Chain, 1988) only 12 years after the discovery of penicillin by Fleming and antimicrobial resistance has only increased in prevalence since.

Two key methods facilitate the prevalence and spread of resistance: spontaneous genetic mutations that give rise to resistance or the acquiring of resistance mechanisms from other species via plasmids. Many species of bacterial pathogens already contain innate resistance genes, which are inadvertently 'selected' for when an antibiotic is used that kills sensitive bacterial cells, leaving only the resistant pathogen (Munita and Arias, 2016). This crisis being faced is worsened, predominantly, by the overuse, misuse and incorrect disposal of antibiotics. In 2018 it was reported that at least 20% of all antibiotics prescribed in primary care in England are inappropriate (Gov.uk, 2018). An incorrect antibiotic may have no effect on the intended target pathogen, but in most cases the pathogen is only partially damaged. This allows for bacterial 'learning' to be accelerated, thus acquiring resistance against bactericidal mechanisms faster than the natural rate. The same outcome is observed when the correct antibiotic is administered, but the treatment course is not fully completed. Again, the pathogen is not completely eliminated, thus can 'learn' how to resist future attacks from the same antibiotic. This ultimately leads to the selection for beneficial AMR genes, which proliferate within bacterial populations, consequently bacteria can begin to withstand antibiotic exposure.

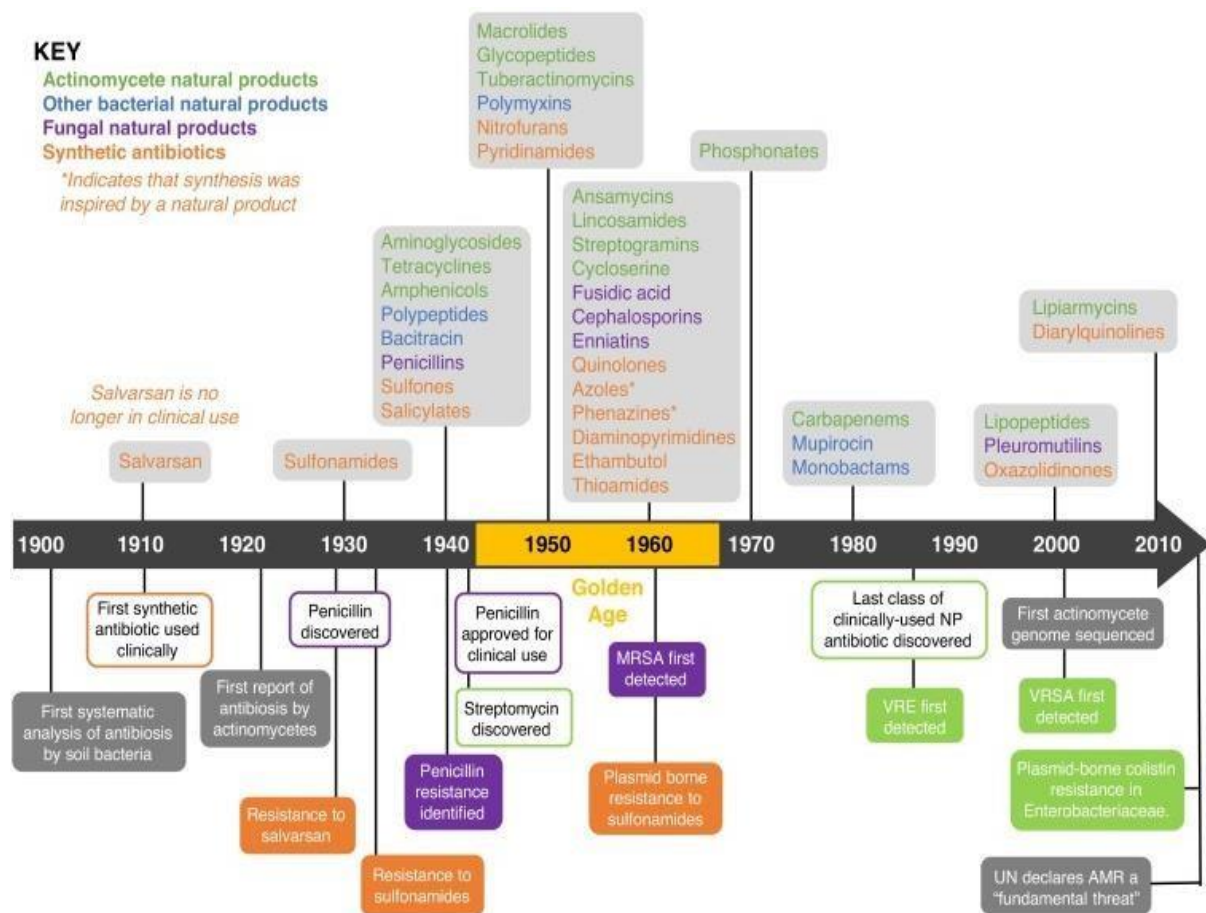


Figure 2. A timeline depicting the discovery of new antibiotic classes. The 'Golden Era' indicated by the yellow section of the timeline saw a huge surge in novel antibiotic discovery across all the antibiotic product types (Hutchings, Truman and Wilkinson, 2019).

An estimated 58,224 people in England had an antibiotic-resistant infection in 2022 and the number of deaths as a result of these antibiotic-resistant infections was 2,202 (Gov.uk, 2023b). Infections caused by antibiotic resistant pathogens in England rose by 4% in 2022 and the amount of antibiotics prescribed also rose by 8.4% (Gov.uk, 2023a). While these numbers may be increasing naturally (larger populations and more individuals in greater risk demographics), there will also be infections caused unnecessarily. The above data still falls below the rates observed during the COVID-19 pandemic, however, a concerning up-tick trend is currently being observed. An increase in infections and hospital admissions places a huge strain on the already thin-drawn NHS, as well as other global healthcare providers. According to the World Health Organisation, in 2019, 4.95 million global deaths were associated with AMR and 1.27 million deaths were directly caused by resistant pathogens (WHO, 2023).

As such, AMR is not only a major and growing threat to public health around the world but also has the potential to immensely impact the economy. The Fleming Fund, subsidised by the UK government, was awarded £210 million to tackle the global crisis and provide aid to local communities and to vulnerable communities overseas (The Fleming Fund, 2023). The UK government has so far published two five-year 'action plans' on how the country will tackle the crisis we are facing. The latest plan outlines four key themes: the overuse of antibiotics, how to optimise the antibiotics we do have, the development of novel antibiotics and spreading awareness of the AMR crisis (HM Government *et al.*, 2024). In particular, the third theme outlined in the 2024-2029 action plan focuses on the reviving of investment into the discovery of novel antibiotics, or modification of existing antibiotics.

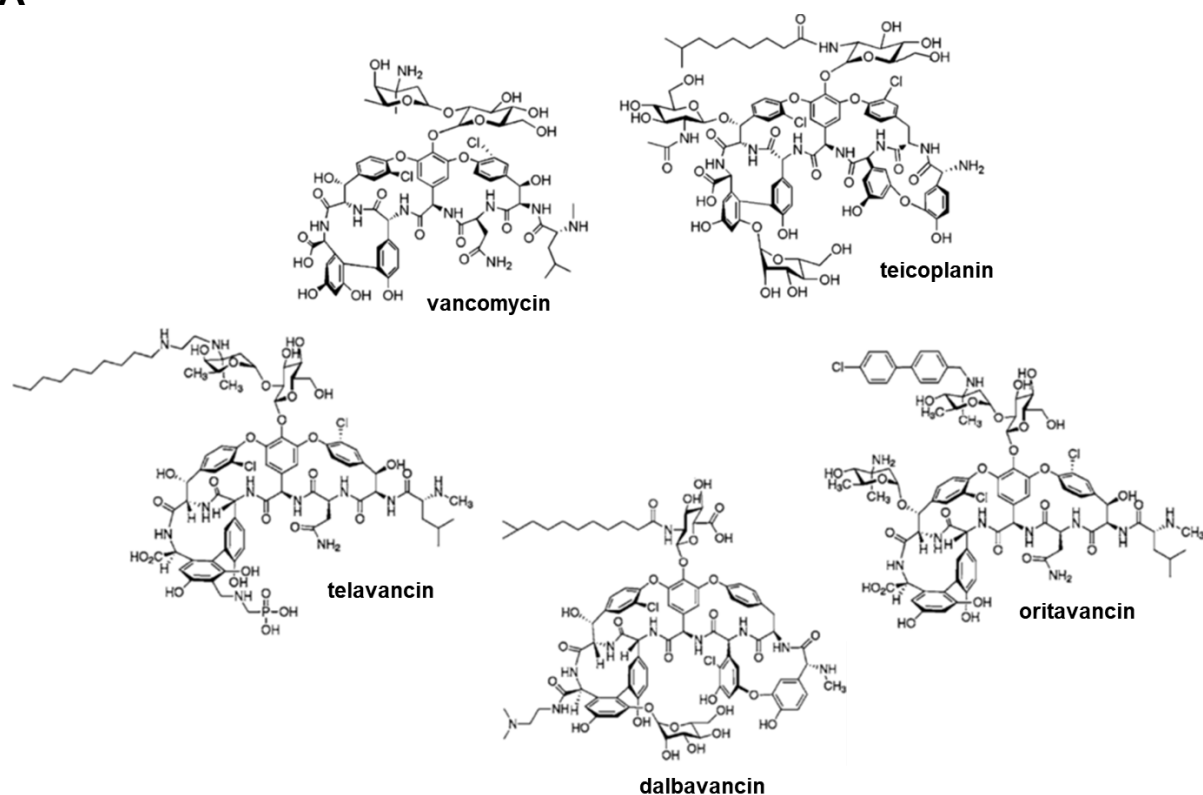
1.3. Glycopeptide antibiotics

The analysis conducted by Murray *et al.*, 2022 saw that the AMR crisis was directly responsible for 1.27 million global deaths and was associated with 4.95 million deaths in 2019 alone. In comparison, the next leading causes of death saw 860,000 deaths directly caused by AIDS/HIV and 640,000 direct deaths caused by Malaria in the same timeframe. Drug resistant pathogens are the leading cause of bacterial infection related deaths. The top six drug resistant bacterial pathogens were directly responsible for 929,000 deaths and associated with 3.57 million deaths in 2019. Of these resistant bacterial deaths, methicillin resistant *Staphylococcus aureus* (MRSA) itself was responsible for over 100,000 deaths. Vancomycin is reserved as a last resort antibiotic to treat infections caused by the multidrug resistant pathogen, MRSA, however, the occurrence of vancomycin resistant MRSA (VRSA) has been isolated in various hospitals in Western US since 2002.

Vancomycin belongs to a glycopeptide antibiotics family. Glycopeptide antibiotics act as bactericidal agents by targeting the lipid II component of bacterial cell walls, which results in an inhibition of bacterial peptidoglycan (PG) synthesis. Glycopeptides specifically bind to the C-terminal L-Lysine-D-Alanyl-D-Alanine subunit of the cell wall PG precursor lipid II. Strong binding between the glycopeptide structure and bacterial PG is established by hydrogen bonding which prevent the bacterial transglycosylation and/or transpeptidation steps required for a mesh-like, three dimensional structure of the mature PG (Zeng *et al.*, 2016). By inhibiting the normal PG synthesis and cross-bridge formation the bacterial cell wall is ultimately weakened, and eventually the cell will lyse. Glycopeptides are composed of several amino acid residues and characterised by their glycosylated aromatic ring and heptapeptide backbone structure (Fig. 3A). Glycopeptides have unique tricyclic or tetracyclic heptapeptide cores that are usually glycosylated. The additional glycosylation of these species increases the binding affinity to target bacterial cell walls, as well as the solubility and overall pharmacokinetics. Glycopeptides also contain a modified hydroxyphenylglycine residue that facilitates binding to bacterial cell walls. The vancomycin structure consists of a heptapeptide core with five glycosylated aromatic rings (Svetitsky, Leibovici and Paul, 2009) and it contains a key L-3,5-dihydroxyphenylglycine residue. This residue facilitates vancomycin's ability to interact with the D-Alanyl-D-Alanine (D-Ala-D-Ala) terminus present within the bacterial cell wall precursor (Fig. 3B). Teicoplanin shares the same heptapeptide core structure as vancomycin, however teicoplanin has an additional aromatic amino acid and an aliphatic N-Methyl-L-Leucine, plus multiple hydroxyphenylglycine units. This structure facilitates greater glycosylation and lipidation, which results in improved activity and pharmacokinetics. Teicoplanin also has an additional lipid tail that facilitates the disruption of PG biosynthesis, thus classified as a lipoglycopeptide (Van Bambeke, 2015). New glycopeptide antibiotics are also being developed in the face of the rising AMR crisis. Three lipoglycopeptide antibiotics have been developed in recent years and have been clinically available since 2014. Dalbavancin and oritavancin are semi-synthetic, and telavancin is entirely synthetic. These new lipoglycopeptides have a wide range of activity and have proven effective at treating gram-positive bacterial infections, including MRSA, as they have been optimised against resistant strains. These new lipoglycopeptides are shown to have improved pharmacokinetics and an enhanced efficacy against resistant Gram-positive bacteria over vancomycin. However, there are issues concerning increased toxicity in liver and kidney cells caused by these new glycopeptides (van Groesen *et al.*, 2022). For this reason, vancomycin and teicoplanin are still used as the most commonly prescribed effective treatments against Gram-positive pathogens, including MRSA and enterococci.

The most significant difference between vancomycin and teicoplanin is the additional lipid tail

A



B

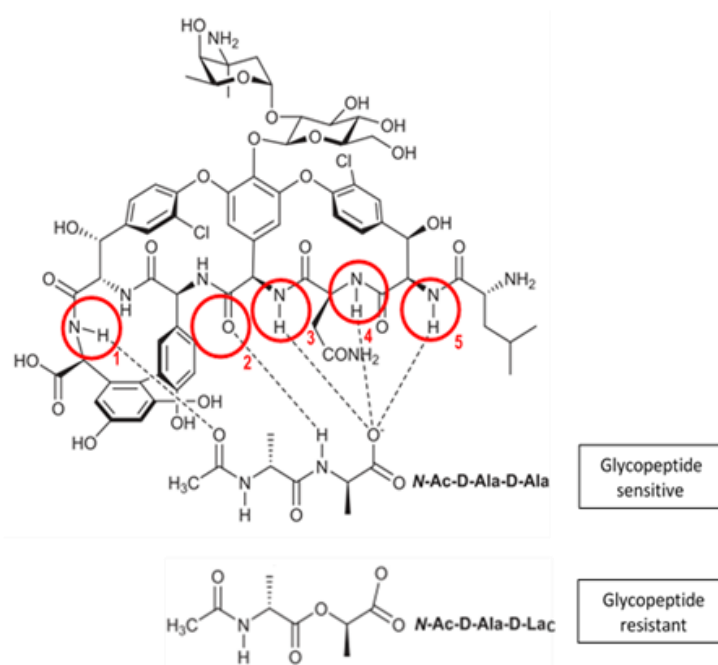


Figure 3. (A) The structures of clinically approved glycopeptide antibiotics. (B) Vancomycin binds to D-Ala-D-Ala terminus of cell wall precursor with five hydrogen bonds.

structure present in teicoplanin. This lipid tail structure is fundamental in the improvements observed in teicoplanin's greater binding affinity and its increased pharmacokinetics over vancomycin. Teicoplanin's lipid tail increases the overall lipophilicity of the structure, which in turn allows teicoplanin to more readily enter lipid-rich tissue and to leave the bloodstream. This ability is greatly improved as compared to vancomycin, especially due to the tail's hydrophobic nature. The tail structure of teicoplanin also facilitates interactions with bacterial membranes and enhances binding to bacterial proteins. When in the bloodstream, teicoplanin exhibits strong binding to plasma proteins, primarily albumin and α 1-acid glycoprotein (Roberts *et al.*, 2013). Teicoplanin's hydrophobic tail also enhances plasma protein binding through interactions with hydrophobic pockets within the plasma proteins. Due to the strong protein binding abilities of Teicoplanin, the free drug concentration within the plasma is relatively low. Once teicoplanin enters tissue, it remains for a greater period, as its stronger binding to lipid-rich tissue slows down renal clearance. This implies an improved therapeutic capability of teicoplanin at lower serum concentrations than vancomycin. These combined factors contribute to the long half-life of Teicoplanin and its greater therapeutic effects, than over vancomycin.

The primary mode of action for teicoplanin is via interactions with the D-Ala-D-Ala terminus of PG precursor as shown in vancomycin, but has an additional mode of action due to its unique lipid tail. The tail differentiates Teicoplanin from vancomycin and enhances its antimicrobial properties. The tail structure has been known to facilitate anchoring into the bacterial membrane within a close proximity to the lipid II substrates. This altered interaction has been found to be energetically favourable over vancomycin's interactions and gives teicoplanin enhanced antibacterial properties (Zeng *et al.*, 2016). Thus, the lipid tail in teicoplanin not only improves the antibiotic's potency but also its pharmacokinetic properties, making it a valuable option in the treatment of Gram-positive bacterial infections, especially in cases where conventional antibiotics might be less effective.

Vancomycin resistance first emerged in the 1980's, with clinical isolates of Vancomycin-resistant *Enterococci* (VRE) strains, followed by the appearance of a series of VRSA in US hospitals over a decade later. All these clinical isolates harbour the *vanSRHAX* resistance cluster, encoding genes required for sensing (via a signalling system encoded by *vanRS*) and coordinating the response to the presence of an antibiotic (via *vanHAX*) consequently reprogramming the termini of cell wall PG precursors from D-Ala-D-Ala to D-Alanyl-D-Lactate (D-Ala-D-Lac) or D-Alanyl-D-Serine (D-Ala-D-Ser). This remodelling of the bacterial cell wall precursors significantly reduces the vancomycin binding affinity (1000-fold for D-Ala-D-Lac, 6fold for D-Ala-D-Ser) thus facilitates vancomycin-resistance, but still allows for the formation of cross-links in the PG layer. The emergence of VRSA was due to a horizontal gene transfer:

MRSA acquired a vancomycin resistance gene, *vanA*, from another common nosocomial Gram-positive pathogen, *Enterococcus*, that had already developed a resistance mechanism against vancomycin. So far, there are 52 VRSA strains that have been identified globally (Cong *et al.*, 2020). The resistance mechanism in VRE, mediated by the *vanA* gene, was transferred to MRSA by transposon Tn1546, which gave rise to VRSA. At the current moment, 90% of VRSA cases can be treated with other antimicrobial agents, but how long will this be the case? The continual emergence of new resistant pathogens highlights the urgent need for new antimicrobial agents or for current antibiotics to be improved so they can overcome these resistance mechanisms.

1.4. Glycopeptide resistance studies in *S. coelicolor*

Cell wall homeostasis studies using *S. coelicolor* as a model system, have begun with the characterisation of the σ^E system and extended to a genome-wide dynamic via functional genomic approaches (Tran *et al.*, 2019). σ^E is an extracytoplasmic function σ factor and is a key regulator of cell envelope stress responses. The expression of genes encoding the σ^E and other genes in the SigE operon are regulated by the transcription levels of the two-component system (TCS) CseB/CseC. Transcriptional profiling found that during cellular stress, upregulation of the σ^E gene occurs and it triggers the activation of many other genes involved with cell envelope maintenance and repair. This σ^E regulon was found to have roles in PG biosynthesis, cell wall remodelling and production of proteins that maintain the integrity of the cell envelope. Also, many of these genes were implicated in antibiotic resistance mechanisms such as the vancomycin resistance gene cluster.

Bacteria that produce antibiotics have evolved mechanisms that provide immunity to the action of the antibiotic, and often these immunity mechanisms have co-evolved with antibiotic biosynthesis genes to protect the producing organisms. Many actinomycete species are antibiotic producers that contain innate antibiotic resistance, which make them a phylum of significant interest in the antibiotic resistance study. Interestingly however, the non-glycopeptide producer strain *S. coelicolor* harbours a cluster of 7 genes (*vanSRJKHAX*) in its genome which confers a strong resistance to vancomycin and some other glycopeptide antibiotics (Fig. 4). Upon vancomycin exposure, this inducible system expresses its *vanRS* TCS which ultimately causes expression of all seven *van* genes. The *vanHAX* genes are responsible for modifying the termini of the PG cell wall precursor from D-Ala-D-Ala to D-Ala-D-Lac (Hong *et al.*, 2004; Hong *et al.*, 2005). Another significant gene within the vancomycin resistance cluster is the actions of *VanK*, which was identified as a novel Fem protein. At the time of the study that characterised *VanK*, there were no equivalent counterparts in other

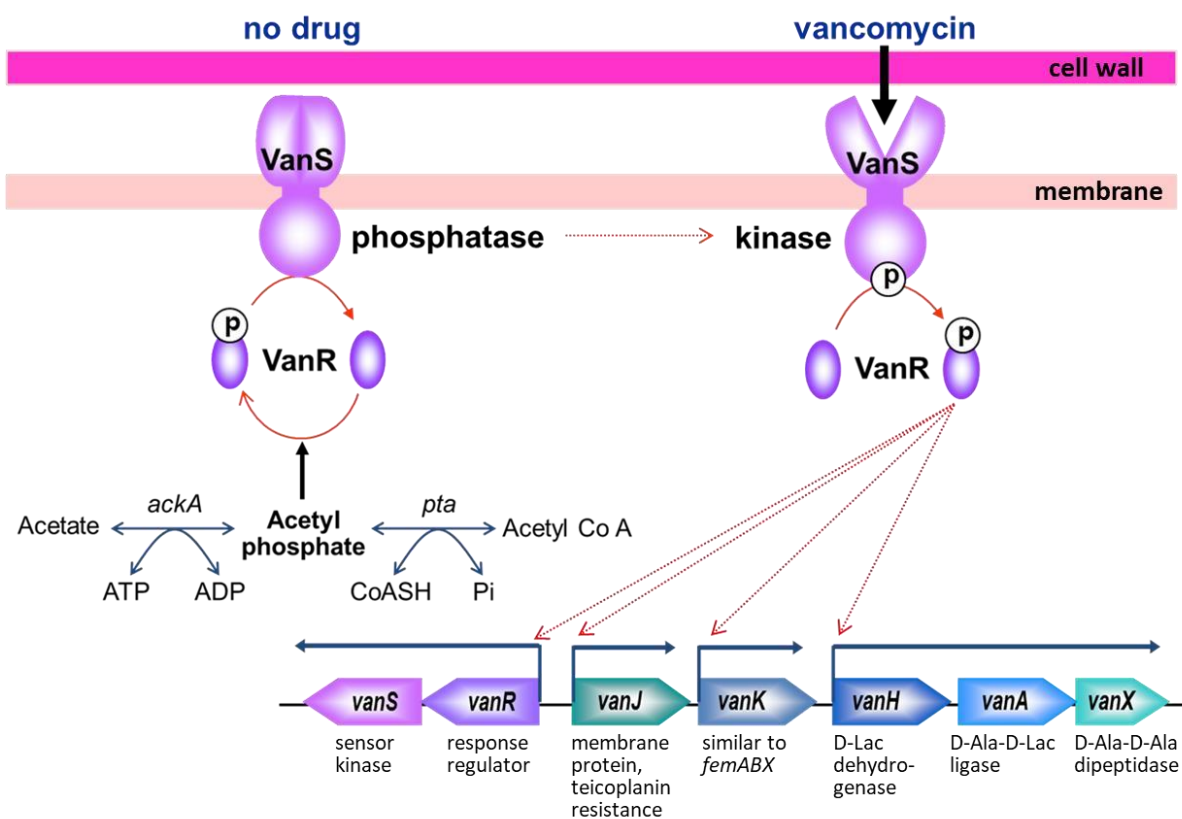


Figure 4. *S. coelicolor* genome segment containing the vancomycin resistance gene cluster (*vanSRJKHAX*). These genes are regulated by the transcription levels of the *vanR/vanS* two component signal transduction system (Hong et al., 2004).

vancomycin resistant pathogens' gene clusters (Hong *et al.*, 2004). However, it has now been discovered that these Fem proteins (FemA, FemB and FemX) are essential in the MRSA mechanism of resistance (Akkiraju *et al.*, 2023). While VanK is specific to vancomycin resistance in *S. coelicolor*, its function is similar to that of the MRSA Fem proteins, which are directly associated with the modified PG precursor sequence and subsequent bacterial cell wall biosynthesis, by adding a single glycine cross-bridge to a modified D-Ala-D-Lac terminating PG precursors.

Previous studies demonstrated that the resistance to vancomycin in *S. coelicolor* is mediated by VanKHAX but not by VanJ (Novotna *et al.*, 2012). Instead, VanJ is required for teicoplanin resistance, and encodes a membrane protein oriented with its C-terminal enzymatic active site exposed to the extracytoplasmic space. VanJ also confers resistance to a broad range of semi-synthetic teicoplanin derivatives, but not generally to vancomycin derivatives (Novotna *et al.*, 2012). Over-expression of *vanJ* in a Δ *vanHAX* mutant strain of *S. coelicolor* produced the same increase in teicoplanin resistance as in the wild-type strain, indicating that this resistance activity is not dependent in any way on the reprogramming of the cell wall precursor biosynthesis encoded by the VanHAX proteins. Homologues of *vanJ*, that facilitate glycopeptide resistance (*vanY*, *vanX*, *vanZ*, *vanK*, *vanA/vanH*) (Beltrametti *et al.*, 2007) are found ubiquitously in actinomycetes, therefore characterization of the divergent enzymatic activity of VanJ and its homologues will contribute to defining the molecular mechanisms important for teicoplanin resistance, and consequently also provide information on its activity.

Further characterisation of the response of *S. coelicolor* cells to teicoplanin using transcriptomics identified an interesting set of genes, SCO6357-53, which are a suspected polycistronic unit (Fig. 5) (Connelly *et al.*, 2021). The transcriptional start site was identified as SCO6357 and the operon is also located immediately downstream of a σ^E regulon promoter sequence. SCO6355 and SCO6356 encode type 2 PAP2 domains which were confirmed to affect the sensitivity of *S. coelicolor* towards a range of antibiotics that target the cell envelope. SCO6354-53 encodes for a histidine kinase and a response regulator in a TCS. Overexpressing one of these genes, SCO6355, produced a hypersensitive-phenotype towards the antibiotic, duramycin, and intermediate levels of sensitivity towards vancomycin-like glycopeptide antibiotics even though *S. coelicolor* shows high-level resistance to vancomycin with its native *van* genes (Connelly *et al.*, 2021). Additional functional genomic studies showed the immense cellular stress induced by teicoplanin exposure. It was found that during teicoplanin exposure, the genes involved with σ^E signal transduction system remain active for the entire exposure duration. But when exposed to other glycopeptides, the expression of the σ^E signal transduction system declines as the *van* genes are induced.

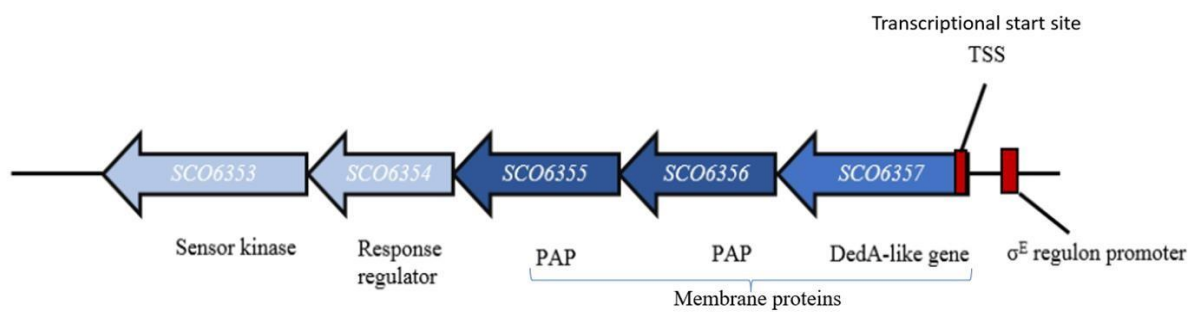


Figure 5. Gene organisation of SCO6357-53 operon. The transcriptional start site (TSS) and σ^E regulon promoter are shown in red. The σ^E promoter is found 88 bp upstream of the TSS (Connolly *et al.*, 2021).

The ultimate goal of this study is to follow on from Connelly's findings by mutagenizing some interesting candidate genes including the SCO6357-53 operon; this will give better insights into the subtle mechanistic differences that arise from structural differences between glycopeptide antibiotics. Further exploration of these mechanisms will be fundamental in the development of more effective glycopeptide antibiotics that work against glycopeptide resistant clinical pathogens. However, the primary focus of this study is to design and construct a simple but efficient CRISPR-Cas9 gene editing system in order to mutagenize these genes.

1.5. The CRISPR-Cas gene editing system

1.5.1. Discovery of CRISPR-Cas

CRISPR is an acronym for Clustered Regularly Interspaced Short Palindromic Repeat. This short DNA sequence was first found through analysis of the *Escherichia coli* genome in 1987 but also exists in the genome of a wide range of bacterial organisms and archaea (Sola *et al.*, 2015). CRISPR is originally derived from DNA fragments of bacteriophages that had previously infected the host organisms. The short palindromic repeats of the CRISPR arrays are often surrounded by unique sequences which are called 'spacers', which when aligned with known DNA libraries, will correspond with bacteriophage DNA. When a prokaryote overcomes a bacteriophage attack, the first stage of the immune response is to capture phage DNA and insert it into a CRISPR locus in the form of a spacer in its own genome. This builds an immune 'memory' for future attacks from the same bacteriophage, hence CRISPR plays a key role in the antiphage defence system of bacteria and archaea (Chen, 2017; Ishino *et al.*, 2018). A typical CRISPR array consists of an AT-rich leader sequence followed by short repeats that are separated by unique spacers. CRISPR repeats can vary widely (23-55 nucleotides) though they typically range in size between 28 and 32 nucleotides. Their partially palindromic or dyad symmetric nature allows them to form a secondary structure such as a hairpin like stem-loop structure. The size of spacers in different CRISPR arrays can also vary (21-72 nucleotides) but they typically range between 32 and 38 nucleotides. New spacers appear rapidly as part of the immune response to phage infection but they are usually less than 50 units of the repeat-spacer sequence in a CRISPR array (Barrangou and Marraffini, 2014; Jiang and Doudna, 2015).

Along with the discovery of CRISPR, the presence of genes associated with CRISPR repeats, so called *cas* (Crispr associated system), was also acknowledged (Cobb *et al.*, 2015). These genes are often found in close proximity to the CRISPR loci as a small cluster and they encode proteins consisting of helicase and nuclease motifs representing endonuclease activity

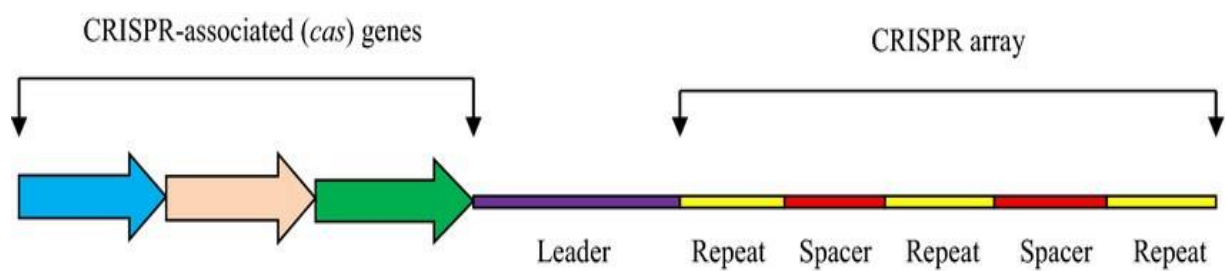


Figure 6. Structure of CRISPR locus. In CRISPR locus, each 'repeat' sequence is flanked by 'spacer' sequence and these spacers match with the genomic sequences found in virus, plasmid, or pathogen. The leader sequence and the CRISPR-associated (*Cas*) genes are located upstream of the CRISPR array (Barman and Chakraborty, 2020).

to cut specific DNA strands indicating their roles in the dynamic structure of the CRISPR loci (Fig. 6). The CRISPR arrays and the associated Cas gene operon is highly conserved across both bacteria and archaea, thus implies that this CRISPR-Cas system is crucial for the survival of these organisms.

CRISPR-Cas systems can be simply divided into two classes. Class 1 uses a complex of multiple Cas proteins to degrade foreign nucleic acids (type I and III). Class 2 utilises a single large Cas protein (type II). Class 1 is a more commonly found system as many organisms contain multiple CRISPR-Cas systems (Jinek *et al.*, 2012).

1.5.2. CRISPR-Cas9

Among the two CRISPR-Cas classes, Class 2 (type II) is a simpler version derived from *Streptococcus pyogenes* which relies on the protein Cas9 for its CRISPR-Cas system (Jinek *et al.*, 2012). Cas9 is an endonuclease enzyme that uses CRISPR sequences as a guide to recognize and open up specific strands of DNA that are complementary to the CRISPR sequence. The native Cas9 endonuclease is a four-component system that eventually forms a complex with CRISPR RNA (crRNA) and associated *trans*-activating CRISPR RNA (tracrRNA). These RNAs are transcribed from the spacer sequences found in the Class 2 (type II) system. crRNA sequences are found in CRISPR loci but there are large variations in tracrRNA location size and sequence identity. The tracrRNA is complementary to the repeat sequences found in pre-crRNA, which enables hybridisation of the two RNAs before making a complex with Cas9 protein. The formation of the crRNA-tracrRNA duplex is then processed by double-stranded RNA-specific ribonuclease RNase III and other possible nucleases. Cas9 endonuclease finally interacts with this duplex to form a complex which then scans for the presence of foreign DNA protospacer adjacent motifs (PAMs) in the target genome. When this complex binds to spacer sequences and finds a PAM, nuclease activity is then triggered by activation of the HNH and RuvC domains in the Cas9 protein (Fig. 7).

The Cas9 protein is a bilobed structure, composed of the NUC lobe and the REC lobe. The NUC lobe constitutes the two nuclease domains (HNH and RuvC) and a C-terminal PAM interacting domain. The second REC lobe is composed of three α -helical domains (REC1-3), and an arginine-rich helix; this long helix bridges the two lobes to form the Cas9 structure (Fig. 8B). Activation of the nuclease domains (HNH and the RuvC) introduces a double-strand break (DSB) in the target sequence; each domain cleaves one strand of the sequence targeted by the complex. This naturally occurring system can be hijacked and manipulated to target specific genes of interest within a target organism (Wilkinson *et al.*, 2019; Zhu, 2022) (Fig. 8).

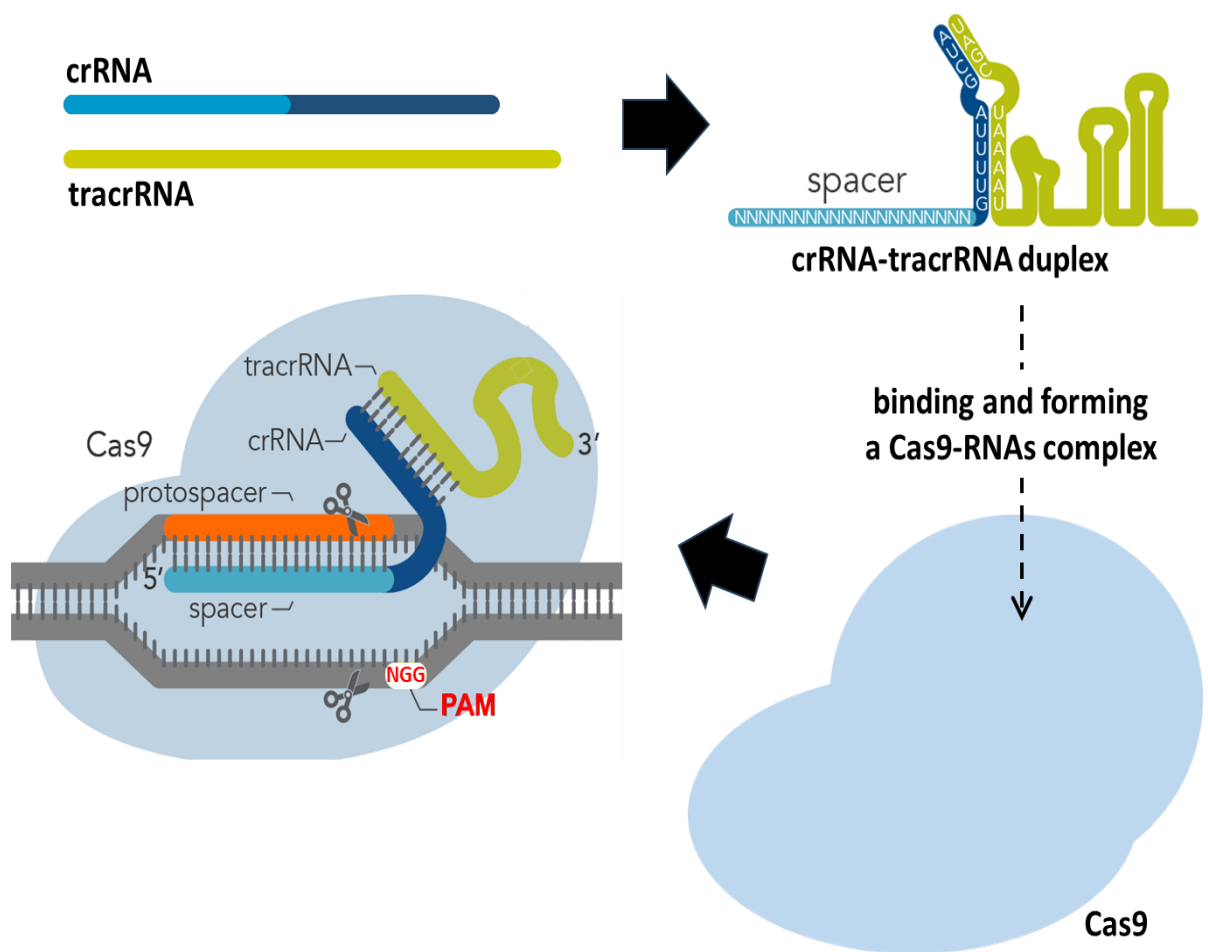
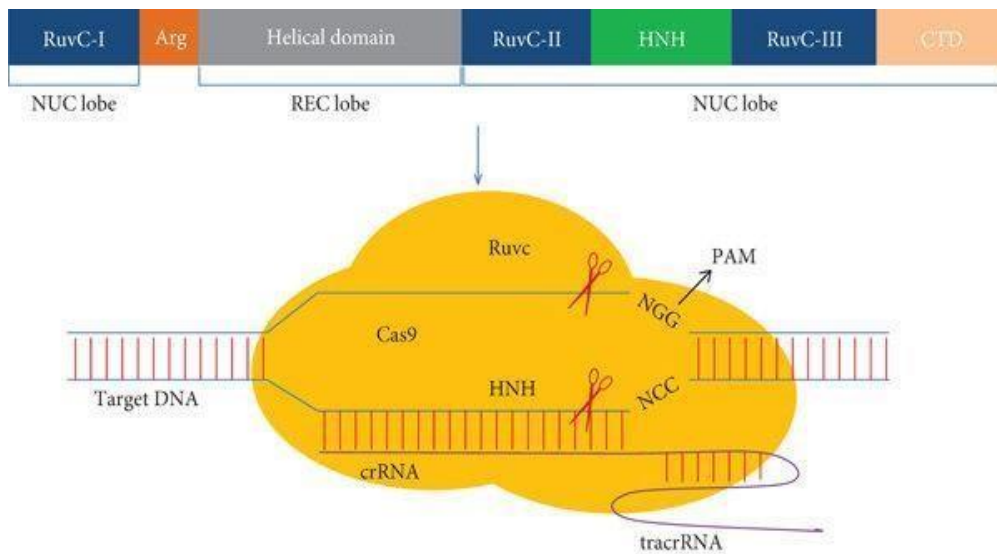


Figure 7. A native Class 2 (type II) CRISPR-Cas9 system in *Streptomyces pyogenes*.

A



B

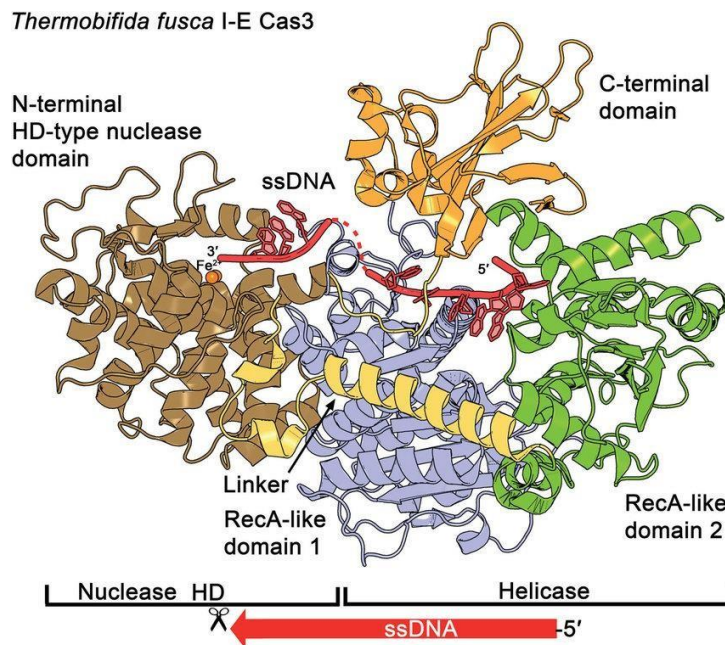


Figure 8. (A) The overall structure of the Cas9 protein and the schematic diagram of the Cas9–sgRNA–target DNA complex in a Class 2 (type II) system. Cas9 protein comprises NUC and REC lobes. The REC lobe contains a long alpha helix and two REC domains. NUC lobe consists of PI, RuvC, and HNH domains. RuvC and HNH domains have cleavage functions, which can recognize the PAM site and cut the target DNA under the guide of the crRNA: tracrRNA complex (adapted from Zhu, 2022). (B) The crystal structure of the type I-E DNA nuclease Cas3 of *Thermobifida fusca*. This reveals two tandem RecA-like domains, one HD-type nuclease domain and a C-Terminal Domain located at the top of the ensemble. The core helicase, containing two RecA-like domains, forms a cleft that locates the residues for the binding of NTP, Mg²⁺ ions and the ssDNA substrate (Plagens *et al.*, 2015).

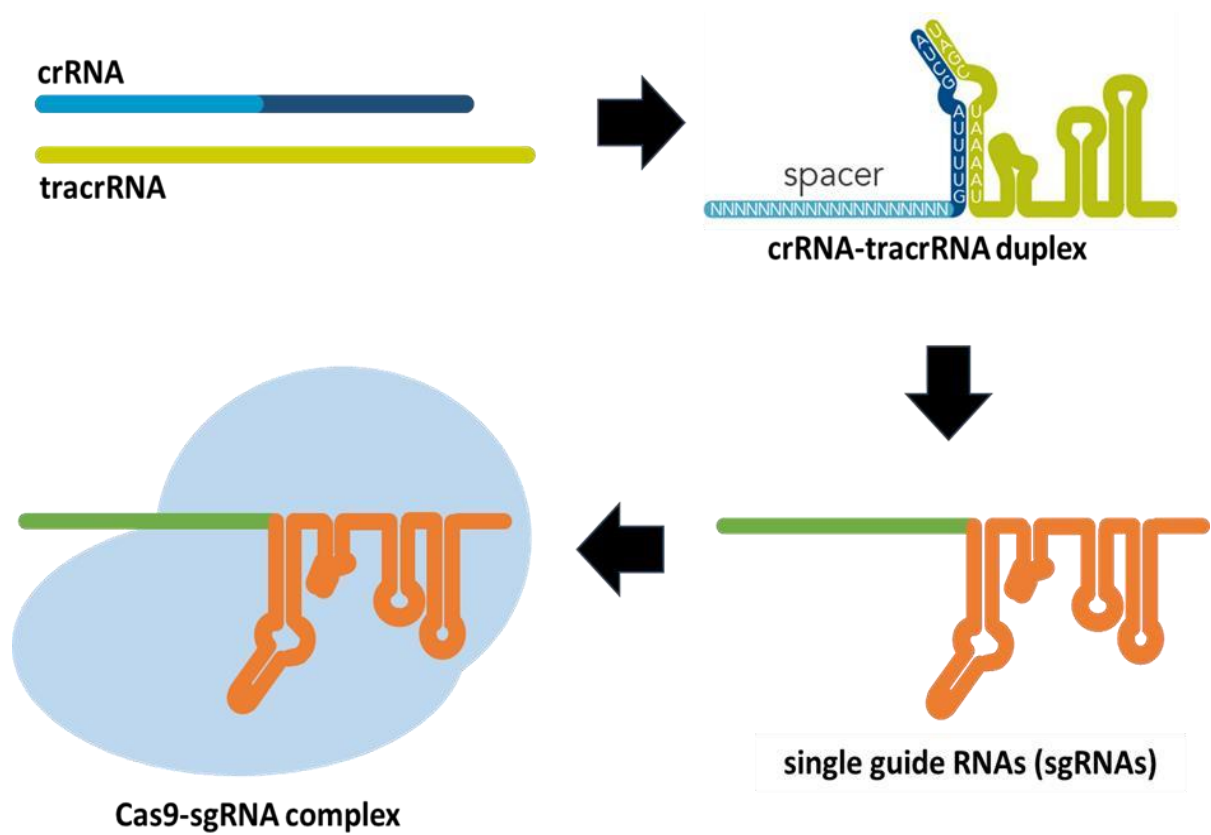


Figure 9. Application of the native CRISPR-Cas9 system to a simple gene editing system. The diagram represents a more manageable 2-component CRISPR-Cas9 system by using a "single-guide RNA". This single-guide RNA is a fused crRNA and tracrRNA chimaera which can be artificially generated.

This complex CRISPR-Cas9 system was eventually developed into a more manageable 2-component system by Emmanuelle Charpentier and Jennifer Doudna in 2012 by fusing the two RNA molecules into a "single-guide RNA". When sgRNA is combined with Cas9, this complex can find and cut target DNA, specified by the guide RNA (Fig. 9). This customised version of CRISPR-Cas9 is currently regarded as the most reliable gene editing tool as it acts as a precise pair of molecular scissors that can cut a target DNA sequence. Thus, a component of the bacterial immune system against bacteriophage infections has been repurposed and developed into a generic and powerful gene editing tool. The technique was recognised by the Nobel Prize in Chemistry, awarded to Emmanuelle Charpentier and Jennifer Doudna in 2020 (Jinek *et al.*, 2012).

1.5.3. CRISPR-Cas9 based gene editing systems developed in *Streptomyces*

The CRISPR-Cas9 system has revolutionised gene editing across a variety of organisms, including bacteria like *Streptomyces*. *Streptomyces* species are important in biotechnology and natural product discovery due to their ability to produce a wide range of secondary metabolites, including antibiotics. However, traditional genetic manipulation methods in *Streptomyces*, such as the Re-Direct (Recombination-Directed) gene editing system, have limitations due to low efficiency, labour-intensive protocols and strain-specific challenges. Moreover, securing an entire cosmid DNA library covering the whole genome of the host cell presents the most prominent disadvantage related to the large size and complexity of the cosmid. This makes it difficult to transfer cosmids, and they may become unstable, making gene editing more laborious and less predictable. The CRISPR-Cas9 system offers an alternative approach that can potentially overcome some of these challenges.

The application of *S. pyogenes* CRISPR-Cas9 system in Actinomycetes, particularly in *Streptomyces* species, sought to engineer their naturally occurring biosynthetic pathways and to influence the expression of secondary metabolites. The first engineered CRISPR-Cas9 system in *Streptomyces* was configured by Cobb *et al.* (2015). In this study, two versions of editing plasmids, pCRISPomyces-1 and -2, were developed as rapid but highly efficient multiplex genome editing tools and tested for three different streptomyces strains. These plasmids contain an integrated codon-optimised *cas9* gene and a temperature-sensitive pSG5 origin for *Streptomyces* expression. The nucleotide sequences of *S. pyogenes* originated *cas9* gene has been altered to account for the difference in G+C content (35% for *S. pyogenes* vs 72% for general *Streptomyces*) and codon usage between *S. pyogenes* and *Streptomyces* species. pSG5 enables replication at temperatures in the low 30°C range or below in *Streptomyces* but this plasmid DNA can be lost from the host cells when culturing them at

37°C. The plasmid provides the rolling circle replication initiator protein; thus, it facilitates continuous genome editing within the cell (Muth, 2018). The *lacZ* cassette enables the assembly of spacer sequences via one-step Golden Gate Assembly and additional editing templates can be inserted at the *Xba*I site. Transcription of the *Cas9* protein is under the control of the strong constitutive promoter *ermE*^{*} (erythromycin resistance promoter) commonly used in actinomycetes (Ye *et al.*, 2020), along with terminators from phages fd, λ, and T7. Inclusion of apramycin resistance gene, *aac(3)-IV*, allows selection in both *E. coli* and *Streptomyces* hosts. The *colE1* origin enables replication in *E. coli*, and the RP4 origin of transfer *oriT* enables conjugative transfer of the plasmid DNAs from *E. coli* to *Streptomyces* hosts. pCRISPomyces-1 plasmid contains both tracrRNA and CRISPR array expression along with a *cas9* gene, whereas pCRISPomyces-2 contains an sgRNA expression cassette and *cas9*.

The pCRISPomyces-2 editing plasmid was found to have a significantly higher editing efficiency (Cobb *et al.*, 2015). Many further developments of *Streptomyces*-optimised gene editing plasmids following a CRISPR-Cas base have been made following the development of the pCRISPomyces plasmids. Most developed plasmids incorporated the CRISPR-Cas9 mechanism of action, however some plasmids utilising the CRISPR-Cas12 system have also been developed. The next plasmid developed, pKCcas99dO (Huang *et al.*, 2015), contained a temperature-sensitive backbone and a codon-optimised *Cas9* nickase (*Cas9D10A*) to introduce ss breaks, thus reducing the risk of off target effects. This plasmid also included a target-specific guide RNA, two HDR templates and induced expression of the *Cas9* through the promoter *tipAp*. This gene-specific editing cassette found between the *Spe*I and *Hind*III sites was designed to target and delete *actII-orf4* in *S. coelicolor*. However, this system can only be utilised in targeting and deleting single target genes or gene clusters.

A separate plasmid was also developed by the Tong group alongside the pCRISPR-Cas9 plasmid. This plasmid implemented the use of a catalytically inactive *Cas9* (d*Cas9*), hence referred to as pCRISPR-d*Cas9*. The plasmid does not have any endonuclease activity due to the inactivation of both nuclease domains (RuvC1 and HNH). As a result, the d*Cas9* can only bind to specific sites, as directed by the gRNA. Instead of inducing DSBs into the genome, the d*Cas9* is fused with a transcriptional repressor domain which can block the transcription of the target gene (CRISPR-interference, CRISPRi) (Tong *et al.*, 2020).

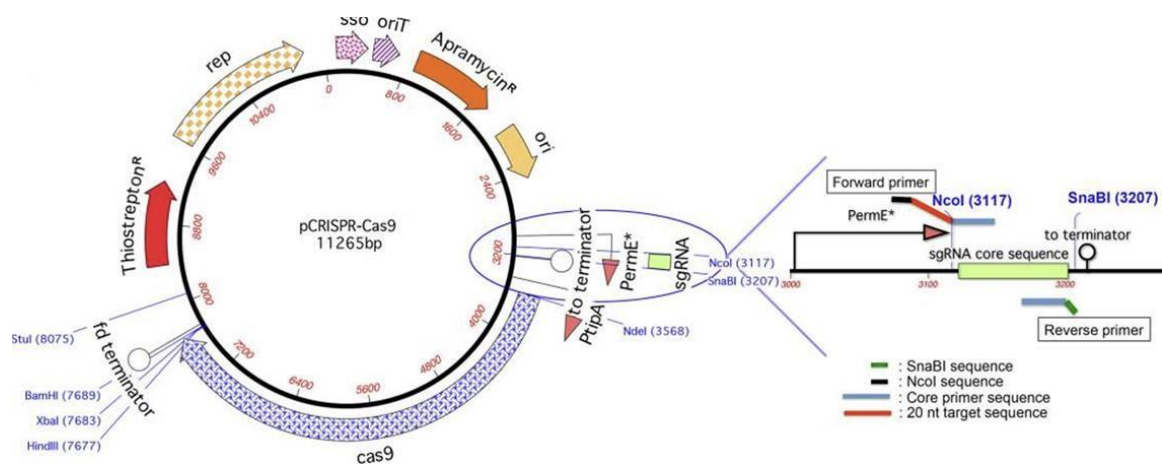
Additionally, other advancements in *Streptomyces*-optimised editing plasmids have seen the rise of CRISPR-Cas based plasmids that utilise the *Cas12* (Cpf1) editing mechanism, rather than *Cas9*. *Cas12* is a class 2 type V CRISPR-Cas system which acts via introducing staggered cuts into the DNA sequence, rather than the blunt-ended cuts seen by *Cas9* (Tong

et al., 2015). The introduction of sticky-ends helps facilitate more precise insertions or deletions of target sequences. The *Cas12* is also directed by a gRNA to target DNA sequences, however, when bound to target sequences *Cas12* can also introduce collateral cleavage activity surrounding the intended target site. This is due to the different PAM requirement of *Cas12*, and can expand the range of targetable sequences, hence why additional nuclease activity may be observed. The multiplexing abilities of CRISPR-Cas12 plasmids have been investigated in the pKCCpf1 plasmid (Li *et al.*, 2018) and in the CRISPR-Cas AsCas12j-2 plasmid (Tan *et al.*, 2024). The CRISPR-Cpf1 based pKCCpf1 plasmid could also be used as a transcriptional repressor, thus can also act via a CRISPRi mechanism. The study by Tan *et al.* (2024) sought to investigate the wide range of natural products, including antibiotics, contained within *Streptomyces* species' genomes. Gene knockouts through use of Cas12j were used to produce desired metabolites and to study the functions of the associated genes.

The gene editing system used in the present study was based on the plasmids pCRISPR-Cas9 and pCRISPR-cBEST, developed by Tong and co-workers (Tong *et al.*, 2015; Tong *et al.*, 2019). The plasmid pCRISPR-Cas9 was created based on plasmid pGM1190 which was derived from the temperature-sensitive plasmid pSG5. pCRISPR-Cas9 contains a 20 base 'single guide RNA' (sgRNA) which the Cas9 uses as a template to identify a DNA sequence to cut (Fig. 10A). This sequence must be followed by a trinucleotide PAM consisting of NGG. A guide sequence including a desired target sequence can be produced by PCR using the plasmid DNA as a template and inserted into the plasmid between the *NcoI* and *SnaBI* restriction sites. The expression of the guide sequence is under the control of a strong and constitutive promoter, *ermE**. The *cas9* gene in this plasmid is also a codon-optimised version for the expression in *Streptomyces*. Transcription of the *cas9* gene is under the control of a strong, thiostrepton inducible promoter, *tipA*. Inducing the *tipA* promoter requires a thiostrepton responsive activator, TipA, however genes encoding this protein do not need to be present in the plasmid because the *tipA* gene is present in the genomes of most *Streptomyces*, including the model actinomycete, *S. coelicolor*. For streptomycetes' host cells to survive exposure to thiostrepton, the plasmid also requires the presence of a thiostrepton resistance gene, *tsr*. The plasmid also contains an apramycin resistance gene, *apmR*, as apramycin is used as a selection marker in both *E. coli* and *Streptomyces* hosts. The origin enables replication in *E. coli* (*ori*) and an origin of transfer (*oriT*) is also required for bacterial conjugation to occur, i.e. the transfer of plasmid DNA from *E. coli* to a recipient *Streptomyces* strain.

The remaining features of pCRISPR-Cas9 plasmid include various promoter sequences, binding sites and scaffolds to facilitate the gene targeting and editing functions. Various primers and promoters are also found within the plasmid sequence: pBR322ori-F, L4440,

A



B

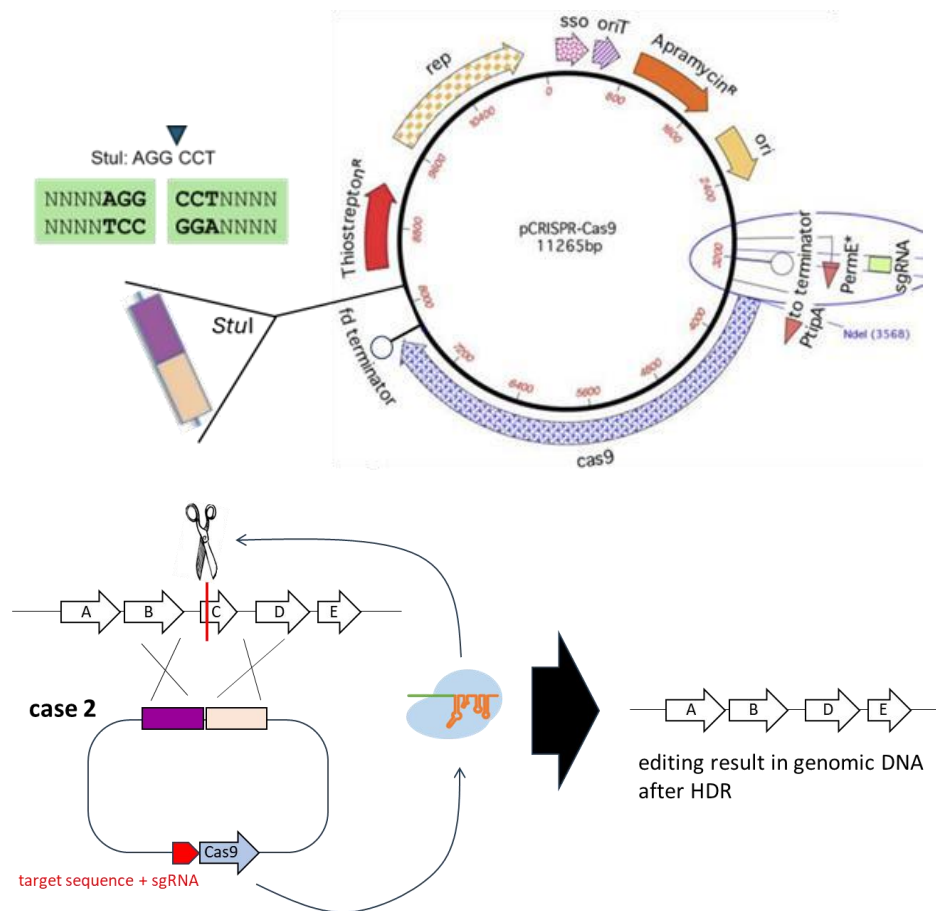


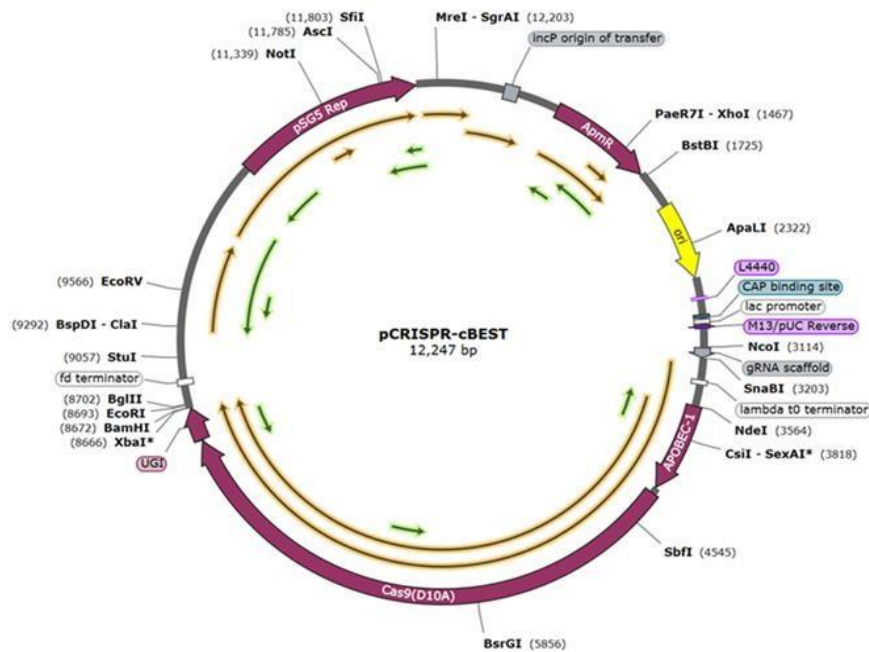
Figure 10. (A) pCRISPR-Cas9 plasmid (Tong *et al.*, 2015). (B) A diagram demonstrating Homology-directed repair (HDR) which requires an insertion of 1 kb long downstream and upstream regions of the targeted deletion into the plasmid at the *StuI* restriction site.

M13/pUC reverse, *lac* promoter and a *lac* operator. The primer sequences facilitate the binding of endonucleases, which are fundamental for the functions of the plasmid. The *lac* promoter, operator and associated repressor facilitate expression and control of the *lac* operon, which is associated with the catabolic activator protein (CAP) site. The *lac* genes in *E. coli* can act as regulatory sites as they can be induced. The presence of lactose or isopropyl-beta-D-thiogalactopyranoside (IPTG) controls the actions of the *lac* promoter. In the absence of these inducer molecules, the *lac* repressor protein binds to DNA, thus prevents the binding of RNA polymerase. However, when expressed, the *lac* promoter will enable the initiation of transcription (Singh, 2014). The primer and promoter sequences are located prior to the *Nco*I binding site, which encourages binding of the Cas9 at the correct site, adjacent to the sgRNA scaffold sequence. Successful transcription is also encouraged by the presence of the CAP binding site which is in close proximity to the sgRNA scaffold. The binding of CAP, when in the presence of cyclic AMP (cAMP), at this key scaffold site triggers transcription (Lawson *et al.*, 2004). The combination of the various primers, scaffold and the CAP binding site provide the optimal conditions for successful binding. Formation of the sgRNA editing scaffold at this desired *Nco*I site endows subsequent successful gene editing (Renneberg, 2023).

The introduction of a target sequence linked to an sgRNA cassette into the pCRISPR-Cas9 plasmid is a very laborious and time intensive process as it requires a series of multiple steps including PCR, restriction, ligation and sequencing. Another issue of using this plasmid as a gene editing system in *Streptomyces* is that it creates a library of deletions with variable sizes around the targeted sequence because the site-specific DNA double-strand breaks (DSBs) caused by Cas9 are repaired through the error-prone nonhomologous end joining (NHEJ) pathway. This issue can be circumvented via a homology directed repair (HDR) system which requires inserting ~ 1 kb long size of both downstream and upstream regions of the targeted deletion into the plasmid using the *Stu*I restriction site (Fig. 10B). In addition, it was also reported that the overexpression of the Cas9 protein could cause a toxic effect in *E. coli* which in turn increases the incidence of off-target effects. To correct these flaws of a pCRISPR-Cas9 based gene editing system, Tong and co-workers (2019) developed a DSB-free, single base pair editing system in *Streptomyces* which is designated as the pCRISPR-base editing system (pCRISPR-BEST) plasmid. The pCRISPR-BEST base gene editing system does not introduce a DSB for gene modifications; instead a single nucleotide modification in a single-strand break is introduced by using deaminase.

Two versions of the pCRISPR-BEST plasmids were developed: a cytidine deaminase-based base editor (pCRISPR-cBEST) enabling the conversion of C:G to T:A, and an adenosine deaminase-based base editor (pCRISPR-aBEST) enabling the conversion of A:T to G:C. The most significant change in the pCRISPR-BEST system is the modification of the Cas9 protein.

A



B

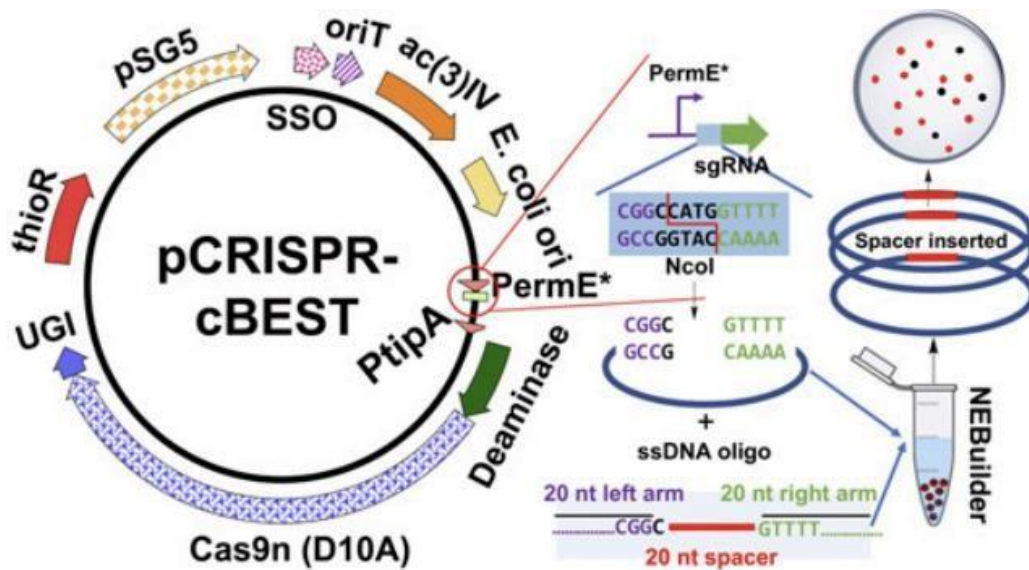


Figure 11. (A) The map of pCRISPR-cBEST plasmid. The restriction sites (black outer labels), key sequences and proteins (illustrated with the red/yellow arrow(s)) are also illustrated. Plasmid viewed using SnapGene software (Addgene plasmid #271202 (Tong et al., 2019)). (B) The map of CRISPR-cBEST plasmid from Tong et al. (2019). The core component is the fusion protein of *S. coelicolor* A3(2) codon-optimised rAPOBEC1, Cas9n (D10A), and UGI. The key component of CRISPR-cBEST is a fusion protein of *S. coelicolor* A3(2) codon optimised ecTadA, and Cas9n (D10A). Expression of the fusion proteins is controlled by a thiostrepton-inducible *tipA* promoter. Using this promoter, leakage expression is enough to carry out the base editing functionality, but a stronger expression can be induced by adding thiostrepton. The sgRNA cassette is under control of *ermE** promoter. A PCR-free, 1-step ssDNA bridging approach can be applied for the 20-bp spacer cloning.

Switching the 10th aspartic acid (D) to alanine (A) of the modified Cas9 nickase (Cas9n(D10A)) in these plasmids results in the loss of its original endonuclease activity. This prevents DSBs but Cas9n (D10A) can still locate the target site by sgRNA and open up the double strand to form an R loop. Nevertheless, many other fundamental elements between pCRISPR-Cas9 and pCRISPR-BEST plasmids are conserved: pSG5 replicon, *oriT* (or also designated as an *incP*), *apmR*, L4440, *lac* promoter, pBR322ori-F, *lac* operator, CAP binding site, M13/pUC reverse, sgRNA scaffold and a *tsr* gene as pCRISPR-BEST was designed and created based on the pCRISPR-Cas9 plasmid (Fig. 11A).

The multiplexed editing capabilities of the pCRISPR-BEST plasmids are enabled by the Cys4based sgRNA processing complexes, which can be modified to target desired sequences. The BE3 system utilised in pCRISPR-cBEST was artificially constructed by fusing the APOBEC1 cytidine deaminase found within rats (rAPOBEC1) with the Cas9n and an uracil Glycosylase inhibitor (UGI). The expression of these fusion proteins is controlled by a thiostrepton inducible *tipA* promoter, hence the need for a thiostrepton resistance gene, *tsr*, in the editing plasmid sequence. The modified sgRNA cassette facilitates the binding to target DNA and enables separation of the double-stranded DNA, then the deamination mechanism facilitates editing of the nucleotide sequence. Due to the non-DSB nature of the system, an R-loop forms when the cassette binds to a target sequence, thus creating a small editing window. The pCRISPR-cBEST deaminase converts target cytidine (C) in a C:G base pair to result in a uracil (U) in a U:G mismatch. The pCRISPR-aBEST deaminase converts target adenine (A) in an A:T base pair and results in an inosine (I) in a I:T mismatch. In DNA replication, U is recognized as T, and I (inosine) is recognized as G. When the base pair mismatch in the DNA is found, the cellular mismatch repair (MMR) becomes the dominant DNA repair pathway and, as a consequence, nucleotide excision repair (NER) will occur. The converted U by pCRISPR-cBEST deaminase is an illegitimate DNA base in DNA replication, therefore it will be recognised and excised by uracil-DNA glycosylase (UDGs), and the conserved NER in the cells leading to the reversion to the original base pair will be initiated. But this process can be inhibited by a uracil-DNA glycosylase inhibitor (UGI). A codon-optimised UGI for *Streptomyces* expression was fused to the C-terminus of Cas9n in the pCRISPR-cBEST plasmid, therefore this will ultimately lead to a targeted G being replaced by an A in the next replication cycle, and for its complementary base to become a T. Thus, pCRISPR-cBEST can efficiently facilitate the conversion of C bases into T. However, no similar inhibitor was reported for NER for the I:T mismatch and this could be one of the reasons pCRISPR-cBEST has a higher editing performance than pCRISPR-aBEST, despite pCRISPR-cBEST causes more off-target effects (Tong *et al.*, 2019).

Overall, the fidelity of the pCRISPR-BEST system was high in *Streptomyces* species, as derived from the genome-wide off-target evaluation conducted by the Tong and co-workers (2019). The efficiency of the MMR repair can be increased by introducing a single strand DNA nick in proximity to the editing site and the illegitimate DNA bases will be repaired during DNA synthesis. This process generates permanent modifications of the target DNA without the requirement of a DSB. Therefore, by careful selection and designing of the target sites, base editors can generate point mutations resulting in amino acid replacements or the introduction of multiple stop codons. In addition to this DSB-free, single-nucleotide-resolution genome editing system, a site to introduce target sequences in these plasmids has been modified to be a labour-saving version of a PCR-free, 1-step ssDNA bridging approach which can be applied for the 20-bp spacer cloning (Fig. 11B). The original sgRNA cassette from pCRISPR-Cas9 was modified to enable compatibility with the single-strand DNA (ssDNA) oligo bridging method in the pCRISPR-BEST plasmids. The modified pCRISPR-BEST editing cassette facilitates the 20 nucleotide-space cloning step and also increases the cloning efficiency due to its highly specific targeting nature. The transcription of the sgRNA is controlled by a strong, constitutive *ermE** promoter and terminated by its associated terminator. These significant modifications of the original pCRISPR-Cas9 editing plasmid enable the CRISPR-BEST plasmids to edit a target nucleotide sequence due to their highly specific editing mechanism. The sgRNA editing cassette can be programmed at the *NcoI* site just after the *ermE** promoter. A 20nt spacer sequence can be inserted that will allow the editing cassette to target a specific gene sequence. The software application 'CRISPy-web' (Available at: [CRISPy-web](#)) was utilised to simplify the protospacer identification process. The application can design sgRNAs for a given desired sequence: it provides a recommended target sequence, the associated PAM and the expected number of off-target mismatches. The search parameters can then be refined by enabling 'CRISPR-BEST mode' which filters cBEST only sequences, aBEST only sequences and sequences that bring about a STOP codon.

1.6. Aims and objectives

The most advanced and appropriate genome editing system for applications in *Streptomyces* with high G+C is the aforementioned plasmids based on pCRISPR-BEST. These gene editing plasmids have been developed for the single gene editing purpose and to be a more efficient cloning system that is less labour intensive than that in pCRISPR-Cas9. However, this advanced system is not suitable for the ultimate goal of this project, which intends to mutagenize an operon consisting of 5 polycistronic genes (SCO6357-53) identified from the previous transcriptomic profiling study conducted by Connelly (2021). The pCRISPR-BEST

system is not suitable as it cannot target multiple genes of interest at once. Therefore, the primary goal of this study is to create a simple but more efficient HDR-oriented gene editing system by utilising two existing CRISPR-Cas9 based gene editing plasmids, pCRISPR-Cas9 and pCRISPR-cBEST. Consequently, the resulting system will incorporate advantages from each editing mechanism. This new editing plasmid will be much more efficient, and labour saving compared to the established genome editing models in *Streptomyces*, due to the incorporated HDR system, which is superior to the NHEJ system used in existing pCRISPR-Cas9. The significance of the establishment of this new editing plasmid is that it will facilitate easier characterisation of genome-wide dynamics within *S. coelicolor*, specifically its response to glycopeptide antibiotics, particularly vancomycin and teicoplanin, through analysis of glycopeptide resistance genes. Ultimately, this study aims to characterise the functions of genes relating to teicoplanin tolerance/resistance displayed by *S. coelicolor*. Gaining insights into the microbial genome and the antimicrobial mechanisms within *S. coelicolor*, a non-pathogenic actinomycete, may also provide insights into how antimicrobial mechanisms arise in pathogenic bacteria found within the same phylum and family. The non-pathogenic nature of *S. coelicolor* makes it an ideal model organism to conduct genetic manipulations on to investigate its antimicrobial mechanisms. Insights gained from analysing *S. coelicolor*'s genome could then allow for pathogenic antimicrobial mechanisms to be understood. Through understanding how resistance arises, this will ultimately help to alleviate the rising AMR crisis as antimicrobial infections can be effectively treated. Additionally, new antimicrobial agents could then be developed to target the underlying resistance mechanisms. By effectively treating bacterial infections or developing novel treatments, the occurrence of resistant pathogens would decrease, as the antimicrobial agents would effectively eliminate the pathogenic strains, which would ultimately help alleviate the rising AMR crisis.

2. Materials and Methods

2.1. Bacterial strains, plasmids and oligos for PCR

Bacterial strains, plasmid DNAs and oligos used in this study are summarised in Table 1-3.

Table 1. Bacterial strains used in this study.

Strain	Description and Genotype	Reference
<i>Streptomyces coelicolor</i> A3(2) M600	Prototrophic, SCP1 ⁻ , SCP2 ⁻ , plasmid free derivative of the original John Innes stock strain 1147	Chakraborty and Bibb, 1997
<i>Escherichia coli</i> ET12567/pUZ8002	F ⁻ <i>dam13::Tn9</i> (Cam ^R) <i>dcm-6 hsdM hsdR recF143 zjj202::Tn10</i> (Tet ^R) <i>rpsL136</i> . Methylation deficient <i>E. coli</i> strain used in intergeneric conjugation into <i>Streptomyces</i>	Flett <i>et al.</i> , 1997
<i>Escherichia coli</i> DH5α	F ⁺ <i>endA1 hsdR17</i> (r _k ⁻ , m _k ⁺) <i>glnV44 thi-1 recA1 gyrA</i> (Nal ^r) <i>relA1 Δ(lacIZYA-argF)</i> U169 <i>deoR</i> (Φ80 <i>dlacΔ</i> (<i>lacZ</i>)M15)	Woodcock <i>et al.</i> , 1989

Table 2. Plasmids used in this study.

Plasmid	Description/ genotype	Reference
pCRISPR-Cas9	<i>Streptomyces</i> codon optimised spCas9, sgRNA cassette; single pSG5 replicon-based plasmid backbone	(Tong <i>et al.</i> , 2015) Addgene Catalog #: 125686 (discontinued)
pCRISPR-cBEST	<i>Streptomyces</i> codon optimised spCas9n (D10A), modified sgRNA cassette, <i>streptomyces</i> codon optimised rAPOBEC1; single pSG5 replicon-based plasmid backbone	(Tong <i>et al.</i> , 2019) Addgene Catalog #: 125689
pCRISPR-Cas9-KM1	0.45kb NdeI-NcoI fragment from pCRISPR-cBEST ligated in 10.8 kb NdeI-NcoI fragment of pCRISPR-Cas9	This Study
pCRISPR-Cas9-KM2	6.7Kb StuI-NdeI fragment from pCRISPR-cBEST ligated with a 4.5Kb NdeI-StuI fragment from pCRISPR-Cas9	This Study

Table 3. Oligonucleotides used in this study. All oligonucleotide stock solutions were made in 100 mM in TE buffer. Oligonucleotide stock was aliquoted at time of use in 10 mM solutions.

Name of oligos	Sequence (5'-3')	Description
pCRISPR-Cas9-KM_FWD_2942	ATTGTGAGCGGATAA CAATTCACACAGGA	sgRNA targeting 2942-2971 (upstream of <i>NcoI</i> site) of cBEST plasmid to select for correct insert at <i>NcoI-NdeI</i> (single 'G' deletion)
pCRISPR-Cas9KM_REV_3513	TCCACGCTGCCTCCT CACGTGACGTGAGGT	sgRNA targeting 3513-3542 (upstream of <i>NdeI</i> site) of cBEST plasmid to select for correct insert at <i>NcoI-NdeI</i> (single 'G' deletion)
pCRISPR-Cas9_FWD_Stu	AGCGGAGTGAGAATA GAAAGGAACAATAA	sgRNA targeting 7800-7829 (upstream of <i>StuI</i> site) of Cas9 plasmid to ensure the presence of Cas9 backbone
pCRISPR-Cas9_REV_Stu	TGACCGGCAATCACC AACGCAAGGGGAAGT	sgRNA targeting 8143-8172 (downstream of <i>StuI</i> site) of Cas9 plasmid to ensure presence of Cas9 backbone
pCRISPRdCas9_test_F	AACACCCAGCTGCAG AACGAGAAGCTGTAC	sgRNA targeting 3351-3380 (within dCas9) of dCas9 plasmid to ensure that Cas9 plasmid is genuine; forward primer
pCRISPRdCas9_test_R	TTGTGCGAACTTGCGC TGGGTGATGAGCTTG	sgRNA targeting 3611-3640 (within dCas9) of dCas9 plasmid to ensure that Cas9 plasmid is genuine; reverse primer
VanK target	GTTTGGTAGGATCGA CGGCGACACGTCGC TTCATATAGGTTTTAG AGCTAGAAATAGC	60 nt sgRNA; 40 nt of complementary pCRISPR-Cas9-KM DNA, encasing a 20 nt region of <i>S. coelicolor</i> 's DNA that targets SCO3593
Van-Up_F	TCGTGCGAAGGCACTA GAAGGTCACCTGCCG GTGTGCGGAG	40 nt sgRNA targeting two locations: first 20 nt targets 8054-8073 (<i>StuI</i> site) in pCRISPR-Cas9-KM; the second 20 nt targets <i>S. coelicolor</i> SCO3589 (<i>vanS</i>) upstream region
Van-Up_R	ATCCTGGGGAACGAC CCCGAGATTCTGAC ACTCGCGGGC	40 nt sgRNA targeting two locations: first 20 nt targets <i>S. coelicolor</i> SCO3596 (<i>vanX</i>) upstream region; the second 20 nt targets <i>S. coelicolor</i> SCO3589 (<i>vanS</i>) downstream region
Van-Down_F	ATCCTGGGGAACGAC CCCGA	20 nt sgRNA targeting <i>S. coelicolor</i> SCO3596 (<i>vanX</i>) upstream region
Van-Down_R	GGTCGATCCCCGCAT ATAGGCTACGCGATG GGAAAGTCGA	40 nt sgRNA targeting two locations: first 20 nt targets 8073-8093 (<i>StuI</i> site) in pCRISPR-Cas9-KM; the second 20 nt targets <i>S. coelicolor</i> SCO3596 (<i>vanX</i>) downstream region

2.2. Growth media

2.2.1. Soya flour mannitol agar (SFM)

SFM was used for the preparation of *Streptomyces* spores; SFM required autoclaving twice and the use of tap water. SFM was made up of the following:

Mannitol	20 g
Soya Flour	20 g
Agar	20 g
Tap water to	1 L

20 g mannitol was dissolved in 1 L of tap water and 200 mL poured into a 500 mL Duran bottle, each containing 2 g soya flour and 2 g agar. Bottles were autoclave at 115°C for 15 minutes twice; all prepared solutions were autoclaved at these parameters, unless otherwise stated.

2.2.2. Luria-Bertani broth (LB)

LB was used for *E. coli* cultures and made up of the following:

Difco Bacto Tryptone	10 g
Difco Bacto Yeast Extract	5 g
NaCl	5 g
Glucose	1 g
Deionised (DI) water to	1 L

The solution was dispensed in 100 mL aliquots and autoclaved.

2.2.3. Luria-Bertani agar (LA)

LA was used for *E. coli* cultures and made up from 1 L of LB as above. 100 mL LB was poured into a 250 mL Duran bottle, each containing 1 g agar and autoclaved.

2.2.4. 2x Yeast extract tryptone medium (2x YT)

2x YT medium was used for *Streptomyces* and *E. coli* conjugation, and made up of the following:

Difco Bacto Tryptone	16 g
Difco Bacto Yeast Extract	10 g
NaCl	5 g
DI water to	1 L

The solution was dispensed in 100 mL aliquots and autoclaved.

2.3. Buffers and solutions

2.3.1. Transformation and storage solution (TSS)

TSS was used for producing chemically competent *E. coli* cells and made up of the following:

LB	85% (v/v)
Polyethylene Glycol (PEG) 6000	10% (w/v)
Dimethyl Sulfoxide (DMSO)	5% (v/v)
MgCl ₂	50 mM
pH	6.5

LB, 1 M stock solution MgCl₂ and DMSO were prepared; to make 100 mL TSS, 10 g of PEG6000 was dissolved in ~80 mL LB. After adding 5 ml of DMSO and 5 mL of 1 M MgCl₂, the total volume was filled to 100 mL with LB. The pH was adjusted to 6.5 prior to autoclaving; the pH was measured using a digital pH meter (Electronic Temperature Instruments LTD).

2.3.2. Super optimal media with catabolic repressor for transformation (SOC)

SOC was used for *E. coli* transformation and made up of the following:

Yeast Extract	500 mg
Difco Bacto Tryptone	2 g
NaCl	58.4 mg
KCl	8.6 mg
MgCl ₂	460 mg
MgSO ₄	240 mg
pH	7.5
DI water to	100 mL

The pH was adjusted to 7.5 prior to autoclaving. At the time of use, 2 mL of sterile 20% (w/v) glucose solution was added into the 100 ml solution.

2.3.3. Solutions for small-scale plasmid extraction from *E. coli*

2.3.3.1. Sodium chloride Tris-EDTA (STE)

Stock solutions of 5 M NaCl, 1 M Tris-HCl (pH8) and 0.25 M Ethylenediaminetetraacetic acid (EDTA) (pH8) were prepared and autoclaved. For 100 mL of STE, the following were mixed:

5 M NaCl	5 mL
1 M Tris-HCl (pH8)	1 mL
0.25 M EDTA (pH8)	0.4 mL
Autoclaved DI water to	100 mL

2.3.3.2. Solution 1

Stock solutions of 20% (w/v) Glucose, 1 M Tris-HCl (pH8) and 0.25 M EDTA (pH8) were autoclaved. For 100 mL of Solution 1, the following were mixed:

20% (w/v) Glucose	4.5 mL
1 M Tris-HCl (pH8)	2.5 mL
0.25 M EDTA (pH8)	4 mL
Autoclaved DI water to	100 mL

2.3.3.3. Solution 2

Stock solutions of 10 N NaOH and 10% (w/v) Sodium Dodecyl-Sulphate (SDS) were prepared and autoclaved. 10 N NaOH solution was carefully prepared on ice, as dissolving the pellets in water is highly exothermic. The solutions were prepared and mixed freshly for each use. For 1 mL of Solution 2, the following were mixed:

10 N NaOH	0.02 mL
10% (w/v) SDS	0.1 mL
Autoclaved DI water to	1 mL

2.3.3.4. Solution 3

A stock solution of 5 M potassium acetate (CH_3COOK) was prepared and autoclaved. For 100 mL of Solution 3, the following were mixed:

5 M CH ₃ COOK	4.5 mL
Glacial Acetic Acid	11.5 mL
Autoclaved DI water to	100 mL

2.3.3.5. TE buffer

Stock solutions of 1 M Tris-HCl (pH8) and 0.25 M EDTA (pH8) were prepared and autoclaved. For 100 mL of TE buffer, the following were mixed:

1 M Tris-HCl (pH8)	1 mL
0.25 M EDTA (pH8)	0.25 mL
Autoclaved DI water to	100 mL

2.3.3.6. Other solutions

Solutions of Phenol:Chloroform (1:1, v/v), 10M CH₃COONH₄, 100% (v/v) and 70% (v/v) Ethanol were used alternatively to the Qiagen QIAprep Spin Miniprep kit (50) (Qiagen, Catalogue # QIA27104) when extracting plasmid DNA from *E. coli* cultures. The Quiagen kit was preferentially used due to the simplified process the kit provided for the extraction of DNA.

2.4. Bacterial culture and storages

2.4.1. *E. coli* strains

All *E. coli* strains used in this study were grown on liquid (LB) or solid (LA) medium supplemented with appropriate antibiotics (see 2.6.2.) at 37°C. Glycerol stocks were made from fresh overnight liquid cultures by adding 40% (v/v) glycerol to an equal volume of culture and storing at -80°C.

2.4.2. Preparation of chemically competent *E. coli* cells

The glycerol stock of *E. coli* cells was streaked onto LA supplemented with antibiotic(s) (Apramycin at 25 mg/mL, Kanamycin at 50 mg/mL and Chloramphenicol at 25 mg/mL; see section 2.6.2) if necessary and incubated overnight to obtain a single colony. A single colony was picked and placed in a universal tube containing 5 mL LB (supplemented with the same antibiotic(s) used for LA) and incubated with shaking at 220 rpm at 37°C overnight (~16-20 hours). 200 µL of this seed culture was then placed in a flask containing 50 mL LB supplemented with the same antibiotic(s) used for the seed culture and cultured in a shaking incubator at 220 rpm at 37°C for 2 hours until the optical density reached 0.3-0.4 at 600 nm (OD₆₀₀) (Jenway 7305 Spectrophotometer – Fisher Scientific). The culture flask was then

cooled on ice for 10 min and the culture broth was moved to a 50 mL conical tube. The broth was centrifuged at 3000 rpm at 4°C for 20 minutes. The supernatant was discarded, the cell pellet was resuspended in the remaining broth (~0.5-1 mL) then the tube was placed on ice. 4.75 mL of ice-cold TSS and 250 µL of ice-cold DMSO were added to the resuspended pellet and mixed gently. The resulting solution was aliquoted in 50 µL into Eppendorf tubes and stored at -80°C. Commercial competent DH5α *E. coli* cells were also purchased from Thermo Scientific (Catalogue #18265017) as an alternative.

2.4.3. *S. coelicolor* cells

S. coelicolor cells were grown on SFM for the preparation of spore stock suspensions. A single colony of *S. coelicolor* on SFM was resuspended in 20 µl autoclaved deionised water and was spread on an SFM plate to yield a confluent lawn. The plate was incubated at 30°C for ~7 days (but no longer than 10 days) until a confluent lawn of grey spores was visible. The surface of the spore lawns was then wiped across by a sterile cotton pad wetted by DI water. After, 1 mL of autoclaved 20% (v/v) glycerol was spread onto the surface of the spore lawn; a sterile cotton pad was used to wipe across the plate surface. The cotton absorbed the glycerol and the bacterial spores. The soaked cotton pads were then placed inside a sterile syringe and pressed into a sterile Eppendorf tube, thus collecting spores suspended in 20% (v/v) glycerol. Established tubes were labelled and stored at -20°C (short-term use) or at -80°C (long-term storage).

2.5. Microbial DNA cloning

2.5.1. Small-scale plasmid DNA extraction from *E. coli*

E. coli cells carrying desired plasmid DNA were inoculated in up to 5 mL LB with appropriate antibiotics and incubated with shaking at 220 rpm at 37°C overnight. Cell pellets were harvested by centrifugation (20 minutes at 3000 rpm at 4°C) in 1.5 mL Eppendorf tubes, then resuspended in 500 µL of STE solution containing a diluted RNase. Cell pellets were harvested again by a quick centrifugation and completely resuspended in 100 µL of ice-cold Solution 1. 200 µL of freshly prepared Solution 2 was added to the suspension, then the solution was gently mixed by inverting the tube (NOT vortexing) several times until the sample became viscous and clear. 150 µL of ice-cold Solution 3 was added and the sample was mixed gently again by inverting and tapping the tube. The sample was left on ice for 5-10 minutes, after which, it produced a white precipitate. The sample was then centrifuged further for another 5 minutes at 4°C and the supernatant was transferred to a fresh 1.5 mL Eppendorf tube. The sample was extracted with equal volumes of phenol:chloroform (1:1, v/v) by vigorous vortexing

for 1 minute and centrifugation for 5 minutes at 13,000 rpm at 4°C. After centrifugation, the aqueous upper layer of the sample was carefully transferred to a fresh Eppendorf tube. For DNA precipitation, 2 volumes of 100% isopropanol and 1/4 volume of 10 M CH₃COONH₄ were added and the sample was mixed by brief vortexing and left at room temperature for 5 minutes (the sample can be preserved at -20°C at this stage until the next step; refrigeration of the sample for a few hours or overnight is recommended to improve the final yield of DNAs). The plasmid DNAs were harvested by centrifugation (20 minutes at 13,000 rpm at 4°C), washed in 500 µL of 70% (v/v) ethanol, dried, and resuspended in ~50 µL Tris-EDTA (TE). The concentration of plasmid DNA was quantified using a Nanodrop 2000c Spectrophotometer (Thermo Scientific) and stored at -20°C. The QIAprep Spin Miniprep Kit (50) (Qiagen, Catalogue # QIA27104) was also utilised alternatively to this protocol to prepare plasmid DNA.

2.5.2. Restriction endonuclease digestion

In order to confirm the presence of the desired plasmids, various restriction enzymes were selected and utilised (*Bam*HI-HF, *Eco*RV-HF, *Nde*I, *Nco*I-HF, *Stu*I, *Xba*I and *Xho*I; all from New England Biolabs (NEB). The restriction maps of the plasmid DNA were obtained by using SnapGene, a viewer software. When establishing restriction reactions with more than one restriction enzyme per plasmid DNA, it was ensured to select reaction conditions including a reaction buffer that was optimised for the restriction activity of both enzymes. To ensure complete digestion, usually less than 1 µg/µL of DNA concentration was used in a final 20-40 µL volume of sample and reactions were often left overnight at the desired reaction temperature.

2.5.3. Gel electrophoresis

Usually, a 1% (w/v) agarose was used, but concentrations down to 0.5% (w/v) and up to 3% (w/v) were also used for analysis of large or small fragments, respectively. Invitrogen SYBR safe (Thermo Fisher Scientific) was added to liquid agarose before creating the gel mould to stain DNA and allow for subsequent DNA visualisation. Agarose gels were prepared and run submerged with 1x TAE buffer. The voltage for the electrophoresis was set depending on the expected size of DNA bands; relatively low voltage (60 V) was set for running larger bands up to 10 kb and high voltages (80-100 V) were set for smaller bands (~0.5 kb). The appropriate ladder was run alongside the samples (e.g. 1 kb or 1 kb plus or 100 bp DNA ladder; NEB) depending on the expected fragment sizes and the need for band size clarification. DNA separation on the gel was viewed and the images were documented using a UV imager (imaging using UV was only conducted when the DNA contained within bands was not required for excision from the gel and subsequent DNA utilisation).

2.5.4. DNA extraction and purification from Agarose gel

Desired DNA fragments from the gel were visualised and cut out using a blue light box or a long-wavelength UV light (~310 nm) to minimise damage to the DNA molecules. The DNA fragments were then extracted using the QIAquick Gel Extraction Kit (Qiagen #28704) following the manufacturer's instructions. The concentration of the extracted DNA was quantified using the Nanodrop 2000c.

2.5.5. Ligation

The ligation protocols were established using the NEBioCalculator (Available at <https://nebiocalculator.neb.com/#!/ligation>). Vector and insert DNA were mixed at a molar ratio of 1:5 or 1:10 or 1:20, with 1/10 volume 10x ligation buffer and 10 U T4 DNA ligase in up to 40 µL total volume. It was ensured that the ligase was added last to the reaction after the other components had been mixed and centrifuged briefly. The mixture was incubated at 18°C for 4-24 hours. The ligated DNA was used to transform competent *E. coli* cells. A successful ligation with the desired insert was confirmed by customised Sanger sequencing from Eurofins.

2.5.6. Transformation of competent *E. coli* cells

For transformation, the frozen competent cells were quickly defrosted on ice. DNA (~100 ng) was added to 50-200 µL of competent cells. When transforming ligation products, 20 µL of reaction sample was used to maximise the occurrence of successful transformation. The sample in the tube was gently mixed by pipetting up and down, not by vortexing, and placed on ice for 30 minutes. The sample was then heat-shocked at 42°C for 90 seconds and placed back on ice to cool for 5 minutes. 1-2 mL of pre-warmed (37°C) SOC media was added to the sample which was then incubated in a shaker (220 rpm, 37°C) for 2 hours. The transformed cells were plated out onto LA containing the appropriate antibiotic(s) and incubated at 37°C for 16-20 hours. Upon colony growth, established colonies were re-streaked onto new LA containing the appropriate antibiotic(s). The confirmation of successful transformation was verified by colony PCR or plasmid extraction and restriction digestion.

2.5.7. PCR

All designed oligonucleotide primers in Table 3 were synthesised by Merck. For most routine PCR reactions to obtain up to 1 kb size product, the reaction mixture contained a 200 µM final concentration of each of the four dNTPs, 2-2.5 U of Taq polymerase (OneTaq DNA polymerase, NEB), 1x reaction buffer, 50 pmol of each primer, ~50 ng of template DNA and 10% (v/v) DMSO in a final volume of 50 µL. For colony PCR, 5 to 10 different single colonies were selected to

be tested. A sterile toothpick was used to pick a small amount of bacterial spores from a single colony on the agar plate and used to streak onto a replica plate. The same toothpick was then used to resuspend the colony into a pre-made PCR master mix. The reaction tubes were mixed by pipetting and centrifuged briefly before placing onto the thermocycler. After denaturation at 95°C for 5 minutes, the samples were subjected to 25-30 cycles of denaturation (95°C, 1 minute), annealing (55-62°C, 1 minute), and extension (72°C, 1 minutes) and then incubated for a further 5 minutes at 72°C. The annealing temperature varied according to the primers utilised. Additionally, the annealing and extension time varied: 45 seconds was used for 200-500 bp long PCR products, and 1 minute was used for 500-1,000 bp long PCR products. PCR products were analysed by agarose gel electrophoresis. For all PCR reactions, Mili-Q water was used.

2.6. Other microbiological methods and materials

2.6.1. Conjugation between *E. coli* and *S. coelicolor*

E. coli DH5 α strain was used as a primary DNA cloning host but the resulting plasmid DNA was transferred to a methylation-deficient strain *E. coli* ET12567 carrying pUZ8002 plasmid (ET12567/pUZ8002) for conjugation with *Streptomyces* so that the resulting conjugation vector was transferred to the final desired host strain of *Streptomyces*. In this study, pCRISPRCas9-KM was the conjugation vector to be transferred to *S. coelicolor* for its genome editing.

The initial ET12567/pUZ8002 before transformation with pCRISPR-Cas9-KM was grown in LB containing two antibiotics, chloramphenicol (50 μ g/mL) and kanamycin (100 μ g/mL) but the donor cultures of ET12567/pUZ8002 carrying pCRISPR-Cas9-KM conjugation vector (ET12567/pUZ8002/pCRISPR-Cas9-KM) were grown at 37°C to an OD₆₀₀ of 0.4-0.6 in LB supplemented with chloramphenicol (50 μ g/mL), kanamycin (100 μ g/mL) and apramycin (50 μ g/mL). The donor cultures were then washed twice using LB and mixed with $\sim 10^8$ *S. coelicolor* spores that had been activated for rapid germination (heated to 50°C for 10 minutes in 2x YT solution then cooled). The mixed sample of 500 μ L donors ($\sim 2 \times 10^8$ ET12567/pUZ8002/pCRISPR-Cas9-KM) and 500 μ L recipients ($\sim 10^8$ *S. coelicolor* spores) was spread on SFM containing 10 mM MgCl₂ then incubated at 30°C overnight. After overnight incubation, the plates were flooded with 1 mL sterile DI water containing 0.5 mg nalidixic acid and 1 mg apramycin then incubated at 30°C. Exconjugants were picked from the plates 5-7 days later and streaked for single colonies on SFM plates containing 25 μ g/mL nalidixic acid and 50 μ g/mL apramycin. Nalidixic acid is an antibiotic that inhibits the growth of *E. coli* but does not harm *Streptomyces* strains.

To test if the desired gene editing in *S. coelicolor* via pCRISPR-Cas9-KM was successfully conducted, selected apramycin resistant exconjugants were streaked on both SFM containing 50 µg/mL apramycin and SFM containing 10 µg/mL vancomycin. If the HDR by the pCRISPRCas9-KM in *S. coelicolor* worked, the whole 7 gene vancomycin resistance cluster in the genome must have been deleted. The resulting *S. coelicolor* strain becomes sensitive to vancomycin therefore the exconjugant will only be viable on apramycin SFM but not on vancomycin SFM.

2.6.2. Miscellaneous reagents, enzymes and kits used for this study

Various stock reagents were prepared to carry out the protocols in the present study; stocks of the following were produced: 1 M MgSO₄ , 1 M CaCl₂ 1 M MgCl₂ , 10 mM dNTP , 10% (v/v) and 20% (v/v) glycerol.

Additionally, for the handling of bacteria, and for the selection of desired bacterial growth, various antibiotics were utilised: 25 mg/mL Apramycin , 50 mg/mL Kanamycin , 25 mg/mL Chloramphenicol , 100 mg/mL Carbenicillin , 50 mg/mL Thiostrepton , 10 mg/mL Vancomycin. All antibiotics were purchased from Sigma-Aldrich.

A range of restriction enzymes were utilised to allow for the visualisation of plasmid DNA, and to determine whether the expected DNA fragments were produced. All restriction enzymes utilised were purchased from New England Biolabs: *Bam*HI-HF (New England Biolabs (NEB), Catalogue # R3136S) , *Eco*RV-HF (New England Biolabs (NEB), Catalogue # R3195S) , *Nde*I (New England Biolabs (NEB), Catalogue #R0111S) , *Nco*I-HF (New England Biolabs (NEB), Catalogue #R3193S) , *Stu*I (New England Biolabs (NEB), Catalogue #R0187S) , *Xba*I (New England Biolabs (NEB), Catalogue #R0145S) , *Xho*I (New England Biolabs (NEB), Catalogue #R0146S).

Lastly, various kits, protocol reagents and cells were obtained for the completion of this study: QIAquick Gel Extraction Kit (50) (Qiagen, Catalogue #28704) , QIAprep Spin Miniprep Kit (50) (Qiagen, Catalogue # QIA27104) , NEBuilder® HiFi DNA Assembly Master Mix (New England Biolabs (NEB), Catalogue #E2621S) , T4 DNA ligase (New England Biolabs (NEB), Catalogue #M0202S) , OneTaq DNA polymerase (New England Biolabs (NEB), Catalogue #M0480S) , Ultrapure Phenol:Chloroform:Isoamyl Alcohol (25:24:1, v/v) (Thermo Scientific, Catalogue #15593031) , Subcloning efficiency DH5α chemically competent cells (Thermo Scientific, Catalogue #18265017) , pCRISPR-Cas9 (this was provided as a kind gift from Prof. Colin Smith and Dr. Giselda Bucca in the University of Brighton) , pCRISPR-cBEST (Addgene, Catalogue # 125689).

3. Results

3.1. Configuration of pCRISPR-Cas9-KM

The present study sought to investigate the teicoplanin resistance-associated gene clusters identified by the work of Connelly (2021). Glycopeptides such as vancomycin or teicoplanin are the only effective group of natural product antibiotics effective against resistant gram-positive pathogens, such as MRSA. However, despite being in the same antibiotic family, teicoplanin does not operate via the same well-known mode of action or resistance mechanism as in vancomycin. This is partly due to the additional long lipid moiety present in teicoplanin, that is absent in the vancomycin structure. This structural distinction in teicoplanin suggests that the mode of teicoplanin is more to do with the inhibiting of the bacterial cell membrane rather than the cell wall (Novotna *et al.*, 2012). Among the many genes and gene clusters identified in response to teicoplanin in Connelly's study (Connelly *et al.*, 2021), a 5 gene cluster at SCO635357-53 is one of the most interesting targets as it is a suspected polycistronic unit under the control of a typical two-component system predicted to be involved in cell membrane homeostasis, therefore is thought to be strongly related to a novel teicoplanin resistance/tolerance cluster in *S. coelicolor*. Genetic manipulations of this cluster within *S. coelicolor*'s genome can be brought about through the use of actinomycete-optimised CRISPR-Cas9 based editing plasmids. The most current actinomycete-optimised plasmids to introduce genetic manipulations is the pCRISPR-BEST system (Tong *et al.*, 2019). However, for the purposes of this study, this system will not suffice as the pCRISPR-BEST system cannot target entire gene clusters, instead can only implement single nucleotide mutations. On the other hand, the pCRISPR-Cas9 plasmid also does not suffice for implementing the required genetic manipulations as use of this plasmid involves a laborious cloning process in order to genetically modify hosts, and the repair mechanism by which modifications are introduced into target sequences uses the error-prone NHEJ pathway.

In this study, the primary goal is to generate a hybrid version of new CRISPR-Cas9 based gene editing plasmids pCRISPR-Cas9 and pCRISPR-cBEST plasmids. The resulting plasmid enables the HDR system which can target then edit multiple genes or gene clusters (adopted from pCRISPR-Cas9) in the *Streptomyces* host strain by a simple 1-step single-strand DNA oligo bridging method (adopted from pCRISPR-cBEST). In order to construct the desired plasmid designated as pCRISPR-Cas9-KM (Fig. 16), the optimal configuration of plasmid DNA from pCRISPR-Cas9 and pCRISPR-cBEST was determined through analysis of their plasmid DNA sequences (Fig. 12).

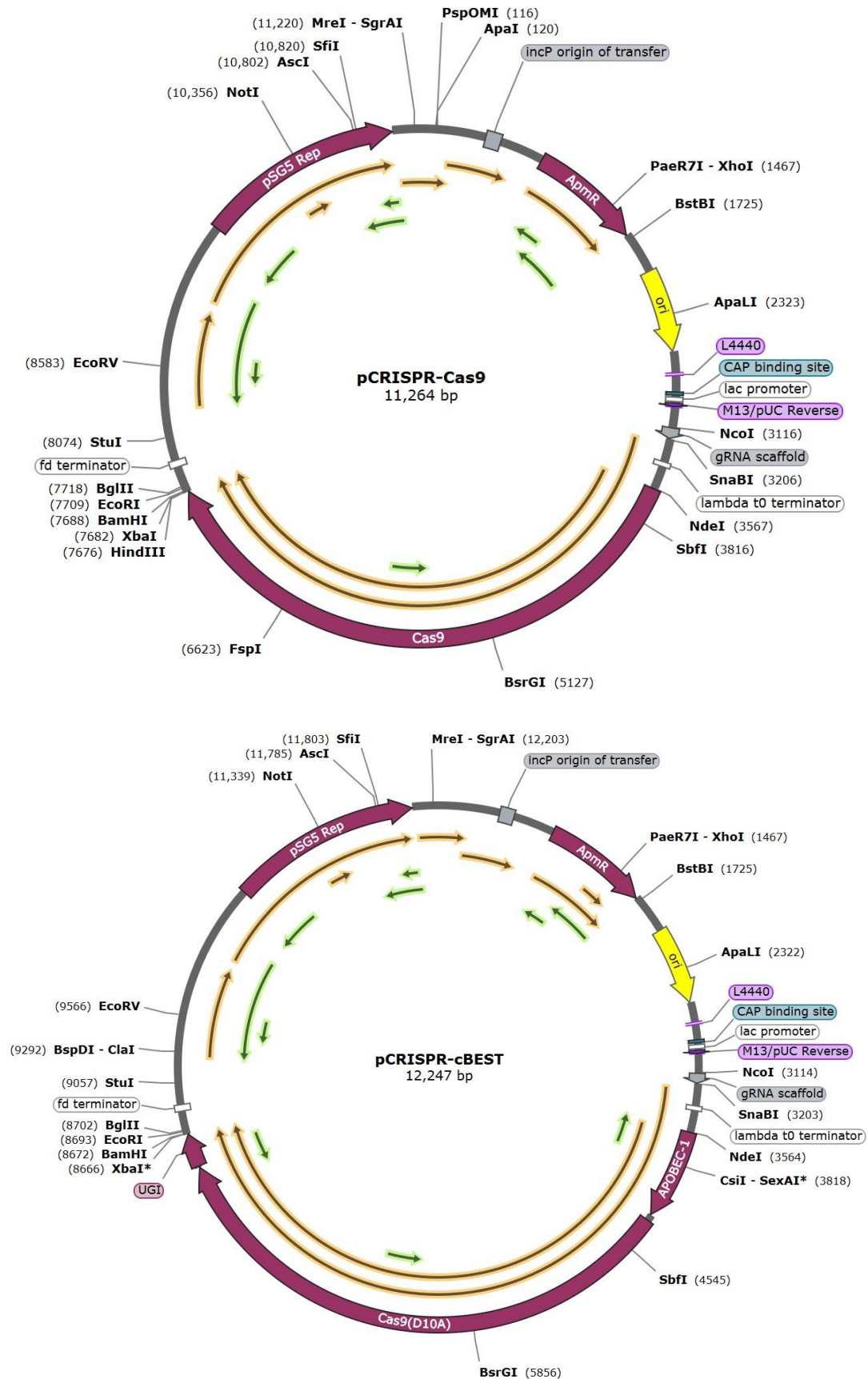


Figure 12. Plasmid maps of pCRISPR-Cas9 and pCRISPR-cBEST. Images from AddGene.

3.1.1. Sequence analysis and confirmation of pCRISPR-Cas9 and pCRISPR-cBEST plasmids by restrictions

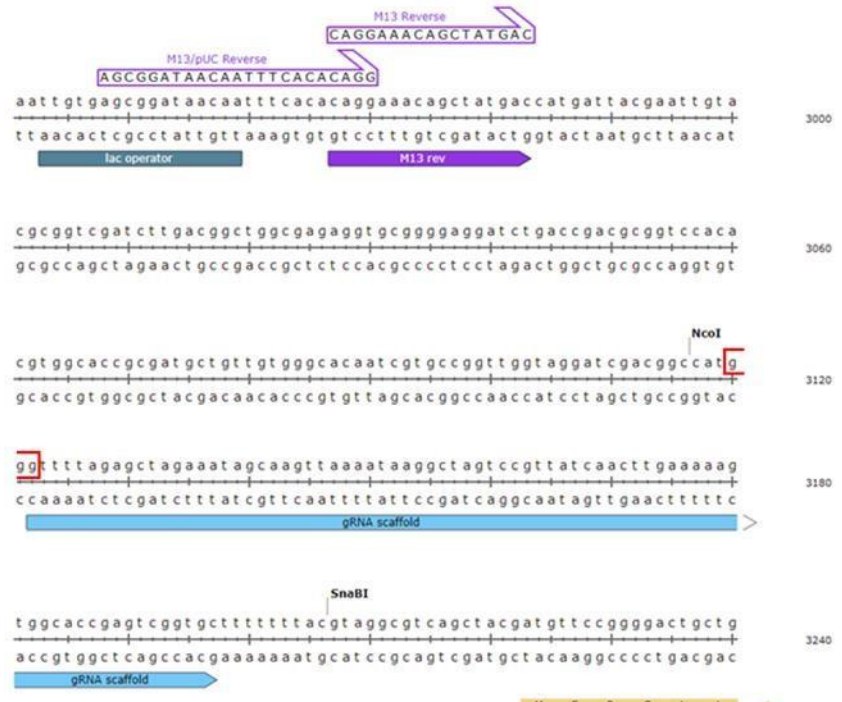
The SnapGene viewer software was utilised to identify restriction sites that would allow for the desired sgRNA cassette and the *Cas9* gene to be extracted from pCRISPR-Cas9 and pCRISPR-cBEST. Initial analysis of the plasmid DNA sequences of pCRISPR-Cas9 and pCRISPR-cBEST was conducted (Fig. 12). The modified sgRNA cassette created for the 1-step single-strand DNA bridging system in pCRISPR-cBEST was only a single nucleotide different in size compared to the same region in pCRISPR-Cas9. This modified sgRNA located between the *NcoI* and *NdeI* restriction sites in pCRISPR-cBEST had a single 'G' deletion located at the beginning of the sgRNA scaffold (Fig. 13). Ultimately, this deletion is what enables the 1-step single-strand DNA oligo bridging method, which significantly simplified the cloning efficiency of target sequences into the sgRNA cassette (Tong *et al.*, 2019). By replacing pCRISPR-Cas9's *NcoI-NdeI* region with that from pCRISPR-cBEST, this would create a desired CRISPR-Cas9 based editing plasmid that also has a modified base-editing sgRNA cassette. This 'programmable' sgRNA cassette can then target the PAM sequence within target organisms and carry out genetic manipulations when it recognises this sequence. The PAM is incorporated into a 60 bp long single-strand DNA oligo sequence which contains one set of 20 bp sequences complementary to each end region of plasmid DNA, encasing a 20 bp of PAM region.

To confirm the presence of pCRISPR-Cas9 and pCRISPR-cBEST, the plasmid DNAs were extracted from *E. coli* cells using QIAprep Spin Miniprep Kit. The resulting plasmid DNA concentrations were quantified by NanoDrop, and up to 1 µg of plasmid DNA was used for restriction digestion. The plasmids were cut using various restriction enzymes; pCRISPRCas9 was cut with *XbaI* and *NdeI*, which was expected to create two bands at 7.1 Kb and 4.1 Kb (Fig. 14 Left). pCRISPR-cBEST was cut with *BamHI* and *NdeI*; this was expected to create two bands at 7.1 Kb and 5.1 Kb (Fig. 14 Right).

3.1.2. Construction of pCRISPR-Cas9-KM

After confirmation of pCRISPR-Cas9 and pCRISPR-cBEST plasmid DNAs by restriction digestion, both plasmids were cut with *NcoI* and *NdeI* restriction enzymes. Both plasmids were expected to produce two different sized bands. For pCRISPR-Cas9, a ~450 bp size of small DNA fragment containing the sgRNA core sequence and the rest of pCRISPR-Cas9 plasmid DNA in a large 10.8 Kb size fragment was expected. Digestion of pCRISPR-cBEST with *NcoI* and *NdeI* restriction enzymes was also expected to produce a ~450 bp size of small fragment containing sgRNA core sequence but with a single 'G' deletion located at the beginning of the

pCRISPR-Cas9



pCRISPR-cBEST



Figure 13. Single nucleotide (G) deletion in pCRISPR-cBEST near *NcoI* restriction site highlighted in red box (Sequence data from AddGene).

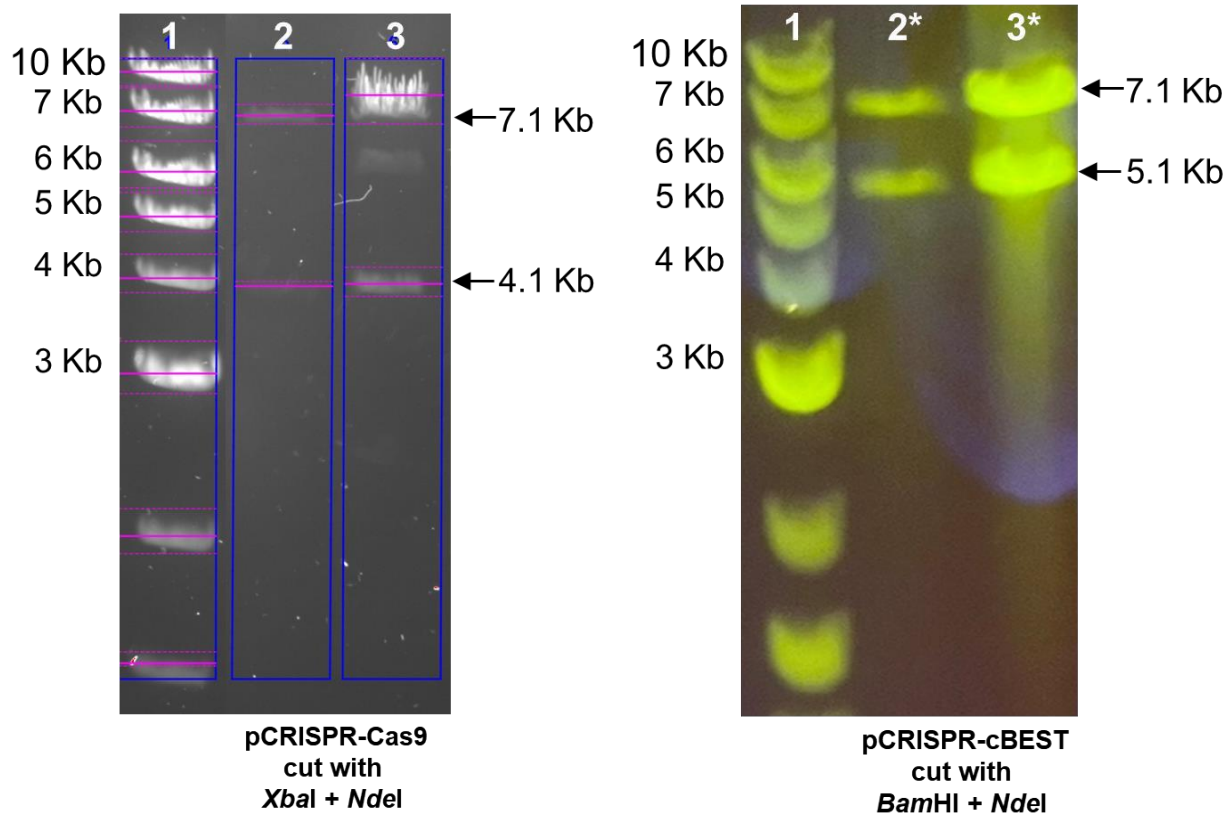


Figure 14. Restriction map of pCRISPR-Cas9 and pCRISPR-cBEST. 1 Kb DNA ladder used as reference in lane 1. Lanes 2-3 show pCRISPR-Cas9 plasmid DNAs extracted from two different *E. coli* colonies and digested with *XbaI* and *NdeI* restriction enzymes (left). Lanes 2*-3* show pCRISPR-cBEST plasmid DNAs extracted from two different *E. coli* colonies and digested with *BamHI* and *NdeI* restriction enzymes (right).

sgRNA scaffold and the rest of pCRISPR-cBEST plasmid DNA in 11.7 Kb sized large fragment (Fig. 15).

The small 450 bp fragment from the digestion of pCRISPR-cBEST with *NcoI* and *NdeI* (Fig. 15) was extracted and purified, so that it could be used as an insert fragment. The large 10.8 Kb fragment of pCRISPR-Cas9 by digestion using *NcoI* and *NdeI* was also extracted and purified, to be used as a vector fragment. The two fragments were ligated using T4 DNA ligase and a reaction protocol was established using the NEBioCalculator (Available at <https://nebiocalculator.neb.com/#!/ligation>). The map of the expected resulting plasmid pCRISPR-Cas9-KM was generated using the SnapGene software (Fig. 16).

Ligation products were transformed into chemically competent DH5 α and incubated on LA plates containing 25 μ g/ml Apramycin. If the ligation was successful, *E. coli* transformants carrying a circular plasmid conferring apramycin resistance gene will appear to grow. After 16 hours of incubation, many *E. coli* colonies appeared on agar medium indicating that they are apramycin resistant due to the ligated circular plasmid DNA.

3.1.3. Confirmation of pCRISPR-Cas9-KM plasmid

To see if these *E. coli* colonies grown on apramycin plates carried desired ligation products, instead of unforeseen self-ligated products due to the contamination of incomplete digested plasmids, colony PCR analysis of some *E. coli* transformants was performed. To confirm the *NcoI-NdeI* region containing the sgRNA core sequence from pCRISPR-cBEST where the desired single 'G' deletion is included, colony PCR with oligos pCRISPR-Cas9KM_FWD_2942 and pCRISPR-Cas9-KM_REV_3513 (Table 3) was performed. An expected size of PCR product in this case was 0.54 Kb. Another colony PCR using oligos pCRISPRCas9_FWD_Stu and pCRISPR-Cas9_REV_Stu (Table 3) was also performed to confirm the *StuI* region from pCRISPR-Cas9 where the 1 Kb upstream and downstream regions of desired target site should be cloned for the HDR oriented gene editing. PCR in this case was expected to produce a 0.42 Kb size product. Five colonies for each case were selected and analysed but only one set of each PCR reaction was shown in the figure. As shown in Fig. 17, both expected sizes of PCR products were obtained from two separate PCR reactions indicating that selected *E. coli* transformants are likely to carry desired resulting plasmid DNA. These *E. coli* transformants were then inoculated in liquid LB medium, containing 25 μ g/ml Apramycin, for plasmid DNA extraction and digestion with restriction enzymes. Plasmid DNAs from each culture were extracted using QIAprep Spin Miniprep Kit and their concentrations were measured by NanoDrop. Plasmid DNA up to 1 μ g was used for all restriction digestions. The extracted DNAs were digested with the restriction enzymes *NdeI* and *XbaI*, which were expected to present two different sized fragments with 7.1 Kb and 4.1 Kb.

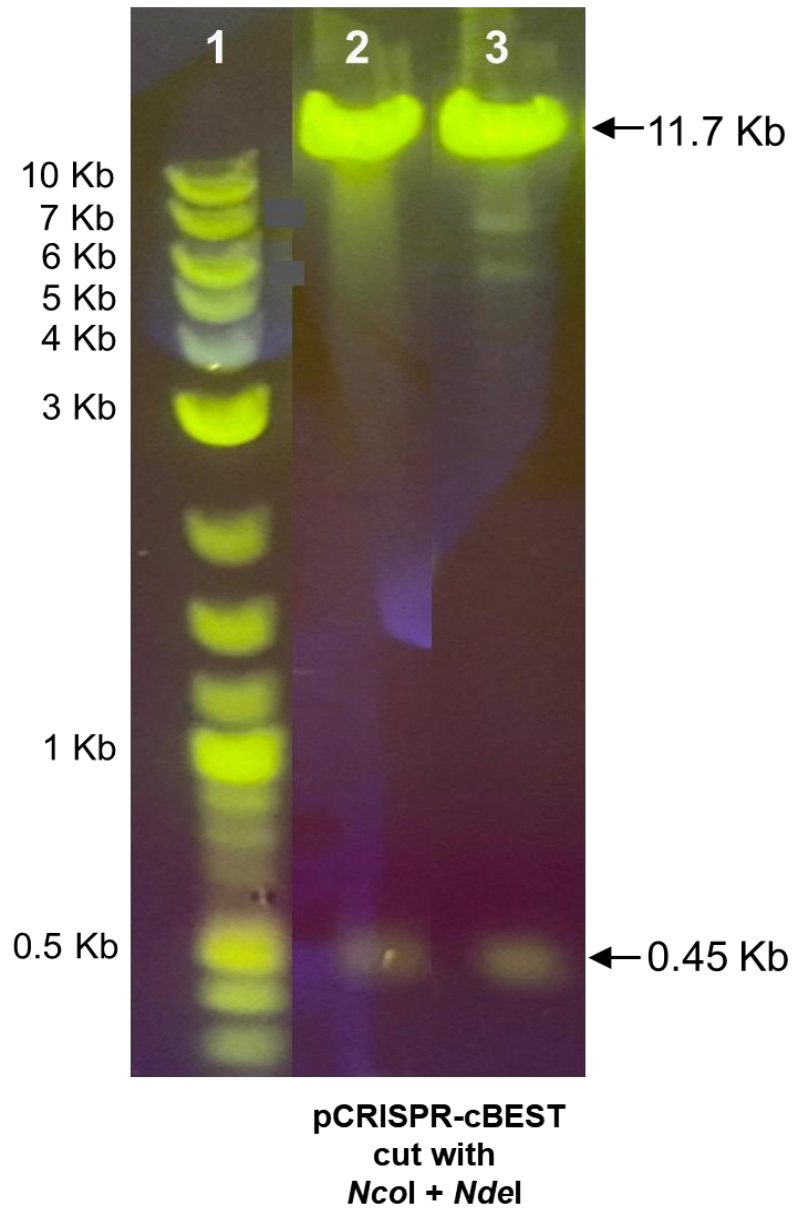


Figure 15. Restriction map of pCRISPR-cBEST. 1 Kb DNA ladder used as reference in lane 1. Lanes 2-3 show pCRISPR-cBEST plasmid DNAs extracted from two different *E. coli* colonies and digested with *NcoI* and *NdeI* restriction enzymes.

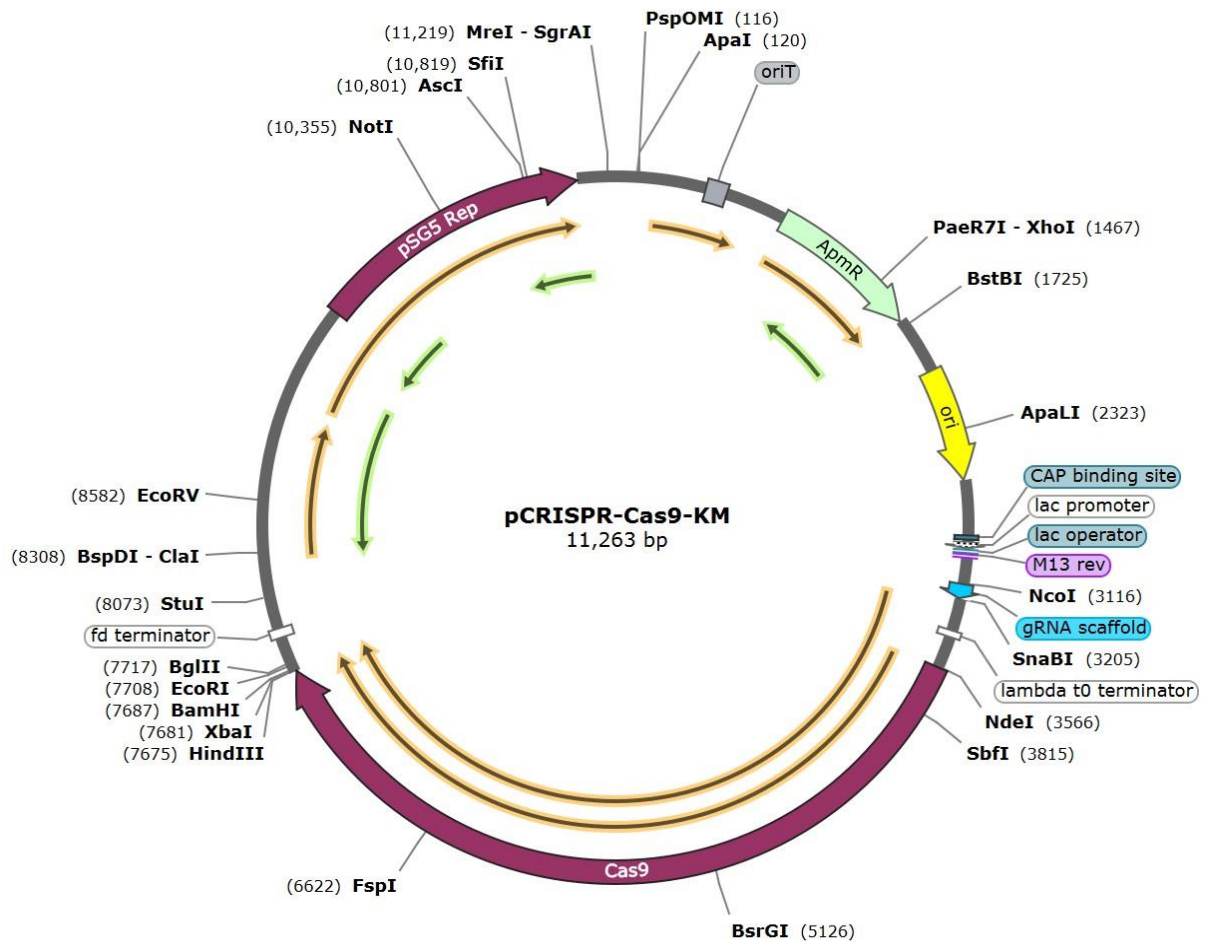


Figure 16. The pCRISPR-Cas9-KM editing plasmid, a chimaera utilising the Cas9 protein from pCRISPR-Cas9 and the sgRNA cassette from pCRISPR-cBEST (Image generated using SnapGene software; own creation).

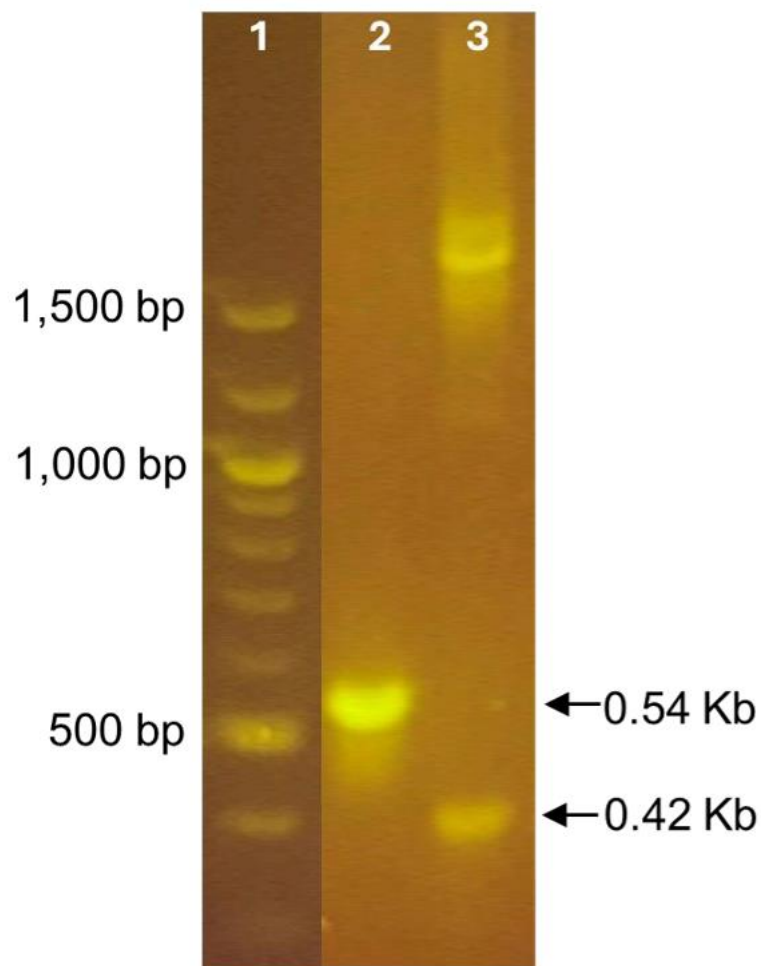


Figure 17. Colony PCR of pCRISPR-Cas9-KM. 100 bp DNA ladder used as reference in lane 1. Lane 2 shows the expected 0.54 Kb PCR fragment for the desired *NcoI-NdeI* region containing the sgRNA core sequence from pCRISPR-cBEST where the single 'G' deletion is included. Lane 3 shows the expected 0.42 Kb fragment for the *StuI* region from pCRISPRCas9 where the 1 Kb upstream and downstream regions of desired target site should be cloned for the HDR oriented gene editing.

Strangely, however, *NdeI* and *XbaI* digestion of pCRISPR-Cas9-KM did not show an expected digestion pattern (Fig. 18A). Among the five plasmid DNAs extracted from five different *E. coli* transformants, four of them produced a 7 Kb size single band (Lane 2 in Fig. 18A) with *NdeI*-*XbaI* digestion and one of them produced two different sized bands, which included one with an expected size of 4.1 Kb but the other one was almost 2 Kb bigger than the expected second band with the size of 9 Kb (Lane 3 in Fig. 18A). Further analysis of the pCRISPRCas9-KM plasmid DNA was conducted through a restriction digestion with *NcoI* to linearise the plasmid and to confirm its full size. Another digestion with *NcoI* and *XhoI* using one of the Lane 2 plasmids in Fig. 18A was also performed to confirm the size of plasmid DNA obtained by the single cut with *NcoI*. However, again the expected digestion pattern was not observed (Fig. 18B). *NcoI* digestion of this plasmid still produced a 7 Kb sized single band (Lane 2-1 in Fig. 18B) which matches with the size of single band from the previous digestion of the plasmid DNA with *NdeI*-*XbaI* (Lane 2 in Fig. 18A). Digestion of the same plasmid with *NcoI*-*XhoI* produced two bands with the size of 5.5 Kb and 1.6 Kb (Lane 2-2 in Fig. 18B) so the estimated total size of this plasmid DNA is 7 Kb which matches with the size of *NcoI* single cut product. But none of these digestion patterns confirmed the expected total size of desired pCRISPRCas9-KM which should be ~11 Kb, thus it was determined that genuine pCRISPR-Cas9-KM was not produced.

3.1.4. Re-configuration of pCRISPR-Cas9-KM

The construction of the desired pCRISPR-Cas9-KM was re-configured; larger ligation fragments were utilised, as the previous ligation method using a large vector and a small insert was stated to have a lower ligation efficiency (Addgene 2024; NEB, 2025a). From these sources, it was thought that the ligation of two larger fragments would facilitate easier successful production of pCRISPR-Cas9-KM, despite more difficult DNA alignment. Plasmid DNA was extracted from new colony growth of both pCRISPR-Cas9 and pCRISPR-cBEST and digested using *StuI* and *NdeI*. 6.7 Kb fragment from pCRISPR-cBEST and 4.5 Kb fragment from pCRISPR-Cas9 were selected for ligation. The ligation products were then introduced into competent *E. coli* DH5 α . A few selected *E. coli* transformants grown on LA containing 25 μ g/ml apramycin were selected and cultured in the liquid LB medium containing the same concentration of apramycin. Extracted plasmid DNAs were tested again with colony PCR and restriction digestion. The resulting plasmid was designated as pCRISPR-Cas9-KM*.

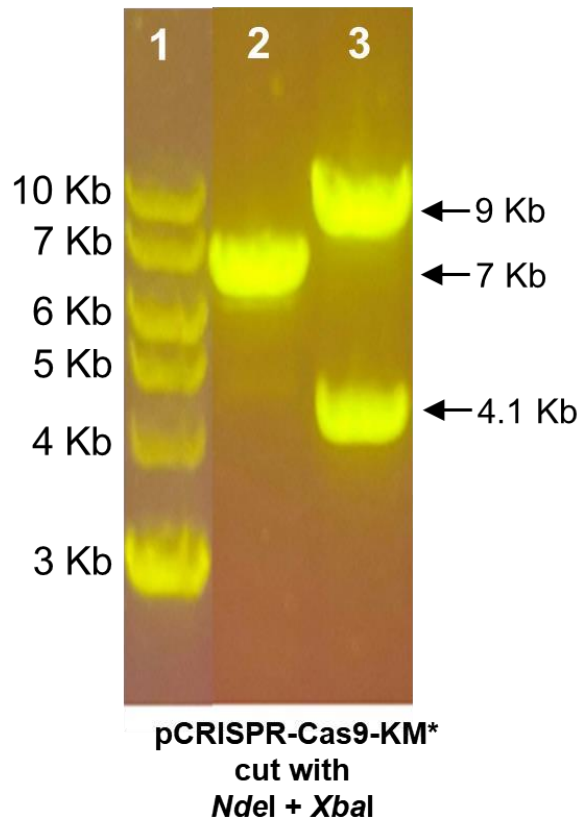
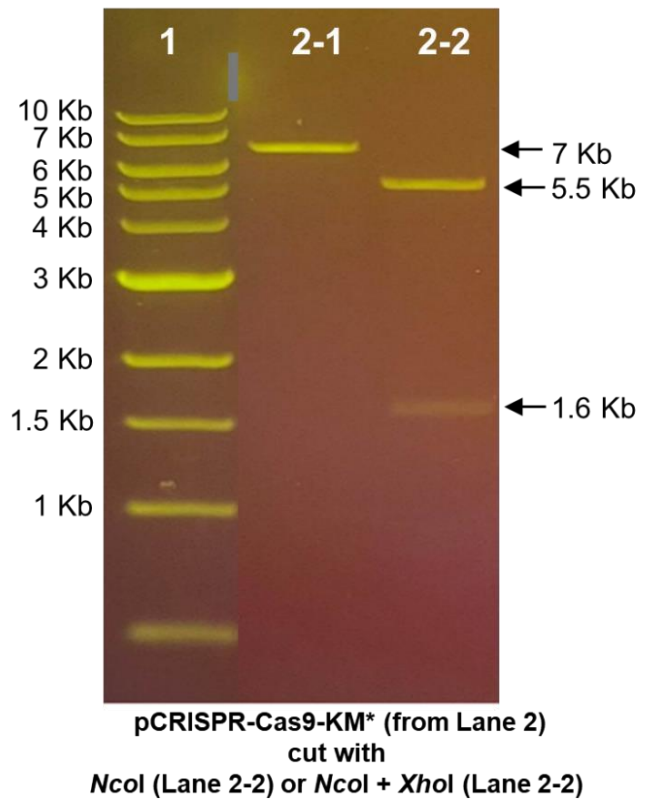
A**B**

Figure 18. Restriction map of pCRISPR-Cas9-KM*. 1 Kb DNA ladder used as reference in lane 1. (A) pCRISPR-Cas9-KM* cut with *NdeI*-*XbaI*. Lane 2 shows only a single 7 Kb sized band. Lane 3 shows two bands with the size of 9 Kb and 4.1 Kb. (B) pCRISPR-Cas9-KM* plasmid DNA of Lane 2 in (A) cut with *NcoI* (Lane 2-1) and *NcoI*-*XhoI* (Lane 2-2). Lane 2-1 shows a single 7 Kb sized band. Lane 2-2 shows two bands with the size of 5.5 Kb and 1.5 Kb.

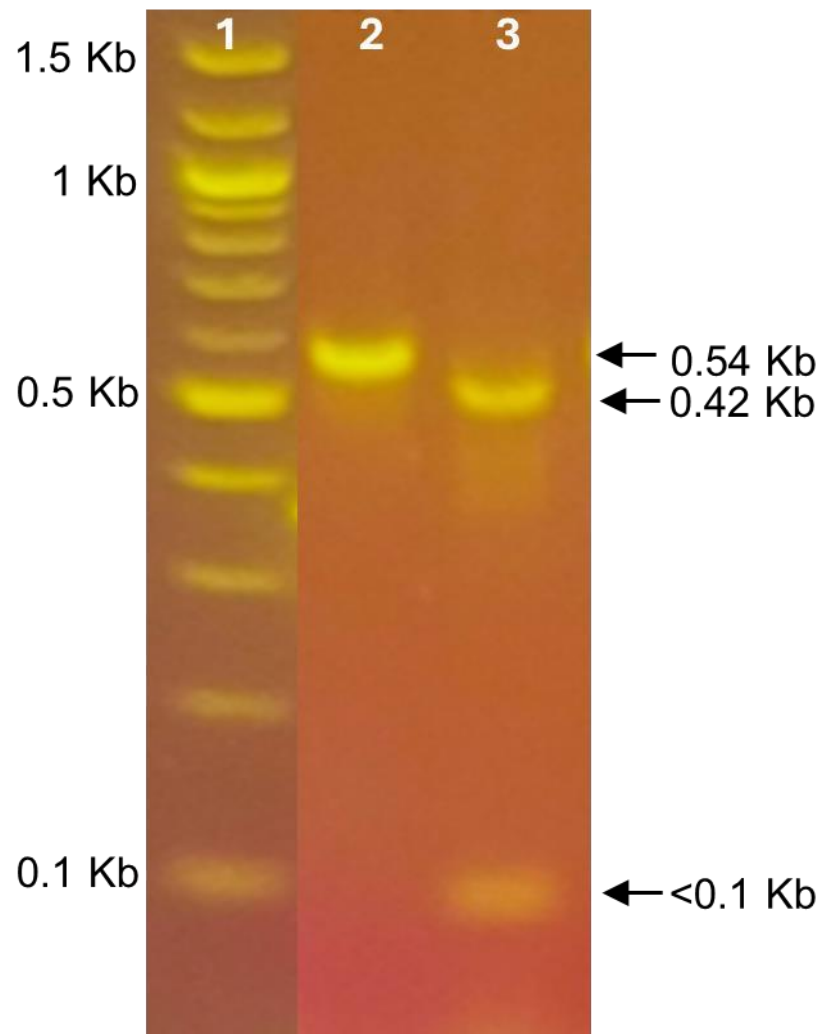


Figure 19. Colony PCR of pCRISPR-Cas9-KM* following plasmid re-configuration. 100 bp DNA ladder used as reference in lane 1. Lane 2 shows the expected 0.54 Kb PCR fragment for the region containing the sgRNA core sequence from 6.7 Kb ligation fragment of pCRISPR-cBEST. Lane 3 shows the expected 0.42 Kb fragment for the *StuI* region from 4.5 Kb ligation fragment of pCRISPR-Cas9 but also shows unexpected additional PCR product with <100 bp sized band.

3.2. Analysis of the inability to create pCRISPR-Cas9-KM

3.2.1. Colony PCR of pCRISPR-Cas9-KM following plasmid re-configuration

Through colony PCR analysis of pCRISPR-Cas9-KM samples following re-configuration of the plasmid arrangement, it is probable that no pCRISPR-Cas9-KM* is present. The fragment sizes observed in Fig. 19 do show banding at the desired 0.54 Kb from the colony PCR using oligos pCRISPR-Cas9-KM_FWD_2942 and pCRISPR-Cas9-KM_REV_3513 (Table 3) to confirm the region containing the sgRNA core sequence from the 6.7 Kb fragment of pCRISPR-cBEST plasmid. However, the other colony PCR using oligos pCRISPRCas9_FWD_Stu and pCRISPR-Cas9_REV_Stu (Table 3) to confirm the *StuI* region from the 4.5 Kb fragment of pCRISPR-Cas9 plasmid produced an unexpected product sized ~0.1 Kb in addition to the desired 0.42 Kb band (Fig. 19). Eighteen colonies were initially tested, and all showed the same banding pattern as shown in Fig. 19. Further colony PCR analysis was conducted on new pCRISPR-Cas9-KM* transformation products, however again, the expected size products are not created. Therefore, it was concluded that desired pCRISPRCas9-KM was not produced again in these trials.

3.2.2. Restriction enzyme analysis of subsequent pCRISPR-Cas9-KM*

Although an unexpected small fragment of PCR product was observed from the colony PCR using the newly configured pCRISPR-Cas9-KM*, the result obtained both expected bands at 0.54 Kb and 0.42 Kb. To investigate this resulting plasmid further, restriction digestion with *StuI* and *NdeI* was performed further. Two expected bands with the size at 6.7 Kb and 4.5 Kb were present this time (Fig. 20) however two unexpected additional bands both with size >10 Kb were also present. Moreover, one of the expected sized bands, 6.7 Kb, appeared very weak on the gel and one of unexpected band sized ~12 Kb appeared strongly, implicating that the 4.5 Kb and 12 Kb fragments are the major products of the complete *StuI-NdeI* digestion of the plasmid and the other two weak bands are the product from the incomplete digestion in the background.

3.2.3. Sequencing analysis of reconfigured pCRISPR-Cas9-KM* plasmid indicates unexpected rearrangements

Following the results observed by colony PCR and restriction digestion, oligonucleotide synthesis sequencing analysis of the reconfigured pCRISPR-Cas9-KM* plasmid DNA was conducted. Five premixed samples containing the same oligonucleotide, pCRISPR-Cas9KM_FWD_2942, complementary to the region at the *NcoI* site from pCRISPR-

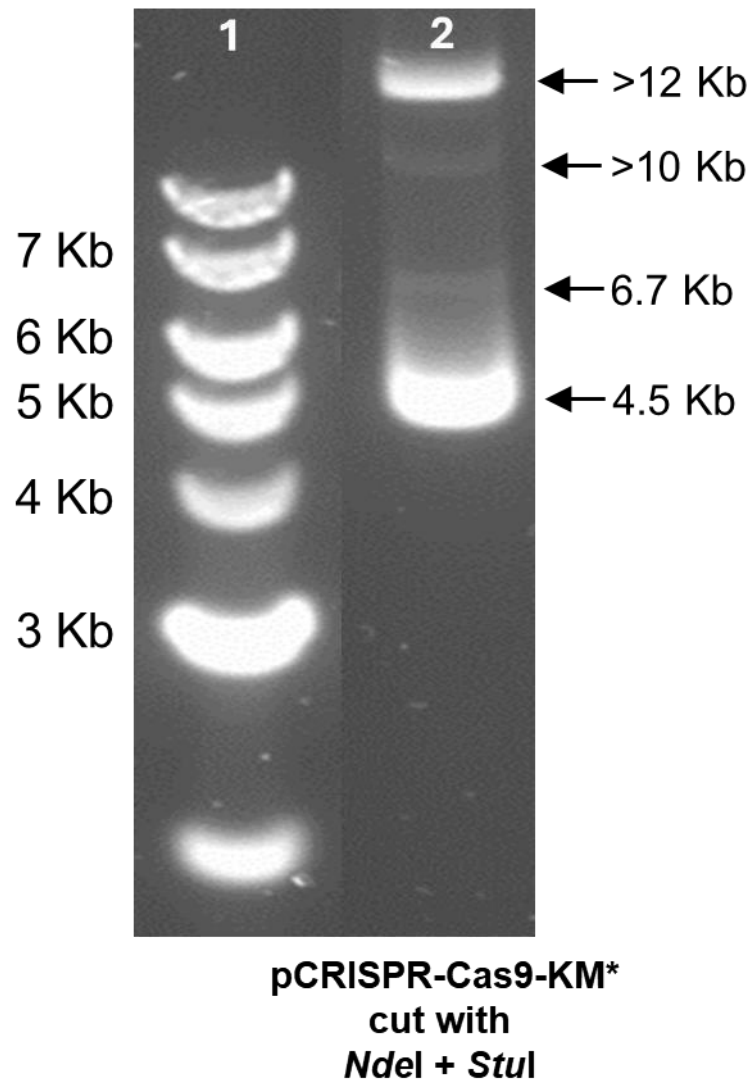


Figure 20. Restriction map of pCRISPR-Cas9-KM*. 1 Kb DNA ladder used as reference in lane 1. Lane 2 is the reconfigured pCRISPR-Cas9-KM* plasmid DNA digested with *NdeI* and *StuI*. This lane shows two strong bands with the size of 4.5 Kb and >12 Kb and two weak bands with the size of 6.7 Kb and >10 Kb.

cBEST, were also used for the former colony PCR analysis and pCRISPR-Cas9-KM* plasmid DNAs were extracted from five different *E. coli* transformants and sent to Eurofins for sequencing.

The analysed sequencing results indicated that the desired 6.7 Kb ligation fragment from pCRISPR-cBEST is present in all five samples, thus the resulting plasmid contains the required modified sgRNA cassette from pCRISPR-cBEST. However, further sequencing analysis of the whole plasmid DNA revealed that the pCRISPR-Cas9-KM* plasmid does not contain the desired 4.5 Kb ligation fragment from pCRISPR-Cas9. The sequences of the same five differently extracted pCRISPR-Cas9-KM* plasmid DNAs were obtained through whole plasmid Sanger sequencing from Eurofins. The SnapGene viewer software was then utilised to align generated whole plasmid DNA sequencing data against the pCRISPR-Cas9-KM* plasmid map to determine the sequences of pCRISPR-Cas9-KM*. Complete sequence analysis has shown that a chunk of over 4.5 Kb sized DNA in the same location is missing from three out of five plasmid DNA samples sent for the sequencing. This missing region includes a large part of gene encoding Cas9 protein therefore only a fragment of *cas9* gene is present in these three pCRISPR-Cas9-KM* plasmid DNAs (Fig. 21). This contradicts the colony PCR result shown in Fig. 19 as the 0.42 Kb PCR band should not have been produced if a complete *cas9* gene is not present. However, the 0.54 Kb sized fragment observed through former colony PCR analysis was also observed by the whole plasmid sequencing data as such the desired *NcoI* region containing the sgRNA core sequence with single 'G' deletion from pCRISPR-cBEST was present in two different plasmid samples.

Interestingly, sequencing data analysis also revealed that one of five sequenced plasmid samples had an insertion of unexpected external DNA fragments in multiple locations in the pCRISPR-Cas9-KM*. One of them carrying a fragment of *b/e* gene and another fragment of FRT mutant next to each other and were found immediately downstream of where the apramycin resistance gene (*apmR*) is located in the plasmid (Fig. 22). The second inserted DNA contains a fragment of *lacI* gene which was located at the immediate downstream of the *ori* gene in the plasmid. The third inserted DNA carries a range of different DNA fragments including a multiple cloning site (MCS), M13 gene VIII and an *insG* which was found right in the middle of the fd terminator locus near where the single *StuI* restriction site is located in the plasmid (Fig. 22). Analysis of the restriction sites in the whole plasmid sequencing data generated provided an explanation as to why unexpected restriction patterns of the pCRISPR-Cas9-KM* plasmid were observed, as the exact plasmid sizes were identified, and it was found

that some plasmid samples had their expected restriction sites absent. These abnormal restriction site patterns explain why unexpected restriction patterns were also observed when digesting samples with restriction enzymes (Fig. 18 & 20).

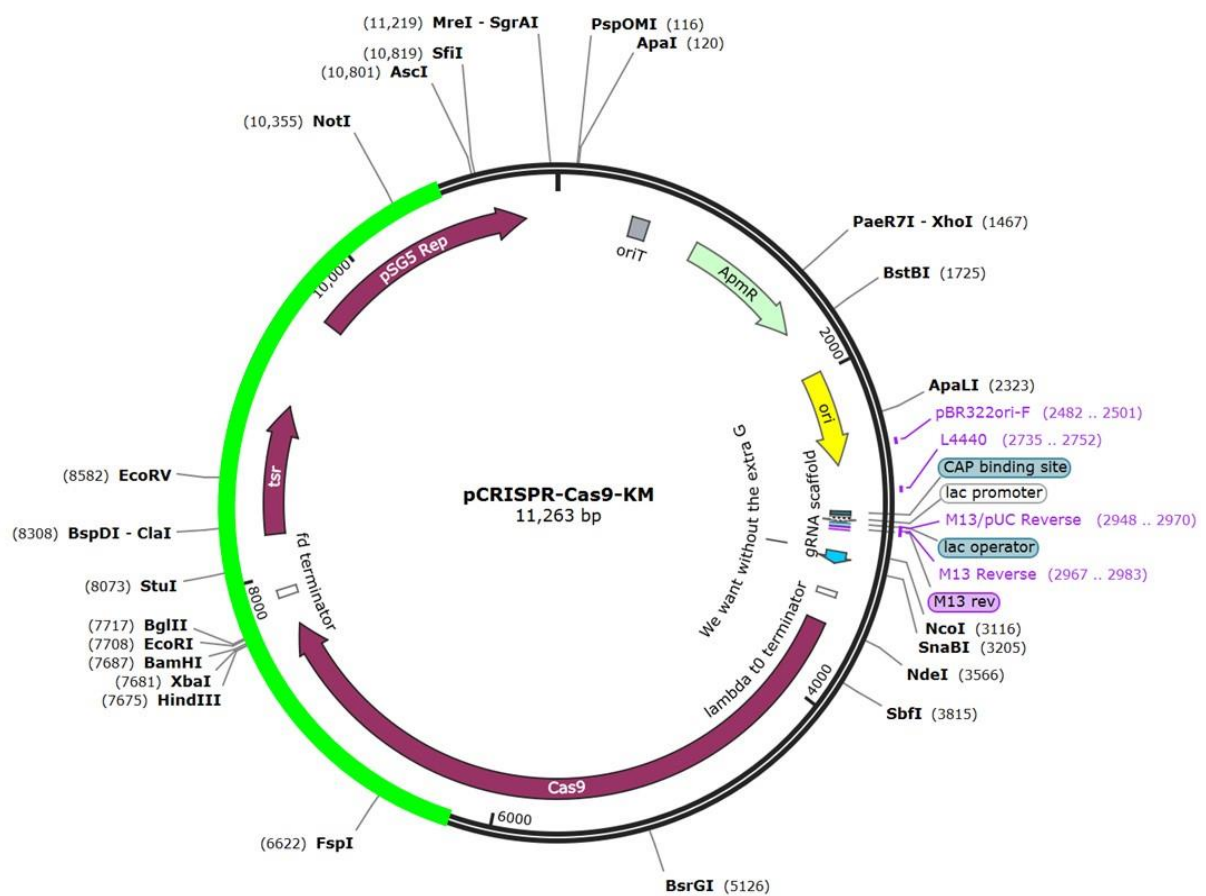


Figure 21. A map of desired the pCRISPR-Cas9-KM indicating the location of a missing 4.5 Kb region highlighted in fluorescent green in the reconfigured pCRISPR-Cas9-KM* plasmid based on the whole plasmid DNA sequencing data.

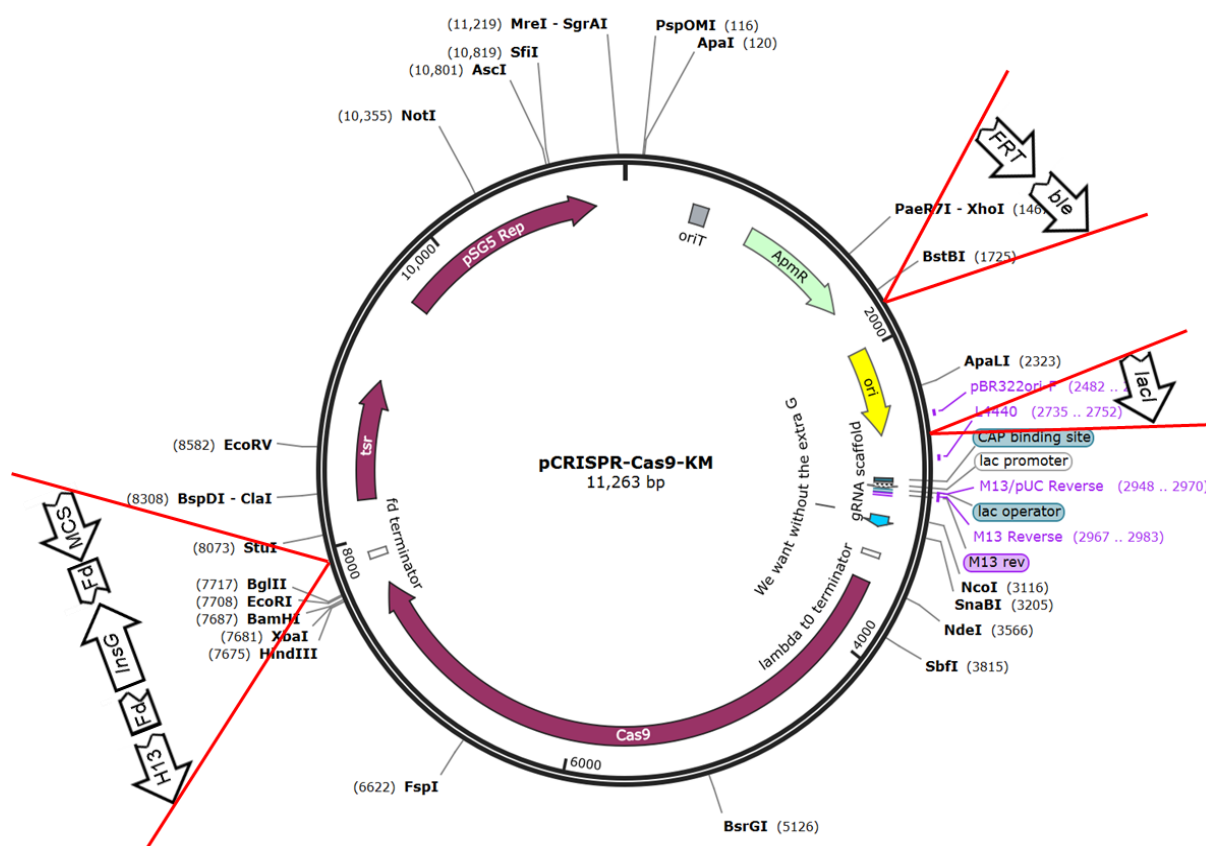


Figure 22. A map of pCRISPR-Cas9-KM indicates the multiple locations where unexpected insertion of external DNA fragments were found in the reconfigured pCRISPR-Cas9-KM* plasmid based on the whole plasmid DNA sequencing data.

4. Discussion

4.1. Inability to produce desired pCRISPR-Cas9-KM plasmid due to unexpected rearrangement of plasmid DNA

During this study, even after the trial of two different configuration methods (Fig. 23), the production of desired pCRISPR-Cas9-KM was not achieved, which is evident due to the occurrence of unexpected results through colony PCR analysis, restriction enzyme digestion and finally confirmed by whole plasmid DNA sequencing results. Analysis of the whole plasmid DNA sequencing data revealed that a range of unexpected rearrangements within the plasmid were undertaken. Three out of five sequences of pCRISPR-Cas9-KM* plasmid samples showed that a large 4,565 bp size fragment, covering the second half of the *cas9* gene, the whole of the thiostrepton resistance gene (*tsr*) and almost all of the temperature sensitive pSG5 Rep region was lost in all three plasmids (Fig. 21). In addition to the loss of DNA chunk from the pCRISPR-Cas9-KM* plasmid, the insertion of a range of different external DNA fragments including *lacI*, *ble*, FRT mutant, multi cloning site, M13 gene VIII and *insG* in multiple loci in the same plasmid but in the different sequenced samples were also observed. Insertion of a 93 bp size fragmented gene encoding a LacI protein was found immediately downstream of where the *ori* gene is present in the plasmid (Fig. 22). *lacI* encodes a lactose repressor protein, LacI, and is derived from the *lac* operon found natively in *E. coli*. LacI protein controls the expression of genes in response to the presence of lactose or IPTG (Phillips *et al.*, 2019). The other inserted external fragments were found immediately downstream of where the apramycin resistance gene, *apmR*, is located. This locus contains two different inserted DNA fragments, a *ble* gene and an FRT mutant in the same orientation. The *ble* is known as a bleomycin resistance gene and is not natively found in *E. coli*, but is used as a selectable marker in some CRISPR-based plasmids (Zhang *et al.*, 2020). Bleomycin belongs to an antineoplastics group of medicines and is used to treat several different types of cancer such as cervix and uterus cancer, head and neck cancer, testicle and penile cancer, and certain types of lymphoma. Bleomycin is also used as a selectable fungal marker and has been shown to support replication of fungal species (Nodvig *et al.*, 2015). The other fragment is a FRT refers to a FLP Recombination Target site, which is a DNA sequence used in site-specific recombination systems. Again, this fragment is not natively found in *E. coli*, but can be found in CRISPR-Cas based plasmids to facilitate scarless genome editing (Swings *et al.*, 2018). When two FRT sites are present in a plasmid or genome, FLP recombinase can mediate recombination between these sites. Thus, InDels or inversions of the genome between these

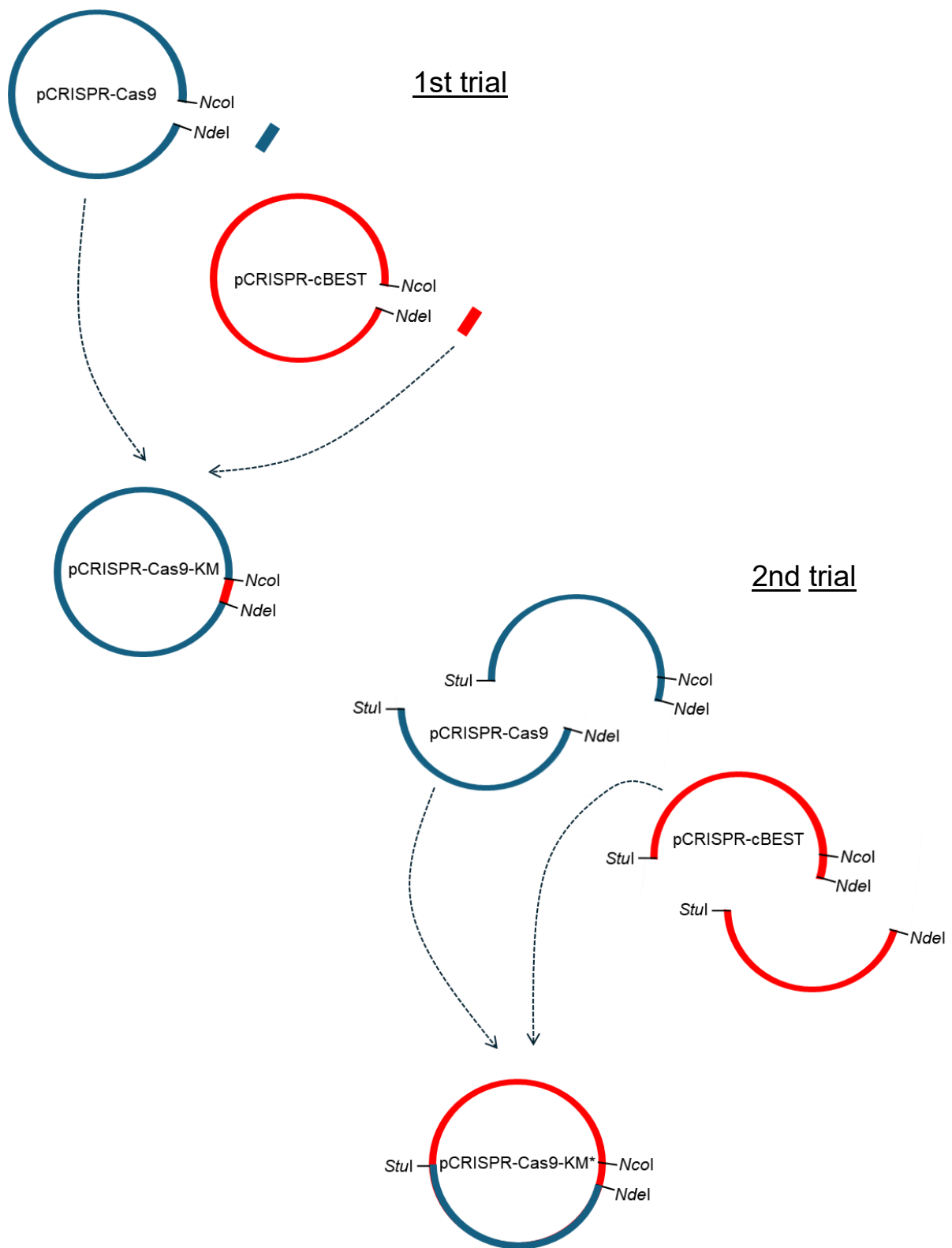


Figure 23. A simple diagram illustrating the two trials of ligation to configure pCRISPR-Cas9KM.

FRT sites are mediated. Mutated FRTs allow for recombination to only occur in identically mutated FRTs, therefore highly specific gene editing is conferred (Li *et al.*, 2020). The third external DNA insertion was observed right in the middle of the fd terminator in between where the *tsr* and *cas9* genes are located (Fig. 22). Inserted DNA contains multiple DNA fragments which were identified as multi cloning site (MCS), M13 gene III and an *insG* gene. The MCS and M13 gene III fragments were located in the same orientation whereas *insG* was located in the middle and facing the opposite direction to them and splitting the fd terminator in the plasmid. The MCS correlates to a region of DNA that contains a range of restriction enzyme recognition sites. This site is usually located in non-essential regions of the plasmid and facilitates the insertion of clone DNA fragments into the plasmid. Specifically, the MCS is used to insert guide RNA sequences and other genetic elements (e.g. fluorescent markers, antibiotic resistance genes or HDR templates) (Kim *et al.*, 2021). The M13 gene III is found in the M13 bacteriophage that attacks *E. coli*. When used in CRISPR technology, this M13 gene III is often used in phage display, whereby peptides or proteins are expressed on bacteriophage surfaces. The bacteriophage can also be utilised in delivering CRISPR components into bacterial cells (Passaretti *et al.*, 2020). However, the designed pCRISPR-Cas9-KM plasmid does not utilise such a mechanism. Lastly, the additional *insG* is a gene encoded by an IS4 transposable element. These small mobile genetic elements can move throughout the genome or to transposases. The *insG* gene within the IS4 element likely encodes for a transposase that mediates the IS4 insertion sequence mobility (Karp *et al.*, 1997). The insertion of this IS4 element at the fd terminator may cause plasmid instability if it disrupts regions essential for maintaining plasmid structural integrity. The disruption of the fd terminator can also stimulate recombination events, which lead to rearrangements of the plasmid, which again may reduce the plasmid stability.

4.2. Critical review of methods

The present study aimed to advance upon the most current CRISPR-Cas9 based systems, by utilising both a pCRISPR-Cas9 plasmid mediated mechanism, and a single nucleotide modification mechanism from pCRISPR-cBEST. However, when constructing this optimised plasmid, many issues were faced: ligation products did not produce the expected banding patterns when analysed via colony PCR and restriction enzyme digestion. The configured plasmid optimised both the HDR oriented Cas9 editing mechanism, suitable for more than one gene deletion at once, and a labour-saving version of a PCR-free, 1-step ssDNA bridging approach, which would allow for easier genome editing. The configured pCRISPR-Cas9-KM

was based on the pCRISPR-Cas9 and pCRISPR-cBEST plasmids (both produced by the Tong group), the same basic protocols were followed as outlined for the previous plasmids.

After the inability to produce the first configuration of pCRISPR-Cas9-KM, extensive sequence analysis of the parent plasmids pCRISPR-Cas9 and pCRISPR-cBEST was conducted. A concern of the pCRISPR-cBEST mechanism discovered by the Tong group (2019) was that genomic instability may result due to multiple uses of the same promoter or terminator that control sgRNA transcription. However, the use of alternative promoters or terminators was also avoided due to concerns of unevenly distributed sgRNAs as the result of different promoter strengths. To circumvent this issue, a Golden Gate Assembly-compatible Cys4based sgRNA self-processing system was designed. Golden Gate assembly is a molecular cloning technique that can efficiently assemble multiple DNA fragments simultaneously (Engler and Marillonnet, 2014). This Golden-gate based system would only requires a single promoter and terminator for multiple sgRNAs separated by Cys4 recognition sites. Cys4 is an RNA endonuclease derived from the *Pseudomonas aeruginosa* CRISPR-Cas system, which recognises specific Csy4 recognition sites in the CRISPR array. The binding of the endonuclease facilitates processing of the precursor sgRNAs and enables sgRNA multiplexing. This system was incorporated into the pCRISPR-cBEST plasmid by the Tong group (Tong *et al.*, 2019), but for the creation of pCRISPR-Cas9-KM, the system was not incorporated.

For the construction of the pCRISPR-Cas9-KM plasmid, both pCRISPR-Cas9's ability to target entire gene clusters and the pCRISPR-BEST multiplexed editing capabilities were desired. It was thought that to facilitate genome editing, and subsequent DNA repair via the HDR pathway, that only the modified sgRNA cassette from pCRISPR-cBEST was required. However, by only incorporating this segment of DNA, the Cys4-based sgRNA processing complexes were omitted. As a result, the lack of this system in the desired plasmid may have prevented the formation of the desired plasmid.

On the other hand, the pCRISPR-cBEST editing plasmid was not used to introduce mutations as the *cas9n* (D10A) it contains was found to be too toxic as a carrier protein for the deaminases in *E. coli*. *E. coli* lacks sufficient DNA repair pathways to circumvent the damage caused by *cas9n* (D10A), and so could not be used as an alternative (Tong *et al.*, 2019).

During the process of transforming ligation products into chemically competent DH5 α , it was noted that the *dcm*- (DNA cytosine methylase) nature of the *E. coli* strain would influence the restriction enzyme sites. Each restriction site selected was confirmed to have activity in the same reaction buffer and have the same incubation temperature to maximise the occurrence of successful restriction, and consequent successful ligation and transformation. It was also

noted that the *Xba*I site can become blocked as the consequence of *dam* methylation (Tong *et al.*, 2019).

To mitigate the occurrence of methylation or blockage of key sites, the optimised configuration of pCRISPR-Cas9-KM was selected to avoid these issues. However, this still resulted in an inability to produce pCRISPR-Cas9-KM. If the present project had allowed for more time in configuring and producing the desired pCRISPR-Cas9-KM, an alternative configuration, possibly utilising *Not*I and *Nde*I restriction sites, could generate the desired plasmid. But again, this configuration poses its own concerns as digestion using *Not*I would digest through the middle of the pSG5 replicon. Even considering the formation of sticky ends by *Not*I digestion, it would not pose very likely for the fragments to anneal as expected, to reproduce a complete pSG5 replicon, due to the utilisation of larger ligation fragments.

In hindsight, one area that the present study could have improved is the method by which the plasmid DNA concentration was quantified. The present study utilised a Nanodrop to obtain sample concentrations, however the use of a QuBit may produce a more accurate plasmid DNA quantification when at lower concentrations. QuBit uses a fluorescence-based dye to bind specifically to DNA and can distinguish between dsDNA and ssDNA. Measurements using a QuBit will also exclude RNA or ssDNA contaminants from the sample concentration value, thus provides a more sensitive and accurate DNA concentration reading. While the QuBit may provide more accurate DNA concentration readings when at lower levels, this method of quantification can be both time- and cost- expensive, and so may not be the optimal quantification method for this study (Versmessen, *et al.*, 2024).

4.3. Trial of Gibson Assembly to ligate 1 Kb up- and down-stream region of *vanK* into *Stu*I site of pCRISPR-Cas9-KM* using NEBuilder HiFi DNA Assembly

Despite the analysis of whole plasmid DNA sequencing data revealing two different unexpected rearrangements of pCRISPR-Cas9-KM*, one of these plasmids did preserve an intact *cas9*. The insertion of this whole gene, along with the additional external DNA fragments, would hopefully preserve the desired labour-saving version of the sgRNA cassette at the *Nco*I site with the single 'G' deletion from pCRISPR-cBEST in addition to other required elements for HDR oriented gene editing. Therefore, it may be plausible to use this plasmid further for HDR oriented gene editing to see if this rearranged version of pCRISPR-Cas9-KM would still work. The vancomycin resistance gene cluster (*vanRSJKHAX*) in *S. coelicolor* was selected as a positive control for this attempt. First of all, PCR should be undertaken to produce two fragments of 1 Kb PCR products complementary to the upstream and downstream regions of

the *van* cluster in *S. coelicolor*. For the upstream region, a pair of Van-Up_F and Van-Up_R oligos were used. For the downstream region, a pair of Van-Down_F and Van-Down_R oligos were used (Table 3). The next step is to utilise the NEBuilder HiFi DNA assembly kit, which was designed based on the principle of Gibson Assembly, whereby multiple DNA fragments are seamlessly annealed. When utilising the HiFi assembly kit, the fragments to be annealed require DNA overhangs of 15-30 base pairs and two pairs of specifically designed oligos (NEB, 2025b). The designed pair of oligos for the pCRISPR-Cas9-KM plasmid facilitated the insertion of two 1 Kb PCR products into the end of a linearised pCRISPR-Cas9-KM* plasmid DNA cut with *StuI* restriction enzyme. Then finally, a 60 bp sized single stranded DNA oligonucleotide targeting the middle of the *vanK* gene, which is located in the middle of the *van* cluster, was inserted into the *NcoI* site of the plasmid. The PCR products inserted into the plasmid at the *StuI* site and the sgRNA including the *vanK* target sequence at *NcoI* site in the plasmid facilitate HDR oriented deletion of the *van* cluster in *S. coelicolor* genome (Fig. 24).

The NEBuilder HiFi DNA assembly kit can seamlessly anneal many fragments with overlapping regions of 15-30 nt, regardless of their length and compatibility due to the actions of its exonuclease, DNA polymerase and DNA ligase. The assembly kit can also be used for the assembly of single-stranded oligonucleotides, provided there is suitable overlap and a T_m of $\geq 48^\circ\text{C}$. The New England Biolabs interactive NEBuilder Protocol Calculator tool was utilised (Available at: [NEBuilder Calculator](#)). For optimal reaction conditions for the three fragments it was ensured that all the fragments had overlapping regions that were at least 20 nucleotides long. The HiFi DNA Assembly kit follows a Gibson Assembly-based joining of fragments to form a continuous DNA molecule. Gibson Assembly implements the use of a 5' exonuclease, DNA polymerase and DNA ligase to seamlessly anneal fragments. As with the HiFi kit, Gibson Assembly requires the presence of overlapping sequences at least 20 nucleotides long and they are required to be complementary. The HiFi Assembly mix contains the same various enzymes as used in Gibson assembly and are provided within the same buffer. The exonucleases create ss 3' overhangs to facilitate annealing of the fragments that share complementarity at one end (overlapping regions). The polymerases fill gaps within annealed fragments, and DNA ligase seals nicks in the assembled DNA. The end result is a double-stranded fully sealed DNA that serves as a template for PCR. The product can also be used in direct transformation of competent *E. coli*.

Unfortunately, however the experiment did not progress further than the second stage as the sequencing analysis of the *StuI* site and *NcoI* site of the resulting product failed to produce the expected inserts. Nevertheless, it only one trial up to this second stage was established, therefore it is worthwhile to attempt more trials using the existing pCRISPR-Cas9-KM*.

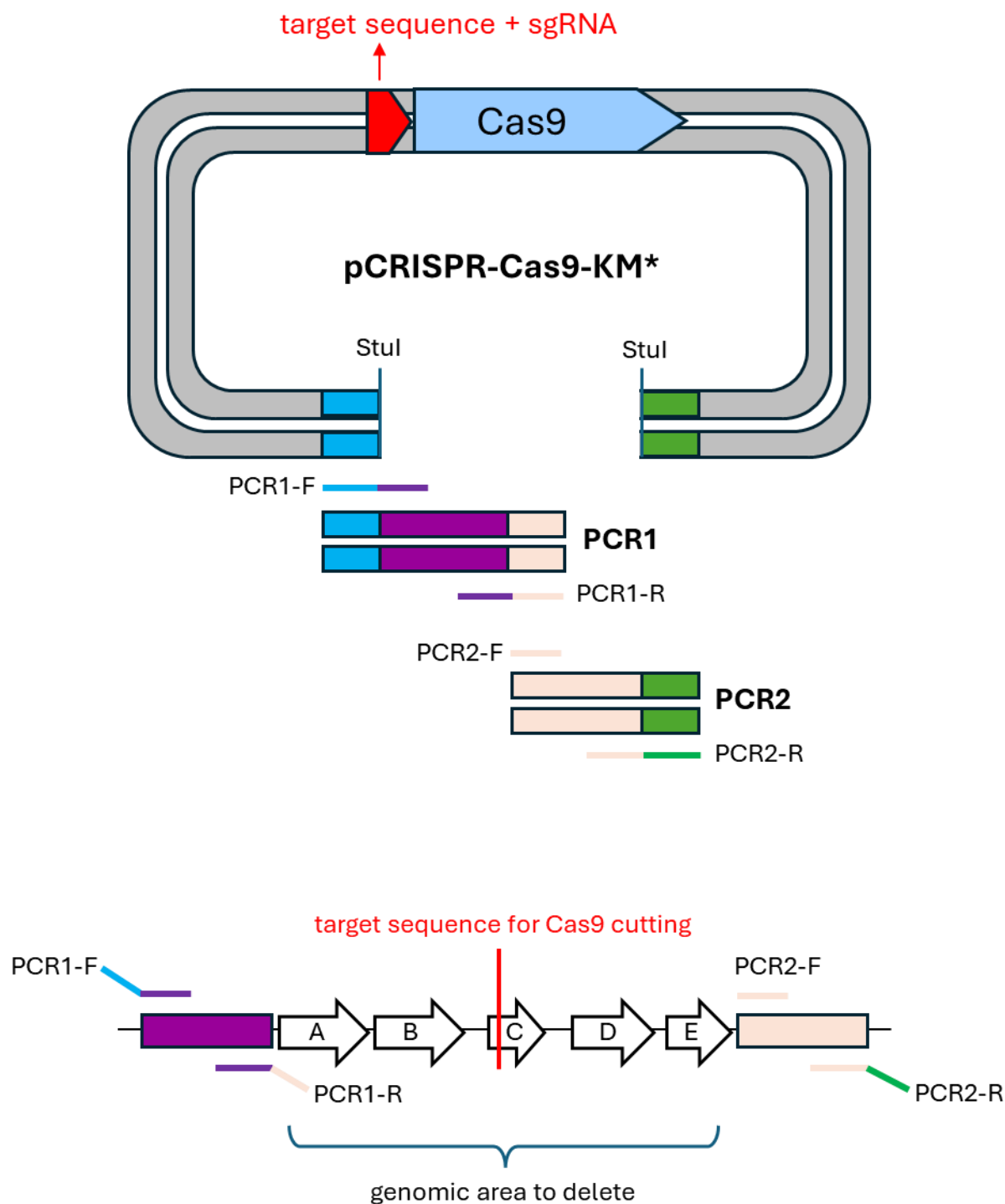


Figure 24. A diagram representing Gibson assembly to introduce two PCR fragments at the *Stul* site and oligo designing. For HDR oriented gene editing, a pair of 1 Kb-PCR products complementing the upstream and downstream of the target gene cluster should be inserted into the blunt end of the linearised plasmid mediated by *Stul* digestion.

4.4. Conclusions and future perspectives

Had the present study successfully generated the desired pCRISPR-Cas9-KM editing plasmid, and if successful introduction of the designed PCR products had been created following HiFi DNA assembly, genetic manipulations of *S. coelicolor* could be carried out. After successful insertion of the required genome targeting PCR products, conjugation between *E. coli* carrying the “programmed” plasmid would be carried out to insert this plasmid into *S. coelicolor* colonies. *E. coli* ETPUZ8002 is used as a vector to insert the genome editing plasmid into *S. coelicolor* due to the easily transformed nature of this strain: ETPUZ8002 replicates with a high copy number and can be isolated with a high yield. The strain ET12567/pUZ8002 is methylation deficient ($\Delta dcm\Delta dam$) and contains a driver plasmid pUZ8002 for conjugation to *Streptomyces* species. They also contain bifunctional vectors that replicate in both *E. coli* and *Streptomyces*, which allows for easier manipulation of DNA. These bifunctional vectors contain an *oriT* that allows for RP4-mediated mobilisation from *E. coli* to streptomyces (Flett *et al.*, 1997).

Following construction of the “programmed” pCRISPR-Cas9-KM editing plasmid, using PCR inserts targeting the *van* cluster, the editing efficacy of the resultant plasmid could be determined. The highly characterised vancomycin resistance cluster (Hong *et al.*, 2004; Hong *et al.*, 2005; Hutchings *et al.*, 2006) can be used as a positive control when determining the ability of plasmids to target and/or remove specific targets. If the vancomycin resistance cluster is successfully targeted and removed, resultant *S. coelicolor* should become sensitive to vancomycin exposure. To confirm vancomycin sensitivity following the *van* cluster deletion, biochemical testing to determine the MIC will be conducted. Sequencing verification of the resultant *S. coelicolor* DNA will also be carried out to confirm the correct mutagenesis of target DNA. Further bioassay and biochemical tests can then be carried out to quantify the phenotype of the successful mutant strain comparative to the wild type *S. coelicolor*. After verification of the pCRISPR-Cas9-KM editing capabilities, subsequent genetic manipulations targeting the suspected teicoplanin resistance cluster SCO6357-53 can be carried out. Again, biochemical testing to determine the MIC and phenotypic variation of mutated *S. coelicolor* will be conducted.

4.4.1. Functional genomic data concerning teicoplanin resistance clusters

The configured pCRISPR-Cas9-KM genome editing plasmid that this project sought to construct was intended to be utilised in gaining functional genomic insights concerning the teicoplanin resistance cluster in *S. coelicolor*. The resistance cluster SCO6357-53 was identified through previous functional genomic data (Connelly *et al.*, 2021), through the

overexpression of this cluster. The present study sought to build upon this work by analysing the effects of removing the cluster, through use of the genome editing plasmid pCRISPRCas9-KM.

Other studies that explore the innate teicoplanin resistant cluster and its biosynthesis in other Actinomycetes species, have shown striking similarities in their resistomes, as compared to *S. coelicolor*. Suspected orthologs of the *vanRS-HAX* genes have been identified in the producer strain *Actinoplanes teichomyceticus* (Yushchuk *et al.*, 2020a), but again, further genomic analysis is required to fully elucidate the teicoplanin metabolic pathway, and its subsequent interactions with other innate resistomes (Yushchuk *et al.*, 2020b).

Following successful deletion of the suspected teicoplanin resistance cluster in *S. coelicolor*, antibiotic susceptibility testing by determining the MIC of resultant colony growth will be determined. By quantifying the MIC of resultant SCO6357-53 knock-out mutants and comparing this MIC data to wild-type *S. coelicolor*, it can then be determined if the removal of this cluster of interest has a direct impact on antibiotic sensitivity. *S. coelicolor* naturally exhibits a high-level of intrinsic resistance to a range of antibiotics. As such, a library of antibiotics will be used; the *S. coelicolor* control strains sc10257m and sc10257v (as used in Connelly's study, 2021) show the same phenotype against most antibiotics, excluding vancomycin and kanamycin. The sc10257v strain is vastly more resistant to kanamycin. Resultant teicoplanin knock-out colonies will be analysed through exposure to teicoplanin, vancomycin, ramoplanin, carbenicillin, D-cycloserine, flavomycin, nisin, duramycin, novobiocin and kanamycin. Establishment of the MIC against a library of antibiotics, comparative to the control strains, will allow for further insights into the resultant genomic effects following removal of the suspected teicoplanin resistance cluster.

However, further complications at this stage must be taken into consideration. The study conducted by Connelly *et al.* (2021), from which this study is based, failed to generate successful knock-out deletions of the target cluster SCO6357-53 due to an inability to conjugate the CRISPR-based plasmid into *S. coelicolor*.

4.4.2. Implications and utilisations of an optimised CRISPR-Cas9 based plasmid

The development of actinomycete-optimised genome editing tools will allow for greater investigation into the innate antibiotic resistance mechanisms present within both producer and non-producer bacterial strains. Similarly, the CRISPR-Cas based editing tools can be utilised to target and activate innate biosynthetic gene clusters which are often silent or minimally expressed. The natural products found within Actinomycete species often contain bioactive products, particularly ones with antimicrobial properties. Expression of these

secondary metabolites could re-vitalise the discovery of novel antibiotics, or the discovery of equally significant anti-cancer or anti-viral compounds (Tao *et al.*, 2018).

Advancements into computational genome analysis can also lead to the identification of regions of interest. The software antiSMASH and various bioinformatic tools (PRISM, ClusterFinder etc.) can be used to reveal biosynthetic gene clusters. 'Genome mining' can be used to identify and predict secondary metabolites, and identify whether they contain antimicrobial, antifungal or anticancer properties (Lee *et al.*, 2020). Computational analysis of bacterial genomes that contain abundant secondary metabolites can implement genetic engineering and heterologous expression to confer genes that are not naturally expressed. By gaining better insights into the innate biosynthetic gene clusters found in actinomycete species, many novel natural products can be extracted and possibly developed into therapeutic drugs. Ultimately, all these efforts will aid in tackling the rising anti-microbial crisis we are currently facing.

References

- Abraham, E.P., Chain, E. (1988) 'An enzyme from bacteria able to destroy penicillin. 1940' *Reviews of Infectious Diseases*, 10 (4), pp.677–678.
- AddGene (2024) 'DNA ligation' *AddGene*. Available at: [Addgene: Protocol - How to Ligate Plasmid DNA](#). (Accessed 09/02/2025).
- Akkiraju, A.G., Badugu, A., Das, A. and Sagurthi, S.R. (2023) 'Molecular docking-based screening of methicillin-resistant *Staphylococcus aureus* FEM proteins with FDA-approved drugs'. *Bioinformation*. 19 (11), pp.1035-1042.
- Barman, A., Deb, B. and Chakraborty, S. (2020) 'A glance at genome editing with CRISPR–Cas9 technology.' *Current genetics*. 66, pp.447-462.
- Barrangou, R. and Marraffini, L.A. (2014) 'CRISPR-Cas systems: Prokaryotes upgrade to adaptive immunity'. *Molecular cell*. 54 (2), pp.234–244.
- Beltrametti, F., Consolandi, A., Carrano, L., Bagatin, F., Rossi, R., Leoni, L., Zennaro, E., Selva, E. and Marinelli, F. (2007) 'Resistance to glycopeptide antibiotics in the teicoplanin producer is mediated by *van* gene homologue expression directing the synthesis of a modified cell wall peptidoglycan' *Antimicrobial agents and chemotherapy*. 51 (4), pp.1135-1141.
- Bentley, S.D., Chater, K.F., Cerdeno-Tarraga, A.M., Challis, G.L., Thomson, N.R., James, K.D., Harris, D.E., Quail, M.A., Kieser, H., Harper, D., Bateman, A. *et al.* (2002) 'Complete genome sequence of the model actinomycete *Streptomyces coelicolor* A3(2)' *Nature*. 417 (6885), pp.141-7.
- Chain, E., Florey, H.W., Gardner, A.D., Heatley, N.G., Jennings, M.A., Orr-Ewing, J. and Sanders, A.G. (1940) 'Penicillin as a Chemotherapeutic Agent'. *Lancet*. 236 (6104), pp. 226–228.
- Chater, K. F. (1972) 'A Morphological and Genetic Mapping Study of White Colony Mutants of *Streptomyces coelicolor*'. *Journal of General Microbiology*. 72 (1), pp.9-28.
- Chater, K.F. and Bibb, M.J. (1997) 'Regulation of bacterial antibiotic production'. Chapter 2 of Biotech 7, *Production of Secondary Metabolites*. Kleinkauf, H. and von Dohren, H. (eds).
- Chen, A. (2017) 'CRISPR Technology explained: towards a CRISPR genome!' *BiteSizeBio*. Available at: [CRISPR Technology Explained: Towards a CRISPR Genome! \(bitesizebio.com\)](#) (Accessed 22/07/2024).
- Cobb, R.E., Wang, Y. and Zhao, H. (2015) 'High-Efficiency Multiplex Genome Editing of *Streptomyces* Species using an Engineered CRISPR/Cas System'. *ACS Synthetic Biology*. 4 (6), pp.723-728.
- Cong, Y., Yang, S. and Rao, X. (2020) 'Vancomycin resistant *Staphylococcus aureus* infections: A review of case updating and clinical features' *Journal of Advanced Research*. 21, pp.169-176.
- Connelly, S., Hong, H-J. and Hesketh, A. (2021) 'Understanding the Genome-Wide Response of *Streptomyces coelicolor* to the Antibiotic Teicoplanin' *Oxford Brookes University – Faculty of Health and Life Sciences*. PhD thesis. Available at: [Understanding-the-Genome-WideResponse-of-Streptomyces-coelicolor-to-the-Glycopeptide-Antibiotic-Teicoplanin.pdf \(researchgate.net\)](#) (Accessed: 21/09/2024).

Engler, C. and Marillonnet, S. (2014) 'Golden gate cloning' *Methods in Molecular Biology* 1116, pp.119-131.

Flett, F., Mersinis, V. and Smith, C.P. (1997) 'High efficiency intergeneric conjugal transfer of plasmid DNA from *Escherichia coli* to methyl DNA-restricting *Streptomyces*' *FEMS Microbiology Letters*. 155, pp.223-229.

Gov.uk (2018) 'Research reveals levels of inappropriate prescriptions in England'. *Public Health England*. Available at: [Research reveals levels of inappropriate prescriptions in England - GOV.UK \(www.gov.uk\)](https://www.gov.uk/government/news/research-reveals-levels-of-inappropriate-prescriptions-in-england) (Accessed 11/07/2024).

Gov.uk (2023a) 'The fight against antimicrobial resistance: progress and challenges'. *UK Health Security Agency*. Available at: [The fight against antimicrobial resistance: progress and challenges – UK Health Security Agency \(blog.gov.uk\)](https://www.blog.gov.uk/2023/07/11/the-fight-against-antimicrobial-resistance-progress-and-challenges/) (Accessed 11/07/2024).

Gov.uk (2023b) 'Antibiotic resistant infection and associated deaths increase' *UK Health Security Agency*. Available at: [Antibiotic resistant infections and associated deaths increase - GOV.UK \(www.gov.uk\)](https://www.gov.uk/government/news/antibiotic-resistant-infections-and-associated-deaths-increase) (Accessed 12/07/2024).

HM Government, Welsh Government, Scottish Government and The Department of Agriculture, Environment and Rural Affairs (2024) 'Confronting antimicrobial resistance 2024 to 2029' *Department of Health*. Available at: [Confronting antimicrobial resistance 2024 to 2029 \(publishing.service.gov.uk\)](https://publishing.service.gov.uk/government/uploads/system/uploads/attachment_data/file/131444/confronting-antimicrobial-resistance-2024-to-2029.pdf) (Accessed 12/07/2024).

Hong, H.-J., Hutchings, M.I., Neu, J.M., Wright, G.D., Paget, M.S.B. and Buttner, M.J. (2004) 'Characterisation of an inducible vancomycin resistance system in *Streptomyces coelicolor* reveals a novel gene (vanK) required for drug resistance'. *Molecular Microbiology*. 52 (4), pp.1107-1121.

Hong, H.-J., Hutchings, M.I., Hill, L.M. and Buttner, M.J. (2005) 'The role of the novel Fem protein VanK in vancomycin resistance in *Streptomyces coelicolor*'. *The journal of Biological chemistry*. 280 (13), pp.13055-61.

Hu, D., Chen, Y., Sun, C., Jin, T., Fan, G., Liao, Q., Mok, K.M. and Lee, M.-Y.S. (2018) 'Genome guided investigation of antibiotics producing Actinomycetales strain isolated from a Macau mangrove ecosystem.' *Scientific Reports*. 8 (14271), pp.1-12.

Huang, H., Zheng, G., Jiang, W., Hu, H. and Lu, Y. (2015) 'One-step high-efficiency CRISPR/Cas9-mediated genome editing in *Streptomyces*'. *Acta biochimica et biophysica Sinica*. 47 (4), pp.231-43.

Hutchings, M.I., Truman, A.W. and Wilkinson, B. (2019) 'Antibiotics: past, present and future' *Current Opinion in Microbiology*. 51, pp72-80.

Ishino, Y., Krupovic, M. and Forterre, P. (2018) 'History of CRISPR-Cas from Encounter with a Mysterious repeated sequence to genome editing technology'. *Journal of Bacteriology*. 200 (7), pp1-14.

Jinek, M., Chylinski, K., Fonfara, I., Hauer, M., Doudna, J.A., and Charpentier, E. (2012) 'A programmable dual-RNA-guided DNA endonuclease in adaptive bacterial immunity'. *Science*. 337, pp816–821.

Jiang, F. and Doudna, J.A. (2015) 'The structural biology of CRISPR-Cas systems'. *Current opinion in structural biology*. 30, pp.100-111.

Karp, P.D., Riley, M., Paley, S.M., Pellegrini-Toole, A. and Krummenacker, M. (1997) 'EcoCyc: Encyclopedia of *Escherichia coli* genes and metabolism'. *Nucleic Acids Research*. 25, pp.43–50.

- Kieser, T., Bibb, M.J., Buttner, M.J., Chater, K.F. and Hopwood, D.A. (2000) 'Practical *Streptomyces* Genetics' Crowes, Norwich, England; The John Innes Foundation.
- Kim, W-N., Kin, H-J., Chung, Y-S. and Kim H-U. (2021) 'Construction of multiple guide RNAs in CRISPR-Cas9 vector using stepwise or simultaneous golden gate cloning: case study for targeting the FAD2 and FATB multigene in Soybean'. *Plants*. 10 (2542), pp.1-17.
- Lawson, C.L., Swigon, D., Murakami, K.S., Darst, S.A., Berman, H.M., Ebright, R.H. (2004) 'Catabolite activator protein: DNA binding and transcription activation'. *Current Opinion in Structural Biology*. 14 (1), pp.10-20.
- Lee, N., Hwang, S., Kim, J., Cho, S., Palsson, B. and Cho, B-K. (2020) 'Mini review: Genome mining approaches for the identification of secondary metabolite biosynthetic gene clusters in *Streptomyces*'. *Computational and Structural Biotechnology Journal*. 18, pp.1548-1556.
- Li, J., Li, Y., Pawlik, K.M., Napierala, J.S. and Napierala, M. (2020) 'A CRISPR-Cas9, Cre-lox, and Flp-FRT cascade strategy for the precise and efficient integration of exogenous DNA into cellular genomes'. *The CRISPR journal*. 3 (5), pp.470-486.
- Li, L., Wei, K., Zheng, G., Liu, X., Chen, S., Jiang, W. and Lu, Y. (2018) 'CRISPR-Cpf1 assisted multiplex genome editing and transcriptional repression in *Streptomyces*'. *American Society for Microbiology – applied and environmental microbiology*. 84 (18), pp.1-18.
- Manyi-Loh, C., Mamphweli, S., Meyer, E. and Okoh, A. (2018) 'Antibiotic use in agriculture and its consequential resistance in environmental sources: potential public health implications'. *Molecules*. 23 (4), pp795.
- Munita, J.M. and Arias, C.A. (2016) 'Mechanisms of Antibiotic Resistance'. *American Society for Microbiology - Microbiology Spectrum*. 4 (2).
- Murray, C.J.L., Ikuta, K.S., Sharara, F., Swetschinski, L. and Aguilar, G.R. (2022) 'Global burden of bacterial antimicrobial resistance in 2019: a systematic analysis' *The Lancet*. 399 (10325), pp.629-655.
- Muth, G. (2018) 'The pSG5-based thermosensitive vector family for genome editing and gene expression in actinomycetes.' *Applied microbiology and biotechnology*. 102 (21), pp.90679080.
- New England BioLabs (2023) 'NEB Catalog' *NEB*. Available at: [NEB_Catalog_2023-2024.pdf \(neb-online.de\)](https://neb-online.de) (Accesses 18/06/2024).
- New England BioLabs (2025a) 'FAQs – Resources' *NEB*. Available at: [Search | NEB](#) (Accessed 09/02/2025).
- New England BioLabs (2025b) 'Gibson Assembly' *NEB*. Available at: [Gibson Assembly® | NEB](#) (Accessed 09/02/2025).
- Nodvig, C.S., Nielsen, J.B., Kogle, M.E. and Mortensen, U.H. (2015) 'A CRISPR-Cas9 system for genetic engineering of filamentous fungi'. *PLOS One*. 10 (7), pp.1-18.
- Novotna, G., Hill, C., Vincent, K., Liu, C. and Hong, H-J. (2012) 'A novel membrane protein VanJ, conferring resistance to teicoplanin'. *Antimicrobial agents and chemotherapy*. 56 (4), pp.1784-96.
- Passaretti, P., Khan, I., Dafforn, T.R. and Oppenheimer, P.G. (2020) 'Improvements in the production of purified M13 bacteriophage bio-nanoparticle'. *Scientific reports*. 10 (18538), pp.1-9.

Pepper, I.L., Gerba, C.P. and Gentry, T.J. (2015) 'Environmental Microbiology' 3rd Edition. *Academic Press*.

Phillips, K.N., Widmann, S., Lai, H.-Y., Nguyen, J., Ray, J.C.J., Balazsi, G. and Cooper, T.F. (2019) 'Diversity in *lac* operon regulation among diverse *Escherichia coli* isolates depends on the broader genetic background but is not explained by genetic relatedness'. *American Society for Microbiology*. 10 (6), pp.1-14.

Plagens, A., Richter, H., Charpentier, E. and Randau, L. (2015) 'DNA and RNA interference mechanisms by CRISPR-Cas surveillance complexes'. *FEMS microbiology reviews*. 39 (3), pp.442-463. Renneberg, R., Süßbier, D., Berkling, V. and Lorch, V. (2023) 'Biotechnology for beginners'. 3rd Edition. *Academic Press*. Singh, V. (2014) 'Recent Advancements in synthetic biology: Current status and challenges'. *Gene*. 535 (1), pp.1-11.

Roberts, J.A., Pea, F. and Lipman, J. (2013) 'The Clinical relevance of plasma protein binding changes'. *Clinical Pharmacokinetics*. 52, pp1-8.

Sola, C., Abadia, E., Le Hello, S. and Weill, F.-X. (2015) 'High-throughput CRISPR Typing of *Mycobacterium tuberculosis* complex and *Salmonella enterica* serotype typhimurium'. *Methods in Molecular Biology*. 1311, pp91-109.

Svetitsky, S., Leibovici, L. and Paul, M. (2009) 'Comparative Efficacy and Safety of Vancomycin versus Teicoplanin: Systematic Review and Meta-Analysis'. *Antimicrobial Agents and Chemotherapy*. 53 (10), pp.4069-4079.

Swings, T., Marciano, D.C., Atri, B., Bosserman, R.E., Wang, C., Leysen, M., Bonte, C., Schalck, T., Furey, I. *et al.* (2018) 'CRISPR-FRT targets shared sites in a knock-out collection for off-the-shelf genome editing'. *Nature communications*. 9 (2231), pp.1-10.

Tan, L.L., Heng, E., Leong, C.Y., Ng, V., Yang, L.K., Seow, D.C.S., Koduru, L., Kanagasundaram, Y., Ng, S.B., Peh, G., Lim, Y.H. and Wong, F.T. (2024) 'Application of Cas12j for *Streptomyces* editing.' *Biomolecules*. 14 (486), pp.1-8.

Tao, W., Yang, A., Deng, Z. and Sun, Y. (2018) 'CRISPR/Cas9-based editing of *streptomyces* for discovery, characterisation, and production of natural products'. *Frontiers in Microbiology*. 9 (1660), pp.1-8.

The Fleming Fund (2023) '£210 million to tackle deadly antimicrobial resistance'. *The Fleming Fund*. Available at: [£210 million to tackle deadly antimicrobial resistance | Fleming Fund](#). (Accessed 12/07/2024).

Tong, Y., Charusanti, P., Zhang, L., Weber, T. and Lee, S.Y. (2015) 'CRISPR-Cas9 Based Engineering of Actinomycetal Genomes'. *ACS Synthetic Biology*. 4 (9), pp.1020-9.

Tong, Y., Whitford, C.M., Robertsen, H.L., Blin, K., Jorgensen, T.S., Klitgaard, A.K., Gren, T., Jiang, X., Weber, T. and Lee, S.Y. (2019) 'Highly efficient DSB-free base editing for Streptomycetes with CRISPR-BEST'. *Proceedings of the National Academy of Sciences of the United States of America*. 116 (41), pp.20366-20375.

Tong, Y., Whitford, C.M., Blin, K., Jorgensen, T.S., Weber, T. and Lee, S.Y. (2020) 'CRISPR-Cas9, CRISPRi and CRISPR-BEST-mediated genetic manipulation in streptomyces.' *Nature Protocols*. 15, pp.2470-2502.

Tran, N.T., Huang, X., Hong, H.-J., Bush, M.J., Chandra, G., Pinto, D., Bibb, M.J., Hutchings, M.I., Mascher, T. and Buttner, M.J. (2019) 'Defining the regulon of genes controlled by σ^E , key

regulator of the cell stress response in *Streptomyces coelicolor*. *Molecular Microbiology*. 112 (2), pp.461-481.

Van Bambeke, F. (2015) 'Lipoglycopeptide Antibacterial Agents in Gram-Positive Infections: A comparative Review'. *Drugs*. 75 (18), pp.2073-95.

Van Groesen, E., Innocenti, P. and Martin, N.I. (2022) 'Recent Advances in the Development of semisynthetic glycopeptide antibiotics: 2014-2022'. *ACS Infectious Diseases*. 8, pp.1381-1407.

Versmessen, N., Van Simaey, L., Negash, A.A., Vandekerckhove, M., Hulpiau, P., Vaneechoutte, M. and Cools, P. (2024) 'Comparison of DeNovix, NanoDrop and QuBit for DNA quantification and impurity detection of bacterial DNA extracts'. *PLoS One*. 19 (6), pp.1-14.

Wildermuth, H. (1970) 'Development and organisation of the aerial mycelium in *Streptomyces coelicolor*'. *Journal of General Microbiology*. 60, pp.43-50.

Wilkinson, R.A., Martin, C., Nemudryi, A.A. and Wiedenheft, B. (2019) 'CRISPR RNA-guided autonomous delivery of Cas9'. *Nature Structural and Molecular Biology*. 26 (1), pp.14-24.

World Health Organisation (2023) 'Antimicrobial Resistance' WHO. Available at: [Antimicrobial resistance \(who.int\)](https://www.who.int/antimicrobial-resistance) (Accessed 12/07/2024).

Ye, S., Enghiad, B., Zhao, H. and Takano, E. (2020) 'Fine-tuning the regulation of Cas9 expression levels for efficient CRISPR-Cas9 mediated recombination in *Streptomyces*'. *Society for Industrial Microbiology and Biotechnology*. 47, pp.413-423.

Yushchuk, O., Homoniuk, V., Ostash, B., Marinelli, F. and Fedorenko, V. (2020a) 'Genetic insights into the mechanism of teicoplanin self-resistance in *Actinoplanes teichomyceticus*'. *The Journal of Antibiotics*. 73, pp.255-259.

Yushchuk, O., Ostash, B., Truman, A.W., Marinelli, F. and Fedorenko, V. (2020b) 'Teicoplanin biosynthesis: unravelling the interplay of structural, regulatory, and resistance genes.' *Applied Microbiology and Biotechnology*. 104, pp.3279-3291.

Zeng, D., Debahov, D., Hartsell, T.L., Cano, R.J., Adams, S., Schuyler, J.A., McMillan, R. and Pace, J.L. (2016) 'Approved Glycopeptide antibacterial drugs: mechanism of action and resistance'. *Cold Spring Harbor Perspectives in Medicine*. 6, pp.1-17.

Zhang, W-W., Lypaczewski, P. and Matlashewski, G. (2020) 'Application of CRISPR/Cas9-mediated genome editing in *Leishmania*'. *Trypanosomatids – Methods in Molecular Biology*. 2116, pp.199-224.

Zhu, Y. (2022) 'Advances in CRISPR/Cas9'. *BioMed research international*. (1), pp.9978571.