# Understanding the Genome-Wide Response of *Streptomyces coelicolor* to the Glycopeptide Antibiotic Teicoplanin

Samuel Connelly

Oxford Brookes University

Health and Life Sciences department

Supervisors: Dr Hee-Jeon Hong, Dr Victor M. Bolanos-Garcia and Dr Andrew Hesketh

# Contents

# Acknowledgements

There have been many people who have made this thesis possible and allowed me to get to this stage of my academic career. I would like to thank my three supervisors for their continued support during the last four years. Firstly, Dr Andrew Hesketh who has been an excellent teacher and provided me with the technical expertise throughout this project. I'd also like to thank him for being so patient and understanding during the first 6 month of my project. Secondly, I would like to thank Dr Victor M. Bolanos-Garcia for supporting my lab work and brightening my weekends spent in the office working on my PhD. Finally, I would like to thank Dr Hee-Jeon Hong who has always been there whenever I needed, but also helped to progress my career outside of her lab. I would also like to thank the support staff at Oxford Brookes including the technicians in the Tonge building, for always doing their best to help in times of difficulty. Finally, my friends and family also deserve to be acknowledged for the continual support during the past four years and being compassionate during my times of difficulty. This work was funded by the Nigel Groome Scholarship provided by the Health and Life Sciences department of Oxford Brookes University.

# Abbreviations

ACC – acetylCoA carboxylase

ACP – acyl carrier protein

ACT – actinorhodin

AMP – antimicrobial peptide

AMR – antimicrobial resistance

CAMP(s) – cyclic antimicrobial peptides

CFU – colony forming units

CL – cardiolipin

CPM – counts per million

DAG - Diacylglycerol

D-Ala-D-Ala – D-alanine-D-alanine

D-Ala-D-Lac – D-alanine-D-lactate

DDL - diacetyldihydrolutidine

DE – differential expression

DEG – differentially expressed gene(s)

DHAP - dihydroxyacetone phosphate

DIZ – diameter of inhibitory zone

DSB – double-stranded breaks

ECF – extracytoplasmic sigma factor

EEP - endonuclease/exonuclease/phosphatase

FC – fold-change

FPP – farnesyl diphosphate

G3P – glycerol-3-phosphate

GDPD - glycerophosphodiester phosphodiesterases

GlcNAc – N-acetylglucosamine

GO – gene ontology

HA – homology arms

IPP – isopentenyl diphosphate

LPG – lysylphosphatidylglycerol

LPS - lipopolysaccharide

MIC – minimum inhibitory concentration

MOP – **m**ultidrug/**o**ligosaccharidyl-lipid/**p**olysaccharide

murNAc – N-acetylmuramic acid

NGS – next generation sequencing

NMMP – Minimal liquid medium

NO – nitric oxide

OD – optical density

OL – ornithine lipid

OM – outer membrane

PA – phosphatidic acid

PAP – phosphatidic acid phosphatase

PBP – penicillin binding protein

PCA – principle component analysis

PE – phosphatidylethanolamine

PG – peptidoglycan

PpG – phosphatidylglyerol

PIM – phosphatidylinositol mannosidase

PS – phosphatidylserine

QC – quality control

RED – undecylprodigiosin

RR – response regulator

SK – sensor kinase

STK – serine threonine/kinase

TCS – two-component system

TMM – trimmed mean of M-values

TREAT – T-tests relative to a threshold

UP – undecaprenyl phosphate

UPP – undecaprenyl pyrophosphate

VISA – vancomycin-intermediate *Staphylococcus aureus*

VRE – vancomycin-resistant enterococci

# Abstract

The glycopeptide antibiotics vancomycin and teicoplanin are clinically important as a second-line therapy to treat nosocomial infections caused by Gram-positive pathogens. Glycopeptide antibiotics universally target the terminal residues, D-Alanyl-D-Alanine on the cell wall peptidoglycan intermediate lipid II, interfering with peptidoglycan biosynthesis and weakening the cell wall. A general resistance mechanism to these antibiotics requires a core set of genes, *vanRSHAX,* that detect a glycopeptide (VanS) and upregulate genes (VanR) which orchestrate the remodelling of D-Ala-D-Ala on lipid II to D-Ala-D-Lactate (VanHAX), reducing glyopeptide affinity by 1000-fold. Our previous study demonstrated that altering the termini of lipid II by VanHAX action is insufficient for providing resistance to teicoplanin in *S. coelicolor*, which is instead mediated mainly by the elusive membrane protein, VanJ.

This study further characterised VanJ by comparing the transcriptomes of a *wt S. coelicolor A3(2) M600* strain and an isogenic Δ*vanJ* knock-out mutant after exposing cells to teicoplanin, identifying that Δ*vanJ* exhibited increased signs of cellular stress that were attributed to a delayed induction of genes involved in the osmotic, redox, and cell envelope stress responses. This dataset led to the functional characterisation of a group of genes with phosphatidic acid phosphatase activity which affected the sensitivity of *S. coelicolor* to a broad range of cell wall targeting antibiotics. One of these genes, *SCO6355*, significantly counteracted the intrinsic *vanRSHAX* resistance system of *S. coelicolor*, lowering its high-level vancomycin resistance (80 μg/mL) by four fold, to intermediate levels (20 μg/mL). This work demonstrates a novel mechanism which can antagonise the function of intrinsic *van* resistance clusters that will be important in the development of strategies that can circumvent glycopeptide resistance in clinical pathogens.

# 1. Introduction

# 1.1 Antimicrobial resistance

## 1.1.1 The origins of antibiotics and the problem of antimicrobial resistance

Many infections easily treated today were rampant at the turn of the 20th century, leading scientists to look for what they coined a 'magic bullet', that would kill pathogens while leaving the host relatively unharmed [1]. In 1928, Alexander Fleming found a Staphylococcal petri dish that had been contaminated with a fungus that showed bacteriolytic properties [2]. This was the first discovery of a natural product produced by an organism with antibacterial properties. Later work, carried out in the 40s by a team led by Howard Florey and Ernst Chain, resulted in the mass production of the natural product which was named penicillin (**Figure 1.1)** [3,4]. The introduction of penicillin revolutionised modern medicine by not only making it easier to treat infections but also making complex procedures safer by reducing the rate of nosocomial infection.

The two decades following the introduction of penicillin could be considered the 'Golden era' of antibiotic discovery (**Table 1.1**), where all the significant antibiotic classes were identified [5]. Pharmaceutical companies shifted from discovery towards the development of the antibiotics already available because the rate of rediscovery was too high to make it economically viable for further discovery initiatives. As a result, many generations of some antibiotics have been developed, and very few novel antibiotics have made it to the market since the 1980s (**Table 1.1**). During this time, some infections became more challenging to treat, requiring higher doses of antibiotics or alternative drugs entirely. Bacteria producing enzymes with the ability to cleave the β-lactam ring in penicillin were identified as one example reducing the biological activity of an antibiotic [2,6,7]. In other cases, some bacteria were innately resistance to antibiotics or acquired resistance through beneficial mutations or horizontal gene transfer.

With, approximately 1,000 tonnes of antibiotics used in the UK in 2013, antibiotic use is increasing globally [8,9]. Use of antibiotics in farms and within the community has been demonstrated to produce large amounts of antibiotic pollution which can contaminate the ecosystem and exacerbates antimicrobial resistance (AMR) by selecting for organisms with beneficial AMR genes [10–13]. As demonstrated in **Figure 1.2**, if these organisms get back into the human population, in some instances they can prove challenging to treat, increasing the rate of mortality, where AMR now accounts for 700,000 premature deaths globally each year [14].

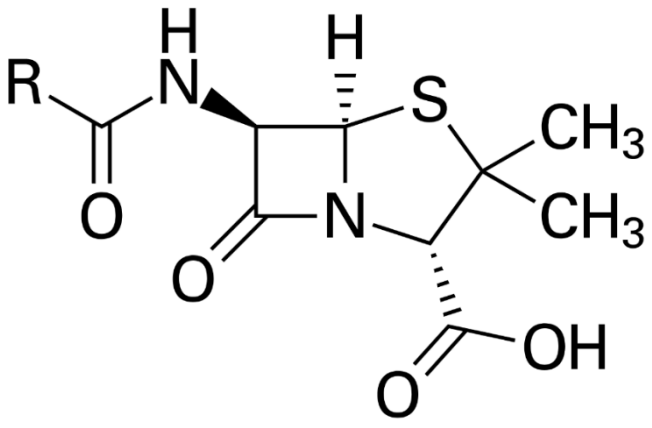



**Figure 1.1** The core structure of Penicillin derived from *Penicillium spp*. The characteristic β-lactam ring is shown next to the variable "R" group.

**Table 1.1** A list of clinically available antibiotics and when they were introduced. The first instance where antibiotic resistance was identified is also shown in the third column. Adapted from Walsh and Wencewicz (2013).

| Antibiotic | Year introduced | Resistance observed |
|---|---|---|
| Sulfonamides[a] | 1930s | 1940s |
| Penicillin | 1943 | 1946 |
| Streptomycin | 1943 | 1959 |
| Chloramphenicol | 1947 | 1959 |
| Tetracycline | 1948 | 1953 |
| Erythromycin | 1952 | 1988 |
| Vancomycin | 1956 | 1988 |
| Methicillin[b] | 1960 | 1961 |
| Ampicillin | 1961 | 1973 |
| Cephalosporins | 1960s | Late 1960s |
| Nalidixic acid[a] | 1962 | 1962 |
| Teicoplanin | 1980s | 1988 |
| Fluoroquinolones[a] | 1980s | 1980s |
| Linezolid[a] | 1999 | 1999 |
| Daptomycin | 2003 | 2003 |
| Retapumulin[b] | 2007 | 2007 |
| Fidaxomicin | 2011 | 2011 |
| Telavancin[b,c] | 2011 | ? |
| Bedaquiline | 2013 | 2009 |
| Dalbavancin[b,c] | 2015 | ? |
| Oritavancin[b,c] | 2015 | ? |

[a] Synthetic antibiotics
[b]Semi-synthetic antibiotics
[c]The semi-synthetic glycopeptide antibiotics currently have no clinical-resistance data available

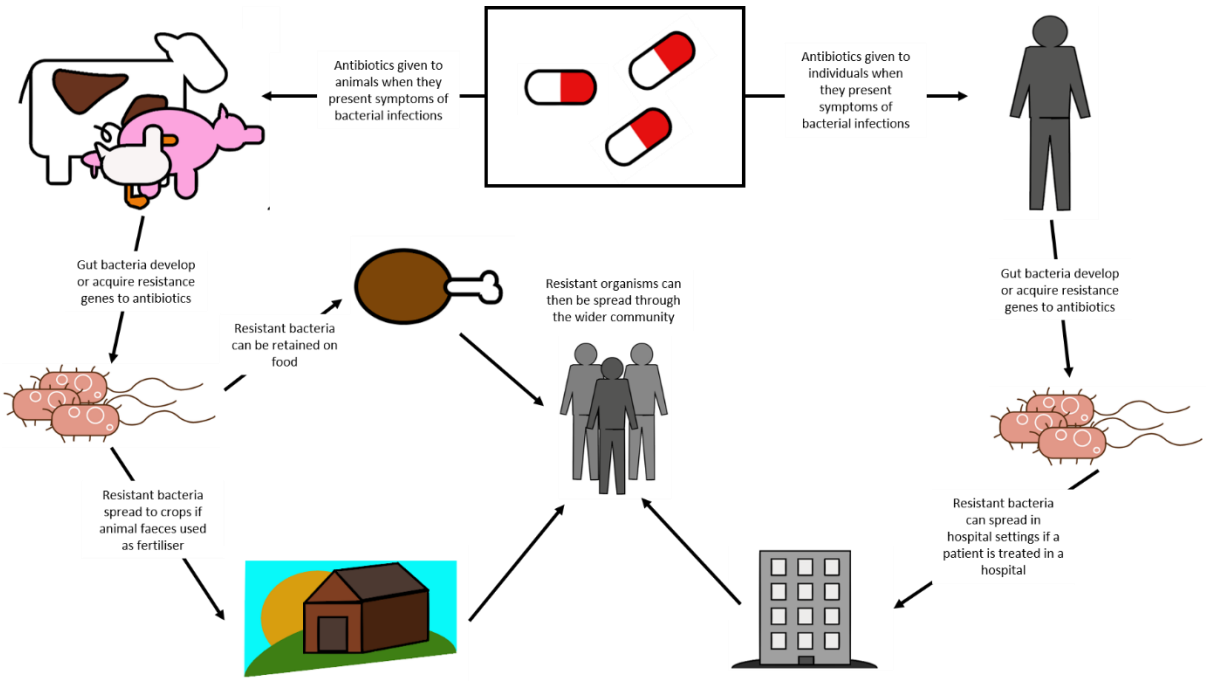



**Figure 1.2** Image shows how AMR is believed to be selected for and introduced back into human populations through the use of antibiotics in both hospital and agricultural settings.

# 1.2 Antibiotics and their targets

## 1.2.1 General mechanisms of antibiotics

Effective antibiotics target unique bacterial components with minimal cross-reactivity with eukaryotic cells. This makes bacterial cells incredibly target poor, with all antibiotics interfering with the normal homeostasis of the cell. For example, penicillin belongs to a group of antibiotics called the β-lactam antibiotics which target penicillin-binding proteins (PBP). These enzymes are involved in maintaining the integrity of the unique peptidoglycan (PG) layer surrounding bacterial cells. By inhibiting these enzymes, the cell struggles to maintain the integrity of the PG, increasing susceptibility to rupturing due to the high osmotic pressure inside of the cell. This is just one example of how antibiotics can affect the homeostasis of a bacterial cell, but all antibiotics within a class of antibiotic broadly share the same cellular target. There are six general targets within bacterial cells. This includes antibiotics that target the steps involved in PG biosynthesis, which include β-lactams, lipopeptides and glycopeptide antibiotics. Others target either DNA replication, transcription or translation slowing the rate of growth of a cell. Antibiotics such as the sulpha drugs target the bacterial specific folate metabolic pathway. Finally, antimicrobial peptides (AMPs) and antibiotics such as daptomycin interact with the negatively charged bacterial membrane surface and associate to form membrane pores, resulting in membrane depolarisation as ions leak out of the cell [5,16]. A more detailed list of the classes and targets of common antibiotics can be found in **Table 1.2**.

## 1.2.2 The bacterial cell envelope and its importance for survival

As previously mentioned, bacterial cells are surrounded by a protective layer of PG which is unique to bacteria. Its function is to provide the cell shape and counteract the high intracellular pressure of several atmospheres [17,18]. PG and all the components embedded within this structure make up the bacterial cell wall, which along with the cell membrane and outer membrane (only Gram-negative bacteria), make up the cell envelope. The cell envelope acts as a barrier to the hostile and unpredictable exogenous environment, protecting internal cellular components. Due to its constant interaction with the external environment, it has become the principal target of many different antibiotics, which interfere with one or multiple components involved in the normal function of this structure. This structure is still an attractive target for drug development because many of the enzymes and intermediates involved in its maturation do not have eukaryotic counterparts [19,20].

**Table 1.2** Antibiotic classes and their cellular targets showing the specific mechanism of action.

| Antibiotic Class | Cellular Target | Mechanism | Reference |
|---|---|---|---|
| β-lactam | Irreversibly bind to penicillin binding proteins (PBPs) | Inhibits cell wall biosynthesis | [7] |
| Sulfonamides | dihydropteroate synthase | Inhibits folate synthesis | [21] |
| Aminoglycosides | Disturbs peptide elongation at the ribosome | Interferes with protein synthesis | [22] |
| Tetracyclines | Block loading of aminoacyl-tRNA in the ribosome | Interferes with protein synthesis | [23] |
| Chloramphenicol | Inhibits protein chain elongation | Interferes with protein synthesis | [24] |
| Macrolides | Prevents peptidyltransferase from adding to the growing peptide chain | Interferes with protein synthesis | [25] |
| Glycopeptides | Bind to lipid II in the bacterial membrane | Inhibits cell wall biosynthesis | [26] |
| Ansamycins | Binds to RNA polymerase | Prevents DNA-dependant RNA synthesis | [27] |
| Quinolones | Covalently bind to DNA-gyrase | Cause double-strand breaks in DNA during DNA replication | [28] |
| Streptogramin | Bind the P-site of the ribosome | Interferes with protein synthesis | [29] |
| Oxazolidinones | Bind the P-site of the ribosome | Interferes with protein synthesis | [30] |
| Lipopeptides | Interact with the cell membrane | Disrupts membrane structure | [31] |

There are substantial differences in the cell envelope organisation of both Gram-positive and Gram-negative bacteria (**Figure 1.3)** even though the structure has a similar role in both. It serves to allow the selective passage of nutrients and expulsion of waste; acting as a site of biochemical reactions and segregating specific molecules into different subcellular compartments [32]. Both Gram-positive and Gram-negative bacteria contain a cell membrane, consisting of a lipid bilayer studded with different proteins, lipoproteins, glycoproteins and lipid-linked polymers. This membrane is then surrounded by the cell wall, which has a layer of PG. The thickness of this layer is only a few nanometres (several layers) in Gram-negative bacteria but can be anywhere between 30 and 100 nm in Gram-positive bacteria with both membrane-anchored and PG-anchored teichoic acids. Gram-negative bacteria also have a secondary membrane called the outer membrane (OM) which surrounds the inner membrane and peptidoglycan layer. The inner leaflet of the OM membrane consists of a phospholipid monolayer and a glycolipid outer leaflet mostly consisting of lipopolysaccharide (LPS). This secondary membrane provides greater resistance to a broad range of harmful compounds, including some bulkier antimicrobial compounds such as the glycopeptide antibiotics [32].

# 1.3 The bacterial cell envelope

## 1.3.1 The PG layer and the processes that are involved in its biosynthesis

### 1.3.1.1 Cytoplasmic steps in PG biosynthesis

In the following sections, there will be more emphasis put into the Gram-positive cell envelope as this project focuses on the interactions of the *Streptomyces coelicolor* cell envelope with different antibiotics, some of which are specific for Gram-positive bacteria. However, much of the work that elucidated the processes that synthesise PG and the cell envelope have been studied primarily in *Escherichia coli*. The need for these processes for bacterial survival means that they are fairly well conserved across bacteria, and some generalisations can be applied to most bacterial species and where gaps in our current knowledge exist, predictions can be applied from knowledge of model systems such as *E. coli*.

The major component of the cell envelope in Gram-positive bacteria is the thick PG layer which is made up of an N-acetylglucoseamine (GlcNAc)-N-acetylmuramic acid (murNAc) disaccharide.

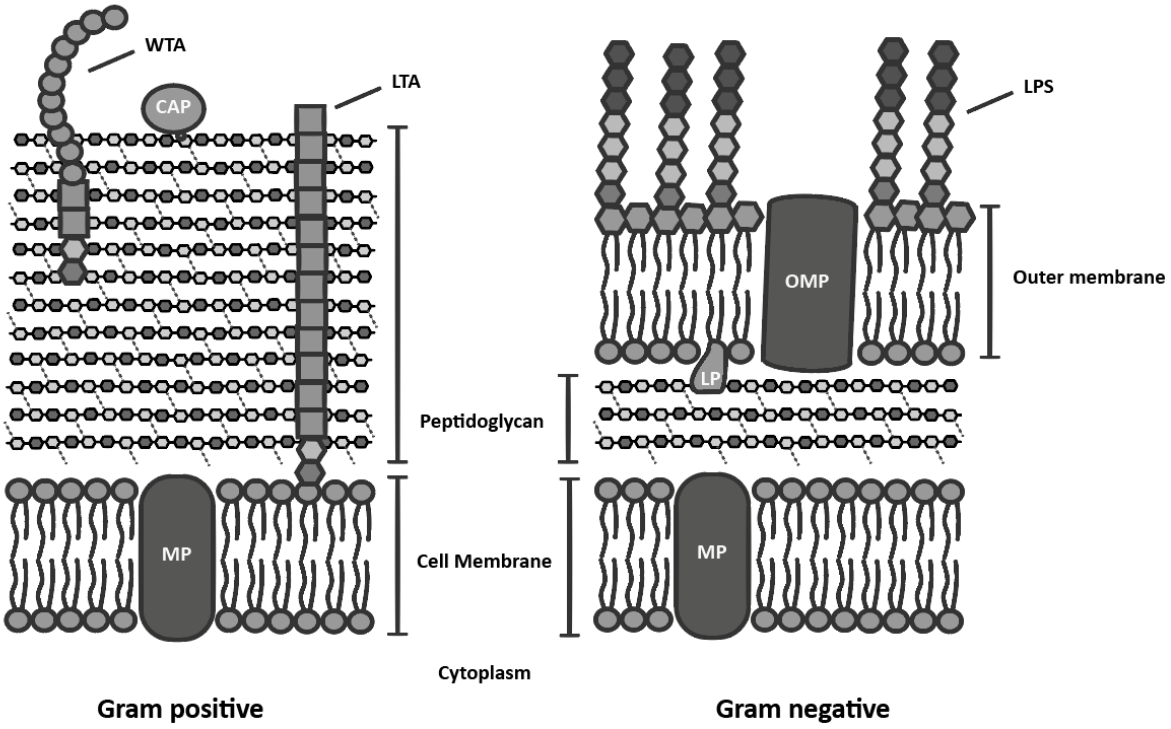



**Figure 1.3** Diagrams showing comparisons of the Gram-positive and Gram-negative cell envelopes. Abbreviations: IMP, inner membrane protein; LTA, lipoteichoic acid; WTA, wall teichoic acid; CAP, covalently attached protein; LP, lipoprotein; OMP, outer membrane protein; LPS, lipopolysaccharide.

The sugars are covalently linked by β1-4 glycosidic bonds to make long-chain polymers with a pentapeptide stem attached to each murNAc. Interglycan cross-linking occurs through the transpeptidation of the pentapeptide chain between neighbouring glycans, strengthening the structure [33]. PG biosynthesis has been the focus of many studies over the past 50 years, and the following sections will summarise what is currently known about this process. The cytoplasmic steps of PG biosynthesis first produce the sugar-peptide moiety in PG (**Figure 1.4**). The first step in this process converts fructose-6-phosphate to glucosamine-6-P through the action of GlmS. The second reaction, catalysed by GlmM, produces glucosamine-1-P. The same bifunctional enzyme, GlmU, catalyses the third and fourth reactions in this pathway, producing N-acetylglucosamine-1-P and then uridine diphosphate (UDP)-GlcNAc. MurA and MurB then convert UDP-GlcNAc to UDP-murNAc, and the MurCDEF ligases are then responsible for sequentially adding amino acid residues to UDP-MurNAc, catalysing the formation of the stem peptide. Typically, the peptide starts with a D-Alanine (D-Ala) which is ligated to a D-glutamic (D-Glu) acid by MurD. Gram-positive bacteria generally follow this with L-lysine which is added by MurE. Finally, a D-alanyl-D-alanine (D-Ala-D-Ala) motif caps the end of the pentapeptide; however, alternative terminal residues such as D-alanyl-D-lactate (D-Ala-D-Lac) are seen under certain circumstances [34].

#### 1.3.1.2 Membrane-bound steps of peptidoglycan biosynthesis

The UDP-MurNAc-pentapeptide is then transferred to a membrane-bound $C_{55}$-lipid carrier, undecaprenyl phosphate (UP). This is a vital lipid carrier not just for PG but for several lipid-bound structures in bacteria, such as teichoic acids and LPS in Gram-negative bacteria [18]. It is an essential molecule that is synthesised from the isoprenoid, isopentenyl diphosphate (IPP). IPP is first converted to dimethylallyl diphosphate (DMAPP) where it is condensed with IPP to produce farnesyl diphosphate (FPP). FPP acts as the building block for most isoprenoids found in bacteria, including UP (**Figure 1.5**). UppS then generates undecaprenyl pyrophosphate (UPP) from FPP within the membrane. UPP has to be dephosphorylated before it can be used for PG biosynthesis, and this is mostly carried out by the UPP-phosphatase (UppP) [18,35]. The UDP-MurNAc-pentapeptide is then transferred to a UP molecule by MurX, releasing uridine monophosphate, generating lipid I. A GlcNAc moiety is then transferred to Lipid I by the glycosyltransferase MurG on the C4 hydroxyl group to produce the $C_{55}$-P-MurNAc-(pentapeptide)-GlcNAc, also referred to as Lipid II (**Figure 1.6**).

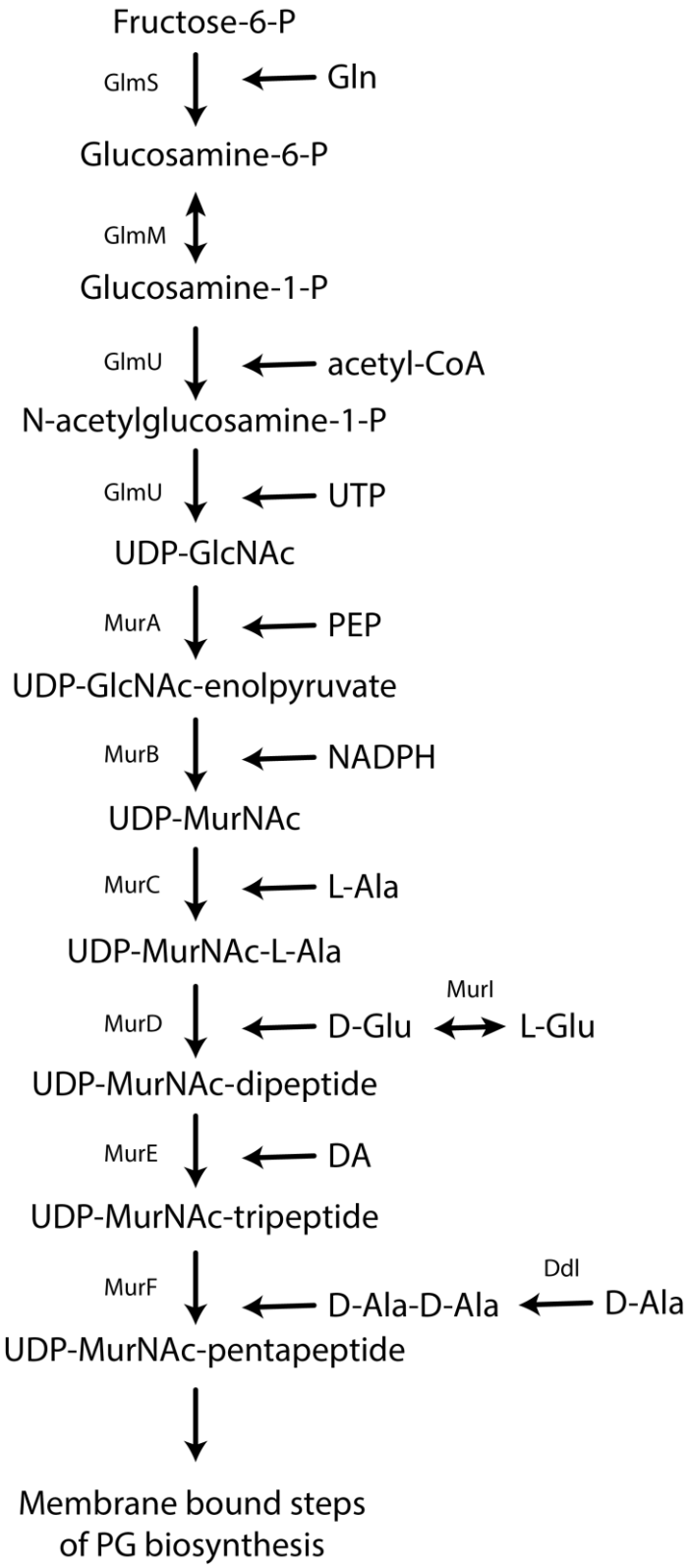



**Figure 1.4** Cytoplasmic steps of peptidoglycan biosynthesis showing the enzymes involved.

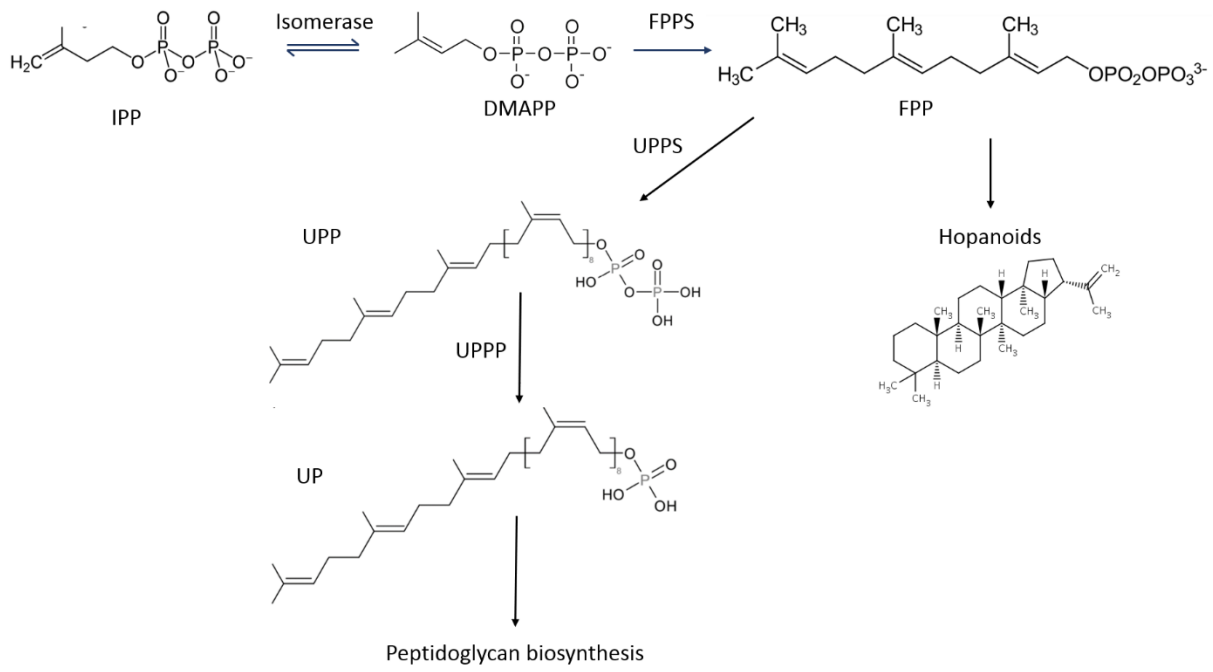



**Figure 1.5** Biosynthetic steps for making the isoprenoid lipids, undecaprenyl phosphate (UP) and hopanoids from farnesyl diphosphate (FPP). Isopentenyl diphosphate (IPP), dimethylallyl diphosphate (DMAPP), farnesyl diphosphate synthase (FPPS), undecaprenyl pyrophosphate synthase (UPPS) and undecaprenyl pyrophosphate phosphatase (UPPP).

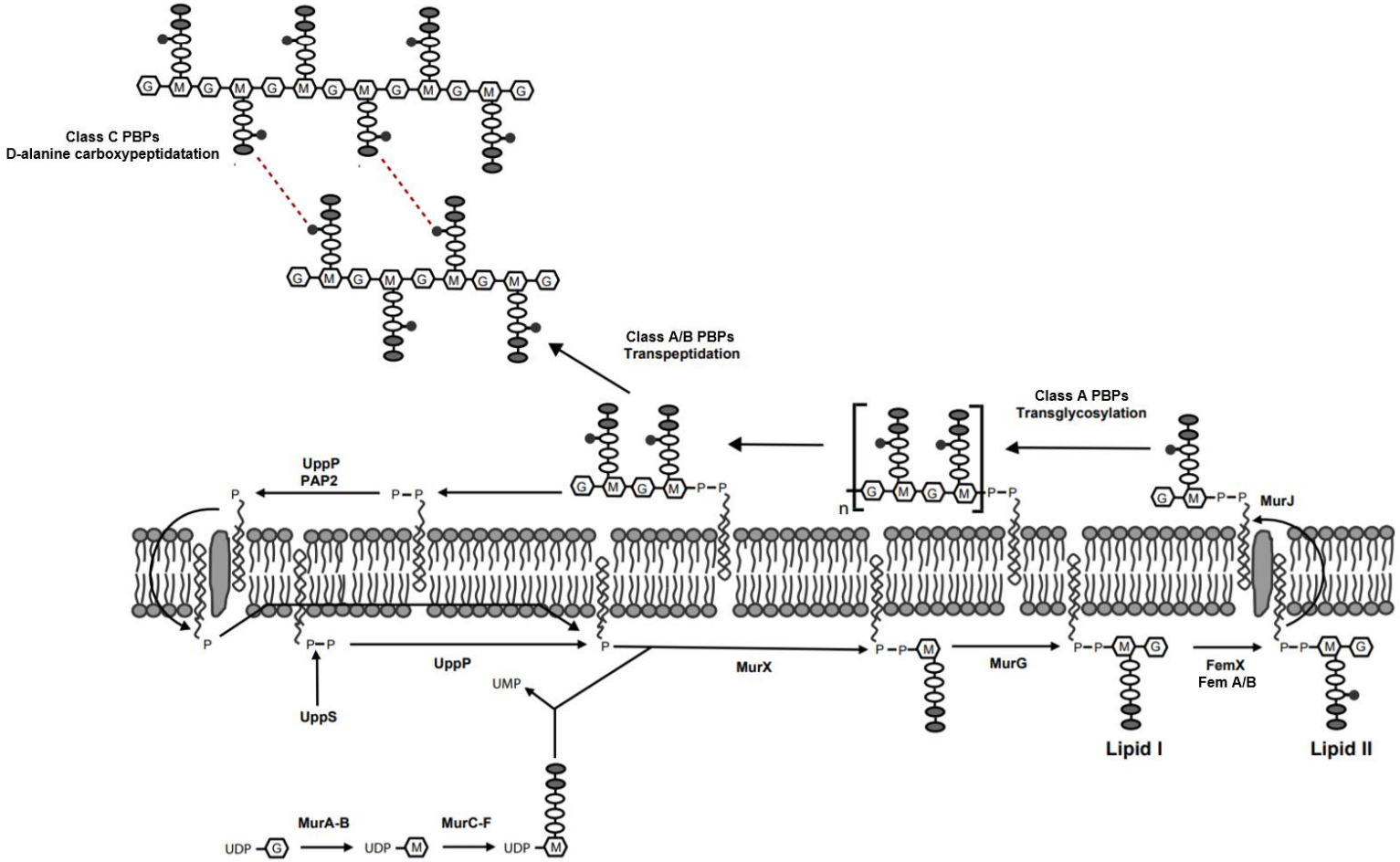



**Figure 1.6** Cytoplasmic, inner membrane and outer membrane reactions of PG biosynthesis in bacteria. Enzymes for each reaction are highlighted above/below arrows.

Most bacteria also contain additional interglycan cross-links between pentapeptide chains of neighbouring glycans (**Figure 1.7**). These cross-links are made between residues that branch off the stem pentapeptide and vary in structure between species. In *S. coelicolor*, only one glycine is present in the cross-bridge; but, in *Staphylococcus aureus*, there are five (**Figure 1.8**). The first residue is added by the non-ribosomal peptidyltransferase, FemX (**Figure 1.6**), which recognises the terminal D-Ala-D-Ala motif [36]. Two other proteins FemA and FemB are required for the additional glycine residues in *S. aureus* but are absent in species like *S. coelicolor* [37].

The next step in this process is the transfer of lipid II from the inner leaflet to the outer leaflet of the lipid bilayer. As lipid II is a highly polar molecule, the rate that this would occur spontaneously would be too slow to be biologically useful, so it has long been hypothesised that there must be a transporter or flippase that fulfils this role. Although several candidates have been suggested, MurJ has the most substantial evidence to support its role in this cellular function [38]. MurJ belongs to a superfamily of multidrug/oligosaccharidyl-lipid/polysaccharide (MOP) exporters, and members of this protein family are involved in the excretion of LPS. As both PG and LPS use UP as a lipid carrier for their intermediates, it seems logical that MOP transporters could be involved in both processes. However, MurJ has also been found to be redundant in *Bacillus subtilis*, suggesting that several proteins may have lipid II flippase activity [39–41].

#### 1.3.1.3 Late stages of peptidoglycan biosynthesis

Once flipped to the external surface of the membrane, lipid II needs to deliver the PG monomer to the nascent glycan termini (**Figure 1.6**). PBPs are responsible for catalysing these reactions and modifying mature PG to cope with the needs of the cell. These enzymes can be grouped broadly into three classes which each have different roles in the cell wall. Class A PBPs possess an N-terminal glycosyltransferase which catalyses the formation of glycosidic bonds between the disaccharide and the PG polymer. They also contain a C-terminal transpeptidase domain which can cross-link the amino acid branches between neighbouring glycans. Class B PBPs only have a C-terminal transpeptidase domain and when studied in *Escherichia coli*, are typically involved in cell division and PG maturation/recycling. The role of class C PBPs is less clear, but they contain a DD-carboxypeptidase and can remove the terminal D-ala from a pentapeptide to prevent any further enzymatic modification of the peptide cross-bridge [17].

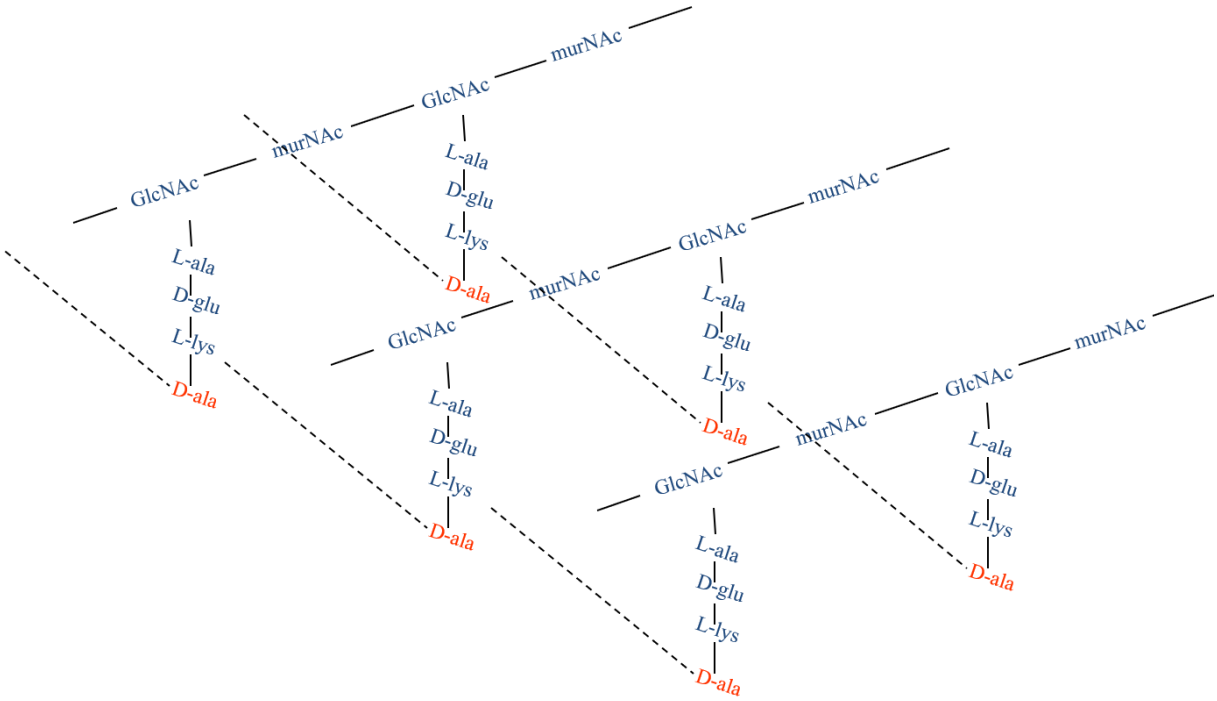



**Figure 1.7** Diagram showing an example of the lattice structure of crosslinked peptidoglycan in bacterial cells. Interpeptide cross-links are displayed as dashed lines.

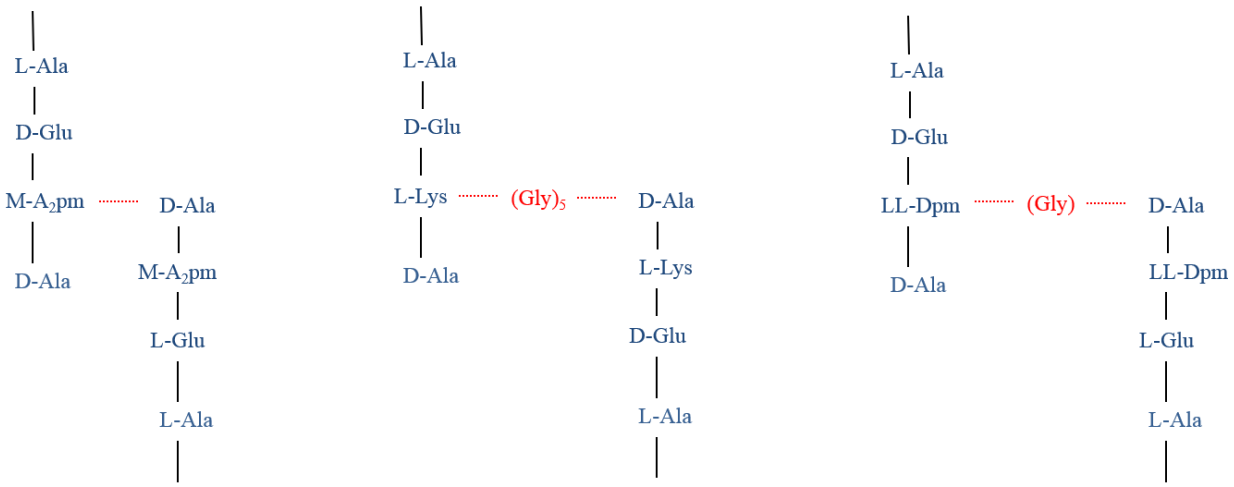



**Figure 1.8** Diagrams of the interpeptide bridges in the stem peptides of three different bacteria. *E. coli* forms bridges directly between its stem peptides, whereas both *S. aureus* and the *Streptomyces spp.* bridge their stem peptides with differing numbers of glycine residues. Adapted from Vollmer *et al*., 2008.

So far, all bacteria with PG have been discovered to have at least one of each class although the number may vary. *E. coli* has 12 PBP while the sporulating bacteria *Bacillus subtilis* and *S. coelicolor* have 16 and 21 respectively. It has been proposed that the higher number of PBPs in the latter organisms may help with the more complicated developmental processes that occur during sporulation [17].

The final stage of PG biosynthesis involves the dephosphorylation of UPP to reconstitute UP for further rounds of PG cycling. Although the flippase that transports the $C_{55}$-carrier lipid back to the inner leaflet of the membrane remains elusive, there is strong evidence to suggest that the BacA membrane protein is the main UPP phosphatase in *E. coli*. Although not essential, BacA has been shown to dephosphorylate UPP in vitro and when deleted, causes hypersensitivity to the UPP targeting antibiotic bacitracin, [42,43]. Other studies have demonstrated that other proteins are also able to fulfil some of the UPP-phosphatase activity in the membrane, including Yeiw and PgpB from *E. coli* and BcrC from *B. subtilis* [44,45]. All belong to the type 2 phosphatidic acid phosphatase (PAP2) superfamily which includes various other membrane phosphatases. It is likely that because of the importance of maintaining the homeostasis of the cell envelope, membrane phosphatases exhibit multi-functionality to ensure that cells can still function in case one enzyme fails [46].

## 1.3.2 The bacterial cell membrane

### 1.3.2.1 Architecture and composition of bacterial cell membranes

This section will discuss the current understanding of the structure and function of the cellular membranes and how they relate to the function of the bacterial cell envelope. This mostly lamellar phospholipid bilayer confers spatial separation between the intra- and extracellular environments [47,48] and early models proposed that this structure existed as a lipid matrix with globular proteins embedded into the structure. [49]. Based on this model, components can diffuse freely in the plane of the membrane while smaller, less-polar molecules can spontaneously flip between the leaflets of the membrane (**Figure 1.9**). Flippases such as MurJ, are required to flip bulkier molecules across the membrane creating a heterogeneous composition [50–52], but there is now compelling evidence to demonstrate that bacteria augment the composition of their membranes to adapt to their environments [53–55]. In doing so, they maintain the integrity of the cell envelope [56], resist against high osmolarity [57]; increase thermotolerance [58,59]; and to adapt to different compounds such as antibiotics [60].

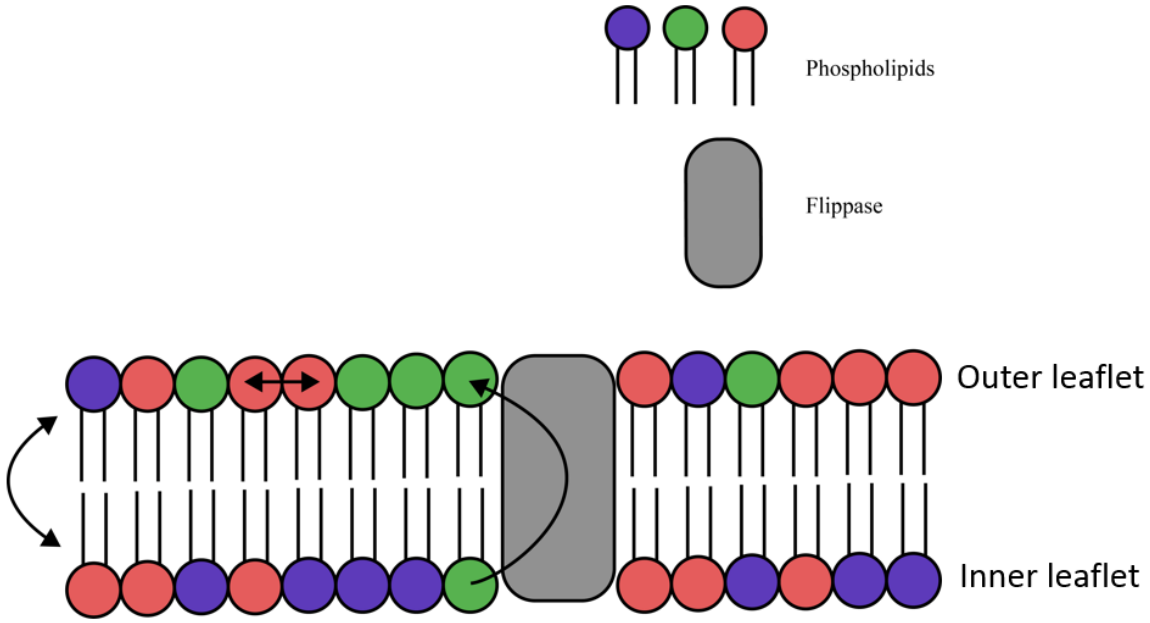



**Figure 1.9** Diagram depicting the movement of lipids in biological membranes. Lipids can diffuse through the plane of both the inner and outer leaflets of the lipid bilayer, and some smaller lipids can spontaneously flip-flop between the leaflets. For larger, more polar phospholipids, flip-flopping is facilitated by membrane flippase.

$C_{55}$-P lipids have already been discussed concerning PG biosynthesis. However, the primary lipids in bacterial membranes are the amphiphilic phospholipids. They consist of an acylated glycerol-3-phosphate (G3P) molecule that may be loaded with several different polar head groups. The composition of these structures can vary between species [61], but common phospholipids include phosphatidic acid (PA), phosphatidylglycerol (PpG), cardiolipin (CL), phosphatidylinositol mannosides (PIM) and phosphatidylethanolamine (PE) (**Figure 1.10**). To understand how the bacterial membrane protects the broad range of environments, the metabolic processes that generate these phospholipids must be understood. Like PG, the initial stages in the biosynthesis of each phospholipid occur in the cytoplasm and are then loaded into the membrane.

The mechanisms that go into creating phospholipids are relatively conserved amongst bacteria and start with the generation of G3P through the phosphorylation of glycerol by GlpK [62]. Simultaneously, fatty acids are produced through the fatty acid biosynthetic pathway. This pathway begins with the conversion of acetyl-CoA to malonyl-CoA by acetyl-CoA carboxylase (ACC). Chain elongation requires the action of four enzymes working cyclically to increase the length of the nascent acyl chain. The malonyl group is transferred to the acyl carrier protein (ACP) by FabD, and the intermediate is condensed with another acetyl-CoA by FabH to form β-ketoacyl-ACP. Branched acyl chains can also be incorporated into this metabolic pathway as is typically the case in Gram-positive bacteria, and bacteria that are undergoing temperature-related stress. FabA then dehydrates the β-ketoacyl-ACP to trans-2-enoyl-ACP, which is finally reduced by FabI to produce an acyl-ACP (**Figure 1.11**).

The membrane phospholipid intermediate, lysophosphatidic acid (LPA), is generated at the membrane surface from the acylation of a G3P by an acyl-transferase such as PlsX/PlsY or sometimes PlsC [58]. A second acyl-chain is then added by PlsB, producing PA. Its small size and neutral charge allow it to spontaneously flip-flop across the membrane, serving as a lipid carrier on either side of the membrane (**Figure 1.11**). PA is then condensed with a cytidine diphosphate to produce cytidine diphosphate-diacylglycerol (CDP-DAG)[61,63]. Bacterial membranes differ in their phospholipid composition, but PE and CL are two of the more common lipids [64]. PE is synthesised from CDP-DAG through the action of two enzymes: phosphatidylserine synthase (Pss) and phosphatidylethanolamine decarboxylase (Psd) which first generate phosphatidylserine (PS) from CDP-DAG followed by and PE (**Figure 1.11**).

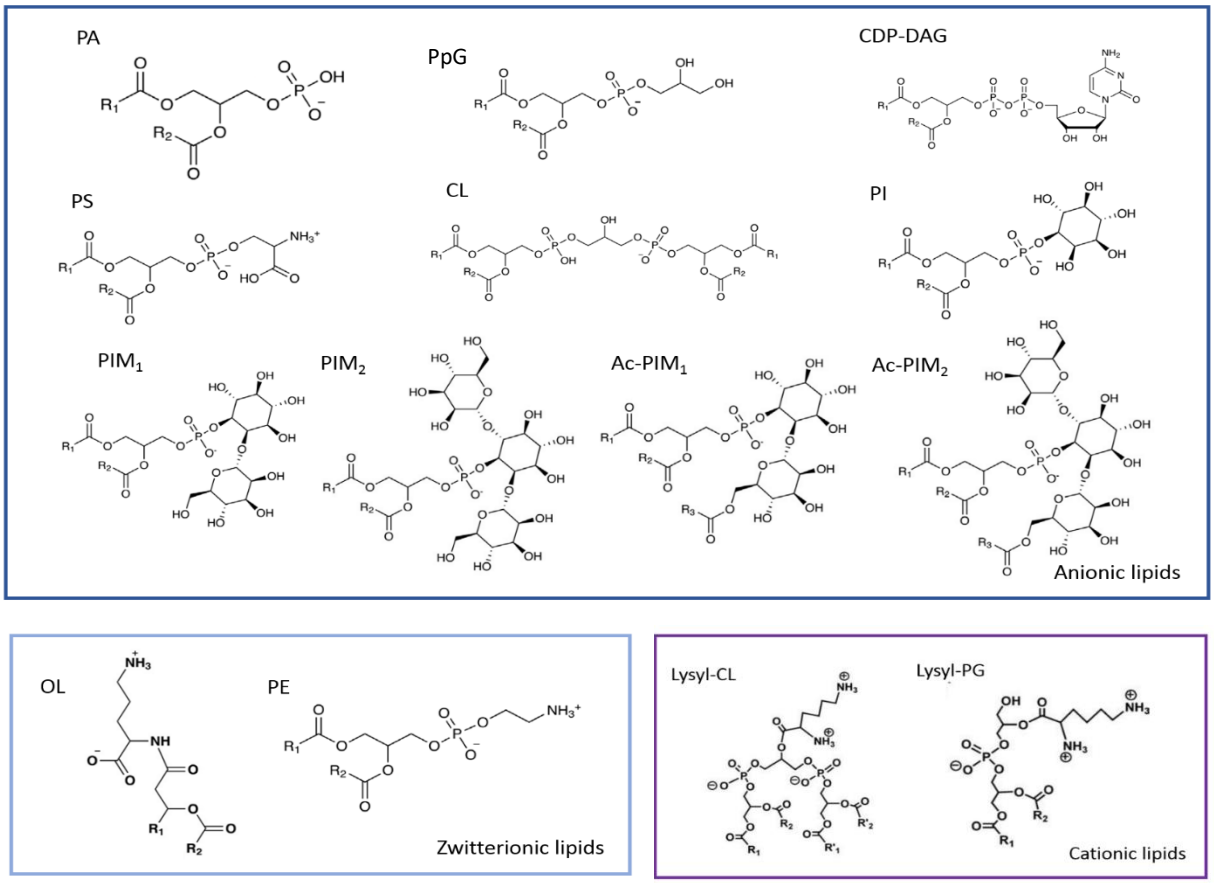



**Figure 1.10** Structure of the main bacterial phospholipids as well as the phosphate-free, ornithine lipid. Each lipid has been grouped based on their overall charge, showing that most bacterial lipids are anionic. Acyl-chains are represented as R-groups.

Alternatively, a G3P can be added to CDP-DAG by a PpG phosphate synthase (PgsA) to form PGP. PGP can then be dephosphorylated by a PpG phosphate phosphatase (PgpA, PgpB or PgpC in *E. coli*) to produce PG. CL is then synthesised from the condensation of a PG and a CDP-DAG via a CL synthase (**Figure 1.11**) [61]. Further modifications of PGP with lysine or alanine have been recognised in some bacteria, providing resistance to cyclic antimicrobial peptides (CAMPs), osmotic stress and bile salts [53,54,65,66].

Other phospholipids seen in bacteria include phosphatidylinositol (PI). PI is more commonly found in eukaryotes, but PI is also widespread throughout the Actinobacteria and δ-proteobacteria. PI is essential for the synthesis of lipoglycans in mycobacteria, but its role in other bacteria has yet to be elucidated [60,67]. Myoinositol-1-phosphate is first added to CDP-DAG by PisA to produce PI-phosphate (PIP). The mechanism by which this molecule is dephosphorylated is unknown, but a membrane phosphatase likely catalyses the reaction [61,67]. In the *Streptomyces spp*., PimA adds a mannose to position-2 of PI to produce PI monomannoside ($PIM_1$). $PIM_1$ can then either have an additional mannose added to its structure by PimB' to produce $PIM_2$, or undergo further acylation, producing acyl-$PIM_1$ (Ac-$PIM_1$). Ac-$PIM_2$ is then generated from either of these intermediates using the converse enzyme (**Figure 1.11**) [67].

### 1.3.2.2 The enzymes involved in recycling phospholipids

Although we have a good understanding of the enzymes and biochemical reactions involved in the recycling of the $C_{55}$-P in peptidoglycan biosynthesis, there is still a poor understanding of the processes that are involved in the degradation and recycling of phospholipids. The PAP family are known to be one group of enzymes involved. These enzymes can be grouped based on whether they are $Mg^{2+}$-dependant (PAP1) or $Mg^{2+}$-independent (PAP2), with the former only found in eukaryotes and the latter predominant in bacteria. PAP2 enzymes show a broader specificity for their targets. The phosphatidic acid phosphatase, PgpB from *E. coli* is known to hydrolyse many pyrophosphate phospholipid derivatives other than phosphatidic acid, including including the PG intermediate, UPP [68]. This also includes BcrC from *Bacillus sp*., which expresses UPP-phosphatase activity in the bacterial membrane increasing resistance to the antibiotic bacitracin by competing for its target, UPP [45,69]. The widespread distribution of predicted membrane phosphatases suggests that these enzymes could also have a diverse and essential role in modifying and recycling phospholipids, although their role in these processes is still poorly understood.

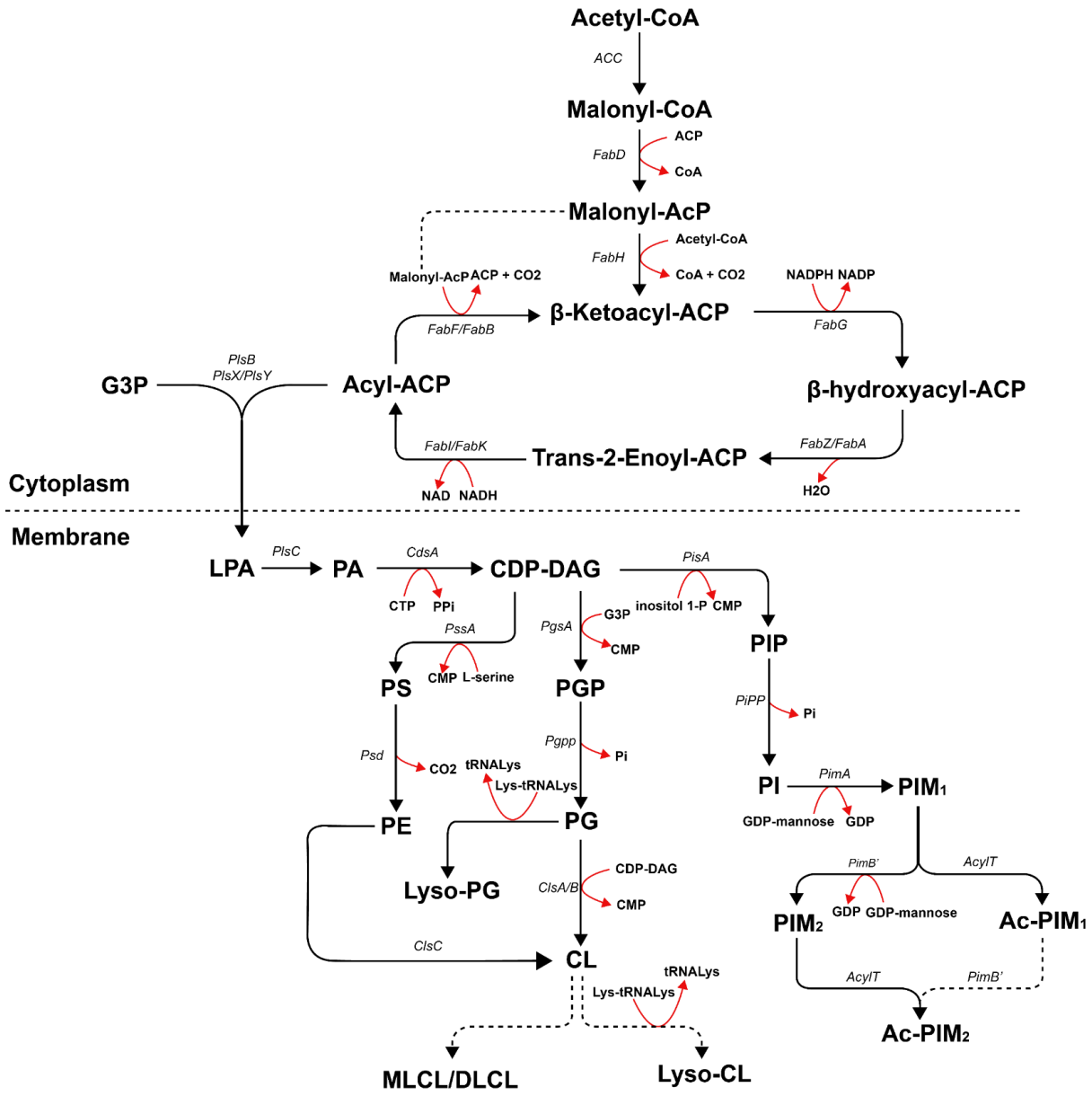



**Figure 1.11** Cytoplasmic and membrane reactions that produce membrane phospholipids that have so far been identified in bacteria. Solid arrows indicate reactions that have been experimentally observed, while dotted arrows have yet to have their mechanisms elucidated.

#### 1.3.2.3 Other important bacterial membrane lipids

Some bacteria have also been found to produce phosphate-free lipids such as ornithine lipid (OL). They have been well characterised in the α-, β-, γ-proteobacteria as well as in the actinomycetes. Recent work on *S. coelicolor* has shown that OL is produced during later stages of development and during phosphate-limited conditions [60]. The pathway is well conserved in the *Streptomyces* and is carried out by two enzymes that acylate one molecule of ornithine. OlsB carries out the first acylation step, producing lyso-OL (LOL). OlsB then carries out a second acylation step to produce OL (**Figure 1.12**).

Unlike eukaryotic membranes, bacterial membranes lack the isoprenoid lipid, cholesterol. Instead, they produce hopanoids which carry out a similar function in bacteria (**Figure 1.5**). These molecules can interact with the acyl chains of phospholipids, reducing their fluidity in the plane of the bilayer. In some bacteria, such as *B. subtilis* and *S. coelicolor*, they become abundant during development, particularly at the point of sporulation [48,70]. More recently these molecules have been found to coalesce with flotillin proteins to form microdomains within bacterial membranes. Although not well understood in bacterial membranes, flotillins can recruit many different proteins to the site of the microdomain, possibly localising particular functions to specific locations within the membrane [48,71].

# 1.4 The glycopeptide antibiotics

## 1.4.1 Glycopeptide antibiotics and their structure

Glycopeptide antibiotics constitute roughly 0.3% of all antibiotics used in the UK yet, they have become instrumental in the second-line treatment of multidrug-resistant, Gram-positive bacteria [9]. All known naturally derived glycopeptide antibiotics originated from the soil-dwelling Gram-positive actinomycetes. Vancomycin (**Figure 1.13A**) was the first glycopeptide to be introduced and was first purified from *Amycolatopsis orientalis* (previously *Streptomyces orientalis*) in 1953 [72]. After its introduction in 1958, it became the drug of choice for penicillin-resistant staphylococcal infections but fell out of favour due to several of its undesirable side effects including; ototoxicity, nephrotoxicity and a characteristic upper-body rash referred to as 'red-man syndrome' [73]. Because they were generally safer and easier to administer, newer generations of β-lactam antibiotics became the preferred line of treatment for Gram-positive infections, and vancomycin fell out of favour. A second glycopeptide antibiotic, teicoplanin (**Figure 1.13B**), was later licensed in the EU in 1988 [74].

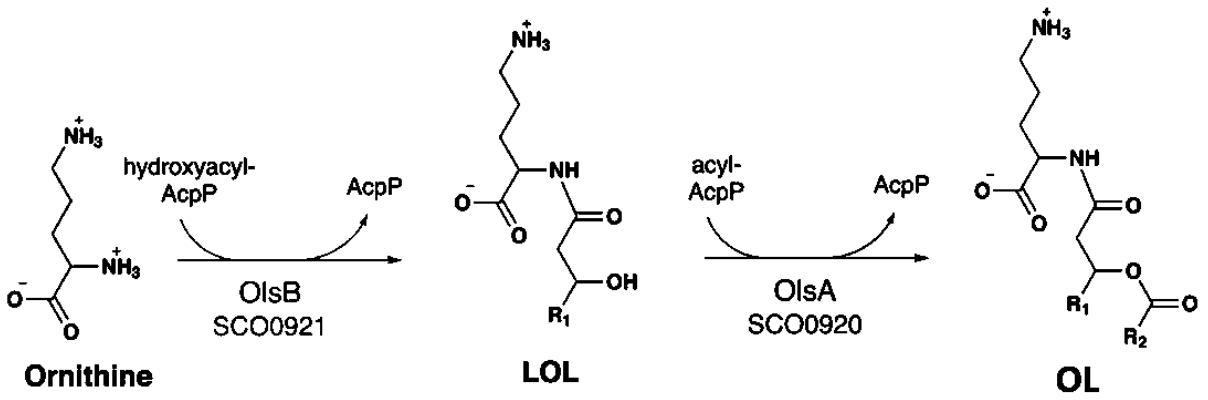



**Figure 1.12** The biosynthesis of OL in *S. coelicolor* through the action of OlsA and OlsB.

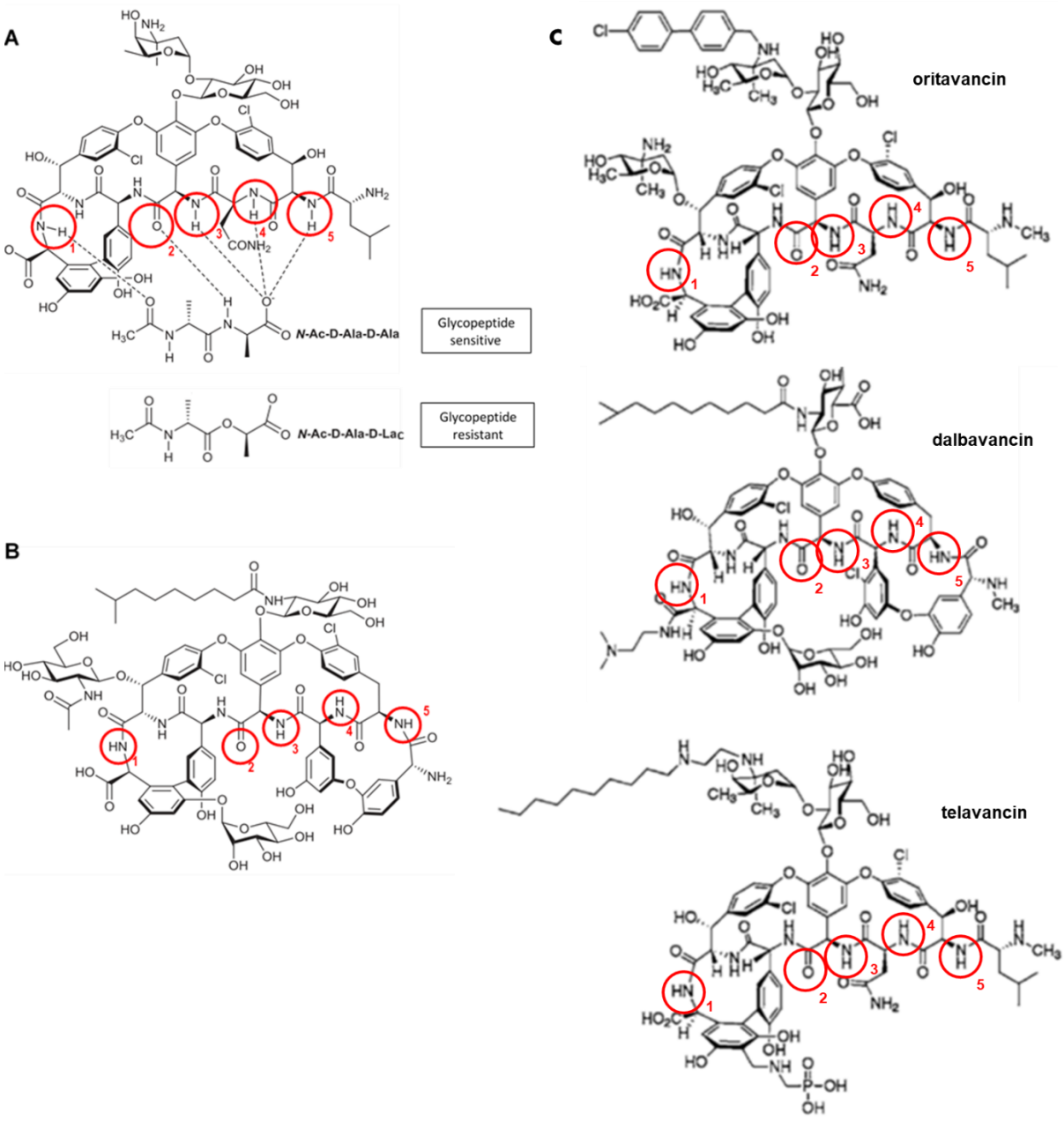



**Figure 1.13** The structures of clinically approved glycopeptide antibiotics. (A) Vancomycin binds to D-Ala-D-Ala with five hydrogen bonds in a glycopeptide sensitive phenotype. The structure of the D-Ala-D-Lac in resistant phenotype is shown underneath. (B) The structure of teicoplanin $A_2$-2. (C) Three semi-synthetic lipoglycopeptide antibiotics that have been approved recently for the use in the EU. Red circles indicate shared hydrogen bonding groups.

It was first purified as a natural product in 1977 from the bacterium *Actinoplanes teichomyceticus* [75]. Like vancomycin, this novel drug was a potent inhibitor of Gram-positive organisms but had a longer half-life within the body while presenting far fewer adverse side effects associated with vancomycin treatments [76,77].

The glycopeptides can be divided into Types I-V depending on the amino acid sequence in their heptapeptide backbone. Depending on the residues, they form either di-, tri- or tetracyclic cores and are synthesised non-ribosomally. Vancomycin belongs to the Type I glycopeptides, which have tricyclic structures with aliphatic residues at positions AA-1 and AA-3 (**Figure 1.13A**), while Types II, III and IV all have aromatic residues in these positions. Type III and IV (which includes teicoplanin) glycopeptides have an additional aryl ester (**Figure 1.13B**) creating tetracyclic structures [78].

The glycopeptide aglycone may also display several structural variations which include different glycosylation patterns, the presence of chlorine atoms to provide rotational stability, and the presence of a fatty-acid chain substituted into a sugar moiety on AA-4 [78]. Teicoplanin is an interesting glycopeptide because it is produced as a mixture of five structurally related compounds ($TA_2$-1 – $TA_2$-5) that all have the same core but have variable alkyl sidechains. It is believed that the improved clinical properties of teicoplanin are attributed to these lipophilic sidechains, and it should be no surprise that many semi-synthetic glycopeptide derivatives have similar modifications. The second-generation of glycopeptides which have been introduced over the past ten years (**Figure 1.11C**) all contain lipophilic side chains and include the chloroeremomycin derivative, oritavancin [79,80]; the vancomycin derivative, telavancin [81,82]; and a derivative of A40926, dalbavancin [83,84].

## 1.4.2 Mechanisms of glycopeptide antibiotics; modes of action and resistance mechanisms

The high molecular weight of glycopeptide antibiotics means that they are unable to pass through the outer membrane of Gram-negative bacteria, making Gram-negatives innately resistant to these drugs. However, they can diffuse through the thick PG layer of Gram-positive bacteria, to reach the membrane surface where they target, the D-Ala-D-Ala motif of lipid II [85]. This interaction is mediated by five hydrogen bonds that are stabilised by many hydrophobic van der Waals contacts [86], allowing them to sterically occlude the enzymes involved in the later stages of PG biosynthesis (**Figure 1.13A**) [87,88].

Glycopeptides are unusual antibiotics in that they target a substrate rather than the enzymes within a biochemical process. This means that the chance of spontaneous resistance to glycopeptides is low as several enzymes would have to develop mutations in unison to produce a modified lipid II. However, reports of glycopeptide resistance were first documented in the 1980s when the first isolates were identified in the enterococci [89,90]. These organisms form part of the healthy microbiota but can become opportunistic pathogens in immunocompromised individuals, leading to urinary tract infections, endocarditis and sepsis. It was found that vancomycin-resistant enterococci (VRE) had acquired mobile elements with gene clusters that allowed them to modify the D-Ala-D-Ala motif in lipid II to produce pentapeptides terminating in D-Ala-D-Lactate. In doing so, the number of hydrogen bonds that could form between lipid II and the glycopeptide was reduced by one, resulting in a 1000-fold decrease in affinity, and significant reduction of biological activity (**Figure 1.13A**) [72]. The health crisis caused by VRE continued during the following decade where nine subtypes of VRE (VanA, -B, -C, -D, -E, -G, -L, -M and -N) were identified based on the genetic organisation of their resistance clusters [72,91]. For a full review of the subtypes of VRE, see Ahmed and Baptise (2017), but the most common types, VanA and -B will be discussed here.

Both VRE VanA and B show high levels of inducible resistance to vancomycin, but only VanA also exhibits high-level teicoplanin resistance. Genetic differences in the resistance clusters are what cause this difference in phenotype. Both have a core set of genes, *vanRSHAX,* that are essential for modifying D-Ala-D-Ala. Additionally, they also both have several accessory genes that are not essential for glycopeptide resistance (**Figure 1.14A**). The enzymatic functions converting D-Ala-D-Ala to D-Ala-D-Lac are encoded for by the genes *vanHAX,* and their induction causes high-level resistance to both vancomycin and teicoplanin. VanH is responsible for converting pyruvate to lactate which is then used by the D-Ala-D-Lac ligase, VanA (denoted VanB in VanB-type). VanX cleaves any remaining D-Ala-D-Ala to prevent any further incorporation into lipid II (**Figure 1.14B**) [91].

*vanRS* encode a two-component system (TCS) which are the primary system of signal transduction in bacteria and are essential in the regulation of transcription in response to a wide variety of environmental conditions. The classic TCS involves a sensory kinase (SK), usually located in the membrane that is responsible for recognising an environmental signal. Once activated, the SK subsequently phosphorylates its cognate response regulator (RR), which then modulates genes expression [92,93].

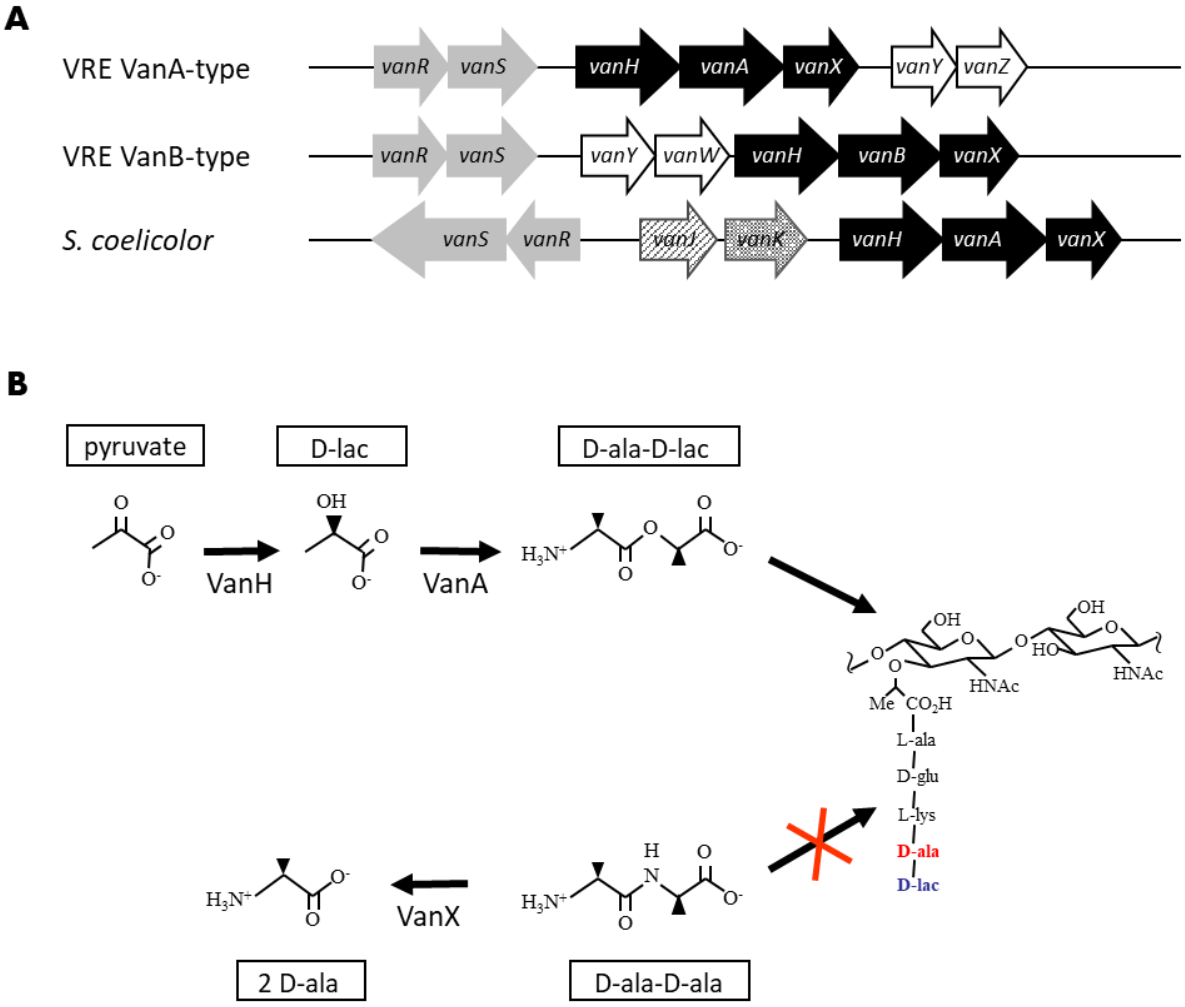



**Figure 1.14** Glycopeptide resistance gene cluster (A) and the proposed function of VanH, VanA and VanX proteins (B).

The TCSs from VanA and B enterococci are distinctly different with evidence suggesting that the ligand for VanA-type enterococci is a teicoplanin-D-Ala-D-Ala complex. It is unclear what the ligand is for in VanB-type but the SK is more truncated than the VanA SK [94]. There is now strong evidence to suggest that the differences observed between the two types is down to the induction of *vanHAX* by the two different SKs and the reduced resistance to teicoplanin in VanB-type enterococci is due to poor induction of the SK [95].

### 1.4.3 Evidence of a dual target for lipoglycopeptide antibiotics

Teicoplanin is empirically more effective than vancomycin even in VanA-type VRE [77,81,83,96–99]. The major structural difference between the two is the lipophilic sidechain of teicoplanin which classifies teicoplanin as a lipoglycopeptide antibiotic. It is known that the structural differences between teicoplanin and vancomycin can affect the induction of *van* genes [100,101], but the role of these structural difference remains controversial. Studies have indicated that vancomycin preferentially blocks the transglycosylation step of PG biosynthesis at the membrane surface [102], but early studies on teicoplanin suggested that it was also able to interfere with membrane integrity, causing depolarisation or lysis, similar to the biological effects seen with the lipopeptide daptomycin [88,103,104].

Challenging evidence has emerged in the past decade where *in vivo* studies using rotational-echo double resonance (REDOR) NMR on live cells have indicated that the lipophilic side chains of oritavancin analogues contact the branch amino acids on the pentapeptide of unlinked PG. This suggests, like vancomycin, lipoglycopeptides can block transglycosylases from polymerising nascent glycan strands nearer the membrane during transglycosylation. But, additionally they can also occlude class II PBPs from accessing unlinked bridge peptides during transpeptidation (**Figure 1.15**). The secondary binding site also means that lipoglycopeptides have a greater affinity for lipid II even in cells expressing the D-Ala-D-Lac phenotype, explaining why these drugs are more active against vancomycin-resistant organisms [88,105,106]. It is also likely that due to the secondary binding site, lipoglycopeptides are more difficult to recognise by certain SKs, leading to reduced induction of the *van* genes by VanRS.

### 1.4.4 Glycopeptide resistance in *S. coelicolor*

Most of the research on glycopeptide resistance has been carried out on VRE, nonetheless, several problems arise when studying clinical isolates. Firstly, they require special handling procedures to ensure their containment.

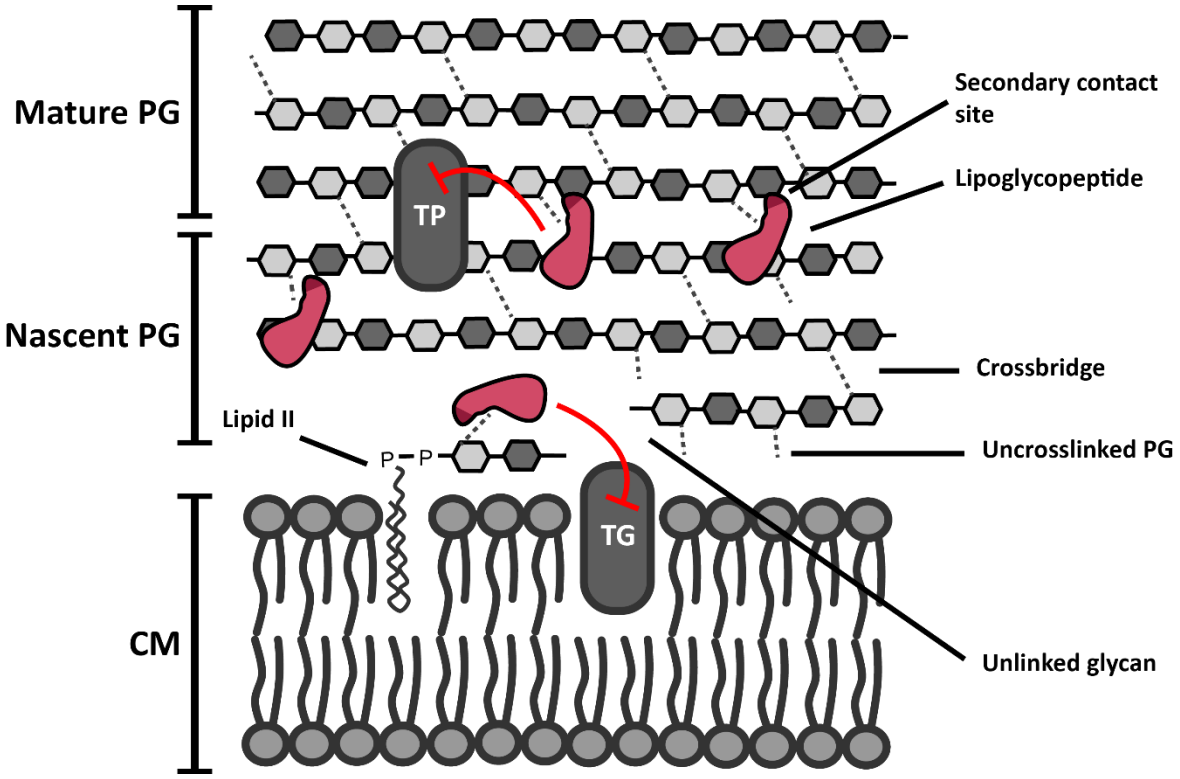



**Figure 1.15** Model to show the proposed dual mechanism of lipoglycopeptides in the cell wall of *S. aureus*. PG pentapeptides are represented as dotted lines, with mature cross-links between glycan chains shown as connected light grey hexagons. Lipoglycopeptides are represented in red with the hydrophobic side chain making the secondary contact site shown in dark red. By binding both D-Ala-D-Ala and the terminal branch amino acids that make up the crossbridge in the stem peptide, both transglycosylases (TG) and transpeptidases (TP) are inhibited, and PG maturation is reduced.

Secondly, their phenotypes can vary drastically between strains, making them challenging to study. Heterologous expression in a non-pathogenic model organism can be used to address the first issue, but heterologous expression can produce spurious phenotypes that may not be a true representation of the parental organism[107]. One comprehensive solution is to use a non-pathogenic organism with an endogenous *van* resistance system. *S. coelicolor* makes a good candidate for this role. It is a member of the actinomycetes, which are soil-dwelling organisms that collectively make approximately two-thirds of all clinically relevant antibiotics and can have up to 15 intrinsic AMR systems [5,108]. *S. coelicolor* does not produce any glycopeptide antibiotics, but it is genetically well-characterised and possesses an inducible glycopeptide resistance system similar to VRE [37](**Figure 1.13A**).

The *S. coelicolor van* cluster is organised into four transcriptional units (*vanRS*, *vanJ*, *vanK*, *vanHAX*) which are all under the control of the VanS/VanR TCS. Our current model for the TCS in *S. coelicolor* indicates that in the absence of vancomycin, the small molecule phosphodonor, acetyl phosphate, phosphorylates VanR, while VanS acts as a phosphatase to deplete the intracellular pool of phospho-VanR (**Figure 1.16**). After exposure to vancomycin, VanS acts as a kinase, instead phosphorylating VanR and increasing the concentration of VanR-P. As a result, there is an increase in expression from each of the four *van* promoters and an increase in resistance to vancomycin (**Figure 1.16**) [109]. The *van* gene cluster in *S. coelicolor* confers high-level resistance to vancomycin and related glycopeptides [110], but like VanB-type enterococci, VanS is poorly induced in the presence of teicoplanin (**Figure 1.17**). *S. coelicolor* is therefore highly resistant to vancomycin but remains susceptible to teicoplanin. Studies have demonstrated that vancomycin directly binds purified VanS sensor proteins [111,112] and that in the case of the VanS from *S. coelicolor*, vancomycin has to be bound to its ligand to induce the resistance system [113]. Work carried out on both VanB-type VRE and *S. coelicolor* has demonstrated that although the core aglycones of both teicoplanin and vancomycin induce the resistance cluster, the presence of a lipophilic side-chain on the former reduces the transcriptomic response [100,114]. This information is important for understanding how glycopeptide antibiotics interact with the bacterial cell-surface. The function of lipophilic side chains in shielding lipoglycopeptide antibiotics from some VanS SK variants in VRE and *S. coelicolor* demonstrates the importance of these secondary modifications and how they can counteract high-level glycopeptide resistance. This information is important for understanding how glycopeptide antibiotics interact with the bacterial cell-surface.

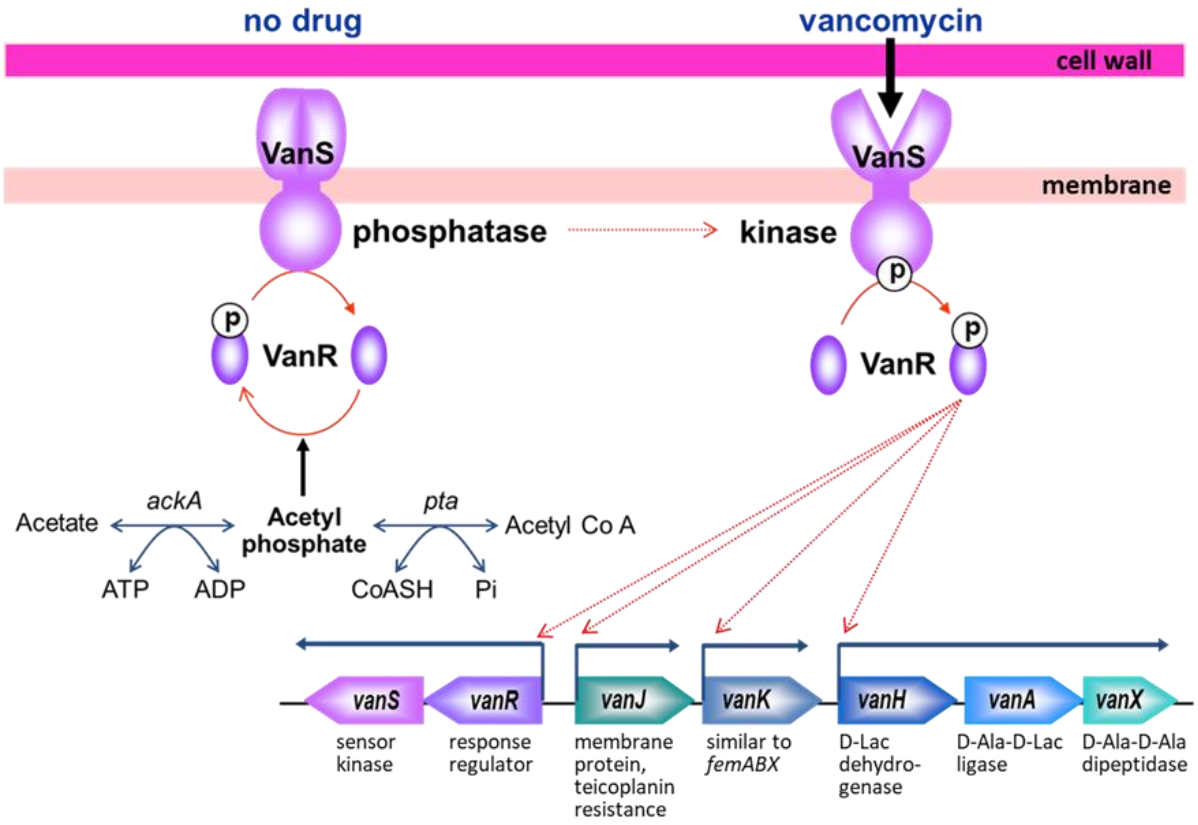



**Figure 1.16** A model illustrating the function and regulation of vancomycin resistance via the VanS/VanR TCS in *S. coelicolor*.

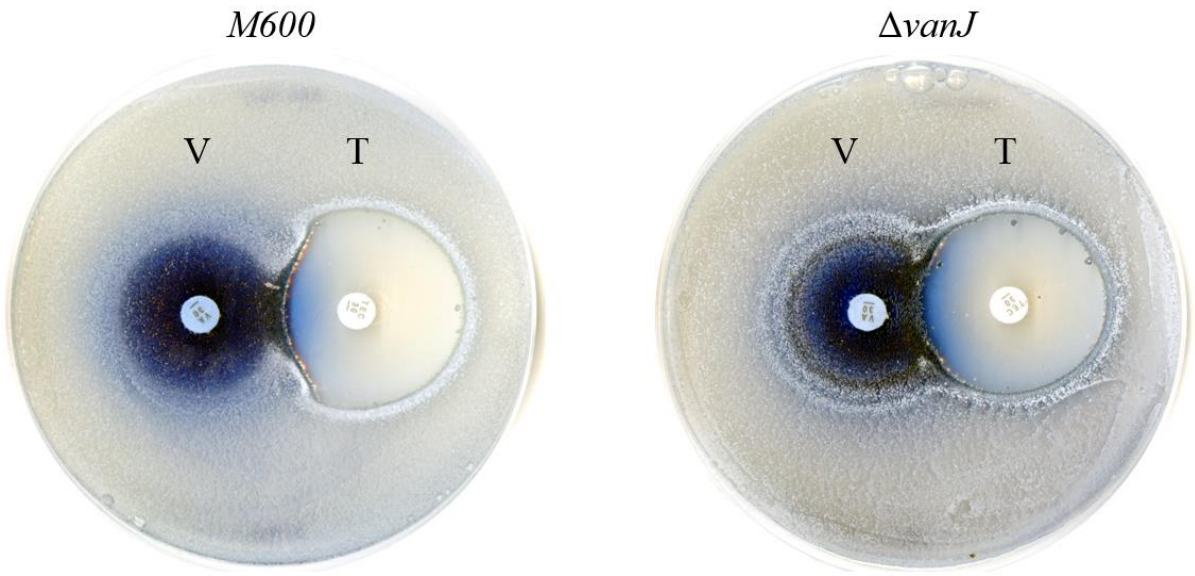



**Figure 1.17** Plates showing the resistant phenotype of *S. coelicolor A3(2)* M600 *wt* and *S. coelicolor A3(2)* Δ*vanJ* strains on MMCGT with vancomycin (30µg) and teicoplanin (30µg) antibiotic disks. In M600 the characteristic D shape in the inhibitory zone is seen because *vanJ* is induced in the presence of vancomycin, increasing resistance to teicoplanin, allowing cells to grow. In the Δ*vanJ* mutant, the zone of inhibition remains circular because the cells remain sensitive to teicoplanin even in the presence of vancomycin.

### 1.4.5 The novel *vanJ* gene from *S. coelicolor* mediates specific resistance to teicoplanin

The *S. coelicolor van* cluster has two additional accessory genes, *vanK* and *vanJ*. The function of *vanK* has been elucidated and has no counterpart in other resistance clusters in pathogenic organisms, but is essential for vancomycin resistance in *S. coelicolor* [110]. It belongs to the Fem family of enzymes which add the branch amino acids to the pentapeptide in PG precursors (**Figure 1.8**). The native FemX enzyme adds a single glycine to pentapeptides terminating in D-Ala-D-Ala but cannot recognise precursors terminating in D-Ala-D-Lac. VanK can recognise D-Ala-D-Lac terminating precursors and therefore can add a single glycine to the pentapeptide when *vanHAX* are expressed.

Work also carried out by Hong's group has demonstrated that the other gene, *vanJ*, is not involved in resistance to vancomycin but provides resistance to teicoplanin-like glycopeptide antibiotics when overexpressed [115]. The gene encodes a membrane protein oriented with its C-terminal predicted active site exposed to the extracytoplasmic space. Over-expression of *vanJ* in a Δ*vanHAX* mutant strain of *S. coelicolor* produces the same increase in teicoplanin resistance as overexpressing *vanJ* in the wild-type strain, indicating that this phenotype is independent of the reprogramming of cell wall precursors encoded by the VanHAX proteins [115].

VanJ belongs to a family of genes with a predicted endonuclease/exonuclease/phosphatase (EEP) domain. Other proteins with this domain have diverse roles in other bacteria including SpnA from *Streptococcus pyogenes*, which acts as a DNase to degrade neutrophil extracellular traps during infections [116]. Others, such as REX1 and REX2, are involved in mRNA maturation within the mitochondria of trypanosomes [117,118]. This superfamily of proteins is heavily implicated in numerous processes in both eukaryotic and prokaryotic cells involving the cleavage of phosphodiesters, but the particular role of VanJ remains elusive. This protein has been shown to be localised to the membrane with an externally facing catalytic domain that does not inactivate teicoplanin or modify PG precursors. One suggestion for the role of VanJ has been the cleavage of the phosphodiester bond of lipid II in the presence of teicoplanin to alleviate the effects of teicoplanin on the cell wall [115]. But this fails to explain why VanJ specifically provides resistance for teicoplanin-like glycopeptides and not vancomycin and its derivatives.

Homologues of *vanJ* are found ubiquitously in actinomycetes, including *staP* from the A47934 producer, *Streptomyces toyocaensis* (**Figure 1.17**). With a high amino acid sequence similarity between VanJ and StaP (>70%), they appear to represent a subset of a larger protein family which have acquired specialist roles in teicoplanin resistance. In comparison, other homologous genes such as *SCO2255* and *SCO7017* encoding putative membrane proteins with much lower amino acid identity to VanJ (50-60%) are found in *S. coelicolor* and other actinomycetes, but are not involved in teicoplanin resistance [115]. Further understanding of the functional role of VanJ and its homologues will improve our understanding of the mechanisms necessary for teicoplanin resistance, and consequently also provide information on its activity why this drug is more biologically effective.

# 1.5 Aims and objectives

This project combines transcriptomics with functional analysis to identify the possible role of the teicoplanin-resistance gene, *vanJ,* by comparing the transcriptomes of a *wt S. coelicolor A3(2) M600* strain and an isogenic Δ*vanJ* knock-out mutant. The specific teicoplanin stress response will also be analysed using the current knowledge on the *S. coelicolor* stress response to other cell wall targeting antibiotics in order to identify possible novel teicoplanin resistance genes. Each proposed gene will be validated for their possible role in altering the resistance of *S. coelicolor* against a library of antibiotics because resistance systems can show cross-reactivity between similarly related drugs. In doing so, this project aims to improve our understanding of how *vanJ* alters resistance to teicoplanin by identifying which systems are affected in the vanJ null mutant. This project also aims to identify novel resistance markers that are involved in teicoplanin resistance that can aid the design of more effective therapeutics.

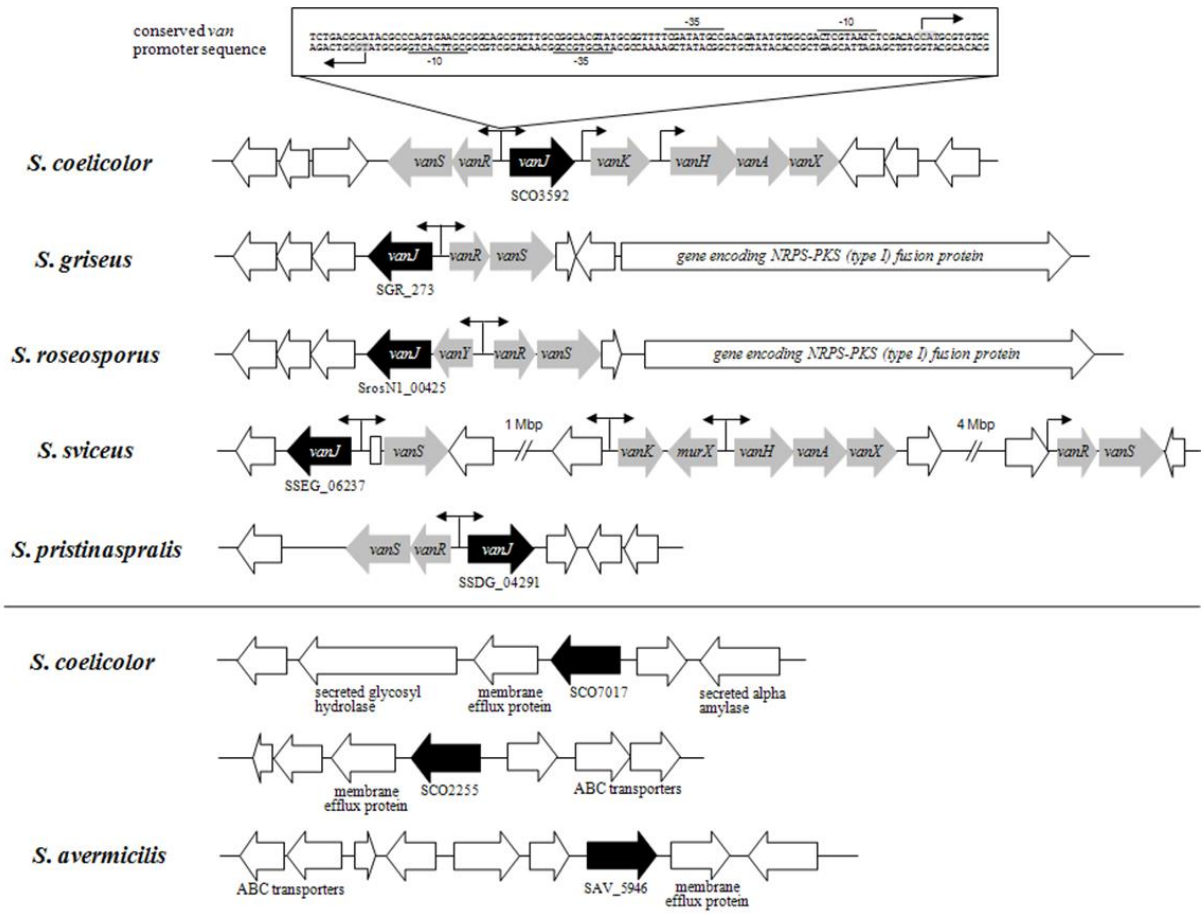



**Figure 1.17** VanJ homologues in *S. coelicolor* and other *Streptomyces* strains (adapted from Novotna *et al*., 2012).

# Chapter 2. Materials and Methods

# 2.1 Bioinformatic analysis of RNA-sequencing data

## 2.1.1 Extraction of RNA from *S. coelicolor* wild type and Δ*vanJ* strains

RNA was prepared and sequenced prior to the start of this project. RNA was prepared as previously described [100], by growing both *S. coelicolor* M600 and *S. coelicolor* Δ*vanJ* null-mutant strains to mid-log phase in NNMP media (~0.5 at $OD_{600}$). Immediately after the first sample was taken (time 0 min), teicoplanin was added to cultures to make a final concentration of 10μg/mL and samples were taken at 30 and 90 minutes using the RNeasy Midi kit (Qiagen). rRNA was depleted from samples using Ribo-Zero (Illumina), and *S. coelicolor* mRNA was then randomly primed to create a strand-specific cDNA library. The cDNA was checked using gel electrophoresis and sample concentration was measured using a ThermoScientific NanoDrop 2000c spectrophotometer. Approximately between 1.5 and 3 mg/mL of cDNA was sent to GATC (acquired by Eurofins in 2017) to be sequenced using an Illumina HiSeq sequencer.

## 2.1.2 Upstream processing of RNA-seq data using the Tuxedo package

FastQC (https://www.bioinformatics.babraham.ac.uk/projects/fastqc/) was used to assess the quality of the raw read data. We employed the pipeline outlined in Pertea *et al*. (2016), which used the HISAT2 (hierarchical indexing for spliced alignments of transcripts) package to align raw sequencing reads to the *S. coelicolor* reference genome (ftp://ftp.ncbi.nih.gov/genomes/refseq/bacteria/Streptomyces_coelicolor/latest_assembly_versions/GCF_000203835.1_ASM20383v1/GCF_000203835.1_*). The reads were then assembled into full and partial transcripts with StringTie, which was also used to identify novel RNA isoforms. The Integrative Genomics Viewer (IGV) was then used to check for coverage [119–121], and the number of reads was quantified using the Rsubread 3.5 Bioconductor package for R [122].

## 2.1.3 Downstream processing of RNA-seq data using the EdgeR and Limma packages for R

Downstream data analysis was carried out in R studio v3.3.2 (R v3.3.2) using the protocols previously laid out by Chen *et al.* (2016) and Law *et al.* (2016). The EdgeR package was used to filter out lowly expressed transcripts from the data which was then normalised using the trimmed mean of M-values method (TMM), to remove any bias created by differences in

library size. The pcaMethods package for R was used to cluster the data to create an unsupervised principle component analysis (PCA) plot. The voom function in Limma was used to calculate the mean-variance relationship in the data, which was then rescaled to remove this relationship. A linear model of the data was then created using the empirical Bayes method in order for us to estimate the gene-wise variability [122,123]. 12 pairwise comparisons were created, which are listed in **Table 2.1**. An adjusted P-value was calculated by applying the **t**-tests **re**lative to **a t**hreshold (TREAT) method to our data to create a stricter definition of significance [124]. TREAT combines empirical Bayes statistics with a fold-change cut-off set by the user (FC = 2) in order for genes to be classified as differentially expressed (DE) [123,124]. All pairwise comparisons can be accessed at https://doi.org/10.24384/4z3s-4x55.

### 2.1.4 Data clustering and functional enrichment analysis

DEGs were hierarchically clustered based on their expression profiles using Pearson's correlation algorithm. Seven clusters of similarly expressed genes were created and plotted onto separate graphs using ggplot2. GO analysis was carried out as previously described using an annotation file from the EBI website (ftp://ftp.ebi.ac.uk/pub/databases/GO/goa/proteomes/84.S_coelicolor.goa) [125]. The Goseq package was used to identify any significantly enriched gene categories within each cluster whilst accounting for length bias. The Wallenius approximation was then used to calculate the number of under- and over-expressed GO categories within each cluster [125,126]. All data tables and plots generated from this dataset can be accessed at https://doi.org/10.24384/4z3s-4x55.

## 2.2 *E. coli* strains and culture methods

### 2.2.1 Making competent *E. coli* cells for bacterial transformations

Competent cells were made for DH5α and ETpUZ strains of *E. coli* for subsequent transformation steps. A modified version of the standard $CaCl_2$ method [127] with a wash step in $MgCl_2$ (100 mM) followed by suspending the cells in 2 mL of a $CaCl_2$ (85 mM), 15% (w/v) glycerol buffer solution. Cells were distributed into 100 µL aliquots and flash frozen in dry ice for storage at -80ºC.

### 2.2.2 Growth conditions and storage for *E. coli* cells

Stocks of *E. coli* were first created by streaking out the bacteria on LB agar (**Table 2.2)** with any required antibiotics (**Table 2.3**) and left to grow overnight at 37ºC.

**Table 2.1** Table describes the list of pairwise comparisons used in this study to identify DEGs. M600 represents *S. coelicolor* M600 wild type and VANJ represents *S. coelicolor* Δ*vanJ* mutant. T denotes the time in minutes after exposure to teicoplanin.

| Sample | Comparison | RNA Datasets compared |
|---|---|---|
| M600_T90-M600_T0 | A | M600 T90 against T0 |
| M600_T90-M600_T30 | B | M600 T90 against T30 |
| M600_T30-M600_T0 | C | M600T30 against T0 |
| VANJ_T90-VANJ_T0 | D | Δ*vanJ* T90 against T0 |
| VANJ_T90-VANJ_T30 | E | Δ*vanJ* T90 against T30 |
| VANJ_T30-VANJ_T0 | F | Δ*vanJ* T30 against T0 |
| VANJ_T90-M600_T90 | G | Δ*vanJ* T90 against M600 T90 |
| VANJ_T30-M600_T30 | H | Δ*vanJ* T30 against M600 T30 |
| VANJ_T0-M600_T0 | I | Δ*vanJ* T0 against M600 T0 |
| (VANJ_T90-VANJ_T0)-(M600_T90-M600_T0) | J | (Δ*vanJ* T90 against T0) against (M600 T90 against T0) |
| (VANJ_T30-VANJ_T0)-(M600_T30-M600_T0) | K | (Δ*vanJ* T30 against T0) against (M600T30 against T0) |
| (VANJ_T90-VANJ_T30)-(M600_T90-M600_T30) | L | (Δ*vanJ* T90 against T30) against (M600T90 against T30) |

**Table 2.2** Growth media for *E. coli.*

| | | |
|---|---|---|
| Lysogeny Broth (LB) – Sigma-Aldrich | Tryptone | 10 g |
| | NaCl | 10 g |
| | Yeast extract | 5 g |
| | Distilled water | Up to 1 L |
| LB agar | Agar | 10 g |
| | Difco Bacto Tryptone | 10 g |
| | NaCl | 5 g |
| | Glucose | 5 g |
| | Distilled water | Up to 1 L |

**Table 2.3** *E. coli* strains used in this study. Chloramphenicol (Chl) and kanamycin (Kan).

| Strain | Description | Resistance markers | Working antibiotic concentration | Reference/Source |
|---|---|---|---|---|
| DH5α | F- Φ80 *lac*ZΔM15Δ(*lac*ZYA-*arg*F) U169 *rec*A1 *end*A1 *hsd*R17(rk-, mk+) *pho*A *sup*E44 *thi*-1 *gyr*A96 *rel*A1 λ- | - | - | Thermofisher |
| ET12567/pUZ8002 | ET12567 containing helper plasmid pUZ8002 | Chl<br>Kan | 25 μg/mL<br>25 μg/mL | [128] |

Bacteria remained viable for up to a month if plates were properly sealed and stored at 4ºC. Single colonies were picked and inoculated into LB with any required antibiotics (**Table 2.3**). Samples were grown overnight at 37ºC in glass universal tubes whilst shaking at 250 rpm. No more than 10 mL of LB was used for each culture, ensuring sufficient aeration. Cultures were pelleted at 4,000 x g in an Eppendorf 5810 R desktop centrifuge and resuspended in 20% glycerol. If the cells were required for DNA preparation, they were resuspended in 500 µL of LB and half the mixture was used for DNA extraction. The remainder diluted with 40% glycerol. Cells were stored at -80ºC for long term storage

# 2.3 Construction of plasmid DNA

## 2.3.1 Cloning *S. coelicolor* target genes into the pIJ10257 Vector

PCR products were amplified from purified *S. coelicolor* genomic DNA with synthesised oligo nucleotides (**Table 2.4**) using Q5 DNA polymerase (NEB) according to the cycle and specifications provided by NEB for amplifying high G+C DNA sequences. PCR fragments were purified using the QIAquick PCR purification kit (Qiagen) and eluted in 35 µL of TE. 20 µL of purified DNA was then digested with the appropriate restriction enzymes overnight at 37ºC, at which time digested DNA was again purified using the QIAquick PCR purification kit to remove all restriction enzymes and small digested fragments of DNA. Purified DNA was kept in TE buffer (pH 8.0) and stored at -20ºC until needed.

Vector stocks of pre-digested DNA were made by digesting 20 µL of the pIJ10257 vector with either NdeI and XhoI or NdeI and PacI overnight at 37ºC. DNA was purified using the QIAquick PCR purification kit and DNA was stored at -20ºC until needed. The cut vector remained viable for several months of use, however when transformation efficiency decreased, more pIJ10257 DNA was cut and purified.

Digested PCR fragments were ligated into the pIJ10257 vector using a standard protocol for T4 DNA ligase (Promega) and left at 4ºC overnight. For transformation, <10% the total volume of the DH5α competent cells were added from the ligation mixture. After incubating cells with the ligated plasmid for 15 minutes, cells were heat-shocked at 42ºC for 60 seconds. After incubating cells on ice for ~60 seconds, 1 mL of LB was added to the cells, and they were then allowed to incubate at 37ºC whilst shaking at 250 rpm for 1 hour. Serial dilutions of DH5α culture were set up for $10^{1}$, $10^{-1}$ and $10^{-2}$.

**Table 2.4** Primers used in this study for genetic manipulations of pIJ10257 vector and its derivative plasmids.

| Primer | Sequence | Description | Size (bp) |
|---|---|---|---|
| pIJ10257_F | TGGTCGATGTCGGACCGGAG | Forward sequencing primer for pIJ10257 | variable |
| pIJ10257_R | CGTCATCTCGTTCTCCGCTC | Reverse sequencing primer for pIJ10258 | |
| **Primers for genes with a predicted PAP2 domain** | | | |
| SCO2335_F | caattaCATATGATGCACCAAAATGGCGTACACG | Reverse primer with *Nde*I restriction site | 2316 |
| SCO2335_R | cgatgcCTCGAGTCAGCTCCCGGTACGGTCCTTCG | Reverse primer with *Xho*I restriction site | |
| SCO2335_iF | GCCGGTCGCCATCAGCCTG | Internal sequencing primer | - |
| SCO2335_iR | GCATCATCTTCGCGGGCAGCGTCA | Internal sequencing primer | |
| SCO2807_F | cagaaaCATATGGTGACCGAGAAGCAGGTAGCG | Reverse primer with *Nde*I restriction site | 1035 |
| SCO2807_R | atcagaCTCGAGTCAGCGAGCCGGTGCCGGAGT | Reverse primer with *Xho*I restriction site | |
| SCO2807_iF | TCATCGGTCTCATCGGCTTCA | Internal sequencing primer | - |
| SCO2807_iR | AATCGCTCACCCGCGGAAGTC | Internal sequencing primer | |
| SCO6355_F | cagaacCATATGGTGAAGCGTGGCGATGTCGCCGAA | Reverse primer with *Nde*I restriction site | 819 |
| SCO6355_R | cagaaaCTCGAGTCACGCAGGGCGTGGTCGTCCTC | Reverse primer with *Xho*I restriction site | |
| SCO6355_iF | GTTTGCTGATCATCGCTGTCCGC | Internal sequencing primer | - |
| SCO6355_iR | GCCCTCGTGCCACGTCCGTCTG | Internal sequencing primer | |
| SCO6356_F | caattaCATATGATGATTATCGCCCTTGACGGCT | Reverse primer with *Nde*I restriction site | 606 |
| SCO6356_R | cagaaaCTCGAGTCACGACACCCCCAGCAAAGGA | Reverse primer with *Xho*I restriction site | |
| SCO6356_F | GAGTGGACGCGGTGCTGAA | Internal sequencing primer | - |
| SCO6356_R | GCAGCGCCGACATGACGAG | Internal sequencing primer | |
| SCO6511_F | cagttaCATATGATGAGTCCCGACGTAGACCTCACC | Reverse primer with *Nde*I restriction site | 1517 |
| SCO6511_R | cgatgcTTAATTAAGGGGAGTGGTGTCAGGAGAACA | Reverse primer with *Pac*I restriction site | |
| SCO6511_iF | GTCTACACCGGCGTCCACTT | Internal sequencing primer | - |
| SCO6511_iR | GCCGAGGCTGAAGGTGTTCA | Internal sequencing primer | |
| **Primers for the genes in the glycerol operon** | | | |
| SCO1659_F | caattaCATATGGTGTCCAGCTCCGACATCTTCA | Reverse primer with *Nde*I restriction site | 795 |
| SCO1659_R | cgatgcCTCGAGTCAGGCGAACGCGACGTTG | Reverse primer with *Xho*I restriction site | |
| SCO1660_F | caattaCATATGCCCGCTTCGGCGATCATCGA | Reverse primer with *Nde*I restriction site | 1623 |

| | | | |
|---|---|---|---|
| SCO1660_R | cgatgcTTAATTAATTACTCCTCGTCCTCGAGCCAGC | Reverse primer with *Pac*I restriction site | |
| SCO1660_iF | CGATCTTGTAGCCGACCGTG | Internal sequencing primer | - |
| SCO1660_iR | CGACCTGCTCTTCGGCACCA | Internal sequencing primer | |
| SCO1661_F | caattaCATATGCTGGCTCAAGGCCGTCGA | Reverse primer with *Nde*I restriction site | 1677 |
| SCO1661_R | cgatgcTTAATTAACTACTTCTTGTCCAGCACGTCCTG | Reverse primer with *Xho*I restriction site | |
| SCO1661_iF | GGGGAGGGCGAAGGTGAT | Internal sequencing primer | - |
| SCO1661_iR | ACGGCGACGACCAGATGAAC | Internal sequencing primer | |

**Primers for genes in the ontology 'response to drug'**

| | | | |
|---|---|---|---|
| SCO2472_F | gtattaCATATGATGCGCCGCCCGAGACTT | Reverse primer with *Nde*I restriction site | 883 |
| SCO2472_R | cgatgcCTCGAGCTCGACGGGCGTCAGATTCA | Reverse primer with *Xho*I restriction site | |
| SCO3910_F | gtattaCATATGTTGGTGCTCATGACACAGGC | Reverse primer with *Nde*I restriction site | 1359 |
| SCO3910_R | cgatgcCTCGAGCACGTGAAACGGTCAGCG | Reverse primer with *Xho*I restriction site | |
| SCO3910_iF | AGACAACACGGTGCCTCCT | Internal sequencing primer | - |
| SCO3910_iR | ATCATGAGGATCGCCCGCA | Internal sequencing primer | |

900 µL of each dilution was then plated onto LB agar supplemented with hygromycin B (150 µg/mL) to ensure sufficient inhibition of non-successful transformants.

### 2.3.2 Purification of plasmids from *E. coli*

Two to three colonies for each transformant were picked and inoculated into LB with selective antibiotic and left to culture overnight at 37ºC while shaking at 250 rpm. Plasmid DNA was isolated using the standard alkaline lysis and ethanol precipitation method [129]. To remove contaminating RNA, 5 µL of RNase A (10 µg/mL) was added to each sample during the lysis steps of the protocol. To reduce other contaminants in each solution, a phenol:chloroform (pH = 7.5) step was included to remove the remaining protein and lipid cell components. The soluble layer was transferred to a new 1.5 mL Eppendorf tube, and two volumes of 100% ethanol and ¼ volume of ammonium acetate were added to the tube. After several inversions, samples were spun at 13,000 x g for 25 minutes at 4ºC using the Eppendorf 5145 R centrifuge. Two wash steps using 70% ethanol were then carried out on the white precipitate formed at the bottom of the tube. All ethanol was removed from the tube and the pellets were allowed to dry between 25 - 30ºC ensuring no residual ethanol contamination. The pellet was solubilised in 30 µL of TE buffer and stored at -20ºC until needed.

The UV-spectra of each plasmid solution was measured using a ThermoScientific NanoDrop 2000c spectrophotometer to ensure the sufficient removal of all contaminating material. Each plasmid was then digested with the appropriate enzymes to confirm the insertion of the correctly sized DNA. If the expected band was present, the gene was amplified using Q5 and our pIJ10257 primers flanking the 5'-upstream and 3'-downstream regions of the gene. The PCR product was purified using the QIAquick PCR purification kit (Qiagen) and sent for sequencing at Eurofins genomics. If a PCR product was >700 nt, internal sequencing primers were designed for sequencing contiguous stretches of DNA. Contiguous reads were assembled using Gene Studio Professional Edition v. 2.2.0.0, and the consensus sequence was compared against the StrepDB BLAST database for the *S. coelicolor* genome. Sequences with a 100% amino acid match were then carried forward for conjugation into wild type *S. coelicolor* M600. All plasmid maps were generated using ApE 10.14.6 for Windows 10. The resulting plasmids are listed in **Table 2.5**.

**Table 2.5** Plasmids used in this study. Hygromycin B (Hyg) and apramycin (Apr).

| Strain | Description | Antibiotic working concentration | Reference/ Source |
|---|---|---|---|
| pIJ10257 plasmids | | | |
| pIJ10257 | 330 bp *ermE*p promoter with ribosome binding site and multicloning site from pIJ8723 cloned into pMS81 and cut with *Kpn*I-*Nsi*I ($Hyg^R$) | Hyg (100 μg/mL) | [37] |
| pIJ10257 derivatives | | | |
| pIJ10257+*SCO1659* | | Hyg (100 μg/mL) | This study |
| pIJ10257+*SCO1660* | | Hyg (100 μg/mL) | This study |
| pIJ10257+*SCO1661* | | Hyg (100 μg/mL) | This study |
| pIJ10257+*SCO2335* | | Hyg (100 μg/mL) | This study |
| pIJ10257+*SCO2472* | | Hyg (100 μg/mL) | This study |
| pIJ10257+*SCO2807* | | Hyg (100 μg/mL) | This study |
| pIJ10257+*SCO3910* | | Hyg (100 μg/mL) | This study |
| pIJ10257+*SCO6354-53* | | Hyg (100 μg/mL) | This study |
| pIJ10257+*SCO6355* | | Hyg (100 μg/mL) | This study |
| pIJ10257+*SCO6356* | | Hyg (100 μg/mL) | This study |
| pIJ10257+*SCO6357* | | Hyg (100 μg/mL) | This study |
| pIJ10257+*SCO6357-53* | | Hyg (100 μg/mL) | This study |
| pIJ10257+*SCO6357-55* | | Hyg (100 μg/mL) | This study |
| pIJ10257+*SCO6511* | | Hyg (100 μg/mL) | This study |
| CRISPR plasmids | | | |
| pCRISPomyces-2 | *gadhp*-promoter adjacent to BbsI cutsite and scaffold for gRNA. Multiple cloning site in *lacZ* gene and Apramycin$^r$. | Apr (50 μg/mL) | [130] |
| pKCcas9dO | | Apr (50 μg/mL) | [131] |
| pCRvanRS-1 | pCRISPomyces-2 + HA | Apr (50 μg/mL) | This study |
| pCRvanRS-2 | pCRISPomyces-2 + gRNA1 + HA | Apr (50 μg/mL) | This study |
| pCRvanRS-2 | pCRISPomyces-2 + gRNA2 + HA | Apr (50 μg/mL) | This study |

### 2.3.3 Synthesis of CRISPR plasmids targeting the *vanRS* cluster

*E. coli* cells harbouring the *Streptomyces* CRISPR plasmid pCRISPomyces-2 [130] were purchased from Addgene (#61737) and cultured to allow for the purification of a high-purity plasmid. The protocol laid out in Cobb *et al*., 2014 was followed in order to design the required oligos (**Table 2.6**) to make the guide RNA (gRNA) and homology arms (HA). The vector was digested with BbsI and purified as in **Section 2.3.2** and then ligated overnight with the annealed gRNAs. The reaction mixture was transformed into DH5α, and a blue-white screen was carried out on LB agar + 0.5M IPTG + 20mg/mL X-gal. White colonies were selected for overnight culture and in L-broth + Apramycin (50 µg/mL) at 37ºC. Plasmids were purified, as stated previously in **Section 2.3.2**.

Simultaneously, two HAs were amplified from the flanking regions of the *vanRS* gene cluster and purified to a high purity as previously described in **Section 2.3.1**. The two amplified HAs were joined using golden-gate assembly, and the reaction was confirmed by gel electrophoresis. The desired unified fragment was excised and purified using the Qiagen gel extraction kit using the manufacturer's instructions. Both HAs and the gRNA-vector backbone were digested with XbaI and dephosphorylated. After one final purification step using the Qiagen PCR purification kit, a high ratio of digested HA DNA was ligated to the gRNA-vector backbone overnight as previously described, followed by a transformation step. Successful transformants were subcultured in L-broth + Apr 50 µg/mL overnight at 37ºC, and plasmids were extracted and purified as described in **Section 2.5.2**. Plasmids were confirmed by DNA digest with XbaI (NEB) and BamHI (NEB) as per the manufacturer's instructions and plasmids used are listed in **Table 2.5**.

**Table 2.6** Primers used to engineer pCRISPomyces-2 to knockout the *vanRS* genes in *S. coelicolor*.

| Primer | Sequence | Description |
|---|---|---|
| vanRSspcrF | ACGCTGGTTGCTCACCAACGAACG | Forward spacer for the gRNA |
| vanRSspcrR | AAACCGTTCGTTGGTGAGCAACCA | Reverse spacer for the gRNA |
| vanRSspcrF2 | ACGCGACACTCAGCAGCTCCAACG | Forward spacer for the gRNA |
| vanRSspcrR2 | AAACCGTTGGAGCTGCTGAGTGTC | Reverse spacer for the gRNA |
| vanRS_HA1F | atctctagaGCGGGAGTCCGTCCAGAA | Homology arm 1 forward primer with XbaI restriction site |
| vanRS_HA1R | AGAGCTACCGCTGAGCCGATAAAC CGCATACGTGCCGG | Homology arm 1 reverse primer with homology arm 2 overlap |
| vanRS_HA2F | CCGGCACGTATGCGGTTTATCGGC TCAGCGGTAGCTCT | Homology arm 2 reverse primer with homology arm 1 overlap |
| vanRS_HA2R | atctctagaTTCCAGGGCCAGTACGTCT TC | Homology arm 2 forward primer with XbaI restriction site |

## 2.4 *S. coelicolor* strains and culture methods

### 2.4.1 Conjugation of *S. coelicolor* using ETpUZ

The conjugation protocol used in this work was mainly based off work that has been laid out in Kieser *et al*. (2000) for the transfer of replicative and integrative plasmid vectors with the non-methylating *E. coli* strain ET12567 with the pUZ8002 helper plasmid (ETpUZ). Successfully engineered pIJ10257 vectors (**Table 2.5**) were transformed into chemically competent ETpUZ using the same protocol as previously stated for DH5α. Cells were left to grow on LB agar with chloramphenicol (25 μg/mL), kanamycin (25 μg/mL) and the working stock of the selective antibiotic for the conjugated plasmid overnight at 37ºC.

Colonies were picked the following day and cultured overnight in LB with chloramphenicol (25 μg/mL), kanamycin (25 μg/mL) and the additional selective antibiotic for the specific plasmid pIJ10257 or pCRISPomyces-2 (See **Tables 2.3** and **2.5**). The following morning, successful cultures were serially diluted to 1:100 in fresh LB with selective antibiotics and grown to an $OD_{600}$ of 0.4-0.6, which would take approximately 6 - 8 hours. Cells were washed twice with fresh LB and resuspended in 1mL of LB. 500 μL of each culture was mixed with 500 μL of 2 x YT broth containing approximately $10^8$ *S. coelicolor* spores that had been pre-germinated by heat shock at 50ºC for 10 minutes. The mixture of both species of bacteria was serially diluted to $10^{-3}$. Both $10^{-2}$ and $10^{-3}$ dilutions were plated on SFM media and allowed to dry. After 16 - 20 hours, plates were overlaid with 1 mL of water containing 0.5 mg of nalidixic acid to kill off any *E. coli* and any other necessary selective antibiotics. Some antibiotics do not mix well with naladixic acid due to its acidity. In these circumstances, naladixic acid was added first, and the plates allowed to dry. Subsequent antibiotics were then added to the plates in the same way and allowed to dry. The plates were then incubated for seven days at 28ºC. Several large colonies were subcultured on SFM supplemented with 0.5 mg of nalidixic acid and other selective antibiotics and grown for a further 7-days. Any cultures that grew poorly after subculturing were discarded due to their likely loss of the appropriate selective markers.

### 2.4.2 Confirming the correct insertion of plasmids into *S. coelicolor* strains

Colony PCRs were carried out on *S. coelicolor* exconjugants to confirm the expected DNA fragment had been inserted into its genome. The PCR was carried out using OneTaq DNA polymerase (NEB) using the suggested requirements for high G+C content DNA. Nonidet-p40 (2%) was also added to the reaction mixture to help lyse the spores. A sterile toothpick or

pipette tip was used to lightly brush any selected colony and inoculate spores into the reaction mixture. DNA was visualised on an agarose (1.2%) gel. If a colony showed the correct band, it was subcultured to produce a lawn of mycelia that could be harvested after incubating at 28ºC, allowing spores to develop. After incubating the plates for more than a week, 1 mL of 20% glycerol was applied to the surface of the plate, and by using a sterile cotton pad, the spores were gently lifted from the plate. The cotton pad with the spore suspension was inserted into a 2.5mL sterile syringe and the plunger depressed, ejecting the spore inoculum. Spore preparations were then homogenised and distributed into 50-100 µL aliquots. Working stocks of spore preparations could be stored at -20ºC for months without a decrease in viability, however master stocks of each bacterial strain were stored at -80ºC. All *S. coelicolor* strains and media used in this study are listed in **Table 2.7** and **Table 2.8**.

### 2.4.3 Quantification of viable *S. coelicolor* spores

The viability of spore stocks was assessed using the serial dilution method, where stocks were diluted to $10^{-8}$ in $dH_2O$ and plated onto SFM media. Plates were incubated at 28ºC for 3 – 4 days and the number of colony-forming units (CFU) were counted, and the CFU/mL was determined for each stock. The CFU/mL was used to determine the required volume of spore stock needed to give an approximate number of spores in a sample.

### 2.4.4 Antibiotic sensitivity assays of *S. coelicolor* strains

Minimum Inhibitory Concentration (MIC) testing was carried out following the EUCAST guidelines [132]. MMCGT plates were made so that they contained the desired concentration of each antibiotic (**Table 2.9**). Each spore stock was diluted in $dH_2O$ to a concentration of 2.0 x $10^3$ spores/µL and vigorously homogenised. 5 µl ($10^4$ spores) of each spore solution was plated onto each antibiotic plate as three technical repeats and allowed to dry. Plates were stored at 28ºC and scored at approximately four days. The MIC was taken to be the minimum concentration of antibiotic that inhibited all growth between 3 and 4 days. Three biological replicates were carried out for each strain and the average MIC taken for each.

Disc diffusion assays were set up on MMCGT plates with lawns of $10^7$ spores and allowed to dry. Oxoid teicoplanin 30µg disks were applied to each plate along with vancomycin papers disks seeded with 100µg of antibiotic using sterile forceps. Plates were scored after 3 days and disk inhibitory zones (DIZ) were measured using image J. Statistical analysis were carried out in R studio v3.3.2 (R v3.3.2).

**Table 2.7** *S. coelicolor* strains used in this study.

| Strain | Description | Reference |
|---|---|---|
| M600 | SCP1- SCP2- | (Kieser *et al.*, 2000) |
| J3220 | Δ*vanJ*::*apr* SCP1- SCP2- | (Hong *et al*., 2002) |
| sc10257m | M600 + pIJ10257 | This study |
| sc10257v | J3220 + pIJ10257 | This study |
| scSCO1659m | M600 + pIJ10257 + *SCO1659* | This study |
| scSCO1659v | J3220 + pIJ10257 + *SCO1659* | This study |
| scSCO1660m | M600 + pIJ10257 + *SCO1660* | This study |
| scSCO1660v | J3220 + pIJ10257 + *SCO1660* | This study |
| scSCO1661m | M600 + pIJ10257 + *SCO1661* | This study |
| scSCO1661v | J3220 + pIJ10257 + *SCO1661* | This study |
| scSCO2335m | M600 + pIJ10257 + *SCO2335* | This study |
| scSCO2335v | J3220 + pIJ10257 + *SCO2335* | This study |
| scSCO2472m | M600 + pIJ10257 + *SCO2472* | This study |
| scSCO2472v | J3220 + pIJ10257 + *SCO2472* | This study |
| scSCO2807m | M600 + pIJ10257 + *SCO2807* | This study |
| scSCO2807v | J3220 + pIJ10257 + *SCO2807* | This study |
| scSCO3910m | M600 + pIJ10257 + *SCO3910* | This study |
| scSCO3910v | J3220 + pIJ10257 + *SCO3910* | This study |
| scSCO6354-53m | M600 + pIJ10257 + *SCO6354-53* | This study |
| scSCO6354-53v | J3220 + pIJ10257 + *SCO6354-53* | This study |
| scSCO6355m | M600 + pIJ10257 + *SCO6355* | This study |
| scSCO6355v | J3220 + pIJ10257 + *SCO6355* | This study |
| scSCO6356m | M600 + pIJ10257 + *SCO6356* | This study |
| scSCO6356v | J3220 + pIJ10257 + *SCO6356* | This study |
| scSCO6357m | M600 + pIJ10257 + *SCO6357* | This study |
| scSCO6357v | J3220 + pIJ10257 + *SCO6357* | This study |
| scSCO6511m | M600 + pIJ10257 + *SCO6511* | This study |
| scSCO6511v | J3220 + pIJ10257 + *SCO6511* | This study |

**Table 2.8** Growth Media for *S. coelicolor.*

| | | |
|---|---|---|
| 2 x YT | Difco Bacto Tryptone | 16 g |
| | NaCl | 10 g |
| | Yeast extract | 5 g |
| | Distilled water | Up to 1 L |
| Soy Flour Mannitol medium (SFM) - also referred to as MS medium | Agar | 20 g |
| | Soya flour (Holland and Barrats) | 20 g |
| | Mannitol | 20 g |
| | Tap water | Up to 1 L |
| Minimal Medium supplemented with Casamino acids, Glucose and Tiger milk (MMCGT) | Agar | 10 g |
| | L-asparagine | 0.5 g |
| | $K_2HPO_4$ | 0.5 g |
| | $MgSO_4.7H_2O$ | 0.2 g |
| | $FeSO_4.7H2O$ | 0.01 g |
| | Distilled water | Up to 1 L |
| | When ready to use, add $dH_2O$ up to 1L | |
| | Glucose (20%) | 25 mL |
| | Difco Casaminoacids (20%) | 30 mL |
| | Tiger milk | 7.5 mL |
| | L-arginine | 750mg |
| | L-cystine | 750mg |
| | L-histidine | 1000mg |
| | DL-homoserine | 750mg |
| | L-leucine | 750mg |
| | L-phenylalanine | 750mg |
| | L-proline | 750mg |
| | Adenine | 150mg |
| | Uracil | 150mg |
| | Nicotinamide | 10mg |
| Minimal Liquid medium (NMMP) | $(NH_4)_2SO_4$ | 2 g |
| | Difco Casaminoacids | 5 g |
| | $MgSO_4.7H_2O$ | 0.6 g |
| | PEG6000 | 50 g |

| | | |
|---|---|---|
| | Minor Elements solution[2] | 1 mL |
| | Distilled water | Up to 0.8 L |
| | When ready to use, add $dH_2O$ up to 1L | |
| | $NaH_2PO_4/K_2HPO_4$ buffer (0.1 M / pH 6.8) | 15 mL |
| | Glucose (20%) | 2.5 mL |
| | Pre-germinated spore inoculum | Variable |
| 2 x YT | Tryptone | 16 g |
| | Yeast | 10 g |
| | NaCl | 5 g |
| | When ready to use, add $dH_2O$ up to 1L | |

**Table 2.9** Table showing the antibiotic concentrations used for MIC testing of bacteria overexpressing pIJ10257 constructs.

| Antibiotic | Target | Final Antibiotic Concentration (μg/mL) | | | | | |
|---|---|---|---|---|---|---|---|
| Teicoplanin | Transglycosylation /Transpeptidation | 0.06 | 0.13 | 0.25 | 0.50 | 1.00 | 2.00 |
| Vancomycin | Transpeptidation | 5.00 | 10.00 | 20.00 | 40.00 | 80.00 | 160.00 |
| Balhimycin | Transpeptidation | 0.50 | 1.00 | 2.00 | 4.00 | 8.00 | 16.00 |
| Chloroeremomycin | Transpeptidation | 0.13 | 0.25 | 0.50 | 1.00 | 2.00 | 4.00 |
| Ristocetin | Transpeptidation | 0.50 | 1.00 | 2.00 | 4.00 | 8.00 | 16.00 |
| Oritavancin | Transglycosylation /Transpeptidation | 0.50 | 1.00 | 2.00 | 4.00 | 8.00 | 16.00 |
| Telavancin | Transglycosylation /Transpeptidation | 0.50 | 1.00 | 2.00 | 4.00 | 8.00 | 16.00 |
| Dalbavancin | Transglycosylation /Transpeptidation | 0.50 | 1.00 | 2.00 | 4.00 | 8.00 | 16.00 |
| Ramoplanin | Transglycosylation | 4.00 | 8.00 | 16.00 | 32.00 | 64.00 | 128.00 |
| Flavomycin | Transglycosylation | 4.00 | 8.00 | 16.00 | 32.00 | 64.00 | 128.00 |
| D-cycloserine | D-Ala-D-Ala ligase /Alanine racemase | 8.00 | 16.00 | 32.00 | 64.00 | 128.00 | 256.00 |
| Duramycin | Phosphatidyl -ethanolamine | 1.00 | 2.00 | 4.00 | 8.00 | 16.00 | 32.00 |
| Carbenicillin | PBP | 16.00 | 32.00 | 64.00 | 128.00 | 256.00 | 512.00 |
| Bacitracin | Undecaprenyl pyrophosphate | 8.00 | 16.00 | 32.00 | 64.00 | 128.00 | 256.00 |
| Nisin | Lipid II | 8.00 | 16.00 | 32.00 | 64.00 | 128.00 | 256.00 |
| Kanamycin | 30S subunit of the bacterial Ribosome | 0.13 | 0.25 | 0.50 | 1.00 | 2.00 | 4.00 |
| Novobiocin | DNA-gyrase | 1.00 | 2.00 | 4.00 | 8.00 | 16.00 | 32.00 |

## 2.5 Phylogenetic analysis of PAP2 domains

### 2.5.1 Acquisition of amino acid sequences of *S. coelicolor* proteins with PAP2 domains

*S. coelicolor* proteins with PAP2 domains were first identified using the Uniprot and StrepDB databases. The boundaries for each PAP2 domain were then determined using BLAST. The StrepDB database was then used to search for further PAP2 domains in genes from *S. lividans* and *S. venezualae*.

### 2.5.2 Phylogenetic analysis of PAP2 domains

The PAP2 domains from the three *Streptomyces* species plus an outgroup from *Bacillus subtilis* were aligned using MUSCLE in MEGA X. A phylogenetic relationship was generated using maximum likelihood and 1000 bootstrap values.

## 2.6 The use of a colourimetric assay to determine the concentration of glycerol in the growth media of *S. coelicolor* cultures

### 2.6.1 Determining a standard curve for glycerol in NMMP media

10x stock solutions of glycerol were first established by mixing >99% pure glycerol with the sterile NMMP culture media. The concentrations that were found to be the most suitable for this assay ranged between 0.01 – 0.25 mM. The reactions were optimised for volumes in a single 1.5mL Eppendorf tube. Solutions of sodium periodate (10mM), potassium iodide (40mM) and sodium thiosulfate (80mM) were all prepared on the day from powder. 0.2M acetylacetone was prepared weekly, and 4.0M ammonium acetate was stored between 2 – 8ºC for up to 2 years.

100µL of each glycerol-NMMP solution was mixed with 100µL volume of sodium periodate (10mM) in 1.5mL Eppendorf tubes, and the samples were put onto a 30ºC shaker for 5-minutes at ~400rpm to ensure complete oxidation of glycerol. Samples were then immediately put on ice to prevent any further oxidation of glycerol. 250µL of potassium iodide (40mM) and 250µL of sodium thiosulphate (80mM) were then added to quench the reaction by oxidising the remaining periodate ions as excess periodate ions inte rfere with colour development in the

final reaction. 200µL of acetylacetone (200mM) and 100µL of ammonium acetate (catalyst) were added to tubes, and 600µL of water was added to make a final volume of 1mL. The tubes were then left to develop for 10 minutes on a 50ºC heat block. Tubes were then put in ice. The absorbance of each sample was read at 410nm on a ThermoScientific NanoDrop 2000c spectrophotometer using the cuvette feature. A negative baseline control with no glycerol was used to adjust the absorbance values for any background noise. All plots were then generated using the ggplot2 package in R.

### 2.6.2 Establishing growth kinetics of *S. coelicolor*

Growth curves were carried out on both M600, and Δ*vanJ* strains to identify the early to the mid-log phase of growth ($OD_{540nm}$ = ~0.4) in liquid cultures. $10^3$ spores were first spun down to remove any glycerol as this can inhibit germination. Spores were then germinated in 625µL of 2 x YT media at 50ºC for 10 minutes. Cells were then cultured in the same way as previously described for RNA extraction in **Section 2.1.1**. The OD of each culture was then read at 0, 12, 15, 18, 24, 36 and 48h. Each culture took approximately 12.5h to read an $OD_{540nm}$ = ~0.4.

### 2.6.3 Establishing a colourimetric test to determine the glycerol concentration produced by cultures of *S. coelicolor* after exposure to different antibiotics

For each colourimetric assay, cultures were set up in the same way as previously described in **Section 2.1.1**. A negative control at 0 minutes was used before the addition of antibiotics, which were then added once cell densities reached an $OD_{540nm}$ = ~0.4 (**Table 2.10**). 1mL of each sample was taken at 2 and 4h and spun down at 4ºC for 10 minutes. Three technical replicates were tested from each sample by taking 100µL of the supernatant forward for the colourimetric test described in **Section 2.5.1**. The absorbance at $OD_{410nm}$ was adjusted for the baseline, and the glycerol concentration was estimated using the NMMP + glycerol standard curve. Each experiment was repeated three times to generate an average value for three biological replicates and plots were generated using ggplot2 in R.

Dosage effects from teicoplanin and kanamycin were identified in the same way as in **Section 2.5.3**, but cells were treated with 0.1, 1.0 and 10 µg/mL of antibiotic instead. Each experiment was repeated three times to generate an average value for three biological replicates and plots were generated using ggplot2 in R.

**Table 2.10** Antibiotics concentrations used in the colourimetric assay to determine the concentration of glycerol in cultures of *S. coelicolor* after the addition of differing antibiotics.

| **Antibiotic** | **Working concentration** | **Target** | **Source** |
|---|---|---|---|
| Teicoplanin | 0.1, 1.0 and 1 0µg/mL | Transglycosylation/ transpeptidation | Sigma-Aldrich |
| Bacitracin | 50 µg/mL | UPP dephosphorylation | Sigma-Aldrich |
| Ramoplanin | 10 µg/mL | Transglycosylation | Sigma-Aldrich |
| Kanamycin | 0.1, 1.0 and 10 µg/mL | Protein synthesis | Sigma-Aldrich |

# Chapter 3. Preparation of RNA-seq Data and Exploratory Data Analysis

# 3.1 Introduction

Understanding how organisms respond and adapt to their environment has been at the core of gene expression studies since the discovery that mRNA was involved in functional changes within cells. Advances in studying changes in RNA through the advent of probe-based technologies like microarray and more recently next-generation sequencing technologies have increased our understanding of the complexities of the regulatory networks that are involved in the cellular response to internal and external stimuli [133–138]. Older technologies that required intimate prior knowledge of both the genome sequence and the organisation of gene transcription were inherently limited by both their dynamic range and susceptibility to signal noise. Next-generation sequencing (NGS) technologies involving amplification of cDNA libraries has revolutionised studying the transcriptional changes, as RNA-seq can be performed without any prior knowledge of functional sequences, has a superior dynamic range and low background noise [135,137,138].

Although NGS has resolved some issues found in earlier technologies, some considerations still need to be addressed to ensure the quality of any RNA-seq dataset (**Table 3.1**). Currently, a consensus is building around a small number of RNA-seq methodologies for carrying out DE analysis. Even though this is the case, choosing the appropriate pipeline and carrying out the necessary quality control (QC) checks is a necessity for generating reliable data. It is also pertinent that RNA-seq studies be followed up with further experimental evidence from proteomics, metabolomics, or any other molecular technique that can give further support for conclusions made from a dataset [139].

This chapter focuses on the pre-processing of an RNA-seq dataset generated from RNA extracted (by Dr Hee-Jeon Hong) from *S. coelicolor* M600 and Δ*vanJ* before and after exposure to teicoplanin. Here we discuss the quality control (QC) checks carried out on the transcriptomic dataset and the two methods used to identify differences in how teicoplanin affects gene expression in the two strains (Summarised in **Table 3.1** and **Figure 3.1**).

**Table 3.1** Table listing QC steps necessary to ensure quality in RNA-seq analysis.

| Limitations of RNA-seq | Measures to remove bias |
|---|---|
| **Sample preparation** | |
| Introduction of unwanted non-coding rRNAs | • Deplete samples of rRNA using a specialised kit |
| Reproducibility | • Carry out >2 biological replicates for each sample.<br>• Include technical replicates<br>• Random sample prep. |
| **RNA-seq and pre-analysis** | |
| Coverage of the whole transcriptome | • Sequence to a greater depth.<br>• Sequence longer reads. |
| Read quality | • QC checks of raw data to check for sequencing defects, failure to remove adapter sequences or contamination.<br>• Filter out poorly sequenced reads. |
| Read bias from lowly expressed transcripts | • Filter out sequences that are poorly expressed. |
| **Core analysis** | |
| Library size bias | • Convert raw counts to a scale that accounts for differences in library size (CPM and log-CPM).<br>• Normalise gene expression distributions. |
| Reproducibility in data | • Carry out unsupervised/supervised cluster analysis. |
| False positives for differential expression analysis | • Use an appropriate statistical method to identify DEGs in data |

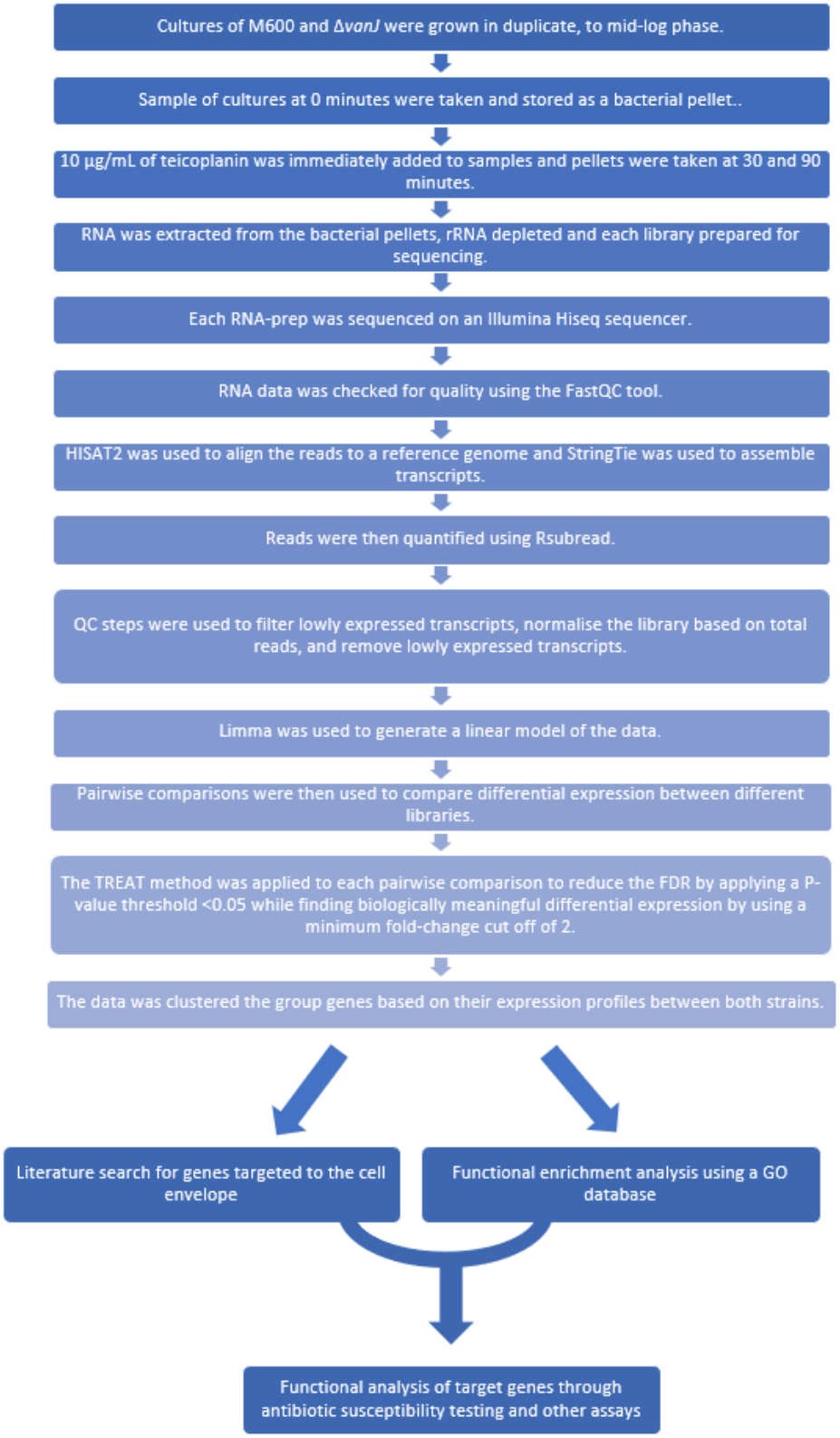



**Figure 3.1** Flow chart for the pipeline used in this thesis from RNA-preparation to functional analysis

# 3.2 Results

## 3.2.1 Examination of the RNA-seq data indicated that it was high quality

The FastQC tool was used to assess the quality of our raw data. All reports can be accessed online (https://doi.org/10.24384/4z3s-4x55) and all had Q-scores >30, indicating that the probability of incorrectly called bases for any of the reads was low (**Figure 3.2B)**. None of the FastQC reports indicated any unexpected issues with the raw data, and the average library size contained 15,644,262 (SD = 2,886,645) reads. The average read length was 51, with no contamination from adapter sequences detected. There were 0 poorly detected reads and 0 bases with Ns called (**Table 3.2**). Over 95% of these reads were aligned to 8,677 sites on the latest reference genome available for *S. coelicolor*. 8,298 of these were open reading frames of genes, 14 were unknown intergenic regions, 342 asRNAs and 23 possible polymerase run-on fragments.

## 3.2.2 QC checks to address bias prior to DE analysis

435 of the 8,677 annotated genomic elements were removed from the analysis as they were not expressed in any sample. Any transcript that was not expressed to a biologically significant level, which was defined as 1 count per million (cpm) in 4/12 samples, was also removed from the data (**Figure 3.3A** and **B**). The remaining 7,756 genes were used for differential expression analysis testing. Both samples VANJ_T30_A and VANJ_T30_B had a higher density of reads than the other libraries. To account for differences in library sizes, each sample was normalised using the TMM, and the differences between pre- and post-normalisation distribution can be seen in **Figure 3.4** along with the scaling factors.

A mean-variance relationship of log-CPM values was created to check for biological variation and to indicate whether poorly expressed genes had been adequately filtered in earlier steps. From this, a 'variance modelling at the observational level' (voom) plot was created. The data showed low biological variation as emphasised by the steep trend represented by the LOWESS curve in **Figure 3.5A**. **Figure 3.5B** validates that the relationship between the mean and variance has been removed.

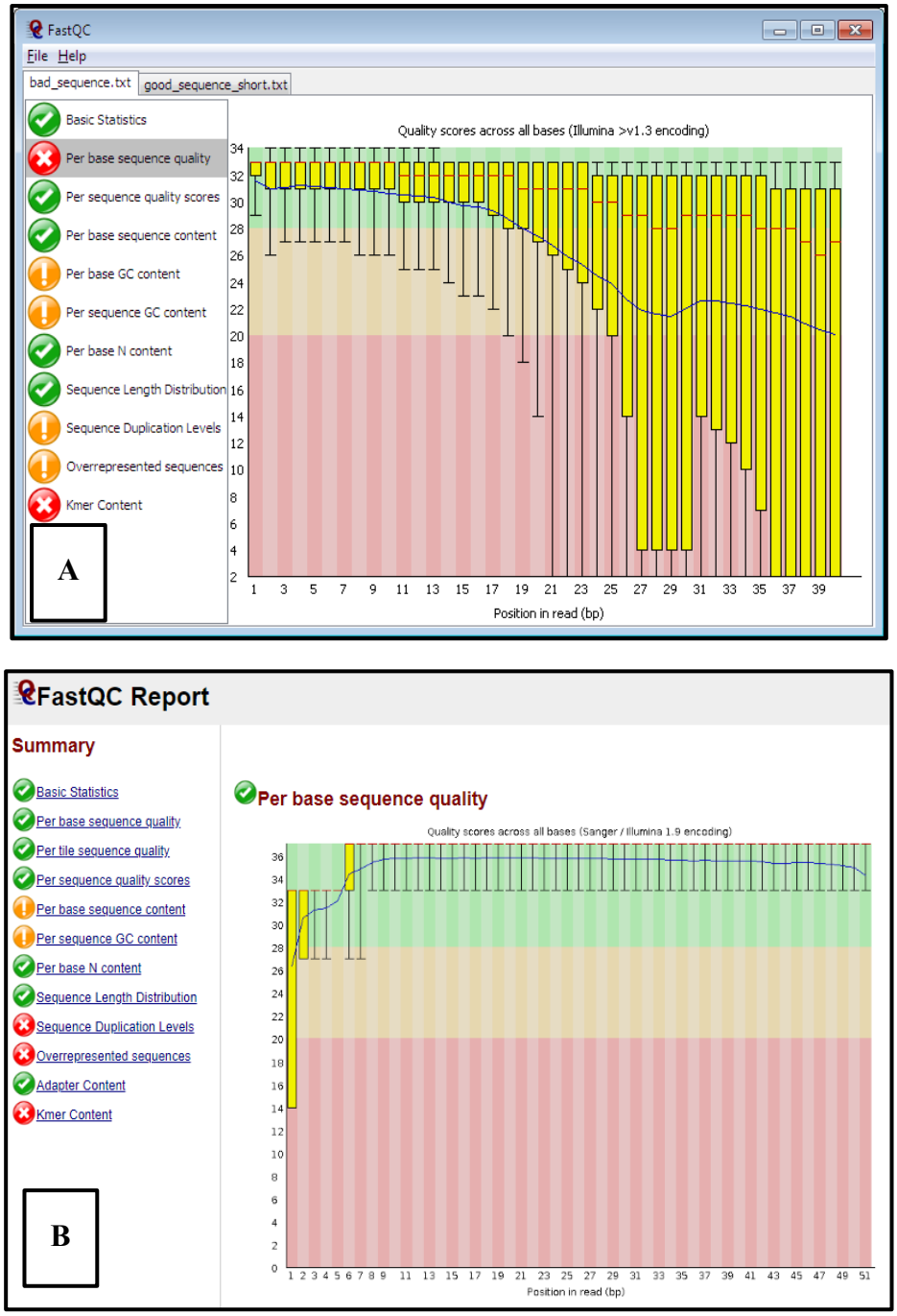



**Figure 3.2** An example FastQC report showing one of the QC checks that FastQC performs on base sequence quality. One low-quality data set taken from the FastQC manual (A) is shown against a report for sample M600_T0_A from the RNA-seq data used in this study (B). The graphs show the quality scores on the y-axis. The y-axis is divided based on quality, with red indicating poor quality reads and green indicating good quality reads.

**Table 3.2** Data quality checks from FastQC tool for each RNA-seq sample used in this study.

| | M600 T0 A | M600 T0 B | M600 T30 A | M600 T30 B | M600 T90 A | M600 T90 B | ΔVANJ T0 A | ΔVANJ T0 B | ΔVANJ T30 A | ΔVANJ T30 B | ΔVANJ T90 A | ΔVANJ T90 B |
|---|---|---|---|---|---|---|---|---|---|---|---|---|
| **Total sequences** | 1.6E+07 | 1.7E+07 | 1.7E+07 | 1.8E+07 | 1.0E+07 | 1.6E+07 | 1.6E+07 | 1.1E+07 | 1.7E+07 | 1.9E+07 | 1.6E+07 | 1.3E+07 |
| **Low quality reads** | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 |
| **Average length** | 51 | 51 | 51 | 51 | 51 | 51 | 51 | 51 | 51 | 51 | 51 | 51 |
| **% GC content** | 68 | 68 | 69 | 70 | 69 | 68 | 68 | 68 | 71 | 71 | 69 | 69 |
| **Bases called as N** | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 |
| **Adapter content** | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 |

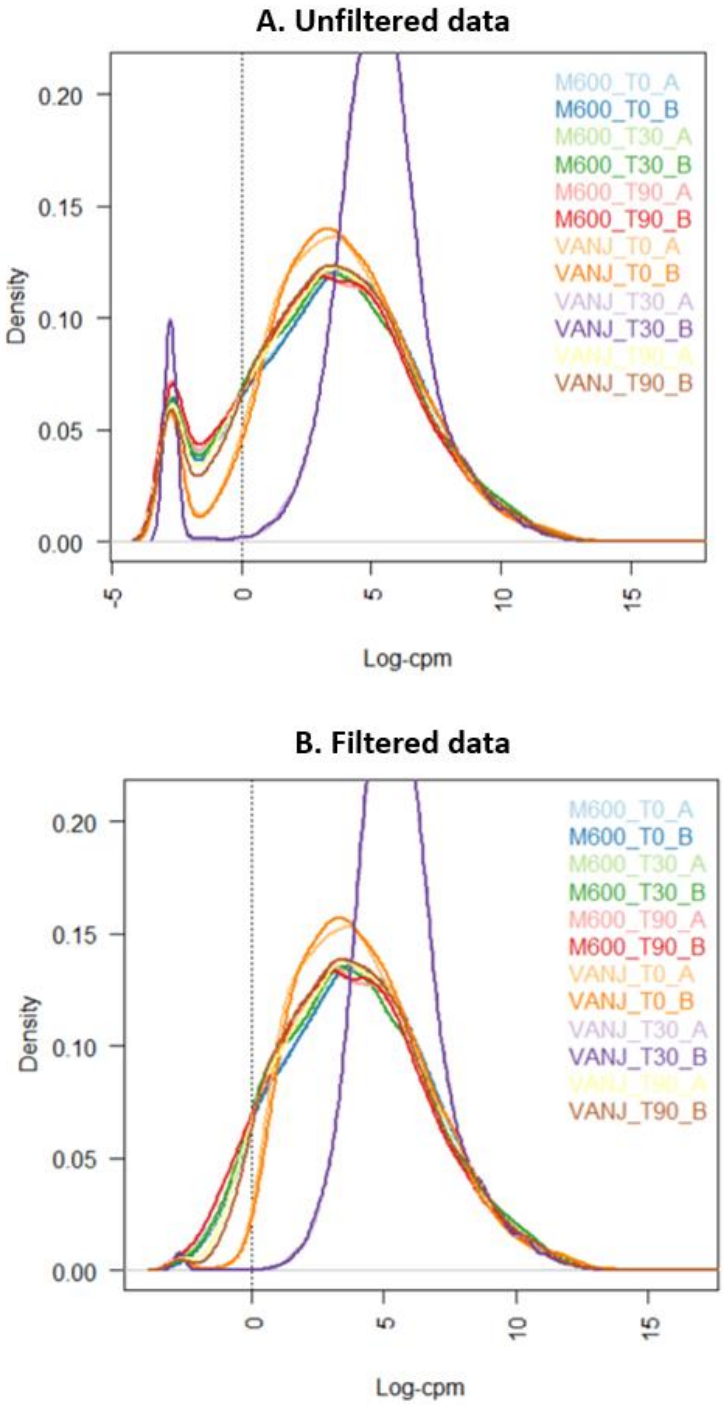
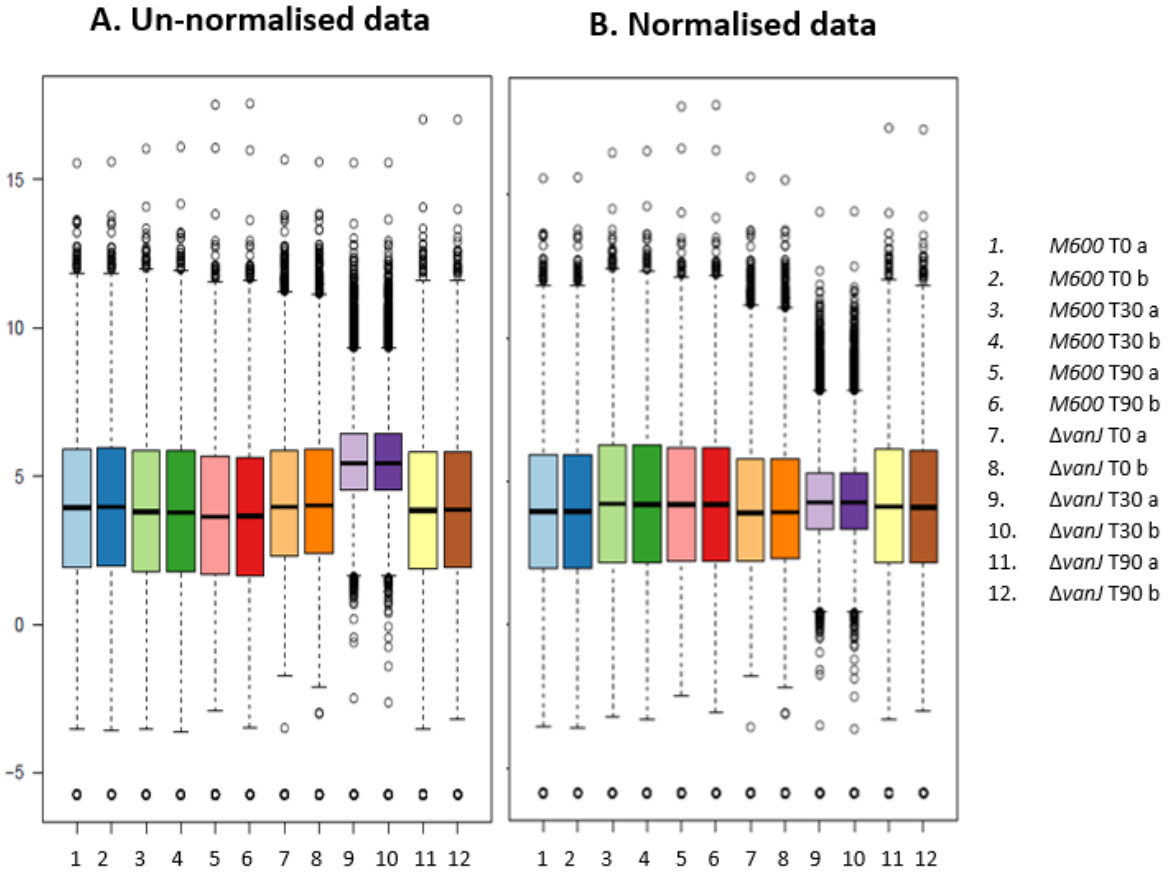



**Figure 3.3** Plots show the density of transcripts before (A) and after (B) filtering data for lowly expressed transcripts. Log-CPM is shown along the x-axis




| Sample | 1 | 2 | 3 | 4 | 5 | 6 | 7 | 8 | 9 | 10 | 11 | 12 |
|---|---|---|---|---|---|---|---|---|---|---|---|---|
| Normalisation factors | 0.991 | 1.002 | 0.748 | 0.750 | 0.687 | 0.680 | 1.041 | 1.063 | 2.248 | 2.241 | 0.817 | 0.844 |

**Figure 3.4** Plots (A) and (B) show how normalising data using the TMM method brings the distribution of the data more in line with the median for all samples. The numbers indicate the dataset listed on the right.

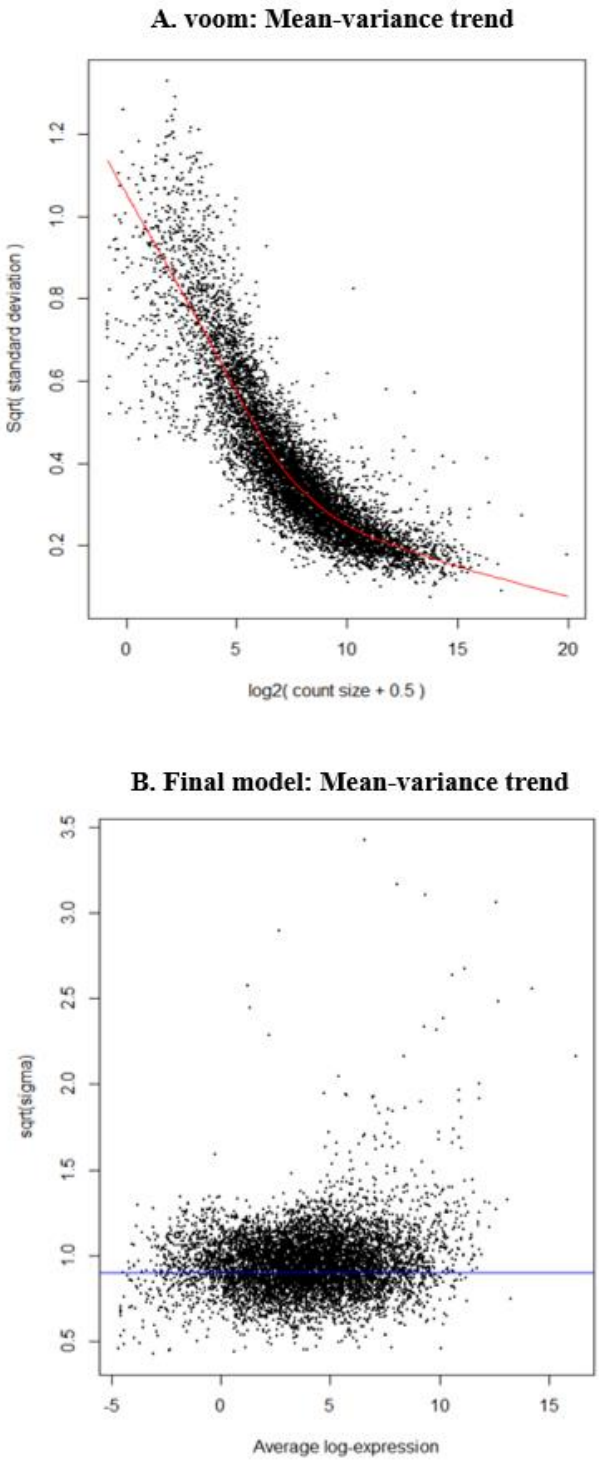



**Figure 3.5** Voom plot (A) showing the mean-variance relationship of RNA-seq data with the LOWESS trend shown as a red line. (B) shows how the trend is removed with the addition of voom's precision weights. The blue line represents the average log2 residual standard deviation. Each black dot represents a single gene

An unsupervised principal component analysis (PCA) was carried out to assess the similarity between samples (**Figure 3.6**). Unsupervised clustering indicates how much differential expression might be detected between two samples. If two points cluster, then there is low biological variation [140]. PC1 separated data based on strain while PC2 based on time. Together they show 84.5% of the variation within these data, so are likely to contribute to differences seen in expression [123]. As expected, replicates cluster closely together, showing that they have little variation, giving us more confidence in our two replicates. We can see from the PCA that there are some differences between the two strains at 0 minutes. For the wild type M600 strain, both the 30 and 90 minute time points cluster closer together than the 0 minute time point. The 90 minute data point for Δ*vanJ* clusters close to the 30 and 90 minute data for the M600, but interestingly, there is a large amount of variation between the 30 minute data points for Δ*vanJ,* strongly indicating that this time point in Δ*vanJ* shows more variation to the 30 and 90 minute time points of M600 and the 90 minute time point of Δ*vanJ.* In all PCAs carried out, replicates clustered together.

### 3.2.3 The TREAT method conservatively reduces the number of genes classified as DE in order to identify biologically meaningful changes in expression

Statistical methods that are used to define DE are good at identifying when expression is different from 0, but this does not always equate to a meaningful biological outcome. The TREAT method applied in this work allows one to assign statistical significance (P-value <0.05) to identify DEGs while also applying a biologically meaningful threshold which in the case of this work, was a fold-change of 2 [124]. This extension to an earlier empirical Bayes moderate t-statistics method [141] is more successful at identifying DEGs with biological significance [124] and helps to reduce FDR.

Pair-wise comparisons were set up between our datasets to identify any changes in expression between samples. We generated three groups of pairwise-comparisons between time points for both strains (**Table 3.3** comparisons 11A-F); pair-wise comparisons of individual time points between the strains (**Table 3.3** comparisons 11G-I); and pair-wise comparisons between the comparisons of time points within a strain (**Table 3.3** comparisons 11J-L). The DEGs (adjusted P-value <0.05, fold-change cut-off = 2) from each comparison are listed in **Table 3.3**, along with the reduction in DEGs after applying the TREAT method.

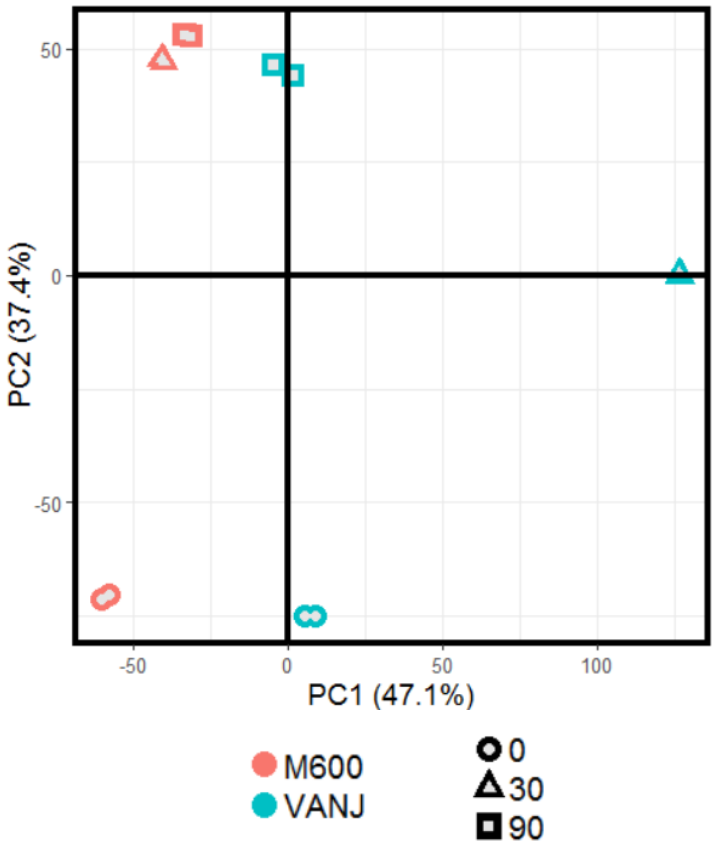



**Figure 3.6** A PCA where PC1 represents strain and PC2 represents time. The amount contributed to variation is indicated in the brackets. Time is indicated as different shapes and strain is represented by either pink (M600) or blue (Δ*vanJ*).

**Table 3.3** Table showing each pair-wise comparison and the number of genes identified as either positively or negatively DE before and after the application of the TREAT method using a P-value < 0.05 and fold-change cut-off = 2.

| Regulation | Comparison | Pairwise comparison | Pre-TREAT | Post-TREAT |
|---|---|---|---|---|
| **Down** | | M600_T90-M600_T0 | 2723 | 674 |
| **NotSig** | A | M600_T90-M600_T0 | 2163 | 5763 |
| **Up** | | M600_T90-M600_T0 | 2870 | 1321 |
| **Down** | | M600_T90-M600_T30 | 2086 | 57 |
| **NotSig** | B | M600_T90-M600_T30 | 4243 | 7599 |
| **Up** | | M600_T90-M600_T30 | 1427 | 100 |
| **Down** | | WT_T30-WT_T0 | 2733 | 528 |
| **NotSig** | C | WT_T30-WT_T0 | 2279 | 5908 |
| **Up** | | WT_T30-WT_T0 | 2744 | 1320 |
| **Down** | | VANJ_T90-VANJ_T0 | 3139 | 680 |
| **NotSig** | D | VANJ_T90-VANJ_T0 | 2118 | 5891 |
| **Up** | | VANJ_T90-VANJ_T0 | 2499 | 1185 |
| **Down** | | VANJ_T90-VANJ_T30 | 3092 | 1751 |
| **NotSig** | E | VANJ_T90-VANJ_T30 | 857 | 4103 |
| **Up** | | VANJ_T90-VANJ_T30 | 3807 | 1902 |
| **Down** | | VANJ_T30-VANJ_T0 | 3207 | 1794 |
| **NotSig** | F | VANJ_T30-VANJ_T0 | 727 | 3663 |
| **Up** | | VANJ_T30-VANJ_T0 | 3822 | 2299 |
| **Down** | | VANJ_T90-M600_T90 | 2092 | 115 |
| **NotSig** | G | VANJ_T9-M600_T90 | 4363 | 7578 |
| **Up** | | VANJ_T90-M600_T90 | 1301 | 63 |
| **Down** | | VANJ_T30-M600_T30 | 4018 | 2511 |
| **NotSig** | H | VANJ_T30-M600_T30 | 732 | 3434 |
| **Up** | | VANJ_T30-M600_T30 | 3006 | 1811 |
| **Down** | | VANJ_T0-M600_T0 | 2066 | 97 |
| **NotSig** | I | VANJ_T0-M600_T0 | 3673 | 7462 |
| **Up** | | VANJ_T0-M600_T0 | 2017 | 197 |
| **Down** | | (VANJ_T90-VANJ_T0)-(M600_T90-M600_T0) | 1853 | 88 |
| **NotSig** | J | (VANJ_T90-VANJ_T0)-(M600_T90-M600_T0) | 4579 | 7595 |
| **Up** | | (VANJ_T90-VANJ_T0)-(M600_T90-M600_T0) | 1324 | 73 |
| **Down** | | (VANJ_T30-VANJ_T0)-(M600_T30-M600_T0) | 3771 | 1984 |
| **NotSig** | K | (VANJ_T30-VANJ_T0)-(M600_T30-M600_T0) | 1475 | 4780 |
| **Up** | | (VANJ_T30-VANJ_T0)-(M600_T30-M600_T0) | 2510 | 992 |
| **Down** | | (VANJ_T90-VANJ_T30)-(M600_T90-M600_T30) | 2632 | 1402 |
| **NotSig** | L | (VANJ_T90-VANJ_T30)-(M600_T90-M600_T30) | 1513 | 4275 |
| **Up** | | (VANJ_T90-VANJ_T30)-(M600_T90-M600_T30) | 3611 | 2079 |

Application of the TREAT method sharply reduces the number of DEGs for all comparisons. Comparison B drops from 3,513 DEGs to 157 or 4.3% of the DEGs called initially. In contrast, the number of DEGs for comparison E is reduced from 6,899 to 3,653 or 52.9% of the DEGs called initially. Together these data support the conclusions made from the PCA; there is less variation between the 30 and 90 minutes time points in M600, but there is much variation between the 30 and 90 minute time points in Δ*vanJ*. We suggest that the two strains respond to teicoplanin differently when looking at comparisons G-I which compare the individual time points. When comparing the 0 and 90 minutes (Comparisons I and G), there are 294 DEGs between the 0 minute time points and 178 between the 90 minute time points. However, there are 4,322 DEGs between the 30 minutes time points (Comparison H), highlighting how much variation there is between these datasets. The gene lists created for each comparison can be viewed online (https://doi.org/10.24384/4z3s-4x55).

### 3.2.4 Biological interpretation of DEGs using hierarchical clustering and functional enrichment analysis

RNA-seq produces large amounts of data. Sorting through all this information can be challenging for anyone trying to identify any functional relationship with the data. Although considering genes significantly up- or down-regulated can reveal genes involved in a response, the process can be labour-intensive and fail to draw any functional relationships between data. A more holistic approach that takes into consideration all the genomic information generated in an experiment is to cluster genes based on their expression patterns. The two primary clustering methods include partitioning such as k-means and hierarchical clustering which groups genes into clusters to form a dendrogram that can be 'cut' at a user-determined place to generate the cluster groups [142–144]. The later was implemented in this work, exploiting the pair-wise relationships generated in the previous section to assess the similarity between genes over the time-course.

The most significant ~1,500 DEGs were hierarchically clustered (**Figure 3.7**) together into seven distinct clusters (**Figure 3.8**). Replicates resembled each other, and the clusters were broadly grouped into four categories based on their response to teicoplanin. This included: *M600*/Δ*vanJ* downregulated (Cluster 1), *M600*/Δ*vanJ* upregulated (Clusters 3, 5 and 7), *M600* constitutive/Δ*vanJ* upregulated (Cluster 2) and *M600* constitutive/Δ*vanJ* downregulated (Clusters 4 and 6).

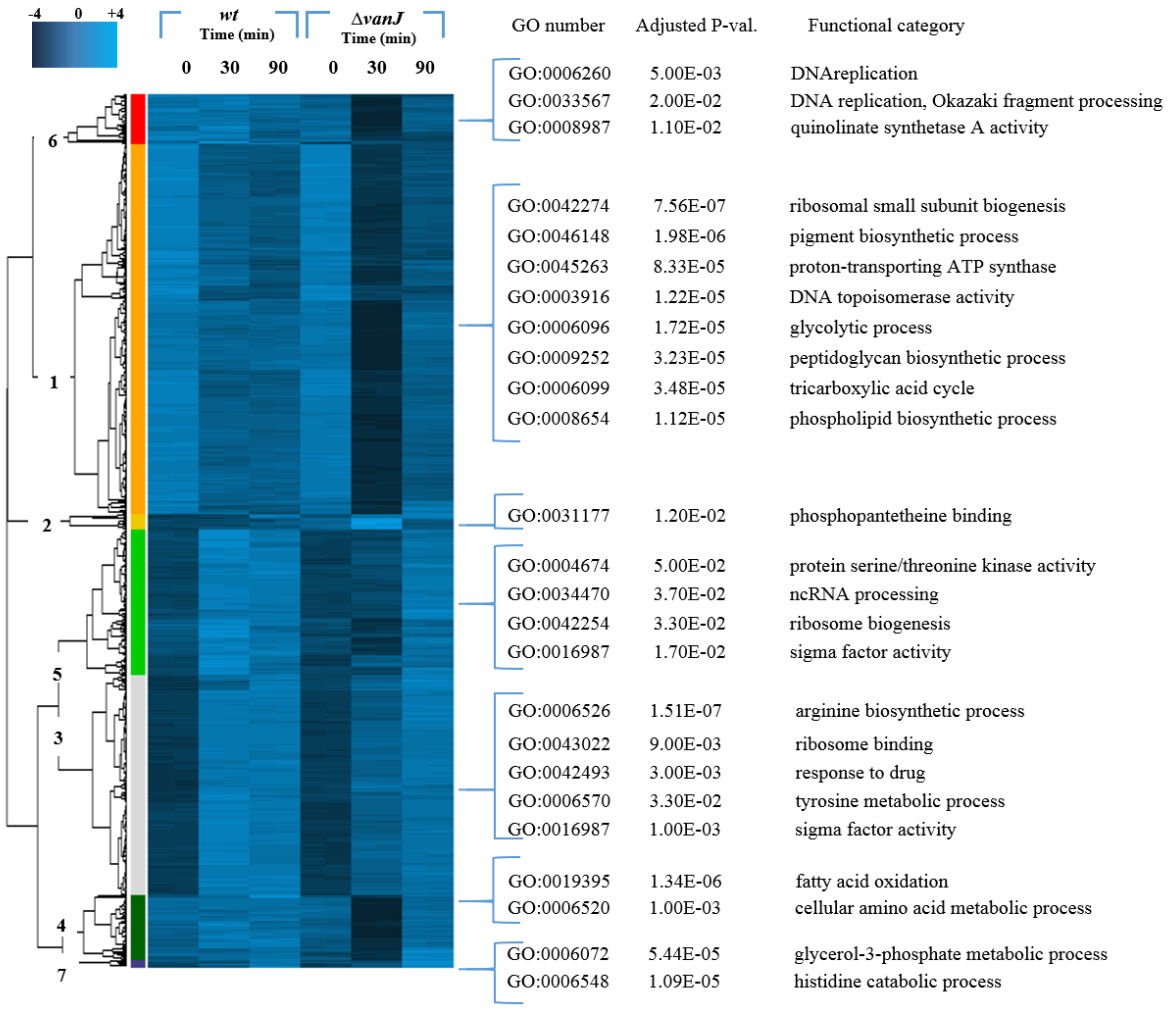



**Figure 3.7** Hierarchical clustering of the 1,557 most significantly DE genes presented as a heat map with the strain and time indicated above. The dendrogram on the left highlights how the clusters were grouped. To the right of the figure are some of the most significantly enriched functional groups pulled out of the data during the GO analysis. GO number, adjusted P-value and category name are all listed with the cluster they were enriched within.

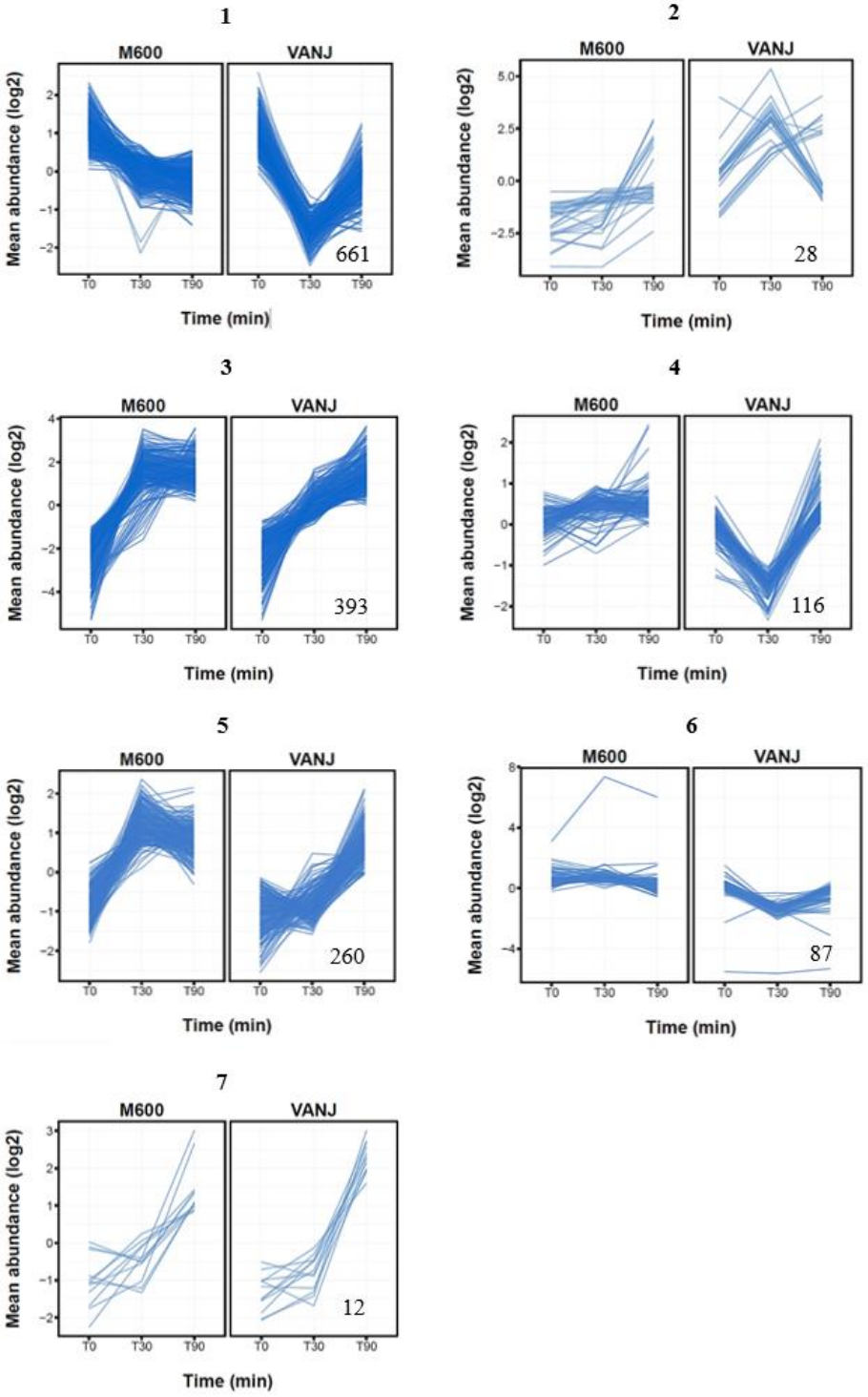



**Figure 3.8** The average $\log_2$ expression for each replicate of the identified DE genes, clustered together based on expression profile for M600 and Δ*vanJ* (VANJ) strains. Each gene is plotted against time in minutes. Numbers above each plot indicate the cluster number and numbers in the bottom right indicate the number of genes in each cluster.

The genetic profile of genes at the 30-minute time point in Δ*vanJ* are more strongly downregulated in Clusters 1, 4 and 6 than in *M600*, while genes in Cluster 2 are induced exclusively in Δ*vanJ*. This confirms our hypothesis that most of the variation observed occurs at the 30-minute time point. It is also clear that Cluster 3 and 5 are both induced at a slower rate in Δ*vanJ* than observed in *M600* (**Figure 3.8**).

The main benefit of clustering data allows for the association of genes that are regulated similarly, as genes that are more similarly regulated are likely to have related functions within the cell. Therefore, if a particular cluster is enriched for specific functional groups, that functional group may be necessary under a particular condition depending on the expression level. In this work, we used a GO database to assign the current functional knowledge for *S. coelicolor* genes, or gene products to the cluster analysis to provide information about how genes function, where their function is located and what biological pathways they are involved with (http://geneontology.org/docs/go-annotations/)[126].

The DEGs from the enrichment analysis for each GO term will be discussed in more detail in the following chapter, but some of GO terms that were enriched for particular genes are displayed in **Figure 3.7**. Cluster 1 and 6 were highly-enriched for key pathways for active growth involved in the production of biomass. Cluster 2 and 4 were highly-enriched for genes relating to phosphopantetheine binding, fatty acid oxidation and amino acid metabolic processes. Clusters 3, 5 and 7 were highly-enriched for genes relating to amino acid metabolism for histidine, tyrosine and arginine, cell signalling including genes for sigma factors and protein serine/threonine kinases (STKs) and glycerol-3-phosphate metabolism. Together, data on Clusters 1, 6 and 7 indicate that there is possibly a shift away from processes that are involved in primary metabolism and the production of biomass. Simultaneously, Clusters 3 and 5 indicate that there may be profound changes on larger parts of the transcriptome as we also observed changes in the expression of genes associated with cell signalling processes.

## 3.3 Summary

In the last few years, a consensus on a few effective methodologies for carrying out RNA-seq analysis has been established. There are several crucial steps that need to be carried out during any RNA-seq pipeline to ensure the reliability of the data, and this section discussed what steps were carried out to ensure this [123,139,145].

Reassuringly, the FastQC reported no problems with the raw RNA-seq data, indicating that there were minimal problems with RNA extraction, cDNA library preparation or sequencing. The average library size was 1.56 x $10^7$ reads with VANJ_T30_A and B having the highest number of reads out of the 12 samples. 921 genes were removed before any analysis for either not being expressed at all or for not being expressed at a biologically meaningful level, reducing interference with statistical analysis [122]. Differences in library size were accounted for by normalising each library using the TMM method, which brought the distributions of VANJ_T30_A and B in line with the other samples. The heteroscedasticity of the data was removed by applying voom to the data so that the mean-variance relationship was removed, allowing the application of Limma-based statistical analysis on the data.

The unsupervised PCA showed that repeats clustered together, demonstrating that there was little variation between replicates. Most of the variation between the two strains occurred at the 30-minute time point, immediately after the cells were exposed to teicoplanin. This indicates that basal levels of VanJ in the membrane of M600 was enough to influence gene expression. Supporting this, differential expression testing was carried out on pair-wise comparisons using the TREAT method in the Limma pipeline. This method required both a fold-change cut off (2-fold) and a P-value (<0.05) threshold for a gene to be classified as DE. This made it easier to focus in on the most significant changes in gene expression. Comparison H, comparing M600 and Δ*vanJ* at 30-minutes, supported this with the most significant number of DEGs shown in this comparison.

The most significant DEGs were hierarchically clustered to produce seven distinct clusters based on the expression patterns of each gene in both strains. Genes in Cluster 1 were found to be downregulated more substantially in Δ*vanJ* than M600, while Clusters 2, 4 and 6 only showed changes in activity in Δ*vanJ*, providing an explanation as to why most of the variation between the two datasets occurs at 30 minutes. Genes in two of the largest clusters, Clusters 3 and 5, were also induced more slowly in Δ*vanJ* demonstrating the profound effects that the loss of VanJ has on *S. coelicolor* gene expression immediately after exposure to teicoplanin.

Higher biological functions were also assigned to genes within each cluster using a GO database. Some of the functionally related genes identified from the GO analysis included the genes relating to primary metabolism such as DNA replication, glucose metabolism and ribosome biogenesis which were all downregulated. Clusters enriched for genes assigned to ontologies relating to both phospholipid and PG biosynthesis were also found to decrease in

activity in response to teicoplanin. In contrast, clusters enriched for genes relating to amino acid metabolism, glycerol metabolism and cell signalling were all upregulated.

In conclusion, we have demonstrated that the quality of this dataset was suitable for differential expression analysis. The following sections will address the conclusions made from this data followed with the experimental data for testing whether any of the DEGs are involved in providing resistance to teicoplanin or other related antibiotics.

# Chapter 4. Interpretation of the Different Responses of the M600 and Δv*anJ* Mutant Strains to Teicoplanin Identified in the Transcriptome Analysis

# 4.1 Introduction

The dedicated glycopeptide resistance system in *S. coelicolor* has been extensively studied, which provides high -level resistance to vancomycin-like glycopeptides when induced [110]. Additional work has also demonstrated the importance of the regulons of genes relating to oxidative [146,147], osmotic [147], and nutritional stress [148] in *S. coelicolor* in aiding the adaptive responses to antibiotics, and how there is often overlap with these regulons and resistance systems in bacteria [149].

Most notably, the cell envelope stress regulon is vital for normal cell envelope function. Although studied for nearly two decades, the regulon has only recently been published for *S. coelicolor* [150]. The regulon is controlled by the global transcription factor $\sigma^{E,}$ which is regulated at the level of transcription by the TSC, CseBC [151]. Unlike VanRS, which is induced by specific glycopeptide antibiotics, the CseBC signal transduction system is broadly activated by a range of compounds that affect the integrity of the cell envelope [151]. $\sigma^{E}$ mutants also exhibit developmental defects and require high levels of magnesium for growth and disruption of any of these genes increases sensitivity to muramidases [152].

Although the specific mechanism by which CseBC sensing damage to the cell envelope is unknown, the majority of the genes within this regulon have a predicted function targeting the cell envelope [150]. This includes genes involved in PG assembly such as PBPs, cell wall teichoic acid deposition, sporulation and membrane modification and maintenance. Although not fully understood, the interplay of these systems along with dedicated resistance mechanisms has been suggested as a method of maximising resistance to antibiotics.

Transcriptomic studies have been useful in establishing the relationships between bacterial stress regulons and antibiotics by providing a method for identifying how drugs influence biological processes [149,153,154]. In this study, we wanted to understand how the presence of *vanJ* influences the genome-wide expression of genes in *S. coelicolor* compared to a Δ*vanJ* null mutant. In the previous section, we discussed how even the minimal induction of *vanJ* by teicoplanin was enough to considerably impact the gene expression of M600 in comparison to Δ*vanJ*. In this chapter, we discuss which genes and regulatory systems are impacted by teicoplanin. We also explore how different regulatory systems are affected between the two strains to provide some further information about a possible role for VanJ. Using literature on the homeostasis of the *S. coelicolor* cell envelope, several subsets of genes were selected to

carry forward for further functional characterisation to validate whether they had a specific role in resistance. Data tables for this section can be found in **Appendix 1**.

# 4.2 Results

## 4.2.1 Teicoplanin induces the expression of many genes known to have roles as transcriptional regulators

Transcriptional regulation involves any method that a cell uses to regulate the synthesis of RNA from DNA. Although bacterial regulatory networks are less complex than eukaryotic signalling cascades, bacteria dedicate a large proportion of their genomes to the regulation of transcription. Approximately 900 gene products (12%) of the *S. coelicolor* genome is predicted to be involved in transcriptional regulation [155], including over 60 sigma factors, 79 TCSs and 40 serine/threonine kinases (STKs) [156,157]. Such regulatory capabilities have likely evolved to allow *S. coelicolor* to strictly regulate the complex mechanisms involved in morphological development and adapt to the fluctuating soil environment.

Clusters 3 (GO:0016987, P-value = $1.00 \times 10^{-3}$) and 5 (GO:0016987, P-value = $1.70 \times 10^{-2}$) from the GO analysis had genes that were significantly enriched for 'sigma factor activity' (**Figure 4.1**). σ-factors are fundamental in directing RNA-polymerases to specific promoter sites to regulate the expression of related groups of genes that carry out a particular function within the cell [156]. Well characterised sigma-factors involved in responding to cellular stress include a series of genes encoding $\sigma^E$ (*SCO3356*), $\sigma^Q$ (*SCO4908*), $\sigma^R$ (*SCO5216*), $\sigma^B$ (*SCO0600*), $\sigma^L$ (*SCO7278*), and $\sigma^M$ (*SCO7314*) that were all significantly upregulated in both strains. The $\sigma^Q$ gene was the most strongly induced gene, increasing by 54 - 180 fold in both strains. This sigma factor has been previously implicated in morphological development and negatively regulating antibiotic production [158]. $\sigma^R$ is implicated in disulphide stress caused by oxidative damage [146] while $\sigma^B$, $\sigma^L$, and $\sigma^M$ form a regulatory cascade with $\sigma^B$ as the master regulator. The cascade regulates genes involved in differentiation, osmotic and oxidative stress responses in *S. coelicolor* [147,156]. All these sigma factors showed a decline in activity in M600 during later time points, suggesting that the inducers for these cascades had declined in activity.

As expected, the cell envelope signal-transduction system was also strongly induced in both strains due to the potent activation of CseBC by glycopeptide antibiotics.

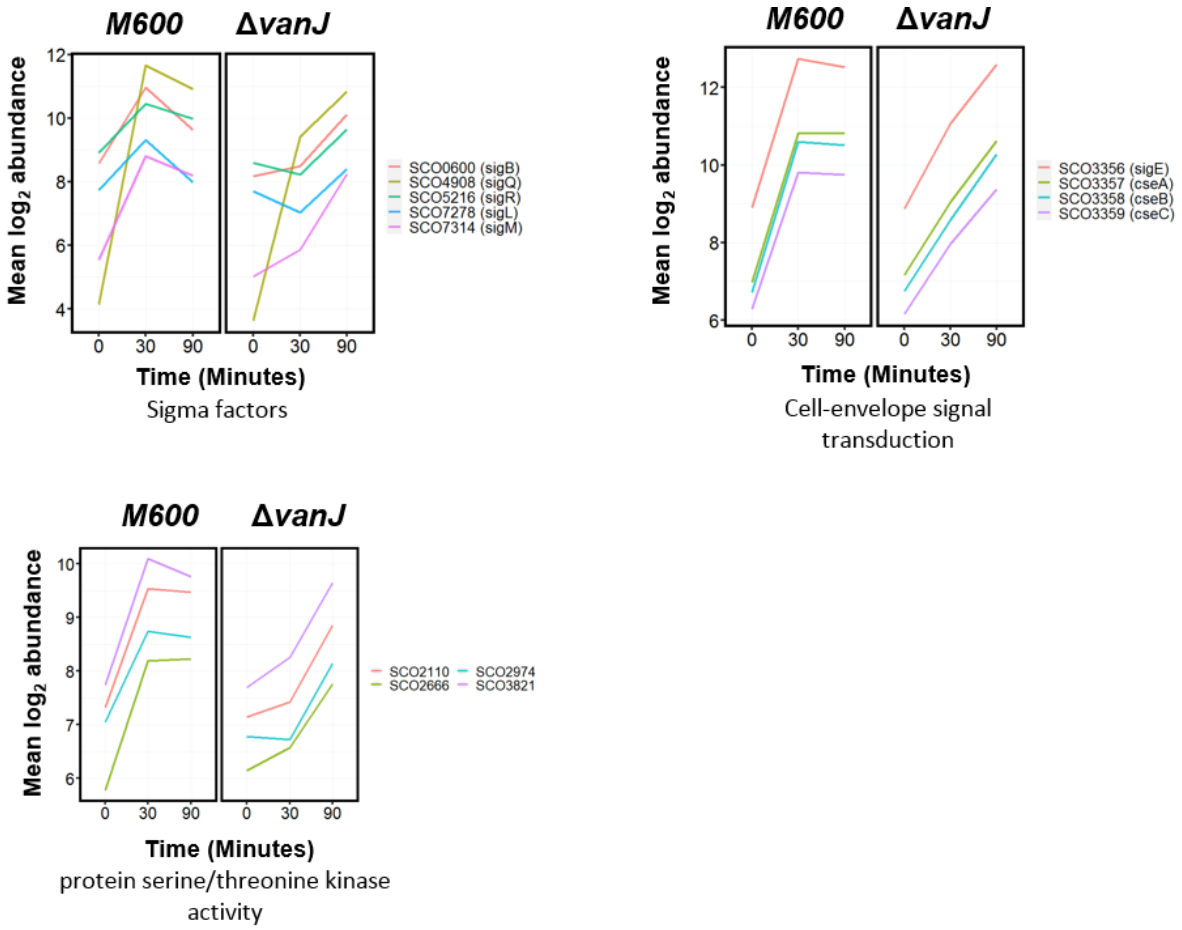



**Figure 4.1** Teicoplanin induces genes involved in cell signal transduction in both strains. Expression profiles show genes encoding sigma factors, SigE signal transduction system and STK activity in *S. coelicolor* are presented. Time in minutes is along the x-axis, and average log2 abundance is along the y-axis. Each point is the average of two replicates.

Glycopeptide antibiotics are known to be potent inducers of the CseBC TCS, but the expression of this system typically declines after exposure to glycopeptide antibiotics that can also induce VanRS such as vancomycin. It is believed that this occurs because the expression of VanHAX alleviates the strain put on cells by the offending glycopeptide. The same response is not observed here presumably because there is minimal activation of the VanRS TCS by teicoplanin in M600. The response seen in Δ*vanJ* was strikingly different with a slower increase in activity over the 90 minutes, indicating that the weak induction of VanJ in M600 could possibly promote the activation of the $\sigma^E$ signal transduction system during earlier time points (**Figure 4.1**). But, as there is still induction of the $\sigma^E$ signal transduction system in Δ*vanJ*, there are likely to be additional mechanisms that lead to the activation of this regulatory system in the presence of teicoplanin.

Cluster 5 was enriched for genes predicted to encode 'STK activity' (GO:0004674, P-value = $4.95 \times 10^{-2}$) which are also involved in signal transduction within bacteria. They resemble the eukaryotic STKs (or Hanks-type kinases) that phosphorylate serine or threonine residues when activated [157,159,160]. All genes within this ontology were significantly upregulated in both strains (**Figure 4.1**). These genes included *pkaF* (*SCO2110*) which has previously been characterised in *S. coelicolor* and encodes a sensory **p**enicillin-binding **a**ssociate **s**erine/**t**hreonine kinase-**a**ssociated (PASTA) domain. *pkaF* has been implicated in negatively regulating the development of aerial hyphae, and actinorhodin production [161] and likely senses an environmental cue relating to PG due to predicted PASTA domain. These genes also exhibited delayed activation in Δ*vanJ* which also suggests that VanJ could be partially involved in triggering their activation and demonstrates the importance of VanJ during the response to teicoplanin.

### 4.2.2 Exposure to teicoplanin affects the expression of genes involved in growth, metabolism and morphological development in *S. coelicolor*

Genome-wide studies on cell wall targeting antibiotics have begun to unearth the systematic effects of antibiotics on cells, with resulting metabolic downshifts and decreased expression of genes involved in cellular growth and the upregulation of genes involved in the cell-envelope, osmotic-shock, heat-shock, redox-stress and the stringent stress-responses. There is also a commonality of the responses between antibiotics with different targets demonstrating the importance of these regulons during times of cellular stress [149,162]. In this section, we explore how the characteristic signs of some of these biological processes are less evident after

exposure to teicoplanin in M600, emphasising the protective role VanJ provides against teicoplanin.

#### 4.2.1.1 The expression of genes relating to the ribosome and protein synthesis is significantly different between the M600 and Δ*vanJ* mutant

Most genes encoding the 30S and 50S ribosome subunits except *rpmF2* (*SCO0436*), *rpmG3* (*SCO0570*) and *rpmE2-2* (*SCO3427*) were significantly downregulated in the Δ*vanJ* mutant by approximately 2 - 10 fold at 30 minutes. In contrast, only 14 out of 56 genes were significantly downregulated in M600 (**Figure 4.2**). Interestingly, the expression level of all these genes was quickly recovered in Δ*vanJ* after 30 minutes (**Figure 4.2**) suggesting that protein synthesis may have been rapidly arrested, and then restarted after the initial exposure to teicoplanin.

Other groups of genes with predicted roles in ribosome biogenesis were also affected by teicoplanin. Clusters 1 (GO:0043022, P-value = $9.3 \times 10^{-3}$) and 3 (GO:0043022, P-value = $9.31 \times 10^{-3}$) had genes enriched for 'ribosome binding' (**Figure 4.2**). Genes found in Cluster 1 include *infC*, *prfB*, *infA* and *prfA* which encode factors involved in peptide chain elongation at the ribosome and their expression levels were also significantly downregulated during earlier time points followed by rapid recovery at 90 minutes only in Δ*vanJ*. Three genes were significantly upregulated by between 6.6 - 28.2 fold in both strains. These encoded the ribosome hibernation factor, *hpf* (*SCO3009*), *SCO4278* which is predicted to encode the peptide chain release factor, and a GTPase with a predicted role in protein synthesis *hflX* (*SCO5796*).

Teicoplanin also caused an increase in the expression of genes relating to stable RNAs. Cluster 5 was over enriched for genes within the ontology 'non-coding RNA (ncRNA) processing' (GO:0034470, P-value = $3.7 \times 10^{-2}$) with genes predicted to be involved in modifying the structural properties of rRNA. The expression of these genes increased by 1.8 - 6.7 fold in both strains. *SCO2533* and *SCO2599* were predicted to encode endoribonuclease activity involved in 16S ribosomal RNA (rRNA) maturation. *SCO5708* is predicted to encode a ribosome binding factor involved in pre-processing of rRNA [163] and *SCO5709* encodes a pseudouridine synthase which is involved in the modification of rRNA to convert uridine to pseudouridine. Psuedouridines are the most common modification seen in structural RNAs and are vital for maintaining transfer RNA (tRNA) and rRNA stability [164] (**Figure 4.2**).

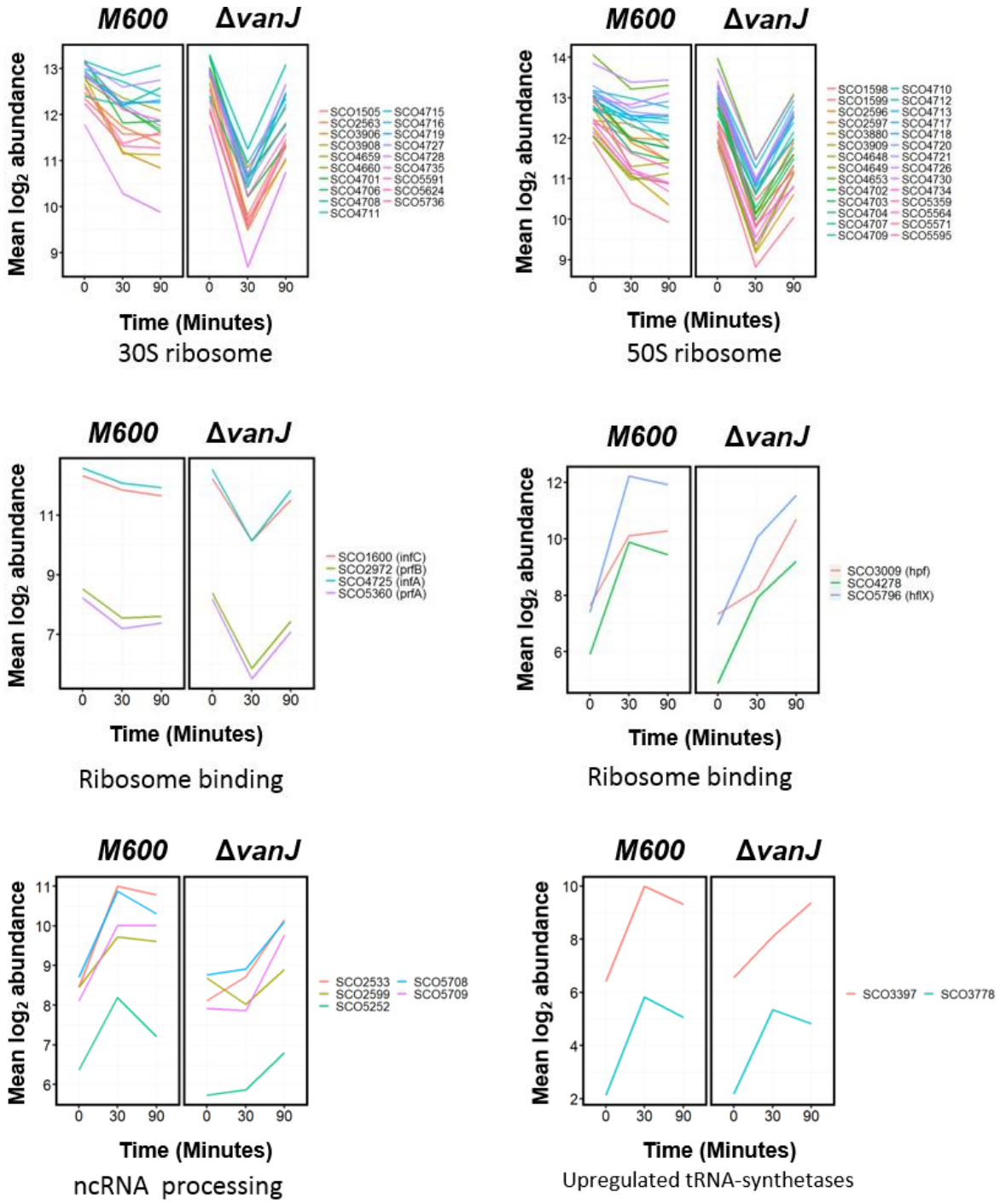



**Figure 4.2** The effects of teicoplanin on the expression of genes involved in proteins synthesis. Expression profiles show genes encoding the 30S and 50S subunits of the ribosome, ribosome binding proteins, ncRNA processing proteins and tRNA-synthetases in *S. coelicolor* are presented. Time in minutes is along the x-axis, and average log2 abundance is along the y-axis. Each point is the average of two replicates.

Some other genes encoding proteins which are involved in the synthesis of other stable RNAs were also affected by teicoplanin (**Figure 4.2**). *SCO3397*, encoding the lysyl-tRNA synthase (*mprF*), was upregulated by 2.9 - 11.8 fold in both strains. While the expression of the gene encoding threonyl-tRNA synthetase, *SCO3778*, was also upregulated by 8.7 - 13 fold in both strains.

Here we present evidence that processes involved in protein synthesis and stable RNAs are affected in both strains. The changes in ribosome biogenesis are transient but are more strongly affected in VanJ in the first 30 minutes. This also includes three Cluster 3 genes encoding ribosome-binding proteins, all five genes encoding ncRNA processing proteins and *SCO3397* encoding the lysyl-tRNA synthase (**Figure 4.2**). *SCO5708*, *SCO5709* and *SCO5796* have all been implicated in the oxidative stress regulon and found to be under the control of $\sigma^{R}$ [165]. Two tRNA-synthetases discussed here, *mprF* and *SCO3778*, are also implicated in antimicrobial resistance. *SCO3778* is essential for resistance to the macrolide antibiotic borrelidin in *E. coli* [149,166]. It may also have a similar role in protecting *S. coelicolor* from the damaging effects of one of its own polyketide natural products, actinorhodin (ACT). *mprF* has been well characterised in several bacteria including *B. subtilis*, *Listeria monocytogenes* and *S. aureus* and is known to add lysine groups to phospholipids. These modifications decrease the overall net negative charge of the membrane, repelling antimicrobial peptides and membrane targeting antibiotics such as daptomycin [53,66,167,168]. This gene also belongs to the $\sigma^{E}$ regulon and is essential for maximal resistance to both vancomycin and bacitracin in *S. coelicolor*, so this gene likely has a vital role in protecting the cell envelope from teicoplanin [149,150].

#### 4.2.1.2 Teicoplanin leads to a reduction in the expression of genes relating to carbon utilisation and DNA replication

There were significant decreases in the expression of genes relating to carbon utilisation and DNA replication in both wild type M600 and Δ*vanJ* mutant. Although expression levels were quickly recovered after 30 minutes, more of these genes were only significantly downregulated in Δ*vanJ* when compared to M600 (**Figure 4.3** and **4.4**). All the genes enriched for 'glycolytic process' (GO:0006096, P-value = $1.72 \times 10^{-5}$), were found to be DE in Δ*vanJ* and were downregulated by approximately 5 fold in the first 30 minutes. There were proportionally more genes also significantly downregulated which related to the 'tricarboxylic acid cycle' (GO:0006099, P-value = $3.48 \times 10^{-5}$), 'ATP synthetase complex' (GO:0045259, P-value = $7.17 \times 10^{-10}$) and 'pigment biosynthetic process' (GO:0046148, P-value = $1.98 \times 10^{-6}$) in Δ*vanJ* than

M600. Genes significantly enriched for the functional process 'pigment biosynthetic process' encode functions related to heme-production for the cytochromes which are involved in oxidative phosphorylation along with genes in making up the ATP synthase [169]. These genes were downregulated by between 1.25 - 5 fold during the first 30 minutes after exposure to teicoplanin in both strains (**Figure 4.3**). This suggests that both anaerobic and aerobic stages of respiration are affected in Δ*vanJ*, but anaerobic respiration may continue in M600. This may relate to the increased levels of cellular stress in relation to oxidative damage in Δ*vanJ* immediately after exposure to teicoplanin.

Cluster 1 was also enriched for genes with functions that relate to DNA biosynthesis, including 'DNA replication' (GO:0006260, P-value = 2.1 x $10^{-3}$), 'DNA topoisomerase activity' (GO:0003916, P-value = 1.21 x $10^{-5}$) and 'quinolate synthetase A' activity (GO:0008987, P-value = 0.01). They were all significantly downregulated in Δ*vanJ* by between 2.5 - 10 fold, but only genes encoding the ribonuclease HII *rnhB,* (*SCO5812*), and a DNA gyrase-like protein (*SCO5836*) were significantly downregulated in both strains (**Figure 4.4**). This information, along with the data on genes functionally involved in carbon utilisation further support the hypothesis that processes relating to growth are more strongly affected in Δ*vanJ* when compared with M600. We suggest that a likely explanation for this is that the Δ*vanJ* strain is unable to cope with the increased cellular stress and resources are likely diverted away from active growth.

#### 4.2.1.3 Teicoplanin induced genes involved in the biosynthesis of arginine

Genes involved in the biosynthetic pathway that converts L-histidine to arginine were strongly induced in response to teicoplanin. This process first begins with the conversion of L-histidine to L-glutamate. Cluster 7 was over- enriched for genes relating to 'Histidine catabolic process' (GO:0006548, P-value = 1.09 x $10^{-5}$) **(Figure 4.5A)** which included *hutH*, *hutU* and *hutI*. *hutI* was significantly upregulated by 3.8 fold in Δ*vanJ*, while *hutU* was upregulated by 3.2 fold in M600 (**Figure 4.5B**). A significant increase of *hutH* was observed in both strains by approximately 2 fold at 30 minutes. **(Figure 4.5A**). The remaining genes in this pathway were found in Cluster 3, which was enriched for genes in the 'arginine biosynthetic process' (GO:0006526, P-value = 1.51 x $10^{-7}$) (**Figure 4.5B**). All the *arg* genes were upregulated at 30 minutes by approximately 5 - 30 fold in both strains and suggests that teicoplanin induces the shift in metabolism. In bacteria, arginine is a hub-metabolite, and it has been found to increase in concentration in response to salt and oxidative stress [170].

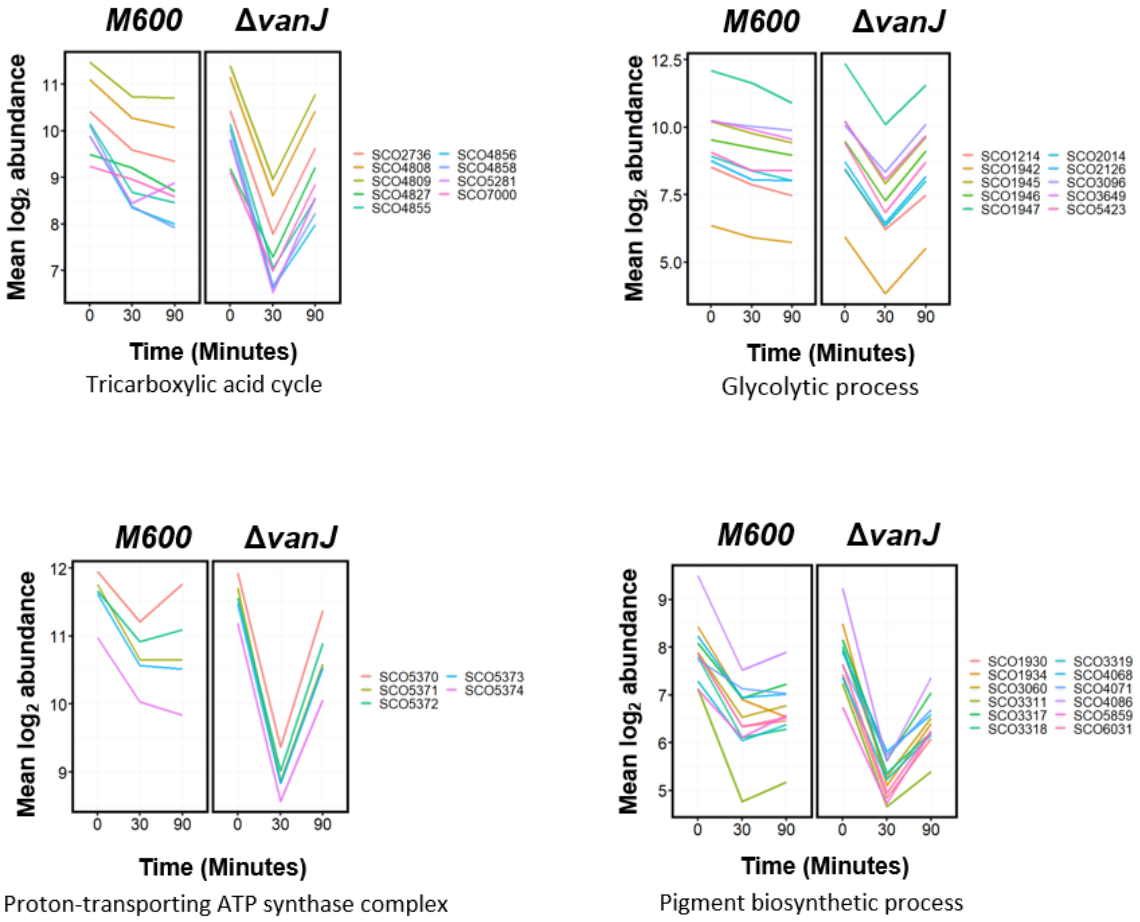



**Figure 4.3** The outcome of exposure to teicoplanin on the expression of genes involved in carbon utilisation. Gene expression profiles for the ontologies: glycolytic process, tricarboxylic acid cycle, protein-transporting ATP synthase complex and pigment biosynthetic process. Time in minutes is along the x-axis, and average $\log_2$ abundance is along the y-axis. Each point is the average of two replicates.

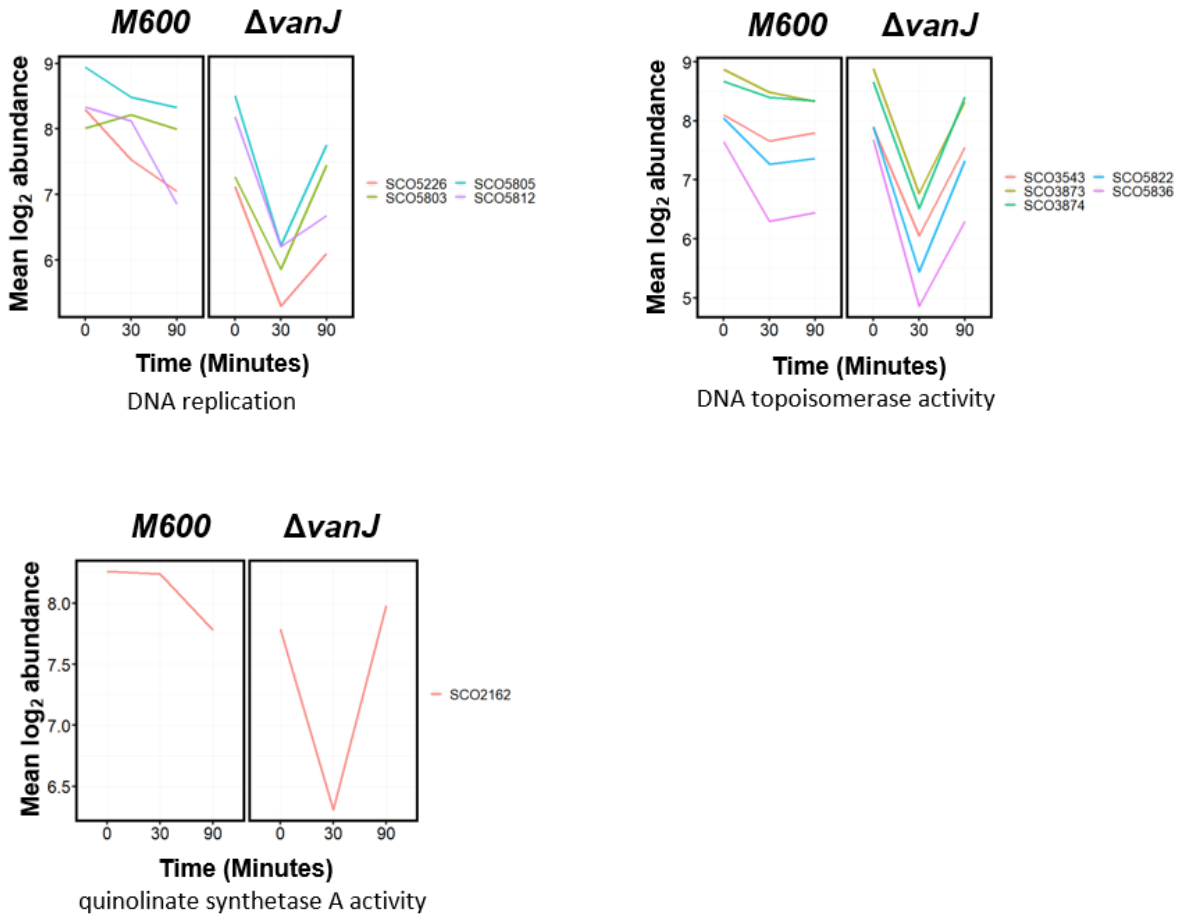
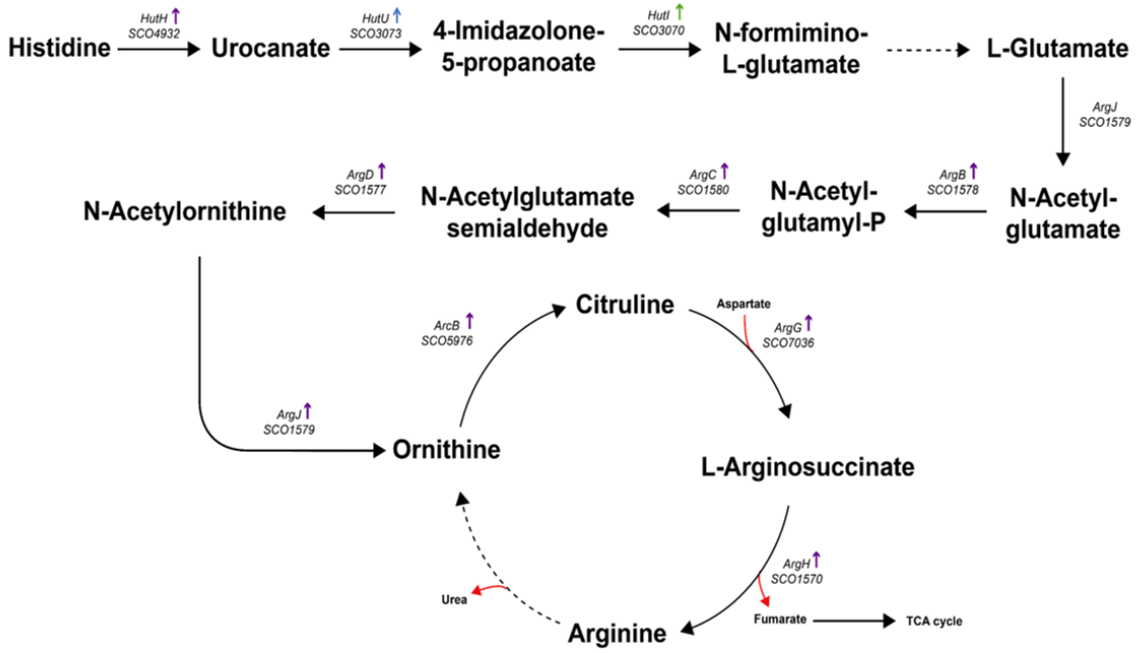



**Figure 4.4** The effects of teicoplanin on the expression of genes involved in DNA replication and biosynthesis. Gene expression profiles for the ontologies: DNA replication, DNA topoisomerase activity, and quinolinate synthetase A activity. Time in minutes is along the x-axis, and average $\log_2$ abundance is along the y-axis. Each point is the average of two replicates.

**A**




**B**

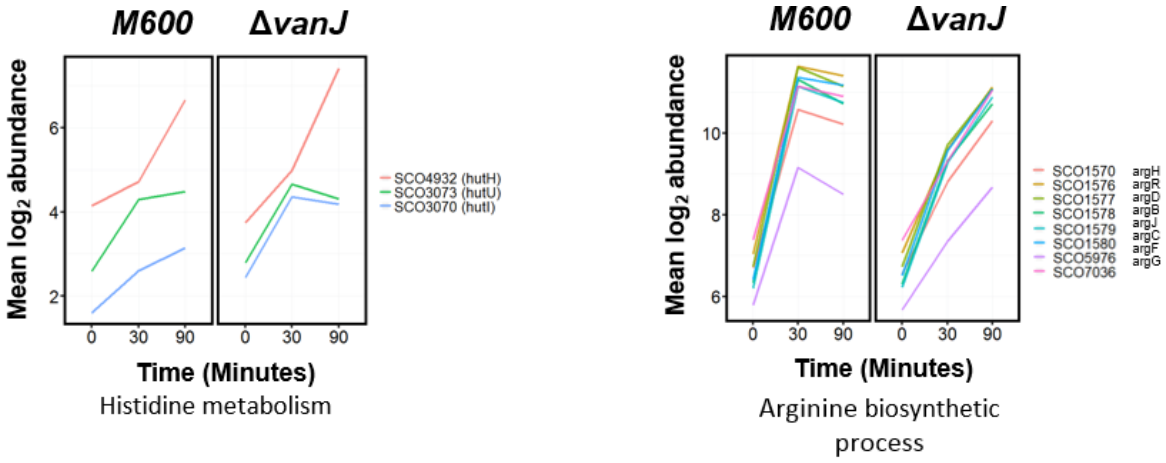



**Figure 4.5** Teicoplanin leads to the induction of genes involved in the pathways involved in the synthesis of arginine. (A) Metabolic pathway of the conversion of histidine to L-glutamate for *S. coelicolor*. Solid arrows indicate characterised steps with known genes, and dashed arrows indicate unknown mechanisms. Coloured arrows indicate the direction of expression after exposure to teicoplanin. Purple arrows indicate DE in both strains. Green arrows indicate DE in just the Δ*vanJ* mutant. (B) Expression profiles for the genes encoding the enzymes involved in the conversion of histidine to glutamate (left). The gene expression profile for the ontology arginine biosynthetic process is also shown (right). Time in minutes is along the x-axis, and average log2 abundance is along the y-axis. Each point is the average of two replicates.

This metabolite is also a precursor for ectoine and the polyamines which are osmoprotectants in *S. coelicolor*. The polyamines have also been documented to protect cells against reactive oxygen species [171], which could suggest that arginine may be vital for protecting cells against the increased cell envelope stress caused by teicoplanin.

#### 4.2.1.4 Induction of genes involved in the synthesis of the hydrophobic coat of the aerial mycelium

*Streptomyces sp*. exhibit complex life cycles with several distinct differential stages, preferring growth on solid substrates, which most likely emulates their environmental niche at the water/air and soil interface [60,172]. Aerial hyphae must undergo several crucial checkpoints before differentiating into chains of unigenomic pigmented spores [173] which can be disseminated into the environment and germinate into new mycelium when the conditions are right. Vegetative mycelia are hydrophilic structures that 'explore' the damp environment of the soil in search of nutrients. However, as they differentiate into their reproductive life cycle, they develop a hydrophobic coat around their aerial mycelia to support their escape from the damp soil environment into the air. The hydrophobic coat, referred to as the rodlet layer, is essential for the formation of aerial mycelia and mature spores. It consists of a mosaic-like structure made up of chaplins, rodlins and SapB that are encoded for by the *chpA-H*, *rdlAB* and *ramCSABR* genes respectively with the *bld* genes playing a role in regulating differentiation [108,173–177]. Most of the genes involved in aerial hyphae development in *S. coelicolor* were not classified as DE in M600; however, the expression of chaplin genes, *chpB*, *chpD*, *chpG*, as well as the *ramCSAB* cluster were significantly, but transiently increased in Δ*vanJ* mutant at 30 minutes (**Figure 4.6**). As many *Streptomyces sp*. do not sporulate in submerged cultures, bar several notable exceptions such as *S. griseus*, *S. albus* and *S. acrimycini*, [178], this was an unusual observation to make.

#### 4.2.1.5 The effects of teicoplanin on the expression of genes involved in sporulation

As the wild type *S. coelicolor* sporulates, colonies change from white to grey as spores mature. Mutants unable to sporulate, remain white as they mature because they lack the grey pigment formed during the later stages of sporulation. These mutants have been termed *whi* (whi-te) mutants because of their lack of colouration as they develop [179]. Aerial hyphae undergo several checkpoints before committing to sporulation which ensures that genomic DNA is correctly partitioned and septal rings are equally spaced along hyphae before septation occurs.

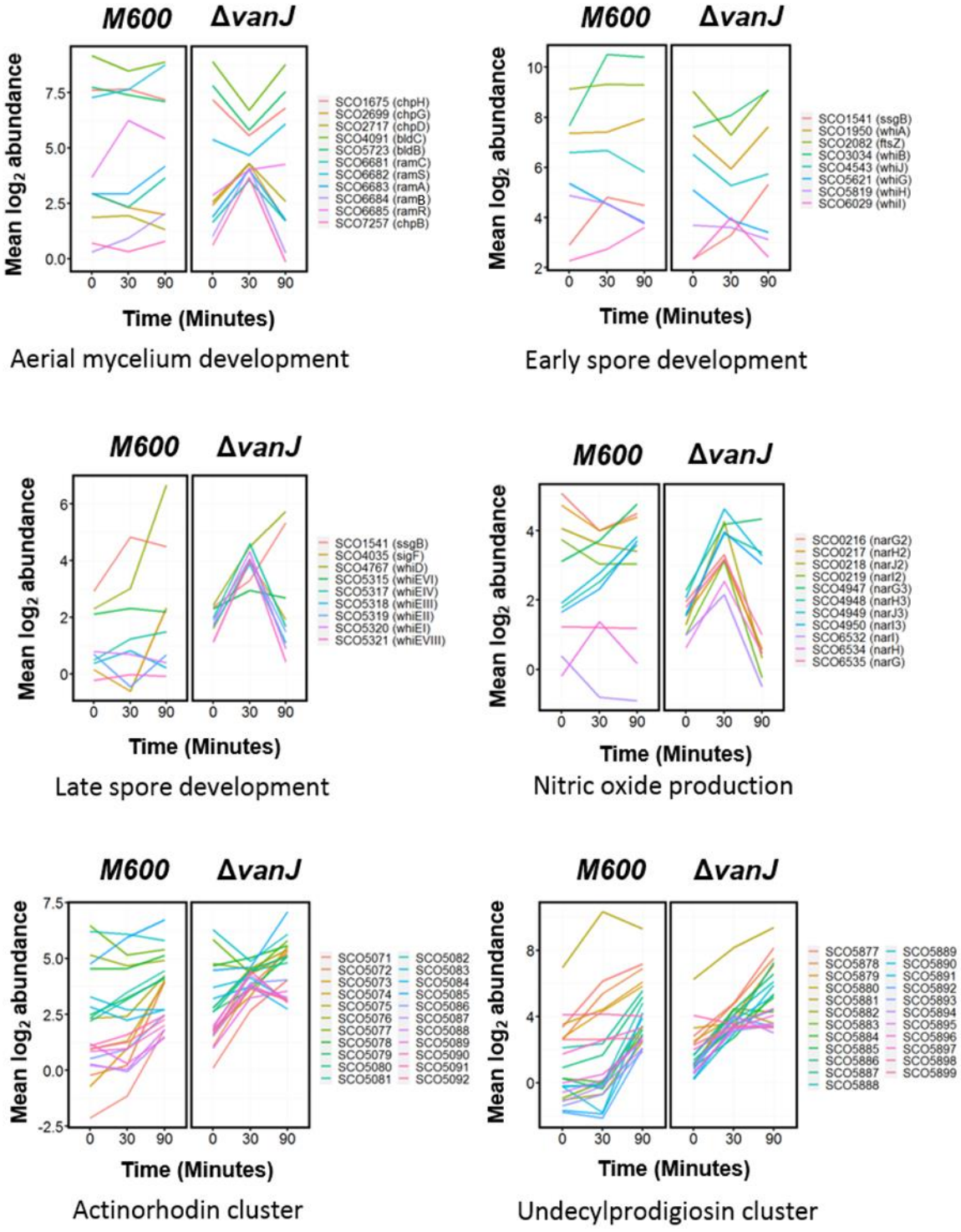



**Figure 4.6** The effects of teicoplanin on the expression of genes involved in developmental processes with *S. coelicolor*. Time in minutes is along the x-axis, and average log2 abundance is along the y-axis. Each point is the average of two replicates.

At least five genes are involved in 'early' spore development, acting as checkpoints during the switch from vegetative to aerial growth. These genes encode the transcriptional regulators, WhiG (*SCO5621*), WhiA (*SCO1950*), WhiB (*SCO3034*), WhiH (*SCO5819*) and WhiJ (*SCO4543*), which are responsible for the correct positioning of septa in each pre-spore compartment and correctly partitioning genomic DNA [180,181].

Few changes were noted in the expression of 'early' genes with a few exceptions. Teicoplanin induced the expression of *whiB* in M600 by 7.1 fold during the first 30 minutes, but no significant changes in expression were observed for the other four genes in this strain. Significant decreases in the expression of *whiJ* and *whiG* were observed in both strains within the 90 minutes time frame. Expression of *whiA* and *ftsZ*, which encodes the septal ring [180], were significantly downregulated only in Δ*vanJ* but the expression of both genes recovered almost back to normal level after 30 minutes. Increases in *ssgB* (*SCO1541*) expression by between 4.5 and 6.1 fold was observed in both strains during the 90 minute time course (**Figure 4.6**). This gene is essential for septation and believed to encode an anchoring protein that targets septal ring to the membrane [180,182]. Interestingly, the expression level of both *whiB* and *ssgB* declined after 30 minutes in M600, but continued to increase in Δ*vanJ* until 90 minutes, suggesting that there could be structural changes in both strains as a result of exposure to teicoplanin.

'Late' developmental genes are induced once the hyphae commit to developing into spores and involves the rounding of pre-spore compartments simultaneously as the spore cell wall is thickened to produce thermoresistant and hydrophobic spores. This process starts with the transcriptional regulator, WhiI which sequentially activates the expression of sigma factor, $\sigma^F$, which in turn activates the *whiE* gene cluster (*SCO5315*, *SCO5317-21*). *whiE* ORFs I-VIII are responsible for producing the grey compound that gives mature spores their pigmentation [183,184]. Significant increases in the expression of both *whiI* and $\sigma^F$ were shown in both strains at 30 minutes, but their expression levels returned to pre-exposure levels at 90 minutes in only Δ*vanJ.* There were also significant but transient increases in the expression of *whiE* ORFs I-IV and VIII (4.6 - 7.6 fold) during earlier time points in Δ*vanJ*, with expression levels decreasing at 90 minutes, providing further support to teicoplanin instigating developmental processes in these strains.

Finally, the *whiD* gene is also an essential gene that is responsible for regulating the processes that thicken the cell wall surrounding spores [185]. The expression of this gene significantly increased in by 14.1 fold in M600 and 6.5 fold in Δ*vanJ* over the 90 minutes time course

(**Figure 4.6**). Together these data suggest that teicoplanin could be influencing some of the processes that are involved in spore development, particularly in Δ*vanJ*. As spores are environmentally resilient, it is possible that some of the structures that make spores so resilient could protect against the damage caused by teicoplanin. It seems likely that the increased cellular stress observed in Δ*vanJ* may explain why more of these processes are induced in this strain, but further study needs to be carried out to identify whether these observations translate into morphological changes.

#### 4.2.1.6 Evidence of nitric oxide generation as a response to teicoplanin in Δ*vanJ*

The metabolic processes that are involved in nitric oxide (NO) biosynthesis are still poorly understood in *S. coelicolor*, but it is believed to be generated from the conversion of nitrite ($NO_2^-$) to NO through an unidentified process. This small compound is present in times of stress and is linked with development processes, acting as an intra- and intercellular signalling molecule [186,187]. Nitrite can be synthesised by nitrate reductases (Nar) which convert nitrate ($NO_3^-$) to nitrite. These data suggest that there could be changes in the levels of intracellular NO as there are changes in the expression of genes relating to the three *nar* clusters in *S. coelicolor*. There were significant increases in all nar genes in Δ*vanJ*, while only the third *nar* cluster along with *narH* were significantly induced in M600. This suggests that the absence of VanJ may lead to increases in the production of NO in the presence of teicoplanin. Increased intracellular NO levels could also provide an explanation as to why particular developmental processes are significantly affected in Δ*vanJ* at the 30-minute time point (**Figure 4.6**) as NO can act as a signalling molecule for some regulatory proteins such as the WhiB-like (wbl) proteins, (WhiB and WhiD) [188].

#### 4.2.1.7 Teicoplanin induces genes for antibiotic production in both strains

Many biological processes are developmentally linked, including the production of many secondary metabolites. Two of the most well studied secondary metabolites in *S. coelicolor* are the antibiotics uncecylprodigiosin (RED), and ACT which are produced as cells switch from vegetative to reproductive growth [155,189]. 20 out of 23 genes from the RED cluster were significantly upregulated in response to teicoplanin in both strains. Three genes were not DE across any of the time points, including *redG* (*SCO5897*), *redF* (*SCO5897*) and *SCO5899*. In contrast, only 6 out of 22 genes from the ACT cluster were DE in M600 compared to 16 out of 22 genes in Δ*vanJ*. *actVA2* (*SCO5077*) was the only gene that was downregulated, which occured in both strains (**Figure 4.6**). These data suggest that antibiotic production may increase

after exposure to teicoplanin, which is likely linked to the increased cellular stress caused by this drug.

## 4.2.3 Effects of teicoplanin on the expression of genes relating to the cell envelope

### 4.2.3.1 Teicoplanin causes changes in the expression of genes related to biosynthesis and maintenance of PG

Significant decreases were observed in the expression of most genes involved in PG biosynthesis in both strains within the first 30 minutes. This aligns with previous data on other cell wall targeting antibiotics. Genes involved in the earliest steps of PG biosynthesis *ddlA* (*SCO5560*), *murA* (*SCO2949*) and *murC* (*SCO6060*) were only significantly downregulated in Δ*vanJ*, while *murB* (*SCO4643*) was not significantly DE in either strain. It should also be noted that *uppS* (*SCO2509*) and the proposed lipid II flippase, *murJ* (*SCO3894*), were also only significantly downregulated in the Δ*vanJ* strain (**Figure 4.7**).

We also explored how teicoplanin affected the 21 different PBPs of S. coelicolor which in most instances are still poorly characterised [17]. Eighteen of these showed some form of DE in either strain. There were similarities between both strains including the upregulation of two class A PBPs (*SCO2897* and *SCO5039*), five class B PBPs (*SCO1875*, *SCO3156*, *SCO3157*, *SCO3771*, and *SCO3847*) and two class C PBPs (*SCO0830* and *SCO4847*) (**Figure 4.7**). These genes increased in activity by between 3.1 - 38.7 fold with *SCO1875* and *SCO3771*, showing the most significant changes in expression. Nine PBPs were only DE in Δ*vanJ* including the class A PBPs *SCO3580* and *SCO3901*; the class B PBPs, *ftsI* (*SCO2090*), *pbp2* (*SCO2608*) and *SCO4013*; and the class C PBP *dacA* (*SCO3811*). These genes were downregulated by between 2.5 - 10 fold within the first 30 minutes. *SCO5301* and *SCO7050* were also significantly induced in Δ*vanJ* by 7.4 and 4.8 fold respectively (**Figure 4.7**). The subtle differences in the expression patterns of the PBPs between the strains could suggest that there are different requirements for PBPs between the strains in how they respond to teicoplanin. The transient changes in some of the PBPs in Δ*vanJ* could also be developmentally linked to the changes observed in sporulation genes. Interestingly, the majority of the PBPs that are influenced by teicoplanin in both strains are class B and C PBPs which operate away from the membrane in the PG layer. This is likely a response to the inhibition of transpeptidation by teicoplanin in the PG layer which may result in the modification of the stem peptides in uncross-linked PG.

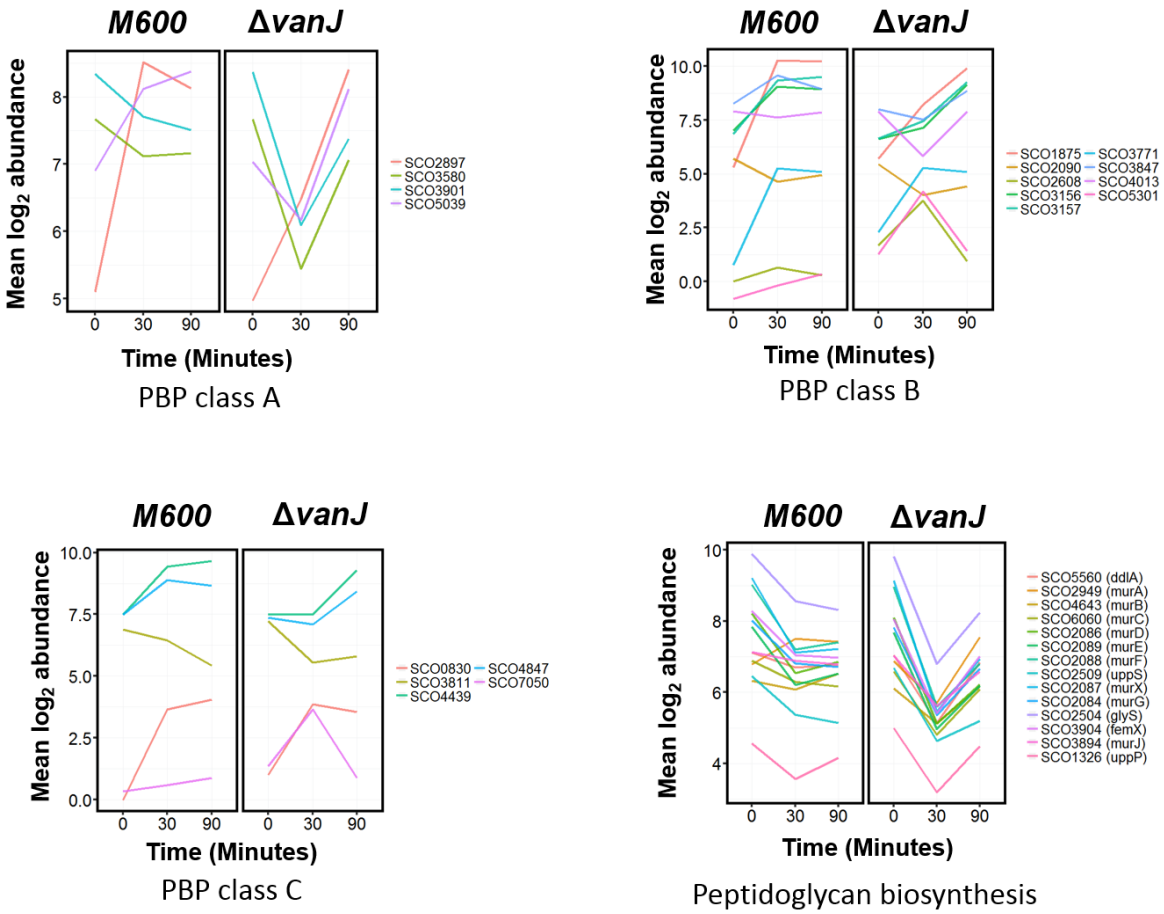



**Figure 4.7** The effects of teicoplanin on the expression of genes involved in processes that synthesise and maintain the PG layer in *S. coelicolor*. Time in minutes is along the x-axis and average log2 abundance is along the y-axis. Each point is the average of two replicates.

#### 4.2.3.2 Both Strains showed a significant increase in the expression of genes involved in transport across the membrane

Cluster 3 was significantly enriched for genes with functions relating to 'response to drug' (GO:0042493, P-value = 2.93 x $10^{-3}$), which contained four genes with predicted transporter activity. Membrane transporters have previously been implicated in the cell-envelope stress responses of *B. subtilis* [190], *S. aureus* [191] and *S. coelicolor* [149], but none of these particular genes have been implicated in resistance to glycopeptides in *S. coelicolor*. All genes enriched for this functional group were significantly upregulated in both M600 and Δ*vanJ* in response to teicoplanin (**Figure 4.8**). *SCO3824* and *SCO4359* belonged to the ATP-binding cassette (ABC) transporter superfamily and were upregulated by 24.6 and 15.2 fold in M600 and 7.3 and 4.2 fold in Δ*vanJ* respectively. *SCO2472* is a member of the *ydcF*-like superfamily of genes that also contain the *sanA* gene which has previously been found to complement a vancomycin-sensitivity mutation in *E. coli* by an unknown mechanism [192]. *SCO2472* was upregulated by an astonishing 98.4 fold in M600 and 16.2 fold in Δ*vanJ* during earlier time points. The final gene, *SCO3910*, belongs the MOP superfamily, which is found throughout all domains of life and includes MurJ. These genes are also strongly related to multidrug resistance in cells and the transport of bulky membrane carriers such as lipid II [193]. *SCO3910* increased in activity by 8.9 and 3.5 fold in M600 and Δ*vanJ,* respectively within the first 30 minutes of the time course (**Figure 4.8**). The products of these membrane transporters are likely involved in the removal or transport of toxic metabolites in response to teicoplanin, however the similarity of *SCO2472* to *sanA* and the predicted MOP domain of *SCO3910* make them attractive genes to study in *S. coelicolor* in relation to glycopeptide resistance.

#### 4.2.3.3 There are observable differences in the expression of genes involved in membrane lipid biosynthesis between the two strains.

How the glycopeptides affect the bacterial membrane is poorly understood. We therefore decided to explore the expression of the known genes involved homeostasis of the *S. coelicolor* cell membrane and its associated lipids. Firstly we observed that *SCO6759-63* and *SCO6766* from the hopanoid cluster were significantly downregulated in Δ*vanJ* (**Figure 4.9**). These molecules are synthesised from isoprenoids and have a similar role as cholesterol in eukaryotic membranes, increasing membrane packing, to form discrete microdomains referred to as lipid rafts. These are usually associated with flotillin proteins, which are known to recruit other proteins involved in secretion and cell signalling [48,71].

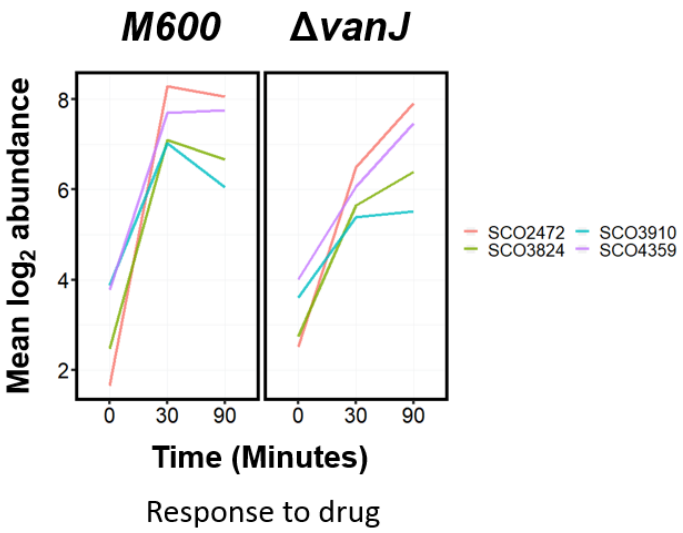



**Figure 4.8** Teicoplanin strongly induces the expression of genes with predicted transporter function that were associated with the ontology 'Response to drug'. Time in minutes is along the x-axis, and average $log_2$ abundance is along the y-axis. Each point is the average of two replicates.

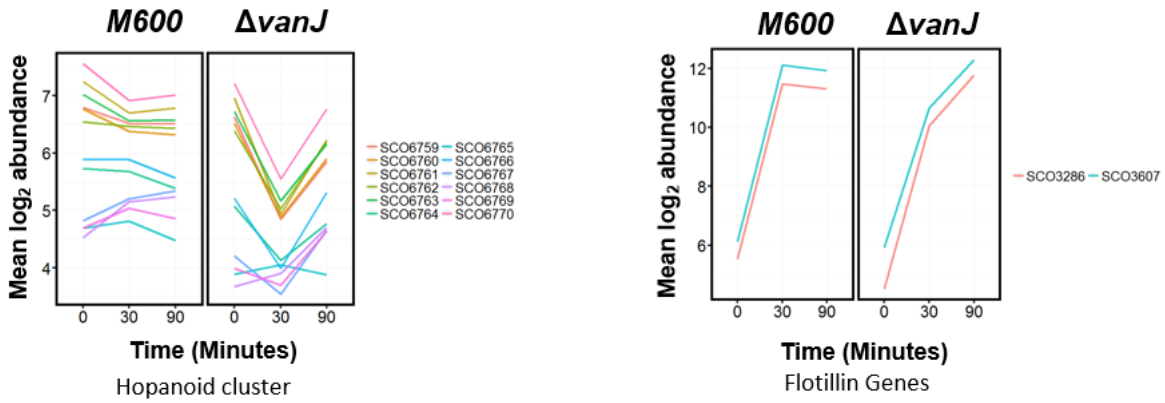



**Figure 4.9** Expression profiles for genes involved in hopanoid and flotillin biosynthesis in *S. coelicolor* are displayed. Time in minutes is along the x-axis and average $log_2$ abundance is along the y-axis. Each point is the average of two replicates.

We also found the expression of two flotillin genes, *SCO3286* and *SCO3607*, increased by up to 60 fold between the two strains at 30 minutes (**Figure 4.9**) which could highlight that teicoplanin leads to discreet changes in the architecture of the membrane, and the formation of flotillin microdomains in both strains.

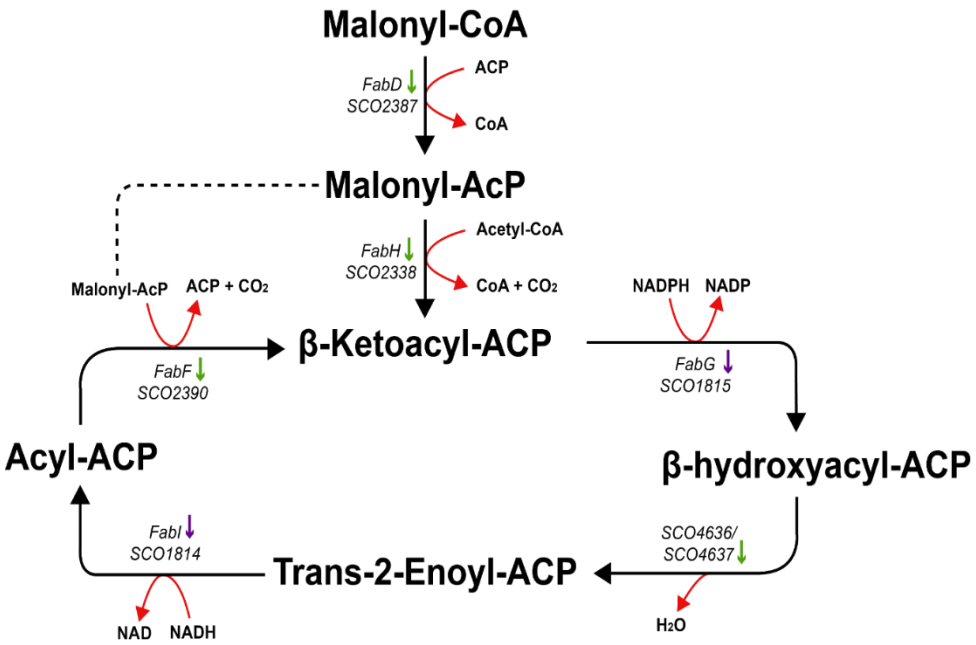
The major constituent of the bacterial membrane are phospholipids. These lipids consist of an acylated glycerophosphodiester that can be modified with many variable fatty acids and alcohol head groups. Initially, the acyl chains that make up the hydrophobic tails of phospholipids are synthesised through the fatty acid biosynthesis (FAB) pathway, which is well conserved within bacteria (**Figure 4.10A**). Teicoplanin was found to cause significant decreases in the expression of genes relating to FAB. *fabI* (*SCO1814*) and *fabG* (*SCO1815*) were downregulated in both strains by between 2.5 - 10 fold while *fabD* (*SCO2387*), *fabH* (*SCO2388*), *fabF* (*SCO2390*), *SCO4636* and *SCO4637* were only downregulated in Δ*vanJ* by between 2.5 - 10 fold within 30 minutes (**Figure 4.10B**).

The latter stages of phospholipid biosynthesis involve loading the core structure with various head groups to generate the diverse array of membrane lipids. The expression of genes that are involved in the biosynthesis of CDP-DAG (*SCO5626*), PS (*SCO6467*), PE (*SCO6468*), PpG (*SCO5753*) and CL (*SCO1389*) were all significantly downregulated in Δ*vanJ* (**Figure 4.11B**). *acylT* (*SCO1526*) and *pisA* (*SCO1527*), which encode the enzymes that synthesise Ac-$PIM_{1-2}$ and PIP (**Figure 4.11A**), were significantly downregulated in both strains. In contrast, the genes that are involved in ornithine biosynthesis, *olsA* and *olsB* (**Figure 4.11A**), were significantly upregulated by 7.3 - 8.9 fold in the Δ*vanJ* mutant within the first 30 minutes (**Figure 4.11B**). Together these data suggest that teicoplanin likely leads to changes in the overall composition of the membrane in Δ*vanJ* with particular decreases in genes relating to positively charged phospholipids, while there are increases in the zwitterionic, ornithine lipid.

#### 4.2.3.4 There is evidence for phospholipid remodelling occurring in the cell membrane in response to teicoplanin

The membrane surface is a dynamic structure that is continuously remodelled to adapt to novel environmental conditions. As previously mentioned, the gene encoding the tRNA synthetase, *mprF*, is strongly upregulated in both strains in response to teicoplanin. The product of this gene can add lysyl-groups to PpG to decrease the net negative charge of the bacterial membrane (**Figure 4.12**).

**A**




**B**

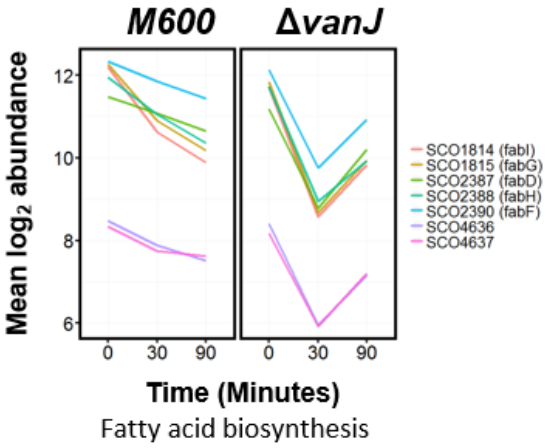



**Figure 4.10** (A) Metabolic pathway involved in fatty acid biosynthesis. Solid arrows indicate steps with known genes and dashed arrows indicate unknown mechanisms with no candidate gene identified. Coloured arrows next to each gene indicate the direction of expression after exposure to teicoplanin, with purple arrows indicating DE in both strains and green arrows indicating DE in just the Δ*vanJ* mutant. (B) Expression profiles for genes encoding the enzymes involved in FAB process in *S. coelicolor* is presented below. Time in minutes is along the x-axis, and average $\log_2$ abundance is along the y-axis. Each point is the average of two replicates.

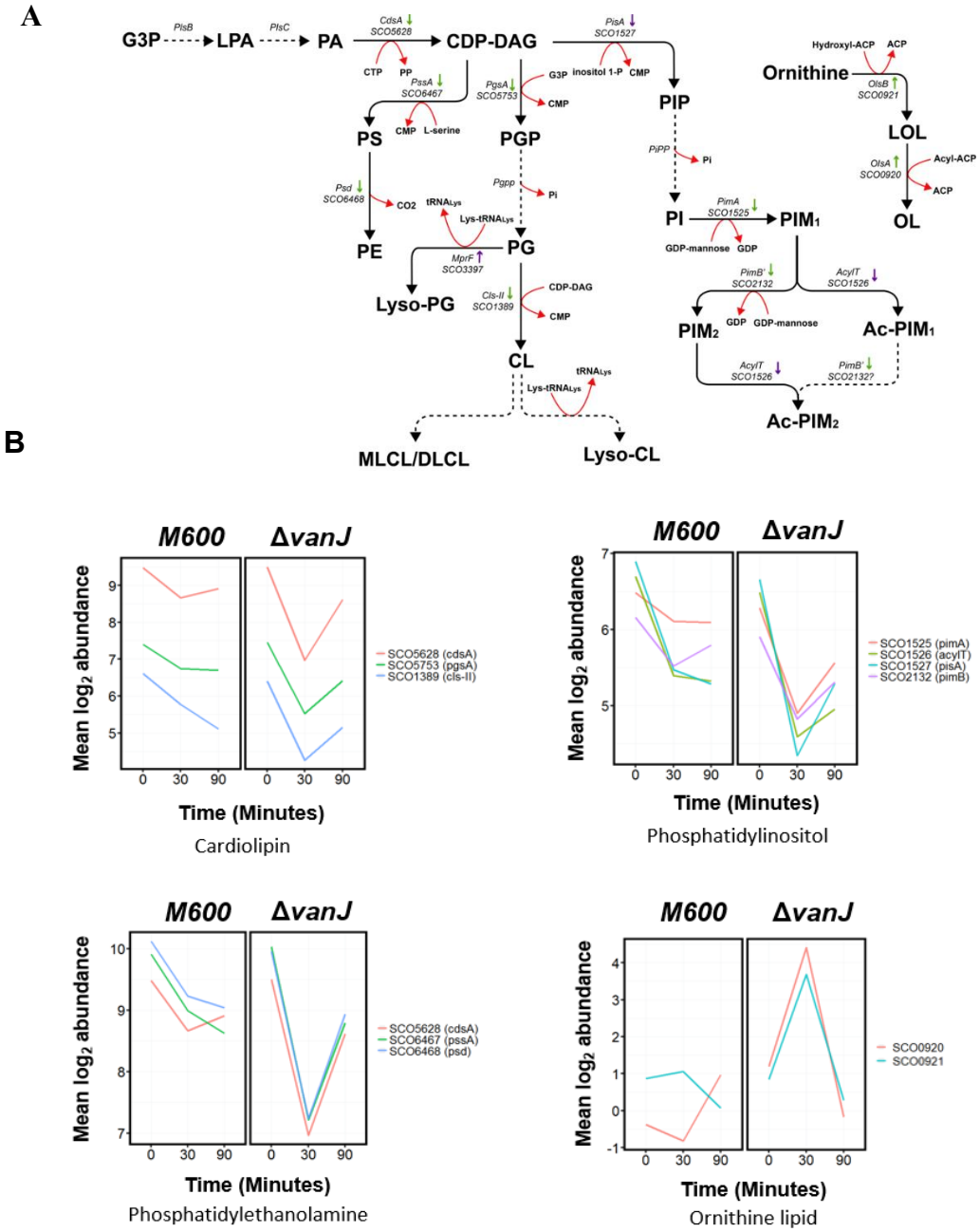



**Figure 4.11** (A) Metabolic pathways involved in membrane lipid biosynthesis. Solid arrows indicate characterised steps with known genes and dashed arrows indicate unknown mechanisms. Coloured arrows next to each gene indicate the direction of expression after exposure to teicoplanin. Purple arrows indicate DE in both strains. Green arrows indicative DE in just the Δ*vanJ* mutant. (B) Expression profile for genes encoding the enzymes involved in membrane phospholipid biosynthesis and ornithine metabolism in *S. coelicolor*. Time in minutes is along the x-axis and average log2 abundance is along the y-axis. Each point is the average of two replicates.

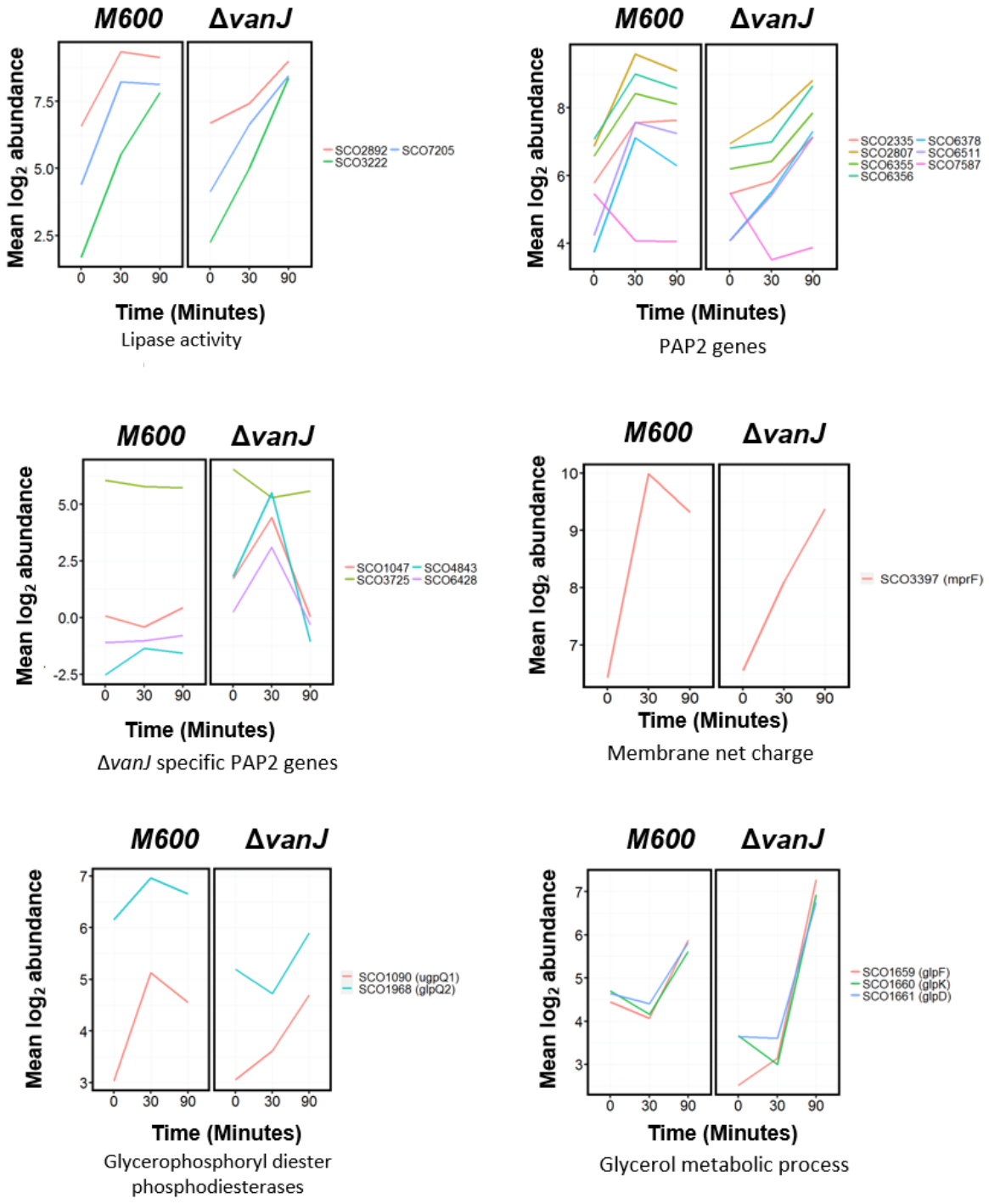



**Figure 4.12** Teicoplanin effects the expression of genes that are involved in degrading and modifying the bacterial membrane. The expression profile for genes in the glycerol operon is also shown. Time in minutes is along the x-axis, and average $log_2$ abundance is along the y-axis. Each point is the average of two replicates.

There are many other enzymes which also can modify the structure of phospholipids, including a large family of PAP all with a characteristic PAP2 domain that cleave phosphodiester bonds (**Figure 4.13**). The diversity of these enzymes in the metabolic processes surrounding phospholipids makes them an appealing target for study because of their possible interaction with lipid II and other molecules with a phosphodiester bond. A BLAST search predicted that 18 genes had the PAP2 domain within the *S. coelicolor* genome. Only 12 of these were found to be DE (**Figure 4.12**). Seven were universal to both strains including *SCO2335*, *SCO2807*, *SCO6355*, *SCO6356*, *SCO6378* and *SCO6511,* all of which showed an increase in expression by approximately 4 – 11 fold over the 90 minutes. Interestingly, the expression of *SCO7587* decreased in both strains by 2.5 - 3.3 fold in the first 30 minutes. *SCO1047, SCO3725, SCO4843* and *SCO6428* were all specifically DE in Δ*vanJ*. *SCO1047*, *SCO4843* and *SCO6428* were all upregulated by 6.5 - 14 fold, and, *SCO3725* was downregulated by 2.5 fold during the first 30-minutes. The transient induction of some of these genes in Δ*vanJ* reminiscent of the induction seen in some developmental genes discussed in in **Section 4.2.2** could suggest that these genes are linked. The converse expression of *SCO7587* in comparison to the other PAP may also suggest that the products of these genes have distinct roles in the membrane in response to teicoplanin and the resulting products of these enzymes may be involved in the adaptive response to the external environment.

In addition to the PAPs, phospholipases also play a vital role in the metabolism of phospholipids, as well as host defence and signal transduction in many bacteria [194–196]. Although poorly understood in *S. coelicolor*, they can be subdivided into four classes ($A_1$, $A_2$, C or D) based on the cleavage site on the phospholipid structure (**Figure 4.13**). Cluster 3 was enriched for genes with the functional group 'lipase activity' (GO:0016298, P-value = 5.0 x $10^{-2}$), encompassing three genes with predicted hydrolysing activity (**Figure 4.12**). *SCO3222* showed strong induction (17-19 fold) during the time course in both strains and belongs to the phospholipase A (PLA) class of phospholipases, sharing 100% identity with $PLA_2$ from *Streptomyces violaceoruber*. $PLA_2$ enzymes cleave the 2-acyl bond of glycerophospholipids to produce a fatty acid and an LPA [194,197]. Additionally, the expression of *SCO2892* gene also increased by 4.7 - 7.8 fold in both strains and is predicted to belong to the thioesterase I/protease I/lysophospholipase (TAP) protein family [198]. This promiscuous family is known to have broad specificity for mono-acyl phospholipids such as LPA which could suggest the products of these genes work synergistically to cleave the 1 and 2-acyl chains from phospholipids in response to teicoplanin (**Figure 4.12**).

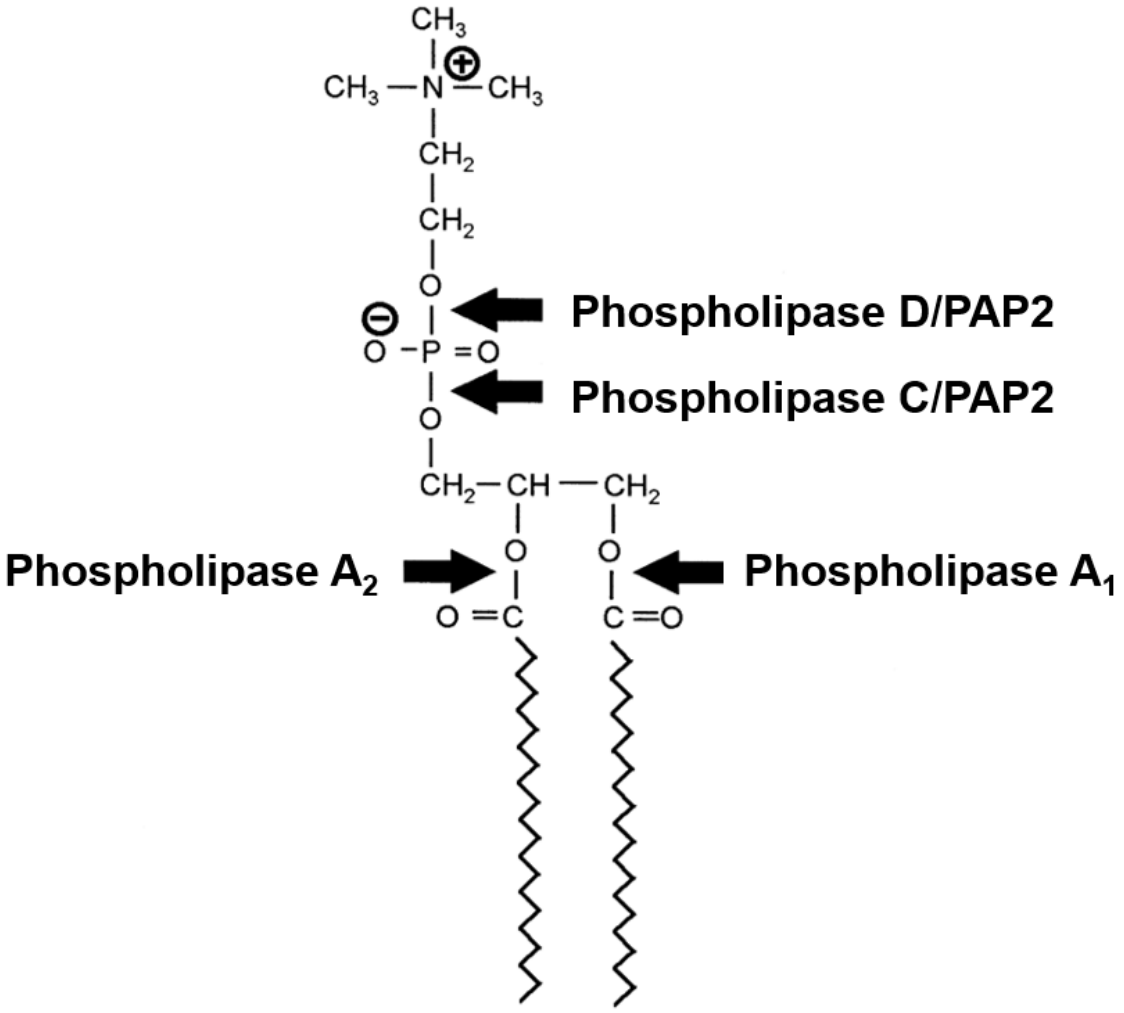



**Figure 4.13** Diagram showing the cleavage sites of phospholipase classes and PAP2 enzymes on phospholipids. Adapted from Titball (1998).

Most of the studies carried out to date on the processes involved in the recycling of phospholipids have used *E. coli* as a model. The complex processes that occur during the complex lifecycle of *S. coelicolor* are only just beginning to be understood. We know from *E. coli* that phospholipids are degraded to produce glycerophosphodiesters which differ based on their alcohol moiety. These can then undergo further degradation through the action glycerophosphodiester phosphodiesterases (GDPD), yielding sn-glycerol-3-phosphate (G3P) and the corresponding alcohol. GlpQ carries out this reaction in the *E. coli* periplasm and UgpQ in the cytosol, but both show broad specificity for glycerophosphodiesters [199,200]. *S. coelicolor* has three homologues of *glpQ* (*glpQ1-3*) and four of *ugpQ* (*ugpQ1-4*), but other than some evidence for substrate-specific regulation in *glpQ1* and *glpQ2*, little is known about the specific role of these GDPDs [201]. Interestingly, only two of the GDPDs were differentially expressed in our dataset. *ugpQ1* was significantly upregulated by 3.6 fold in M600, and *glpQ2* was significantly upregulated by 3 fold in Δ*vanJ* (**Figure 4.12**). Together these data suggest that there could be increased GDPD activity occurring in response to teicoplanin, but why there are differences between the two strains is still unclear.

In *E.coli*, the final stages of phospholipid metabolism involve transporting G3P back into the cell via the transporter GlpT, where G3P can be converted to dihydroxyacetone phosphate (DHAP) and shuttled into glycolysis or used for further rounds of membrane biosynthesis [62]. *S. coelicolor* lacks a G3P transporter so G3P must instead be dephosphorylated to generate glycerol, before being transported across the membrane (personal communication with Colin P. Smith). We found that Cluster 7 was enriched for genes with functional groups involved in the 'glycerol metabolic process' (GO:0006071, P-value = 2.38 x $10^{-4}$). This included the three genes in the glycerol operon, *glpFKD*. These genes encode a membrane facilitator that increases the rate of glycerol diffusion across the membrane, a kinase that converts glycerol to G3P and a dehydrogenase that converts G3P to DHAP [202]. The operon was significantly induced in both strains after exposure to teicoplanin and showed increases in activity between 2.7 - 18 fold (**Figure 4.12**). The operon is known to be induced by G3P, suggesting that exposure to teicoplanin leads to increased concentrations of G3P or glycerol in the surrounding media [203]. We speculate that if a response to teicoplanin leads to increased phospholipid degradation through the action of some of the gene products discussed in this section, it is plausible that the increased activity of the glycerol operon is due to increased concentration of breakdown products due to the action of the enzymes discussed here. And, as far as we're

aware, this is currently an unexplored mechanism of teicoplanin and the glycopeptide antibiotics.

## 4.3 Summary

In this section, we presented the genome-wide comparisons of genes in *S. coelicolor* after exposure to teicoplanin. There were expected key similarities between both strains including increased activity in the regulators of the cell envelope, osmotic and oxidative stress responses in *S. coelicolor*. This reaffirms the data previously presented on the overlap between global stress response systems induced by antibiotics that target the cell wall in *S. coelicolor* [149]. Unexpectedly, the $\sigma^E$ signal transduction system, the cell envelope stress regulon and the induction of several STKs were induced more slowly in Δ*vanJ*. This suggests that the weak induction of VanJ in M600 is enough to affect the transcription of other genes and may highlight that VanJ is involved in induction of the $\sigma^E$ signal transduction system.

Currently the ligand for the CseC SK is unknown, but the system is triggered by a wide range of antibiotics and enzymes that affect the integrity of the cell wall [151]. It has been suggested that either a common intermediate produced by cell wall targeting antibiotics may be the cognate ligand or instead, a biophysical change in the cell envelope is detected. The former explanation makes more logical sense with our data and would provide an explanation as to why induction of $\sigma^E$ signal transduction system is delayed rather than inhibited in the Δ*vanJ* strain. VanJ may therefore be involved in generating this ligand in the presence of teicoplanin. However, if other pathways also generate this ligand, there would still be observable increases in these pathways, albeit at a slower rate. This delayed response appears to also be tied with the onset of the production of cellular stress markers that were observed during earlier time points in Δ*vanJ*. This indicates that even the limited induction of *vanJ* in M600 cells after exposure to teicoplanin, is profoundly effective at reducing the overall effects of cellular stress caused by teicoplanin.

There were clear indications that processes relating to growth and development were more affected by teicoplanin in Δ*vanJ*. Notably, genes involved in ribosome biogenesis, protein translation, carbon utilisation, DNA replication, and cell envelope biosynthesis were all significantly downregulated during earlier time points but almost returned to pre-treatment levels by 90 minutes. Simultaneously, in Δ*vanJ*, there were transient increases in genes relating to the generation of the signalling molecule, NO and genes such as the *whiE* genes that relate

to spore maturation. It is not entirely clear whether these transcriptomic changes relate to physiological outcomes, but active growth is likely reduced to prevent cellular death by teicoplanin. What we do know is that WhiB and WhiD are involved in controlling hyphal growth and thickening the spore wall respectively, and the former makes up part of the cell envelope stress regulon [150]. Although speculative, in the presence of teicoplanin and particularly in the absence of VanJ, these processes may reduce the damaging effects of teicoplanin through modifying cell ultrastructure [185], however further study needs to be carried out to confirm this.

Most genes involved in the biosynthetic processes that convert histidine to arginine were strongly induced in both strains. Arginine has been implicated in the oxidative and salt stress responses in *S. coelicolor* and other bacteria in previous studies and is likely a metabolic precursor for ectoine and the polyamines, which can act as osmoprotectants under severe salt stress [170,171,204]. Therefore, it is likely that arginine could provide a protective role against teicoplanin and the stress it exerts on the cell.

Finally, teicoplanin appears to affect genes involved in the homeostasis of membrane lipids. Many genes involved in lipid modification and degradation were upregulated, while genes involved in the synthesis of negatively charged lipids were downregulated. There was also strong activity seen in the glycerol operon which is under strict catabolite repression in the absence of its positive inducer, G3P which is a breakdown product of phospholipids [203]. From this information, it seems plausible that teicoplanin increases the rate of phospholipid turnover and G3P may accumulate in the growth media. It is also likely that S. coelicolor salvages this metabolite for further rounds of phospholipid biosynthesis or glycolysis through the activity of the glycerol operon. To our knowledge, this is currently an unreported mechanism of glycopeptide antibiotics, and so would be an exciting avenue to explore further in relation to antibiotic resistance.

In the following sections, we characterise some of the genes in this section to identify which might be essential for resisting the effects of teicoplanin, and build a better model of how teicoplanin might interact with the cell surface, and potentially provide future drug targets for development.

# Chapter 5. Functional Analysis of Candidate Genes to Identify Their Impact on The Phenotype of *S. coelicolor* to Various Antimicrobial Agents

# 5.1 Introduction

There were a substantial number of genes identified in Chapter 4 that were differentially expressed. Many of these genes appeared to relate to broader stress responses, but several groups of genes with previously uncharacterised roles in glycopeptide resistance were identified that had localised functions relating to the cell envelope. This section will tackle the question of whether any of these genes functionally contribute to antibiotic resistance towards a library of antibiotics, or if the genes likely represented a collective stress response caused by teicoplanin. To understand any relationship between AMR and these target genes, an antibiotic library was selected encompassing antibiotics that target different stages of PG biosynthesis, the cell membrane, protein synthesis and DNA replication (Summarised in **Table 5.1**). A small selection of glycopeptide antibiotics were also used for further characterisation of genes that demonstrated interesting phenotypes against the general antibiotic library. These glycopeptide antibiotics had a variety of secondary and tertiary modifications, including various glycosylation patterns and lipophilic side chains (**Figure 5.1**).

To test for changes in phenotype, the agar dilution method is the 'golden standard' for measuring the sensitivity of microorganisms to antibiotics and has been particularly important when assessing the efficacy of novel or modified agents in comparison to reference drugs. The method can also be used to compare the sensitivity of different species or strains of microorganisms that may have different genetic backgrounds in order to identify phenotypic changes caused by mutations or the expression of particular genes. The European Committee for Antimicrobial Susceptibility Testing (EUCAST) first established the procedures to record the MIC of bacteria accurately in the early 2000s, with minor amendments made in following years [132,205]. The method involves agar plates with serial dilutions of the selected antibiotic that are then seeded with known concentrations of the target bacteria usually in 5 - 10 µL aliquots. Either positive or negative changes in sensitivity indicate that the independent variable being tested has influenced the phenotype. These studies can inform which modifications are the most successful and can aid in future drug development.

The purpose of this study was to identify which genes are involved in response to teicoplanin by genetically manipulating *S. coelicolor* M600 and Δ*vanJ* background strains to identify any changes in the phenotype.

**Table 5.1** List of antibiotics used in this study and their bacterial targets.

| Antibiotic | Target | Effect | Ref. |
|---|---|---|---|
| Glycopeptides:<br>• Vancomycin<br>• Ristocetin<br>• Chloroeremomycin<br>• Balhimycin | Bind lipid II and inhibit transglycosylation | Weakens the cell wall and makes cells susceptible to lysis | [102] |
| Lipoglycopeptides:<br>• Teicoplanin<br>• Oritavancin<br>• Dalbavancin<br>• Telavancin | Bind lipid II and inhibit transglycosylation and transpeptidation | Weakens the cell wall and makes cells susceptible to lysis | [206] |
| Ramoplanin | Binds both MurG and transglycosylases | Weakens the cell wall and makes cells susceptible to lysis | [207] |
| Carbenicillin | Bind to PBPs, blocking transpeptidation | Weakens the cell wall and makes cells susceptible to lysis | [208] |
| D-cycloserine | Inhibits D-Ala-D-Ala ligase | Weakens the cell wall and makes cells susceptible to lysis | [209] |
| Flavomycin | Inhibits transglycosylases | Weakens the cell wall and makes cells susceptible to lysis | [210] |
| Nisin | Binds to the lipid pyrophosphate | Creates pores in the membrane causing depolarisation | [211] |
| Duramycin | Binds to PE | Creates pores in the membrane causing depolarisation | [212] |
| Novobiocin | Binds to DNA gyrase | Inhibiting DNA replication | [213] |
| Kanamycin | Binds to the ribosome | Inhibits protein synthesis | [22] |

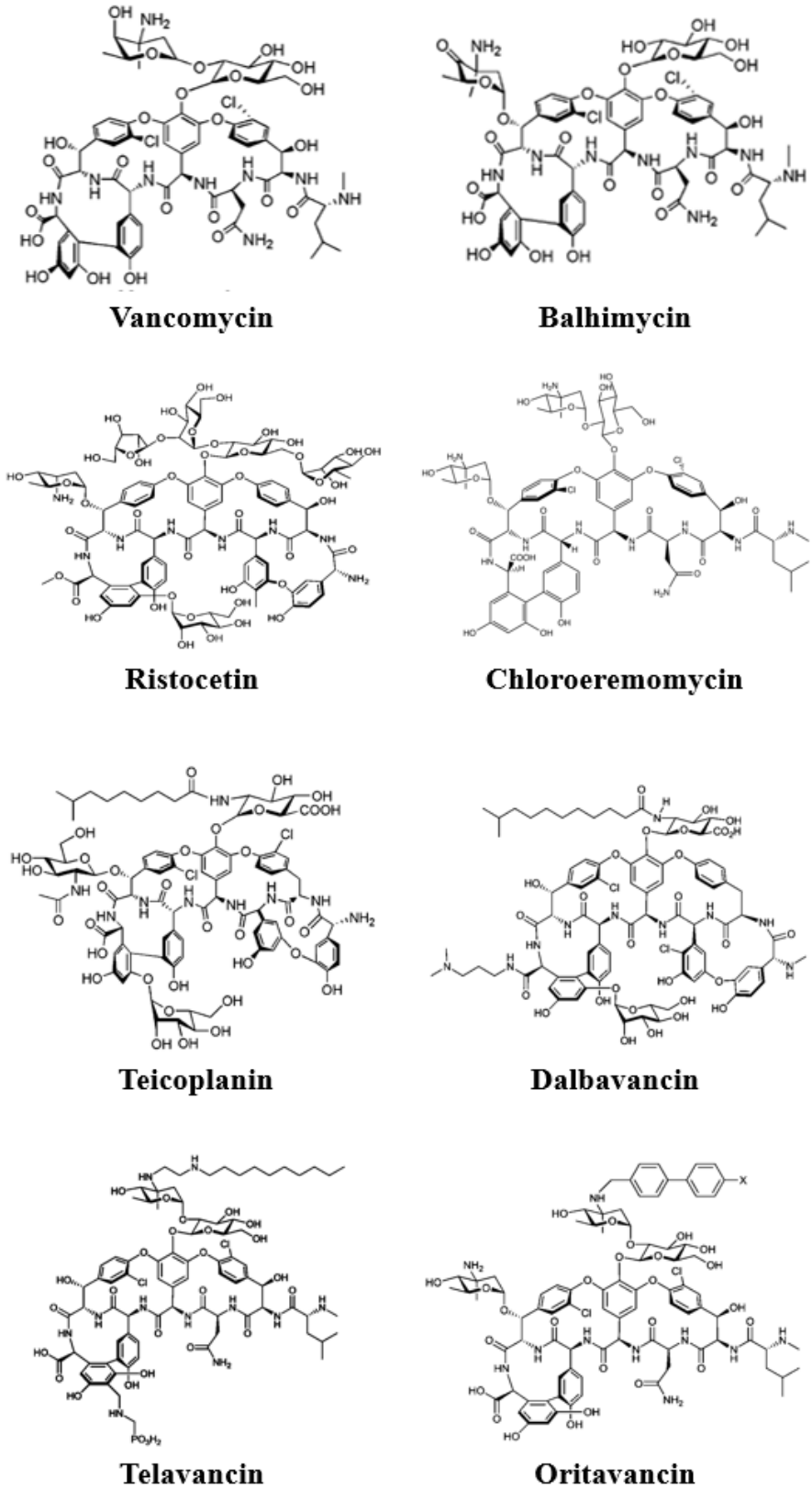



**Figure 5.1** Chemical structures of the glycopeptide antibiotics used in this study. The four lipoglycopeptide antibiotics are displayed at the bottom showing their lipophilic side chains.

The two most common methods for understanding the function of individual genes are through the generation of gene overexpression strains and gene knockout mutants. Having both phenotypes is essential for assigning function to a particular gene and can indicate the importance of specific genes in different environmental conditions. Two major vector systems were employed in this work. The first was the integrative pIJ10257 plasmid (**Figure 5.2**) which strongly expresses genes that are put under the control of its *ermE*-promoter (*ermE*-p) adjacent to a multiple cloning site [37]. This promoter originates from *Saccharopolyspora erythrea* [214] and allows for high-level expression of genes in *S. coelicolor* and has previously been used to demonstrate that *vanJ* is required for teicoplanin resistance [115].

The second system employed in this project is the pCRISPomyces-2 plasmid that has been engineered with a codon optimised *cas9* gene for *Streptomyces* [130]. The plasmid also has a scaffold sequence that facilitates binding to the Cas9 protein and a *BbsI* cut site where a short nucleotide sequence can be inserted next to the scaffold to produce a gRNA. This gRNA is what functions to guide the Cas9 protein to a specific DNA sequence where it cuts the DNA, causing a double-stranded break (DSB). The plasmid can also be engineered with a homologous region flanking the DSB so that when the cell tries to repair the DNA damage, the homologous region is used as a template for repair and the DNA sequence is modified [215]. A second CRISPR plasmid was also used in this work to try and generate knockout mutants, and both are shown in **Figure 5.3**.

# 5.2 Results

## 5.2.1 Selection of genes to construct overexpression plasmid DNAs in *S. coelicolor*

A total of ten genes were selected to test for their ability to confer antimicrobial resistance to *S. coelicolor* wild type M600 and Δ*vanJ* mutant strains. As there was no way of predicting which genes would influence the sensitivity of *S. coelicolor* to the library of antibiotics, genes were chosen from several groups of genes involved in different aspects of the cell envelope. This included, seven PAP2 genes, the three glycerol operon genes, and two genes enriched with the 'response to drug' functional group, and all successfully cloned in pIJ10257 and introduced into M600 and Δ*vanJ* background.

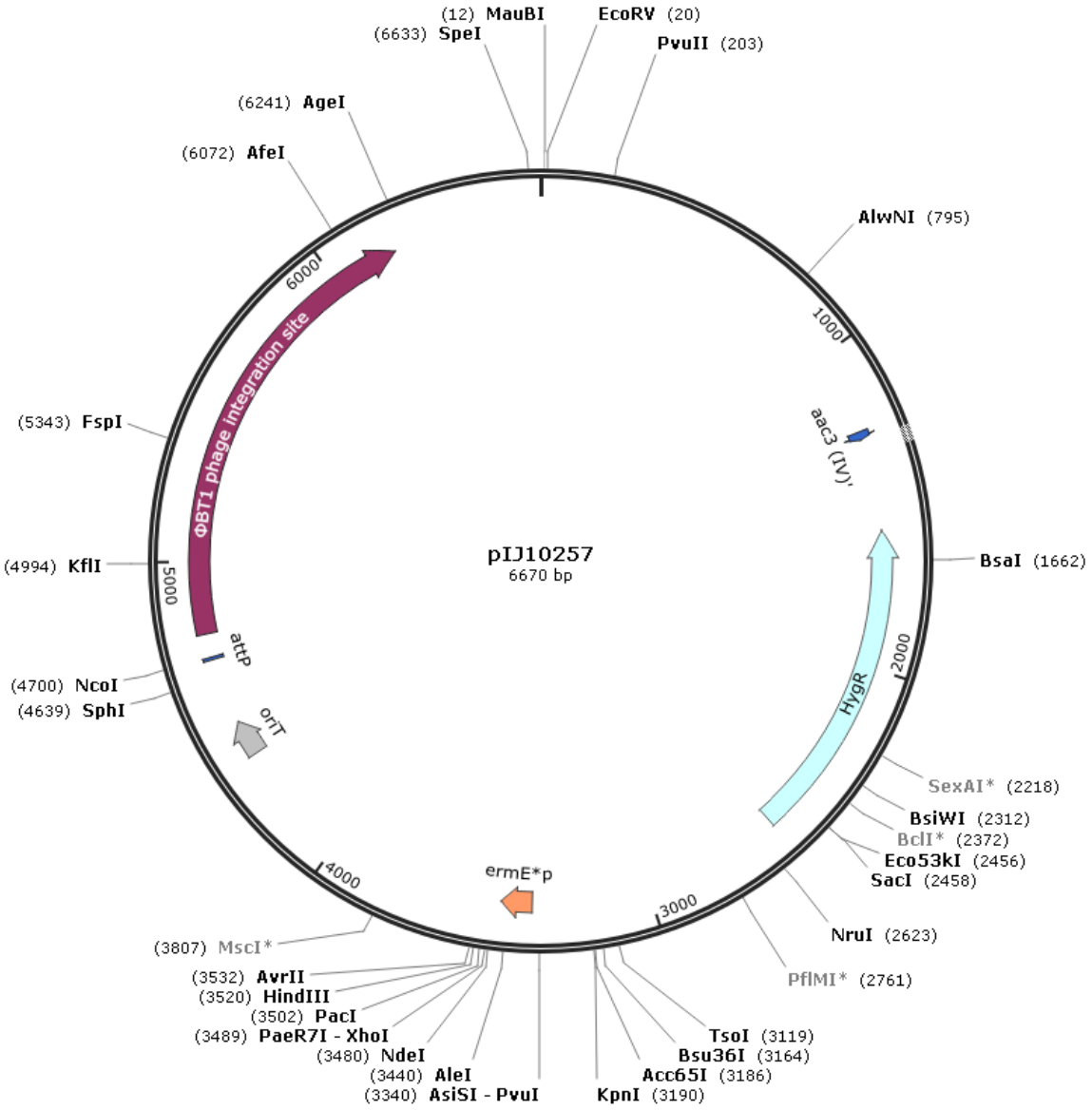



**Figure 5.2** Plasmid map for the integrative *Streptomyces* vector, pIJ10257.

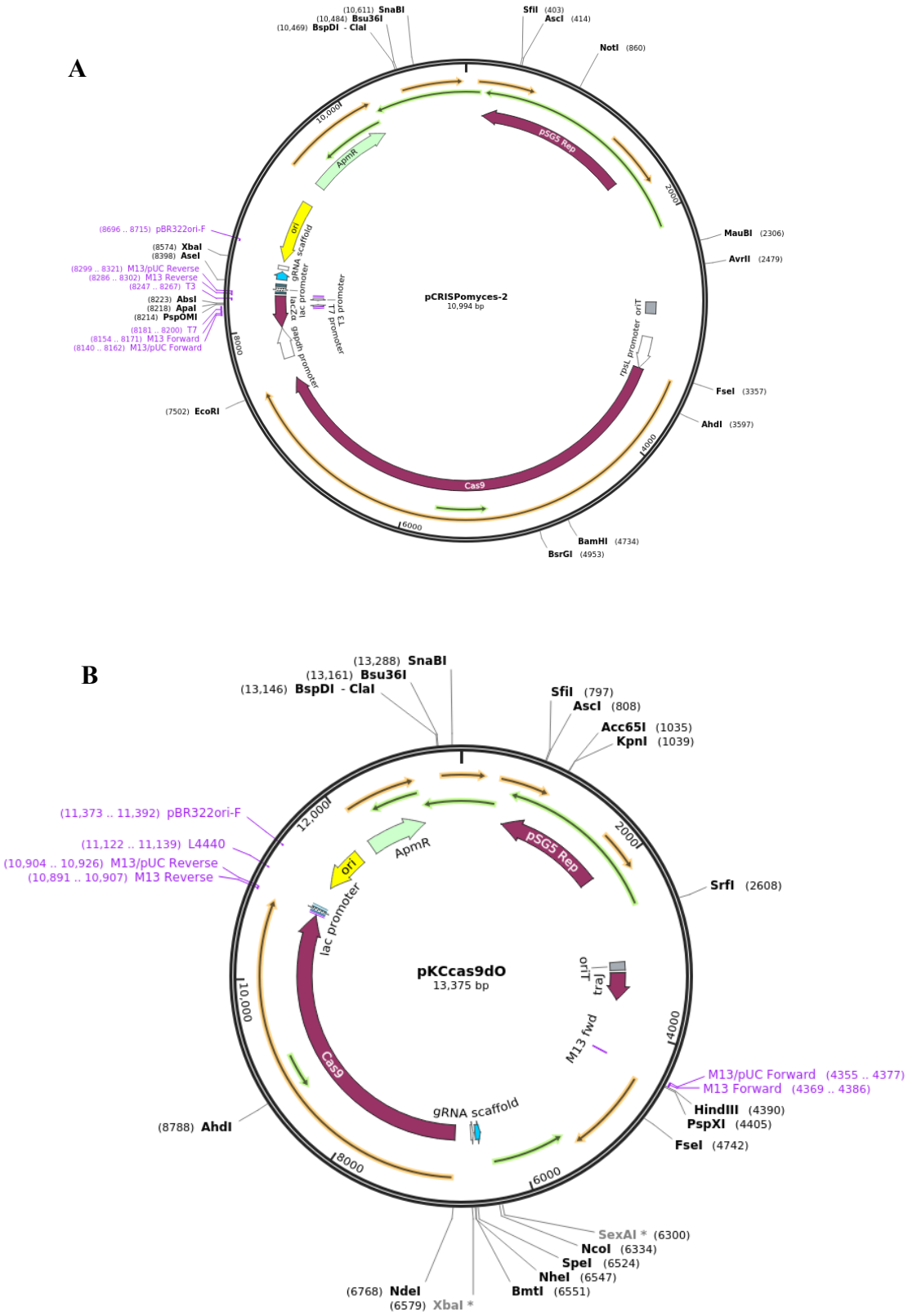



**Figure 5.3** Plasmids maps for the two CRISPR plasmids pCRISPomyces-2 (A) and pKCcas9dO (B) used in this study.

Two of the PAP genes, *SCO6355* and *SCO6356* were adjacent in an operon that also contained *SCO6357* which belongs to the widespread *dedA*/*tvp38* family of putative proton transporters that are important for sustaining the membrane potential in *E. coli* [216–218]. The operon also contained a predicted TCS *SCO6354-53*, indicating the potential for autoregulation. Both *SCO6357* and *SCO6354-53* were also introduced into *S. coelicolor* and are discussed later in the chapter.

Several trials attempted to clone *SCO2335* into pIJ10257, but all resulting plasmids were found to contain the same substitutions at 743bp (T → C) and 2077bp (T → C) which would result in a conversion from a valine to an alanine and a conversion of a tryptophan to an arginine, respectively. Polymerases with proof-reading capabilities were used in the amplification of DNA, so it was unlikely that this was caused by errors in the replication. It was possible that this mutation is a sequence variant in M600 that is not found in the published M145 strain of *S. coelicolor*. All other sequences aligned perfectly with the annotated *S. coelicolor* genomic DNA sequences. All resulting plasmids used for this study are summarised and demonstrated in **Figure 5.4**. Integration of each resulting plasmid into the genomic DNA of the background strain was confirmed through colony PCR. The CFU/µL was then used to estimate the volume of spores needed to make the selected concentrations used for MIC testing. The CFU/µL of each strain is listed in **Table 5.2**.

Several failed attempts to clone both CRISPR plasmids were carried out, and it was eventually decided that both plasmids must be toxic to *S. coelicolor* as the empty vectors produced no successful exconjugants. These experiments were abandoned and since this work, several new S. coelicolor optimised plasmids have been developed which are mentioned in the discussion.

### 5.2.2 MIC test of engineered *S. coelicolor* strains against the library of antibiotics

*S. coelicolor*, like many other actinomycetes, exhibits high-level intrinsic resistance to many antibiotics. This has likely resulted from an evolutionary arms race between neighbouring organisms within the complex soil microbiome where organisms are competing for the same nutrients. Both control strains of sc10257m and sc10257v showed the same phenotype against most antibiotics tested with the exception of kanamycin and vancomycin (**Table 5.3)**. The sc10257v strain was vastly more resisatant to kanamycin which is likely due to the apramycin resistance cassette used to inactivate the *vanJ* gene in Δ*vanJ* background strain.


**Figure 5.4** Plasmid maps for pIJ10257 derivative plasmids.

**Table 5.2** CFU/µL concentrations for each strain conjugated with pIJ10257 derivative plasmids.

| Strain | Average CFU | Plated | Dilution | Stock (CFU/µL) |
|---|---|---|---|---|
| sc10257m | 64 | 0.9 | 1.00E+08 | 7.15E+06 |
| sc10257v | 75 | 0.9 | 1.00E+08 | 8.33E+06 |
| scSCO1659m | 46 | 0.9 | 1.00E+08 | 5.07E+06 |
| scSCO1659v | 54 | 0.9 | 1.00E+08 | 6.00E+06 |
| scSCO1660m | 57 | 0.9 | 1.00E+08 | 6.33E+06 |
| scSCO1660v | 69 | 0.9 | 1.00E+08 | 7.63E+06 |
| scSCO1661m | 120 | 0.9 | 1.00E+08 | 1.33E+07 |
| scSCO1661v | 126 | 0.9 | 1.00E+08 | 1.40E+07 |
| scSCO2335m | 36 | 0.9 | 1.00E+08 | 4.04E+06 |
| scSCO2335v | 322 | 0.9 | 1.00E+07 | 3.58E+06 |
| scSCO2472m | 139 | 0.9 | 1.00E+08 | 1.54E+07 |
| scSCO2472v | 29 | 0.9 | 1.00E+08 | 3.19E+06 |
| scSCO2807m | 491 | 0.9 | 1.00E+07 | 5.46E+06 |
| scSCO2807v | 174 | 0.9 | 1.00E+07 | 1.93E+06 |
| scSCO3910m | 81 | 0.9 | 1.00E+08 | 9.04E+06 |
| scSCO3910v | 102 | 0.9 | 1.00E+08 | 1.13E+07 |
| scSCO6355m | 143 | 0.9 | 1.00E+08 | 1.59E+07 |
| scSCO6354-53m | 49 | 0.9 | 1.00E+08 | 5.48E+06 |
| scSCO6354-53v | 90 | 0.9 | 1.00E+07 | 9.96E+05 |
| scSCO6355v | 33 | 0.9 | 1.00E+08 | 3.63E+06 |
| scSCO6356m | 27 | 0.9 | 1.00E+08 | 3.04E+06 |
| scSCO6356v | 63 | 0.9 | 1.00E+08 | 7.00E+06 |
| scSCO6357m | 120 | 0.9 | 1.00E+07 | 1.34E+06 |
| scSCO6357v | 88 | 0.9 | 1.00E+07 | 9.78E+05 |
| scSCO6511m | 53 | 0.9 | 1.00E+08 | 5.85E+06 |
| scSCO6511v | 54 | 0.9 | 1.00E+07 | 6.00E+05 |

**Table 5.3** List of MIC values for all overexpressing constructs used in this study. Values are the average of three biological replicates and values indicate µg/mL of antibiotic needed to prevent any growth on the media. The two control strains were modified to contain the empty pIJ10257 vector. T (teicoplanin), V (Vancomycin), R (ramoplanin), C (carbenicillin), D-c (D-cycloserine), F (flavomycin), Ni (nisin), Du (duramycin), N (novobiocin), and K (Kanamycin).

| Strain | T | V | R | C | D-c | F | Ni | Du | N | K |
|---|---|---|---|---|---|---|---|---|---|---|
| Controls | | | | | | | | | | |
| sc10257m | 0.5 | 80 | 4 | >512 | 256 | >128 | 256 | 8 | 16 | <0.125 |
| sc10257v | 0.5 | 40 | 4 | >512 | 256 | >128 | 256 | 8 | 16 | 0.5 |
| Glycerol operon | | | | | | | | | | |
| scSCO1659m | 0.5 | 80 | 4 | >512 | 256 | >128 | 256 | 8 | 16 | <0.125 |
| scSCO1659v | 0.25 | 80 | 4 | >512 | 256 | >128 | 256 | 8 | 16 | 0.5 |
| scSCO1660m | 0.5 | 80 | 4 | >512 | 256 | >128 | 256 | 8 | 16 | <0.125 |
| scSCO1660v | 0.25 | 80 | 4 | >512 | 256 | >128 | 256 | 8 | 16 | 0.5 |
| scSCO1661m | 0.5 | 80 | 4 | >512 | 256 | >128 | 256 | 8 | 16 | <0.125 |
| scSCO1661v | 0.5 | 80 | 4 | >512 | 256 | >128 | 256 | 8 | 16 | 0.5 |
| Response to drug | | | | | | | | | | |
| scSCO2472m | 0.25 | 40 | | >512 | | >128 | | | | |
| scSCO2472v | 0.25 | 40 | | >512 | | >128 | | | | |
| scSCO3910m | 0.25 | 40 | | >512 | | >128 | | | | |
| scSCO3910v | 0.25 | 40 | | >512 | | >128 | | | | |
| PAP2 | | | | | | | | | | |
| scSCO2335m | 0.5 | 80 | 2 | >512 | 128 | >128 | 256 | 8 | 16 | <0.125 |
| scSCO2335v | 0.25 | 80 | 2 | >512 | 128 | >128 | 128 | 8 | 16 | 0.5 |
| scSCO2807m | 0.5 | 80 | 2 | >512 | 128 | >128 | 256 | 8 | 16 | <0.125 |
| scSCO2807v | 0.25 | 80 | 2 | >512 | 64 | >128 | 128 | 8 | 16 | 0.5 |
| scSCO6355m | 0.5 | 20 | 4 | >512 | 256 | >128 | 256 | <2 | 16 | <0.125 |
| scSCO6355v | 0.5 | 10 | 4 | >512 | 256 | >128 | 128 | <2 | 16 | 0.5 |
| scSCO6356m | 0.5 | 80 | 2 | >512 | 128 | >128 | 128 | 8 | 16 | <0.125 |
| scSCO6356v | 0.5 | 80 | 2 | >512 | 128 | >128 | 128 | 8 | 16 | 0.5 |
| scSCO6511m | 0.5 | 80 | 2 | >512 | 128 | >128 | 128 | 8 | 16 | <0.125 |
| scSCO6511v | 0.25 | 80 | 2 | >512 | 128 | >128 | 128 | 8 | 16 | 0.5 |

Both apramycin and kanamycin are aminoglycoside antibiotics, and the resistance gene likely provides resistance towards both. The increased sensitivity toward vancomycin observed in sc10257v has also been seen in other Δ*vanJ* strains conjugated with pIJ10257, although it is unclear why this occurs [115]. Both strains exhibited high-level resistance to both carbenicillin and flavomycin, and the true MIC exceeded the antibiotic concentrations that were possible to test (**Table 5.3**).

#### 5.2.2.1 Overexpression of the glycerol catabolic cluster affected the MIC towards vancomycin and teicoplanin

The glycerol cluster includes the glycerol operon, consisting of *glpFKD*, which are regulated by the negative regulator, *gylR* (**Figure 5.5A**). These genes are involved in the uptake and metabolism of glycerol from the surrounding media (**Figure 5.5B**). Overexpressing some genes from this cluster affected the MIC of vancomycin and teicoplanin in *S. coelicolor*. Overexpression of *SCO1659* and *SCO1660* genes in the Δ*vanJ* background (scSCO1659v and scSCO1660v) increased the sensitivity of *S. coelicolor* 2 fold against teicoplanin but no changes were observed in M600 background strains overexpressing any of the genes. Interestingly, overexpression of all three genes from this cluster increased resistance of Δ*vanJ* background to vancomycin as the MIC increased by 2 fold (**Table 5.3**). This may suggest that these genes complement the more sensitive vancomycin phenotype observed in Δ*vanJ* background strains. This could also indicate that this sensitive phenotype has something to do with glycerol metabolism.

#### 5.2.2.2 Overexpression of genes enriched for the functional group 'response to drug' increased sensitivity of *S. coelicolor* toward vancomycin and teicoplanin

We have already discussed how membrane transporters are likely to be involved in the removal of toxic metabolites in the presence of glycopeptide antibiotics. But we became interested in *SCO2472* and *SCO3910* because the former is a homologue of the *sanA* gene from *E. coli* and *sfiX* from *S. typhimurium* which are believed to be involved in the barrier function towards cell envelope damaging reagents, including vancomycin [192,219]. Previous studies have shown how these genes can restore intrinsic vancomycin resistance in vancomycin-susceptible strains of both bacteria. *SCO3910*, on the other hand, was predicted to encode a MOP domain similar to MurJ [220], suggesting that *SCO3910* could be involved in the transport of cell envelope polymers.

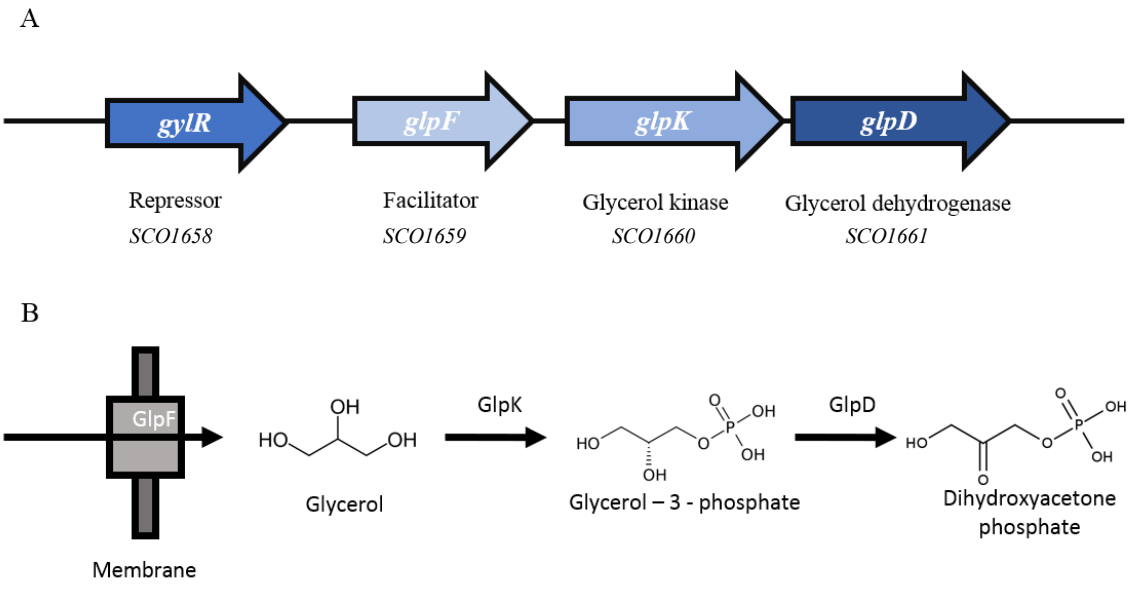



**Figure 5.5** (A) Diagram showing the genetic organisation of the glycerol operon and (B) shows the functional processes that each gene of the glycerol operon encodes.

We did not observe any indications that over-expression of either of these genes positively affected the phenotype of either M600 or Δ*vanJ* background strains towards teicoplanin or vancomycin, but instead increased sensitivity to teicoplanin and vancomycin (Δ*vanJ* background only) by 2-fold (**Table 5.3**). It is therefore likely that these genes have a role not related specifically to glycopeptide antibiotics, and may be involved in rectifying general cellular stress or become increase cellular stress when expressed above a certain threshold.

#### 5.2.2.3 Overexpression of PAP genes diversely affects the susceptibility profile of *S. coelicolor*

Overexpression of PAP genes in *S. coelicolor* produced no changes in the MIC towards kanamycin, novobiocin, flavomycin or carbenicillin but noticeable differences were seen in the MIC of some other antibiotics that target cell wall biosynthesis. Overexpression of *SCO2335*, *SCO2807* and *SCO6511* increased sensitivity of *S. coelicolor* toward teicoplanin in Δ*vanJ* background by 2 fold, with no change in phenotype observed for in M600 background strains (**Table 5.3**). Overexpression of *SCO6356* increased resistance toward vancomycin in Δ*vanJ* by 2 fold. Strains overexpressing *SCO6355* showed a 2 – 4 fold increase in sensitivity towards both vancomycin and nisin. Expression of all PAP2 candidate genes, except *SCO6355*, increased the sensitivity of *S. coelicolor* toward ramoplanin or D-cycloserine. However, interestingly this was the only gene to significantly increase the sensitivity of *S. coelicolor* toward a lantibiotic, duramycin, by more than 4 fold (**Table 5.3**).

### 5.2.3 Phylogenetic analysis of the PAP2 domains from *Streptomyces* PAP suggests possible functional subclasses

A phylogenetic relationship was determined based on the protein sequences available for all genes with PAP2 domains in *S. coelicolor* and the two closely related *Streptomyces* strains, *Streptomyces lividans* and *Streptomyces venezualae* (**Figure 5.6**). Each strain had at least one protein sequence in each major phylum, suggesting that there are distinct classes of PAP2 containing proteins in *Streptomyces*. PAP2 domains from genes that produced more similar MIC profiles when overexpressed in *S. coelicolor,* including *SCO2335* and *SCO2807* as well as *SCO6511* and *SCO6356,* were found in similar clades. The clade containing the PAP2 domain of *SCO6355* also contained the three other *S. coelicolor* genes, *SCO1102* (*Ippα*), *SCO1753* (*Ippβ*) and *SCO5674* which were not differentially expressed in the RNA-seq dataset in either strain.

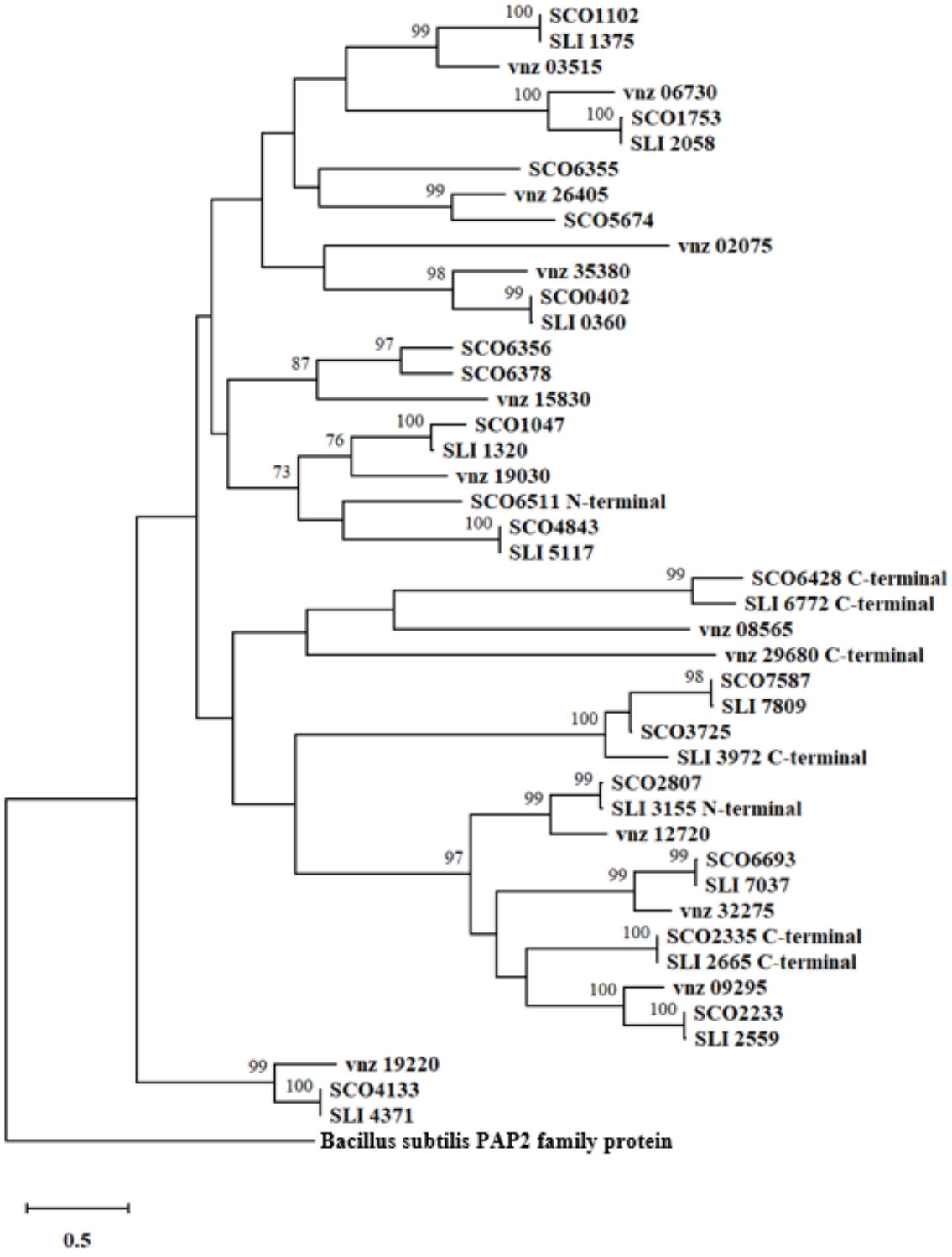



**Figure 5.6** The evolutionary history inferred using the maximum likelihood method for the protein-coding sequences of PAP2 domains in *S. coelicolor* (SCO), *S. lividans* (SLI) and *S. venezualae* (vnz). Forty-five amino acid sequences were used, and bootstrap values (>70%) are shown above the branches. The *B. subtilis* PAP2 family protein was used as the outgroup.

These data corroborate with data in the previous section, demonstrating that the PAP2 genes within *S. coelicolor* could possibly fall into distinct classes that may have specific roles during particular cellular events. We have also demonstrated that these classes could be widely distributed throughout the Actinomycetes, which could indicate their diverse roles within the membranes of these organisms.

### 5.2.4 Further characterisation of the *SCO6357-53* operon

Two PAP2 genes, *SCO6355* and *SCO6356,* were found within a suspected polycistronic unit including three other genes *SCO6357*, *SCO6354* and *SCO6353* (**Figure 5.7**). *SCO6357* has been confirmed to be the transcriptional start site [221], and the downstream products follow the same transcriptional pattern characteristic of an operon (**Figure 5.7**). *SCO6357* has also been confirmed to be downstream from a predicted $\sigma^E$ regulon promoter sequence [150], meaning that this operon could have a functional role cell envelope stress response. *SCO6357* is predicted to encode a DedA/Tvp38-like protein which are commonly found throughout prokaryotic genomes, but their specific function remains elusive. There is some evidence that it functions as proton-motive force-driven antiporter in *E. coli*, and disruption can lead to sensitivity against cationic compounds [216,217]. *SCO6354-53* are predicted to encode for a histidine kinase and a response regulator in a TCS. The presence of this TCS suggests the possibility for autoregulation or that its involvement in a regulatory cascade.

Overexpressing *SCO6357* or *SCO6354-53* did not affect the susceptibility of *S. coelicolor* toward teicoplanin, vancomycin and most of the other cell wall targeting antibiotics (**Table 5.4**). Overexpression of SCO6357 did increase the resistance of *S. coelicolor* to the lantibiotics nisin and duramycin by 2 fold in the M600 background. However, the overexpression of both *SCO6357* and *SCO6354-53* in Δ*vanJ* background increased the sensitivity of *S. coelicolor* toward duramycin by between 2-4 fold. *SCO6357* also increased the resistance against kanamycin by 2 fold in the Δ*vanJ* background. Although kanamycin does not have a mode of action that targets the cell envelope or its biosynthesis, it is plausible based on their predicted function, that this gene plays a role in the barrier function of the cell membrane, possibly reducing the rate of diffusion of kanamycin across the membrane (**Table 5.4**).

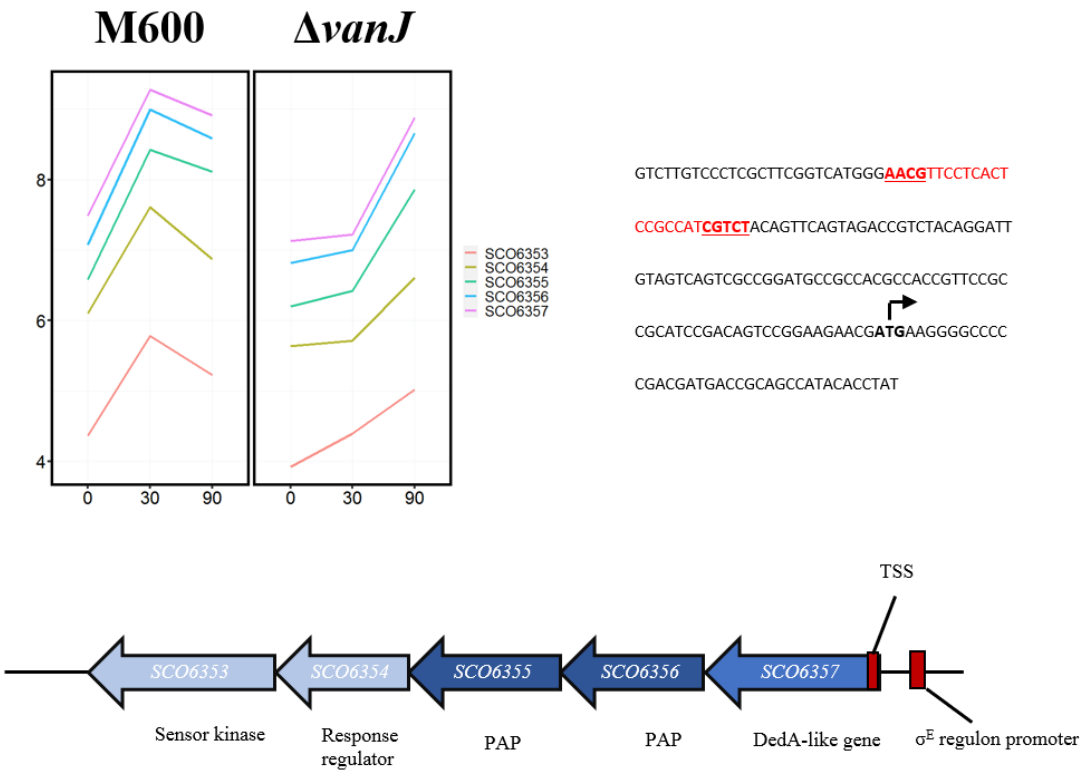



**Figure 5.7** Plot showing the expression of the *SCO6357-53* operon over 90 minutes after exposure to teicoplanin, for both M600 and Δ*vanJ*. Each point is the average of two biological replicates. The genomic sequence with the transcriptional start site (TSS) and location of the $\sigma^E$ regulon promoter shown 88bp upstream of the operon are shown. The genomic organisation of the operon is also displayed at the bottom of the image.

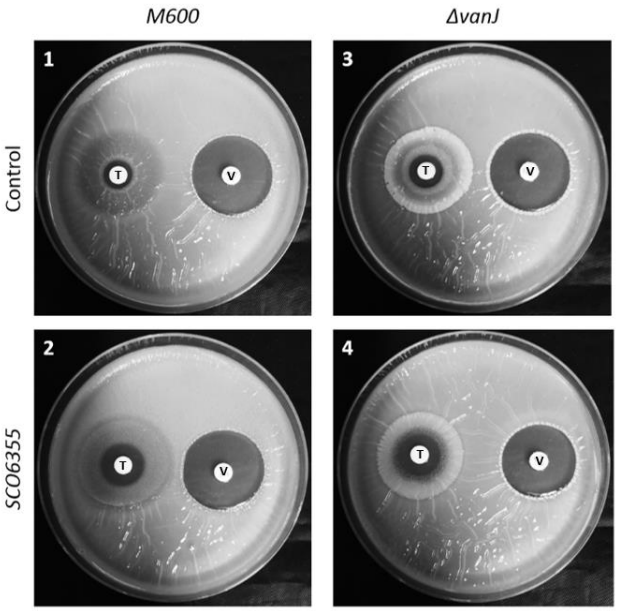
**Table 5.4** MIC values determined for the antibiotic library for strains overexpressing genes from the *SCO6357-53* operon. T (teicoplanin), V (Vancomycin), R (ramoplanin), C (carbenicillin), D-c (D-cycloserine), F (flavomycin), Ni (nisin), Du (duramycin), N (novobiocin), K (Kanamycin) and B (bacitracin).

| Strain | T | V | R | C | D-c | F | Ni | Du | N | K | B |
|---|---|---|---|---|---|---|---|---|---|---|---|
| sc10257m | 0.5 | 80 | 4 | >512 | 256 | >128 | 256 | 8 | 16 | <0.125 | 64 |
| sc10257v | 0.5 | 40 | 4 | >512 | 256 | >128 | 256 | 8 | 16 | 0.5 | 64 |
| scSCO6357m | 0.5 | 80 | 4 | >512 | 256 | >128 | 512 | 16 | 16 | <0.125 | 64 |
| scSCO6357v | 0.5 | 40 | 4 | >512 | 256 | >128 | 256 | 4 | 16 | 1 | 64 |
| scSCO6354-53m | 0.5 | 80 | 4 | >512 | 256 | >128 | 256 | 8 | 16 | <0.125 | 64 |
| scSCO6354-53v | 0.5 | 40 | 4 | >512 | 256 | >128 | 256 | <2 | 16 | 1 | 64 |

### 5.2.5 Understanding the role of *SCO6355* in antibiotic resistance

A series of vancomycin and teicoplanin disc diffusion assays were carried out using *SCO6355* overexpression constructs. There were significant differences ($P<0.0005$) between the average diameters of the inhibitory zones (DIZ) of the two control strains against vancomycin but no significant differences against teicoplanin. As predicted, the DIZ around both antibiotic discs of all tested strains overexpressing *SCO6355* were larger in Δ*vanJ* background than M600 background. The average DIZ of *SCO6355* overexpressing *S. coelicolor* constructs, regardless of its background, was significantly bigger around vancomycin disc compared to those of the corresponding controls (**Figure 5.8**). These results are consistent with the MIC results shown in **Table 5.3** as the overexpression of *SCO6355* significantly increased the sensitivity of *S. coelicolor* toward vancomycin in both M600 and Δ*vanJ* background, indicating that this gene likely antagonises the endogenous vancomycin-resistance system in *S. coelicolor*.

*S. coelicolor* is highly resistant to vancomycin but displays varying susceptibility to a structurally diverse range of glycopeptide antibiotics. We now understand that this is reliant on the level of detection by the VanS protein [100]. Further MIC test for a small library of glycopeptide antibiotics was also carried out using the control strains along with scSCO6355m and scSCO6355v, to identify relationships between *SCO6355* and glycopeptide structure. Both control strains exhibited between intermediate or high-level resistance to vancomycin, balhimycin, ristocetin, oritavancin and telavancin but they exhibited relatively high sensitivity toward chloroeremomycin and teicoplanin as shown with very low MIC scores (**Table. 5.5** and **Figure 5.9**). The overexpression strains exhibited 4 fold increased sensitivity toward vancomycin and balhimycin and 2 fold toward ristocetin (**Table 5.5)**. No changes in sensitivity were observed against chloroeremomycin or any of other lipoglycopeptides tested. We also didn't observe any MIC changes for vancomycin derivatives with putrescine substituted onto the seventh amino acid nor were they observed for this vancomycin-putrescine derivative with a chlorobiphenyl group substituent (data not shown). Together these data suggest that overexpressing *SCO6355* increases the sensitivity of *S. coelicolor* toward vancomycin and other glycopeptide structurally similar glycopeptides. However, the same phenotype is not seen with glycopeptides such as the lipoglycopeptide antibiotic or chloroeremomycin. We note that all these glycopeptide antibiotics poorly induce the van resistance system, so from this we can suggest that the enzymatic function of *SCO6355* may interfere with the sensory capabilities of VanS to detect certain glycopeptide antibiotics like vancomycin. As a result, there is weaker induction of the *vanHAX* operon and these cells become more sensitive to these antibiotics.




| Sample | Strain | Teicoplanin Average diameter (cm) | Vancomycin Average diameter (cm) |
|---|---|---|---|
| 1 | M600 + pIJ10257 | 2.71 ± 0.03 | 0.96 ± 0.03 |
| 2 | M600 + SCO6355 | 2.68 ± 0.03 | 1.44 ± 0.01 |
| 3 | ΔvanJ + pIJ10257 | 2.74 ± 0.03 | 1.11 ± 0.02 |
| 4 | ΔvanJ + SCO6355 | 2.47 ± 0.04 | 1.40 ± 0.02 |

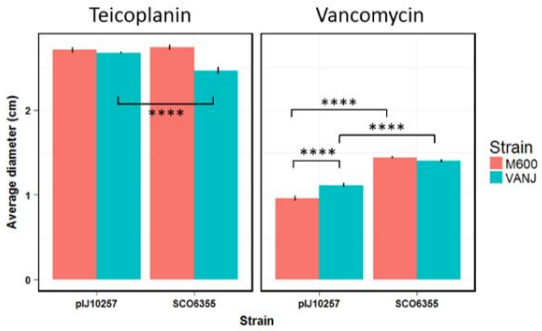



**Figure 5.8** Disk diffusion assay of sc10257m (1), sc10257v (2), scSCO6355m (3) and scSCO6355v (4) against V (vancomycin 100 µg) and T (teicoplanin 30 µg). The diameter of the inhibitory zones (DIZ) was an average of four random measurements shown in the table in the top right with SD and plotted on the bar graph at the bottom of the table. An unpaired T-test was carried out to show significance was <0.0005.

**Table 5.5** List of MIC values of glycopeptide antibiotics for the control strains and strains overexpressing *SCO6355*. Values are the average of three biological replicates and indicate μg/mL of antibiotic needed to prevent any growth on the media. Red values indicate decreased resistance when compared to controls. V (Vancomycin), T (teicoplanin), Bal (balhimycin), Ri (ristocetin), Ch (chloroeremomycin), Tel (telavancin), Dal (dalbavancin) and O (oritavancin).

| Strain | V | T | Bal | Ri | Ch | Tel | Da | O |
|---|---|---|---|---|---|---|---|---|
| sc10257m | 80 | 0.5 | 8 | 2 | 0.5 | 16 | >16 | 8 |
| sc10257v | 40 | 0.5 | 8 | 4 | 0.5 | 16 | >16 | 8 |
| scSCO6355m | 20 | 0.5 | 2 | 1 | 0.5 | 16 | >16 | 8 |
| scSCO6355v | 10 | 0.5 | 2 | 2 | 0.5 | 16 | >16 | 8 |

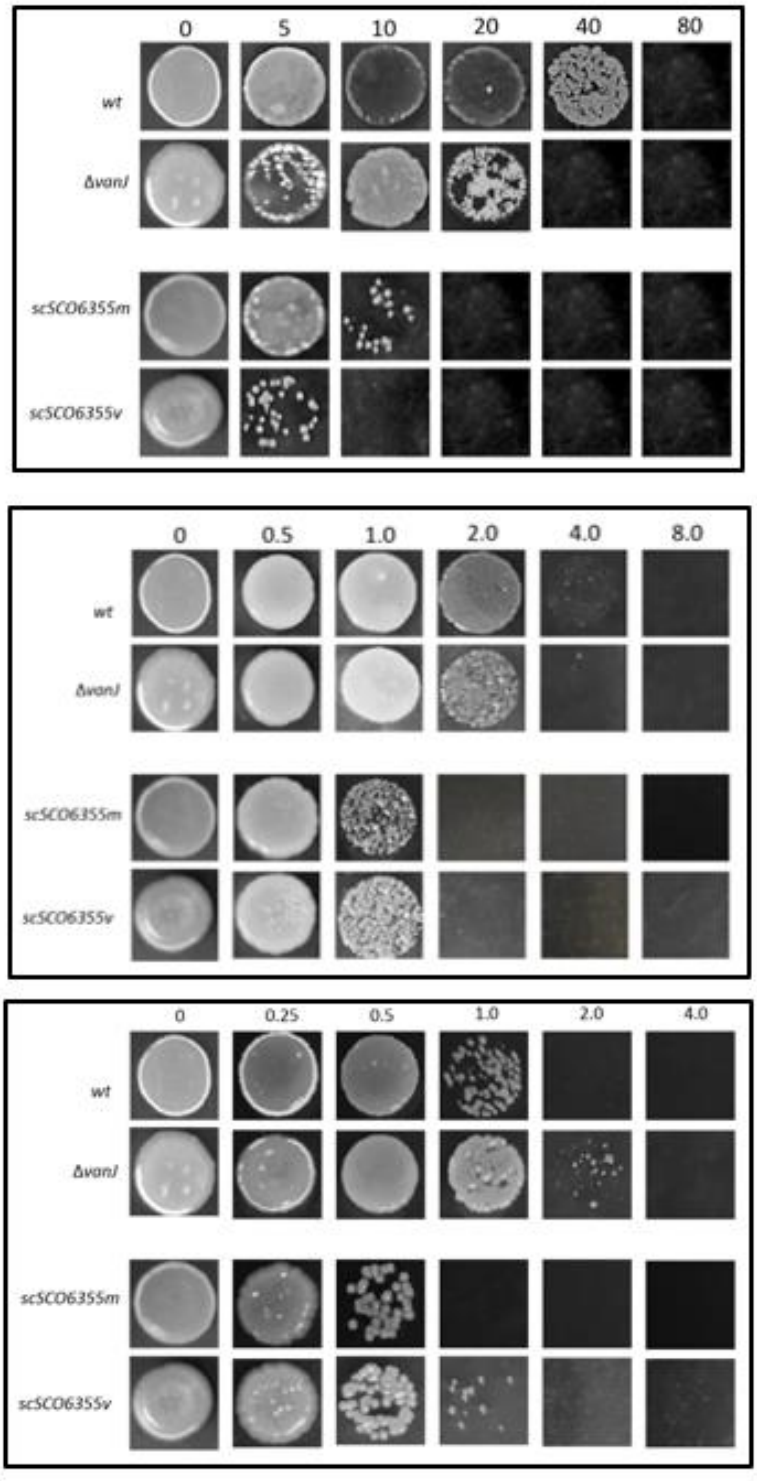



**Figure 5.9** Image of colonies on plates showing the MIC of each strain against the glycopeptide antibiotics vancomycin (Top), balhimycin (Middle) and ristocetin (Bottom). Antibiotic concentration in μg/mL is shown along the top, and the strains are listed on the left.

## 5.3 Summary

This section focused on the functional characterisation of the teicoplanin-responsive DE genes that were discussed in **Chapter 4**. The selected genes came from three different categories: the three genes in the glycerol operon that relate to glycerol metabolism; two strongly induced membrane transporters; and five genes with predicted PAP2 activity. The additional genes within the *SCO6357-53* operon, *SCO6357* and *SCO6354-53* were also cloned into pIJ10257 and conjugated into *S. coelicolor* to characterise their possible involvement in resistance to the glycopeptide antibiotics. Unfortunately, we failed to successfully conjugate either CRISPR plasmid into *S. coelicolor* regardless of any genetic manipulation, and therefore we were not able to generate knock-out mutants for any of the selected genes. But, a total of 14 plasmids were used in this study (including two pIJ10257 controls) to generate 28 strains of *S. coelicolor* in two different background strains, M600 and Δ*vanJ*.

The two background controls (sc10257m and sc10257v) both showed the same pattern of sensitivity towards the antibiotic library, but sc10257v was 2 fold more sensitive to vancomycin. Overexpressing *SCO2472* or *SCO3910* increased sensitivity of *S. coelicolor* to teicoplanin regardless of its background but increased sensitivity toward vancomycin was only observed in the M600 background strains. Overexpressing any of the three genes from the glycerol operon complemented the vancomycin sensitive phenotype seen only in Δ*vanJ* background. Both scSCO1659v and scSCO1660v were 2 fold more sensitive to teicoplanin but only in M600 background.

There were significant changes in the phenotype of *S. coelicolor* in strains overexpressing genes with predicted PAP2 domains. Overexpression constructs within this category demonstrated diverse phenotypes towards different cell envelope targeting antibiotics suggesting that the products of these genes correlate with the physiology of the cell envelope. The amino acid sequences of these genes were compared with the sequences of the predicted PAP2-containing protein sequences in closely related *Streptomyces*. Unsurprisingly, genes that produced similar MIC profiles against the antibiotic library were grouped more closely together. The distribution of these classes seems to be prevalent in the Streptomyces species compared in this study, but further work needs to be done to identify how prevalent these classes of PAP are in the actinomycetes and their specific roles in the bacterial membrane.

*SCO6355* became a focus of this section because overexpression of this single gene significantly increased sensitivity to vancomycin and to other cell wall-specific antibiotics such as the lantibiotic duramycin. Further MIC and disc diffusion assay tests indicated that overexpression of *SCO6355* significantly contributed to the sensitive phenotype of *S. coelicolor* toward glycopeptide antibiotics structurally similar to vancomycin. This suggests that *SCO6355* antagonises the intrinsic *vanHAX* resistance system in *S. coelicolor* possibly through affecting the composition of the bacterial cell membrane. Understanding the mechanistic relationship between *SCO6355* and vancomycin will be vital for developing strategies that are able to circumvent glycopeptide resistance in glycopeptide resistance pathogenic bacteria.

# Chapter 6. Investigation of the Production of Glycerol as a result of Exposing *S. coelicolor* to Teicoplanin

# 6.1 Introduction

Although the MIC data presented in **Chapter 5** did not indicate that expressing genes within the glycerol operon strongly increases the fitness of *S. coelicolor* in the presence of the antibiotic library tested, we still wanted to explore the possible involvement of glycerol in the response to teicoplanin as this is still an unexplored and unreported outcome of glycopeptide antibiotics [149]

The glycerol operon is positively induced in the presence of glycerol in the media [203] and it has been speculated that it is the G3P generated from the phosphorylation of glycerol by GlpK that acts as a positive regulator of the promoter sites in the glycerol operon and the *gylR* gene (**Figure 6.1**) [222]. The glycerol operon is regulated by the product of the *gylR* gene, which is induced by glycerol and weakly repressed by glucose generating a negative feedback loop in the presence of glycerol. In *S .coelicolor*, glycerol is likely the final product in the breakdown of phospholipids, being released from G3P before it can be transported back into cells, where it can be shuttled into glycolysis or used for further rounds of phospholipid biosynthesis (Personal communication with Colin P. Smith, 2020).

We previously speculated that the large increase in the activity of the glycerol operon could be due to the increased degradation of phospholipids, leading to increased concentrations of G3P and glycerol in the culture media used to grow *S. coelicolor*. In this section, we investigate the likelihood of phospholipid degradation through the application of a colorimetric assay that is able to detect changes in glycerol concentration. To test this hypothesis, a 2-step colourimetric assay was designed based on the work of Bompelly and Skaf (2014) who laid out the foundations for determining the concentration of glycerol in biodiesel [223]. This method relies on the oxidation of one molecule of glycerol by periodate ($IO_4^-$) to produce two molecules of formaldehyde (**Figure 6.2A**) which can then be quantified using the colourimetric, Nash test (**Figure 6.2B**) [224]. The Nash test involves reacting formaldehyde with the ester, acetylacetone using ammonium acetate as a catalyst to produce the soluble yellow compound diacetyldihydrolutidine (DDL). The amount of light absorbed by DDL can then be used to determine the concentration of glycerol in the original sample.

Further understanding of the molecular processes that bacterial cells carry out in response to stressful environments such as those that are created by antibiotics is crucial for indicating how drugs fail in the clinic, and how we can avoid these outcomes in future.

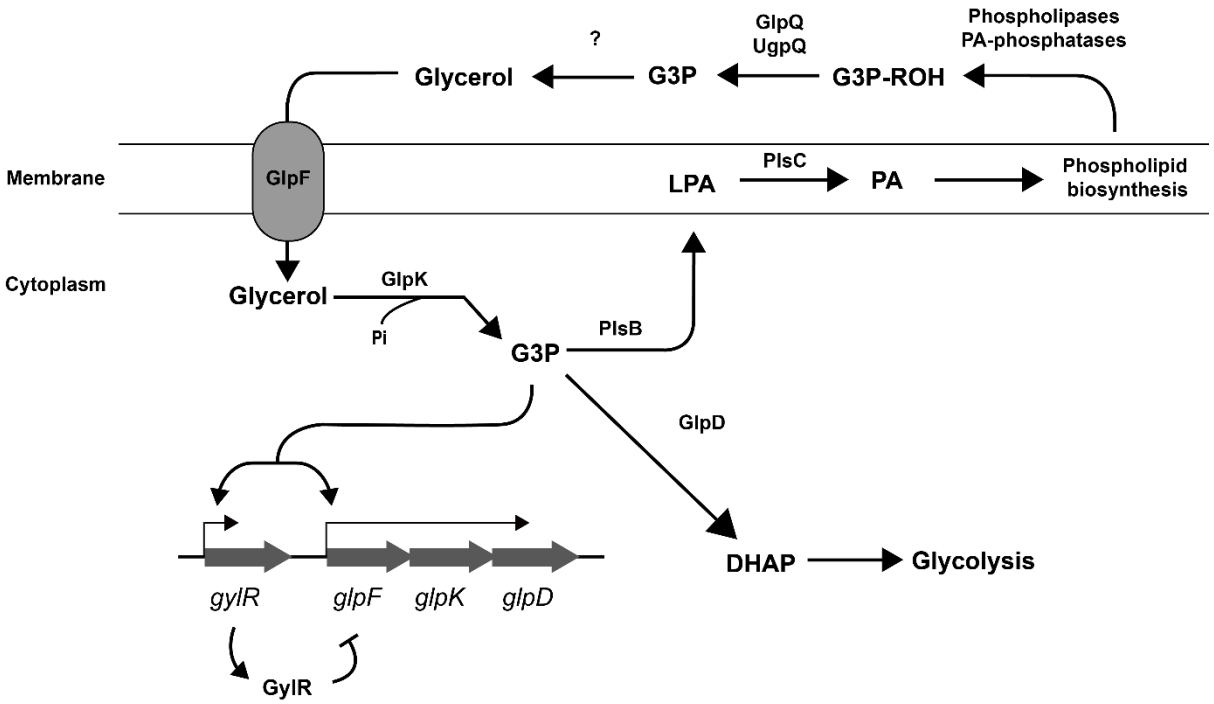



**Figure 6.1** The metabolic pathways for glycerol and how these are possibly linked with glycolysis and phospholipid biosynthesis in *S. coelicolor*. Genes involved in biosynthetic processes are listed above arrows.

A

$$C_3H_8O_3 + IO_4^- \rightarrow HCHO + CH_3COOH + IO_3^- + H_2O,$$

$$CH_3COOH + IO_4^- \rightarrow HCHO + HCOOH + IO_3^-.$$

B

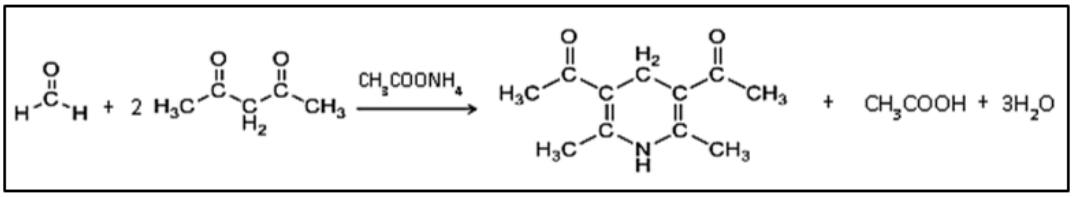



**Figure 6.2** The two-step reaction to produce two molecules of formaldehyde from glycerol under oxidative conditions (A). The reaction found in (B) shows the reaction to produce DDL from formaldehyde, acetylacetone and ammonium acetate.

# 6.2 Results

## 6.2.1 Standard curves to determine the limitations of the colourimetric assay

Standard curves were produced using both formaldehyde and NMMP spiked with glycerol to observe the stoichiometric relationship between the two reactions (**Figure 6.3A and B**). Although both reactions showed strong linear relationships ($R^2 > 0.95$) with clear linear boundaries, the curve for NMMP spiked with glycerol was less steep than that observed for formaldehyde. There also appeared to be increased background noise in NMMP, indicating that one component in NMMP could be reacting with one of the components in either the first or second steps of the assay. As a result, the second standard curve with NMMP + glycerol was used for the indirect measurement of the concentration of glycerol in subsequent reactions (**Figure 6.3B**).

## 6.2.2 Understanding the growth kinetics of *S. coelicolor*

A growth curve was carried out for both strains of *S. coelicolor* to determine the $OD_{540nm}$ of cultures grown for 24h in NMMP culture media. Both *M600* and Δ*vanJ* strains were observed to have similar growth kinetics with Δ*vanJ* initially growing more slowly. Both strains reached mid-log phase at approximately 18h and reached a maximum observable growth at around 24h, after which cultures started to produce antibiotics and it became impossible to get an accurate OD reading (**Figure 6.4A**).

We also carried out the colourimetric test on the growth media of cultures grown for 12, 15, 18 and 21h to indicate whether cellular growth influenced the outcome of the assay. Although there was a slight decline in the absorbance ($OD_{410nm}$) in the colorimetric assay at 18h when using media from Δ*vanJ* cultures compared to cultures taken from M600. This demonstrated that providing the background noise is taken into consideration, this assay should be suitable for indirectly measuring any increases in glycerol produced as a result of antibiotic exposure (**Figure 6.4B**). And, any changes seen in the assay are likely to be a result of exposing cells to antibiotic rather than because of cellular growth.

.

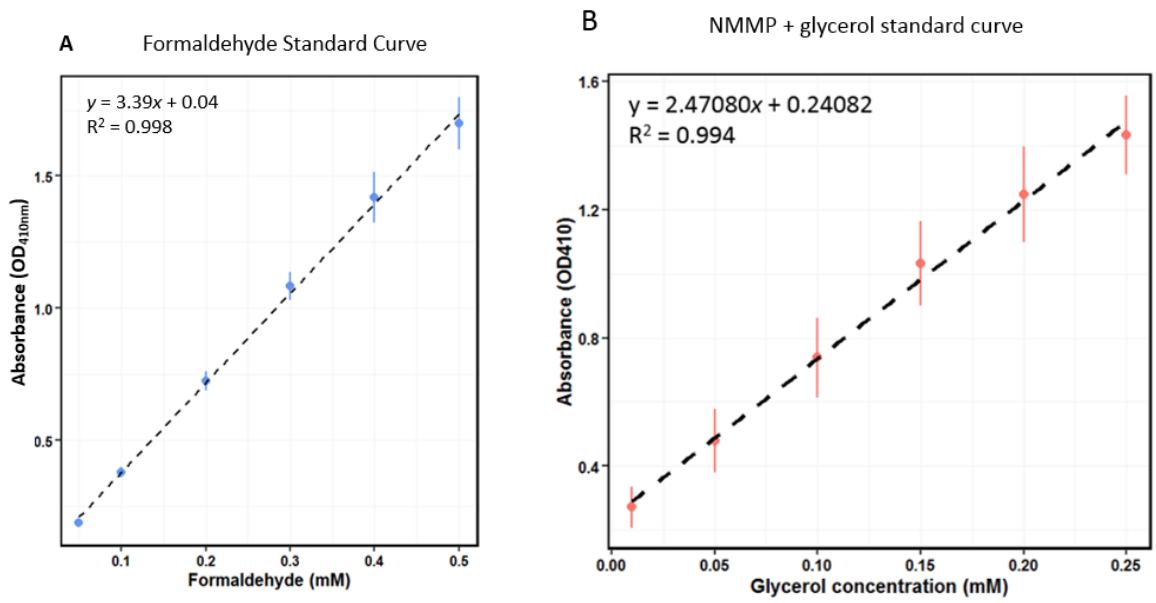



**Figure 6.3** (A) shows the standard curve for formaldehyde concentration (mM) against absorbance ($OD_{410nm}$). (B) The standard curve for NMMP spiked with corresponding glycerol concentrations (mM) against absorbance ($OD_{410nm}$). The formulae for the curves are displayed along with the $R^2$ value in the top right hand corner of each plot. Each point is the average of three biological replicates and standard error bars are also shown in B.

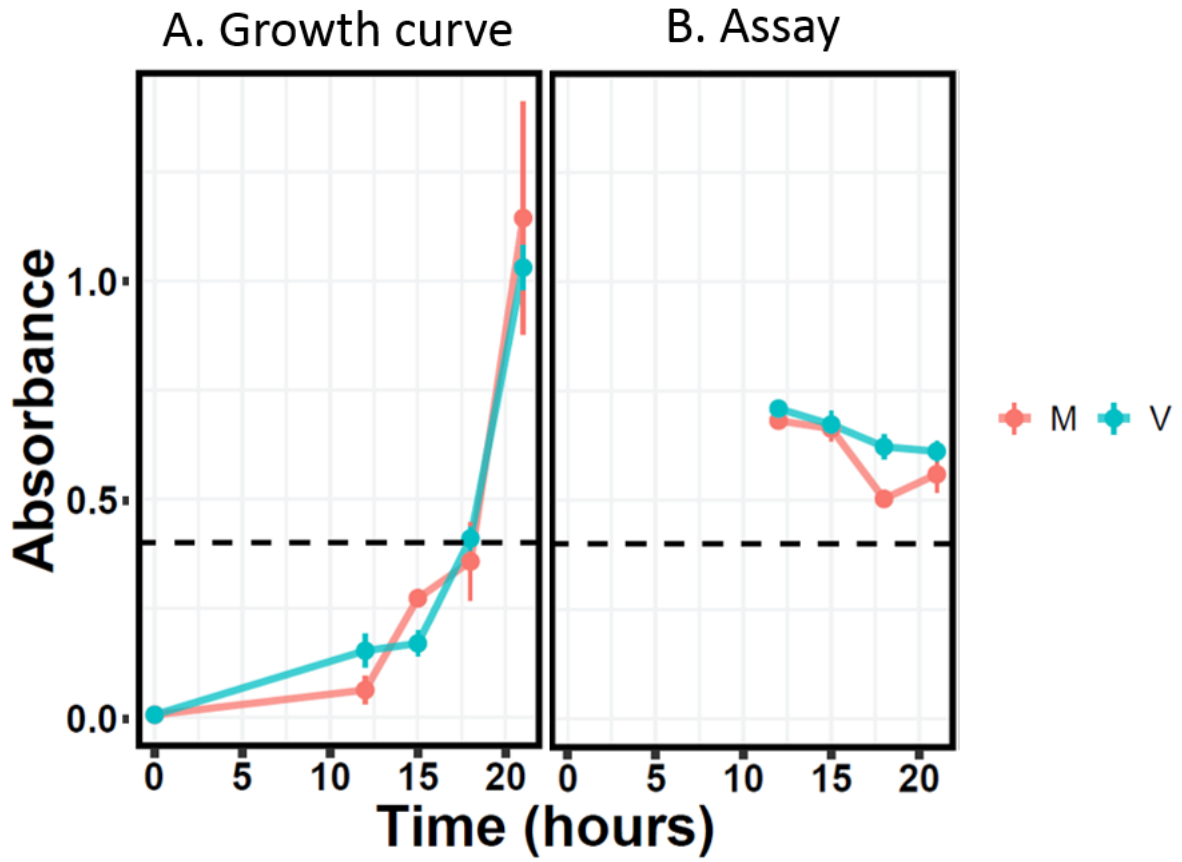



**Figure 6.4** Absorbance ($OD_{540nm}$) for the growth curve of *S. coelicolor* (A) for M600 (M) and Δ*vanJ* (V) over 24h. The second plot (B), shows the absorbance observed (OD410nm) for the colourimetric assay carried out on the supernatants of cultures grown in time (hours) along the x-axis. Absorbance readings are the average of three biological replicates, and the horizontal dashed line indicates an $OD_{540nm}$ = 0.4 (mid-log growth).

### 6.2.3 Indirect measurement of glycerol production as a result of exposure to teicoplanin

Four antibiotics were assayed against *S. coelicolor* to identify whether glycerol production was specific to teicoplanin. Bacitracin and ramoplanin were selected because they both have different targets to teicoplanin, but also target the later stages of PG biosynthesis. Kanamycin was selected as a non-cell envelope targeting antibiotic negative control. Lethal concentrations of each antibiotic were used to ensure all cells were appropriately 'challenged' by each antibiotic. **Figure 6.5** demonstrates that glycerol concentration increased at 2h and 4h by up to 8 µM after exposing M600 to teicoplanin, bacitracin and ramoplanin. The effects were minimal for the kanamycin control, suggesting that this could be a result of cell envelope targeting antibiotics that typically cause cell lysis.

The same assay was carried out on the Δ*vanJ* mutant, and although less glycerol was observed in the media during earlier time points, glycerol concentrations increased to 8 - 10µM for teicoplanin, bacitracin and teicoplanin by 4h. Glycerol concentration also increased by 2 - 4µM over the 4h time course after exposing cells to kanamycin, indicating that the observed increase in glycerol is likely to be the result of stress and evidently cell death caused by antibiotics. The lower perceived glycerol concentration in Δ*vanJ* at earlier time points could also be a result of a decrease in the expression of genes relating to growth when cells are exposed to teicoplanin that we presented in **Chapter 4**. If increases in glycerol are the result of cell death, then less cell death may occur in Δ*vanJ* while processes involved in growth are slowed, if similar changes in gene expression are also caused by the other three antibiotics.

To assess whether the results observed in **Figure 6.5** were the result of cell death caused by antibiotic dosage, assays were carried out using different concentrations of teicoplanin and kanamycin (**Figure 6.6**). Sub-lethal concentrations of both kanamycin and teicoplanin cause decreases in the concentration of glycerol in the media in comparison to the negative control. Presumably, this is due to continued cellular growth as a result of sub-lethal concentrations used to challenge cells. In contrast, lethal concentrations (10 µg/mL) of both antibiotics caused increases in glycerol concentration even after 1h of treatment. These data explain why we saw increased expression of genes of the glycerol operon in response to teicoplanin, but have not observed this in previous studies which have used sub-lethal antibiotic concentrations [149].

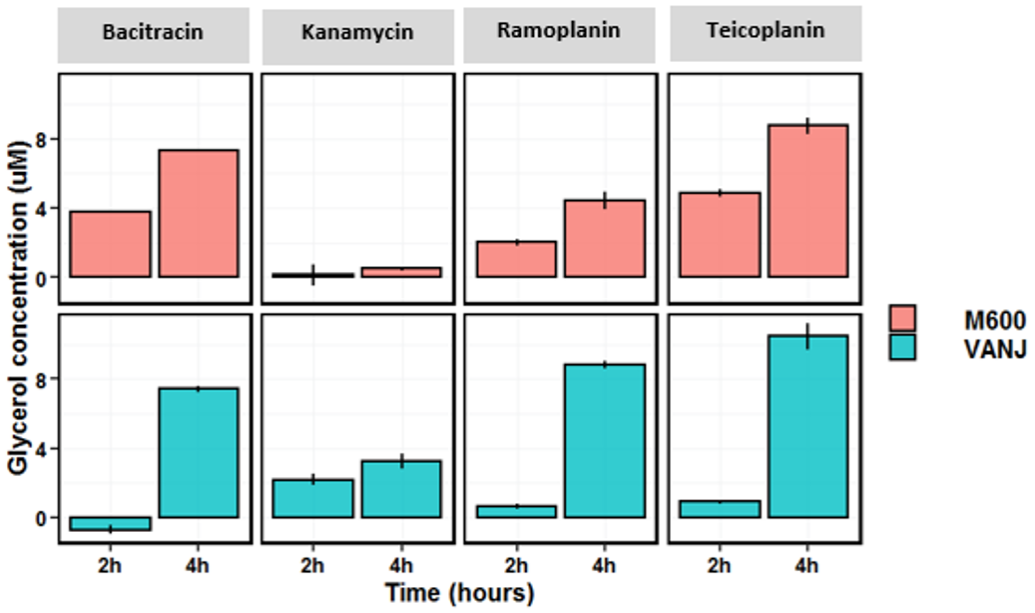



**Figure 6.5** Plots showing the estimated glycerol concentration in M600 and Δ*vanJ* (VANJ) cultures at 2 and 4h after exposing cells to one of four different antibiotics bacitracin, kanamycin, ramoplanin and teicoplanin. Experiments are the average of three biological replicates, and the standard error bars are shown on each plot.

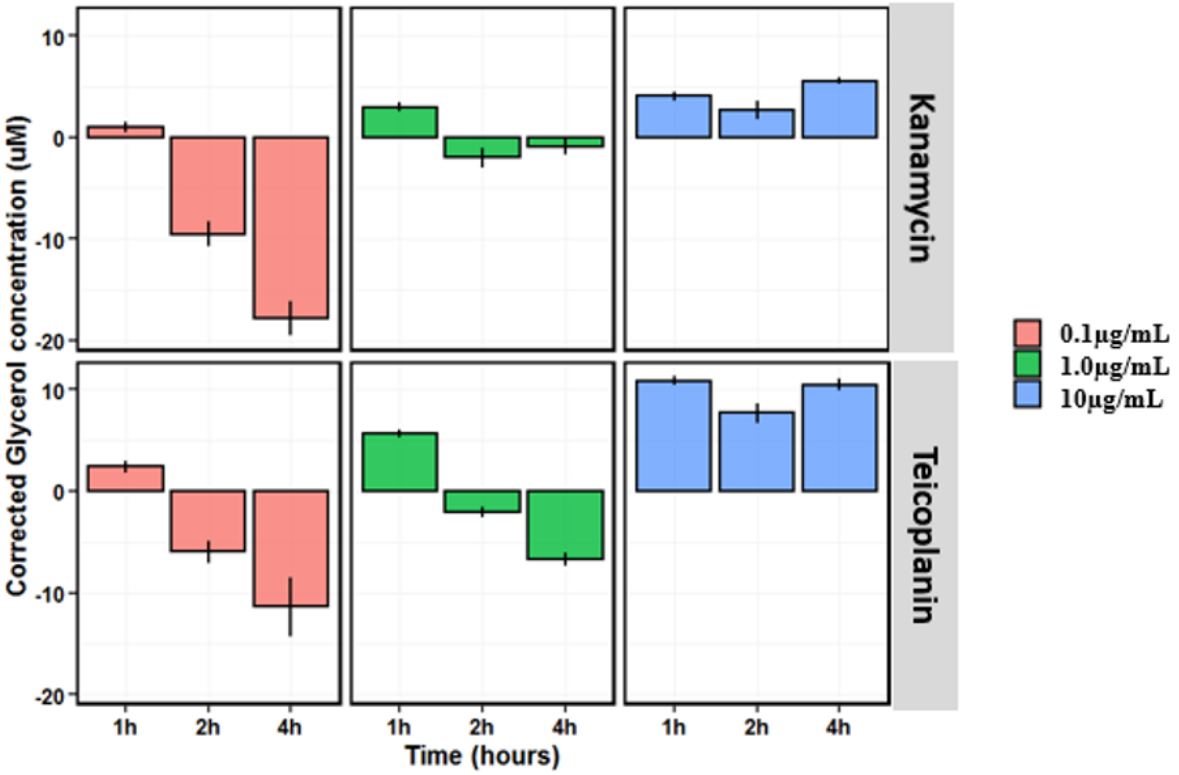



**Figure 6.6** Plot showing how the dosage-response of kanamycin and teicoplanin affect the estimated glycerol concentration after exposing M600 to each antibiotic for 1, 2 and 4h. Antibiotic concentration are shown in the legend and experiments are the average of three biological replicates, with the standard error bars are shown on each plot.

## 6.3 Summary

This chapter has demonstrated that the two-step colourimetric assay first involving the oxidation of glycerol, to produce two molecules of formaldehyde, followed by the esterification of formaldehyde with acetylacetone to produce DDL, can be used to indirectly measure glycerol concentration in cultures of *S. coelicolor*. Although the assay did produce some background noise, this was minimally affected in actively growing cultures. The results from these experiments also demonstrated that glycerol was produced as a result of exposure to teicoplanin, but similar effects were also observed for both ramoplanin and bacitracin, and with higher doses of kanamycin. This dose dependence on glycerol production suggests that glycerol production is likely the result of cell death caused by lethal concentrations of antibiotics, rather than a specific mechanism as a result of teicoplanin. This would also explain why the glycerol operon became more active in the Δ*vanJ* strain, which is now known to exhibit greater signs of cellular stress as a result of teicoplanin, so was likely to experience higher rates of cellular death. We also conclude that this effect has not been observed previously because earlier studies have used sub-lethal concentrations of cell wall targeting antibiotics [149].

# Chapter 7. Discussion

# 7.1 Pre-analysis of the teicoplanin induced RNA-seq dataset

NGS technologies have progressed the field of RNA biology, giving researchers access to the transcriptome in unprecedented detail to identify novel RNAs, and accurately quantify transcripts with an almost infinite dynamic range, all without prior knowledge of functional genes [136,225]. Moreover, providing care is taken during sample preparation and experimental design, experimental replicates are typically more congruent [226]. In recent years, a consensus of a few pipelines for carrying out transcriptomic studies has arisen. That being said, there are still steps that must be taken to ensure that minimal bias is introduced into the pipeline.

The primary source of bias in NGS data comes from the loss of material and quantitative bias, but every step in the analysis pipeline can introduce bias into the dataset [138]. To prevent the loss of material, ideally, RNA molecules would be sequenced in their entirety. Third-generation sequencing technologies, such as the Oxford Nanopore Technologies or the Pacific Biosciences SMRT systems overcome assembly problems, allowing the recovery of full-length transcripts [121,145]. However, the Illumina technology employed in this work still uses fragmented RNA (~50 bp) from purified samples that are then amplified and sequenced as individual fragments. These short fragments are aligned to a reference and assembled back into transcripts in silico, followed by quantification using statistical software. Additionally, short-read sequencing technologies are still more cost-effective and provide much higher accuracy per base. Previous groups have attempted to lay out a comprehensive survey for the best practices to date involved in RNA-seq [145], and this section will discuss the QC steps taken to ensure confidence in the findings. These steps are summarised in **Table 7.1**.

## 7.1.1 Addressing the bias introduced through library construction and NGS

In this study, two experimental questions were addressed. What were the significant differences in the transcriptomes of M600 and Δ*vanJ*? As well as, what genes are involved in the direct transcriptional response to teicoplanin? The experimental set up used to answer these questions involved sequencing RNA sampled at 0, 30 and 90 minute time points for M600 and Δ*vanJ* mutant strains after being exposed to teicoplanin (10 μg/mL). Before any core analysis, the quality of the sequence data was assessed to ensure that there were minimal sequencing errors.

Illumina sequencing has come a long way in recent years, but errors commonly occur after guanine, and adenines are sometimes substituted for cytosines [227–230].

**Table 7.1** QC steps that were undertaken during this project to ensure the reliability of the RNA-seq dataset.

| Problem | Quality Control Step |
|---|---|
| Sequence Quality | Ran data through the FastQC tool to assess data quality |
| Unaligned reads | Removed ~5% of reads which failed to align to the reference genome |
| Lowly expressed reads | Reads expressed at less than 1 log-CPM were removed from the data |
| Inactive genes | 435 genes removed from analysis for inactivity |
| Differences in library size | Each library was normalised using TMM |
| Filtered genes correctly removed | The voom plot provides a good assessment of whether poorly expressed genes were effectively removed from the data. |
| Replicability and sample organisation | The PCA plot showed that replicates clustered so samples were kept in order and the data was reproducible |
| FDR | The TREAT method was applied to the data which uses both a fold-change and P-value cut-off to conservatively reduce the FDR. |

Newer HiSeq sequencers released in 2010 improved upon their previous machine, decreasing error rates, while increasing data acquisition [228]. Even with these advances, the technology can still be prone to errors, particularly with sequences that have a high G+C content which would be expected in *S. coelicolor*. The FastQC tool did not indicate any concerns with the quality of the data presented in this work, demonstrating that the RNA-extraction produced high-quality material for sequencing.

Each library was sequenced in duplicate, and to a depth of more than 10 million reads, theoretically covering the entire *S. coelicolor* genome (**Section 3.2.1**). These two factors are essential to consider together because even though RNA-seq is usually replicable at the level of sequencing, increasing the number of biological replicates improves the statistical inference. Sequencing samples to a greater depth captures information from rare or lowly expressed transcripts [145], but models have demonstrated that there is a theoretical saturation point for any RNA-seq library where increasing depth is limited in the amount of additional statistical power it provides. Instead, resources should be put into increasing the number of biological replicates, as this will improve the chances of identifying biologically relevant DE [231]. This study lacked a third replicate, however we discuss how this was dealt with in later sections.

### 7.1.2 QC checks for the pre-analysis of the RNA-seq dataset

Raw RNA-seq data is never used directly for differential expression analysis because library differences inherently bias the data. Normalising libraries based on their size is the most common method to remove differences in sequence depth. After the transcripts were assembled, any counted at less than 1-CPM were removed from the library (**Figure 3.3**). This meant that any reads counted less than 10 - 20 times in any given sample were removed from the data, which is justified both statistically and biologically because these RNAs are unlikely to be biologically relevant to gene expression. And, failure to remove them would interfere with the statistical power of any tools used to identify DE [122]. The sequencing process can also introduce systemic technical effects leading to differences in the size of each library. Each library was normalised using the TMM method to scale each dataset based on their size (**Figure 3.4**). Samples VANJ_T30_A and _B both had higher densities of reads than the other samples, but after applying the TMM method, the differences in these samples were noticeably reduced. The TMM method was used in this work for its efficacy and ease with the limma pipeline [123,232], but several other suitable normalising methods also exist.

The limma + voom pipeline was implemented for this study as limma performs on par with other DE statistical software such as edgeR, NOISeq and DESeq2, while effectively controlling FDR even with smaller sample sizes making it appropriate for our smaller dataset [123,233,234]. Limma was initially used for linear modelling in microarray analysis but has recently been updated to make it accessible for count-based RNA-seq data. One thing to take into consideration from this pipeline was that statistical analysis for microarray data worked on the basis of fluorescence data which was assumed to be normally distributed. RNA-seq data consists of direct counts, and it has been argued that count-based statistics should be used. A property of count-based data is that it has unequal variability even after log transformation, and larger counts have more significant standard deviations [234]. The voom tool provides a way of removing the mean-variance relationship (**Figure 3.5**), giving a visual representation of the mean and variance are not independent of each other before applying the algorithm. This heteroscedasticity of the data is decreased by the incorporation of voom's precision weights within limma's empirical Bayes statistics. A secondary feature of the voom plots also indicates whether lowly expressed genes were filtered adequately, as lowly expressed genes would accumulate nearer the origin [234]. As this did not occur, it was assumed that the filtering step had been carried out effectively.

### 7.1.3 Most of the variation observed between the two strains occurs at the 30 minute time point

Previous work has demonstrated that VanJ is an anchored protein with an N-terminal membrane-spanning region and an external-facing catalytic domain. There are no phenotypic differences in the sensitivity of both M600 and Δ*vanJ* towards teicoplanin because teicoplanin is a poor-inducer for VanS in *S. coelicolor* [115]. The PCA in Chapter 3 (**Figure 3.6**) indicated that there are substantial differences in the gene expression of both strains at the 30 minute time point. The PCA highlighted that the 0 and 90 minute time points are more similar than the 30 minute time points suggesting the initial genetic response to teicoplanin differs in both strains. This was an unexpected outcome because *van* gene transcription is tightly controlled by the VanR/VanS TCS in *S. coelicolor*. Because of this, the concentration of VanJ is believed to be negligible when no or very poor inducers are present.

### 7.1.4 Using the TREAT method to identify DEGs

Identifying DEGs is a challenge in every RNA-seq pipeline. Mathematically, any gene that differs in its expression level from one condition to another is statistically significant if enough replicates are carried out. However, the biological significance of such changes may be trivial if the change in abundance is low relative to other genes [124]. On top of this, it is also becoming more apparent that changes in transcription do not always equate to physiological outcome [235,236]. It is essential to implement thresholds that define the limits of what can be classified as a DEG to control the FDR or biologically insignificant DE.

Historically, log-fold changes and P-values have been used, but independently they both suffer type I error or do not take gene-wise variability into account [124,237]. The TREAT method combines elements from previous methods, to test for a more representative measure of differential expression [124]. Although considered relatively conservative, the TREAT method has been shown to outperform other t-test based methods for differential expression analysis, while effectively controlling the FDR. The number of DEGs were drastically reduced in each pairwise comparison demonstrating the profound effect this method has on the data. A more stringent 2-fold cut-off was used as a means of compensating for having fewer biological repeats, and making it more likely that biologically significant genes would be identified.

The differences seen in both strains at 30 minutes was an unexpected outcome as we knew that teicoplanin is a weak inducer for the expression of *van* genes, so it was assumed that deleting *vanJ* would not have a significant impact on the teicoplanin susceptibility of *S. coelicolor* [110]. However, we noted that there were significant differences in the number DEGs between the 30 minute time points of M600 and Δ*vanJ* strains, as shown by comparison E and H (**Table 3.3**). There were substantially more DEGs in Δ*vanJ*, demonstrating the pleiotropic effects caused by the loss of *vanJ* when cells are exposed to teicoplanin. This supports the hypothesis proposed from the PCA in **Figure 3.6**, that most of the variation between the datasets occurs at the 30 minute time points and that VanJ may play a role in transcriptional regulation within *S. coelicolor*.

Here we present for the first time, evidence that there is activity of VanJ in the presence of teicoplanin in *S. coelicolor* even though teicoplanin is a poor inducer of the *van* system. Secondly, we also show that it is likely that VanJ has a role to play in the immediate response to teicoplanin as the expression profiles of both strains differs most at 30 minutes after exposure

to the antibiotic. However, as there is less variation between the strains at later time points, the loss of VanJ must be compensated for during later time points in Δ*vanJ*.

### 7.1.5 A cluster analysis demonstrated the profound differences in the initial response to teicoplanin between the two strains

To elaborate on a possible role for VanJ in *S. coelicolor*, DEGs were clustered into seven distinct clusters in order to make them more manageable than a larger dataset. A second benefit of clustering is that genes that are regulated similarly are more likely to be grouped together as genes that are regulated in a similar way are more likely to share similar regulatory pathways and be involved in corresponding functional processes [142].

Again, this data offers further supported that there were significant differences in how the two strains responded to teicoplanin. From the data, seven distinct clusters were produced that could be classified into four different groups depending on their regulation within the two strains. Clusters 2, 4 and 6 all contained genes that were mostly DE in Δ*vanJ,* confirming that there is a different response to teicoplanin in the absence of *vanJ* (**Figure 3.7**). Clusters 3, 5 and 7 contained genes that were induced in both strains, but it was clear from Cluster 3 and 5 that genes within these clusters are induced more slowly within Δ*vanJ* (**Figure 3.7**).

Though the specific role of VanJ remains elusive, the evidence presented here demonstrates that the weak induction of VanJ caused by teicoplanin is enough to profoundly affect the gene expression of M600 in comparison to Δ*vanJ*. VanJ belongs to a broad class of EEP enzymes which all cleave phosphate groups, but can be involved in many different cellular functions. Other bacterial members of this family include YafD from *Samonella enterica*, which provides resistance to egg yolk albumin [238], and *Spy0747* from *Streptococcus spp*. which is believed to be a nuclease involved in virulence [239]. The enzymatic processes encoded for by EEP domains, could suggest that VanJ may have some role in cell signalling, but how this occurs is still unclear. We discuss this further in the following section and how this relates to the regulatory systems of *S. coelicolor*, but further work will need to be carried out to elucidate the mechanisms underlying the differences between M600 and Δ*vanJ*.

## 7.2 Genome-wide transcriptional analysis of the response of M600 and Δ*vanJ* towards teicoplanin

To interpret the DE analysis and identify higher biological functions that were affected by teicoplanin exposure, two approaches were used to assign meaning to the RNA-seq dataset. The goal of using these two methods was not for comparison, but instead to gain a better idea of how teicoplanin affected the gene expression in both strains. The first method used information currently available on the genome of *S. coelicolor*. The second attempted to assign a higher biological function to each of the genes found within each cluster through an enrichment analysis using a gene ontology (GO) database. The former method is more laborious, requiring extensive knowledge of the genetics of *S. coelicolor*, however it allows more freedom when you want to obtain information about biological processes with limited information. The second method was a more efficient method for assigning function to genes but relies on curated databases which can quickly become outdated for organisms are not well studied. This section will look at the genes that were identified using either of the two methods and what can be inferred from the DEGs [122,125,126].

### 7.2.1 Differences observed in signal transduction systems and how this gives insight into a possible role for VanJ

In the previous section, it became clear that there were critical differences between the responses of the two strains to teicoplanin, and analysis of the DEGs highlighted that this likely occurs because of the deregulation of certain transcriptional regulators. Both strains showed increased activity in regulators for cell envelope, oxidative and osmotic stress responses, with declines in the latter during later time points. This is supported by data on other cell wall targeting antibiotics and shows the overlap of stress responses for seemingly different types of cell stress[149].

What was apparent between the two strains was that genes involved in the $\sigma^E$ signal transduction system remained active across the 90 minutes. This confirmed previous work and demonstrated the immense cell wall stress caused by teicoplanin in *S. coelicolor* [100]. In the presence of glycopeptides such as vancomycin, expression of the $\sigma^E$ signal transduction system declines as the *van* genes are induced. This gradually decreases cell wall stress over time until the damage can be rectified in cells. The lack of decline observed in M600 and more so in

Δ*vanJ*, demonstrates how potent teicoplanin is to these cells. We were surprised to see that the loss of *vanJ*, strongly affected the rate of induction for the $\sigma^E$ signal transduction system which may suggest that VanJ is involved in activating CseC SK. Currently, the nature of ligand perceived by the CseC sensor protein is unknown, but due to its ability to respond to a broad range of cell envelope stress factors, it is improbable that each of these factors can act as the direct ligand. Instead, it has been suggested that CseC may detect an intermediate that accumulates as a result of cell envelope damage. Alternatively, a biophysical property of the cell envelope, such as turgor pressure could activate CseC [151]. Our data is more logically explained by the first theory, as an intermediate generate through the action of VanJ could interact with CseC increasing the rate of induction. However, as there is still induction of the signal transduction system, the intermediate could still be accumulated through other unknown mechanisms, leading to the delayed response observed in Δ*vanJ*.

### 7.2.2 The $\sigma^E$ regulon and its likely importance in response to teicoplanin

The $\sigma^E$ regulon within *S. coelicolor* was recently published, and it emphasised the pleiotropic nature of this sigma factor [150]. It has been found to regulate over 50 targets with direct implications in PG biosynthesis, cell wall teichoic acid deposition, sporulation, and membrane surface modifications. Unsurprisingly, a number of the genes mentioned in Chapter 4 also belonged to the $\sigma^E$ regulon (**Figure 7.1**), with all genes showing increased activity in both strains after exposure to teicoplanin. As previously discussed, genes in Δ*vanJ* were induced at a slower rate than in M600, but the activity for most of these genes was similar to that of M600 by 90 minutes.

The regulon included four PBPs (*SCO1875*, *SCO2897*, *SCO4847*, and *SCO5039*) were included in this list with *SCO1875*, *SCO2897* and *SCO5039* previously implicated in resistance to vancomycin [149]. The *mprF* gene also belongs to this regulon and has been found to protect cells against the osmotic stress caused by bacitracin, vancomycin, and CAMPs by modifying the net charge of the membrane by adding lysyl-groups to PG, reducing the interactions at the membrane surface [53,54,149]. Finally, several genes with predicted roles in phospholipid metabolism were also found in the $\sigma^E$ regulon. Of these, this included a gene encoding predicted phospholipase (*SCO2892*) as well as genes encoding PAPs (*SCO2807*, *SCO4133* and the *SCO6357-53* operon containing *SCO6356* and *SCO6355*).

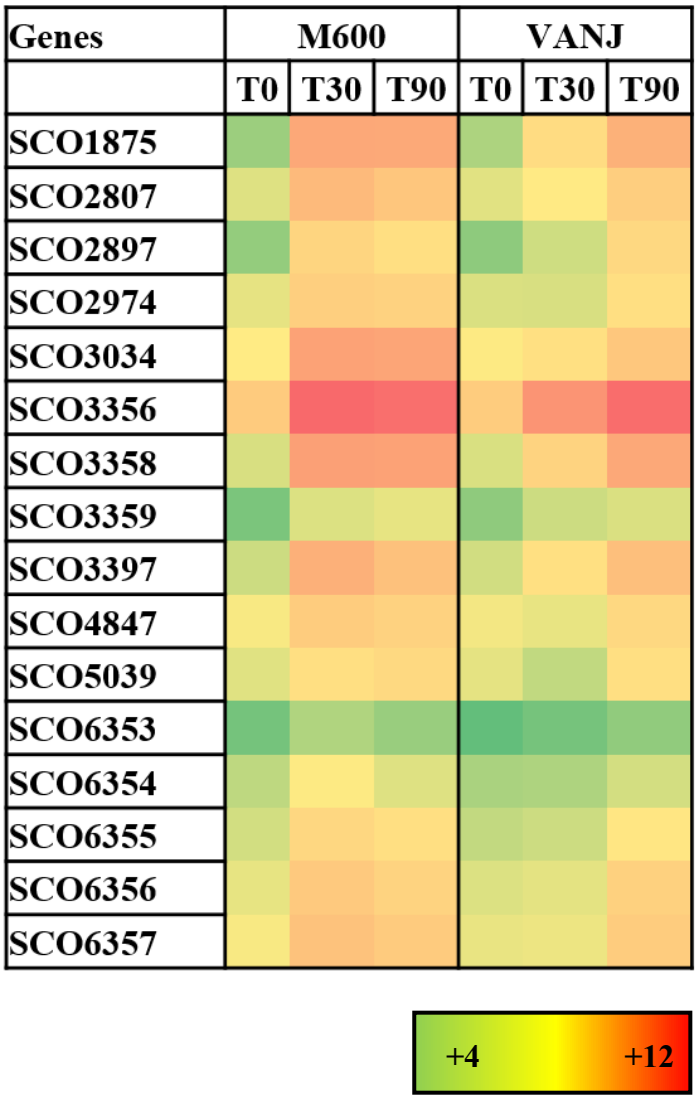



**Figure 7.1** Heatmap showing the relative changes in $log_2$ fold expression for genes in the published σE regulon mentioned in this work between M600 and Δ*vanJ* strains. Key in the bottom right shows which colours represent which log2 values. *SCO3356* represents $\sigma^E$.

Together it is likely that there is a broad response to teicoplanin with many genes working synergistically to enact a defence against this drug. Further functional characterisation of some of these individual genes will be required to pinpoint which genes if any are more important in response to teicoplanin, however the ones tested in this study did not affect the sensitivity of *S. coelicolor* towards teicoplanin (**Section 7.3**).

### 7.2.3 Δ*vanJ* mutant exhibits more significant cellular stress than the wild type M600 in response to teicoplanin

*Streptomyces* are non-motile bacteria that cannot move to new sources of nutrition during times of starvation. Instead, *Streptomyces* switch from vegetative growth to reproductive growth, producing secondary metabolites and spores which are more capable of surviving in adverse conditions than vegetative cells. There are many factors involved that control these intricate processes [240,241]. However, one of the more important factors is the highly phosphorylated guanine nucleotide (p)ppGpp master regulator, known as the stringent factor. The stringent factor is generated predominantly by the ribosome-bound RelA, which synthesises the stringent factor from ATP and GTP when an uncharged tRNA binds the acceptor site of a translating ribosome during times of nutritional stress [148,242–244]. The increased intracellular pool of stringent factor directly binds the RNA-polymerase, directing gene transcription away from genes involved in growth, including stable RNAs (**Figure 4.2**), ribosomal proteins (**Figure 4.2**), DNA replication (**Figure 4.4**), FAB (**Figure 4.10**), PG biosynthesis (**Figure 4.7**) and membrane lipids (**Figure 4.11**). Conversely, genes involved in other stress responses (**Figure 4.1**), proteolysis, amino acid metabolism (**Figure 4.5**), cellular communication, antibiotic resistance and morphological development (**Figure 4.6**) are all known to be positively regulated by the stringent factor in *Streptomyces* [148,149,245].

Although no significant changes in the expression the ppGpp synthetase, *relA* were observed (data not shown), the data presented here indicate that teicoplanin affects similar processes as ppGpp. Hesketh *et al*., (2011) have previously commented on this effect, where vancomycin, bacitracin and moenomycin all modulated the gene expression of *S. coelicolor* as if ppGpp was being synthesised. However, ppGpp could not be detected after exposure to vancomycin, so it was suggested that cell wall targeting antibiotics could modulate some of the same genes as ppGpp or trigger the same signalling cascades. It is not clear why these processes are more significantly affected in Δ*vanJ*, but the absence of VanJ and the subsequent delay in activating the cell envelope stress response could mean Δ*vanJ* relies on other stress responses to cope

with the stress caused by the drug immediately after exposure until alternative pathways can be switched on. This further demonstrates the important role of VanJ in the presence of teicoplanin.

### 7.2.4 Developmental genes that were influenced by teicoplanin exposure

As we have discussed, teicoplanin affects *S. coelicolor* similarly to ppGpp, which is activated under nutritional stress. We also found evidence that some genes involved spore development were also induced in response to teicoplanin including the chaplin genes, the *ramCSAB* operon, *ssgB* and the late spore developmental genes *whiI*, *sigF* and *whiE* orfs *I-IV* and *VIII*. Sporulation occurs in *S. coelicolor,* usually during a nutritional downshift, leading to a colony-wide shift from vegetative growth to reproductive growth [172,178]. *Streptomyces* rarely sporulate in submerged cultures, and it is believed that *S. coelicolor* fails to progress from the pre-sporulation stage of development to form mature spores in submerged culture. Transcriptomic analysis of cultures grown on solid and liquid substrates have shown a surprisingly small number of differences in the activity of genes in the two different conditions. Failure to progress developmentally appears to relate to the inactivity of genes involved in the hydrophobic coat or spore pigment and includes the *chpF*, *rdlA* and the *whiE* genes [246].

Studies have demonstrated that modifying the media with calcium and proline [178], mechanically breaking large mycelial clumps into smaller fragments [172], over-expressing *ssgA* [247], or knocking out *argR* [248] can all lead to the production of spore-like structures in *S. coelicolor*. However, even though these structures are morphologically related to spores, they are often less stable than true spores [178]. Although the complexities of sporulation in submerged cultures are still poorly understood in *S. coelicolor*, there was evidence to suggest that teicoplanin induced several systems involved in the production of spores in the absence of VanJ. Interestingly, there were indications that some of the late sporulation genes mainly involved in spore pigment production and hydrophobic coat formation were transiently activated in response to teicoplanin in Δ*vanJ* (**Figure 4.6**). This information, along with the induction of several genes involved in hydrophobic coat formation and late spore development, could implicate that a developmental shift from mycelial growth to sporulation is a survival strategy for Δ*vanJ,* however further work will need to be carried out to identify whether these transcriptomic changes relate to phenotypic changes in spores. .

Remarkably, both strains also prominently overexpressed the *wbl* genes, *whiB* and *whiD*. *whiB* encodes a developmental regulator that regulates the length of aerial hyphae and location of sporulation septa and the regulator encoded by *whiD*, which is involved in the maturation of cell wall in developing pre-spores [185]. Curiously, *whiB* is also part of the $\sigma^E$ regulon, and some *whiB* homologues in *M. tuberculosis* can be induced in response to particular cell wall stress [249] indicating that *whiB* could be broadly involved in the cell wall stress response. Both *whiB* and *whiD* also belong to an actinomycete specific gene family which encode Wbl-proteins with [4Fe-4S] clusters, that have redox properties [188]. The Wbl proteins are known to interact with the gaseous molecules of nitric oxide (NO) which has an essential role in many bacterial processes including host defence, AMR and cellular communication during development in the *Streptomyces* [188,250,251]. We did note that after exposure to teicoplanin, genes involved in the NO metabolic processes were induced in both strains. If NO is generated as a result of the stressful conditions caused by teicoplanin, it could act as one of the regulatory ligands for the Wbl-proteins and promote the development of spores [241]. Both proteins may broadly slow the growth of hyphae and thicken the cell walls of spores, which we hypothesise could counteract the effects of teicoplanin by preventing lysis in a subset of the population, and possibly ensure the survival of some spores from the colony until teicoplanin is no longer at lethal concentrations.

### 7.2.5 Changes in arginine metabolism could be linked to osmoprotection

Genes relating to the regulation and biosynthesis of arginine were strongly induced in both strains (**Figure 4.5**). This also included the pleiotropic negative regulator of the *arg* genes, ArgR, which has also been demonstrated to positively regulate the expression of ACT and RED through comparisons of an *argR* knockout strain with wild type cells in *S. coelicolor* [248]. Arginine itself is an essential precursor for several secondary metabolites in the *Streptomyces*, including clavulanic acid from *Streptomyces clavuligerus* NRRL 3585 and streptomycin from *Streptomyces griseus* [252,253]. There is little information available on the role of arginine in secondary metabolite production in *S. coelicolor* and to our knowledge it is not involved in the synthesis of ACT or RED. Nevertheless, in *S. coelicolor*, arginine has been found to increase in concentration in response to increased salt stress, suggesting that it could act as an osmoprotectant alongside ectoine and proline [170]. Previously published studies have demonstrated that vancomycin increases the concentration of arginine residues in the cytosolic fraction supporting the evidence that arginine could be acting as an osmolyte in the presence

of glycopeptide antibiotics [204]. It would be interesting to develop this further and identify whether arginine is necessary for tolerating glycopeptides and why this is the case. It would also be interesting to identify whether similar outcomes are also observed in clinical pathogens which could identify whether the arginine pathway could be a future drug target providing the enzymes involved in this pathway were sufficiently different from eukaryotic counterparts.

### 7.2.6 Elevated transcription of *olsAB* in Δ*vanJ* further suggests that teicoplanin induces developmental processes

Ornithine is one of the intermediates in arginine biosynthesis, and it can also be used in the production of clavulanic acid in *S. clavuligerus* [252]. In *S. coelicolor*, it is involved in the production of the phosphate-free membrane lipid, OL. The data here not only presented increases in the expression of genes involved in arginine biosynthesis but also *olsAB* in the Δ*vanJ* background (**Figure 4.11**). *olsAB* are required for the production of OL in bacterial membranes, and this membrane lipid has been identified in *S. coelicolor* cultures that are in the later stages of development or under phosphate limiting conditions [60]. There were no indications of phosphate starvation in either strain of *S. coelicolor* (data not shown) therefore it seems likely that the increased activity of these genes could be linked to the increased activity seen in genes relating to spore maturation and development. Further work would need to be done to understand whether these genes are related to antibiotic resistance as the literature currently does not indicate any relationship between OL production and the glycopeptide antibiotics, but with OL being a zwitterionic phospholipid, this does not rule out that there could be additional changes in membrane charge brought about in Δ*vanJ* after exposure to teicolanin.

### 7.2.7 Changes in membrane transport

Membrane transporters are essential for cellular function, allowing the selective transport of metabolites, either actively or passively across the membrane. They are implicated in protein secretion [254], provide either self- or acquired-resistance to antibiotics [29,255], regulate the membrane potential [256] and are also involved in the flippase activity required to transport charged lipid carriers involved in the biosynthesis of cell wall polymers [46]. Both *SCO3824* and *SCO4359* encode proteins belonging to the ABC-transporter superfamily, which use ATP to actively transport a broad range of substrates (**Figure 4.8**). Many drugs such as macrolides and bacitracin are transported across the membrane in this manner [29,255]. Neither gene has been

reported to be involved in the general cell envelope stress response of *S. coelicolor*, but ABC-transporters are frequently identified in cell envelope stress regulons for different antibiotics [191,257–259]. Although it is not clear what role these transporters play in response to glycopeptide antibiotics, they are possibly involved in the removal of toxic metabolites such as free radicals, from the cell, if they cannot be metabolised.

A gene encoding the third membrane transporter, *SCO3910*, was also upregulated in response to teicoplanin. Predicted to encode a MOP transporter (**Figure 4.8**), which are widespread amongst eukaryotes and prokaryotes and couple substrate export with the flow of $K^+/Na^+$ ions down an electrochemical gradient [193,260]. Not only are they prolific transporters for toxic metabolites, they are also implicated in the transport of intermediates that are used in the generation of cell wall glycans such as PG, LPS and the capsular polysaccharide in bacteria [261]. Although there were no increases seen in the activity of genes involved in PG biosynthesis in response to teicoplanin, vancomycin-intermediate *S. aureus* (VISA) has been known to have unusually thick cell walls that prevent the drug from reaching the membrane surface. As there were no increases observed for genes involved in PG biosynthesis, it is unlikely that SCO3910 is exporting PG precursors, but there is still the possibility that it may play a role in the transport of other lipid carriers that could prevent the access of teicoplanin to the membrane [257].

*SCO2472* has not previously been implicated in the cell envelope stress response to other antibiotics in *S. coelicolor* [149] (**Figure 4.8**). However, its homologues *sanA* from *E. coli* and *sfiX* from *Salmonella typhimurium* are both implicated in restoring intrinsic vancomycin resistance in vancomycin-sensitive mutant strains. The expression of *sanA* is also essential for determining the sensitivity of *E. coli* to detergents and bile salts, suggesting that these genes could be broadly specific for in environments that affect the outer membrane of Gram-negative bacteria [192,219]. *SCO2472* showed an extraordinary increase in activity in response to teicoplanin in both M600 and Δ*vanJ* mutant strains, suggesting that *SCO2472* may also play a role in the barrier function of *S. coelicolor* towards teicoplanin, however it is unclear whether this gene will have the same function in *S. coelicolor* as it lacks an outer membrane.

### 7.2.8 Evidence of phospholipid remodelling in *S. coelicolor* cell membrane in response to teicoplanin

Bacterial membranes are highly dynamic structures and are remodelled continuously in response to environmental changes [60]. Clear evidence for phospholipid remodelling was

identified in this study, showing that this is one response to teicoplanin exposure (**Figure 4.12**). Other genes have previously been identified, which remodel the membrane surface affecting its properties towards cell wall stress. *mprF* was strongly upregulated in this dataset and is known to form part of the $\sigma^E$ regulon and provides resistance to both vancomycin and bacitracin in *S. coelicolor* [149,150]. The product of *mprF* was first identified as an enzyme that could reduce the overall negative charge at the surface of bacterial membranes by adding lysine groups to PpG [53]. In doing so, bacteria can broadly prevent molecules such as cyclic antimicrobial peptides, daptomycin and vancomycin from reaching their target sites at the membrane surface [53,262,263].

Several other genes that are likely to play roles in the modification and turn-over of membrane phospholipids were also upregulated. Two major groups of enzymes have been identified to be involved in phospholipid turn-over in bacteria, and they prevent the accumulation of phospholipid degradation products that could affect the membrane architecture [58]. These include the phospholipases and PAP2 enzymes. A group of genes encoding lipases were identified in the cluster analysis including *SCO3222* which was predicted to encode a $PLA_2$ enzyme responsible for cleaving the 2-acyl bond of glycerophospholipids, producing a fatty acid and an LPA [194,197]. The gene encoding $PLA_1$ responsible for cleaving the 1-acyl bond of glycerophospholipids in *S. coelicolor* is currently unknown, but one other gene from this ontology, *SCO2892*, could fulfil this role. The *SCO2892* protein is uncharacterised but has the SGNH-fold that is found in a promiscuous group of enzymes that have broad hydrolytic activities including the *E. coli* thioesterase I/protease I/lysophospholipase (TAP) protein, with broad specificity for mono-acyl phospholipids, such as LPA [198]. The products of *SCO3222* and *SCO2892* may work in unison to liberate fatty acids from phospholipids in response to teicoplanin.

The role of enzymes containing PAP2 domains is becoming more prominent in membrane homeostasis because of their prolific phosphatase activity. Other members of this family include BcrC from *Bacillus sp.* and PgpB from *E. coli*. BcrC is a UPP-phosphatase that competes with bacitracin for its target, UPP, increasing bacitracin resistance when expressed [45]. PgpB is more diverse in its phosphatase activity and has not only demonstrated UPP-phosphatase activity but the ability to cleave a wide range of phosphodiesters [68]. Three genes with PAP2 domains have previously been described in *S. coelicolor* including *ippα* (*SCO1102*) and *ippβ* (*SCO1753*) which are involved in the production of DAG from PA in the

triacylglycerol (TAG) biosynthesis pathway [264]. The product of *SCO6355* is also predicted to belong to this family; however, no functional role has previously been assigned [67]. A BLAST search revealed that 15 other genes encoded PAP2 domains in the *S. coelicolor* genome. 12 out of 15 genes were differentially expressed in response to teicoplanin with six positively regulated in both strains. The genes, *ippα* and *ippβ* were not differentially expressed, but *SCO6355* was along with the *SCO6537-53* operon. There appeared to be converse regulation of *SCO7587* in comparison to those positively induced PAP2 genes highlighting that these genes could have distinct functions within the membrane. The strain specificity observed in the increased expression of *SCO1047*, *SCO4843* and *SCO6428* also supported that these genes may have specific roles in each strain and that the transient inducted of these genes in particular may also be linked to the increase in developmental genes associated with spore development previously discussed in the Δ*vanJ* background.

Further support for the remodelling of the cell membrane came from an increase seen in the two GDPD encoding genes and the glycerol operon. GDPDs carry out the final step of phospholipid recycling by degrading glycerophosphodiesters that are released during the catabolism of phospholipids through lipases and PAP2s [199,200]. This allows cells to recycle the liberated G3P which can then be used for further phospholipid biosynthesis or as a carbon source if converted to DHAP [62]. The process that transports G3P back into cells is not known in *S. coelicolor*, but the GlpT transporter carries out this role in *E. coli*. It is also not entirely clear if *S. coelicolor* possess a G3P-phosphatase to degrade G3P into glycerol and Pi, but we know that glycerol is passively transported back into cells aided by the facilitator, GlpF. The kinase, GlpK and the G3P-dehydrogenase, GlpD are encoded for by adjacent genes in the glycerol operon and are responsible for the phosphorylation of glycerol to G3P and the dehydration of G3P to DHAP. The operon these genes belong to is strongly induced in the presence of glycerol [202,203], which we observed in our dataset suggesting that there are significant increases in glycerol which likely originated from the membranes of bacteria through either phospholipid degradation or lysis of neighbouring cells. Further understanding of the mechanisms behind this observation will be important in developing our understanding of how *S. coelicolor* adapts to and counteracts the stress caused by teicoplanin.

## 7.3 Functional analysis of selected genes

### 7.3.1 Challenges with generating knock out mutants using CRISPR

Both overexpression and knock out strains are essential tools in functional genomics for studying the roles of specific genes. The methods for overexpressing genes in *S. coelicolor* have been well established, but the historical methods of gene editing were time-consuming and somewhat limited [265]. They also did not allow for adequate multiplexing of knockouts which can be necessary for *Streptomyces* due to the high level of redundancy within their genomes. At the time of these experiments, four CRISPR toolkits were available for the *Streptomyces* spp. which have previously been reviewed [266].

We pursued two novel CRISPR systems that were supposedly optimised for the *Streptomyces* spp. The pCRISPomyce-2 plasmid has been used with high success to engineer several species of *Streptomyces* including *S. lividans* [130]. The novel *Streptomyces* pCRISPomyces-2 plasmid system was engineered to target the *vanRS* cluster in M600. Theoretically, knocking out *vanRS* should strongly sensitise strains of *S. coelicolor* to vancomycin so identification of desired mutants would be easy to identify with replica plating on selective and non-selective media. Unfortunately, we were unable to obtain any successful exconjugants as confirmed by PCR, for the pCRISPomyces-2 CRISPR vector nor its derivative plasmids pCRvanRS-1 and -2. Since the time of this study, Wang *et al*., (2016) have published a more comprehensive protocol for gene editing with this system in the *Streptomyces* [265].

We also tested a second CRISPR plasmid, pKCcas9dO because it has previously been successfully conjugated into *S. coelicolor* M145 [131]. However, we also failed to conjugate this plasmid into M600. It has been suggested that some CRISPR plasmids are toxic to strains of *S. coelicolor* such as M600 (M. Hutchings, personal communication) due to the generation of DSB at off-targets [266]. To curtail this problem, a novel CRISPR system has been developed, which involves DSB-free editing. The CRISPR-cBEST system instead can convert a single cytidine to thymine, allowing the insertion of stop codons into DNA sequences with high-fidelity [267]. Theoretically, this system simplifies the amount of genetic manipulation needed to edit the genome and can be multiplexed to improve the outcome of CRISPR in *Streptomyces*. There is currently the intention for this work to be carried on through a future MRes project to confirm whether it is possible to use either of the two DSB or DSB-free CRISPR editing

methods in *S. coelicolor* M600. This will become an invaluable tool for understanding the relationship of the genome and the phenotype of cells in the presence of different antibiotics.

### 7.3.2 The influence of the glycerol operon on the susceptibility of *S. coelicolor* towards glycopeptide antibiotics

The glycerol operon consists of *glpFKD* and is known to contribute to escalating the rate of diffusion of glycerol across the membrane via a membrane transporter encoded by *glpF* (**Figure 5.6**). Once glycerol diffuses across the membrane, it is phosphorylated by GlpK to produce G3P. G3P can be used for phospholipid biosynthesis or shuttled into the glycolytic pathway after being converted to DHAP by the product of *glpD* [62]. These genes in the glycerol operon were selected for functional study in *S. coelicolor* M600 and Δ*vanJ* background strains for MIC testing against a library of antibiotics. The library consisted of six antibiotics that targeted PG biosynthesis, two antibiotics that targeted the cell membrane, one antibiotic that inhibited DNA replication and one antibiotic that inhibited protein synthesis (**Table 5.1**).

The glycerol operon of *S. coelicolor* is subject to catabolite repression, being inhibited in the presence of glucose, but induced in the presence glycerol [203]. It has been proposed that the system is in place to help maintain intracellular concentrations of G3P, which acts as a building block for cell wall teichoic acids and phospholipids, or can be converted to DHAP and shuttled into glycolysis [202,203]. The glycerol operon is not known to be involved in the response to vancomycin in *S. coelicolor* nor is it involved in the response to other cell wall targeting antibiotics. GlpD has been shown to be essential for the tolerance for the beta-lactam, ampicillin and ofloxacillin, and deletion of this gene decreases the amount of persister cells in chronic *E. coli* infections which is important in the development of multidrug tolerance [268]. Overexpression of *glpF* or *glpK* results in *S. coelicolor* becoming more sensitive to teicoplanin, but this result was observed only in the Δ*vanJ* background. Intriguingly, integration of plasmids containing any gene from the glycerol operon led to the suppression of the vancomycin-sensitive phenotype in the Δ*vanJ* strain (**Table 5.3**). It is not clear why this occurs, but possibly the same mechanism that promotes persister cell formation in *E. coli*, may also benefit Δ*vanJ* in the presence of vancomycin, possibly through increasing the amount of DHAP available for glycolysis [202]. Alternatively, glycerol can also act as an osmoprotectant in some species such as yeast [269], and may therefore provide protection from the increased osmotic stress caused by teicoplanin. Both of these hypotheses could be involved in counteracting the negative impact of the loss of VanJ, restoring high-level vancomycin resistance to Δ*vanJ* background strains

### 7.3.3 The results of the colourimetric test suggest that lethal concentrations of antibiotics lead to increased concentration of glycerol in the media

MIC testing suggested that some of these glycerol operon genes could be decreasing the sensitivity to vancomycin in the Δ*vanJ* background mutant, so we were still interested in exploring the mechanisms behind the strong induction of the glycerol operon in response to teicoplanin. Using a colourimetric assay developed by Bompelli and Skaf (2014), we were able to quantify the concentration of glycerol in cultures of *S. coelicolor* after exposure to antibiotics in **Chapter 6**. The assay has previously shown to be reasonably robust in complex mixtures of biodiesel containing complex macromolecules [223]. The NMMP media used to grow submerged *S. coelicolor* cultures was also found to generate background noise, which could be offset with absorbance values that were adjusted to take into consideration a 'media control'. We also found that the assay was replicable, providing the incubation periods were strictly followed, and samples were stored on ice while setting up each reaction.

The data generated in **Figure 6.5** supported the hypothesis that glycerol concentration did increase in the media after exposure to teicoplanin, but this was also observed for the other cell wall targeting antibiotics bacitracin and ramoplanin. Furthermore, a dosage response indicated that the concentration of antibiotic applied to the sample affected the concentration of glycerol in the media, with higher concentrations of both teicoplanin and kanamycin causing the same response (**Figure 6.6**). Together these data suggest that the increase in glycerol concentration might be the result of lethal concentrations of antibiotic which in turn lead to increased activity on the glycerol operon rather than a specific mechanism involved in the response to teicoplanin. The greater increase in these genes in Δ*vanJ* is likely due to the loss of VanJ in its membrane, which we now know leads to increased cellular stress, and therefore are likely to exhibit greater cellular death in response to teicoplanin.

### 7.3.4 The negative impact of *SCO2472* and *SCO3910* on glycopeptide sensitivity

Two genes from the 'response to drug' ontology were also selected and overexpressed to see their effect on the MIC of glycopeptide antibiotics. *SCO2472* was chosen for its strong resemblance to the *E. coli*, *sanA* and *S. typhimurium*, *sfi* genes. Although these Gram-negative bacteria usually show intrinsic resistance to vancomycin, both of these genes have been implicated in complementing a vancomycin-sensitive phenotype in their respective strains

[192,219]. Overexpression of the *sanA* gene also abolishes sensitivity to detergents, and Rida *et al.* (1996) suggest that *sanA* could, therefore, be broadly implicated in the barrier function of *E. coli*. Unexpectedly, overexpressing *SCO2472* increased sensitivity to teicoplanin in both background strains and increased sensitivity to vancomycin in the M600 background strain. It can be hypothesised that *SCO2472* may have a different function in the cell envelope in Gram-positive bacteria, or both genes form part of the general cell stress response. The other gene selected for overexpression study was *SCO3910*. With a predicted MOP domain, this gene belongs to a family of genes encoding membrane transporters that have broad functionality. This includes the predicted lipid II flippase, MurJ, the *E. coli* multidrug resistance protein, MtdK, and the glycolipid transporters for cell wall glycans [38,220,261]. The same phenotypes were observed in strains overexpressing *SCO3910* as *SCO2472* (**Table 5.3**). Further work still needs to be done to identify whether overexpression of these genes increases the susceptibility of *S. coelicolor* to other cell wall specific antibiotics. However, it is clear that they both negatively impact the function of the cell envelope and likely increase cellular stress in the presence of glycopeptide antibiotics.

### 7.3.5 The involvement of PAPs in the maintenance of the cell membrane in the presence of cell envelope targeting antibiotics

PAPs are a large family of enzymes found in prokaryotes and eukaryotes, with roles in cleaving phosphodiester bonds found in many membrane lipids. They all can be characterised by a KXXXXXXRP-($X_{12-54}$)-PSGH-($X_{31-54}$)-SHXXXD (PAP2) motif [270]. Prominent members of this family include BcrC and PgpB which are both involved in resistance to bacitracin [45,68,69,271]. Ippα and Ippβ have been characterised in *S. coelicolor* for their role in TAG biosynthesis, cleaving PA to produce DAG in the TAG biosynthetic pathway [264]. The diverse functionality of these enzymes made them an exciting group of genes to pursue for their potential involvement in their response to teicoplanin.

The data presented here demonstrated that overexpressing some of the PAP encoding genes in *S. coelicolor* widely affects sensitivity towards a range of cell wall targeting antibiotics. *SCO2335*/*SCO2807* and *SCO6356*/*SCO6511* showed close phylogenetic relationships along with highly similar MIC profiles against the antibiotic library when overexpressed, indicating that the pairs are functional homologues. The subtle differences in the MIC patterns between these gene pairings also support the hypothesis that there are subclasses of PAP in the *S. coelicolor* genome.

The overexpression of any of these PAP genes increased the sensitivity of *S. coelicolor* to the majority of cell envelope targeting antibiotics. We note that this is in stark contrast to other membrane modifying enymes such as MprF which is known to add lysyl groups to PG decreasing the negative charge of the membrane and broadly increasing resistance to osmotic stress caused by CAMPs, vancomycin, bacitracin, moenomycin and daptomycin [53,54,66,272]. It is plausible that the PAPs presented here could be carrying out an opposing role that increases the negative charge of the membrane by generating negatively charged membrane lipids. Decreased concentrations of PpG in the membrane are associated with high resistance to antibiotics such as vancomycin and daptomycin [273], so it is rational to assume that overexpressing these genes could be generating a higher concentration of PpG in the membrane. Alternatively, as these PAP genes were all upregulated in response to teicoplanin; one cannot also rule out the possibility that these genes may also work cooperatively, which could explain why any strain expressing genes encoding PAPs, including the three genes (*SCO2807*, *SCO6355* and *SCO6356*) that belonged to the $\sigma^E$ regulon [150], did not show any increase in resistance to teicoplanin. Understanding these mechanisms is important because these enzymes are widely distributed amongst bacteria, including clinical pathogens, and therefore may serve as future drug targets to exploit.

### 7.3.6 The proposed role of *SCO6355*

Although the function of the *SCO6357-53* operon remains to be understood, overexpression of *SOC6355* appears to sensitise *S. coelicolor* to a narrow range of glycopeptide antibiotics (vancomycin, balhimycin and ristocetin) which it normally displays high-level resistance to and the lantibiotic duramycin. Interestingly, however, overexpression of this gene did not change the susceptibility of *S. coelicolor* toward the lipoglycopeptides (e.g. teicoplanin, telavancin, dalbavancin and oritavancin), chloroeremomycin (**Table 5.5**) or vancomycin derivatives with putrescine substituted onto the seventh amino acid and a vancomycin-putrescine derivative with a chlorobiphenyl group (data not shown). The protein-coding sequence of *SCO6355* also showed a close phylogenetic relationship with Ippα and Ippβ (**Figure 5.10**). From this information, several hypotheses can be proposed about the possible function of *SCO6355* in *S. coelicolor*.

Firstly, overexpressing *SCO6355* made *S. coelicolor* hypersensitive to the class I type B lantibiotic duramycin. This 19 amino acid tetracyclic peptide that is effective against both Gram-positive bacteria and some fungi, and it is known to specifically target zwitterionic

ethanolamine phospholipids which includes one of the most abundant bacterial phospholipids, PE [274,275]. Annotated sequence data has shown that *SCO6355* lacks the phosphatidylserine decarboxylase domain present in Psd that is used to convert PS to PE. Instead, its product may generate a more unstable phospholipid such as PA from either PG or CL, that is shuttled along the PE biosynthetic pathway indirectly increasing the abundance of PE [67]. It is also possible that the product of *SCO6355* could carry out a multiplicity of reactions similar to PgpB from *E. coli*, generating several membrane lipids, including PE, that make *S. coelicolor* more sensitive to these specific antibiotics.

Secondly, the specificity of the glycopeptide sensitive phenotype observed in *SCO6355* overexpressing construct is not affected by the core heptapeptide backbone structure. The core structure of ristocetin is more similar to teicoplanin with the heptapeptide backbone containing an asparagine and leucine that is replaced by two arylglycine residues (1, 3). In contrast, vancomycin shares only a macrocyclic tetrapeptide (4-7) with ristocetin which is invariant except for the extent of chlorination of ring 6. Additionally, vancomycin, balhimycin and chloroeremomycin all share the same aglycone, yet the sensitivity to chloroeremomycin does not change in strains overexpressing *SCO6355* (**Figure 5.1**) [72,276,277]. The presence of either an aryl- or alkyl- hydrophobic side chain is an important factor in determining whether or not there is a change observed in the MIC as all lipoglycopeptide antibiotics, including a vancomycin putricine-derivative, showed no changes in MIC when comparing controls with strains overexpressing SCO6355.

Other than lipophilic side chains, glycopeptides are commonly modified with the halogens (Br or Cl) and can be glycosylated with an assortment of sugars that can range from the modest aglycone structure to the elaborately decorated ristocetin which adorns six sugar residues [78]. These modifications can improve solubilisation, provide structural rigidity and alter the propensity of some glycopeptides to form cooperative dimers [277–279]. The sugars that are present on position 6 are more important than those in position 4 for increasing the likelihood of forming dimers, but the sugar residues help to form a sophisticated network of hydrogen bonds on the ‘back-face’ of the molecule (**Figure 1.15**). There is speculation that this improves the *in vitro* activity of some glycopeptides such as chloroeremomycin which strongly dimerises, over other glycopeptides like vancomycin, which weakly associates with itself [279–281]. Dimerisation also helps to bury hydrophobic surfaces, but if the expression of *SCO6355* affects

electrochemical properties near the surface of the membrane the propensity of some glycopeptides to dimerise at the membrane surface may increase, improving their activity.

Together these two bits of information suggest overexpressing *SCO6355* may increase the amount of zwitterionic phospholipid PE in the membrane by decreasing the overall negative charge of the cell surface. As a result, *S. coelicolor* becomes more sensitive to duramycin, which targets PE, sequestering it in the membrane. Many modifications to the cell envelope also affect the sensitivity of cells to other drugs, and in this case vancomycin, balhimycin and ristocetin, which all show increased activity against strains of *S. coelicolor* expressing *SCO6355*. What sets these antibiotics apart from other glycopeptides like teicoplanin is that they lack a hydrophobic side chain. In general, glycopeptide antibiotics containing hydrophobic side chains poorly dimerise in comparison with other counterparts such as chloroeremomycin. It seems possible that the surface modification that is induced by the product of *SCO6355* increases the ability of these three glycopeptide antibiotics to dimerise at the cell surface, improving their biological activity (**Figure 7.2**). Alternatively, this modification could also interfere with the ability of *S. coelicolor* to identify the cell wall damage through existing sensory systems or even the VanS sensor kinase to recognise these drugs as ligands. Further experiments will need to be carried out to identify whether either of these hypotheses is correct. Together, this information could also be used to develop strategies for improving the efficacy of vancomycin-like glycopeptides in glycopeptide-intermediate and resistant pathogens; however, more work still needs to be done to understand the mechanisms behind the phenotypes observed in this work.

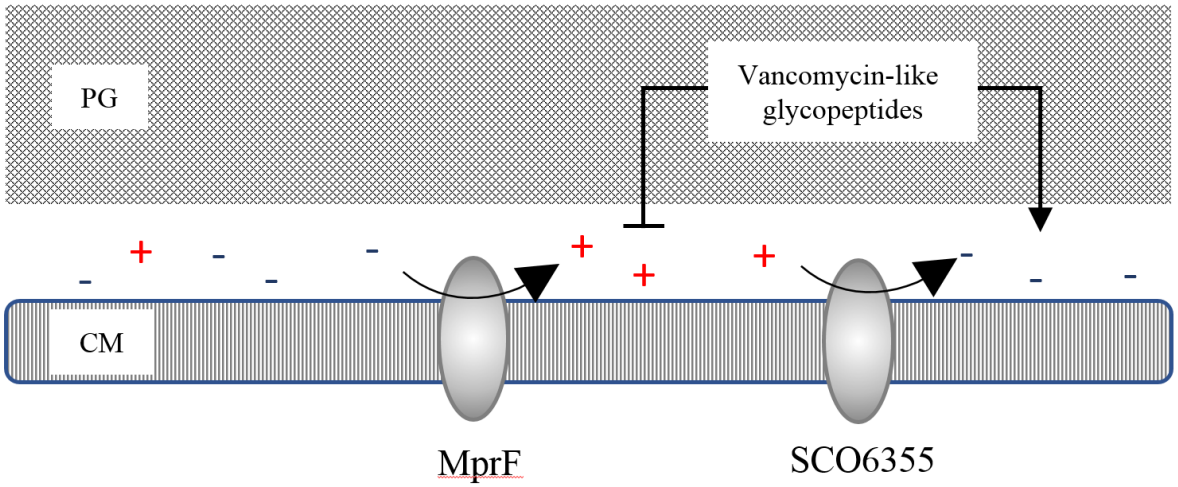



**Figure 7.2** Diagram to show the possible mechanisms of both MprF and the *SCO6355* proteins in the *S. coelicolor* cell membrane. MprF is known to decrease the overall negative charge of the membrane by converting PpG to lysl-PpG, which increases resistance to high osmotic stress and drugs that target the cell envelope. We propose that in this work that the *SCO6355* PAP could be carrying out an opposing role, as cells overexpressing *SCO6355* became hypersensitive specifically to duramycin and exhibited intermediate level resistance towards a narrow group of vancomycin like glycopeptide antibiotics. By affecting the electrochemical properties of the bacterial membrane, the activity of these antibiotics is modulated demonstrating the importance of the bacterial membrane in the mechanism of action of many drugs that target the cell envelope. PG, peptidoglycan, CM, cell membrane, +, positively charged membrane lipids, -, negatively charged membrane lipids.

## 7.4 Conclusions and future perspectives

The first aim of this study was to identify the transcriptomic differences between *S. coelicolor* M600 and a Δ*vanJ* mutant strain in response to the lipoglycopeptide antibiotic, teicoplanin. In doing so, it was hoped that a possible role for the teicoplanin specific resistance marker, VanJ, could be elucidated. The evidence presented here demonstrated that there were profound differences in the transcriptomic responses of the two strains to teicoplanin. This was regardless of teicoplanin acting as a poor inducer of the endogenous glycopeptide resistance cluster TCS, VanR/VanS. From this, we can conclude that even significantly marginal levels of the VanJ protein in the wild type *S. coelicolor* membrane can still influence drastic transcriptomic changes different from a background lacking *vanJ*.

There were approximately double the number of DEGs in Δ*vanJ* compared to M600 after 30 minutes when cells were exposed to teicoplanin. Transcriptomic profiling also demonstrated that Δ*vanJ* cells presented more signs of cellular stress, with systems involved in active growth downregulated during earlier time points. Activation of transcriptional regulators involved in the cell envelope induction system was delayed in Δ*vanJ* when compared with M600, which indicated that VanJ could be involved in the activation of CseC. As transcription of the system did still occur in Δ*vanJ*, VanJ is likely not the only mechanism involved in activating the induction system in the presence of teicoplanin. We also found that some of the genes involved in late spore development were unexpectedly induced in Δ*vanJ,* and two developmental regulators *whiB* and *whiD* were induced in both strains. We propose that the processes regulated by these genes may have a role in counteracting the effects of teicoplanin

To further understand these dynamics, we want to carry out several quantitative Real-Time quantitative Reverse Transcription PCR (qRT-PCR) studies on how the expression of some of these genes are affected between wild type M600, Δ*vanJ* mutant and a *vanJ* overexpressing strain of *S. coelicolor*. Further studies also need to be carried out to distinguish whether the developmental genes involved in sporulation and development that were transiently upregulated in Δ*vanJ* truly relate to morphological changes. Developmental changes such as antibiotic production can be seen with the naked eye in broth, but morphological changes should be visible with light or fluorescent microscopy to visualise spore development. There is also an intention to extend the transcriptomic data presented here to include data on the three other clinically approved lipoglycopeptide antibiotics oritavancin, dalbavancin and telavancin,

to compare with the data in response to teicoplanin in this study. Together, these data would indicate how similarly lipoglycopeptides affect the transcriptome of *S. coelicolor* and how VanJ may protect *S. coelicolor* from teicoplanin.

The second aim of this project was to investigate novel genes for their potential involvement in the adaptive response to teicoplanin. A list of genes were selected for their potential involvement in cell envelope homeostasis. Although overexpression strains were generated from this list of genes, it proved impossible to generate any CRISPR knock out mutants using both the pCRISPomyces-2 and pKCcas9dO vector systems as both appeared lethal in M600. However, a newer DSB-free CRISPR system that involves less genetic manipulation in cells has been developed that we hope will work in M600 [267]. If this system works in *S. coelicolor*, it will also allow the multiplexing of knockouts to identify the architecture of antimicrobial resistance networks. This will prove invaluable for studying whether any of the genes in this study could have been working synergistically with other genes. As no single gene was demonstrated to increase resistance to teicoplanin, it could be that this particular gene list required the interplay of other factors to affect the sensitivity profile to teicoplanin.

We found a family of PAP that specifically increase the susceptibility of *S. coelicolor* to different groups of cell envelope targeting antibiotics. Genes with the PAP2 functional domain have been implicated in several functional roles within bacterial membranes, including bacitracin resistance and phospholipid turnover. Phenotypic data based off of the PAP genes studied indicate that there are at least three different subclasses of PAP within *S. coelicolor*. Combined with the phylogenetic analysis, this value could be much higher. There are indications that the functional subclasses are essential for membrane functionality due to the sensitisation of *S. coelicolor* to a broad range of antibiotics when overexpressing different members of this gene family.

One member of this family, *SCO6355*, was particularly interesting as it generated a hypersensitive phenotype toward the lantibiotic duramycin while creating intermediate levels of sensitivity to several glycopeptide antibiotics, including vancomycin. Because the overexpression of *SCO6355* made cells hypersensitive to duramycin, it is likely its product increases the concentration of PE in the membrane as this is the direct target for duramycin. The close phylogenetic relationship *SCO6355* has with *Ippα* and *Ippβ* indicates that *SCO6355* could also be involved in the synthesis of DAG. It is not uncommon for PAP to produce

multiple products, but further work will need to be carried out to understand the mechanism of this enzyme.

Understanding how SCO6355 increases sensitivity towards vancomycin will be vital in understanding how to counteract endogenous resistance systems such as the *van* cluster. Further experiments will include the purification of the SCO6355 protein for the characterisation of its enzyme kinetics, allowing us to identify its target specificity. We also want to use mass spectrometry or thin-layer chromatography to further characterise the lipid profile of *S. coelicolor* in response to teicoplanin and other glycopeptide antibiotics so that we can highlight whether changes in the phospholipid makeup of the membrane is involved in responding to teicoplanin. We also would extend this work to cover strains overexpressing PAP to identify their involvement in membrane lipid homeostasis.

## Final thoughts

Although more work needs to be carried out to understand the mechanism of VanJ, this study has stressed the importance of VanJ in the presence of teicoplanin within *S. coelicolor*. Furthermore, the action of VanJ is likely to involve multiple systems within *S. coelicolor* which work synergistically to maximise resistance towards teicoplanin-like glycopeptide antibiotics. We also demonstrate that the second group of membrane enzymes, the PAP, are implicated in the barrier function of the cell. Increasing their expression levels harms the fitness of *S. coelicolor* in the presence of different groups of cell wall targeting antibiotics. We became particularly interested in one of these genes, *SCO6355* because its overexpression produced intermediate levels of resistance towards vancomycin, ristocetin and balhimycin, as well as a hypersensitive phenotype against duramycin. *S. coelicolor* usually exhibits high-level resistance to vancomycin due to its native *van* resistance system indicating that overexpressing *SCO6355* interferes with this mechanism, resulting in increased potency of vancomycin and similar glycopeptides. Although we cannot conclude the exact mechanism by which this occurs, it is clear the expression of *SCO6355* improves the *in vivo* activity of these drugs. Understanding how this occurs could be instrumental in developing novel vancomycin derivatives that are more potent in vancomycin-resistant pathogenic bacteria.

# Bibliography


1. Aminov, R. I. A brief history of the antibiotic era: Lessons learned and challenges for the future. *Front. Microbiol.* **1**, (2010).
2. Fleming, A. On the Antibacterial Action of Cultures of A Penicillium, With Special Reference to Their Use in The Isolation of B. influenzae.pdf. *Br. J. Exp. Pathol.* **10**, 226–236 (1929).
3. Chain, E. *et al.* Penicillin as a Chemotherapuetic Agent. *Lancet* 226–228 (1940).
4. American Chemical Society & Royal Society of Chemistry. The Discovery and Development of Penicillin 1928-1945. *The Alexander Fleming Laboratory Museum, UK.* 1–10 (1999).
5. Walsh, C. T. & Wencewicz, T. A. Prospects for new antibiotics: a molecule-centered perspective. *J. Antibiot. (Tokyo).* **67**, 7–22 (2013).
6. Abraham, E. P. & Chain, E. An Enzyme From Bacteria Able to Destroy Penicillin. *Nature* 837 (1940).
7. Tomasz, A. The Mechanisms of the Irreversible Antimicrobial Effects of Penicillins: How the Beta-Lactam Antibiotics Kill and Lyse Bacteria. *Annu. Rev. Microbiol.* (1979).
8. Van Boeckel, T. P. *et al.* Global antibiotic consumption 2000 to 2010: An analysis of national pharmaceutical sales data. *Lancet Infect. Dis.* **14**, 742–750 (2014).
9. Department of Health. *Antimicrobial Resistance Empirical and Statistical Evidence-Base*. (2016).
10. Hirsch, R., Ternes, T., Haberer, K. & Kratz, K. L. Occurrence of antibiotics in the aquatic environment. *Sci. Total Environ.* **225**, 109–118 (1999).
11. Christian, T. *et al.* Determination of antibiotic residues in manure, soil, and surface waters. *Acta Hydrochim. Hydrobiol.* **31**, 36–44 (2003).
12. Baquero, F., Martínez, J. L. & Cantón, R. Antibiotics and antibiotic resistance in water environments. *Curr. Opin. Biotechnol.* **19**, 260–265 (2008).
13. Martinez, J. L. Environmental pollution by antibiotics and by antibiotic resistance determinants. *Environ. Pollut.* **157**, 2893–2902 (2009).
14. O'Neill, J. Tackling drug-resistant infections globally: final report and recommendations. *Rev. Antimicrob. Resist.* 84 (2016). doi:10.1016/j.jpha.2015.11.005
15. CDC. Antibiotic resistance threats. *U.S. Dep. Heal. Hum. Serv.* 22–50 (2013). doi:CS239559-B
16. Kohanski, M. A., Dwyer, D. J. & Collins, J. J. How antibiotics kill bacteria: from targets to networks. *Nat. Rev. Microbiol.* **8**, 423–435 (2010).
17. Sauvage, E., Kerff, F., Terrak, M., Ayala, J. A. & Charlier, P. The penicillin-binding proteins: Structure and role in peptidoglycan biosynthesis. *FEMS Microbiol. Rev.* **32**, 234–258 (2008).
18. Bouhss, A., Trunkfield, A. E., Bugg, T. D. H. & Mengin-Lecreulx, D. The biosynthesis

of peptidoglycan lipid-linked intermediates. *FEMS Microbiology Reviews* (2008). doi:10.1111/j.1574-6976.2007.00089.x

19. Bugg, T. D. H., Braddick, D., Dowson, C. G. & Roper, D. I. Bacterial cell wall assembly: Still an attractive antibacterial target. *Trends in Biotechnology* (2011). doi:10.1016/j.tibtech.2010.12.006

20. Teo, A. & Roper, D. Core Steps of Membrane-Bound Peptidoglycan Biosynthesis: Recent Advances, Insight and Opportunities. *Antibiotics* **4**, 495–520 (2015).

21. Yun, M.-K. *et al.* Catalysis and Sulfa Drug Resistance in Dihydropteroate Synthase. *Sci. Reports* **335**, (2012).

22. Zierhut, G., Piepersberg, W. & Bock, A. Comparative Analysis of the Effect of Aminoglycosides on Bacterial Protein Synthesis in vitro. *Eur. J. Biochem.* **583**, 577–583 (1979).

23. Chopra, I. & Roberts, M. Tetracycline Antibiotics : Mode of Action , Applications , Molecular Biology , and Epidemiology of Bacterial Resistance. *Microbiol. Mol. Biol. Rev.* **65**, 232–260 (2001).

24. Wisseman, C. L., Smadel, J. E., Hahn, F. E. & Hopps, H. E. Action of Chloramphenicol on Assimilation of Ammonia and on Synthesis of Proteins and Nucleic acids in *Escherichia coli*. *Mode Action Chloramphenicol* (1953).

25. Wolfe, A. D. & Hahn, F. E. Erythromycin: Mode of Action. *Science (80-. ).* 1445–1447 (1964).

26. Anderson, J. S., Matsuhashi, M., Haskin, M. A. & Strominger, J. L. Lipid-Phosphoacetylemuramyl-Pentapeptide and Lipid-Phosphodisaccharide-Pentapeptide: Presume Membrane Transport Intermediates in Cell Wall Synthesis. *Biochemistry* **53**, (1965).

27. Wehrli, W. Ansamycins Chemistry , Biosynthesis and Biological Activity. *Med. Chem. (Los. Angeles).* (1977).

28. Fisher, L. M. *et al.* Ciprofloxacin and the Fluoroquinolones. *Am. J. Med.* **87**, 2–8 (1989).

29. Johnston, N. J., Mukhtar, T. A. & Wright, G. D. Streptogramin Antibiotics : Mode of Action and Resistance. *Curr. Drug Targets* **3**, 335–344 (2002).

30. Bozdogan, B. & Appelbaum, P. C. Oxazolidinones: activity, mode of action, and mechanism of resistance. *Int. J. Antimicrob. Agents* **23**, 113–119 (2004).

31. Straus, S. K. & Hancock, R. E. W. Mode of action of the new antibiotic for Gram-positive pathogens daptomycin: Comparison with cationic antimicrobial peptides and lipopeptides. **1758**, 1215–1223 (2006).

32. Silhavy, T. J., Kahne, D. & Walker, S. The Bacterial Cell Envelope. 1–16 (2010).

33. Meroueh, S. O. *et al.* Three-dimensional structure of the bacterial cell wall peptidoglycan. *Proc Natl Acad Sci U S A* **103**, 4404–4409 (2006).

34. Barreteau, H. *et al.* Cytoplasmic steps of peptidoglycan biosynthesis. **32**, 168–207 (2008).

35. Teng, K. & Liang, P. Structures , mechanisms and inhibitors of undecaprenyl

diphosphate synthase : A cis -prenyltransferase for bacterial peptidoglycan biosynthesis. *Bioorg. Chem.* **43**, 51–57 (2012).

36. Rohrer, S., Ehlert, K., Tschierske, M., Labisdhinkski, H. & Berger-Bachi, B. The essential *Staphylococcus aureus* gene *fmhB* is involved in the first step of peptidoglycan pentaglycine interpeptide formation. *Microbiology* **96**, 9351–9356 (1999).

37. Hong, H. J., Hutchings, M. I., Hill, L. M. & Buttner, M. J. The role of the novel fem protein VanK in vancomycin resistance in *Streptomyces coelicolor*. *J. Biol. Chem.* **280**, 13055–13061 (2005).

38. Ruiz, N. Bioinformatics identification of MurJ (MviN) as the peptidoglycan lipid II flippase in *Escherichia coli*. *Proc. Natl. Acad. Sci. U. S. A.* **105**, 15553–7 (2008).

39. Meeske, A. J. *et al.* MurJ and a novel lipid II flippase are required for cell wall biogenesis in *Bacillus subtilis*. *Proc. Natl. Acad. Sci. U. S. A.* **112**, 6437–42 (2015).

40. Mohammadi, T. *et al.* Identification of FtsW as a transporter of lipid-linked cell wall precursors across the membrane. *EMBO J.* **30**, 1425–1432 (2011).

41. Mohammadi, T. *et al.* Specificity of the transport of lipid II by FtsW in *Escherichia coli*. *J. Biol. Chem.* **289**, 14707–14718 (2014).

42. El Ghachi, M., Bouhss, A., Blanot, D. & Mengin-Lecreulx, D. The *bacA* gene of *Escherichia coli* encodes an undecaprenyl pyrophosphate phosphatase activity. *J. Biol. Chem.* **279**, 30106–30113 (2004).

43. Cain, B. D., Norton, P. J., Eubanks, W., Nick, H. S. & Allen, C. M. Amplification of the *bacA* gene confers bacitracin resistance to *Escherichia coli*. *J. Bacteriol.* **175**, 3784–3789 (1993).

44. Ghachi, M. El, Derbise, A., Bouhss, A. & Mengin-lecreulx, D. Identification of Multiple Genes Encoding Membrane Proteins with Undecaprenyl Pyrophosphate Phosphatase (UppP) Activity in *Escherichia coli*. *J. Biol. Chem.* **280**, 18689–18695 (2005).

45. Bernard, R., El Ghachi, M., Mengin-Lecreulx, D., Chippaux, M. & Denizot, F. BcrC from *Bacillus subtilis* acts as an undecaprenyl pyrophosphate phosphatase in bacitracin resistance. *J. Biol. Chem.* **280**, 28852–28857 (2005).

46. Manat, G. *et al.* Deciphering the Metabolism of Undecaprenyl-Phosphate: The Bacterial Cell-Wall Unit Carrier at the Membrane Frontier. *Microb. Drug Resist.* **20**, 199–214 (2014).

47. Simons, K. & Sampaio, J. L. Membrane Organization and Lipid Rafts. *Endocrinologist* **3**, 443 (1993).

48. Bramkamp, M. & Lopez, D. Exploring the Existence of Lipid Rafts in Bacteria. *Microbiol. Mol. Biol. Rev.* **79**, 81–100 (2015).

49. Singer, S. J. & Nicolson, G. L. The Fluid Mosaic Model of the Structure of Cell Membranes. *Science (80-. ).* **175**, (1972).

50. Holthuis, J. C. M. & Levine, T. P. Lipid traffic: floppy drives and a superhighway. *Nat. Rev. Mol. Cell Biol.* **6**, 209–20 (2005).

51. Ogushi, F., Ishitsuka, R., Kobayashi, T. & Sugita, Y. Rapid flip-flop motions of diacylglycerol and ceramide in phospholipid bilayers. *Chem. Phys. Lett.* **522**, 96–102

(2012).

52. Sharom, F. J. Flipping and flopping-lipids on the move. *IUBMB Life* **63**, 736–746 (2011).

53. Staubitz, P. & Peschel, A. MprF-mediated lysinylation of phospholipids in *Bacillus subtilis*-protection against bacteriocins in terrestrial habitats. *Microbiology* **148**, 3327–3328 (2002).

54. Slavetinsky, C. J., Peschel, A. & Ernst, C. M. Alanyl-phosphatidylglycerol and lysyl-phosphatidylglycerol are translocated by the same MprF flippases and have similar capacities to protect against the antibiotic daptomycin in *Staphylococcus aureus*. *Antimicrob. Agents Chemother.* **56**, 3492–3497 (2012).

55. Sutterlin, H. A., Zhang, S. & Silhavya, T. J. Accumulation of phosphatidic acid increases vancomycin resistance in *Escherichia coli*. *J. Bacteriol.* **196**, 3214–3220 (2014).

56. Rowlett, V. W. *et al.* Impact of Membrane Phospholipid Alteration in *Escherichia coli* on Cellular Function and Bacterial Stress Adaptation. *J. Bacteriol.* **199**, 1–22 (2017).

57. Lopez, C. S., Heras, H., Ruzal, S. M., Rivas, C. S. & Rivas, E. A. Variations of the Envelope Composition of *Bacillus subtilis* During Growth in Hyperosmotic Medium. **36**, 55–61 (1998).

58. Zhang, Y.-M. & Rock, C. O. Membrane lipid homeostasis in bacteria. *Nat. Rev. Microbiol.* **6**, 222–233 (2008).

59. Sinensky, M. Homeoviscous Adaptation-A Homeostatic Process that Regulates the Viscosity of Membrane Lipids in *Escherichia coli*. **71**, 522–525 (1974).

60. Sandoval-Calderón, M. *et al.* Plasticity of *Streptomyces coelicolor* Membrane Composition Under Different Growth Conditions and During Development. *Front. Microbiol.* **6**, 1–13 (2015).

61. Sohlenkamp, C. & Geiger, O. Bacterial membrane lipids: Diversity in structures and pathways. *FEMS Microbiol. Rev.* **40**, 133–159 (2015).

62. Lemieux, M. J., Huang, Y. & Wang, D. N. Glycerol-3-phosphate transporter of *Escherichia coli*: Structure, function and regulation. *Res. Microbiol.* **155**, 623–629 (2004).

63. Jeucken, A., Helms, J. B. & Brouwers, J. F. Cardiolipin synthases of *Escherichia coli* have phospholipid class specific phospholipase D activity dependent on endogenous and foreign phospholipids. *BBA - Mol. Cell Biol. Lipids* **1863**, 1345–1353 (2018).

64. Nishibori, A., Kusaka, J., Hara, H., Umeda, M. & Matsumoto, K. Phosphatidylethanolamine domains and localization of phospholipid synthases in *Bacillus subtilis* membranes. *J. Bacteriol.* **187**, 2163–2174 (2005).

65. Ernst, C. M. & Peschel, A. Broad-spectrum antimicrobial peptide resistance by MprF-mediated aminoacylation and flipping of phospholipids. *Mol. Microbiol.* **80**, 290–299 (2011).

66. Dare, K., Shepherd, J., Roy, H., Seveau, S. & Ibba, M. LysPGS formation in *Listeria monocytogenes* has broad roles in maintaining membrane integrity beyond antimicrobial peptide resistance. *Virulence* **5**, 534–546 (2014).

67. Sandoval-calderon, M., Guan, Z. & Sohlenkamp, C. Knowns and unknowns of membrane lipid synthesis in streptomycetes. *Biochimie* 1–9 (2017). doi:10.1016/j.biochi.2017.05.008

68. Touzé, T., Blanot, D. & Mengin-Lecreulx, D. Substrate specificity and membrane topology of *Escherichia coli* PgpB, an undecaprenyl pyrophosphate phosphatase. *J. Biol. Chem.* **283**, 16573–16583 (2008).

69. Bernard, R., Joseph, P., Guiseppi, A., Chippaux, M. & Denizot, F. YtsCD and YwoA, two independent systems that confer bacitracin resistance to *Bacillus subtilis*. *FEMS Microbiol. Lett.* **228**, 93–97 (2003).

70. Pralla, K., Muth, G. & Hartner, T. Hopanoids are formed during transition from substrate to aerial hyphae in *Streptomyces coelicolor A3(2)*. *FEMS Microbiol. Lett.* **189**, 93–95 (2000).

71. López, D. & Kolter, R. Functional microdomains in bacterial membranes. *Genes Dev.* **24**, 1893–1902 (2010).

72. Butler, M. S., Hansford, K. A., Blaskovich, M. A. T., Halai, R. & Cooper, M. A. Glycopeptide antibiotics : Back to the future. *J. Antibiot. (Tokyo).* **67**, 631–644 (2014).

73. Levine, D. P. Vancomycin : A History. *Clin. Infect. Dis.* **42**, 5–12 (2006).

74. Marinelli, F. & Genilloud, O. Antimicrobials: New and old molecules in the fight against multi-resistant bacteria. **9783642399**, (2014).

75. Parenti, F., Beretta, G., Berti, M. & Arioli, V. Teichomycins, New Antibiotics From *Actinoplanes Teichomyceticus J. Antibiot.* 276–283 (1978).

76. Pitkin, D. H., Mico, B. A., Sitrin, R. D. & Nisbet, L. J. Charge and Lipophilicity Govern the Pharmacokinetics of Glycopeptide Antibiotics. *Antimicrob. Agents Chemothapy* **29**, 440–444 (1986).

77. Van Der Auwera, P., Aoun, M. & Meunier, F. Randomized study of vancomycin versus teicoplanin for the treatment of gram-positive bacterial infections in immunocompromised hosts. *Antimicrob. Agents Chemother.* **35**, 451–457 (1991).

78. Nicolaou, K. C., Boddy, C. N. C., Bräse, S. & Winssinger, N. Chemistry , Biology , and Medicine of the Glycopeptide Antibiotics. *Angew. Chemie - Int. Ed.* (1999).

79. Rodriguez, M. J. *et al.* Novel Glycopeptide Antibiotics: N-Alkylated Derivatives Active Against Vancomycin-Resistant Enterococci. *J. Antibiot. (Tokyo).* **51**, (1998).

80. European Medicines Agency. *Orbactiv report*. (2015).

81. Leadbetter, M. R. *et al.* Hydrophobic Vancomycin Derivatives with Improved ADME Properties : Discovery of Telavancin ( TD-6424 ) BAZZINI. *J. Antibiot. (Tokyo).* **57**, (2004).

82. European Medicines Agency. *Vibativ*. (2011).

83. Candiani, G., Abbondi, M., Borgonovi, M., Romano, G. & Parenti, F. In-vitro and in-vivo antibacterial activity of BI 397, a new semi-synthetic glycopeptide antibiotic. *J. Antimicrob. Chemother.* **44**, 179–192 (1999).

84. European Medicine Agency. *Xydalba*. (2015).

85. Nieto, M. & Perkins, H. R. Modifications of the Acyl-D-alanyl-D-alanine Terminus Affecting Complex-Formation with Vancomycin. *Biochem. J.* **143**, 789–803 (1971).

86. Xie, J., Pierce, J. G., James, R. C., Okano, A. & Boger, D. L. A Redesigned Vancomycin Engineered for Dual D-Ala-D-Ala and D-Ala-D-lac Binding Exhibits Potent Antimicrobial Activity Against Vancomycin-Resistant Bacteria. *J. Am. Chem. Soc.* **133**, 13946–13949 (2011).

87. Durocher, Y. & Butler, M. Expression systems for therapeutic glycoprotein production. *Curr. Opin. Biotechnol.* **20**, 700–707 (2009).

88. Kim, S. J. *et al.* Oritavancin Exhibits Dual Mode of Action to Inhibit Cell-Wall Biosynthesis in *Staphylococcus aureus*. *J. Mol. Biol.* **377**, 281–293 (2008).

89. Laclercq, R., Derlot, E., Duval, J. & Courvalin, P. Plasmid-mediated Resistance to Vancomycin and Teicoplanin in *Enterococcus Faecium*. *N. Engl. J. Med.* **319**, 157–161 (1988).

90. Uttley, A. H. C., George, R. C. & Naidoo, J. High-level vancomycin-resistant enterococci causing hospital infections. 173–181 (1989).

91. Ahmed, M. O. & Baptiste, K. E. Vancomycin-Resistant Enterococci : and Perspectives of Human and Animal Health. *Microb. Drug Resist.* **00**, 1–17 (2017).

92. Kofoid, E. C. & Parkinson, J. S. Transmitter and receiver modules in bacterial signaling proteins. *Proc. Natl. Acad. Sci.* **85**, 4981–4985 (1988).

93. Parkinson, J. S. & Kofoid, E. C. Communication modules in bacterial signaling proteins. *Annu. Rev. Genet.* 71–112 (1992).

94. Arthur, M. & Courvalin, P. Genetics and Mechanisms of Glycopeptide Resistance in Enterococci. *Antimicrob. Agents Chemother.* **37**, 1563–1571 (1993).

95. Park, I. J., Lee, W. G., Shin, J. H., Lee, K. W. & Woo, G. J. VanB phenotype-vanA genotype *Enterococcus faecium* with heterogeneous expression of teicoplanin resistance. *J. Clin. Microbiol.* **46**, 3091–3093 (2008).

96. Mckay, G. A. *et al.* Time – kill kinetics of oritavancin and comparator agents against *Staphylococcus aureus* , *Enterococcus faecalis* and *Enterococcus faecium*. *J. Antimicrob. Chemother.* 1191–1199 (2009). doi:10.1093/jac/dkp126

97. Popovic, N. *et al.* Oral teicoplanin for successful treatment of severe refractory *Clostridium difficile* infection. *J. Infect. Dev. Ctries.* **9**, 1062–1067 (2015).

98. Cooper, R. D. G. *et al.* Reductive Alkylation of Glycopeptide Antibiotics: Synthesis and Antibacterial Activity. *J. Antibiot. (Tokyo).* **49**, 575–581 (1996).

99. Neu, H. C. & Labthavikul, P. In Vitro Activity of Teichomycin Compared with Those of Other Antibiotics. *Antimicrob. Agents Chemother.* **24**, 425–428 (1983).

100. Kwun, M. J. & Hong, H. J. The activity of glycopeptide antibiotics against resistant bacteria correlates with their ability to induce the resistance system. *Antimicrob. Agents Chemother.* **58**, 6306–6310 (2014).

101. Arthur, M., Depardieu, F. & Courvalin, P. Regulated interactions between partner and non-partner sensors and response regulators that control glycopeptide resistance gene expression in enterococci. *Microbiology* **145**, 1849–1858 (1999).

102. Cegelski, L. *et al.* Rotational-Echo Double Resonance Characterization of the Effects of Vancomycin on Cell Wall Synthesis in *Staphylococcus aureus*. *Biochemistry* **41**, 13053–13058 (2002).

103. Chmara, H., I P A, S. R., Mignini, F. & Borowski, E. Bacteriolytic effect of teicoplanin. *J. Gen. Microbiol.* **137**, (1991).

104. Higgins, D. L. *et al.* Telavancin, a multifunctional lipoglycopeptide, disrupts both cell wall synthesis and cell membrane integrity in methicillin-resistant *Staphylococcus aureus*. *Antimicrob. Agents Chemother.* **49**, 1127–1134 (2005).

105. Patti, G. J. *et al.* Vancomycin and Oritavancin Have Different Modes of Action in *Enterococcus faecium*. *J. Mol. Biol.* **392**, 1178–1191 (2009).

106. Kim, S. J., Tanaka, K. S. E., Dietrich, E., Rafai Far, A. & Schaefer, J. Locations of the hydrophobic side chains of lipoglycopeptides bound to the peptidoglycan of *Staphylococcus aureus*. *Biochemistry* **52**, 3405–3414 (2013).

107. Ulijasz, A. T., Grenader, A. & Weisblum, B. A vancomycin-inducible LacZ reporter system in *Bacillus subtilis*: Induction by antibiotics that inhibit cell wall synthesis and by lysozyme. *J. Bacteriol.* **178**, 6305–6309 (1996).

108. Yagüe, P. *et al.* Transcriptomic Analysis of *Streptomyces coelicolor* Differentiation in Solid Sporulating Cultures: First Compartmentalized and Second Multinucleated Mycelia Have Different and Distinctive Transcriptomes. *PLoS One* **8**, (2013).

109. Hutchings, M. I., Hong, H. & Buttner, M. J. The vancomycin resistance VanRS two-component signal transduction system of *Streptomyces coelicolor*. *Mol. Microbiol.* **59**, 923–935 (2006).

110. Hong, H.-J. *et al.* Characterization of an inducible vancomycin resistance system in *Streptomyces coelicolor* reveals a novel gene (vanK) required for drug resistance. *Mol. Microbiol.* **52**, 1107–1121 (2004).

111. Phillips-jones, M. K. *et al.* Hydrodynamics of the VanA-type VanS histidine kinase: an extended solution conformation and first evidence for interactions with vancomycin. *Sci. Rep.* 1–12 (2017). doi:10.1038/srep46180

112. Koteva, K. *et al.* A vancomycin photoprobe identifies the histidine kinase VanSsc as a vancomycin receptor. *Nat. Chem. Biol.* **6**, 327–329 (2010).

113. Kwun, M. J., Novotna, G., Hesketh, A. R., Hill, L. & Hong, H. J. In vivo studies suggest that induction of VanS-dependent vancomycin resistance requires binding of the drug to D-Ala-D-Ala termini in the peptidoglycan cell wall. *Antimicrob. Agents Chemother.* **57**, 4470–4480 (2013).

114. Dong, S. D. *et al.* The structural basis for induction of VanB resistance. *J. Am. Chem. Soc.* **124**, 9064–9065 (2002).

115. Novotna, G., Hill, C., Vincent, K., Liu, C. & Hong, H. A Novel Membrane Protein , VanJ, Conferring Resistance to Teicoplanin. *Antimicrob. Agents Chemother.* **56**, 1784–1796 (2012).

116. Chalmers, C. *et al. Streptococcus pyogenes* nuclease A (SpnA) mediated virulence does not exclusively depend on nuclease activity. *J. Microbiol. Immunol. Infect.* **53**, 42–48 (2017).

117. Mian, I. S., Worthey, E. A. & Salavati, R. Taking U out, with two nucleases? *BMC Bioinformatics* **10**, 1–10 (2006).

118. Aphasizheva, I. & Aphasizhev, R. U-Insertion/Deletion mRNA-Editing Holoenzyme: Definition in Sight. *Trends Parasitol.* **32**, 144–156 (2016).

119. Roberts, A., Pimentel, H., Trapnell, C. & Pachter, L. Identification of novel transcripts in annotated genomes using RNA-seq. *Bioinformatics* **27**, 2325–2329 (2011).

120. Trapnell, C. *et al.* Transcript assembly and quantification by RNA-Seq reveals unannotated transcripts and isoform switching during cell differentiation. *Nat. Biotechnol.* **28**, 511–515 (2010).

121. Pertea, M., Kim, D., Pertea, G. M., Leek, J. T. & Salzberg, S. L. Transcript-level expression analysis of RNA-seq experiments with HISAT, StringTie and Ballgown. *Nat Protoc.* **11**, 1650–1667 (2016).

122. Chen, Y., Lun, A. T. L. & Smyth, G. K. From reads to genes to pathways: differential expression analysis of RNA-Seq experiments using Rsubread and the edgeR quasi-likelihood pipeline. *F1000Research* **5**, 1438 (2016).

123. Law, C. W., Alhamdoosh, M., Su, S., Smyth, G. K. & Ritchie, M. E. RNA-seq analysis is easy as 1-2-3 with limma, Glimma and edgeR. *F1000Research* **5**, 1408 (2016).

124. Mccarthy, D. J. & Smyth, G. K. Testing significance relative to a fold-change threshold is a TREAT. *Bioinformatics* **25**, 765–771 (2009).

125. Young, M. D., Wakefield, M. J., Smyth, G. K. & Oshlack, A. Gene ontology analysis for RNA-seq: accounting for selection bias. *Genome Biol.* **11**, R14 (2010).

126. Young, M. D., Wakefield, M. J. & Smyth, G. K. goseq : Gene Ontology testing for RNA-seq datasets Reading data. *Gene* 1–21 (2010).

127. Mandel, M. & Higa, A. Calcium-dependent bacteriophage DNA infection. *J. Mol. Biol.* **53**, 159–162 (1970).

128. Paget, M. S. B., Chamberlin, L., Atrih, A., Foster, S. J. & Buttner, M. J. Evidence that the Extracytoplasmic Function Sigma Factor sigE Is Required for Normal Cell Wall Structure in *Streptomyces coelicolor A3 (2)*. *J. Bacteriol.* **181**, 204–211 (1999).

129. H.C.Birnboim & J.Doly. A Rapid Alkaline Extraction Procedure for Screening Recombinant plasmid DNA. **7**, 1513–1524 (1979).

130. Cobb, R. E., Wang, Y. & Zhao, H. High-Efficiency Multiplex Genome Editing of Streptomyces Species Using an Engineered CRISPR/Cas System. *ACS Synth. Biol.* 141208152636004 (2014). doi:10.1021/sb500351f

131. Huang, H., Zheng, G., Jiang, W., Hu, H. & Lu, Y. One-step high-efficiency CRISPR/Cas9-mediated genome editing in Streptomyces. *ABBS* (2015). doi:10.1093/abbs/gmv007

132. EUCAST. Determination of minimum inhibitory concentrations (MICs) of antibacterial agents by agar dilution. European Committee for Antimicrobial Susceptibility Testing **6**, (2000).

133. Schena, M., Shalon, D., Davis, R. W. & Brown, P. O. Quantitative monitoring of gene expression patterns with a complementary DNA microarray. *Science (80-. ).* **270**, 467–

470 (1995).

134. Velculescu, V. E. *et al.* Characterization of the yeast transcriptome. *Cell* **88**, 243–251 (1997).

135. Hurd, P. J. & Nelson, C. J. Advantages of next-generation sequencing versus the microarray in epigenetic research. *Briefings Funct. Genomics Proteomics* **8**, 174–183 (2009).

136. Sims, D., Sudbery, I., Ilott, N. E., Heger, A. & Ponting, C. P. Sequencing depth and coverage: Key considerations in genomic analyses. *Nat. Rev. Genet.* **15**, 121–132 (2014).

137. Sîrbu, A., Kerr, G., Crane, M. & Ruskin, H. J. RNA-Seq vs Dual- and Single-Channel Microarray Data: Sensitivity Analysis for Differential Expression and Clustering. *PLoS One* **7**, (2012).

138. Thermes, C. Ten years of next-generation sequencing technology. *Trends Genet.* **30**, 418–426 (2014).

139. Dillies, M. A. *et al.* A comprehensive evaluation of normalization methods for Illumina high-throughput RNA sequencing data analysis. *Brief. Bioinform.* **14**, 671–683 (2013).

140. Ritchie, M. E. *et al.* limma powers differential expression analyses for RNA-sequencing and microarray studies. *Nucleic Acids Res.* **43**, e47 (2015).

141. Smyth, G. K. Linear models and empirical bayes methods for assessing differential expression in microarray experiments. *Stat. Appl. Genet. Mol. Biol.* **3**, (2004).

142. D'Haeseleer, P. How does gene expression clustering work? *Nat. Biotechnol.* **23**, 1499–1501 (2005).

143. Nelson, C. & Gleave, M. Cluster-based network model for time-course gene. 507–525 (2007). doi:10.1093/biostatistics/kxl026

144. Eisen, M. B., Spellman, P. T., Brown, P. O. & Botstein, D. Cluster analysis and display of genome-wide expression patterns. *Proc. Natl. Acad. Sci.* **95**, 14868–14868 (1998).

145. Conesa, A. *et al.* A survey of best practices for RNA-seq data analysis. *Genome Biol.* **17**, 13 (2016).

146. Paget, M. S. B., Molle, V., Cohen, G., Aharonowitz, Y. & Buttner, M. J. Defining the disulphide stress response in *Streptomyces coelicolor A3(2)*: Identification of the sigR regulon. *Mol. Microbiol.* **42**, 1007–1020 (2001).

147. Lee, E. J. *et al.* A master regulator σB governs osmotic and oxidative response as well as differentiation via a network of sigma factors in *Streptomyces coelicolor*. *Mol. Microbiol.* **57**, 1252–1264 (2005).

148. Hesketh, A., Chen, W. J., Ryding, J., Chang, S. & Bibb, M. The global role of ppGpp synthesis in morphological differentiation and antibiotic production in *Streptomyces coelicolor A3(2)*. *Genome Biol.* **8**, (2007).

149. Hesketh, A. *et al.* Genome-wide dynamics of a bacterial response to antibiotics that target the cell envelope. *BMC Genomics* **12**, 226 (2011).

150. Tran, N. T. *et al.* Defining the regulon of genes controlled by σE, a key regulator of the

cell envelope stress response in *Streptomyces coelicolor*. *Mol. Microbiol.* **0**, 1–21 (2019).

151. Hong, H. J., Paget, M. S. B. & Buttner, M. J. A signal transduction system in *Streptomyces coelicolor* that activates the expression of a putative cell wall glycan operon in response to vancomycin and other cell wall-specific antibiotics. *Mol. Microbiol.* (2002). doi:10.1046/j.1365-2958.2002.02960.x

152. Paget, M. S. B., Chamberlin, L., Atrih, A., Foster, S. J. & Buttner, M. J. Evidence that the extracytoplasmic function sigma factor $\sigma^E$ is required for normal cell wall structure in *Streptomyces coelicolor A3(2)*. *J. Bacteriol.* **181**, 204–211 (1999).

153. Song, Y., Lunde, C. S., Benton, B. M. & Wilkinson, B. J. Further Insights into the Mode of Action of the Lipoglycopeptide Telavancin through Global Gene Expression Studies. *Antimicrob. Agents Chemothapy* **56**, 3157–3164 (2012).

154. Utaida, S. *et al.* Genome-wide transcriptional profiling of the response of *Staphylococcus aureus* to cell-wall- active antibiotics reveals a cell-wall-stress stimulon. *Microbiology* 2719–2732 (2003). doi:10.1099/mic.0.26426-0

155. Bentley, S. *et al.* Complete genome sequence of the model actinomycete *Streptomyces coelicolor A3(2)*. *Nature* **417**, 141–147 (2002).

156. Lee, E., Cho, Y., Kim, H., Ahn, B. & Roe, J. Regulation of sigB by an Anti- and an Anti-Anti-Sigma Factor in *Streptomyces coelicolor* in Response to Osmotic Stress. *J. Bacteriol.* **186**, (2004).

157. Neu, J. M., MacMillan, S. V., Nodwell, J. R. & Wright, G. D. StoPK-1, a serine/threonine protein kinase from the glycopeptide antibiotic producer *Streptomyces toyocaensis* NRRL 15009, affects oxidative stress response. *Mol. Microbiol.* **44**, 417–430 (2002).

158. Shu, D. *et al.* afsQ1-Q2-sigQ is a pleiotropic but conditionally required signal transduction system for both secondary metabolism and morphological development in *Streptomyces coelicolor*. *Appl. Microbiol. Biotechnol.* **81**, 1149–1160 (2009).

159. Neu, J. M. & Wright, G. D. Inhibition of sporulation, glycopeptide antibiotic production and resistance in *Streptomyces toyocaensis* NRRL 15009 by protein kinase inhibitors. *FEMS Microbiol. Lett.* **199**, 15–20 (2001).

160. Kennelly, P. J. & Potts, M. Fancy Meeting You Here! a Fresh Look at " Prokaryotic " Protein Phosphorylation. *J. Bacteriol.* **178**, 4759–4764 (1996).

161. Oh, E. A. *et al.* Characterization of the autophosphorylating kinase, PkaF, in *Streptomyces coelicolor A3(2)* M130. *Arch. Microbiol.* **193**, 845–856 (2011).

162. Abranches, J. *et al.* The Cell Wall-Targeting Antibiotic Stimulon of *Enterococcus faecalis*. *PLoS One* **8**, (2013).

163. Bylund, G. O., Wipemo, L. C., Lundberg, L. A. C. & Wikström, P. M. RimM and RbfA are essential for efficient processing of 16S rRNA in *Escherichia coli*. *J. Bacteriol.* **180**, 73–82 (1998).

164. Hamma, T., Ferre, A. R., Hutchinson, F. & Families, P. S. Pseudouridine Synthases Review. *Chem. Biol.* 1125–1135 (2006). doi:10.1016/j.chembiol.2006.09.009

165. Kim, M. S. *et al.* Conservation of thiol-oxidative stress responses regulated by SigR orthologues in actinomycetes. *Mol. Microbiol.* **85**, 326–344 (2012).

166. Fröhler, J., Rechenmacher, A., Thomale, J., Nass, G. & Böck, A. Genetic analysis of mutations causing borrelidin resistance by overproduction of threonyl transfer ribonucleic acid synthetase. *J. Bacteriol.* **143**, 1135–1141 (1980).

167. Fisher, A. *et al.* Daptomycin resistance mechanisms in clinically derived *Staphylococcus aureus* strains assessed by a combined transcriptomics and proteomics approach. *J. Antimicrob. Chemother.* 1696–1711 (2011). doi:10.1093/jac/dkr195

168. Palmer, K. L. *et al.* Genetic Basis for Daptomycin Resistance in Enterococci Genetic Basis for Daptomycin Resistance in Enterococci. *Antimicrob. Agents Chemother.* **55**, 3345–3356 (2011).

169. Koch, H. G., Winterstein, C., Saribas, A. S., Alben, J. O. & Daldal, F. Roles of the ccoGHIS gene products in the biogenesis of the cbb3-type cytochrome c oxidase. *J. Mol. Biol.* **297**, 49–65 (2000).

170. Kol, S. *et al.* Metabolomic characterization of the salt stress response in *Streptomyces coelicolor*. *Appl. Environ. Microbiol.* **76**, 2574–2581 (2010).

171. Barrientos-moreno, L., Molina-henares, M. A., Pastor-garcía, M. & Ramos-gonzález, M. I. crossm Arginine Biosynthesis Modulates Pyoverdine Production and Release in Pseudomonas putida as Part of the Mechanism of Adaptation to Oxidative Stress. *J. Bacteriol.* **201**, 1–15 (2019).

172. Rioseras, B., López-garcía, M. T., Yagüe, P., Sánchez, J. & Manteca, Á. Mycelium differentiation and development of *Streptomyces coelicolor* in lab-scale bioreactors: Programmed cell death , differentiation , and lysis are closely linked to undecylprodigiosin and actinorhodin production. *Bioresour. Technol.* **151**, 191–198 (2014).

173. Claessen, D., de Jong, W., Dijkhuizen, L. & Wösten, H. A. B. Regulation of Streptomyces development: reach for the sky! *Trends Microbiol.* **14**, 313–319 (2006).

174. Claessen, D. *et al.* The formation of the rodlet layer of streptomycetes is the result of the interplay between rodlins and chaplins. *Mol. Microbiol.* **53**, 433–443 (2004).

175. Keijser, B. J. F., Noens, E. E. E., Kraal, B., Koerten, H. K. & Van Wezel, G. P. The *Streptomyces coelicolor ssgB* gene is required for early stages of sporulation. *FEMS Microbiol. Lett.* **225**, 59–67 (2003).

176. Keijser, B. J. F., Wezel, G. P. Van, Canters, G. W. & Vijgenboom, E. Developmental Regulation of the *Streptomyces lividans ram* Genes: Involvement of RamR in Regulation of the *ramCSAB* Operon. *J. Bacteriol.* **184**, 4420–4429 (2002).

177. Berardo, C. Di *et al.* Function and Redundancy of the Chaplin Cell Surface Proteins in Aerial Hypha Formation, Rodlet Assembly, and Viability in *Streptomyces coelicolor*. *J. Bacteriol.* **190**, 5879–5889 (2008).

178. Daza, A., Martin, J. F., Dominguez, A. & Gil, J. A. Sporulation of Several Species of Streptomyces in Submerged Cultures after Nutritional Downshift. *J. Gen. Microbiol.* **135**, 2483–2491 (1989).

179. Chatter, K. F. A Morphological and Genetic Mapping Study of White Colony Mutants

of *Streptomyces coelicolor*. *J. Gen. Microbiol.* **72**, (1972).

180. Willemse, J., Mommaas, A. M. & Wezel, G. P. van. Constitutive expression of *ftsZ* overrides the *whi* developmental genes to initiate sporulation of *Streptomyces coelicolor*. *Antonie van Leeuwenhoek, Int. J. Gen. Mol. Microbiol.* 619–632 (2012). doi:10.1007/s10482-011-9678-7

181. Ainsa, J. A., Parry, H. D. & Chater, K. F. A response regulator-like protein that functions at an intermediate stage of sporulation in *Streptomyces coelicolor A3 (2)*. *Mol. Microbiol.* **34**, 607–619 (1999).

182. Xu, Q. *et al.* Structural and Functional Characterizations of SsgB, a Conserved Activator of Developmental Cell Division in Morphologically Complex Actinomycetes. *J. Biol. Chem.* **284**, 25268–25279 (2009).

183. Kelemen, G. H. *et al.* Developmental Regulation of Transcription of *whiE*, a Locus Specifying the Polyketide Spore Pigment in *Streptomyces coelicolor A3(2)*. *J. Bacteriol.* **180**, 2515–2521 (1998).

184. Tian, Y., Fowler, K., Findlay, K., Tan, H. & Chater, K. F. An Unusual Response Regulator Influences Sporulation at Early and Late Stages in *Streptomyces coelicolor*. *J. Bacteriol.* **189**, 2873–2885 (2007).

185. Molle, V., Palframan, W. J. & Findlay, K. I. M. C. WhiD and WhiB , Homologous Proteins Required for Different Stages of Sporulation in *Streptomyces coelicolor A3 (2)*. *J. Bacteriol.* **182**, 1286–1295 (2000).

186. Sasaki, Y., Oguchi, H., Kobayashi, T., Kusama, S. & Sugiura, R. Nitrogen oxide cycle regulates nitric oxide levels and bacterial cell signaling. *Nat. Publ. Gr.* 1–11 (2016). doi:10.1038/srep22038

187. Yukioka, Y. *et al.* A role of nitrite reductase (NirBD) for NO homeostatic regulation in *Streptomyces coelicolor A3(2)*. *FEMS Microbiol. Lett.* 1–6 (2017). doi:10.1093/femsle/fnw241

188. Crack, J. C., Thomson, A. J. & Brun, N. E. Le. Iron − Sulfur Clusters as Biological Sensors: The Chemistry of Reactions with Molecular Oxygen and Nitric Oxide. *Acc. Chem. Res.* (2014).

189. Manteca, A., Alvarez, R., Salazar, N. & Yagu, P. Mycelium Differentiation and Antibiotic Production in Submerged Cultures of *Streptomyces coelicolor*. *Appl. Environ. Microbiol.* **74**, 3877–3886 (2008).

190. Mascher, T., Margulis, N. G., Wang, T., Ye, R. W. & Helmann, J. D. Cell wall stress responses in *Bacillus subtilis*: The regulatory network of the bacitracin stimulon. *Mol. Microbiol.* **50**, 1591–1604 (2003).

191. Utaida, S. *et al.* Genome-wide transcriptional profiling of the response of *Staphylococcus aureus* to cell-wall-active antibiotics reveals a cell-wall-stress stimulon. *Microbiology* **149**, 2719–2732 (2003).

192. Rida, S., Caillet, J. & Alix, J.-H. Amplification of a novel gene , sanA , abolishes a vancomycin-sensitive defect in *Escherichia coli*. *J. Bacteriol.* **178**, 94–102 (1996).

193. Moriyama, Y., Hiasa, M., Matsumoto, T. & Omote, H. Multidrug and toxic compound extrusion (MATE)-type proteins as anchor transporters for the excretion of metabolic

waste products and xenobiotics. *Xenobiotica* **38**, 1107–1118 (2008).

194. Romero, S. *et al.* Purification and characterization of the second Streptomyces phospholipase A2 refolded from an inclusion body. *Protein Expr. Purif.* **50**, 82–88 (2006).

195. Titball, R. W. Bacterial Phospholipases C. *Microbiol. Rev.* **57**, 347–366 (1993).

196. Titball, R. W. Bacterial phospholipases. *J. Appl. Microbiol.* 127–137 (1998).

197. Sugiyama, M. *et al.* A Novel Prokaryotic Phospholipase A2. *J. Biol. Chem.* **277**, 20051–20058 (2002).

198. Schell, S. *et al.* Probing Enzyme Promiscuity of SGNH Hydrolases. *Chembiochem* 2158–2167 (2010). doi:10.1002/cbic.201000398

199. Tommassen, J. *et al.* Characterization of two genes, *glpQ* and *ugpQ*, encoding glycerophosphoryl diester phosphodiesterases of *Escherichia coli*. *MGG Mol. Gen. Genet.* **226**, 321–327 (1991).

200. Ohshima, N. *et al. Escherichia coli* Cytosolic Glycerophosphodiester Phosphodiesterase (UgpQ) Requires $Mg^{2+}$ , $Co^{2+}$ , or $Mn^{2+}$ for Its Enzyme Activity. *J. Bacteriol.* **190**, 1219–1223 (2008).

201. Santos-Beneit, F., Rodrigues-Garcia, A., Apel, A. K. & Martin, J. F. Phosphate and carbon source regulation of two PhoP-dependent glycerophosphodiester phosphodiesterase genes of *Streptomyces coelicolor*. *Microbiology* 1800–1811 (2009). doi:10.1099/mic.0.026799-0

202. Smith, C. P. & Chater, K. F. Structure and regulation of controlling sequences for the *Streptomyces coelicolor* glycerol operon. *J. Mol. Biol.* **204**, 569–580 (1988).

203. Hindle, Z. & Smith, C. P. Substrate induction and catabolite repression of the *Streptomyces coelicolor* glycerol operon are mediated through the GylR protein. *Mol Microbiol* **12**, 737–745 (1994).

204. Hesketh, A., Deery, M. J. & Hong, H.-J. High-Resolution Mass Spectrometry Based Proteomic Analysis of the Response to Vancomycin-Induced Cell Wall Stress in *Streptomyces coelicolor A3(2)*. *J. Proteome Res.* (2015). doi:10.1021/acs.jproteome.5b00242

205. Wiegand, I., Hilpert, K. & Hancock, R. E. W. Agar and broth dilution methods to determine the minimal inhibitory concentration (MIC) of antimicrobial substances. *Nat. Protoc.* **3**, 163–175 (2008).

206. Kim, S. J., Tanaka, K. S. E., Dietrich, E., Rafai Far, A. & Schaefer, J. Locations of the hydrophobic side chains of lipoglycopeptides bound to the peptidoglycan of *Staphylococcus aureus*. *Biochemistry* **52**, 3405–3414 (2013).

207. Fang, X., Tiyanont, K., Zhang, Y., Wanner, J. & Walker, S. The mechanism of action of ramoplanin and enduracidin. *Mol. Biosyst.* 69–76 (2006). doi:10.1039/b515328j

208. Park, J. T. & Strominger, J. L. Mode of Action of Penicillin. *Science (80-. ).* **125**, 99–101 (1957).

209. Batson, S. *et al.* phosphorylated form of the antibiotic D-cycloserine. *Nat. Commun.* 1–7 (2016). doi:10.1038/s41467-017-02118-7

210. Ostash, B. & Walker, S. Moenomycin family antibiotics : chemical synthesis , biosynthesis , and biological activity. *2010* 1594–1617 (2010). doi:10.1039/c001461n

211. Hsu, S.-T. D. *et al.* The nisin-lipid II complex reveals a pyrophosphate cage that provides a blueprint for novel antibiotics. *Nat. Struct. Mol. Biol.* **11**, 963–967 (2004).

212. Boakes, S. & Wadman, S. The therapeutic potential of lantibiotics. *Innov Pharm Technol* 22–25 (2009). doi:10.1517/13543784.9.9.2103

213. Smith, D. H. & Davis, B. D. Mode of Action of Novobiocin in *Escherichia coli*. *J. Bacteriol.* **93**, 71–79 (1967).

214. Schmitt-john, T. & Engels, J. W. Promoter constructions for efficient secretion expression in *Streptomyces lividans*. *Appl. Microbiol. Biotechnol.* 493–498 (1992).

215. Shalem, O., Sanjana, N. E. & Zhang, F. High-throughput functional genomics using CRISPR-Cas9. *Nat. Rev. Genet.* **16**, 299–311 (2015).

216. Doerrler, W. T., Sikdar, R., Kumar, S. & Boughner, L. A. New functions for the ancient DedA membrane protein family. *J. Bacteriol.* **195**, 3–11 (2013).

217. Kumar, S. & Doerrler, W. T. Members of the conserved DedA family are likely membrane transporters and are required for drug resistance in *Escherichia coli*. *Antimicrob. Agents Chemother.* **58**, 923–930 (2014).

218. Kumar, S. & Doerrler, W. T. *Escherichia coli* YqjA, a member of the conserved DedA/Tvp38 membrane protein family, is a putative osmosensing transporter required for growth at alkaline pH. *J. Bacteriol.* **197**, 2292–2300 (2015).

219. Mouslim, C., Cano, D. A. & Casadesus, J. The *sfiX*, *rfe* and *metN* genes of *Salmonella typhimurium* and their involvement in the Hisc pleiotropic response. *Mol. Genet.* 200–211 (1997). doi:10.1016/j.jcjd.2013.02.046

220. Hvorup, R. N. *et al.* The multidrug/oligosaccharidyl-lipid/polysaccharide (MOP) exporter superfamily. *Eur. J. Biochem.* **270**, 799–813 (2003).

221. Jeong, Y. *et al.* The dynamic transcriptional and translational landscape of the model antibiotic producer *Streptomyces coelicolor A3(2)*. *Nat. Commun.* (2016). doi:10.1038/ncomms11605

222. Seno, E. T. & Chater, K. F. Glycerol catabolic enzymes and their regulation in wild-type and mutant strains of *Streptomyces coelicolor A3(2)*. *J. Gen. Microbiol.* **129**, 1403–1413 (1983).

223. Bompelly, R. & Skaf, D. W. Optimization of a Colorimetric Test Method for Quantifying Glycerol in Aqueous Solution. *J Am Oil Chem Soc* **91**, 1605–1610 (2014).

224. Nash, T. The Colorimetric Estimation of Formaldehyde by Means of the Hantzsch Reaction. 416–421 (1953).

225. Hrdlickova, R., Toloue, M. & Tian, B. RNA-Seq methods for transcriptome analysis. *Wiley Interdiscip. Rev. RNA* **8**, (2017).

226. Agarwal, A. *et al.* Comparison and calibration of transcriptome data from RNA-Seq and tiling arrays. *BMC Genomics* **11**, (2010).

227. Dohm, J. C., Lottaz, C., Borodina, T. & Himmelbauer, H. Substantial biases in ultra-

short read data sets from high-throughput DNA sequencing. *Nucleic Acids Res.* **36**, (2008).

228. Minoche, A. E., Dohm, J. C. & Himmelbauer, H. Evaluation of genomic high-throughput seuqencing data generated on illumina HiSeq and Genome Analyzer systems. *Genome Biol.* (2011). doi:10.1186/gb-2011-12-11-r112

229. Kircher, M., Stenzel, U. & Kelso, J. Improved base calling for the Illumina Genome Analyzer using machine learning strategies. *Genome Biol.* **10**, (2009).

230. Kircher, M. & Kelso, J. High-throughput DNA sequencing - Concepts and limitations. *BioEssays* **32**, 524–536 (2010).

231. Liu, Y., Zhou, J. & White, K. P. RNA-seq differential expression studies: More sequence or more replication? *Bioinformatics* **30**, 301–304 (2014).

232. Robinson, M. D. & Oshlack, A. A scaling normalization method for differential expression analysis of RNA-seq data. *Genome Biol.* **11**, (2010).

233. Costa-Silva, J., Domingues, D. & Lopes, F. M. RNA-Seq differential expression analysis: An extended review and a software tool. *PLoS One* **12**, 1–18 (2017).

234. Law, C. W., Chen, Y., Shi, W. & Smyth, G. K. voom: precision weights unlock linear model analysis tools for RNA-seq read counts. *Genome Biol.* **15**, R29 (2014).

235. Lybecker, M., Bilusic, I. & Raghavan, R. Pervasive transcription: detecting functional RNAs in bacteria. *Transcription* **5**, 1–5 (2014).

236. Palazzo, A. F. & Lee, E. S. Non-coding RNA: What is functional and what is junk? *Front. Genet.* **5**, 1–11 (2015).

237. DeRisi, J. *et al.* Use of a cDNA microarray to analyse gene expression patterns in human cancer. *Nat. Genet.* **14**, 409–416 (1996).

238. Lu, S., Killoran, P. B. & Riley, L. W. Association of Salmonella enterica Serovar *Enteritidis* YafD with Resistance to Chicken Egg Albumen. *Infect. Immun.* **71**, 6734–6741 (2003).

239. Hasegawa, T. *et al.* Characterization of a virulence-associated and cell- wall-located DNase of *Streptococcus pyogenes*. **370**, 184–190 (2010).

240. Liu, G., Chater, K. F., Chandra, G., Niu, G. & Tan, H. Molecular regulation of antibiotic biosynthesis in streptomyces. *Microbiol. Mol. Biol. Rev.* **77**, 112–143 (2013).

241. Chater, K. F. Recent advances in understanding Streptomyces. *F1000Research* **5**, 1–16 (2016).

242. Ochi, K. Metabolic Initiation of Differentiation and Secondary Metabolism by *Streptomyces griseus*: Significance of the Stringent Response (ppGpp) and GTP Content in Relation to A Factor. *J. Bacteriol.* **169**, 3608–3616 (1987).

243. Chakraburtty, R. & Bibb, M. The ppGpp Synthetase Gene (relA) of *Streptomyces coelicolor A3(2)* Plays a Conditional Role in Antibiotic Production and Morphological Differentiation. *J. Bacteriol.* **179**, 5854–5861 (1997).

244. Potrykus, K., Murphy, H., Philippe, N. & Cashel, M. ppGpp is the major source of growth rate control in E . coli. *Environ. Microbiol.* **13**, 563–575 (2011).

245. Magnusson, L. U., Farewell, A. & Nystro, T. ppGpp : a global regulator in *Escherichia coli*. *Trends Microbiol.* **13**, (2005).

246. Yague, P. *et al.* Transcriptomic Analysis of Liquid Non-Sporulating *Streptomyces coelicolor* Cultures Demonstrates the Existence of a Complex Differentiation Comparable to That Occurring in Solid Sporulating Cultures. *PLoS One* **9**, (2014).

247. Van Wezel, G. P., Meulen, J. Van Der, Luiten, R. G. M. & Koerten, H. K. *ssgA* Is Essential for Sporulation of *Streptomyces coelicolor A3(2)* and Affects Hyphal Development by Stimulating Septum Formation. *J. Bacteriol.* **182**, 5653–62 (2000).

248. Botas, A., Pérez-redondo, R. & Rodríguez-garcía, A. ArgR of *Streptomyces coelicolor* Is a Pleiotropic Transcriptional Regulator : Effect on the Transcriptome , Antibiotic Production , and Differentiation in Liquid Cultures. *Front. Microbiol.* **9**, 1–18 (2018).

249. Geiman, D. E., Raghunand, T. R., Agarwal, N. & Bishai, W. R. Differential Gene Expression in Response to Exposure to Antimycobacterial Agents and Other Stress Conditions among Seven Mycobacterium tuberculosis *whiB*- Like Genes. *Antimicrob. Agents Chemother.* **50**, 2836–2841 (2006).

250. Gusarov, I., Shatalin, K., Starodubtseva, M. & Nudler, E. Endogenous Nitric Oxide Protects Bacteria Against a Wide Spectrum of Antibiotics. *Science (80-. ).* (2009). doi:10.1126/science.1175439

251. Sasaki, Y., Takaya, N., Morita, A. & Nakamura, A. Nitrite formation from organic nitrogen by *Streptomyces antibioticus* supporting bacterial cell growth and possible involvement of nitric oxide as an intermediate. *Biosci. Biotechnol. Biochem.* **78**, 1599–1606 (2014).

252. Romero, J., Liras, P. & Martin, J. F. Utilization of ornithine and arginine as specific precursors of clavulanic acid. *Appl. Environ. Microbiol.* **52**, 892–897 (1986).

253. Baños, S., Párez-Redondo, R., Koekman, B. & Liras, P. Glycerol utilization gene cluster in *Streptomyces clavuligerus*. *Appl. Environ. Microbiol.* **75**, 2991–2995 (2009).

254. Widdick, D. A. *et al.* The twin-arginine translocation pathway is a major route of protein export in *Streptomyces coelicolor*. *Proc. Natl. Acad. Sci. U. S. A.* **103**, 17927–17932 (2006).

255. Wecke, T., Veith, B., Ehrenreich, A. & Mascher, T. Cell Envelope Stress Response in *Bacillus licheniformis* : Integrating Comparative Genomics , Transcriptional Profiling , and Regulon Mining To Decipher a Complex Regulatory Network. *J. Bacteriol.* **188**, 7500–7511 (2006).

256. Krulwich, T. A., Sachs, G. & Padan, E. Molecular aspects of bacterial pH sensing and homeostasis. *Nat. Rev. Microbiol.* **9**, 330–343 (2011).

257. Kuroda, M., Kuwahara-arai, K. & Hiramatsu, K. Identification of the Up- and Down-Regulated Genes in Vancomycin-Resistant *Staphylococcus aureus* Strains Mu3 and Mu50 by cDNA Differential Hybridization Method. *Biomed. Biophys. Res. Commun.* **490**, 485–490 (2000).

258. Haas, W., Kaushal, D., Sublett, J., Obert, C. & Tuomanen, E. I. Vancomycin Stress Response in a Sensitive and a Tolerant Strain of *Streptococcus pneumoniae*. *J. Bacteriol.* **187**, 8205–8210 (2005).

259. Muthaiyan, A., Silverman, J. A., Jayaswal, R. K. & Wilkinson, B. J. Transcriptional Profiling Reveals that Daptomycin Induces the *Staphylococcus aureus* Cell Wall Stress Stimulon and Genes Responsive to Membrane Depolarization. *Antimicrob. Agents Chemother.* **52**, 980–990 (2008).

260. Kumar, S., Rubino, F. A., Mendoza, A. G. & Ruiz, N. The bacterial lipid II flippase MurJ functions by an alternating-access mechanism. *J. Biol. Chem.* jbc.RA118.006099 (2018). doi:10.1074/jbc.RA118.006099

261. Liston, S. D., Mann, E. & Whitfield, C. Glycolipid substrates for ABC transporters required for the assembly of bacterial cell-envelope and cell-surface glycoconjugates. *Biochim. Biophys. Acta - Mol. Cell Biol. Lipids* **1862**, 1394–1403 (2017).

262. Arias, C. A. *et al.* Genetic Basis for In Vivo Daptomycin Resistance in Enterococci. *N. Engl. J. Med.* **365**, 892–900 (2011).

263. Ruzin, a, Lindsay, J. & Novick, R. P. Molecular genetics of SaPI1 - a mobile pathogenicity island in *Staphylococcus aureus*. *Mol. Microbiol.* **41**, 365–377 (2001).

264. Comba, S., Menendez-Bravo, S., Arabolaza, A. & Gramajo, H. Identification and physiological characterization of phosphatidic acid phosphatase enzymes involved in triacylglycerol biosynthesis in Streptomyces coelicolor. *Microb Cell Fact* **12**, 9 (2013).

265. Wang, Y., Cobb, R. E. & Zhao, H. High-Efficiency Genome Editing of Streptomyces Species by an Engineered CRISPR / Cas System. *Synthetic Biology and Metabolic Engineering in Plants and Microbes Part A: Metabolism in Microbes* **575**, (Elsevier Inc., 2016).

266. Alberti, F. & Corre, C. Editing streptomycete genomes in the CRISPR/Cas9 age. *Nat. Prod. Rep.* 1237–1248 (2019). doi:10.1039/c8np00081f

267. Tong, Y. *et al.* Highly efficient DSB-free base editing for streptomycetes with CRISPR-BEST. *PNAS* **116**, (2019).

268. Spoering, A. L., Vulic, M. & Lewis, K. GlpD and PlsB Participate in Persister Cell Formation in *Escherichia coli* GlpD and PlsB Participate in Persister Cell Formation in Escherichia coli. *J. Bacteriol.* (2015). doi:10.1128/JB.00369-06

269. Burg, M. B. & Ferraris, J. D. Intracellular Organic Osmolytes: Function and Regulation. *J. Biol. Chem.* **283**, 7309–7313 (2009).

270. Stukey, J. & Carman, G. M. Identification of a novel phosphatase sequence motif. *Protein Sci.* 469–472 (1997).

271. Podlesek, Z., Comino, A., Herzog-Velikonja, B. & Grabnar, M. The role of the bacitracin ABC transporter in bacitracin resistance and collateral detergent sensitivity. *FEMS Microbiol. Lett.* **188**, 103–106 (2000).

272. Ruzin, A. *et al.* Inactivation of mprF affects vancomycin susceptibility in *Staphylococcus aureus*. *Biochim. Biophys. Acta - Gen. Subj.* **1621**, 117–121 (2003).

273. Hachmann, A. *et al.* Reduction in Membrane Phosphatidylglycerol Content Leads to Daptomycin Resistance in Bacillus subtilis. *Antimicrob. Agents Chemother.* **55**, 4326–4337 (2011).

274. Iwamoto, K. *et al.* Curvature-dependent recognition of ethanolamine phospholipids by

duramycin and cinnamycin. *Biophys. J.* **93**, 1608–1619 (2007).

275. Epand, R. M., Walker, C., Epand, R. F. & Magarvey, N. A. Molecular mechanisms of membrane targeting antibiotics. *Biochim. Biophys. Acta - Biomembr.* **1858**, 980–987 (2016).

276. Kahne, D., Leimkuhler, C., Lu, W. & Walsh, C. Glycopeptide and Lipoglycopeptide Antibiotics. *ACS Chem. Rev.* (2005).

277. Thaker, M. N. & Wright, G. D. Opportunities for Synthetic Biology in Antibiotics: Expanding Glycopeptide Chemical Diversity. *ACS Synth. Biol.* (2012). doi:10.1021/sb300092n

278. Mackay, J. P. *et al.* Glycopeptide Antibiotic Activity and the Possible Role of Dimerization: A Model for Biological Signaling. *J. Am. Chem. Soc.* **116**, 4581–4590 (1994).

279. Gerhard, U., Mackay, J. P., Maplestone, R. A. & Williams, D. H. The Role of the Sugar and Chlorine Substituents in the Dimerization of Vancomycin Antibiotics. *J. Am. Chem. Soc.* 232–237 (1993).

280. Cooper, M. A., Fiorini, M. T., Abell, C. & Williams, D. H. Binding of Vancomycin Group Antibiotics to D-Alanine and D-Lactate Presenting Self-Assembled Monolayers. *Bioorg. Med. Chem.* **8**, 2609–2616 (2000).

281. Allen, N. E., LeTourneau, D. L. & Hobbs Jr, J. N. The role of hydrophobic side chains as determinants of antibacterial activity of semisynthetic glycopeptide antibiotics. *J. Antibiot. (Tokyo).* **50**, 677–84 (1997).

# Appendix

**Appendix:** Genes identified as differentially expressed in either strain of this RNA-seq dataset comparing T30/T0 or T90/T30 time points. Fold-change (FC) is also shown next to the adjusted P-value which is highlighted in red if value <0.05.

| SCO_ID | Annotation | Gene | Function | M600 T30 - T0 | | M600 T90 - T30 | | VANJ T30 - T0 | | VANJ T90 - T30 | |
|---|---|---|---|---|---|---|---|---|---|---|---|
| | | | | FC | adj.P.val | FC | adj.P.val | FC | adj.P.val | FC | adj.P.val |
| **Carbon utilisation** | | | | | | | | | | | |
| **SCO1214** | ATP-dependent 6-phosphofructokinase 3 | pfkA3 | Glycolysis | 0.6 | n/s | 0.8 | n/s | 0.2 | 3.0E-21 | 2.4 | 4.3E-06 |
| **SCO1930** | Cytochrome c oxidase assembly protein subunit 15 | | cytochrome biosynthesis | 0.3 | 1.9E-10 | 1.1 | n/s | 0.2 | 2.2E-21 | 2.4 | 6.0E-04 |
| **SCO1934** | Protoheme IX farnesyltransferase | | cytochrome biosynthesis | 0.3 | 3.3E-10 | 0.8 | n/s | 0.1 | 2.2E-25 | 2.4 | 1.5E-04 |
| **SCO1942** | Glucose-6-phosphate isomerase 2 | pgi2 | Glycolysis | 0.7 | n/s | 0.9 | n/s | 0.2 | 3.1E-16 | 3.2 | 9.0E-11 |
| **SCO1945** | Triosephosphate isomerase | tpiA | Glycolysis | 0.7 | n/s | 0.8 | n/s | 0.2 | 4.4E-24 | 3.3 | 6.0E-17 |
| **SCO1946** | Phosphoglycerate kinase | pgk | Glycolysis | 0.8 | n/s | 0.8 | n/s | 0.2 | 1.8E-22 | 3.6 | 6.1E-18 |
| **SCO1947** | Glyceraldehyde-3-phosphate dehydrogenase | gap1 | Glycolysis | 0.7 | n/s | 0.6 | n/s | 0.2 | 3.6E-26 | 2.7 | 1.1E-14 |
| **SCO2014** | Pyruvate kinase | pyk1 | Glycolysis | 0.7 | n/s | 0.8 | n/s | 0.2 | 6.1E-21 | 3.2 | 2.4E-15 |
| **SCO2126** | Glucokinase | glk | Glycolysis | 0.6 | n/s | 1 | n/s | 0.2 | 1.9E-21 | 3.3 | 7.0E-15 |
| **SCO2736** | citrate synthase | citA | TCA cycle | 2.8 | 4.7E-05 | 0.8 | n/s | 1.1 | n/s | 1.9 | n/s |
| **SCO3096** | Enolase 1 | eno1 | Glycolysis | 0.9 | n/s | 0.9 | n/s | 0.3 | 1.2E-18 | 3.4 | 1.3E-18 |
| **SCO3311** | Delta-aminolevulinic acid dehydratase | hemB | cytochrome biosynthesis | 0.2 | 1.2E-15 | 1.3 | n/s | 0.2 | 1.8E-20 | 1.7 | n/s |
| **SCO3317** | Putative uroporphyrin-III C-methyltransferase/uroporphyrinogen-III synthase | | cytochrome biosynthesis | 0.4 | 2.2E-03 | 1.4 | n/s | 0.2 | 1.0E-13 | 3 | 5.5E-07 |
| **SCO3318** | Porphobilinogen deaminase 1 | hemC1 | cytochrome biosynthesis | 0.3 | 2.4E-12 | 1.1 | n/s | 0.2 | 6.9E-23 | 1.8 | n/s |
| **SCO3319** | putative glutamyl-tRNA reductase | hemA | cytochrome biosynthesis | 0.4 | 5.2E-05 | 1.3 | n/s | 0.2 | 1.3E-19 | 1.9 | n/s |
| **SCO3649** | Fructose-bisphosphate aldolase | fba | Glycolysis | 0.8 | n/s | 0.8 | n/s | 0.2 | 1.3E-22 | 3.1 | 1.8E-15 |
| **SCO4808** | succinyl-CoA synthetase beta chain | sucC | TCA cycle | 0.6 | n/s | 0.9 | n/s | 0.2 | 5.3E-28 | 3.5 | 2.4E-20 |
| **SCO4809** | succinyl CoA synthetase alpha chain | sucD | TCA cycle | 0.6 | n/s | 1 | n/s | 0.2 | 1.3E-26 | 3.6 | 9.3E-20 |
| **SCO4827** | malate dehydrogenase | mdh | TCA cycle | 0.8 | n/s | 0.7 | n/s | 0.3 | 4.2E-18 | 3.8 | 9.5E-18 |
| **SCO4855** | Succinate dehydrogenase iron-sulfur subunit | dhsB | TCA cycle | 0.4 | 3.5E-12 | 0.9 | n/s | 0.1 | 1.8E-27 | 2.8 | 2.8E-11 |
| **SCO4856** | putative succinate dehydrogenase flavoprotein subunit | dhsA | TCA cycle | 0.3 | 7.8E-15 | 0.8 | n/s | 0.1 | 5.6E-29 | 2.6 | 1.8E-09 |
| **SCO4858** | putative succinate dehydrogenase membrane subunit | | TCA cycle | 0.3 | 1.1E-13 | 0.7 | n/s | 0.1 | 9.5E-26 | 3 | 1.6E-09 |
| **SCO5226** | ribonucleotide-diphosphate reductase large chain | nrdL | Nucleotide biosynthesis | 0.6 | n/s | 0.7 | n/s | 0.3 | 7.8E-16 | 1.7 | n/s |
| **SCO5281** | putative 2-oxoglutarate dehydrogenase | | TCA cycle | 0.4 | 9.9E-15 | 1.4 | n/s | 0.1 | 2.1E-30 | 4.1 | 8.4E-22 |
| **SCO5370** | ATP synthase delta chain | atpH | ATP production | 0.6 | n/s | 1.5 | n/s | 0.2 | 4.1E-27 | 4 | 1.5E-21 |
| **SCO5371** | ATP synthase alpha chain | atpA | ATP production | 0.5 | 6.9E-04 | 1 | n/s | 0.1 | 1.4E-29 | 3.3 | 5.1E-19 |
| **SCO5372** | ATP synthase gamma chain | atpG | ATP production | 0.6 | n/s | 1.1 | n/s | 0.2 | 1.1E-26 | 3.7 | 1.2E-19 |
| **SCO5373** | ATP synthase beta chain | atpD | ATP production | 0.5 | n/s | 1 | n/s | 0.2 | 1.3E-28 | 3.3 | 1.2E-18 |
| **SCO5374** | ATP synthase epsilon chain | atpC | ATP production | 0.5 | n/s | 0.9 | n/s | 0.2 | 3.5E-25 | 2.8 | 7.6E-12 |

| | | | | | | | | | | | |
|---|---|---|---|---|---|---|---|---|---|---|---|
| **SCO5423** | Pyruvate kinase | pyk2 | Glycolysis | 0.6 | n/s | 1 | n/s | 0.2 | 8.7E-25 | 3.7 | 1.2E-17 |
| **SCO5859** | ferrochelatase | hemH | cytochrome biosynthesis | 0.5 | n/s | 1.4 | n/s | 0.2 | 1.2E-16 | 2.8 | 4.7E-09 |
| **SCO6031** | uroporphyrinogen decarboxylase | hemE | cytochrome biosynthesis | 0.3 | 9.2E-10 | 1.2 | n/s | 0.2 | 1.1E-21 | 2.5 | 8.8E-05 |
| **SCO7000** | isocitrate dehydrogenase | idh | TCA cycle | 0.8 | n/s | 0.8 | n/s | 0.2 | 4.6E-23 | 3.6 | 2.0E-19 |
| | | | | | | | | | | | |
| **DNA Replication** | | | | | | | n/s | | | | |
| **SCO2162** | Quinolinate synthetase A | nadA | Nucleotide biosynthesis | 1 | n/s | 0.7 | n/s | 0.4 | 9.8E-12 | 3.2 | 1.1E-14 |
| **SCO3543** | DNA topoisomerase I | topA | DNA replication | 0.7 | n/s | 1.1 | n/s | 0.3 | 1.9E-18 | 2.8 | 1.3E-12 |
| **SCO3873** | DNA gyrase subunit A | gyrA | DNA replication | 0.8 | n/s | 0.9 | n/s | 0.2 | 6.2E-23 | 2.9 | 1.1E-14 |
| **SCO3874** | DNA gyrase subunit B | gyrB | DNA replication | 1.2 | n/s | 1 | n/s | 0.2 | 6.6E-22 | 3.7 | 1.7E-18 |
| **SCO5803** | SOS regulatory protein | lexA | DNA repair | 1.2 | n/s | 0.9 | n/s | 0.4 | 3.6E-08 | 3 | 5.5E-11 |
| **SCO5805** | ribonucleotide reductase | nrdJ | DNA replication | 0.7 | n/s | 0.9 | n/s | 0.2 | 2.7E-24 | 2.9 | 2.9E-14 |
| **SCO5812** | Ribonuclease HII | rnhB | DNA replication | 0.9 | n/s | 0.4 | 6.0E-03 | 0.3 | 2.9E-17 | 1.4 | n/s |
| **SCO5822** | putative DNA gyrase subunit B | gyrB2 | DNA replication | 0.6 | n/s | 1.1 | n/s | 0.2 | 1.6E-23 | 3.7 | 3.2E-17 |
| **SCO5836** | DNA gyrase | | DNA replication | 0.4 | 4.6E-08 | 1.1 | n/s | 0.1 | 8.5E-25 | 2.7 | 3.6E-09 |
| | | | | | | | | | | | |
| **Translation** | | | | | | | | | | | |
| **SCO0436** | 50S ribosomal protein L32-2 | rpmF2 | Translation | 0.9 | n/s | 1.8 | n/s | 0.8 | n/s | 1.1 | n/s |
| **SCO0569** | 50S ribosomal protein L36 2 | rpmJ2 | Translation | 1.7 | n/s | 0.9 | n/s | 0.6 | n/s | 2.8 | 6.4E-04 |
| **SCO0570** | 50S ribosomal protein L33 3 | rpmG3 | Translation | 1.4 | n/s | 1.1 | n/s | 0.6 | n/s | 2.1 | n/s |
| **SCO1150** | 50S ribosomal protein L31 type B 1 | rpmE2-1 | Translation | 6.6 | 6.6E-11 | 0.5 | n/s | 2.2 | n/s | 1.3 | n/s |
| **SCO1505** | 30S ribosomal protein S4 | rpsD | Translation | 0.6 | n/s | 1 | n/s | 0.2 | 7.3E-27 | 3.2 | 3.4E-17 |
| **SCO1598** | 50S ribosomal protein L20 | rplT | Translation | 0.8 | n/s | 0.7 | n/s | 0.3 | 1.1E-20 | 2.2 | 5.1E-04 |
| **SCO1599** | 50S ribosomal protein L35 | rpmI | Translation | 0.9 | n/s | 0.7 | n/s | 0.3 | 2.1E-18 | 2.4 | 2.2E-08 |
| **SCO1600** | Translation initiation factor | infC | Initiating translation | 1 | n/s | 0.9 | n/s | 0.7 | n/s | 1.4 | n/s |
| **SCO2533** | Endoribonuclease | ybeY | Maturation of rRNA | 5.8 | 8.5E-24 | 0.9 | n/s | 1.5 | n/s | 2.7 | 2.3E-11 |
| **SCO2563** | 30S ribosomal protein S20 | rpsT | Translation | 0.6 | n/s | 0.8 | n/s | 0.2 | 2.1E-26 | 2.9 | 2.0E-14 |
| **SCO2596** | 50S ribosomal protein L27 | rpmA | Translation | 0.4 | 5.2E-05 | 0.7 | n/s | 0.1 | 5.9E-28 | 2.1 | n/s |
| **SCO2597** | 50S ribosomal protein L21 | rplU | Translation | 0.5 | n/s | 1 | n/s | 0.1 | 6.7E-29 | 3.5 | 3.3E-19 |
| **SCO2599** | Ribonuclease E | | Maturation of rRNA | 2.4 | 4.4E-10 | 1.4 | n/s | 4.2 | 4.6E-10 | 1.8 | n/s |
| **SCO2972** | Peptide chain release factor 2 | prfB | Terminates translation | 0.5 | n/s | 1.1 | n/s | 0.2 | 2.2E-21 | 3 | 1.2E-10 |
| **SCO3009** | Ribosome hibernation promoting factor | hpf | Arresting translation | 5.5 | 2.6E-22 | 1.1 | n/s | 1.8 | n/s | 5.7 | 1.1E-23 |
| **SCO3397** | putative integral membrane lysyl-tRNA synthetase | mprF | lysyl-tRNA synthesis | 11.8 | 2.7E-30 | 0.6 | n/s | 2.9 | 2.2E-13 | 2.4 | 1.3E-09 |
| **SCO3425** | putative 30S ribosomal protein S18 | rpsR2 | Translation | 1.6 | n/s | 1.1 | n/s | 2.7 | n/s | 0.3 | n/s |
| **SCO3427** | 50S ribosomal protein L31 type B 2 | rpmE2-2 | Translation | 1.4 | n/s | 1.4 | n/s | 1.7 | n/s | 0.7 | n/s |
| **SCO3429** | 50S ribosomal protein L28-2 | rpmB2 | Translation | 0.8 | n/s | 1.2 | n/s | 3.5 | 3.8E-02 | 0.2 | 9.8E-03 |
| **SCO3430** | putative 30S ribosomal protein S14 | rpsN | Translation | 1.1 | n/s | 1.2 | n/s | 3 | 1.3E-02 | 0.1 | 4.1E-06 |
| **SCO3778** | putative threonyl tRNA synthetase | | threonyl-tRNA synthesis | 13 | 2.3E-17 | 0.6 | n/s | 8.9 | 9.4E-14 | 0.7 | n/s |
| **SCO3880** | 50S ribosomal protein L34 | rpmH | Translation | 0.4 | 5.6E-04 | 0.6 | n/s | 0.1 | 2.9E-26 | 2.7 | 1.1E-09 |
| **SCO3906** | putative 30S ribosomal protein S6 | rpsF | Translation | 0.3 | 8.5E-16 | 0.8 | n/s | 0.1 | 9.8E-29 | 2.4 | 3.7E-07 |

| | | | | | | | | | | | |
|---|---|---|---|---|---|---|---|---|---|---|---|
| **SCO3908** | putative 30S ribosomal protein S18 | rpsR | Translation | 0.3 | 8.6E-12 | 1 | n/s | 0.1 | 2.0E-26 | 3.3 | 6.3E-14 |
| **SCO3909** | 50S ribosomal protein L9 | rplI | Translation | 0.4 | 7.5E-12 | 1.1 | n/s | 0.1 | 1.2E-27 | 2.8 | 1.8E-11 |
| **SCO4278** | Ribosome associated protein | | Peptide release | 15.6 | 3.0E-23 | 0.7 | n/s | 8.1 | 8.9E-15 | 2.5 | 1.6E-05 |
| **SCO4635** | 50S ribosomal protein L33 2 | rpmG2 | Translation | 0.4 | 3.2E-04 | 0.6 | n/s | 0.1 | 1.2E-27 | 1.9 | n/s |
| **SCO4648** | 50S ribosomal protein L11 | rplK | Translation | 0.5 | n/s | 1.1 | n/s | 0.2 | 1.5E-27 | 3.7 | 9.3E-20 |
| **SCO4649** | 50S ribosomal protein L1 | rplA | Translation | 0.5 | n/s | 0.9 | n/s | 0.2 | 1.2E-27 | 3.5 | 2.1E-19 |
| **SCO4653** | 50S ribosomal protein L7/L12 | rplL | Translation | 0.6 | n/s | 1.1 | n/s | 0.2 | 4.5E-28 | 3.1 | 7.6E-18 |
| **SCO4659** | 30S ribosomal protein S12 | rpsL | Translation | 0.7 | n/s | 0.8 | n/s | 0.2 | 7.7E-25 | 2.5 | 6.8E-10 |
| **SCO4660** | 30S ribosomal protein S7 | rpsG | Translation | 0.7 | n/s | 0.7 | n/s | 0.2 | 3.3E-24 | 2.2 | 4.2E-05 |
| **SCO4701** | 30S ribosomal protein S10 | rpsJ | Translation | 0.4 | 1.7E-10 | 1 | n/s | 0.1 | 3.2E-28 | 2.2 | 3.1E-03 |
| **SCO4702** | 50S ribosomal protein L3 | rplC | Translation | 0.5 | 3.4E-03 | 0.9 | n/s | 0.1 | 7.8E-29 | 2.7 | 7.1E-13 |
| **SCO4703** | 50S ribosomal protein L4 | rplD | Translation | 0.5 | 7.7E-05 | 0.9 | n/s | 0.1 | 5.2E-29 | 2.9 | 4.5E-15 |
| **SCO4704** | 50S ribosomal protein L23 | rplW | Translation | 0.6 | n/s | 0.9 | n/s | 0.2 | 6.5E-27 | 2.8 | 1.8E-14 |
| **SCO4706** | 30S ribosomal protein S19 | rpsS | Translation | 0.5 | n/s | 0.7 | n/s | 0.2 | 3.5E-25 | 2.2 | 5.4E-03 |
| **SCO4707** | 50S ribosomal protein L22 | rplV | Translation | 0.6 | n/s | 0.8 | n/s | 0.2 | 1.1E-24 | 2.9 | 5.0E-14 |
| **SCO4708** | 30S ribosomal protein S3 | rpsC | Translation | 0.9 | n/s | 1.3 | n/s | 0.2 | 8.2E-18 | 4.2 | 6.4E-18 |
| **SCO4709** | 50S ribosomal protein L16 | rplP | Translation | 0.9 | n/s | 1 | n/s | 0.3 | 3.7E-17 | 3.7 | 1.1E-15 |
| **SCO4710** | 50S ribosomal protein L29 | rpmC | Translation | 0.9 | n/s | 0.9 | n/s | 0.2 | 9.8E-19 | 3.2 | 1.1E-13 |
| **SCO4711** | 30S ribosomal protein S17 | rpsQ | Translation | 0.8 | n/s | 1.2 | n/s | 0.2 | 1.9E-19 | 3.6 | 2.5E-16 |
| **SCO4712** | 50S ribosomal protein L14 | rplN | Translation | 0.9 | n/s | 0.9 | n/s | 0.3 | 3.6E-16 | 2.7 | 3.0E-08 |
| **SCO4713** | 50S ribosomal protein L24 | rplX | Translation | 0.7 | n/s | 0.7 | n/s | 0.2 | 3.4E-13 | 2.3 | n/s |
| **SCO4715** | 30S ribosomal protein S14 | rpsN | Translation | 0.8 | n/s | 0.8 | n/s | 0.3 | 6.0E-16 | 2.6 | 2.3E-07 |
| **SCO4716** | 30S ribosomal protein S8 | rpsH | Translation | 0.6 | n/s | 1.1 | n/s | 0.2 | 2.0E-22 | 3.2 | 1.9E-13 |
| **SCO4717** | 50S ribosomal protein L6 | rplF | Translation | 0.7 | n/s | 1 | n/s | 0.2 | 8.7E-26 | 3.6 | 4.5E-20 |
| **SCO4718** | 50S ribosomal protein L18 | rplR | Translation | 0.7 | n/s | 1 | n/s | 0.2 | 3.4E-24 | 3.2 | 1.1E-15 |
| **SCO4719** | 30S ribosomal protein S5 | rpsE | Translation | 0.7 | n/s | 1 | n/s | 0.2 | 4.0E-26 | 3.3 | 2.6E-19 |
| **SCO4720** | 50S ribosomal protein L30 | rpmD | Translation | 0.7 | n/s | 1.1 | n/s | 0.2 | 1.8E-20 | 3.6 | 1.3E-16 |
| **SCO4721** | 50S ribosomal protein L15 | rplO | Translation | 0.7 | n/s | 0.9 | n/s | 0.2 | 3.2E-25 | 3 | 2.5E-16 |
| **SCO4725** | Translation initiation factor IF-1 | infA | Starts translation | 0.7 | n/s | 0.9 | n/s | 0.2 | 1.7E-24 | 3.2 | 3.2E-16 |
| **SCO4726** | 50S ribosomal protein L36 1 | rpmJ1 | Translation | 0.7 | n/s | 1.1 | n/s | 0.2 | 4.3E-22 | 3 | 1.8E-13 |
| **SCO4727** | 30S ribosomal protein S13 | rpsM | Translation | 0.5 | n/s | 0.7 | n/s | 0.2 | 1.9E-25 | 2.5 | 2.3E-09 |
| **SCO4728** | 30S ribosomal protein S11 | rpsK | Translation | 0.7 | n/s | 1.1 | n/s | 0.2 | 1.6E-23 | 3.9 | 5.6E-20 |
| **SCO4730** | 50S ribosomal protein L17 | rplQ | Translation | 0.9 | n/s | 1.2 | n/s | 0.2 | 3.9E-24 | 3.2 | 6.2E-19 |
| **SCO4734** | 50S ribosomal protein L13 | rplM | Translation | 0.5 | 2.3E-05 | 0.8 | n/s | 0.1 | 4.8E-29 | 2.7 | 5.6E-13 |
| **SCO4735** | 30S ribosomal protein S9 | rpsI | Translation | 0.4 | 1.6E-14 | 0.7 | n/s | 0.1 | 2.6E-29 | 4.2 | 2.3E-21 |
| **SCO5252** | Dihydrofolate reductase | | Nucleic acid biosynthesis | 3.6 | 2.5E-14 | 0.5 | n/s | 1.1 | n/s | 1.9 | n/s |
| **SCO5359** | 50S ribosomal protein L31 | rpmE | Translation | 0.5 | n/s | 0.8 | n/s | 0.2 | 8.2E-24 | 3 | 1.7E-13 |
| **SCO5360** | Peptide chain release factor 1 | prfA | Terminates translation | 0.5 | n/s | 1.1 | n/s | 0.2 | 1.5E-23 | 3 | 1.7E-11 |
| **SCO5564** | 50S ribosomal protein L28-1 | rpmB1 | Translation | 0.3 | 3.5E-18 | 0.7 | n/s | 0.1 | 3.5E-30 | 1.9 | n/s |
| **SCO5571** | 50S ribosomal protein L32-1 | rpmF1 | Translation | 0.6 | n/s | 0.8 | n/s | 0.2 | 5.7E-24 | 3 | 1.3E-12 |

| | | | | | | | | | | | |
|---|---|---|---|---|---|---|---|---|---|---|---|
| **SCO5591** | 30S ribosomal protein S16 | rpsP | Translation | 0.6 | n/s | 0.9 | n/s | 0.2 | 3.9E-28 | 2.6 | 5.9E-11 |
| **SCO5595** | 50S ribosomal protein L19 | rplS | Translation | 0.4 | 1.4E-13 | 0.7 | n/s | 0.1 | 1.3E-27 | 2.3 | 1.1E-05 |
| **SCO5624** | 30S ribosomal protein S2 | rpsB | Translation | 0.5 | n/s | 1 | n/s | 0.2 | 1.6E-27 | 3.4 | 7.6E-19 |
| **SCO5708** | Ribosome binding factor | | Maturation of rRNA | 4.5 | 1.2E-20 | 0.7 | n/s | 1.1 | n/s | 2.3 | 4.1E-05 |
| **SCO5709** | Pseudouridine synthase | | rRNA modification | 3.7 | 3.1E-18 | 1 | n/s | 1 | n/s | 3.7 | 1.9E-18 |
| **SCO5736** | 30S ribosomal protein S15 | rpsO | Translation | 0.4 | 5.2E-06 | 1.2 | n/s | 0.1 | 3.0E-26 | 3.3 | 7.7E-15 |
| **SCO5796** | GTPase | hflX | Possible role in protein synthesis | 28.2 | 8.7E-35 | 0.8 | n/s | 8.6 | 4.1E-27 | 2.8 | 1.1E-16 |
| | | | | | | | | | | | |
| **Transcription** | | | | | | | | | | | |
| **SCO0600** | RNA polymerase sigma factor sig8 | $\sigma^{B}$ | Sigma factor | 5.3 | 1.4E-24 | 0.4 | 3.3E-08 | 1.3 | n/s | 3.1 | 5.2E-16 |
| **SCO0803** | putative RNA polymerase sigma factor | | Sigma factor | 4.1 | 3.9E-17 | 1 | n/s | 1.3 | n/s | 2.5 | 2.9E-07 |
| **SCO0870** | putative two-component system response regulator | | Histidine kinase response regulator | 0.2 | 4.5E-12 | 0.7 | n/s | 0.1 | 1.0E-19 | 1.8 | n/s |
| **SCO0871** | putative two-component sensor protein | | Histidine kinase response regulator | 0.3 | 4.6E-12 | 0.9 | n/s | 0.1 | 8.2E-23 | 1.7 | n/s |
| **SCO1876** | putative RNA polymerase sigma factor | | Sigma factor | 48.4 | 1.3E-24 | 1.5 | n/s | 7.6 | 1.4E-16 | 3.6 | 3.5E-16 |
| **SCO2110** | putative eukaryotic-type serine/threonine protein kinase | pkaF | Serine/threonine response regulator | 4.6 | 4.6E-22 | 1 | n/s | 1.2 | n/s | 2.7 | 1.3E-11 |
| **SCO2639** | putative RNA polymerase sigma factor | | Sigma factor | 28.2 | 2.3E-18 | 1.3 | n/s | 6.8 | 7.4E-11 | 2.4 | 4.1E-04 |
| **SCO2666** | putative serine/threonine protein kinase | | Serine/threonine response regulator | 5.3 | 5.0E-22 | 1 | n/s | 1.4 | n/s | 2.3 | 4.2E-05 |
| **SCO2974** | serine/threonine protein kinase | pkaA | Serine/threonine response regulator | 3.2 | 1.3E-15 | 0.9 | n/s | 1 | n/s | 2.7 | 9.5E-11 |
| **SCO3202** | RNA polymerase principal sigma factor | $\sigma^{hrdD}$ | Sigma factor | 22.6 | 6.0E-29 | 0.9 | n/s | 4.9 | 2.6E-18 | 0.6 | n/s |
| **SCO3356** | ECF sigma factor | $\sigma^{E}$ | Sigma factor | 14.2 | 2.9E-32 | 0.9 | n/s | 4.6 | 1.9E-22 | 2.9 | 3.9E-17 |
| **SCO3358** | two-component system response regulator | cseB | Histidine kinase response regulator | 14.8 | 1.5E-28 | 0.9 | n/s | 3.6 | 2.0E-14 | 3.2 | 3.1E-16 |
| **SCO3613** | putative RNA polymerase sigma factor | | Sigma factor | 6 | 9.3E-24 | 0.8 | n/s | 1.6 | n/s | 2.2 | 2.0E-04 |
| **SCO3821** | serine/threonine protein kinase | pksC | Serine/threonine response regulator | 5.1 | 3.7E-24 | 0.8 | n/s | 1.5 | n/s | 2.6 | 3.4E-12 |
| **SCO4005** | putative RNA polymerase sigma factor | | Sigma factor | 74.6 | 1.0E-26 | 0.5 | n/s | 26.5 | 1.4E-20 | 1.7 | n/s |
| **SCO4155** | putative two-component system sensor | | Histidine kinase response regulator | 13.7 | 7.0E-25 | 1.3 | n/s | 4.4 | 8.2E-15 | 2.9 | 6.2E-13 |
| **SCO4156** | putative two-component systen response regulator | | Histidine kinase response regulator | 21.8 | 1.3E-24 | 1.3 | n/s | 5 | 9.6E-14 | 4.4 | 1.8E-18 |
| **SCO4895** | putative ECF sigma factor | | Sigma factor | 4.7 | 1.2E-17 | 0.7 | n/s | 2 | n/s | 1.8 | n/s |
| **SCO4908** | putative RNA polymerase sigma factor | $\sigma^{Q}$ | Sigma factor | 182.9 | 5.1E-25 | 0.6 | n/s | 54.8 | 7.2E-18 | 2.7 | 6.2E-08 |
| **SCO5147** | putative ECF-subfamily sigma factor | | Sigma factor | 9.5 | 1.3E-27 | 0.9 | n/s | 2.6 | 6.2E-09 | 4.2 | 2.2E-21 |
| **SCO5216** | RNA polymerase sigma factor | $\sigma^{R}$ | Sigma factor | 2.9 | 1.4E-13 | 0.7 | n/s | 0.8 | n/s | 2.7 | 4.5E-11 |
| **SCO5934** | putative sigma factor | | Sigma factor | 9.2 | 1.2E-14 | 1.8 | n/s | 3.2 | 1.2E-04 | 3.3 | 3.5E-10 |
| **SCO6520** | putative RNA polymerase sigma factor | | Sigma factor | 7.2 | 4.0E-20 | 0.5 | n/s | 1.9 | n/s | 2.1 | n/s |
| **SCO7278** | putative RNA polymerase sigma factor | $\sigma^{L}$ | Sigma factor | 3 | 5.2E-13 | 0.4 | 1.5E-05 | 0.6 | n/s | 2.6 | 1.2E-08 |
| **SCO7314** | probable RNA polymerase sigma factor | $\sigma^{M}$ | Sigma factor | 9.6 | 2.1E-23 | 0.7 | n/s | 1.8 | n/s | 5.2 | 2.3E-19 |
| **SCO7327** | putative two-component system sensory histidine kinase | | Histidine kinase response regulator | 0.5 | 3.1E-02 | 0.7 | n/s | 0.2 | 5.0E-19 | 0.6 | n/s |
| | | | | | | | | | | | |
| **Amino acid metabolism** | | | | | | | | | | | |
| **SCO1570** | argininosuccinate lyase | argH | Arginine biosynthesis | 14.7 | 1.5E-29 | 0.8 | n/s | 4.8 | 1.2E-19 | 2.8 | 8.3E-14 |
| **SCO1576** | arginine repressor | argR | Arginine biosynthesis | 24 | 5.4E-31 | 0.9 | n/s | 5.8 | 7.9E-21 | 2.9 | 1.4E-14 |

| | | | | | | | | | | | |
|---|---|---|---|---|---|---|---|---|---|---|---|
| **SCO1577** | acetonitrile aminotransferase | argD | Arginine biosynthesis | 29.6 | 1.4E-33 | 0.6 | n/s | 3.4 | 3.7E-02 | 2.6 | 2.1E-13 |
| **SCO1578** | acetylglutamate kinase | argB | Arginine biosynthesis | 31.6 | 3.9E-32 | 0.7 | n/s | 8 | 1.2E-23 | 2.6 | 5.4E-12 |
| **SCO1579** | putative glutamate N-acetyltransferase | argJ | Arginine biosynthesis | 30.7 | 4.2E-33 | 0.8 | n/s | 8.2 | 4.0E-25 | 3.1 | 1.0E-17 |
| **SCO1580** | N-acetyl-gamma-glutamyl-phosphate reductase | argC | Arginine biosynthesis | 31.2 | 4.3E-31 | 0.9 | n/s | 8.4 | 3.0E-23 | 2.9 | 6.5E-14 |
| **SCO3070** | Imidazolonepropionase | hutI | Histidine metabolism | 2 | n/s | 1.4 | n/s | 3.8 | 5.0E-07 | 0.9 | n/s |
| **SCO3073** | urocanate hydratase | hutU | Histidine metabolism | 3.2 | 1.8E-09 | 1.1 | n/s | 1.1 | n/s | 0.8 | n/s |
| **SCO3221** | Prephenate dehydrogenase | | Tyrosine metabolism | 14.9 | 3.2E-12 | 6.2 | 1.9E-15 | 8.8 | 7.3E-12 | 4.8 | 2.8E-18 |
| **SCO3229** | Putative 4-hydroxyphenylpyruvic acid dioxygenase | | Tyrosine metabolism | 13.1 | 1.5E-11 | 7.3 | 2.0E-16 | 7.4 | 5.4E-11 | 6.3 | 3.7E-21 |
| **SCO4932** | histidine ammonia-lyase | hutH | Histidine metabolism | 1.5 | n/s | 3.8 | 1.7E-11 | 2.3 | 1.1E-02 | 5.4 | 3.7E-21 |
| **SCO5976** | ornithine carbamoyltransferase | argF | Arginine biosynthesis | 10.3 | 1.4E-25 | 0.6 | n/s | 3.2 | 2.4E-11 | 2.5 | 2.9E-08 |
| **SCO7036** | argininosuccinate synthase | argG | Arginine biosynthesis | 13.5 | 7.0E-32 | 0.8 | n/s | 3.8 | 1.5E-19 | 3.4 | 1.9E-20 |
| | | | | | | | | | | | |
| **Sporulation** | | | | | | | | | | | |
| **SCO1541** | sporulation-specific cell devision protein | ssgB | Early Sporulation | 3.7 | 5.9E-05 | 0.8 | n/s | 2 | n/s | 4.1 | 9.0E-09 |
| **SCO1950** | probable cell division protein | whiA | Early Sporulation | 1 | n/s | 1.4 | n/s | 0.4 | 3.1E-07 | 3.2 | 8.8E-13 |
| **SCO2082** | cell division protein | ftsZ | Early Sporulation | 1.2 | n/s | 1 | n/s | 0.3 | 1.3E-17 | 3.5 | 3.7E-18 |
| **SCO3034** | transcriptional regulator | whiB | Early Sporulation | 7.1 | 8.5E-23 | 0.9 | n/s | 1.4 | n/s | 2 | n/s |
| **SCO4035** | RNA polymerase sigma factor | sigF | Late Sporulation | 0.6 | n/s | 7.5 | n/s | 5.5 | 9.6E-06 | 0.2 | 2.4E-04 |
| **SCO4543** | | whiJ | Early Sporulation | 1.1 | n/s | 0.6 | n/s | 0.4 | 3.0E-04 | 1.4 | n/s |
| **SCO4767** | transcriptional regulator | whiD | Late Sporulation | 1.6 | n/s | 12.5 | 6.6E-10 | 4.2 | 4.2E-04 | 2.3 | 4.3E-02 |
| **SCO5314** | | whiE ORFVII | Late Sporulation | 1 | n/s | 0.9 | n/s | 2.3 | n/s | 0.5 | n/s |
| **SCO5315** | polyketide cyclase | whiE ORFVI | Late Sporulation | 4.4 | 7.9E-07 | 1 | n/s | 1.5 | n/s | 0.8 | n/s |
| **SCO5316** | acyl carrier protein | whiE ORFV | Late Sporulation | 1 | n/s | 0.7 | n/s | 2.1 | n/s | 0.3 | n/s |
| **SCO5317** | polyketide beta-ketoacyl synthase beta | whiE ORFIV | Late Sporulation | 1.7 | n/s | 1.2 | n/s | 6.3 | 2.9E-10 | 0.1 | 1.2E-09 |
| **SCO5318** | polyketide beta-ketoacyl synthase alpha | whiE ORFIII | Late Sporulation | 1.4 | n/s | 0.7 | n/s | 4.7 | 6.6E-07 | 0.1 | 2.2E-07 |
| **SCO5319** | | whiE ORFII | Late Sporulation | 0.5 | n/s | 2.4 | n/s | 4.6 | 2.5E-04 | 0.2 | 1.3E-04 |
| **SCO5320** | | whiE ORFI | Late Sporulation | 0.9 | n/s | 0.8 | n/s | 6.1 | 1.1E-08 | 0.1 | 1.2E-08 |
| **SCO5321** | polyketide hydroxylase | whiE ORFVIII | Late Sporulation | 1.2 | n/s | 1 | n/s | 7.6 | 2.0E-11 | 0.1 | 6.8E-11 |
| **SCO5621** | RNA polymerase sigma factor | whiG | Early Sporulation | 0.6 | n/s | 0.6 | n/s | 0.4 | 3.3E-02 | 0.7 | n/s |
| **SCO5819** | sporulation transcription factor | whiH | Early Sporulation | 0.8 | n/s | 0.6 | n/s | 0.9 | n/s | 0.7 | n/s |
| **SCO6029** | two-component regulator | whiI | Early Sporulation | 1.4 | n/s | 1.8 | n/s | 3.1 | 2.6E-03 | 0.3 | 7.9E-03 |
| | | | | | | | | | | | |
| **Aerial mycelium** | | | | | | | | | | | |

| | | | | | | | | | | | |
|---|---|---|---|---|---|---|---|---|---|---|---|
| **SCO1674** | chaplin-C | chpC | Aerial mycelium | 1.3 | n/s | 1 | n/s | 0.8 | n/s | 1.6 | n/s |
| **SCO1675** | chaplin-H | chpH | Aerial mycelium | 1 | n/s | 0.7 | n/s | 0.3 | 2.3E-08 | 2.4 | 7.1E-03 |
| **SCO1800** | chaplin-E | chpE | Aerial mycelium | 1.4 | n/s | 0.7 | n/s | 1.1 | n/s | 1.3 | n/s |
| **SCO2699** | chaplin-G | chpG | Aerial mycelium | 0.7 | n/s | 0.8 | n/s | 3.6 | 8.9E-03 | 0.2 | 5.1E-04 |
| **SCO2705** | chaplin-F | chpF | Aerial mycelium | 0.9 | n/s | 1.8 | n/s | 2.3 | n/s | 1 | n/s |
| **SCO2717** | chaplin-D | chpD | Aerial mycelium | 1.1 | n/s | 0.7 | n/s | 3.5 | 2.5E-02 | 0.3 | n/s |
| **SCO2719** | rodlin protein | rdlB | Aerial mycelium | 2.7 | n/s | 1.3 | n/s | 2.5 | n/s | 0.8 | n/s |
| **SCO3323** | putative RNA-polymerase sigma factor | bldN | Aerial mycelium | 1 | n/s | 0.5 | n/s | 1 | n/s | 1 | n/s |
| **SCO4091** | putative DNA binding protein | bldC | Aerial mycelium | 0.5 | n/s | 1.3 | n/s | 0.2 | 5.7E-24 | 4.2 | 9.5E-17 |
| **SCO4768** | putative two-component regulator | bldM | Aerial mycelium | 1 | n/s | 0.9 | n/s | 0.5 | n/s | 1.4 | n/s |
| **SCO5723** | putative regulator | bldB | Aerial mycelium | 0.8 | n/s | 0.8 | n/s | 0.2 | 8.7E-14 | 3.4 | 3.0E-10 |
| **SCO6681** | probable SapB synthase | ramC | Aerial mycelium | 0.7 | n/s | 2.5 | n/s | 3.7 | 6.8E-08 | 0.3 | 5.1E-07 |
| **SCO6682** | lanthionine-containing peptide SapB precursor | ramS | Aerial mycelium | 1.3 | n/s | 2.2 | n/s | 0.6 | n/s | 2.6 | 2.9E-02 |
| **SCO6683** | ABC transporter ATP binding protein | ramA | Aerial mycelium | 1 | n/s | 2.4 | n/s | 4.5 | 5.8E-09 | 0.2 | 2.0E-08 |
| **SCO6684** | ABC transporter ATP binding protein | ramB | Aerial mycelium | 1.5 | n/s | 2.2 | n/s | 8.5 | 6.5E-11 | 0.1 | 2.6E-10 |
| **SCO6685** | response regulator | ramR | Aerial mycelium | 5.9 | 4.9E-13 | 0.6 | n/s | 2.2 | n/s | 1.2 | n/s |
| **SCO7257** | chaplin-B | chpB | Aerial mycelium | 0.8 | n/s | 1.3 | n/s | 7.8 | 1.1E-05 | 0.1 | 2.8E-06 |
| | | | | | | | | | | | |
| **Antibiotic production** | | | | | | | | | | | |
| **SCO5071** | hydroxylacyl-CoA dehydrogenase | | ACT production | 2 | n/s | 8.5 | n/s | 6.4 | 2.6E-03 | 2.6 | n/s |
| **SCO5072** | hydroxylacyl-CoA dehydrogenase | | ACT production | 1.4 | n/s | 13.2 | 1.0E-06 | 6.8 | 2.0E-07 | 2.9 | 4.6E-07 |
| **SCO5073** | putative oxidoreductase | | ACT production | 3.2 | n/s | 8.1 | 2.5E-05 | 7 | 3.9E-08 | 1.8 | n/s |
| **SCO5074** | putative dehydratase | | ACT production | 1.3 | n/s | 6.3 | 1.2E-03 | 2.9 | 4.5E-02 | 4.6 | 3.2E-11 |
| **SCO5075** | putative oxidoreductase | ORF4 | ACT production | 1.1 | n/s | 3.6 | 3.6E-03 | 3.4 | 1.1E-05 | 1.8 | n/s |
| **SCO5076** | integral membrane protein | actVA1 | ACT production | 0.7 | n/s | 1.2 | n/s | 0.8 | n/s | 1.5 | n/s |
| **SCO5077** | uncharacterised protein | actVA2 | ACT production | 0.4 | 9.8E-03 | 1.2 | n/s | 0.3 | 1.0E-05 | 2.8 | 1.4E-04 |
| **SCO5078** | uncharacterised protein | actVA3 | ACT production | 1 | n/s | 1.5 | n/s | 1.3 | n/s | 1.5 | n/s |
| **SCO5079** | | actVA4 | ACT production | 1.7 | n/s | 1.9 | n/s | 2 | n/s | 2.8 | 1.5E-05 |
| **SCO5080** | putative hydrolase | actVA5 | ACT production | 1.9 | n/s | 2 | n/s | 3.6 | 1.4E-06 | 1.3 | n/s |
| **SCO5081** | | actVA6 | ACT production | 2.2 | n/s | 2 | n/s | 3.2 | 1.0E-02 | 1.5 | n/s |
| **SCO5082** | putative transcriptional regulatory protein | actII-1 | ACT production | 0.9 | n/s | 0.8 | n/s | 0.4 | 6.5E-06 | 2.3 | 9.3E-03 |
| **SCO5083** | putative actinorhodin transporter | actII-2 | ACT production | 0.7 | n/s | 1 | n/s | 1.4 | n/s | 0.5 | n/s |
| **SCO5084** | putative membrane protein | actII-3 | ACT production | 0.7 | n/s | 1.4 | n/s | 1.5 | n/s | 0.5 | n/s |
| **SCO5085** | actinorhodin cluster activator protein | actII-4 | ACT production | 2.3 | 4.5E-02 | 1.7 | n/s | 1.1 | n/s | 5.6 | 1.5E-18 |
| **SCO5086** | ketoacylreductase | actIII | ACT production | 1.4 | n/s | 2.8 | n/s | 4.6 | 1.4E-05 | 1.1 | n/s |
| **SCO5087** | actinorhodin polyketide beta-ketoacyl synthase alpha subunit | actIORF1 | ACT production | 0.8 | n/s | 3.1 | n/s | 4.4 | 5.2E-07 | 0.7 | n/s |
| **SCO5088** | actinorhodin polyketide beta-ketoacyl synthase beta subunit | actIORF2 | ACT production | 0.9 | n/s | 3.1 | n/s | 5 | 9.0E-07 | 0.5 | n/s |
| **SCO5089** | actinorhodin polyketide synthase acyl carrier protein | actIORF3 | ACT production | 0.5 | n/s | 2.3 | n/s | 4.5 | 1.9E-02 | 1.3 | n/s |

| | | | | | | | | | | | |
|---|---|---|---|---|---|---|---|---|---|---|---|
| **SCO5090** | actinorhodin polyketide synthase bifunctional cyclase/dehydratase | actVII | ACT production | 1.1 | n/s | 2.1 | n/s | 4.3 | 2.2E-05 | 0.6 | n/s |
| **SCO5091** | cyclase | actIV | ACT production | 1.4 | n/s | 1.8 | n/s | 5.8 | 1.4E-08 | 0.4 | 5.1E-03 |
| **SCO5092** | actinorhodin polyketide putative dimerase | actVB | ACT production | 1.2 | n/s | 2 | n/s | 3.9 | 3.3E-03 | 0.7 | n/s |
| **SCO5877** | transcriptional regulator | redD | Undecylprodigiosin | 6.8 | 1.0E-15 | 2 | n/s | 3.9 | 1.3E-07 | 10.1 | 2.9E-24 |
| **SCO5878** | polyketide synthase | redX | Undecylprodigiosin | 6.7 | 1.5E-17 | 2.9 | 9.1E-09 | 5.2 | 1.4E-13 | 6.9 | 1.3E-23 |
| **SCO5879** | acyl-coa dehydrogenase | redW | Undecylprodigiosin | 3 | 1.5E-04 | 2.9 | 1.3E-04 | 2.7 | 4.7E-03 | 8.9 | 2.8E-22 |
| **SCO5880** | | redY | Undecylprodigiosin | 1.9 | n/s | 3.2 | 9.3E-03 | 1.2 | n/s | 12.2 | 3.1E-18 |
| **SCO5881** | response regulator | redZ | Undecylprodigiosin | 10.4 | 9.8E-28 | 0.5 | n/s | 3.8 | 3.7E-15 | 2.3 | 1.4E-06 |
| **SCO5882** | | redV | Undecylprodigiosin | 1.3 | n/s | 8.7 | 1.4E-03 | 10.8 | 1.7E-10 | 0.6 | n/s |
| **SCO5883** | hypothetical protein | redU | Undecylprodigiosin | 1.9 | n/s | 6.7 | 1.9E-02 | 10.7 | 1.9E-08 | 0.8 | n/s |
| **SCO5884** | hypothetical protein | | Undecylprodigiosin | 0.9 | n/s | 10.5 | 1.7E-04 | 4.6 | 1.5E-04 | 3.5 | 2.4E-08 |
| **SCO5885** | putative membrane protein | | Undecylprodigiosin | 0.7 | n/s | 12.9 | 4.3E-03 | 3.4 | n/s | 6.4 | 4.5E-11 |
| **SCO5886** | 3-oxoacyl-[acyl-carrier-protein] synthase II | redR | Undecylprodigiosin | 1.7 | n/s | 11.3 | 7.7E-11 | 6.6 | 1.1E-09 | 4.5 | 1.8E-16 |
| **SCO5887** | acyl carrier protein | redQ | Undecylprodigiosin | 1.2 | n/s | 9.3 | 7.3E-05 | 3.7 | n/s | 14.7 | 1.2E-16 |
| **SCO5888** | 3-oxoacyl-[acyl-carrier-protein] synthase | redP | Undecylprodigiosin | 1.1 | n/s | 30.7 | 1.6E-09 | 5.4 | 1.1E-07 | 4 | 7.9E-14 |
| **SCO5889** | hypothetical protein | redO | Undecylprodigiosin | 0.3 | n/s | 65.6 | 1.3E-06 | 6.5 | 5.4E-03 | 7.8 | 1.8E-11 |
| **SCO5890** | putative 8-amino-7-oxononanoate synthase | redN | Undecylprodigiosin | 2.8 | 5.2E-03 | 0.9 | n/s | 3 | 8.7E-05 | 2.8 | 6.1E-08 |
| **SCO5891** | putative peptide synthase | redM | Undecylprodigiosin | 0.8 | n/s | 28.5 | 9.3E-07 | 9.3 | 2.8E-09 | 1.9 | n/s |
| **SCO5892** | polyketide synthase | redL | Undecylprodigiosin | 2 | n/s | 18.4 | 6.6E-09 | 9 | 2.2E-17 | 0.1 | 1.5E-17 |
| **SCO5893** | oxidoreductase | redK | Undecylprodigiosin | 1.6 | n/s | 6.2 | 3.9E-02 | 6 | 5.2E-06 | 1.2 | n/s |
| **SCO5894** | thioesterase | redJ | Undecylprodigiosin | 1.9 | n/s | 4.9 | n/s | 7.8 | 1.1E-07 | 0.5 | n/s |
| **SCO5895** | putative methyltransferase | redI | Undecylprodigiosin | 1.6 | n/s | 5.9 | 4.1E-03 | 6.7 | 1.3E-06 | 1.2 | n/s |
| **SCO5896** | phosphoenolpyruvate-utilising enzyme | redH | Undecylprodigiosin | 1.4 | n/s | 4.1 | 1.1E-03 | 6.3 | 4.7E-10 | 1.1 | n/s |
| **SCO5897** | putative oxidase | redG | Undecylprodigiosin | 1.8 | n/s | 1.7 | n/s | 2.5 | n/s | 0.5 | n/s |
| **SCO5898** | probable membrane protein | redF | Undecylprodigiosin | 1 | n/s | 0.9 | n/s | 0.7 | n/s | 1.9 | n/s |
| | | | | | | | | | | | |
| **Nitrate Reductase** | | | | | | | | | | | |
| ***SCO0216*** | putative nitrate reductase alpha chain | *narG2* | Nitrate reductase | 0.5 | n/s | 1.4 | n/s | 2.6 | 1.5E-03 | 0.4 | n/s |
| ***SCO0217*** | putative nitrate reductase beta chain | *narH2* | Nitrate reductase | 0.6 | n/s | 1.3 | n/s | 3.1 | 3.2E-03 | 0.2 | 1.9E-05 |
| ***SCO0218*** | putative nitrate reductase delta chain | *narJ2* | Nitrate reductase | 0.7 | n/s | 0.9 | n/s | 8.1 | 1.2E-07 | 0.1 | 1.4E-07 |
| ***SCO0219*** | putative nitrate reductase gamma chain | *narI2* | Nitrate reductase | 0.6 | n/s | 1 | n/s | 4.5 | 3.9E-04 | 0.1 | 1.3E-05 |
| ***SCO4947*** | putative nitrate reductase alpha chain | *narG3* | Nitrate reductase | 1.5 | n/s | 2.1 | n/s | 3.7 | 7.3E-11 | 1.1 | n/s |
| ***SCO4948*** | putative nitrate reductase beta chain | *narH3* | Nitrate reductase | 1.7 | n/s | 2.1 | n/s | 5.1 | 3.5E-09 | 0.7 | n/s |
| ***SCO4949*** | putative nitrate reductase delta chain | *narJ3* | Nitrate reductase | 1.8 | n/s | 2.1 | n/s | 5.8 | 3.8E-07 | 0.4 | 2.8E-02 |
| ***SCO4950*** | putative nitrate reductase gamma chain | *narI3* | Nitrate reductase | 1.6 | n/s | 2.6 | n/s | 5.1 | 1.0E-05 | 0.5 | n/s |
| ***SCO6532*** | putative nitrate reductase gamma chain | *narI* | Nitrate reductase | 0.4 | n/s | 0.9 | n/s | 2.4 | n/s | 0.2 | 3.5E-03 |
| ***SCO6534*** | putative nitrate reductase beta chain | *narH* | Nitrate reductase | 3 | n/s | 0.4 | n/s | 3.9 | 4.3E-04 | 0.2 | 6.6E-04 |
| ***SCO6535*** | putative nitrate reductase alpha chain | *narG* | Nitrate reductase | 1 | n/s | 1 | n/s | 2.6 | 4.8E-04 | 0.2 | 3.9E-09 |

| | | | | | | | | | | | |
|---|---|---|---|---|---|---|---|---|---|---|---|
| **Membrane Transport** | | | | | | | | | | | |
| **SCO2472** | putative integral membrane protein | | Membrane transport | 98.4 | 3.9E-18 | 0.9 | n/s | 16.2 | 4.5E-13 | 2.7 | 2.9E-07 |
| **SCO3824** | putative ABC-transporter ATP-binding component | | Membrane transport | 24.6 | 1.9E-20 | 0.7 | n/s | 7.3 | 6.1E-13 | 1.7 | n/s |
| **SCO3910** | putative membrane protein | | Membrane transport | 8.9 | 1.3E-20 | 0.5 | n/s | 3.5 | 6.5E-09 | 1.1 | n/s |
| **SCO4359** | putative ABC transport system ATP-binding protein | | Membrane transport | 15.2 | 1.3E-21 | 1 | n/s | 4.2 | 1.1E-10 | 2.7 | 8.5E-08 |
| | | | | | | | | | | | |
| **Peptidoglycan** | | | | | | | | | | | |
| **SCO1326** | undecaprenyl pyrophosphate phosphatase | uppP | Peptidoglycan biosynthesis | 0.1 | 1.5E-13 | 1 | n/s | 0.1 | 6.2E-18 | 0.7 | n/s |
| **SCO1875** | putative secreted penicillin binding protein | | Class B PBP | 31.7 | 5.6E-31 | 1 | n/s | 5.8 | 2.2E-20 | 3.2 | 1.2E-16 |
| **SCO2084** | UDP-N-acetylglucosamine-N-acetylmuramyl-(pentapeptide) pyrophosphoryl-undecaprenol N-acetylglucosamine transferase | murG | Peptidoglycan biosynthesis | 0.4 | 3.8E-04 | 0.9 | n/s | 0.2 | 6.2E-22 | 2.5 | 4.4E-06 |
| **SCO2086** | UDP-N-acetylmuramoylalanine-D-glutamate ligase | murD | Peptidoglycan biosynthesis | 0.3 | 3.1E-13 | 1.2 | n/s | 0.1 | 1.3E-24 | 2.1 | n/s |
| **SCO2087** | phospho-N-acetylmuramoyl-pentapeptide-transferase | murX | Peptidoglycan biosynthesis | 0.2 | 4.6E-19 | 1.1 | n/s | 0.1 | 3.9E-28 | 2.8 | 4.4E-09 |
| **SCO2088** | UDP-N-acetylmuramoylalanyl-D-glutamyl- 2,6-diaminopimelate- D-alanyl-alanyl ligase | murF | Peptidoglycan biosynthesis | 0.3 | 6.5E-17 | 1.2 | n/s | 0.1 | 7.2E-28 | 2.3 | 3.2E-04 |
| **SCO2089** | UDP-N-acetylmuramoylalanyl-D-glutamate- 2,6-diaminopimelate ligase | murE | Peptidoglycan biosynthesis | 0.3 | 1.0E-12 | 1.2 | n/s | 0.2 | 1.4E-23 | 2.4 | 1.8E-04 |
| **SCO2090** | cell division protein | ftsI | Class B PBP | 0.5 | n/s | 1.2 | n/s | 0.4 | 4.3E-08 | 1.3 | n/s |
| **SCO2504** | glycyl-tRNA synthetase | glyS | Peptidoglycan biosynthesis | 0.4 | 8.0E-11 | 0.8 | n/s | 0.1 | 8.2E-29 | 2.7 | 7.3E-12 |
| **SCO2509** | undecaprenyl phosphate synthetase | uppS | Peptidoglycan biosynthesis | 0.5 | n/s | 0.9 | n/s | 0.2 | 3.0E-13 | 1.5 | n/s |
| **SCO2608** | penicillin binding protein | pbp2 | Class B PBP | 1.6 | n/s | 0.8 | n/s | 0.9 | n/s | 0.1 | 2.9E-16 |
| **SCO2897** | probable secreted penicillin-binding protein | | Class A PBP | 10.7 | 1.0E-26 | 0.8 | n/s | 2.8 | 1.2E-09 | 3.8 | 1.2E-18 |
| **SCO2949** | UDP-N-acetylglucosamine transferase | murA | Peptidoglycan biosynthesis | 1.6 | n/s | 0.9 | n/s | 0.4 | 1.8E-03 | 3.6 | 6.6E-15 |
| **SCO3156** | putative penicillin-binding protein | | Class B PBP | 4.1 | 2.7E-19 | 0.9 | n/s | 1.4 | n/s | 4 | 1.2E-19 |
| **SCO3157** | putative penicillin-binding protein | | Class B PBP | 5.5 | 1.2E-23 | 1.1 | n/s | 1.7 | n/s | 3.5 | 2.2E-18 |
| **SCO3580** | putative transpeptidase | | Class A PBP | 0.7 | n/s | 1 | n/s | 0.2 | 1.6E-20 | 3.1 | 1.0E-12 |
| **SCO3771** | putative secreted penicillin binding protein | | Class B PBP | 22.5 | 2.2E-16 | 0.9 | n/s | 7.8 | 1.4E-14 | 0.9 | n/s |
| **SCO3847** | putative secreted penicillin-binding protein | | Class B PBP | 2.5 | 1.1E-09 | 0.6 | n/s | 0.7 | n/s | 2.5 | 1.4E-10 |
| **SCO3894** | Putative peptidoglycan lipid II flippase | murJ | Peptidoglycan biosynthesis | 0.8 | n/s | 0.9 | n/s | 0.4 | 2.1E-10 | 2 | n/s |
| **SCO3901** | putative secreted penicillin-binding protein | | Class A PBP | 0.6 | n/s | 0.9 | n/s | 0.2 | 7.2E-23 | 2.4 | 2.4E-07 |
| **SCO3904** | Lipid II:glycine glycyltransferase | femX | Peptidoglycan biosynthesis | 0.4 | 7.0E-06 | 1 | n/s | 0.2 | 1.1E-23 | 3 | 4.5E-12 |
| **SCO4013** | putative secreted penicillin-binding protein | | Class B PBP | 0.8 | n/s | 1.2 | n/s | 0.2 | 1.0E-18 | 4.3 | 2.7E-18 |
| **SCO4643** | UDP-N-acetylenoylpyruvoylglucosamine reductase | murB | Peptidoglycan biosynthesis | 0.8 | n/s | 1.4 | n/s | 0.5 | n/s | 2.1 | n/s |
| **SCO5039** | putative penicillin-binding protein | | Class A PBP | 2.3 | 3.3E-05 | 1.2 | n/s | 0.5 | n/s | 3.9 | 2.9E-18 |
| **SCO5301** | putative secreted penicillin-binding protein | | Class B PBP | 1.5 | n/s | 1.3 | n/s | 7.4 | 9.6E-10 | 0.1 | 7.6E-09 |
| **SCO5560** | D-alanine-D-alanine ligase | ddlA | Peptidoglycan biosynthesis | 0.7 | n/s | 1.1 | n/s | 0.3 | 7.5E-16 | 3.2 | 1.2E-12 |
| **SCO6060** | UDP-N-acetylmuramoyl-L-alanine ligase | murC | Peptidoglycan biosynthesis | 0.7 | n/s | 0.9 | n/s | 0.3 | 2.1E-14 | 2.4 | 3.0E-05 |
| **SCO0830** | putative penicillin binding protein | | Class C PBP | 12.7 | 4.47E-09 | 1.3 | n/s | 7.3 | 8.04E-08 | 0.8 | n/s |
| **SCO4847** | putative D-alanyl-D-alanine carboxypeptidase | | Class C PBP | 2.6 | 2.83E-10 | 0.8 | n/s | 0.8 | n/s | 2.5 | 1.05E-08 |

| | | | | | | | | | | | |
|---|---|---|---|---|---|---|---|---|---|---|---|
| **SCO7050** | putative D-alanyl-D-alanine carboxypeptidase | | Class C PBP | 1.2 | n/s | 1.2 | n/s | 4.8 | 1.79E-06 | 0.1 | 6.04E-07 |
| **SCO3811** | putative D-alanine-D-alanine carboxypeptidase | dacA | Class C PBP | 0.7 | n/s | 0.5 | n/s | 0.3 | 2.21E-13 | 1.2 | n/s |
| | | | | | | | | | | | |
| **Membrane Homeostasis** | | | | | | | | | | | |
| **SCO0920** | Lyso-ornithine lipid acyltransferase | olsA | Ornithine lipid biosynthesis | 0.7 | n/s | 3.6 | n/s | 8.9 | 1.4E-08 | 0 | 8.8E-09 |
| **SCO0921** | Ornithine-acyl[acyl carrier protein] N-acyltransferase | olsB | Ornithine lipid biosynthesis | 1.1 | n/s | 0.5 | n/s | 7.3 | 5.5E-07 | 0.1 | 4.8E-07 |
| **SCO1047** | putative integral membrane protein | | Phosphatidic acid phosphatase enzyme | 0.7 | n/s | 1.7 | n/s | 6.5 | 1.9E-07 | 0.0 | 3.4E-08 |
| **SCO1090** | putative phosphodiesterase | ugpQ1 | Glycerophosphoryl diester phosphodiesterase | 4.3 | 5.9E-08 | 0.7 | n/s | 1.5 | n/s | 2.1 | n/s |
| **SCO1389** | Putative CDP-diacylglycerol--glycerol-3-phosphate 3-phosphatidyl-transferase | cls-II | Cardiolipin biosynthesis | 0.6 | n/s | 0.6 | n/s | 0.2 | 2.2E-14 | 1.9 | n/s |
| **SCO1419** | putative phosphodiesterase | ugpQ2 | Glycerophosphoryl diester phosphodiesterase | 1.4 | n/s | 1.2 | n/s | 0.8 | n/s | 1.4 | n/s |
| **SCO1423** | Dolichol-phosphate mannosyltransferase | ppm1 | Polyprenylphosphate mannose biosynthesis | 0.6 | n/s | 1.2 | n/s | 0.2 | 6.3E-20 | 2.7 | 3.0E-07 |
| **SCO1525** | Phosphatidylinositol alpha-mannosyltransferase | pimA | Phosphatidylinositol biosynthesis | 0.8 | n/s | 1 | n/s | 0.4 | 1.1E-06 | 1.6 | n/s |
| **SCO1526** | putative acyltransferase | acylT | Phosphatidylinositol biosynthesis | 0.4 | 6.3E-04 | 1 | n/s | 0.3 | 4.6E-13 | 1.3 | n/s |
| **SCO1527** | CDP-diacylglycerol inositol 3-phosphatidyltransferase | pisA | Phosphatidylinositol biosynthesis | 0.4 | 1.1E-05 | 0.9 | n/s | 0.2 | 1.9E-16 | 1.9 | n/s |
| **SCO1565** | putative phosphodiesterase | glpQ1 | Glycerophosphoryl diester phosphodiesterase | 0.5 | n/s | 0.9 | n/s | 1.2 | n/s | 0.5 | n/s |
| **SCO1658** | Regulatory gene | gylR | Transcriptional regulation of the glycerol operon | 0.5 | n/s | 1.9 | n/s | 0.6 | n/s | 3.4 | 1.7E-09 |
| **SCO1659** | putative glycerol uptake facilitator protein | glpF | Glycerol metabolic process | 0.8 | n/s | 3.5 | 6.2E-06 | 1.6 | n/s | 17.7 | 8.3E-23 |
| **SCO1660** | putative glycerol kinase | glpK | Glycerol metabolic process | 0.7 | n/s | 2.7 | 5.2E-04 | 0.6 | n/s | 15.3 | 2.3E-24 |
| **SCO1661** | putative glycerol-3-phosphate dehydrogenase | glpD | Glycerol metabolic process | 0.9 | n/s | 2.7 | 2.6E-04 | 1 | n/s | 8.9 | 4.3E-23 |
| **SCO1814** | putative enoyl-(acyl-carrier-protein) reductase | fabI | Fatty acid biosynthesis | 0.3 | 2.8E-16 | 0.6 | n/s | 0.1 | 5.6E-29 | 2.4 | 5.6E-07 |
| **SCO1815** | probable 3-oxacyl-(acyl-carrier-protein) reductase | fabG | Fatty acid biosynthesis | 0.4 | 4.0E-12 | 0.6 | n/s | 0.1 | 3.5E-29 | 2.4 | 2.7E-07 |
| **SCO1968** | putative phosphodiesterase | glpQ2 | Glycerophosphoryl diester phosphodiesterase | 1.7 | n/s | 0.8 | n/s | 0.7 | n/s | 2.3 | 1.9E-02 |
| **SCO2132** | putative glycosyl transferase | pimB | Phosphatidylinositol biosynthesis | 0.6 | n/s | 1.2 | n/s | 0.5 | n/s | 1.4 | n/s |
| **SCO2255** | Putative membrane protein | | Endonuclease/Exonuclease/Phosphatase | 22.7 | 8.5E-30 | 0.7 | n/s | 7.1 | 5.5E-21 | 2.4 | 9.4E-08 |
| **SCO2335** | putative integral membrane protein | | Phosphatidic acid phosphatase enzyme | 3.4 | 8.2E-16 | 1.1 | n/s | 1.3 | n/s | 2.5 | 5.1E-08 |
| **SCO2387** | malonyl CoA:acyl carrier protein malonyltransferase | fabD | Fatty acid biosynthesis | 0.8 | n/s | 0.7 | n/s | 0.2 | 1.0E-25 | 2.7 | 2.2E-12 |
| **SCO2388** | 3-oxoacyl-[acyl-carrier-protein] synthase III | fabH | Fatty acid biosynthesis | 0.5 | n/s | 0.4 | n/s | 0.1 | 2.9E-29 | 2 | n/s |
| **SCO2390** | 3-oxoacyl-[acyl-carrier-protein] synthase II | fabF | Fatty acid biosynthesis | 0.7 | n/s | 0.7 | n/s | 0.2 | 3.1E-26 | 2.2 | 2.8E-05 |
| **SCO2807** | putative integral membrane protein | | Phosphatidic acid phosphatase enzyme | 6.6 | 1.9E-24 | 0.7 | n/s | 1.7 | n/s | 2.2 | 3.1E-03 |
| **SCO2892** | putative secreted protein | | Hydrolase | 6.9 | 3.2E-25 | 0.9 | n/s | 1.7 | n/s | 3.0 | 6.6E-15 |
| **SCO3154** | Dolichyl-phosphate-mannose-protein mannosyltransferase | pmt | Polyprenylphosphate mannose biosynthesis | 0.7 | n/s | 0.9 | n/s | 0.7 | n/s | 1.1 | n/s |
| **SCO3222** | putative secreted protein | | Hydrolase | 14.3 | 1.1E-11 | 5.0 | 9.5E-13 | 6.8 | 6.1E-08 | 10.3 | 3.8E-23 |
| **SCO3286** | Putative membrane protein YqiK | | Flottilin | 60.7 | 4.8E-34 | 0.9 | n/s | 46.5 | 2.6E-29 | 3.3 | 4.7E-20 |

| | | | | | | | | | | | |
|---|---|---|---|---|---|---|---|---|---|---|---|
| **SCO3592** | Putative membrane protein | vanJ | Endonuclease/Exonuclease/Phosphatase | 1.2 | n/s | 2.1 | n/s | 0.4 | n/s | 2.2 | n/s |
| **SCO3607** | Flotillin | | Flottilin | 63.1 | 2.6E-34 | 1.1 | n/s | 7 | 2.8E-13 | 3.1 | 3.4E-19 |
| **SCO3725** | putative conserved membrane protein | | Phosphatidic acid phosphatase enzyme | 0.8 | n/s | 1 | n/s | 0.4 | 3.4E-05 | 1.2 | n/s |
| **SCO3976** | putative phosphodiesterase | ugpQ3 | Glycerophosphoryl diester phosphodiesterase | 0.5 | n/s | 0.8 | n/s | 0.9 | n/s | 0.5 | n/s |
| **SCO4133** | putative integral membrane protein | | Phosphatidic acid phosphatase enzyme | 2.7 | 2.0E-11 | 0.6 | n/s | 0.8 | n/s | 2.1 | n/s |
| **SCO4636** | predicted 2-enoyl-CoA hydratase 2 | | Fatty acid biosynthesis | 0.7 | n/s | 0.8 | n/s | 0.2 | 2.0E-19 | 2.3 | 2.7E-03 |
| **SCO4637** | Acyl dehydratase | | Fatty acid biosynthesis | 0.7 | n/s | 0.9 | n/s | 0.2 | 5.2E-17 | 2.4 | 5.1E-04 |
| **SCO4843** | putative integral membrane protein | | Phosphatidic acid phosphatase enzyme | 2.2 | n/s | 0.9 | n/s | 14 | 1.7E-08 | 0 | 2.0E-07 |
| **SCO5628** | Phosphatidate cytidylyltransferase | cdsA | CDP-diacylglycerate biosynthesis | 0.6 | n/s | 1.2 | n/s | 0.2 | 1.1E-24 | 3.1 | 1.1E-14 |
| **SCO5661** | putative phosphodiesterase | ugpQ4 | Glycerophosphoryl diester phosphodiesterase | 0.8 | n/s | 1.3 | n/s | 0.6 | n/s | 1.4 | n/s |
| **SCO5753** | CDP-diacylglycerol--glycerol-3-phosphate 3-phosphatidyltransferase | pgsA | Cardiolipin biosynthesis | 0.6 | n/s | 1 | n/s | 0.3 | 3.0E-14 | 1.9 | n/s |
| **SCO6355** | putative integral membrane protein | | Phosphatidic acid phosphatase enzyme | 3.6 | 2.2E-13 | 0.8 | n/s | 1.2 | n/s | 2.7 | 7.3E-08 |
| **SCO6356** | putative integral membrane protein | | Phosphatidic acid phosphatase enzyme | 3.8 | 1.4E-16 | 0.8 | n/s | 1.1 | n/s | 3.2 | 6.4E-14 |
| **SCO6378** | putative membrane protein | | Phosphatidic acid phosphatase enzyme | 10.4 | 5.0E-13 | 0.6 | n/s | 2.8 | 1.3E-02 | 3.5 | 2.6E-09 |
| **SCO6428** | putative secreted protein | | Phosphatidic acid phosphatase enzyme | 1.1 | n/s | 1.1 | n/s | 7.2 | 2.2E-14 | 0.1 | 7.9E-14 |
| **SCO6467** | putative phosphatidylserine synthase | pssA | Phosphatidylethanolamine biosynthesis | 0.5 | n/s | 0.8 | n/s | 0.1 | 1.7E-25 | 3 | 1.7E-12 |
| **SCO6468** | Phosphatidylserine decarboxylase | psd | Phosphatidylethanolamine biosynthesis | 0.5 | n/s | 0.9 | n/s | 0.2 | 8.2E-26 | 3.2 | 3.1E-15 |
| **SCO6511** | Undecaprenyl pyrophosphatase | | Phosphatidic acid phosphatase enzyme | 10.1 | 1.2E-23 | 0.8 | n/s | 2.6 | 6.3E-05 | 3.3 | 4.7E-14 |
| **SCO6759** | putative phytoene synthase | hpnC | Hopanoid biosynthesis | 0.8 | n/s | 1 | n/s | 0.3 | 4.4E-12 | 2 | n/s |
| **SCO6760** | putative phytoene synthase | hpnD | Hopanoid biosynthesis | 0.8 | n/s | 1 | n/s | 0.3 | 2.7E-11 | 6.3 | 2.4E-22 |
| **SCO6761** | Uncharacterized protein | | Hopanoid biosynthesis | 0.7 | n/s | 1 | n/s | 0.2 | 1.8E-08 | 2.4 | n/s |
| **SCO6762** | putative phytoene dehydrogenase | hpnE | Hopanoid biosynthesis | 0.9 | n/s | 1 | n/s | 0.4 | 1.1E-06 | 2.3 | 5.9E-03 |
| **SCO6763** | putative polyprenyl synthatase | fpps | Hopanoid biosynthesis | 0.7 | n/s | 1 | n/s | 0.3 | 2.5E-11 | 2 | n/s |
| **SCO6764** | putative squalene-hopene cyclase | shc | Hopanoid biosynthesis | 1 | n/s | 0.8 | n/s | 0.5 | n/s | 1.6 | n/s |
| **SCO6765** | putative lipoprotein | hpnG | Hopanoid biosynthesis | 1.1 | n/s | 0.8 | n/s | 1.1 | n/s | 0.9 | n/s |
| **SCO6766** | Hopanoid biosynthesis associated radical SAM protein HpnH | hpnH | Hopanoid biosynthesis | 1 | n/s | 0.8 | n/s | 0.4 | 5.7E-03 | 2.5 | 5.7E-04 |
| **SCO6767** | GcpE protein homolog, conserved hypothetical protein | ispG | Hopanoid biosynthesis | 1.3 | n/s | 1.1 | n/s | 0.6 | n/s | 2.2 | n/s |
| **SCO6768** | probable transketolase | dxps1 | Hopanoid biosynthesis | 1.5 | n/s | 1.1 | n/s | 1.2 | n/s | 1.7 | n/s |
| **SCO6769** | probable aminotransferase | hpnO | Hopanoid biosynthesis | 1.3 | n/s | 0.9 | n/s | 0.8 | n/s | 1.9 | n/s |
| **SCO6770** | putative DNA-binding protein | | Hopanoid biosynthesis | 0.6 | n/s | 1.1 | n/s | 0.3 | 4.1E-11 | 2.3 | 2.2E-03 |
| **SCO7017** | putative membrane protein | | Endonuclease/Exonuclease/Phosphatase | 34.9 | 5.1E-23 | 0.8 | n/s | 7.4 | 2.1E-15 | 2.9 | 7.3E-11 |
| **SCO7205** | putative hydrolase | | Hydrolase | 14.4 | 4.2E-23 | 0.9 | n/s | 5.7 | 2.3E-14 | 3.5 | 2.4E-15 |
| **SCO7587** | putative integral membrane protein | | Phosphatidic acid phosphatase enzyme | 0.4 | 2.7E-04 | 1 | n/s | 0.3 | 4.1E-13 | 1.3 | n/s |