Strathclyde Institute of Pharmacy and Biomedical Sciences

Characterising the biosynthesis and mechanism of action of aurodox from *Streptomyces goldiniensis*

Rebecca Elizabeth McHugh

Thesis presented in fulfilment of the requirement for the degree of Doctor of Philosophy

2020

Declaration

This thesis is the result of the author's original research. It has been composed by the author and has not been previously submitted for examination which has led to the award of a degree. The copyright of this thesis belongs to the author under the terms of the United Kingdom Copyright Acts as qualified by University of Strathclyde Regulation 3.50. Due acknowledgement must always be made of the use of any material contained in, or derived from, this thesis.

The figures used in this thesis are a result of work carried out by myself with the exception of *Figure 3-3B* which was generated in collaboration with Josephine Giard whilst she was working under my supervision in the laboratory and *Figure 6-7A* which was generated by Tom Parker. The gel electrophoresis images in *Figure 5-2(III)* and *Figure 5-3* were provided by Bio S&T (Canada).

Signed:

Date:

Acknowledgments.

Over the last four years, I have been fortunate enough to start my research career in my home city of Glasgow. I have worked with two outstanding research groups, whilst also being surrounded by my family and friends. I will always be grateful to the Glasgow MVLS DTP, the University of Glasgow and the University of Strathclyde for this experience which I will cherish for the rest of my life. I must express my sincere gratitude to each and every member of my close circle of family, friends and colleagues who have helped me to stay focussed and driven while also providing the escapes that are so vital for any PhD student. Of course, there are many individuals who deserve to be mentioned, so the following section is rather comprehensive, but I feel it is required to express my gratitude.

First and foremost, my primary supervisor Professor Paul Hoskisson. Paul has provided me with much guidance whilst giving me scope to drive my own project and explore my own ideas. As a group leader, Paul consistently ensures each of us has a positive PhD experience, and he has been successful in doing this for me. Paul values the role of the wider research community, which has introduced me to, to my great benefit, and for these reasons I will always be grateful. I truly hope that we can continue to bounce ideas around, long into the future. I would also like to thank my secondary supervisor Professor Andrew Roe, who has always made me feel like an integral part of the group, despite being a 'visitor'. I would also like to thank the academics of the Microgroup, namely, Paul Herron, Nick Tucker, Kate Duncan, Iain Hunter, Arnaud Javelle, Morgan Feeney, Gail McConnell and Leighton Pritchard. All of the above have influenced my work and their feedback has been of great value. I look forward to the day when this pandemic is over and we can sit in Todd's over a few (or many) watered down Gin and Tonics, and I hope we can continue to collaborate in the future.

I would like to thank each and every member of the Microgroup, past and present for their continued help, advice and support. Their training and encouragement changed me from the shy student who was petrified of messing up to the experienced person who can help with the very same problems. So to John, Kirsty, Sarah, Jana, Stuart, Walid, Leena, Gillian, Alejandro, Silvia, Gaetan, Anna, Josi, Liam, Iain, Jordan, Molly, Emily, Bea, Ainsley, Talal, Charlie, Eilidh, Lis, Tiago, David, Jonny, Elmira, Gordon, Adriana, Ally, Laia, Parra, Darren and Lily I truly could not have gotten through this it without you all, thank you for making every day a joy. To every Masters or

Undergraduate student, I have learned something from each of you… Special thanks must go to Celtic Women's number 9 Josi Giard who was a joy to share the lab and helped me carry out some of the early fermentation studies in this thesis. Thank you to my colleagues in Bacteriology at the University of Glasgow, Nicky, James, Natasha, Liyana, James, Jeni, Cat, Khedidja, Willybob, Claire and members of the Walker group.

I have to specifically thank some individuals who have become so special to me over the years. When I started this journey I really did not anticipate that I would form such close and long-lasting friendships with my colleagues, but the conference trips have been bonding experiences, and the long days in the lab and office we have shared have never been dull with you all around. Ainsley Beaton, thank you for being the most loyal and supportive friend. You have been there since day one, ready to have a snack time whenever it was required. I honestly don't know how I would have coped without you over the last year. We have laughed and cried and prosecco'd and just, thank you for everything. Liam Rooney where do I start? I end up with actual pains in my sides because of how you much you make me laugh. I am so glad that our paths finally crossed at this point in our life. I knew from the first night of MicroSoc 2017 that we would become good friends and you have not disappointed (especially in the mesolab). Lis, even if you don't know it, you are my lab mother and I'm so grateful for your thoughtful acts that cheer us all up, I can always count on you for advice. You are the Cloning Queen, thanks for all the reagents I ran out of.

Thank you to my friends, thank you for taking me away from work when you knew when I needed to have a break. Denise, your always there, with Roisin Jenna aka Big Mix, what a laugh we have had. Thanks to Annie for carpool, Jamie for weekends at the loch, Amy for the laughter, Aimee for our London times, Louise for the best surprise ever with your beautiful baby and just being you. Eilidh thanks for wee messages. Thank you all for learning about *Streptomyces* so that I could talk about my science with you!

There are too many family members to mention by name. To my Aunties, Uncles, Cousins.. your constant interest in what I do has been the biggest motivation for me, thank you for reminding me of what's important. To Hannah, Emily and Laura, thanks for the holidays and I hope there are t'others in the future.

Thank you to Celtic Football Club for winning every single domestic trophy throughout my PhD. For the home games, the invincibles season, the European trips away (thank

you Kristina), the victory parades the Treble-Treble, and possibly 9 in a row and ANOTHER treble this year. The greatest years of Celtic's history have coincided with my PhD and have provided the best distraction possible when times were hard. Thank you. Hail Hail.

Finally, and most importantly, thank you to my Mum, Dad and Laura. Thank you for putting up with me being in education for the last 21 years (it can't have been easy) I promise to get a real job now and move out. Thank you for picking me up from the train station at all hours when the days in the lab turned in to nights and always showing an interest. Thank you for trying to understand what I do, even when you get it wrong I really appreciate you trying. Laura and Marcus, thanks for being around to have a good laugh with. Hopefully we all get our holiday together soon.

.

Abstract

Aurodox, a specialised metabolite from the soil bacterium *Streptomyces goldiniensis* was discovered in 1973 and was originally investigated for its antibacterial properties. However, aurodox has been recently identified from large-scale compound screens as an inhibitor of the Enteropathogenic *Escherichia coli* (EPEC) Type III Secretion System (T3SS). Therefore, to gain an understanding of its mechanism of action and to assess the suitability of this molecule for repurposing as an anti-virulence compound, a multidisciplinary approach to understanding aurodox was used. The biosynthesis of aurodox by *S. goldiniensis* was investigated through sequencing the whole genome of *S. goldiniensis* to enable the identification of the putative aurodox biosynthetic gene cluster (BGC). This BGC was then cloned and expressed in multiple heterologous hosts including *Streptomyces coelicolor* M1152, confirming that this BGC is responsible for aurodox production. In-depth analysis of the BGC supports a model of a polyketide synthase pathway involving a combination of both *cis* and *trans*-acyltransferases which synthesise the aurodox polyketide backbone before tailoring enzymes are recruited to form the final aurodox precursor, kirromycin, which is also produced by *S. collinus.* These studies suggest that the SAM-dependent *O*-methyltransferase AurM* is responsible for catalysing the conversion of kirromycin to aurodox in *S. goldiniensis* through the methylation of the pyridone moiety. Furthermore, multiple aurodox resistance genes at distinct loci have been identified and their role in aurodox resistance has been explored. The mechanism of action of aurodox has also been investigated. Whole transcriptome analysis, cell infection and GFP-reporter assays were used to demonstrate that aurodox transcriptionally downregulates the expression of the Locus of Enterocyte Effacement (LEE) pathogenicity island- which encodes for the T3SS, acting via its master regulator, Ler. We have also observed these effects across other enteric pathogens carrying a

homologous T3SS such as Enterohemorrhagic *Escherichia coli* (EHEC). Significantly, unlike traditional antibiotics, aurodox does not induce the production of shiga toxin.

Publications

Part of the Introduction to this thesis has been published in the following review article:

Pollock, J., Low, A., **McHugh, RE.,** Muwonge, A., Stevens, M., Corbishley, A. and Gally, D., 2020. Alternatives to antibiotics in a One Health context and the role genomics can play in reducing antimicrobial use. *Clinical Microbiology and Infection*,.

Doi: 10.1016/j.cmi.2020.02.028

The results of *Chapter 6* and some excerpts from the Introduction in this thesis have been published in the following research article:

McHugh, RE., O'Boyle, N., Connolly, J., Hoskisson, P. and Roe, A., 2019. Characterization of the Mode of Action of aurodox, a Type III Secretion System Inhibitor from *Streptomyces goldiniensis*. *Infection and Immunity*, 87(2). (Appendix C)

Doi: 10.1128/IAI.00595-18

List of Tables

List of Figures

Table of Contents

Chapter 1 : Introduction

1.1 Actinobacteria and the genus *Streptomyces*: An Overview

1.1.1 Introduction to *Streptomyces*.

The *Streptomycetaceae* is a family of Gram-positive bacteria within the largest bacterial phylum, the Actinobacteria*,* which was described by Waksman and Henrici in 1943 after the initial isolation of *Streptomyces griseus* (Waksman and Henrici, 1943; Schatz, Bugle and Waksman, 1944*).* The largest genus within the family *Streptomycetaceae* is the *Streptomyces,* consisting of around 600 validly described species which are found ubiquitously in terrestrial and aquatic environments (Porter, 1971). *Streptomyces* species are generally saprophytic, aerobic and morphologically diverse. They have evolved the capacity to produce a plethora of bioactive molecules that have antibacterial, antifungal, immunosuppressive, anti-helminthic, anti-cancer, herbicidal and antivirulence properties (Gomez-Escribano, Alt and Bibb, 2016). The ability to produce these molecules provides *Streptomyces* with a competitive advantage over competing soil organisms. As a result of this metabolic diversity, *Streptomyces* are extensively exploited in medicine and industry.

1.1.2 *Streptomyces* life cycle, morphology and physiology

S*treptomycetes* have the capacity to survive and disperse across a range of environments, which is facilitated by their complex life cycle. *Streptomyces* species are prolific spore formers, and the formation of these spores is tightly regulated in response to environmental factors (Chakraburtty and Bibb, 1997).

Initial germination in *Streptomyces* is led by the emergence of a single germ tube which develops into vegetative hyphae in response to various signals in the surrounding environment. In contrast to unicellular bacteria such as *Escherichia coli,* these hyphae grow via tip extension, achieving exponential growth through branching

(Nieminen et al 2013); During this process, two coiled-coil proteins (DivIVA and Scy) located at the pole of the cell, referred to as the polarisome, interact with the cell wall biosynthesis machinery to form new branch points. This process is regulated by a Serine/Threonine protein kinase, *asfK,* which activates DivIVA via phosphorylation to initiate branching. Therefore, colony formation is initiated under nutrient-rich conditions (Claessen *et al.*, 2003; Hempel *et al.*, 2012)

Following this, often as a result of nutrient limitation or stress, the formation of aerial hyphae is initiated. Firstly, a hydrophobic sheath formed of chaplin and rodlin proteins is formed over the polarisome, and when broken, hair like structures form (Elliot *et al.*, 2003). In addition, a surfactant lantipeptide, SapB is required to break surface tension and allow aerial hyphae to emerge (Kodani *et al.*, 2005; Willey and Gaskell, 2011)) At this point, the aerial hyphae compartments contain several copies of the entire *Streptomyces* species genome until they are divided by multiple septa into a long chain of unigenomic pre-spores (Ohnishi, Seo and Horinouchi, 2002) . The cell walls of these pre-spores proceed to thicken with peptidoglycan and form individual spores (Flärdh and Buttner, 2009). When disturbed, these spores are released into the environment, and hence, sporulation in *Streptomyces* facilitates not only survival in harsh conditions, but also dispersal (Du *et al.*, 2012). The observed morphological differentiation as a multicellular unit is distinctive among the bacteria (Yagüe *et al.*, 2016). Yet, this process has enabled the evolution of the genus as sporulation has facilitated endurance in unfavourable environments in addition to their spread geographically, which is a vital process in what is otherwise a non-motile organism (Flärdh and Buttner, 2009).

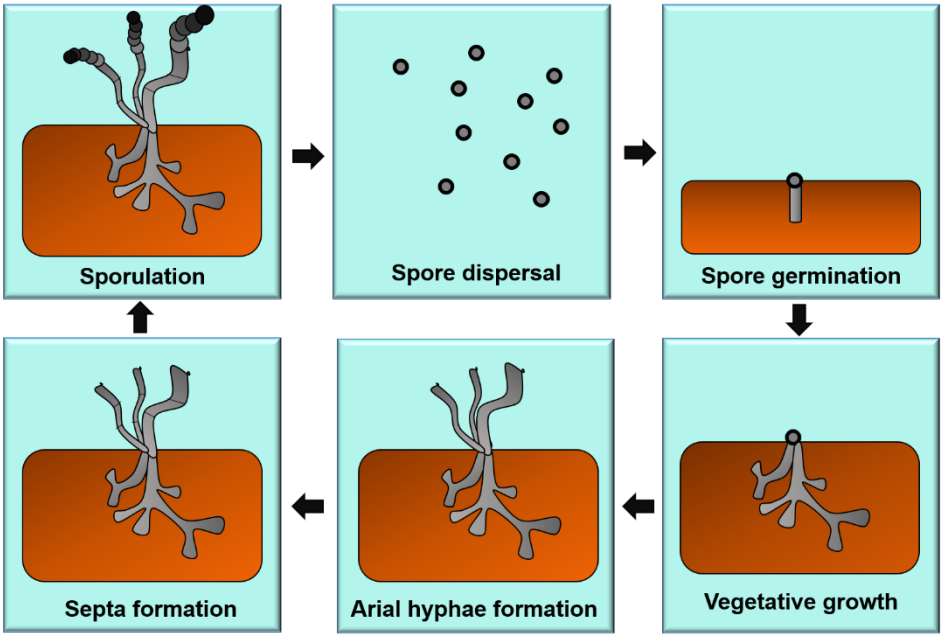


Figure 1-1: **Schematic diagram depicting life cycle of *Streptomyces*.** Adapted from Bush *et al,* 2015.

The life cycle of *Streptomyces* is tightly controlled by two classes of regulators: bald (*bld*) and white (*whi*) (Chater and Chandra, 2006). The *bld* genes generally regulate early events in *Streptomyces* differentiation, specifically the formation of aerial hyphae (Elliot *et al.*, 1998, 2001). The names of these proteins are derived from the phenotype which arises when they are deleted- a bald phenotype - with no aerial hyphae present (Champness, 1988; Salerno *et al.*, 2013). Conversely, the *whi* genes generally control the transition of spore chains to mature spores. The networks which control these regulators are complex and are often closely related to genes that sense and respond to environmental signals (Chater, 1972). However, more recently, the involvement of cyclic-di-AMP as a secondary messenger which can interact with both bald and white genes to control *Streptomyces* differentiation and development has been characterised (Bush *et al.*, 2015; McLean, Lo, *et al.*, 2019). Importantly, the role of the phosphodiesterase AtaC has been investigated as it has been shown to hydrolyse cyclic-di-GMP, whilst binding it to remove it from the cytoplasm and consequently allowing development to continue (Latoscha *et al.*, 2020)

1.1.3 Taxonomy and Phylogeny of *Streptomyces*

Streptomyces species are phylogenetically diverse (Labeda *et al.*, 2012). Throughout evolution across diverse environments, they have evolved great variation, whilst also maintaining the taxonomic species boundaries. These species boundaries are often difficult to define, with studies by Stackebrandt in 1994 demonstrating that the original definitions of these boundaries were largely artificial due to the absence of a definition involving reproduction, as with higher organisms (Stackebrandt and Goebel, 1994). In *Streptomyces,* there is evidence of extensive horizontal gene transfer events resulting in complex taxonomic issues (Tidjani *et al.*, 2019). The work of Stackebrandt *et al*, endorsed a method for phylogenetic studies with higher resolution than morphological observation in Actinobacteria, building of the work of Woese and the

use of 16S rRNA gene (Woese, Blanz and Hahn, 1984). This allowed for the initial characterisation of species based on the sequences of their 16S rRNA gene, a conserved housekeeping gene with species-specific polymorphisms. In the years since this study, 16S rRNA gene sequencing has been the dominant method of initial phylogenetic identification in the bacteria (Tringe and Hugenholtz, 2008). Using the 16S rRNA gene sequencing approach, approximately 615 *Streptomyces* species were divided in to 130 clades (Labeda *et al.*, 2012). The 16S genes from well-studied *Streptomyces* species, as well as those relevant to this study, have been analysed using Timura-Neil Maximum Likelihood Model to build a phylogenetic tree (Kumar, Stecher and Tamura, 2016, *Figure 1-2*). Recently, multiple *Streptomyces* housekeeping genes have been sequenced to improve the low taxonomic resolution which occurs as a result of 16S sequence similarities among *Streptomyces.* These include the *ssgA*/*ssgB* genes which facilitate easier identification of the species boundaries whilst also acting as a morphological indicator, with *Streptomyces* which can and cannot sporulate in liquid culture separated by distinct phylogenetic clades (Girard *et al.*, 2013)

Moreover, recent research has indicated that taxonomically separate species of *Streptomyces* with distinct specialised metabolomes have identical 16S genes (Antony-Babu *et al.*, 2017) and therefore, whole genome is sequencing is superseding 16S analysis to define species boundaries (Alam *et al.*, 2010). In modern phylogenetic analysis, it is widely viewed that phylogenetic studies may benefit from robust whole-genome analysis and multi locus sequence typing to successfully decipher complex taxa such as *Streptomyces* (Sangal *et al.*, 2016). For example, AutoMLST software designed in studies by Alanjary *et al* use an Automated Multi-Locus Species Tree approach to determining phylogenetic relatedness among the Actinobacteria (Alanjary, Steinke and Ziemert, 2019). Despite the widespread use of

Average Nucleotide Identity of an indicator of species boundaries, currently, physical methods such as DNA-DNA hybridisation and biochemical tests are still regarded as the ‘Gold Standard’ for prokaryote species definition (Konstantinidis, Ramette and Tiedje, 2006; Richter and Rosselló-Móra, 2009), although recent comprehensive studies of related Actinobacterial taxa suggest that genome based taxonomy can resolve these issues (Sangal et al., 2016).

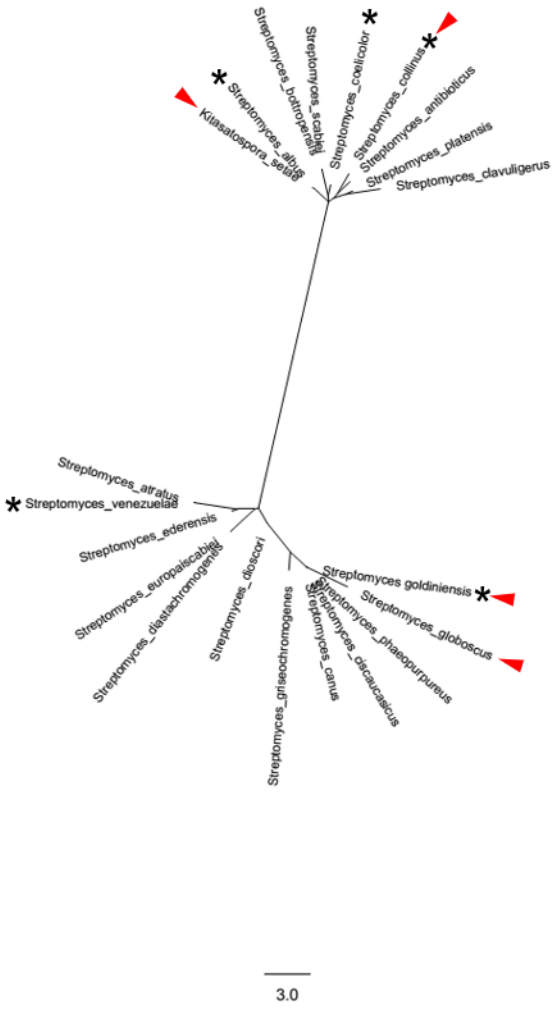


Figure 1-2: **16S rRNA gene phylogenetic analysis of selected *Streptomyces* species by maximum-likelihood method**. Evolutionary history was inferred by using the Maximum Likelihood method based on the Tamura-Neil model. The tree with the highest log likelihood (-4638.5672) is shown. The tree is drawn to scale, with branch lengths measured in the number of substitutions per site. The analysis involved 21 nucleotide sequences. All positions containing gaps and missing data were eliminated. There were a total of 1283 positions in the final dataset. Evolutionary analyses were conducted in MEGA7. Species used for experimental work in this thesis are indicated by *, species known to encode elfamycins are indicated with a red triangle.

1.2 The Metabolome of *Streptomyces:* Nature's Medicine Cabinet

1.2.1 Primary Metabolism in *Streptomyces*

The term 'primary metabolism' encompasses the catabolic and anabolic processes which are essential for survival in bacteria. This includes elements of carbon and nitrogen metabolism as well as reproductive processes. Primary metabolites are typically produced during the exponential phase of growth and incorporate essential compounds for the growth and building of a typical cell (Hodgson, 2000; Chevrette *et al.*, 2020). In *Streptomyces,* the presence of several copies of genes which encode the same biological function can be observed. This has evolved as an adaptive response to complex niches such as soil. Studies have shown that this gene family expansion has facilitated the evolution of specialised metabolism in these bacteria. This is evidenced by the disproportionate number of gene expansion events in specialised metabolite-producing bacteria (Schniete *et al.*, 2018). Consequently, the relationship between primary metabolism and specialised metabolism in *Streptomyces* is well studied and has been modelled computationally for *Streptomyces coelicolor* A(32) (Amara, Takano and Breitling, 2018).

1.2.2 Introduction to Specialised Metabolism in *Streptomyces*

Secondary metabolites are typically produced in the stationary phase of growth and have been historically regarded as non-essential for survival (Van Wezel and McDowall, 2011). These compounds are diverse in structure and function and many provide bioactive properties which confer evolutionary advantages over competing species during ecological warfare (Maplestone, Stone and Williams, 1992). Moreover, each species has its own 'parvome' or species specific catalogue of small molecules, with the expression of such molecules coinciding with the stationary phases of growth (Davies, 2013). Pleiotropic switching, the simultaneous expression of sporulation and antibiotic biosynthesis genes in *Streptomyces,* can be widely

observed. It is this property that lead to these compounds being termed ‘Secondary Metabolites’, with their expression occurring at the latter phases of growth. However, as the production of these compounds has been most intensively studied under laboratory monoculture, the native triggers for expression in the environment could not be defined, and hence non-essentiality to the organism cannot consistently be confirmed. Therefore, debate within the scientific community ensued and resulted in amendment of the preferred terminology to ‘specialised metabolites’. This reflects recent discoveries which suggest such compounds are not necessarily accessory but are in fact essential for survival under specific environmental considerations (Chevrette *et al.*, 2020). For example, in iron-limited conditions, the specialised metabolite siderophore Griseobactin of *Streptomyces griseus* is required for growth and therefore is not considered ‘accessory’ (Patzer and Braun, 2010). Some compounds previously identified as secondary metabolites have specialised roles in the lifecycle of the organism such as the roles played by the lantipeptide SapB (the polyketide spore pigment encoded by the *whiE* locus (Kodani *et al.*, 2004). The exact ecological role for most specialised metabolites has yet to be elucidated (de Lima Procópio *et al.*, 2012).

Specialised metabolite natural products are important both industrially and medicinally, with small molecules from *Streptomyces* contributing significantly to clinically used antibiotics . For example, chloramphenicol, used to treat conjunctivitis is produced by *Streptomyces venezuelae,* although synthetically manufactured (Bibb *et al.*, 2014), and streptomycin can be produced through the fermentation of *S. griseus* (Schatz, Bugle and Waksman, 1944). In addition, *Streptomyces* species produce many other molecules of interest including antitumor and immunosuppressant compounds. With the potential to produce a plethora of molecules with an array of structures and properties, *Streptomyces* can be exploited as a source of new drug

molecules, and libraries of their metabolites can be screened for a specific purpose. In addition, there is a constant search to find new *Streptomyces* species and other species of the phylum Actinobacteria in order to find unique new molecules. This process has become particularly relevant in recent times due to the increased search for novel antibiotic compounds in an aim to combat the looming antibiotic resistance crisis (Jackson, Czaplewski and Piddock, 2018).

Traditionally, culture dependent methods of antibiotic discovery were predominant. This typically involved isolating a species from a specific environment, facilitating its growth in the laboratory and using disk diffusion assays or solvent extracts from liquid culture to test against indicator species such as *E. coli* and *B. subtilis* (Katz and Baltz, 2016). Although many of our clinically used antibiotics were discovered via this route, there are disadvantages to this method. For example, the rate of rediscovery is high and only a limited number of species can be cultured on isolation media, limiting the available microbial resource (Aminov, 2010). Furthermore, these traditional techniques neglect the role of unculturable bacteria in drug discovery. New technologies such as the iCHIP, have been used to study previously unculturable bacteria, and have facilitated the discovery of new species and bioactive molecules, such as Teixobactin, a specialised metabolite of *Eleftheria terrae* (Ling *et al.*, 2015; Homma *et al.*, 2016).

1.2.3 The Genetic Basis of Specialised Metabolites: Biosynthetic Gene Clusters

The contiguous stretches of DNA which encode specialised metabolites in bacteria are termed Biosynthetic Gene Clusters (BGCs) and contain the required genes for the assembly of small molecules. Typically, these clusters contain genes encoding enzymes for substrate conversion and accessory enzymes which catalyse the addition and conversion of accessory groups. These gene clusters can be categorised

into 60 subclasses, according to antiSMASH (Blin *et al.*, 2016). To date, more than 1920 gene clusters have been recorded in the BGC database MIBiG (Minimum Information about a Biosynthetic Gene Cluster, (Medema *et al.*, 2015), including clusters from bacteria, fungi and plants.

1.2.3.1 Regulation of biosynthetic gene cluster expression

The expression of BGCs is tightly regulated and occurs during pleiotropic switching, coinciding with the onset of aerial hyphae formation (van Wezel and McDowall, 2011). Regulation can be global or cluster-specific. The most widespread BGC regulatory systems are one component systems. The most common cluster-situated regulator (CSR) in *Streptomyces* is the *Streptomyces* Antibiotic Regulatory Protein (SARP), which can be found in many types of BGC including Polyketide Synthases (PKSs) and Non-Ribosomal Peptide Synthetases (NRPSs; Bibb, 2005). These positive regulators activate biosynthesis by binding to heptameric repeat sequences in the promoter regions of the cluster they regulate (Okamoto *et al.*, 2009). LuxR-type regulators (LALs) can also be found in *Streptomyces* and are known to activate rapamycin biosynthesis in *Streptomyces hygroscopicus* (Kuščer *et al.*, 2007). There are also regulators which are involved in the adaptive response to specific molecules. TetR-type regulators contain a ligand-binding domain in addition to a helix-turn-helix DNA binding domain. On ligand binding, a conformational change is induced which results in a reduced affinity for DNA binding, and hence this enables negative feedback of expression. This class of regulator can be found in the kirromycin biosynthetic gene cluster (Robertsen *et al.*, 2018). Expression of BGCs can also be regulated at a global level. In *Streptomyces,* the global, developmentally-associated regulator BldD can activate BGC expression, through activation of a CSR (McLean, Wilkinson, *et al.*, 2019) or by directly activating the expression of biosynthetic genes through the targeting of TTA codons, such as the erythromycin genes in

Saccharopolyspora erythraea (Chng *et al.*, 2008). As well as the aforementioned one component systems, two component systems have also be shown to play a role in BGC regulation. These systems utilise distinct sensor and DNA-binding proteins to respond succinctly to changing environmental stimuli. Two-component system can act globally or specifically on an individual BGC (McLean, Lo, *et al.*, 2019). Finally, it must be stated that regulation of specialised metabolite expression in *Streptomyces* involves a complex network of processes and other factors can also contribute such as the highly phosphorylated guanosine nucleotide ppGpp and PHO boxes (Chakraburtty and Bibb, 1997). These processes are summarised in *Figure 1-3.*

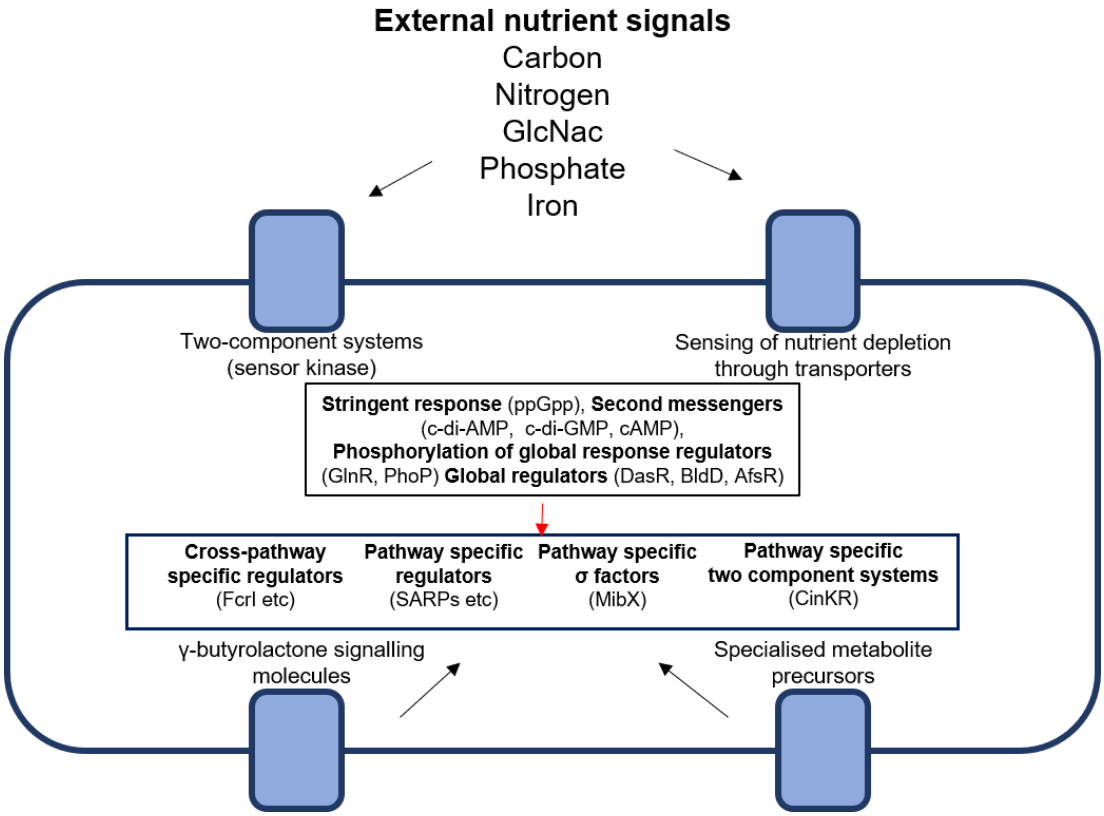


Figure 1-3: **Schematic representation of the role of global and specific regulators in specialised metabolism in *Streptomyces.*** Adapted from Hoskisson & Fernandez-Martinez (2018).

1.2.3.2 Biosynthetic gene cluster associated immunity genes

It is vital that strains producing specialised metabolites have an export mechanism to confer immunity to antibiotic natural products, particularly if the compound has a conserved target which is present in the producing organism, as well as their competitor. Consequently, resistance mechanisms have evolved with the compounds themselves and are often encoded within BGCs. This is the case for most efflux pumps (Thaker, Spanogiannopoulos and Wright, 2010; Blanco *et al.*, 2016). There are rare examples of synergistic resistance mechanisms located at individual loci, which will be discussed later in this chapter. The most common resistance gene within *Streptomyces* gene clusters is the Major-Facilitator Superfamily-type (MFS) multidrug efflux pump, representing almost 40% of all *Streptomyces* transport proteins (Zhou *et al.*, 2016). Proteins from this family are diverse and capable of exporting a wide range of molecules, yet each individual protein has a limited export capacity, which is evidenced by the presence of multiple efflux pumps in specialised metabolite producers. The conformation of MFS proteins is conserved and is characterised by the presence of 12 or 14 transmembrane helices held together by a cytoplasmic loop (Saidijam *et al.*, 2006).

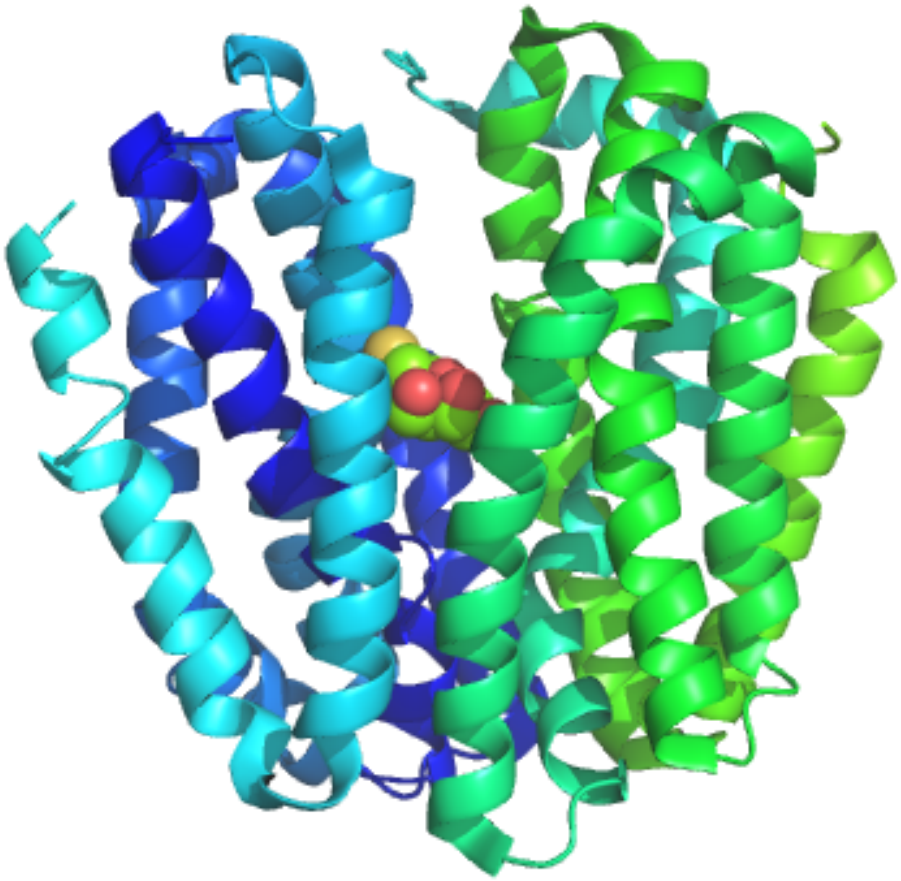


Figure 1-4: **Structure of the model MFS-type exporter, the lactose permease LacY (*E. coli*).** Taken from Abrahamson et al, 2003 (PDB: 1PV6)

1.3 *Escherichia coli*

1.3.1 *E. coli:* The model bacterium

Escherichia coli is a Gram-negative, rod-shaped bacterium which belongs to the family *Enterobactericae.* The first *E. coli* strain to be identified was isolated from the human colon by Austrian physician Theodor Escherich in 1885 and was originally named *Bacterium coli*. However, in 1954, its name was changed to honour its discoverer (Friedmann, 2006). *E. coli* is a facultative anaerobic organism which grows optimally at 37 ^{o}C (Dunne *et al.*, 2017). These growth conditions mimic the lower intestine of warm-blooded mammals, where *E. coli* can be extensively isolated (Karp *et al.*, 2007).

E.coli is regarded as the best-studied free-living organism on earth, with both its physiology and genetics well characterised. *E. coli* K-12, a widely-used laboratory strain isolated from an infected patient was one of the first organisms to have its whole genome sequenced (Blattner *et al*., 1997). From this study it was determined that the *E. coli* K-12 genome is 4.6 MB in size and contained mainly in one large chromosome. The genome has 4300 genes which have almost all been assigned a function (Baba *et al.*, 2006). The extensive understanding of *E. coli* physiology and genetics has cemented its role as a model organism in the study of bacteria.

1.3.2 Pathogenic *E. coli*

In most cases, *E. coli* forms part of the human gut microflora, and provides benefits to the host such as the production of vitamin K (Bentley and Meganathan, 1982). However, Theodore Escherich identified *E. coli* as a causative agent of some diarrhoeal disease found in children (Friedmann, 2006). In 1982, Karmali and their co-workers characterised the links between some *E. coli* strains and acute, systemic symptoms (Karmali *et al.*, 1983). In addition, *E. coli* can infect other niches in the human body, for example, 90% of urinary tract Infections are caused by

Uropathogenic *E. coli* (UPEC) (Wiles, Kulesus and Mulvey, 2008), the circulatory system and CSF, with sepsis and meningitis also be caused by *E. coli* (Bonacorsi and Bingen, 2005).

Historically, the standard method for the identification of pathogenic *E. coli* used antibodies to test for surface antigens. These are specifically the O-polysaccharide antigens, flagellar H-antigens, and capsular K-antigens. Currently, more than 186 O-antigens and 53 H-flagellar antigens have been identified and consequently, this method of identification is often complex (Fratamico *et al.*, 2016). Therefore, recent identification methods have focussed on bacteriophage typing and whole genome sequencing (Banjo *et al.*, 2018). However, identification of *E. coli* serotype remains an important tool in infection diagnostics as there are specific serotypes associated with more acute disease pathology (Donnenberg, 2013). Pathogenicity mechanisms in diarrhoeagenic *E. coli* are largely serotype dependent and much variation can be observed between serotypes (Clements *et al.*, 2012).

Many infectious *E. coli* serotypes can be found within animal reservoirs, with modern farming methods often facilitating their eventual contamination of the food chain and subsequently, human infection (Ferens and Hovde, 2011). Patients acquire these infections through the faecal-oral route, however, the specific source of the pathogen is often serotype linked (Wasteson, 2001). Enteric pathogenic (diarrhoeagenic) *E. coli* can be divided in to six categories based on their virulence characteristics, which are summarised in *Table 1-A.* The work in this thesis largely focusses on Enteropathogenic *E. coli* (EPEC), a pathogen of the large intestine and Enterohaemorrhagic *E. coli* (EHEC, also VTEC or STEC), a more acute pathogen of the small intestine.

Table 1-A: **Summary of the six major enteric *E. coli* pathogens and their epidemiology (Clements et al., 2012).**

E. coli strain	Symptoms	Epidemiology
Enterotoxigenic *E. coli* (ETEC)	Watery diarrhoea, abdominal pain	Common in developing countries. Contaminated water often responsible. Faecal oral route most likely mode of transmission.
Enteropathogenic *E. coli* (EPEC)	Watery diarrhoea, abdominal pain, vomiting	Common in developing countries. Contaminated water often responsible. Faecal oral route most likely. Person to person spread can occur.
Enteroinvasive *E. coli* (EIEC)	Watery diarrhoea, abdominal pain, vomiting, fever	Worldwide. Contaminated food main factor. Faecal oral route most likely- person to person spread likely.
Enterohaemorrhagic *E. coli* (EHEC)	Bloody diarrhoea, vomiting, severe abdominal cramps. Can result in haemolytic uremic syndrome	Worldwide outbreaks, contaminated meat often responsible. Faecal oral route most likely.
Enteroadherent *E. coli* (EAgEC)	Chronic diarrhoea, vomiting and fever	Common in travellers, ingestion of food. Person to person spread unlikely.
Verotoxogenic *E. coli* (VTEC)	Bloody diarrhoea, vomiting, abdominal cramps	Common in travellers, ingestion of food. Person to person spread unlikely.

1.3.3 Enteropathogenic *E. coli* (EPEC).

Enteropathogenic *E. coli* (EPEC) is an important global pathogen and causative agent of chronic diarrhoea. EPEC is accountable for approximately 800,000 deaths *per anum*, with children under five in developing countries most likely to become ill (Liu *et al.*, 2012). EPEC was first characterised in 1954 when it was associated with geographically contained outbreaks of diarrhoea (NETER *et al.*, 1955) however, these infections were not found to consistently comply with Koch's postulates as patients carrying with EPEC are regularly found to do so asymptomatically (Tozzoli and Scheutz, 2014). Sources of EPEC infection are most commonly contaminated water supplies (Ochoa and Contreras, 2011). EPEC strains generally fall in to two categories, typical EPEC (tEPEC) and atypical EPEC (aEPEC). Traditionally, tEPEC strains were characterised by their ability to attach and efface the gut microvilli, forming microcolonies. The genes which facilitate this attachment can be found on the *E. coli* adherence factor plasmid (pEAF) and include genes encoding for bundle forming pilli (*bfp*) and the virulence regulator Per. Typical EPEC strains include the widely used laboratory strain E2348/69, with O126:H6 serotype (Iguchi *et al.*, 2009). On the contrary, this plasmid is absent in aEPEC which typically adheres more diffusely. In addition, aEPEC strains can be isolated ubiquitously from both animal and human reservoirs, whereas tEPEC is typically isolated exclusively from human sources (Hernandes *et al.*, 2009).

Currently, treatment of EPEC is problematic for several reasons. Firstly, pathogens which reside in the gut often display an increased instance of antimicrobial resistance, due to horizontal gene transfer between the microbiota (Modi, Collins and Relman, 2014). In addition, the use of broad-spectrum antibiotics for the treatment of EPEC infections can have negative consequences for the gut microbiota, which may become depleted. It is known that a healthy gut microbiota is essential for the innate

immune system to overcome EPEC infections and therefore, its disruption can have acute consequences for the host. Therefore, in most cases treatment is limited to rehydration therapy (Sperandio *et al.*, 2000a).

1.3.4 Enterohaemorrhagic *E. coli*

Enterohaemorrhagic *E. coli* (EHEC, also known as STEC) are acute pathogens of the small intestine which are responsible for foodborne outbreaks of bloody diarrhoea. The most common and best studied serotype, *E. coli* O157: H7 can be consistently isolated from the rumen of cattle, and therefore infection can occur from the ingestion of contaminated food products of cattle origin such as undercooked beef and unpasteurised milk (Ferens and Hovde, 2011). In 1993, 732 people in California were infected with this serotype during the infamous 'Jack in the Box' outbreak associated with undercooked beef burgers. In addition, contamination of horticultural irrigation facilities with bovine waste can result in the contamination of salad vegetables, such as the outbreak associated with American romaine lettuce in 2018 (CDC, 2018). In recent years, as food supplies travel further distances before reaching consumers, EHEC outbreaks have become more wide-reaching. For example, an outbreak of the chimeric O104:H4 EHEC strain associated with organic, sprouting vegetables from Germany resulted in over 2000 cases over 14 different countries (Mellmann *et al.*, 2011).

Symptoms associated with EHEC are generally more severe than those associated with EPEC. This is largely due to the characteristic presence of prophage-encoded shiga toxins with in the EHEC genome (Nguyen and Sperandio, 2012). There are two main groups of shiga toxins, Stx1 and Stx2 and both of which can have severe consequences for the host if they contact organs such as the kidney or brain. During EHEC infections, shiga toxins can facilitate the dissemination of EHEC from the gut, resulting in systemic infection. When shiga-toxins are expressed during this phase of

pathogenesis, severe damage can be caused as Stx can inhibit protein synthesis by cleaving the 60S unit of the ribosome, with cytotoxic effects, particularly in the kidney. This pathology is known as Haemolytic Uremic Syndrome (HUS) and was first characterised by Karmali and their co-workers in the 1980's (Karmali, 1989). Patients who contract HUS often require lifelong treatment such as dialysis or kidney transplantation. Furthermore, severe cases of HUS can result in death, particularly in children and the elderly (Tarr, Gordon and Chandler, 2005). It is estimated that ten percent of patients who acquire EHEC infections go on to develop Haemolytic Uremic Syndrome. Patients who survive this often require life-long treatment which may involve but is not limited to dialysis, use of antiplatelet and thrombolytic agents and periodically, kidney transplantation solutions which are costly to health services. In 2006, it was estimated the average cost of treatment for a HUS patient was 7.35 million US dollars per year (Goldwater and Bettelheim, 2012).

In Scotland, there is a specific, yet somewhat undefined problem with EHEC infections. On average, three times more EHEC infections are reported *per capita* in Scotland than in the rest of the U.K (Herbert *et al.*, 2014). It is thought that many factors influence this including rural lifestyles, direct contact with cattle, genetics and a higher prevalence of EHEC in cattle carrying prophage associated with a 'super-shedding' phenotype (Chase-Topping *et al.*, 2007).

As with EPEC, there are issues with using antimicrobial agents for the treatment of EHEC, with the problems previously described for EPEC staying true for EHEC. Furthermore, the disadvantages of the use of traditional antibiotics extend to the induction of the bacterial SOS response, which is induced by the denaturation of genomic DNA. Single stranded DNA is present in high numbers in cells treated with traditional antibiotics, including Ciprofloxacin. The presence of single stranded DNA can induce the expression of *recA,* which is involved in the bacterial SOS response.

In turn, RecA is able to cleave LexA repressors which are responsible for the regulation of Stx expression (*Figure 1-5*, Fuchs *et al.*, 1999). In summary, DNA damage caused by antibiotics results in the activation of the Stx-encoding phage. Hence, Stx production is upregulated and symptom severity increases (Huerta-Uribe *et al.*, 2016). Therefore, an alternative therapy that circumvents the induction of *recA* is desirable.

Due to the hazardous implications associated with treatment of EHEC with traditional antibiotics, the majority of patients currently receive treatment that is limited to supportive care. This often includes rehydration therapy and total blood volume expansion (Goldwater and Bettelheim, 2012). When HUS occurs, patients may require lifelong treatment such as dialysis and kidney transplantation. In children, EHEC fatality rates can be as high as 10%, and with up to 2200 cases in the USA per in 2017, Shiga-toxin producing *E. coli* pose a risk to the health of the general population. Therefore, the introduction of novel strategies with alternative mechanisms to traditional antibiotics could prove to be a useful intervention for the treatment of these infections.

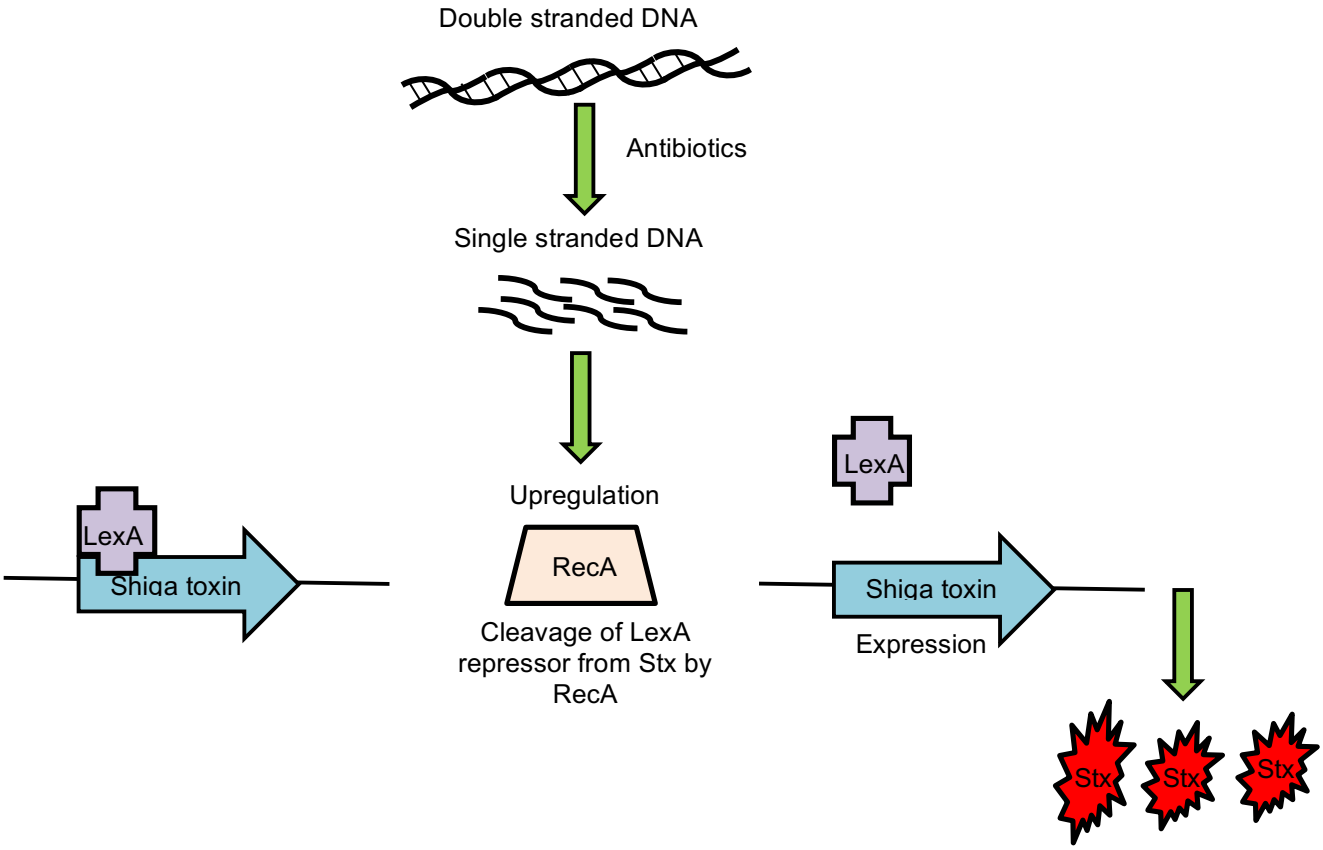


Figure 1-5: **Schematic diagram of the RecA mediated induction of shiga toxin**. Adapted from Fuchs *et al*, 1999.

1.3.5 The role of the Type III Secretion System in EPEC and EHEC infections

Bacteria have evolved a wide array of mechanisms to facilitate the transport of small molecules and proteins to the extracellular space or target host cells (Costa *et al.*, 2015). Often these take the form of molecular nanomachines, known as secretion systems. These systems vary largely in size and function and can be broadly categorised in to seven groups (Type I-Type VII Secretion Systems). In Gram-negative bacteria, Type I, II, III, IV and VI secretion systems span both the outer and inner membrane, whereas Type V and Type VII systems span only the cytoplasmic membrane, with the latter found only Gram-positive bacteria (Gerlach and Hensel, 2007; Famelis *et al.*, 2019). Secretion Systems that span the outer membrane only also include assembly apparatus for surface appendages such as Type I Pilli and curli (Lee *et al.*, 2014).

The Type Three Secretion System (T3SS) or 'injectosome' is a proteinaceous appendage present on the outer membrane of certain Gram-negative bacteria (*Figure 1-6*). In many pathogenic strains of *E. coli,* the T3SS is a needle-like apparatus which facilitates the translocation of effector proteins to the epithelial cells of the gut. Through the injection of these proteins, that include the Translocated Intimin Receptor (Tir) and the Mitochondria Associated Protein (Map), EHEC and EPEC are able to polymerise actin filaments and form intimate junctions with epithelial cells, where several other virulence factors are expressed to destabilize cellular processes (Mills *et al.*, 2008).

The T3SS of EHEC and EPEC is encoded by the highly-conserved Locus of Enterocyte Effacement (LEE) Island (Deng *et al.*, 2004). This pathogenicity island encodes 42 genes on five conserved operons and is regulated by the H-NS-type master regulator, Ler. In turn, *ler* expression is regulated by specific regulators such as GrlA and GrlR, in addition to global regulators which mediate LEE expression in

response to environmental stimuli in EHEC. In EPEC, *ler* is also regulated by Per, a regulator encoded on the virulence associated plasmid pEAF. Ler is also responsible for the regulation of multiple virulence genes outside of the LEE pathogenicity island. The *ler* regulon has been characterised in EHEC and extends to biofilm and extracellular polysaccharide genes (Bingle *et al.*, 2014).

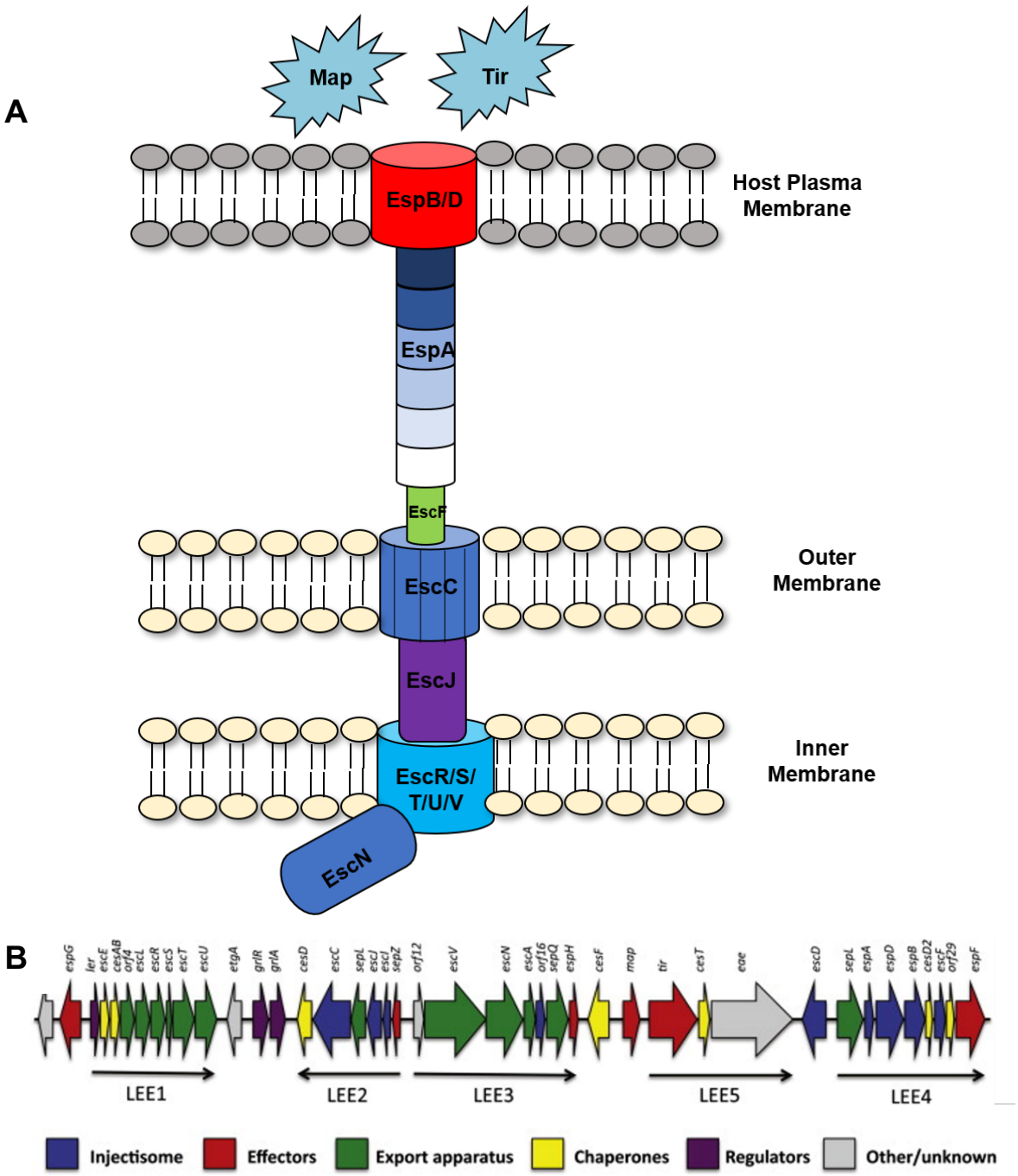


***Figure 1-6*: Structure of the Lee-encoded Type III Secretion System**. (A) Schematic diagram depicting the structure of EHEC T3SS injectosome system. Adapted from Garmendia *et al,* 2005. (B) Genetic architecture of the LEE pathogenicity island in EHEC, taken from Crawford *et al*, 2002.

1.3.6 The Type III Secretion System: a target for antivirulence drugs?

Significant issues with the use of antimicrobial drugs to treat EPEC and EHEC infections have resulted in a lack of effective treatment for the symptoms of these conditions (Pacheco and Sperandio, 2012). As previously discussed, this is due to the negative consequences associated with treatment such as microbial dysbiosis of the gut and increased Stx production in EHEC (Nguyen and Sperandio, 2012). Therefore, novel treatment concepts are required.

One such strategy is the use of 'Anti-virulence' (AV) compounds. Mechanisms of AV drugs fundamentally differ from those of traditional antibiotics. Where antibiotics are used to kill pathogens or inhibit their key metabolic processes, distinctively, anti-virulence drugs are deployed to disarm the infecting bacteria by blocking virulence factors which are required to colonise and persist in their host. Consequently, resistant organisms are less profoundly selected for in the population (Cegelski *et al.*, 2008), as the mutations which result in resistance to AV compounds do not convey the same fitness benefits as those associated with AMR. In most cases, AV mutations in bacteria are more nuanced, with colonisation efficiency often sacrificed to acquire AV resistance (Allen *et al.*, 2014). It has therefore been suggested these drugs could represent an 'evolution-proof' strategy for infections caused by characterised pathogens, with AV resistance mutations often appearing more stochastically and reversibly (Sperandio *et al.*, 2000b). The targets for anti-virulence drugs are varied and can represent intracellular or extracellular structure, or in fact, a process such as quorum sensing (Calvert, Jumde and Titz, 2018). However, all suitable targets must incorporate factors which if inactivated, will result in a significant reduction in the ability of the pathogen to establish and maintain infection in the host. Thus, providing opportunity for the infection to be cleared by the native immune system (Zambelloni, Marquez and Roe, 2015; Mühlen and Dersch, 2016).

Bacterial infections which occur within the gut may benefit from treatment with AV drugs, as the regular excretion of waste from the intestines can facilitate the elimination of pathogenic bacteria from the gut following the inhibition of infection (Totsika, 2016; Calvert, Jumde and Titz, 2018). Anti-virulence therapies for the opportunistic pathogen of the gut, *Clostridium difficile*, have been investigated (Beilhartz *et al.*, 2016; Tam *et al.*, 2018; AbdelKhalek *et al.*, 2019). Patients who acquire *C. difficile* infections often do so as a consequence of the dysbiosis of the gut microbiome which occurs as a result of broad spectrum antibiotic regimens (Fordtran, 2006). Hence, this infection is typically nosocomial, with 2.24 per 1000 hospital inpatients contracting the disease in the US (Balsells *et al.*, 2019). Virulent strains of *C. difficile* require the production of toxins known as Toxin A and Toxin B to modify host target 'rho' proteins, in order to facilitate destruction of the host cytoskeleton and cause cell death (Kuehne *et al.*, 2010, Lyras *et al.*, 2009). Therefore, anti-virulence strategies designed to neutralise these toxins have been investigated. In 2017, the FDA approved the monoclonal antibody Bezlotoxumab – which targets *C. difficile* Toxin B – as a therapy for patients with recurring *C. difficile* infections (Navalkele and Chopra, 2018). Published clinical data has identified that Bezlotoxumab has a rate of cure which is comparable with current frontline therapies, however, the antibody is significantly more successful in preventing recurring infections (Wilcox *et al.*, 2017). Hence, the successful implementation of this AV strategy is a promising foundation for future therapies of similar premise.

Currently, a range of STEC anti-virulence strategies have been investigated including the use of Shiga toxin neutralisers and biofilm inhibitors (Kitov *et al.*, 2000; Lee *et al.*, 2014; Huerta-Uribe *et al.*, 2016; Mühlen and Dersch, 2016). For example, high throughput screening of synthetic compounds lead to the identification of compounds which reduce Shiga-toxin production in a dose dependent manner (Huerta-Uribe *et*

al., 2016). However, the Type III Secretion System (T3SS) has remained a primary target for AV therapy as it is an essential colonisation factor in both humans and ruminants (K Kimura *et al.*, 2011; Beckham and Roe, 2014; Zambelloni, Marquez and Roe, 2015; Pendergrass and May, 2019). This injectosome system facilitates the delivery of bacterial effector proteins into host epithelial cells (Cornelis, 2006), resulting in formation of attaching and effacing lesions (Jarvis *et al.*, 1995; Iyoda *et al.*, 2006; Schmidt, 2010). Notably, synthetic salicylidene acylhydrazide derivatives have been used to inhibit Type Three Secretion effectively, but were found to have multiple targets which may result downstream effects (Veenendaal, Sundin and Blocker, 2009; Wang *et al.*, 2011). One promising Type III Secretion System Inhibitor which has recently been investigated is aurodox, a specialised metabolite of the soil bacterium *Streptomyces goldiniensis,* and hence, its properties will be investigated throughout this work.

1.4 Aurodox: Teaching an old drug new tricks

1.4.1 Aurodox: A specialised metabolite from *Streptomyces goldiniensis* with anti-Gram-positive activity

Streptomyces goldiniensis was first isolated from soil of the Caribbean island of Bermuda in 1969 (Maehr *et al.*, 1980). At this point, classical culture-dependent methods of compound isolation were carried out in order to identify potentially novel bioactive compounds. Subsequently, it was shown that when semi-pure extracts of *S. goldiniensis* culture supernatant were given to mice which had been inoculated with lethal doses of *Streptococcus pyogenes*, the rodents were protected and able to recover from the infection (Berger *et al.*, 1973a). Moreover, a range of compounds were purified through activity guided fractionation and chromatography, and a strong antibiotic effect from the compound initially known as X-5108 and subsequently named aurodox was observed (Berger *et al.*, 1972). Aurodox exhibits a strong inhibition of the growth of many Gram- positive pathogens, including *Streptococcus pyogenes, Micrococcus luteus* and other *Streptomyces* species. Weak activity against Gram-negative bacteria was also observed. Significantly, resistance to aurodox did not arise until after 64 generations of growth. Hence, the results of these experiments led to the identification of aurodox as a potential antibiotic drug candidate (Berger *et al.*, 1973b). During these studies, a fermentation method for aurodox production in *S. goldiniensis* was designed and patented. Aurodox fermentation can be carried out in undefined production media, and can be extracted from culture supernatant after 156 hours (Maehr *et al.*, 1979).

1.4.2 Chemical properties and structure of aurodox

In the seminal works by Berger and their co-workers (Berger *et al.*, 1972), the chemical formula of aurodox was assigned as $C_{44}H_{62}N_2O_{12.}$ However, the structure of the molecule was not derived until later work by Maehr *et al* (1979) and is shown in

Figure 1-7. This study used 13-Carbon Nuclear Magnetic Resonance to demonstrate that the structure of aurodox is highly methylated and contains two unsaturated carbon chains which chemically connect leucine and glycine derived groups. In addition, it was shown that transmethylation is responsible for the addition of the methyl groups. In attempts to synthesise synthetic analogues of aurodox, the alteration of the pyridine moiety lead to a decrease in activity suggesting this was group was responsible for antibiotic activity (Maehr *et al,* 1972)

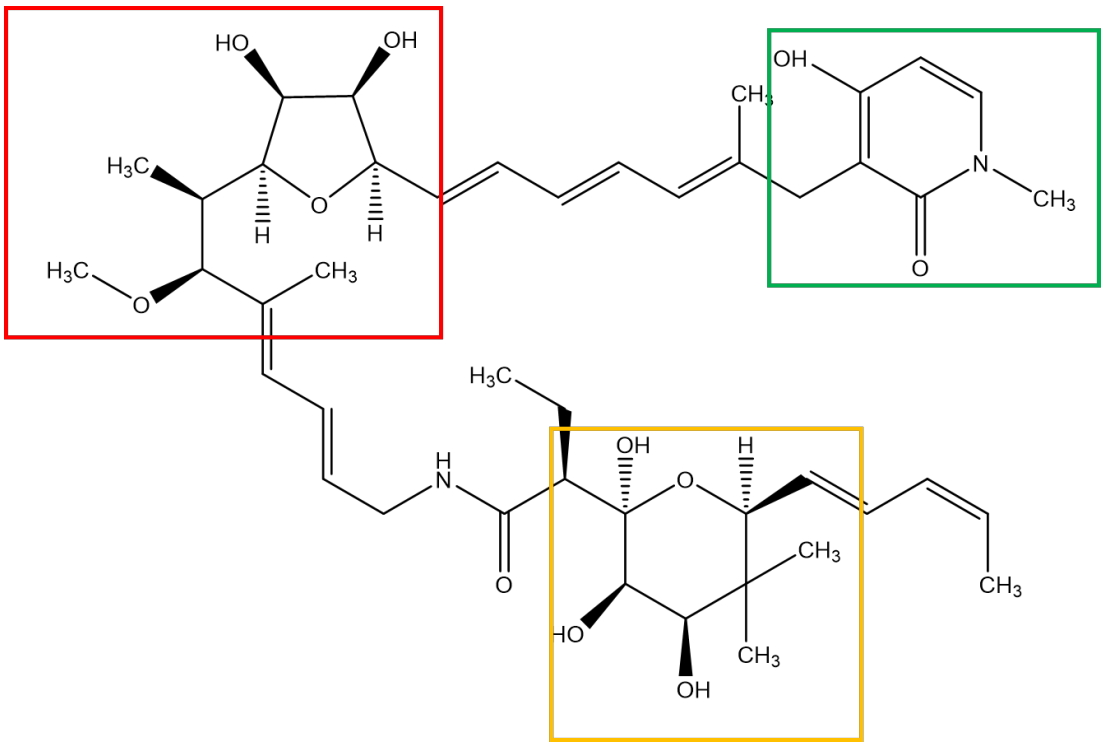


Figure 1-7: **Chemical structure of aurodox**. Green box highlights methylated pyridone ring, yellow box indicated goldinoic acid domain, and red box shows tetrahydrofuran ring.

1.4.3 Mechanism of action: Interaction of aurodox with Elongation Factor Thermo-unstable

EF-Tu (Elongation factor Thermo-unstable) plays an essential role in protein biosynthesis and it is consequently highly conserved across the prokaryotes. The crystal structure of this protein was published by Berchtold *et al.*, 1993, with domains identified which bind to both tRNA and the ribosome, therefore identifying its function as catalyst for their interaction. It has since been shown that during protein synthesis in prokaryotes, EF-Tu delivers charged tRNA to the ribosome where GTP hydrolysis will induce a conformational change in the protein, facilitating its release from the ribosome, and allowing the cycle to continue (Berchtold *et al.*, 1993). It has been shown that aurodox can block this interaction by forming a complex with EF-Tu and GDP at a 1:1:1 ratio (*Figure 1-8*) and hence, blocking GTP hydrolysis. Therefore, initiating an irreversible conformational change leading EF-Tu to become stuck in complex with the ribosome, and hence preventing proteins from being synthesised (L Vogeley *et al.*, 2001). These bactericidal effects of aurodox can be observed both *in vitro* and *in vivo*. Despite these strong effects on Gram-positive growth, aurodox has never been licenced for use in humans. This is a result of its narrow spectrum of activity. The compound was, however, used as a growth promoter in livestock (Berger *et al.*, 1973a).

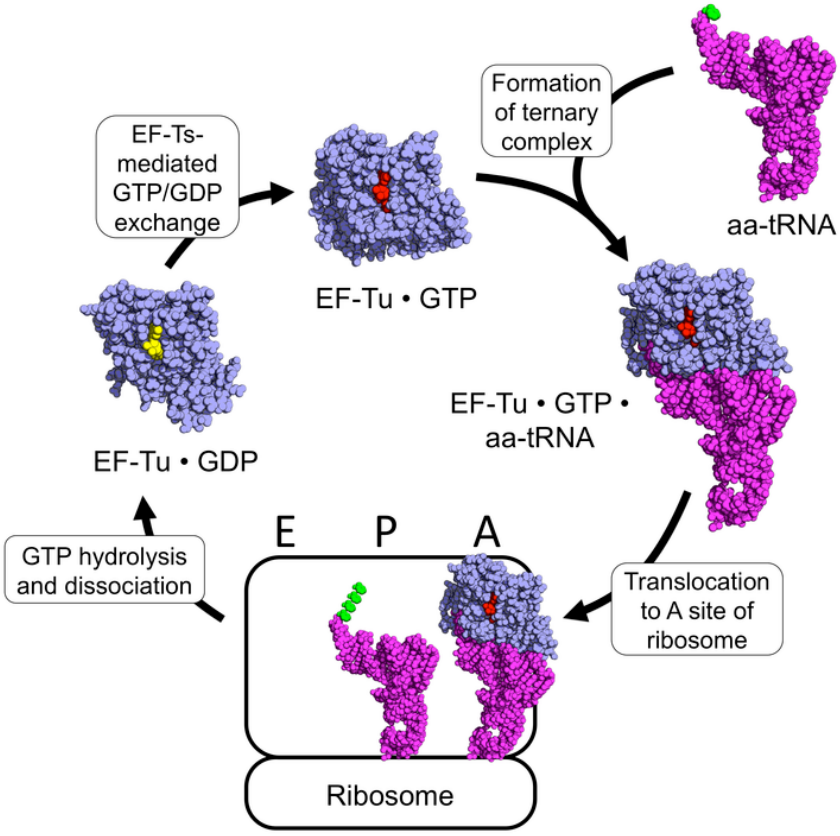


Figure 1-8: **The role of Elongation Factor Thermo-unstable (EF-Tu) in prokaryotic protein synthesis.** Taken from Vogely *et al*, 2001.

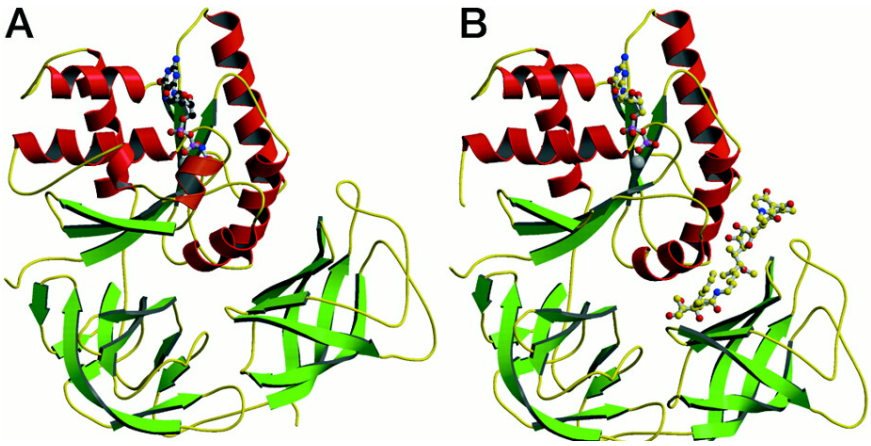


***Figure 1-9**:* **Interaction between aurodox and EF-Tu.** Alpha helices and beta sheets indicated in red and green repeatedly indicate tertiary structure of EF-Tu, with the ball and stick models representing aurodox of GTP. (A) Native confirmation of EF-Tu with GTP bound (B) Aurodox-bound EF-Tu blocking GTP hydrolysis. Adapted from Vogely *et al*, 2001.

1.4.4 Inhibition of the *E. coli* Type three secretion system by aurodox

The reliance of EHEC and EPEC on the T3SS to initiate infection has identified it as a target for novel therapies to fight infection. Typically, these are part of a wider anti-virulence approach in which the aim is to prevent infection by the inhibition of a single virulence factor without inducing a reduction in growth (Zambelloni, Marquez and Roe, 2015). Currently, treatment of EHEC infections with traditional antibiotics is not recommended due to stimulation of the bacterial SOS response (Imamovic and Muniesa, 2012). As EHEC and EPEC infections are typically cleared naturally by the host, anti-virulence approaches to treatment of EHEC and EPEC represent an exciting new strategy for the treatment of these infections. In addition, compounds that do not affect bacterial growth or survival reduce the evolutionary selective pressure on strains resistant to the treatment, enhancing the long-term viability of the therapy (Allen *et al.*, 2014).

Several anti-virulence compounds are actually the natural products of other bacterial species (Beckham and Roe, 2014). Aurodox, a specialised metabolite of *Streptomyces goldiniensis* was shown by Kimura et al (2011) to inhibit the translocation of T3SS encoded effectors in EPEC. Low concentrations (1.5 μM) were shown to inhibit secretion and abolition of detectable effector proteins was observed at 6 μM. Moreover, Kimura and his co-workers tested the effect of the compound *in vivo* through the use of a *Citrobacter rodentium* murine infection model where it was shown that mice treated with the compound survived lethal infections with limited effects on the intestinal tract (K Kimura *et al.*, 2011).

1.5. Scope of this project.

Specialised metabolites from *Streptomyces* have proved to be indispensable additions to the catalogue of molecules used in modern medicine. These compounds which have evolved as a result of inter-species conflict have been exploited for their anti-bacterial, anti-viral and anti-fungal properties and hence, their native roles in ecological conflict can confer benefits to human health (Chevrette *et al.*, 2020). The ability of *Streptomyces* to produce this plethora of molecules with an array of activity has augmented their role as an essential resource for drug discovery, with many avenues yet to be explored (van der Meij *et al.*, 2017).

In studies by Kimura *et al,* (2011) aurodox, a specialised metabolite of *S. goldiniensis* was shown to specifically inhibit the EPEC T3SS, with this property validated both *in vitro* and *in vivo.* The results of this study indicate that aurodox may contribute to a potential anti-virulence approach to the treatment of enteric pathogens such EPEC and EHEC and therefore, the suitability of the compound for repurposing should be investigated. Consequently, in this project, a central focus will be the characterisation of aurodox biosynthesis in an aim to optimise and diversify aurodox production for future experiments. A genetic approach will be taken, with sequencing of *S. goldiniensis* forming the basis for future experiments which will include the identification and characterisation of the BGC. This will therefore facilitate the heterologous expression of the aurodox gene cluster in an aim to improve aurodox production. Finally, the effect of aurodox on Type III Secretion in a range of enteric pathogens including EHEC and EPEC will be characterised as part of a larger aim to identify the mechanism of T3SS inhibition by aurodox. In summary, this work in the following thesis will aim to take aurodox forward as a candidate anti-virulence drug through gaining and understanding of its biosynthesis and mechanism of action.

It was hypothesised that aurodox may be used as an anti-virulence compound for the treatment of EPEC and EHEC and that understanding the mode of action of the T3SS inhibition effect and the biosynthesis of this molecule will help repurpose the molecule for clinical use.

1.6 Specific aims

- Characterisation of *S. goldiniensis* physiology and morphology and subsequent optimisation of aurodox fermentation
- Sequencing of the *S. goldiniensis* whole genome.
- Identification of the aurodox biosynthetic gene cluster and elucidation and characterisation of aurodox biosynthesis.
- Heterologous expression of aurodox production in *Streptomyces* hosts
- Understanding the role of *tuf2* in aurodox resistance
- Characterisation of the effect of aurodox on T3S in EPEC and EHEC and identification of the mechanism of T3SS inhibition.

Chapter 2 : Material and Methods

All chemicals used in this project were of analytical grade and were purchased from Sigma-Aldrich, Qiagen, Fischer Scientific or Invitrogen unless otherwise stated in the text. Sequencing reagents were purchased from each sequencer provider. Aurodox was purchased in its pure form (Enzo Life Sciences, Farmingdale, NY). A 1-mg/ml stock solution was prepared by dissolving the compound in dimethyl sulfoxide (DMSO). The final concentration of DMSO was less than 0.5% in all experiments.

2.1 Strains and plasmids

The strains used in this study are detailed in *Table 2-A*. Plasmids used and their antibiotic selection are described in *Table 2-B.* Working concentrations of selective markers for these plasmids are detailed in *Table 2-C.*

***Table 2-A*: Strains used in experimental work.**

Strain	Description	Genotype	Reference
E. coli DH5α	*E. coli* K12 derivative	*huA2* (*argF-lacZ*)U169 *phoA glnV44* 80 (*lacZ*)M15 *gyrA96 recA1 relA1 endA1 thi-1 hsdR17*	(Grant *et al.*, 1990)
E. coli DH10β Top10 ©	*E. coli* K12 derivative	*mcrA*, Δ(*mrr-hsdRMS-mcrBC*), Φ*lacZ*(del)*M15, ΔlacX74*, *deoR*, *recA1*, *araD139*, Δ(*ara-leu*)7697, *galU*, *galK*, *rpsL*(*SmR*), *endA1*, *nupG*	Invitrogen
E. coli BW21153	*E. coli* K12 derivative	lacI+ , *rrnBT14*, *ΔlacZWJI6*, *hsdR514, ΔaraBADAH33, ΔrhaBADLD78*	(Datsenko and Wanner, 2000)
E. coli ET12567	*E. coli* K12 derivative	*dam13*::Tn9, dcm6, hsdM, hsdR, recF143, zij201::Tn10, *galK2, galT22, ara14, lacY1, xylS, leuB6, thi-1, tonA31, rpsL136, hisG4, tsx78, mtli, glnV44*,	(MacNeil *et al.*, 1992)
E. coli E2348/69	Wild type EPEC O127:H6	NA	(Campellone, Robbins and Leong, 2004)
E. coli ΔescC	Derivative of EPEC E2348/69	*ΔescC*	(Campellone, Robbins and Leong, 2004)
E. coli ΔescF	Derivative of EPEC E2348/69	*ΔescF*	(Campellone, Robbins and Leong, 2004)
E. coli ΔescR	Derivative of EPEC E2348/69	*ΔescR*	(Campellone, Robbins and Leong, 2004)
E. coli ΔescV	Derivative of EPEC E2348/69	*ΔescV*	(Campellone, Robbins and Leong, 2004)
C. rodentium ICC168	Wild type *Citrobacter rodentium*	Wild-type	(Mallick *et al.*, 2012)
C. rodentium BS100	Shiga-toxin expressing *C. rodentium*	Φ *1720a02ΔRz*::cat *ΔStx2dactAB*::kan, CmR,	(Mallick *et al.*, 2012)
Enterococcus faecium ATCC 51299	*Enterococcus faecium* WT, clinical isolate	Wild-type	ATCC
Staphylococcus aureus ATCC 43300	*Staphylococcus aureus* WT, clinical isolate	Wild-type	ATCC
Klebsiella pneumoniae ATCC 700603	*Klebsiella pneumoniae* WT, clinical isolate	Wild-type	ATCC
Acinetobacter baumannii ATCC 19606	*Acinetobacter baumannii* WT, clinical isolate	Wild-type	ATCC
Pseudomonas aeruginosa ATCC 27853	*Pseudomonas aeruginosa* WT, clinical isolate	Wild-type	ATCC
Escherichia coli ATCC 25922	*Escherichia coli* WT, clinical isolate	Wild-type	ATCC

Streptomyces *goldiniensis* ATCC 21386	Wild type, aurodox-producing strain	Wild-type	ATCC
Streptomyces collinus tü 365	Wild type, kirromycin-producing strain	Wild-type	Gift from Wolfgang Wolheben
Streptomyces ramocissimus ATCC 27529	Wild type, kirromycin-producing strain	Wild-type	Gift from Gilles van Weizel
Streptomyces collinus REM1	pESAC-13A-aurodox	AprR,	This study
Streptomyces collinus REM2	pMS82tuf2	HygR	This study
Streptomyces collinus REM3	pESAC-13A-aurodox + pMS82_tuf2	AprR, HygR	This study
Streptomyces collinus REM4	pMS82	HygR	This study
Streptomyces collinus REM5	pESAC-13A-aurodox + pMS82	AprR, HygR	This study
Streptomyces collinus REM6	pESAC-13A	AprR,	This study
Streptomyces collinus REM7	pESAC-13A + pMS82	AprR, HygR	This study
Streptomyces collinus REM30	pIJ6902	AprR,	This study
Streptomyces collinus REM31	pIJ6902aurM*	AprR,	This study
Streptomyces albus G148	Wild Type *Streptomyces albus*	Wild-type	Gifted from Iain Hunter
Streptomyces albus REM8	pESAC-13A-aurodox	AprR,	This study
Streptomyces albus REM9	pMS82tuf2	HygR	This study
Streptomyces albus REM10	pESAC-13A-aurodox + pMS82_tuf2	AprR, HygR	This study
Streptomyces albus REM11	pMS82	HygR	This study
Streptomyces albus REM12	pESAC-13A-aurodox + pMS82	AprR,HygR	This study
Streptomyces albus REM13	pESAC-13A	AprR,	This study
Streptomyces albus REM14	pESAC-13A + pMS82	AprR,HygR	This study
Streptomyces venezuelae ATCC 10712	*Streptomyces venezuelae* wild type	Wild-type	DSMZ 40230
Streptomyces venezuelae REM15	pESAC-13A-aurodox	AprR,	This study
Streptomyces venezuelae REM16	pMS82tuf2	HygR	This study
Streptomyces venezuelae REM17	pESAC-13A-aurodox + pMS82_tuf2	AprR, HygR	This study

Streptomyces venezuelae REM18	pMS82	HygR	This study
Streptomyces venezuelae REM19	pESAC-13A-aurodox + pMS82	AprR,HygR	This study
Streptomyces venezuelae REM20	pESAC-13A	AprR,	This study
Streptomyces venezuelae REM21	pESAC-13A + pMS82	AprR, HygR	This study
Streptomyces coelicolor M1152	*Streptomyces coelicolor* 'superhost'	Δact Δred Δcpk Δcda ΔrpoB	(Gomez-Escribano and Bibb, 2012)
Streptomyces coelicolor REM22	pESAC-13A-aurodox	AprR,	This study
Streptomyces coelicolor REM23	pMS82tuf2	HygR	This study
Streptomyces coelicolor REM24	pESAC-13A-aurodox + pMS82_tuf2	AprR, HygR	This study
Streptomyces coelicolor REM25	pMS82	HygR	This study
Streptomyces coelicolor REM26	pESAC-13A-aurodox + pMS82	AprR, HygR	This study
Streptomyces coelicolor REM27	pESAC-13A	AprR,	This study
Streptomyces coelicolor REM28	pESAC-13A + pMS82	AprR, HygR	This study
Streptomyces coelicolor REM29	pESAC-13A-aurodox ΔAurP1	AprR, HygR	This study

***Table 2-B*: Plasmids and Phage Artifical Chromosomes used in experimental work**

Plasmid name	Description	Resistance	Reference
pESAC-13A	Phage Artifical Chromosome, ΦC31 integrase, *lacZ*,	apramycin, carbenicillin	(Sosio, 2001)
pESAC-13A- AurI	Phage Artifical Chromosome, ΦC31 integrase, triple positive clone for aurodox cluster	apramycin	This study
pESAC-13A-Aur2	Phage Artifical Chromosome, ΦC31 integrase, triple positive clone for aurodox cluster	apramycin	This study
pESAC-13A-AurIΔAurP1	Phage Artifical Chromosome, ΦC31 integrase, triple positive clone for aurodox cluster, *ΔaurP1*	apramycin, hygromycin	This study
pMS82	ΦBT1 integrase	hygromycin	(Gregory, Till and Smith, 2003)
pMS82_*tuf2*	ΦBT1 integrase, *S. goldiniensis tuf2* expression vector	hygromycin	This study
PIJ6902	ΦC31 integrase containing vector with Thiostrepton resistance and a MCS containing and upstream thiostrepton inducible promoter	apramycin, thiostrepton	(Hao *et al.*, 2014)
PIJ6902_*aurM**	ΦC31 integrase, AurM* expression vector	apramycin, thiostrepton	This study
PIJ10790	λ-RED (gam, bet, exo), *cat, araC, rep101ts*	chloramphenicol	(Gust *et al.*, 2003)
PIJ10700	Re-direct template	hygromycin	(Gust *et al.*, 2003)
pAJR70	Empty plasmid vector	chloramphenicol	(Roe. *et al.*, 2004)
pler-gfp	*ler*:gfp transcriptional fusion	chloramphenicol	(Roe. *et al.*, 2004)
prpsM-gfp	*rpsM*: gfp transcriptional fusion	chloramphenicol	(Roe. *et al.*, 2004)
pVS45	Plasmid encoding Ler with arabinose inducible promoter.	Ampicillin	
precA-gfp	*recA*:: gfp translational fusion. Chloramphenicol resistance marker.	chloramphenicol	(Sperandio *et al.*, 2000a)
ptir-gfp	*tir*::gfp translational fusion. Chloramphenicol resistance marker.	chloramphenicol	(Sperandio *et al.*, 2000a)

Table 2-C: Antibiotics and concentrations used in experimental work

Antibiotic	Class	Stock concentration (mg/ ml)	Working Concentration (µg/ ml)	Solvent
Aurodox	Elfamycin, elongation factor-tu binding antibiotic	1	100	DMSO
Ampicillin	β-lactam antibiotic	100	100	H_2O
Apramycin	Aminoglycoside antibiotic	50	50	H_2O
Carbenicillin	β-lactam antibiotic	100	100	H_2O
Chloramphenicol	N-dichloroacylphenylpropanoid antibiotic	25	25	96% EtOH
Hygromycin B	Substituted Aminoglycoside antibiotic	100	100	H_2O
Kanamycin	Aminoglycoside antibiotic	50	50	H_2O
kirromycin	Elfamycin, elongation factor-tu binding antibiotic	1	100	DMSO
Nalidixic acid	DNA gyrase inhibitor antibiotic	25	25	0.1M NaOH
Thiostrepton	Macrocyclic thiopeptide antibiotic	25	0.1	MeOH

2.2 Growth media

Unless otherwise stated, all liquid media was prepared in 200 ml batches in 500 ml Duran bottles and sterilised by autoclaving at 121^{o}C at 100 kPa. For Enteropathogenic and Enterohaemorrhagic *E. coli* and *C. rodentium* growth, MEM-HEPES and DMEM media were purchased from Sigma- Aldrich and L-glutamine levels were adjusted to 0.568 g l^{-1} before use. International *Streptomyces* Project media (ISP2-7) were purchased from ATCC ©. Bacterial culture media used in this study was prepared according to *Table 2-D* which provides instructions for the preparation of 1 L of liquid media. For solid media, 20 g l^{-1} of agar was added prior to autoclaving. To prepare soft agar, agar component was reduced to 10 g per litre.

Table 2-D: Growth media used in this study

Media	Components	pH
Modified Yeast Extract Malt Extract (YEME) (Hoskisson, Hobbs and Sharples, 2001)	10 g glucose, 5 g Oxoid malt extract, 5 g Difco Bacto-peptone, 3 g yeast extract	7.3
Soya Flour Mannitol (SFM) (Hobbs *et al.*, 1989)	20 g mannitol, 20 g soya flour (Holland & Barrett), 10 mM CaCl	7.1
Aurodox Production Media (AP) (Berger *et al.*, 1973b)	20 g glycerol, 1 g glucose, 5 g corn steep powder, 2 g meat extract, 1 g soybean flour	7.2
Lysogeny Broth (LB) (Sambrook, Fritsch and Maniatis, 1989)	10 g tryptone, 5 g NaCl, 5 g yeast extract	7.5
Glucose Yeast Malt Medium (GYM) (Kieser *et al.*, 2000a)	10 g malt extract, 4 g glucose, 4 g yeast extract	7.2
V6 Media (Pavlidou *et al.*, 2011) (Marinelli, 2009)	10 g glucose, 5 g corn steep solids, 5 g yeast extract, 5 g peptone, 3 g casein hydrolysate, 1.5 g NaCl	7.5
Difco nutrient media (Difco)	15g Difco Nutrient Broth powder, reduce agar to 6 g/L for Soft Nutrient Agar	6.8

2.3 Growth and maintenance of bacteria

Unless otherwise stated, *Streptomyces* species were grown at 30 ^{o}C and all *E. coli* strains and ESKAPE pathogens were grown at 37 ^{o}C. Unless otherwise stated, liquid cultures of bacteria were prepared in 50 ml of media in 250 Erlenmeyer flasks. For *Streptomyces,* baffled flasks were used. To prepare *Streptomyces* spore stocks, a single colony from a *Streptomyces* culture was used to inoculate a confluent lawn of growth on agar. After 7-10 days of growth and after the formation of spores, 5 ml of sterile 50% glycerol solution was applied to the plate and the spores were disturbed. The resulting spore solutions were passed through sterile 5 ml syringes containing cotton wool to remove cell debris. An equal volume of sterile dH_20 was added to the suspensions, resulting in a final glycerol concentration of 25%. For glycerol stocks of *E. coli, C. rodentium* and the ESKAPE pathogens, 10 ml overnight cultures were inoculated from a single colony. 1 ml of this culture was removed and added to 1 ml of sterile 50% glycerol solution resulting in a final glycerol concentration of 25%. Glycerol stocks were preserved at – 80 ^{O} C for long term storage.

2.4 Microscopy of *Streptomyces goldiniensis*

2.4.1 Stereomicroscopy

An SMZ1500 stereomicroscope coupled with a mercury arc lap (Nikon, Japan) was used to image *Streptomyces* colonies grown on ISP2-7 for 10 days at 4x magnification. Images were acquired using a DFK 33UX264 CMOS camera (The Imaging Source Europe GmbH, Germany).

2.4.2 Widefield Epifluorescence Microscopy

A dense *Streptomyces* spore stock (1 x10^5 spores per ml) was diluted 1/100 and 5 µl was inoculated on to MS media under a coverslip inserted at a 45^{o} angle. Cultures were grown for two and five days respectively. To visualise the bacteria, Schwedock Fluoroscein Wheat Germagglutinin/Propidium Iodide staining was used (Schwedock

et al., 1997) . Cells were fixed with 500 µl fixative per coverslip and allowed to dry before being washed in PBS, air dried and rehydrated in PBS for five minutes. Coverslips were then incubated with 2 mg/ml lysozyme in RTP buffer for two minutes before being washed with PBS. The coverslips were incubated for a further five mins in 2% BSA before being transferred to a dark room. The samples were then incubated with Stain (2µg/ml fluorescein-WGA, 10µg/ml propidium iodide, 2% BSA in PBS) and left for three hours before being washed with 10µl/ml propidium iodide in PBS eight times. Two washes with equilibrium solution molecular probes solution were executed, followed by a final 5-minute wash step and the removal of excess solution. The slides were then coated in 8 µl of 40% glycerol solution before being mounted to a slide and sealed with clear nail varnish.

A TE2000 inverted wide field microscope (Nikon, Japan), coupled to a mercury arc lamp (Nikon, Japan) was then used to acquire images of the strains. Images were acquired with a 100x/1.3NA Plan Fluor Ph3 DLL objective lens (Nikon, Japan) at 490, 512 and 617nm respectively. An Orca-100 CCD (Hamamatsu, Japan) camera was used to detect the wavelengths. ImageJ was used to analyse and scale the images.

2.5 Gravimetric dry-weight analysis of *S. goldiniensis* growth

S. goldiniensis spores were pre-germinated in Aurodox Production Media for 48 hours and inoculated 1/100 in to 12 Erlenmeyer flasks containing 50 ml GYM media. These were then incubated at 30^0C for twelve hours before 1 ml culture from three 'growth' flasks was removed and transferred on to a Whatman glass microfiber filter paper that had been previously dried in a desiccator at room temperature. A vacuum system was used to pull the culture through a Buchner funnel, leaving only cell mass on the filter paper. Filter papers were dried for 12 hours in a 55 ^{O}C oven before being weighted. Additionally, at each time point, one flask was removed from the incubator, centrifuged at 4000 xg to remove cells, and the supernatant was frozen at -20 ^{o}C for future LC

MS and bioactivity analysis (Kieser *et al.*, 2000a). This was then repeated every six hours until growth reached stationary phase. Growth curves were plotted using Graphpad Prism v6. 0.2.

2.6 Solvent extraction of aurodox from *Streptomyces* cultures

Culture supernatants from *Streptomyces* cultures were filter sterilised through a 0.2 µM Millipore™ filter. Using glass equipment throughout, the cultures were mixed with equal volume of Chloroform and a separation funnel was used to remove the lower, solvent phase. Samples were then dried under nitrogen gas and extracts were dissolved in ethanol for LC-MS and bioactivity analyses.

2.7 Liquid chromatography Mass Spectrometry

LC-MS was carried out on an Aglient 1100 HPLC instrument in conjunction with a Waters Micromass ZQ 2000/4000 mass detector. Electrospray ionization (ESI) was used in all cases. The RP-HPLC analysis was conducted on a Zorbax 45mm x 150mm C18 column at 40 ˚C. Ammonium acetate buffers were used as follows: Buffer A (5 mM Ammonium acetate in Water) and Buffer B (5mM Ammonium acetate in acetonitrile). Positive and negative electrospray methods were applied with of 100 to 1000 AMU positive, 120-1000 AMU negative with a scanning time of 0.5 seconds. The UV detection was carried out at a wavelength of 254 nm.

2.8 Aurodox Bioassays

Overnight cultures of ESKAPE pathogens were inoculated in 10 ml of LB media in sterile universal tubes. These were then inoculated to 3 ml of Soft Nutrient Agar to an OD_{600} of 0.005 and poured on to a sterile petri dish containing 20 ml of LB agar. Subsequently, 10 µl of culture extract was added to a sterile filter paper disk and left to dry. These were added to each plate in triplicate. Cultures were incubated at 37 ^{o}C overnight and the zones of inhibition were measured using a ruler.

2.9 Standard curves for determination of aurodox activity against *K. pneumoniae.*

Serial 1 in 2 dilutions were carried out using aurodox standard in DMSO (Enzo). Bioassays (as described in 2.7) were carried out and the results were plotted using Origin Pro. The line equation was then used to estimate aurodox concentration in culture extracts.

2.10 Whole Genome Sequencing of *Streptomyes goldiniensis*

2.10.1 Genomic DNA extraction from *Streptomyces* (Kieser et al., 2000a)

A 50 ml liquid culture was centrifuged at 4200 rpm for 10 minutes to prepare a pellet. The pellet was the resuspended 3 ml TES buffer containing 4 mg/ml lysozyme and 50 µg/ml RNase and was incubated for 1 hour at 37˚C. Furthermore, 120 µl of 10% SDS and 75ul of 1 mg/ml Proteinase K were added and the solution was incubated at 55˚C until the broth cleared (for *S. goldiniensis* this step was 3.25hrs). Subsequently, 3 ml of phenol-chloroform was added, and the falcon tube and it was centrifuged at 4200 rpm for 10 minutes and the upper layer removed. This process was repeated three times. After the final extraction, equal volume of ethanol was added to precipitate the DNA and stored at -20 ^{o}C overnight. The tube was then centrifuged again at 4200 rpm and the pellet resuspended in nuclease-free water. DNA concentrations and ratios were determined using Nanodrop ™ spectrophotometer.

2.10.2 Sequencing of *S. goldiniensis* using Oxford Nanopore minION™

A genomic DNA library was prepared in accordance with the Nanopore™ 1D ligation protocol using 1 µg genomic DNA. From this, 75 µl was loaded into a MinION© flow cell and placed in the MinION sequencer, connected to a PC by USB cable. The sequencer was run for twelve hours, and the raw data was converted to sequence data via MinKnow base calling software, provided by Nanopore™.

2.10.3 Sequencing of *S. goldiniensis* using Illumina Sequencing Technology

A mycelial preparation of *S. goldiniensis* was sent to MicrobesNG, University of Birmingham. The genome was sequenced using Illumina HiSeq 2500 sequencing platform. DNA extraction and sequencing were carried out according to the following protocol provided by Microbes NG.

Three beads were washed with extraction buffer containing lysozyme (or lysostaphin for Staphylococcus sp.) and RNase A, incubated for 25 min at 37^{o}C. Proteinase K and RNaseA were added and incubated for 5 min at 65^{o}C. Genomic DNA was purified using an equal volume of SPRI beads and resuspended in EB buffer.

DNA was quantified in triplicates with the Quantit dsDNA HS assay in an Eppendorf AF2200 plate reader. Genomic DNA libraries were prepared using NextEra XT Library Prep Kit (Illumina, San Diego, USA) following the manufacturer's protocol with the following modifications: two nanograms of DNA instead of one were used as input, and PCR elongation time was increased to 1 min from 30 seconds. DNA quantification and library preparation were carried out on a Hamilton Microlab STAR automated liquid handling system. Pooled libraries were quantified using the Kapa Biosystems Library Quantification Kit for Illumina on a Roche light cycler 96 qPCR machine. Libraries were sequenced on the Illumina HiSeq using a 250bp paired end protocol. Reads were adapter trimmed using Trimmomatic 0.30 (Bolger, Lohse and Usadel, 2014) with a sliding window quality cut off of Q15.

2.10.4 Sequencing of *S. goldiniensis* using PacBio

Carried out by Nu-omics at the University of Northumbria, high quality *S. goldiniensis* gDNA (>1.8 A260/280) was size sheared using Covaris g-Tube (~9 Kbp) and the size distribution was visualised using Agilent Bioanalyzer. To remove single stranded DNA and to provide blunt ended double stranded DNA, ExoVII was used. Blunt hairpin SMRTbell adapters were ligated to the polished DNA and ExoIII treatment was used

to remove failed ligation products. Sequencing primers were annealed to the adapter templates via the PacBio Polymerase binding reaction. The PacBio Sequel instrument was used to sequence this library using ten-hour movie capture. BAM files of CCS reads were generated for use for assembly in HGAP4.

2.10.5 Assembly of the *Streptomyces goldiniensis* genome

The *Streptomyces goldiniensis* genome was assembled using the k-mer dependent method provided by SPAdes™ (Bankevich *et al.*, 2012). For combined assemblies, PacBio contigs provided by NU-omics were used as trusted contigs. From the combined PacBio/Illumina/Nanopore assembly (77 contigs), AutoMLST (Alanjary, Steinke and Ziemert, 2019) was used to determine the closest neighbour. This genome (*Streptomyces bottropensis)* was then used as a scaffold for MeDuSa (Bosi *et al.*, 2015). Quality analysis was carried out using QUAST (Gurevich *et al.*, 2013).

2.10.6 Annotation of *Streptomyces goldiniensis* genome

Annotation of the *Streptomyces goldiniensis* genome was created using Prokka and can be accessed via NCBI (embargoed until June 2020). Whole genome comparison maps were created using .gff files generated by Prokka (Seemann, 2014) of *Streptomyces goldiniensis* and *Streptomyces bottropensis* using CCT function of CGviewer (Grant, Arantes and Stothard, 2012). Orthovenn2 (Xu *et al.*, 2019) was used for comparison of coding sequences between strains.

2.11 Bioinformatic identification and characterisation of the aurodox Biosynthetic Gene Cluster

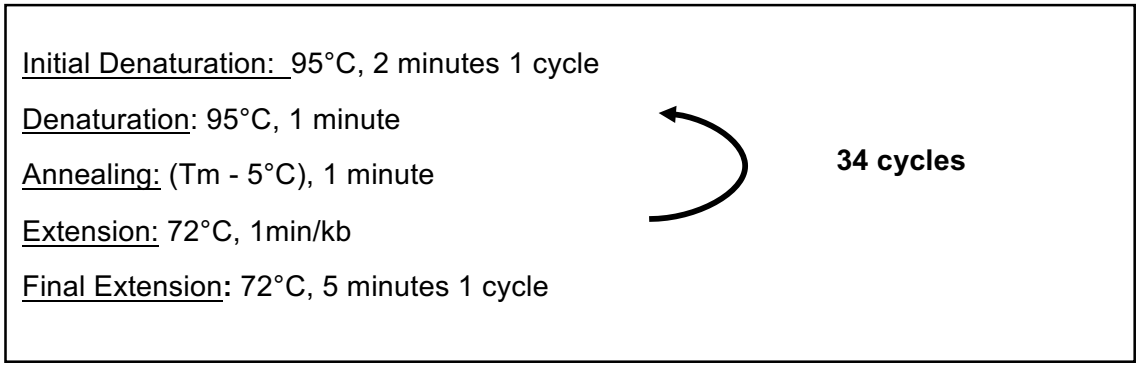
To identify putative biosynthetic gene clusters from the *Streptomyces goldiniensis* genome, antiSMASH bacterial version 5.0.0 (Machado, Eva C. Sonnenschein, *et al.*,

2015) was used. Identification of kirromycin homologs and of the modular enzymatic domains in the aurodox biosynthetic gene cluster was carried out using the PKS/NRPS detailed functional domain annotation, also from antiSMASH bacterial version 5.0.0 (Machado, Eva C. Sonnenschein, *et al.*, 2015). In addition, antiSMASH was used to generate gene cluster comparison maps between *Streptomyces goldiniensis, Streptomyces collinus* and *Streptomyces bottropensis.* Artemis Comparison Tool (Carver *et al.*, 2005) was used to create aurodox/kirromycin gene cluster percentage similarity maps. Easy Fig 2.0 was used to generate aurodox and kirromycin comparison maps.

2.12 Polymerase Chain Reaction

Primers used in this study were designed using SnapGene™. A list of these primers can be found in *Table 2-E.*

For diagnostic and confirmation PCRs, GoTaq ® green DNA polymerase (Promega) was used under the following conditions:



For cloning and gene deletions, KOD HotStart polymerase (Novagen) was used.

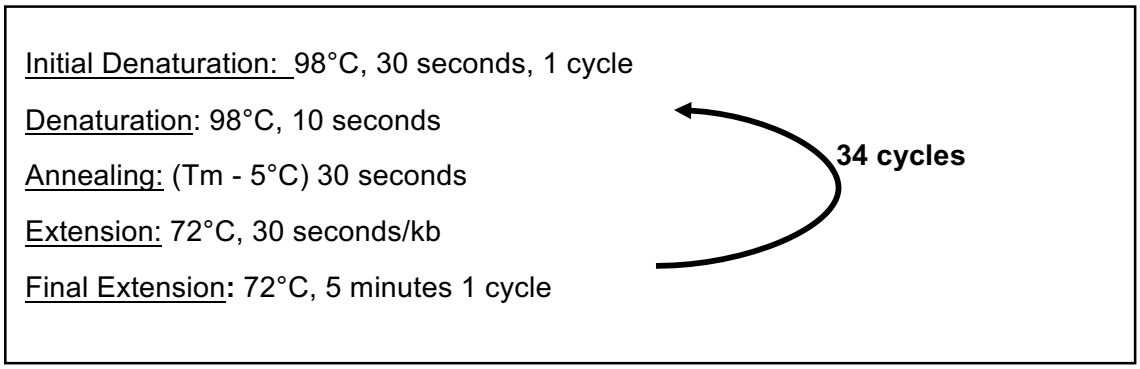


Table 2-E: **List of primers used in this study**

Primer name	Sequence	Tm (^{O}C)
Cluster_poscheckF	CCAGACGCAGGTCCGCTTCGGACG	68.1
Cluster_poscheckR	CCATCGTGGGGATCGCAG	60.9
Cluster_poscheckF	AGGATGTTCCAGTCGGCTCTCACTCCG	65.8
Cluster_poscheckR	CGAGGTCGCCCGGCATGTGGA	61.2
ClusterCheckAF	GAGCTACCAACTCTTTTTCC	50.9
ClusterCheckAR	CCGCGGATGACCGCGTGGACGTGA	69.2
ClusterCheckBF	CATCGGCAGCCTCTTCTCCGC	63.7
ClusterCheckBR	CGGCTCATGCCGGAAGTCTCCCGA	66.8
ClusterCheckCF	TTCCGCTCCCTGGATCGGCAT	66.7
ClusterCheckCR	AGCACTGAACGCCGGCACGGAAGA	67
MTGibsonF	ATTCGAGCTCGGTACGGATGGAAATCCTCAAGCTCGCC	70.9
Methyltransferase_GibsonR	GTCAGAGAAGGGAGCGGACATCAGGAGGCACGCAGGGA	71.8
PIJ6902MT_checkF	GAGATCGGCGTCGTACAT	54.3
PIJ6902MT_checkR	GAGCTACCAACTCTTTTTCC	50.9
HrdB_f	GAGCGGGGAAAGGCTGA	57.8
HrdB_r	GCACTGACCATCAGCGT	55.4
RedirectPKS1F	TTCCATGCCGCGATTCCGGCCGGACCGTGTGTCCGCATGATTCCGGGGACCGTCGA	76.5
RedirectPKS1R	GAACGAGCGCGCCGGGGCGCGTGAGGAAGGACCGGGTTATGTAGGCTGGAGCTGCTT	76.2
Hyg_F	GTGACACAAGAATCCCTGTTACTTCTC	62.7
Hyf_R	CTCGAACACCTCGAAGTCCAGGAA	65.9
PKSredirectcheckF	GGCCTGATCATGCGGTTGGGCGCT	69.5
PKSredirectcheckR	TGGAGGACATCCGGCGCT	66

2.13 Agarose Gel Electrophoresis

Unless otherwise stated, gels containing 1.2 % agarose were used for agarose gel electrophoresis. Agarose concentration was increased to 1.5 % for PCR products smaller than 500 bp. Agarose was dissolved in TAE buffer (40 mM Tris-acetate, 1 mM EDTA) using aghh microwave. The gel mixture was then cooled before Ethidium Bromide was added to a final concentration of 100 µg/ml and poured into a casting tray, sealed with rubber stoppers. DNA samples were mixed with 6x blue/orange loading dye (Promega) and relevant molecular weight markers were chosen according to band size (Promega). Throughout this project, agarose gels were run using Bio-Rad gel dock systems, with 1 x TAE used as the buffer. In general, gels were run at 90 V for between 45 and 90 minutes, with the run time being increased for larger gels. DNA was visualised and imaged under UV excitation using Syngene Bioimaging Ingenius trans-illuminator. For extraction of DNA for cloning, a UV trans-lluminator was used on low energy and a scalpel was used to cut the required band from the gel. DNA was extracted from the bands using the Wizard® SV Gel and PCR Clean-Up kit (Promega).

2.14 Restriction digests

Restriction enzymes were sourced from Promega and the appropriate digest conditions and buffers were determined via the Promega Restriction Enzyme Tool (https://www.promega.co.uk/resources/tools/retool/).

2.15 *In silico* cloning and plasmid construction

In silico cloning and plasmid construction was carried out using SnapGene ™. Plasmid maps were constructed and annotated using the 'show features' function and exported as tifs.

2.16 Plasmid constriction using Gibson assembly

Plasmid backbones of interest were linearised using desired restriction enzymes and Gibson overlap primers were designed according to the Gibson Assembly Protocol (Gibson, 2011). To select plasmids with successful insertions, plasmids were transformed into NEB® 5-alpha competent cells (High Efficiency) according to manufacturer's protocol. For blue/white screening, transformed cells were plated on to LB containing 80 µg/ml X-gal and 0.3 mM IPTG. The presence of the insert of interest was confirmed by colony PCR and sequencing (Eurofins, Cologne, Germany).

2.17 Tri-Parental Mating (Jones *et al.*, 2013)

Overnight cultures of *E. coli* DH10β (neo) containing plasmid of interest (supplied by Bio S&T) *E. coli* Top10 containing the driver plasmid pR9604 and ET12567 with appropriate antibiotics (see *Table 2-2*) were used to inoculate fresh, individual cultures of each *E. coli* strain. These were grown at 37^{O}C and 220 rpm until an OD_{600} of 0.6 was reached. Cells were washed three times in fresh LB by centrifugation at 4000 rpm to remove antibiotic and resuspended in 1 ml of LB media. Subsequently, 20 µl of each strain was plated in the centre of an LB plate and incubated at 37^{O}C overnight. The following day, the colony was re-streaked on to the required antibiotic selection and the presence of the conjugating plasmid in ET12567 was confirmed via colony PCR.

2.18 Preparation of Electrocompetent *E.coli*

A 10 ml overnight culture of *E. coli* was prepared in LB media + relevant antibiotics. The following day, cultures were diluted 1/50 and grown to an OD_{600} of 0.6. The culture was chilled on ice for five minutes before being centrifuged a 4000 xg for five minutes. The cell pellet was then resuspended in 25 ml of sterile , ice cold 50% glycerol solution and centrifuged at 4000 xg for 5 minutes. The washing procedure was then repeated

twice. The pellets were resuspended in 400 µl of sterile water with 10% glycerol (Sambrook, Fritsch and Maniatis, 1989).

2.19 Intergenic conjugation of integrating vectors into *Streptomyces* (Kieser *et al.*, 2000b)

An overnight culture of ET12567/pUZ8002 transformed with the relevent plasmid was used to innoculate a 50 ml culture which was grown to an OD_{600} of 0.6. Cultures were centrifuged at 4000 xg and washed twice with 10 ml Lysogeny Broth to remove antibiotics before being resuspended in 1 ml of Lysogeny Broth in an Eppendorf tube. Approximately 10^6 *Streptomyces* spores were each added to 500 µl LB media in an Eppendorf tube before being heat-shocked in a 50 ºC heat block for ten minutes. The heat shocked spores were then left for ten minutes at room temperature to recover and 500 µl of ET12567/pUZ8002/ plasmid was added to each of the Eppendorf tubes containing the spores. The tube was then vortexed briefly and most of the supernatant was removed before the cells were re-suspended in the residual liquid. Cultures were plated out on MS agar containing 10mM of Magnesium Chloride and incubated for 20 hours. The plate was then overlaid with 1 ml sterile water containing 0.5 mg Nalidixic Acid to kill remaining *E. coli* and incubation was continued for a further 42 hours. Exconjugant colonies were then picked off and innoculated in patches on on selective MS agar containing Hygromycin/Apramycin and Nalidixic acid at a concentration of 50 µg/ml (Sambrook, Fritsch and Maniatis, 1989).

2.20 Minimum Inhibitory Concentration Assays

Serial dilutions (one in two) of aurodox from 100 µg/ml to 1.5 µg/ml, as well as an untreated control were prepared in GYM media in a 96-well plate. Cultures were inoculated with 10 µl of a dense *Streptomyces* spore stock, sealed with a Breath-easy™ membrane and grown at 30°C and 220 rpm for three days. The optical density

was measured using a Synergy HT plate reader and minimum inhibitory concentrations were calculated.

2.21 Redirect-PCR targeting system in *Streptomyces* for lambda-*red* mediated gene deletions on Phage Artifical Chromosomes

Primers for redirect deletion mutants were designed according to PCR targeting system in *Streptomyces coelicolor* A3(2), derived from (Datsenko and Wanner, 2000). For homologous recombination, these overlap primers were used with pIJ10700 as a template for amplification of the disruption cassette which was purified from agarose gel electrophoresis. PCR products were extracted and transformed into *E. coli* Top10/PIJ10790/pESAC-13A-AurI*,* plated on to LB + Apramycin, Hygromycin and Carbenicillin with arabinose to induce the lambda-*red* recombinase and incubated at 37^{o}C to facilitate the loss of the temperature sensitive PIJ10790-lamba-red recombinase plasmid. Colonies were then picked and used in tri-parental mating to transfer the disrupted PAC to ET12567. Disrupted PACs were then conjugated into *Streptomyces coelicolor* M1152 according to 2.17 (Gust *et al.*, 2003)

2.22 Analysis of secreted proteins from EPEC, EHEC and *C. rodentium*

Overnight cultures were used to inoculate each strain into the appropriate prewarmed medium to an initial OD_{600} of 0.05. The strains were grown with increasing concentrations of aurodox to an OD_{600} of 0.7 to 0.9 before being centrifuged at 3,800 rpm for 10 min to separate the cell mass. Cell lysates were prepared using BugBuster protein extraction buffer (Merck, NJ). The supernatant was removed, and proteins were precipitated by the addition of trichloroacetic acid to a concentration of 10% and incubated overnight at 4°C. The suspension was centrifuged at 3,800 x g for 1 h to form a protein pellet which was suspended in 100 µl of Tris-HCl (pH 8.0). The samples were then mixed 1:1 with loading dye (Novex) and run on a 4 to 12% SDS-PAGE gel at 120 V for 1 h. Gels were subsequently stained with Coomassie brilliant blue stain

(Novex) and destained in water overnight. When required, bands were excised for subsequent in-gel digestion and analysis. Proteins analysed by tandem mass spectrometry were assigned a MASCOT score to indicate the probability of the identification being correct.

2.23 Analysis of the effect of aurodox on in vitro growth and cell viability.

Overnight cultures of EHEC (TUV93-0), EPEC (E2348/69), and *C. rodentium* (ICC186) were used to inoculate 200-ml Erlenmeyer flasks containing 50 ml of LB broth with or without 5 g/ml of aurodox to an initial OD_{600} of 0.05. This was carried out in triplicate. At each time point, 100 µl of culture was removed and diluted 1/10 in LB medium, and the OD600 was measured by spectrophotometry. The standard error mean of optical densities was calculated, and a standard curve was plotted using GraphPad Prism. The trendlines were fitted by joining the points and error bars were plotted as standard deviations from the means. P-values were determined using an unpaired t-test. For cell viability assays, each strain was grown to and OD_{600} of 3.0 both with and without aurodox, and CFU were determined through the serial dilution method.

2.24 Infection of HeLa epithelial cells with EHEC O157:H7.

A 24-well tissue culture plate containing coverslips was seeded with 10^7 HeLa cells in MEM-HEPES containing 0.562 g/litre of L-glutamine and 10% foetal calf serum. These were left to grow overnight. In parallel, two cultures of TUV93-0 prpsM::gfp (with or without 5 g/ml of aurodox) were grown in 5 ml volumes under T3SS-inducing conditions until an OD_{600} of 0.6 was reached. The HeLa cells were washed once with phosphate-buffered saline (PBS) before 500 µl of fresh MEM-HEPES (with or without 5 g/ml of aurodox) was added, in addition to 15 µl of bacterial culture. The plate was then centrifuged at 250 x g for 3 min and was left to incubate initially for one hour. Serial dilutions of the inoculum were carried out and spotted onto LB agar to

determine the CFU (as described above). After the initial hour, cells were washed three times with sterile PBS, and 600 µl of fresh MEM-HEPES with or without 5 g/ml of aurodox was added before incubation for a further three h. To quantify the adherent EHEC bacteria on treated and untreated cells, the cells were washed four times with sterile PBS and lysed through the addition of 1% Triton in PBS and incubation at room temperature for 10 min. Serial dilutions were spotted onto LB agar to determine the adherent CFU. Colonisation efficiency was calculated as a percentage of the initial inoculum. To visualize the infections using wide-field epifluorescent microscopy, cells on coverslips were fixed with 4% paraformaldehyde (20 min at room temperature) before permeabilisation with 0.1% Triton in PBS (5 min at room temperature) and staining with phalloidin-Alexa Fluor 555 (Thermo Fisher; 1 in 500 dilution and 1 h at room temperature). Cells were mounted using Vectashield with 4=,6-diamidino-2-phenylindole (DAPI) and sealed with clear nail polish. Images were acquired at a magnification of x 400 on a Zeiss Axioimager M1. For representative images, 11 z-slices of 0.55 m were acquired before deconvolution using default settings in Zeiss Zen Pro software. Host cells were classified as infected if any bacteria were seen to be associated with them.

2.25 mRNA extraction

EHEC (TUV93-0) was cultured as described above in triplicate, both with and without 5 µg/ml of aurodox. The cultures were mixed with 2x volumes of RNAprotect reagent (Qiagen) at room temperature before being centrifuged at 3,800 x g to harvest a cell pellet. RNeasy kit (Qiagen) was used to extract total RNA before TURBO DNase (Ambion, Carlsbad, CA) was used to remove genomic DNA. Furthermore, a MICROBExpress mRNA enrichment kit (Ambion) was used to enrich samples for mRNA. The quality of the mRNA was determined using an Agilent 2100 bioanalyzer at the University of Glasgow Polyomics Facility. Transcriptome analysis using RNA

sequencing. cDNA synthesis and sequencing were performed at the University of Glasgow Polyomics Facility (Illumina NextSeq 500), obtaining 75- or 100-bp single end reads. Treated and untreated samples were sequenced in triplicate. FastQC (Babraham Bioinformatics, Cambridge, UK) was used to for quality control. Reads were trimmed accordingly using CLC Genomics Workbench (CLC Bio, Aarhus, Denmark). Trimmed reads were mapped to the EDL933 reference genome (NCBI accession number NC_002655.2), allowing for 3 mismatches per read and at least 5 reads per feature. Analysis of differential expression was performed using the Empirical analysis of DGE tool, which implements the EdgeR Bioconductor tool. Differentially expressed genes were identified using a positive or negative absolute fold change of x 1.5 and a corrected P value of 0.05 (false-discovery rate of 5%). Gene Ontology functional grouping was summarized according to information available on Colibase and the RegulonDB. Figures were generated using GraphPad Prism and Microsoft Excel. The RNA sequence reads in this study have been deposited in the European Nucleotide Archive under study PRJEB29967.

2.26 *In vitro* GFP fusion reporter assays

Electrocompetent EHEC, EPEC, and *C. rodentium* cells were transformed with the promoter-gfp reporter plasmids listed in *Table 2-B*. The transformants were inoculated in to 10 ml of the appropriate medium and cultured as previously stated. Samples were removed at the desired time points to take measurements of the OD_{600}. To determine the overall fluorescence, 200 µl of culture was transferred into a black 96-well plate and fluorescence was read with excitation of 485 nm and emission at 550 nm, using the FLUOstar Optima fluorescence plate reader system (BMG Labtech, UK). This was carried out in triplicate. To determine the normalized fluorescence value, the background fluorescence intensity from untransformed cells was first subtracted and the subsequent values were normalized by dividing fluorescence by

OD_{600}. The represents the means for the three samples, and error bars represent the standard deviations (Roe *et al.*, 2003).

2.27 Overexpression of Ler

Electrocompetent EHEC cells were transformed with pVS45 plasmid containing the *ler* gene under the control of an arabinose-induced promoter and grown on 50 µg/ml of ampicillin. A transformant colony was then used to prepare an overnight culture, which was subsequently inoculated into T3SS-inducing conditions as described above. At an OD_{600} of 0.2, 2% arabinose was added. The supernatant protein fractions were then precipitated and analyzed using SDS-PAGE as previously explained. Proteins for Western blot analysis were run on a 4 to 12% bis-Tris gel and transferred to an Amersham ECL nitrocellulose membrane (GE Healthcare) using the Nupage Novex gel transfer system (Invitrogen). Blocking was then carried out using 5% skimmed milk powder in PBS-Tween (PBST). The membrane was then incubated with the anti-Tir primary antibody overnight in 1% milk powder–PBST buffer at a 1:1,000 dilution. Antibodies for GroEL (Abcam) were used as a control for bacterial cell lysis at the same dilution. The following morning, the membrane was washed three times with PBST for 10 minutes before being incubated for 1 h with anti-mouse horseradish peroxidase (HRP)-conjugated secondary antibody at a 1:2,000 dilution in 1% milk in PBST. The membrane was again washed three times with PBST. Finally, the membrane was developed with SuperSignal West Pico chemiluminescent ECL (Thermo Fisher) substrate for 5 min before being transferred to a dark room in a cassette and exposed to X-ray film for 30 s (Zaslaver *et al.*, 2006).

2.28 Detection of Shiga toxin expression in *C. rodentium* DBS100 by Western blotting

Adapted from (Mallick *et al.*, 2012), *Citrobacter rodentiu*m DBS100 containing the Stx phage was cultured in 2 ml of DMEM at 37°C and 200 rpm until an OD600 of 0.6 was reached. The whole-cell fraction was removed by centrifugation at 3,800 x g for 10 min. Secreted proteins were harvested as previously described. Proteins for Western blot analysis were run on a 4 to 12% bis-Tris gel and transferred to an Amersham ECL polyvinylidene difluoride (PVDF) membrane (GE Healthcare, Chicago, IL) using the Nupage Novex gel transfer system (Invitrogen, Carlsbad, CA). No blocking was carried out. The membrane was incubated with anti-Stx (beta subunit) antibody (1/1,000 concentration) overnight at 4°C and washed three times with PBST. The membrane was incubated for 1 hour with anti-mouse HRP-conjugated secondary antibody at a 1:2,000 dilution in 1% milk in PBST and washed three times with PBST. Finally, the membrane was developed with SuperSignal West Pico chemiluminescent ECL (Thermo Fisher) substrate for 5 min before being transferred to a dark room in a cassette and exposed to X-ray film for 30 min. The sequence reads in this paper have been deposited in the European Nucleotide Archive under study PRJEB29967.

Chapter 3 Physical and genomic characterisation of the aurodox producer, *Streptomyes goldiniensis.*

3.1 Introduction

Streptomyces goldiniensis was isolated from Bermudian soil in 1972 as part of a project to isolate novel antimicrobials targeting Gram-positive bacteria (Berger *et al.*, 1973b). During this work, aurodox was identified and some basic characterisation of *S. goldiniensis* was performed. The studies which have followed however, have largely focussed on the chemical properties of aurodox and its mechanism of action, rather than on the biology of the producing organism (Maehr *et al.*, 1978; Lutz Vogeley *et al.*, 2001). This has resulted in gaps in the literature regarding species morphology and physiology. Furthermore, the fermentation protocols published in seminal papers largely use historic and undefined media components, which are not readily found in modern research laboratories and fail to take advantage of recent developments in fermentation for *Streptomyces*. Moreover, these protocols are designed for industrial procedures and are largely incompatible for use with the small scale (50 ml – 2L) fermentation and extraction processes required for this project. Hence, one aim of this chapter is to use stereo and epifluorescence microscopy to characterise the morphology of *S. goldiniensis.* In addition, *S. goldiniensis* growth under fermentation will be analysed with the initial aim to optimise aurodox production for the small scale utilised in the laboratory.

Despite the transition of *Streptomyces* research into the genomic era, prior to this study, the whole genome of *S. goldiniensis* had never been sequenced and the genes encoding aurodox production had not been identified. Therefore, using the most advanced available next generation sequencing technologies, specifically PacBio,

Nanopore and Illumina, a central aim of this chapter was to sequence the whole genome of *Streptomyces goldiniensis* and use this to identify the specialised metabolite biosynthetic gene clusters using the genome mining software, antiSMASH with a view to identifying the aurodox BGC enabling further characterisation of biosynthesis.

3.2 Results

3.2.1 Visualisation of *S. goldiniensis* phenotypes on International *Streptomyces* Project medias using Stereomicroscopy.

In order to visualise the morphological characteristics of the strain when grown in differential media conditions and identify optimal media for strain propagation, *S. goldiniensis* was grown on International *Streptomyces* Media (ISP2-7; Shirling & Gottlie, 1966) for ten days (Figure 3-1). Following incubation colonies were imaged using Stereomicroscopy. It was observed that growth of *S. goldiniensis* on ISP4 (Inorganic Salt Starch Medium; Shirling and Gottlieb, 1966) and ISP7 (Tyrosine Agar; Ronald, 2006) was optimal for *S. goldiniensis* sporulation based on the appearance of mature spores. Furthermore, the presence of clear zones around colonies on ISP7 is indicative of L- tyrosine hydrolysis, which is a physiological feature historically used as a taxonomic marker to distinguish *Streptomyces* from other member of the Steptomycetacae family such as *Nocardia.* It can also be observed that sporulation is limited on rich media such as ISP3 (Oat Meal Agar; Shirling and Gottlieb, 1966) and ISP5 (Glycerine Asparagine Agar Base; Shirling and Gottlieb, 1966), which in an effect observed among many *Streptomyces* species as nutrient limitation is an important environmental signal in the cascade involved in the sporulation genes. Finally, the appearance of black colonies on ISP6 (Peptone Yeast Extract Iron Agar; (Shirling and Gottlieb, 1966) is indicative of melanin production and it is therefore probable that, like many *Streptomyces* species, *S. goldiniensis* encodes a melanin biosynthetic gene cluster.

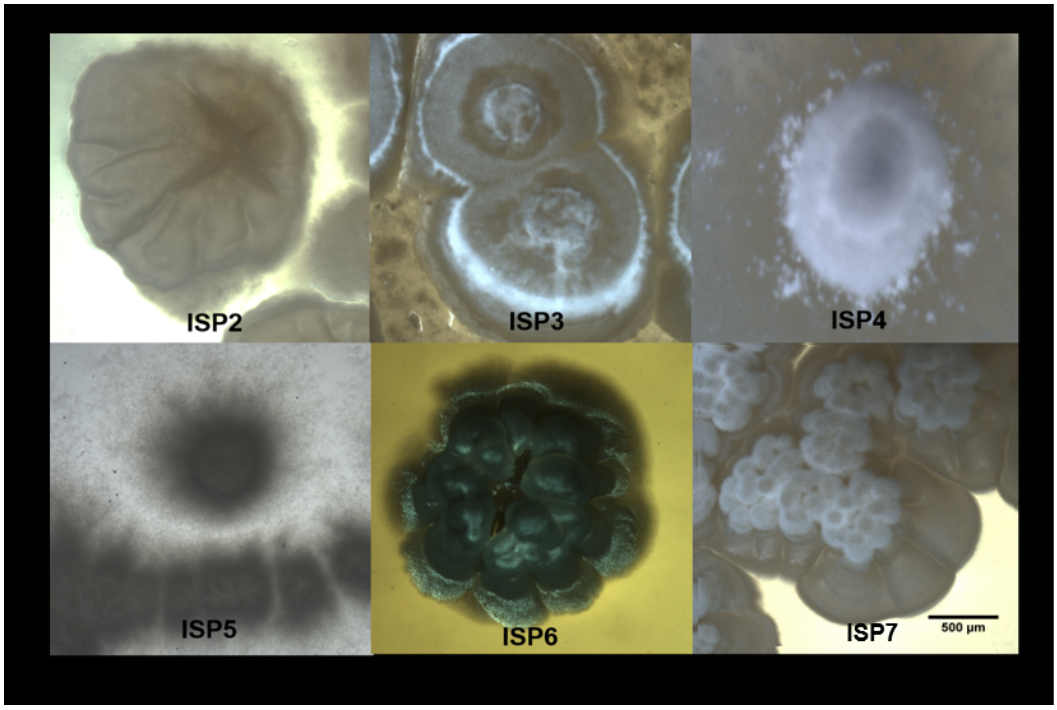
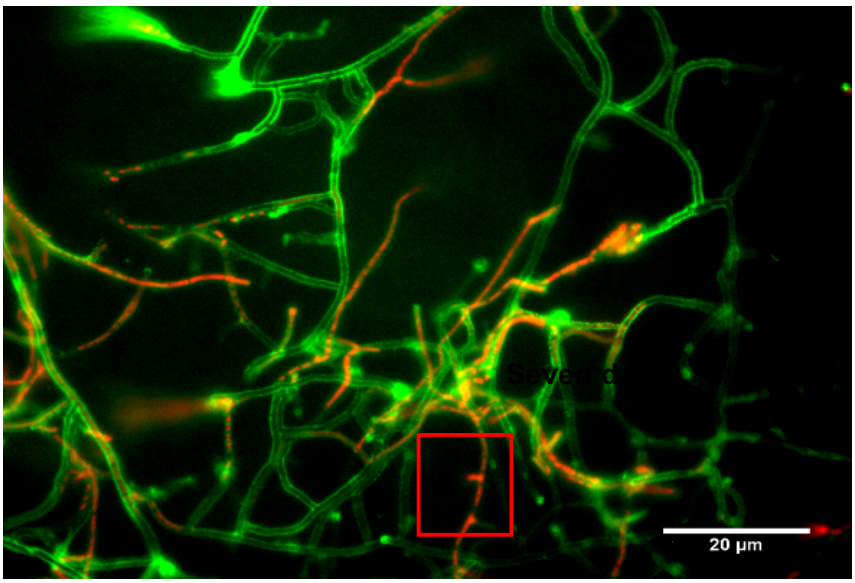


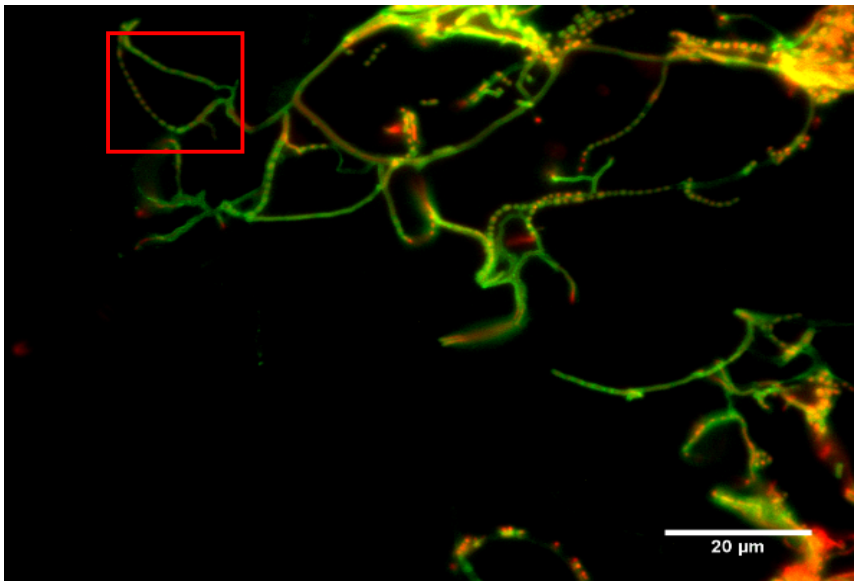
***Figure 3-1**:* **Stereomicroscopy images of *S. goldiniensis* on ISP2-7.** Images were taken at 4 x magnification after ten days of incubation at 30 ^{o}C. An SMZ1500 stereomicroscope coupled with a mercury arc lap was used to generate images which were acquired using a DFK 33UX264 CMOS camera system.

3.2.2 Epifluorescence microscopy imaging of *S. goldiniensis* reveals sporulation after seven days of growth.

In order to visualise the stages of the *S. goldiniensis* lifecycle, Widefield Epifluorescence Microscopy using Fluoroscein Wheat Germagglutinin (FWG)/ Propidium Iodide (PI) staining was used to acquire images after two and seven days growth on solid media (Figure 3-2). In these results, the presence of dense hyphal growth can be observed after two days of incubation (Figure 3-2A). As signified by the red box, there is branching of the vegetative mycelia which is indicated by cross wall formation. The cross walls appear green as they have been stained by Fluoroscein Wheat Germagglutinin which binds to the peptidoglycan present. In these images, nucleic acids are stained by propidium iodide and therefore appear red. Hence, the beginning of aerial hyphae septation can be observed as there are multiple distinct chromosomes in chains along hyphal bodies (Schwedock *et al.*, 1997). However, it is clear that these are not mature spore chains with spore walls not yet becoming thickened with peptidoglycan. Conversely, after seven days of growth (Figure 3-2B) mature spores are present. These bodies are co-stained with both PI and FGW, and the rounded spores are comparable to light microscopy images of *S. goldiniensis* spore chains previously published by (Berger *et al.*, 1973a). In conclusion, these results show that *S. goldiniensis* sporulation occurs after seven days of growth on MS media and the morphology of the strain is typical of *Streptomyces* species and has a growth cycle akin to that of the model organism *Streptomyces coelicolor* (Kieser *et al.*, 2000b).



Two days



Seven days

Figure 3-2: **Widefield Epifluorescence Microscopy Images of *S. goldiniensis* after two and seven days growth on MS agar.** Fluoroscein Wheat Germagglutinin/Propidium Iodide staining were used to stain peptidoglycan and nucleic acids respectively. Images were acquired using A TE2000 inverted wide field microscope coupled to a mercury arc lamp was then used to acquire images of the strains. Images were acquired with a 100x/1.3NA Plan Fluor Ph3 DLL objective lens at 490, 512 and 617nm respectively. An Orca-100 CCD camera was used to detect the wavelengths. ImageJ was used to analyse and scale the images.

3.2.3 Effect of aurodox on the growth of ESKAPE pathogens

Following the discovery of aurodox by Berger and co-workers, the inhibitory effect of the compound on the growth of several bacterial species was characterised. The results reported include zone of inhibition data for bioactivity indicator organisms including *B. subtilis* and *E. coli* as well as murine *in vivo* activity profiles for pathogens such as *K. pneumoniae* and *Staphylococcus aureus* (Berger *et al.*, 1973b). However, these results are not directly comparable to current studies as there are ambiguities in the methods and changes to taxonomic nomenclature. Therefore, it was important to gain an understanding of the effect of aurodox on defined, clinically relevant pathogens. Hence, the effects of aurodox on the growth of the ESKAPE pathogens were determined through bioassays and the diameter of the associated zones of inhibition was measured. These bacteria are commonly used as test species in drug discovery due to their clinical relevance and decreased susceptibility to common antimicrobials (Katz and Baltz, 2016). Consistent with the findings of Berger et al, *Figure 3-3A* shows that aurodox exhibits strong effects on the growth of *K. pneumoniae, E. coli* and *S. aureus.* Berger *et al* also reported limited effects on *Pseudomonas aeruginosa* growth, which is consistent with the results of this study. This work also shows that the strongest effects of aurodox on the bacteria tested were on *Acinetobacter baumannii* with a zone of inhibition measuring 12 mm. There was no effect of aurodox on the growth of *E. faecalis.*

In order to estimate the concentration of aurodox in fermentation extracts, a standard curve was plotted based on the zone of inhibition diameter and aurodox concentration. It was determined that *K. pneumoniae* is a suitable indicator species for use in these diagnostic bioassays as the effects of the compound are significant and concentration-dependent. (*Figure 3-3*). Therefore, a range of aurodox concentrations from 2 mg/ml to 10 µg/ml were used in bioassays and the diameter of the resulting

zones of inhibitions were determined. These were then plotted using OriginPro and the equation of the line was generated. This was used to estimate aurodox concentrations in extracts throughout this project.

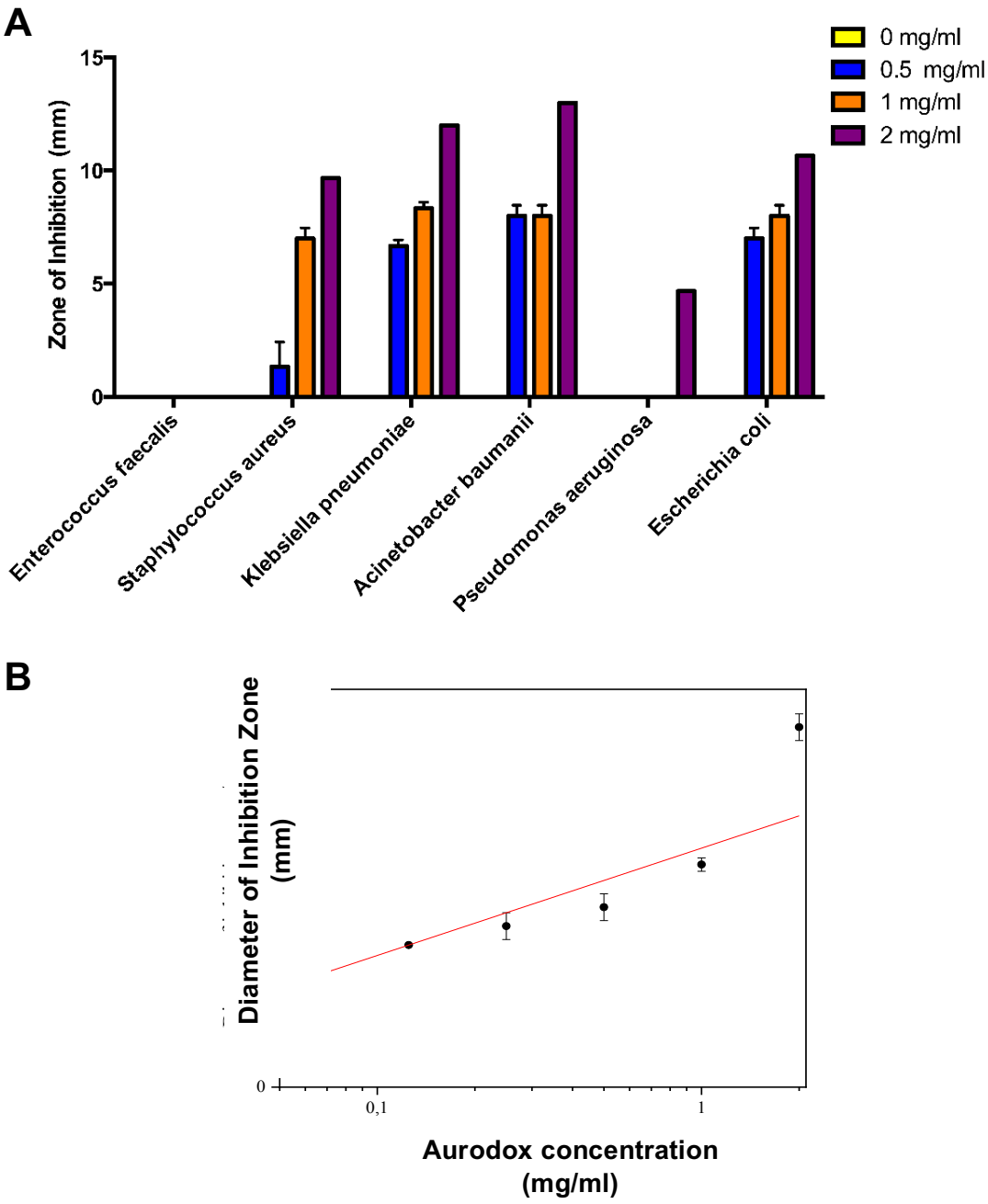


Figure 3-3: **Effect of Aurodox on Growth of ESKAPE pathogens**. (A) Zone of Inhibition associated with decreasing concentrations of aurodox standard. Error bars represent standard deviation from the mean. Aurodox was suspended in DMSO, with DMSO used as 0 mg/ml control. (B) Aurodox dilution series bioassay against *K. pneumoniae.* Points represents mean zone of inhibition (mm) from triplicate experiments. Error bars represent standard deviation from the mean.

3.2.4 Determination of suitable media for laboratory aurodox production by *S. goldiniensis* fermentation.

In previous work by Berger et al, fermentation media for aurodox production was designed and optimised to maximise aurodox yields from *S. goldiniensis* fermentations. As a result, this media is described in the aurodox patent (Berger *et al.*, 1973a). However, this media is undefined and is optimised for use in industrial fermenters, and hence, preliminary yields from shake flask fermentations using this media were low. Furthermore, the presence of Soya Flour in this 'Aurodox Production' media causes difficulty when separating biomass from the media for further procedures such as DNA or RNA extraction. Therefore, batch cultures in various defined media were carried out in order to optimise small-scale aurodox production. The rationale behind the media chosen was their validated use in industrial *Streptomyces* fermentations. Lysogeny Broth was used as a 'negative control', as *Streptomyces* typically do not reach high levels of biomass when grown in LB media, (Kieser *et al.*, 2000b) and therefore the aurodox yield is likely to be low. In *Figure 3-4A*, the extracts from these samples are shown, and in *Figure 3-4B,* LC-MS reveals the presence of aurodox in the culture extracts from GYM and surprisingly, LB media. In addition, a peak at the correct retention time of 7.14 minutes (based on aurodox standard trace) could be found for aurodox production media, but no corresponding aurodox mass was detected in either the negative or positive ion scan in the MS analysis. No aurodox was detected in the extract from the V6 fermentation. These results can be confirmed, as when tested in *K. pneumoniae* bioassays, GYM and LB extracts produced the largest zones of inhibition (10.67 mm and 5.67 mm respectively), shown in *Table 3A.* The results of these bioassays were used to estimate aurodox concentrations via the equation generated in *Figure 3-3.* Results from both LC-MS and bioassay analysis confirm that of the media chosen, GYM facilitates the highest level of aurodox production. In addition, this HPLC analysis

reveals that the GYM extract has a similar level of purity to the aurodox standard. Therefore, GYM was used in aurodox cultures throughout this project.

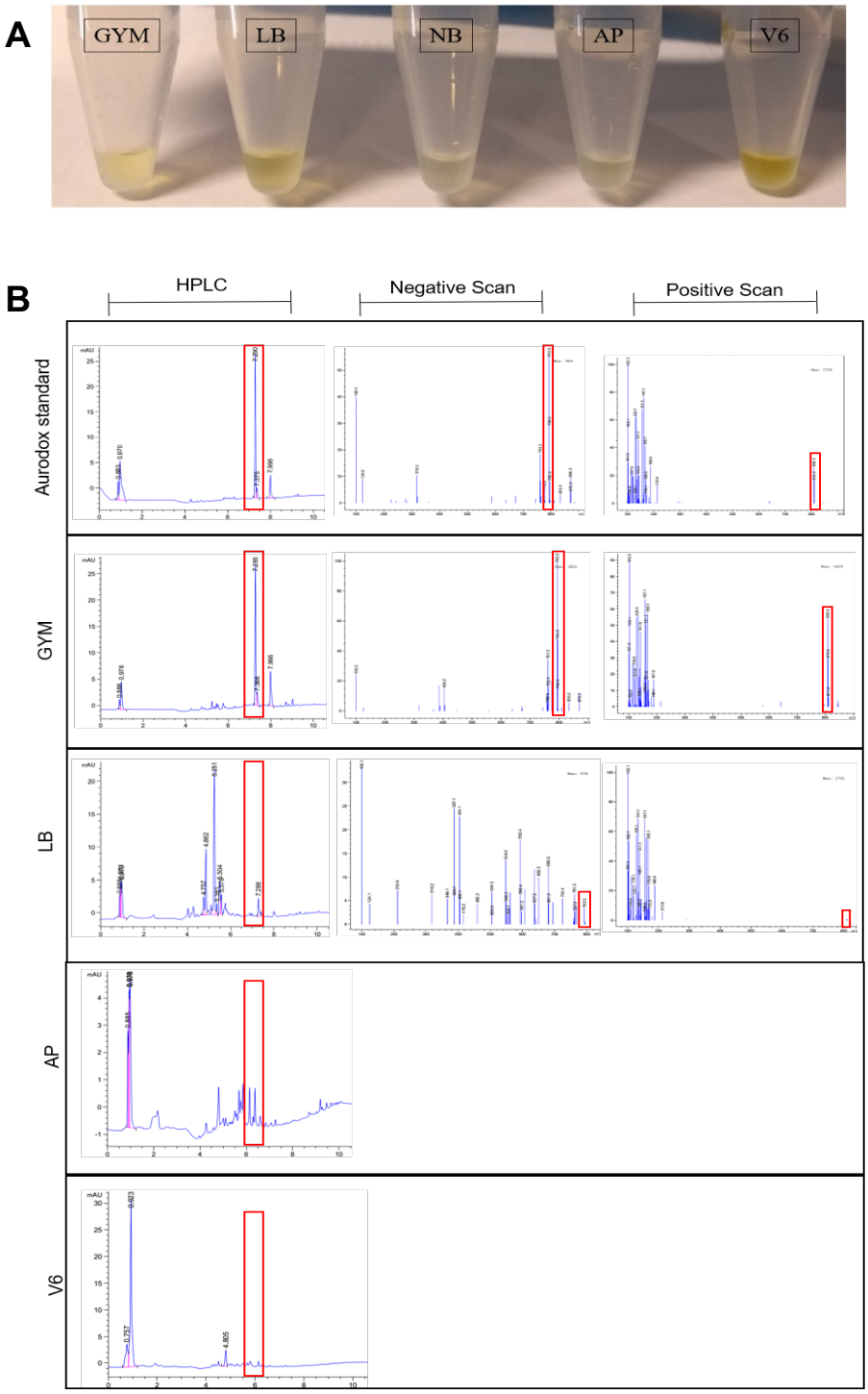


***Figure 3-4*: Analysis of aurodox production efficiency in Aurodox in GYM, LB and AP medium**. (A) Images of chloroform extract from each media (indicated). (B) High Performance Liquid Chromatography traces, and negative and positive scan Mass Spectrometry traces of aurodox standard versus chloroform extracts. Peaks corresponding to aurodox are indicated by red boxes.

Table 3-A: **Summary of estimated aurodox concentrations in extracts from shake flask fermentations**

Medium	Diameter of inhibition (mm)	Aurodox concentration (mg/ml)
Aurodox production media	3.17	0.041
GYM	10.67	3.968
LB	5.67	0.188
Nitrate Broth	5.00	0.125
V6	-	-
DMSO	-	-

3.2.5 Aurodox is produced during stationary phase of *S. goldiniensis* growth in GYM.

In order to understand the growth of *S. goldiniensis* in GYM media, dry cell weight growth analysis was carried out. This method was selected as the widely used Optical Density measurement commonly used to measure bacterial growth is inaccurate for organisms which exhibit mycelial growth (Hobbs *et al.*, 1989). Primarily, this data was used to determine at which point the stationary phase of the growth cycle was reached and hence, predict the onset of secondary metabolite gene expression (de Lima Procópio *et al.*, 2012). From the results of this experiment, shown in *Figure 3-5,* it can be determined that in GYM media, *S. goldiniensis* reaches stationary phase of growth after approximately 20 hours. To complement this data, chloroform extracts from time points throughout the *S. goldiniensis* life cycle were prepared and analysed through bioassays and LC-MS. In extracts taken from 50 ml *S. goldiniensis* shake-flask cultures, aurodox was not detected until `154 hours of growth. This was confirmed by a characteristic 7.2 minute peak in the HPLC trace, in addition to detection of the aurodox mass in the negative and positive scans from Mass Spectrometry and the onset of bioactivity.

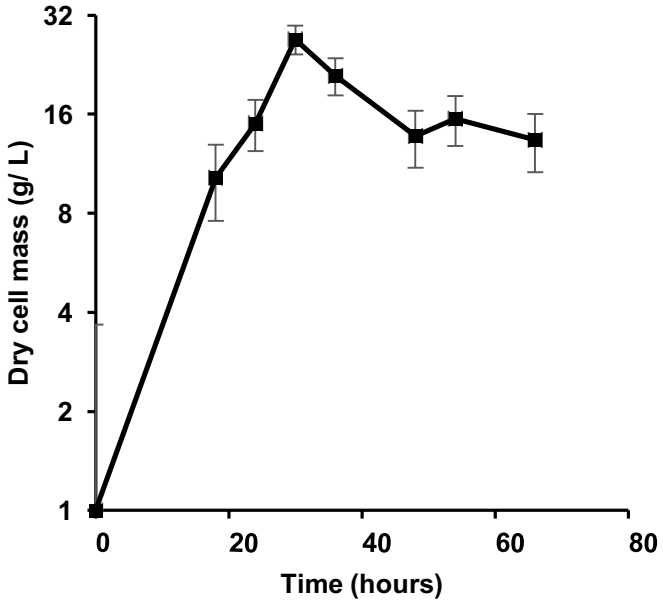


Figure 3-5: Dry weight was determined through gravimetric separation of cell mass from media. Experiments were carried out in triplicate and mean dry cell weight was plotted using Microsoft Excel, with series trendline function. Error bars represent standard deviation from the mean. Specific growth rate was calculated between and 12 hours and was determined as 0.138 $h^{-1.}$.

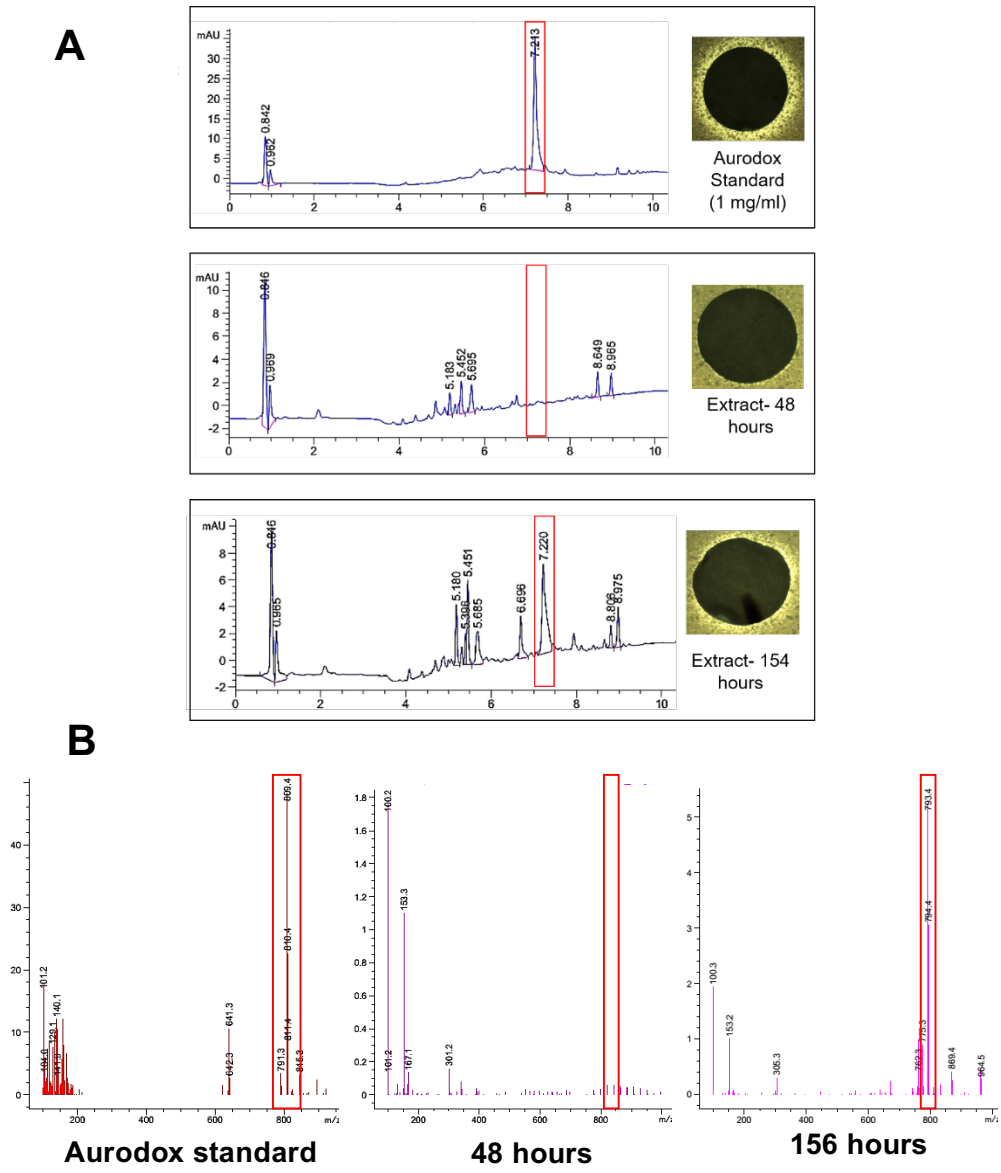


Figure 3-6: **Analysis of Aurodox production by *S. goldiniensis* over time**. (A) HPLC traces of *S. goldiniensis* chloroform extracts and the corresponding disk diffusion bioassay result against *K. pneumoniae* bioassay. (B) Mass Spectrometry (negative scan mode) results for retention time 7.22 minutes (aurodox peak determined from standard).

3.2.6 Whole genome sequencing of *Streptomyces goldiniensis*

Streptomyces genomes have several unique features which can result in increased difficulty in obtaining a whole genome sequence. These include high GC content, linear chromosomes and plasmids, and repetitive regions, particularly around polyketide synthase and non-ribosomal peptide synthetase gene clusters (Gomez-Escribano *et al.*, 2015) . Therefore, a multi-facetted approach to sequencing the *S. goldiniensis* genome was used, with several next generation sequencing technologies deployed. A schematic summary diagram of the processes used in this study, for sequencing *S. goldiniensis* can be found in *Figure 3-7.*

3.2.6.1 Sequencing with Illumina MiSeq

Illumina MiSeq provides paired-end sequencing which produces read lengths of 2 x 300 bases. Although this sequencing technique can readily produce reliable sequence lengths of up 500 nucleotides, this is largely insufficient to resolve the repetitive sequences in large PKS and NRPS genes, especially in *de novo* sequencing (Gomez-Escribano, Alt and Bibb, 2016). Therefore, with a view to gaining preliminary *S. goldiniensis* sequencing data and establish neighbours for use in assembly software such as MeDuSa, Illumina MiSeq sequencing of the *S. goldiniensis* genome was carried out. In the service provided by Microbes NG, a minimum of 30 x coverage of the genome was guaranteed, and the trimmed reads were returned, as well as a *de novo* assembly using SPades. In order to obtain an overview of the general quality of the Illumina MiSeq results obtained, the raw .fasta files from the SPades assembly were uploaded to the Quality assessment Tool for Genome Assemblies (QUAST) server. This tool establishes fundamental information including closely related strains and contig number. In addition, QUAST can be used to calculate N50 and L50 values- which are intrinsically indicative of genome quality. The N50 value describes the number which half of the genome sequence is in contigs

larger than or equal to, and therefore, a large number is suggestive of a high quality genome. The L50 value is defined as the number which corresponds to the number of contigs required to make up half of the genome, hence, a small number is signifies high assembly quality. A summary of these results can be found in *Table 3-B.*

From the results of QUAST analysis shown in *Table 3-B,* the GC content of *Streptomyces goldiniensis* is estimated to be 70.3%, which is typical of species from the *Streptomyces* family. In addition, genetic relatedness to both *Streptomyces scabeii* and *Streptomyces avermitillis MA-4680* was predicted. Despite obtaining 30 x coverage of the genome, the contig number of was 2854, splitting in to 2430 scaffolds. This high number of contigs, in addition to a low N50 value (13064) and a high L50 value (223) demonstrates that as expected, Illumina MiSeq sequencing alone is insufficient for the purpose of obtaining a high quality *S. goldiniensis* genome for identification of the aurodox gene cluster.

3.2.6.2 Sequencing with Pacific Biosciences SMRT technology (PacBio)

To enhance the data obtained from Illumina sequencing, genomic DNA from *S. goldiniensis* was sent to the NU-omics facility at Northumbria University to be sequenced with Pacific Bioscience Single Molecule Real Time sequencing technology (PacBio). In bacteria with low GC content such as *E. coli,* PacBio sequencing alone can be used for whole genome sequencing (Rhoads and Au, 2015), however, high GC content and repetitive regions in *Streptomyces* increase the likelihood of errors and therefore, a combination of short and long read sequencing is recommended for *de novo* assembly. First demonstrated by Gomez-Escribano *et al* for the sequencing of *Streptomyces leeuwenhoekii,* (Gomez-Escribano *et al.*, 2015), combined Illumina MiSeq and PacBio assemblies have been used for the effective assembly of Actinomycete genomes for mining for specialised metabolites. PacBio sequencing

complements Illumina MiSeq data, as in contrast to the short Illumina reads, PacBio sequencing provides read lengths of approximately 20 kb. In total, 19,194 reads were returned in .fastq format, resulting in a coverage depth of 42 x. Furthermore, an assembly was generated from these reads through HGAP. From this, the number of contigs returned was 117, with a much improved N50 of 193883 and L50 of 16 *(Table 3-B)*.

3.2.6.3 Oxford Nanopore Sequencing

In addition to Illumina MiSeq and PacBio sequencing, Oxford Nanopore Sequencing was carried out using the MinION portable sequencing platform. This technology provides long reads of up to 500 kb (Mikheyev and Tin, 2014), which can support the bridging of gaps between the contigs generated by Illumina and PacBio sequencing. The primary results of minION sequencing can be found in *Figure 3-8*. With a total of 53096120 bases read, minION sequencing provided a coverage depth of 5.8 x, and therefore this method was insufficient to generate a standalone assembly, with assembly software typically requiring a minimum of 10 x coverage. However, with 70 reads greater than 20 kb, and reads reaching a maximum of 121 kb, these reads helped to improve the overall *S. goldiniensis* genome assembly quality (*Table 3-8B)*.

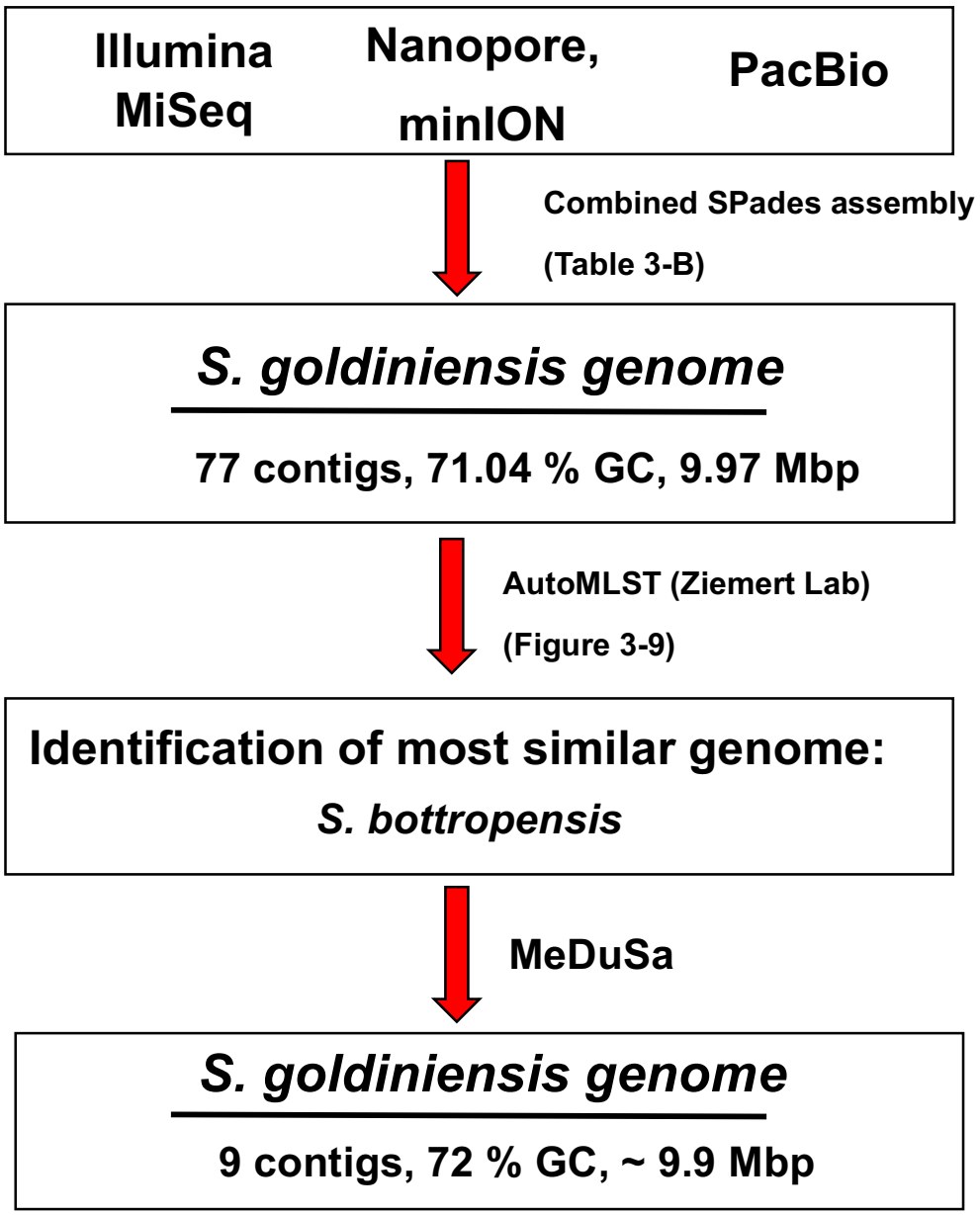
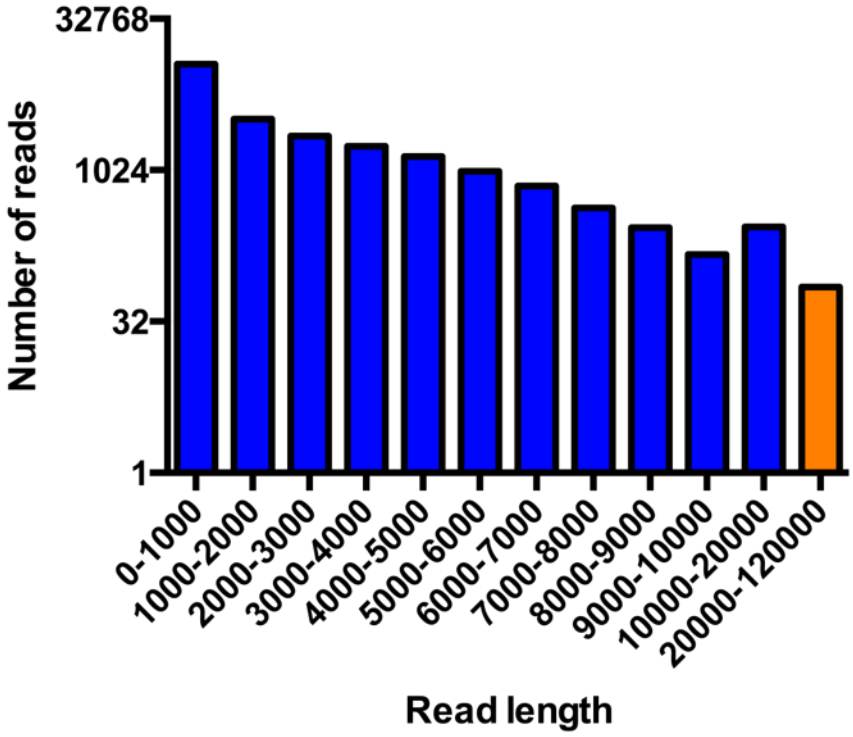


Figure 3-7: **Schematic Diagram describing the processes involved in the sequencing and assembly of the *S. goldiniensis* genome.**

A



B

Minimum read length (bases)	1
Maximum read length (bases)	121380
Number of reads	33346
Mean read length (bases)	1580
Total bases read	53096120

Figure 3-8: **Summary of results from sequencing of *S. goldiniensis* with Oxford Nanopore minION technology**. (A) Histogram describing read lengths of sequenced reads. Reads are grouped according to length. (B) Summary chart describing overall output from minION sequencing

.

3.2.6.5 Combined Assembly using St Petersburg genome assembler (SPades).

To construct a comprehensive assembly of the *S. goldiniensis* genome, a combined assembly using PacBio, Illumina MiSeq and Oxford Nanopore MinION data was executed. After optimisation of SPades parameters, the most effective assembly was generated using the SPades k-mer dependent method (Koren *et al.*, 2016), with a k-mer value of 55. The 117 contigs returned from HGAP assembly of PacBio reads were used as trusted contigs, and the Illumina and MinION reads were mapped. A summary of this assembly can be found in *Table 3-B.* Again, there was a marked improvement in the assembly, indicated by a reduction in contig number to 77, an increase in N50 value to 421525 and a reduction of L50 to 7. As this was determined to be the most effective assembly possible from the sequencing data obtained, this assembly was carried forward for phylogenetic analysis using Auto Multi-Locus Sequence Tree, known as AutoMLST (Alanjary, Steinke and Ziemert, 2019).

AutoMLST is an automated web server designed to model the evolution of strains by highlighting their similarity to known species. AutoMLST generates a phylogenetic tree with 100 members, which are chosen based on Average Nucleotide Identity (ANI)- generated by the MASH algorithm (Alanjary, Steinke and Ziemert, 2019). In addition to the essential, orthologous genes used in traditional multi-locus sequence typing analysis, AutoMLST includes genes which are involved in secondary metabolite biosynthesis. The inclusion of these genes has resulted in improved phylogenetic resolution in the *Actinomycetes* (Alanjary, Steinke and Ziemert, 2019). In *Figure 3-9,* the output of this software can be observed. From this phylogenetic tree, which clusters species with their close relatives, it is apparent that the closest related strain to *S. goldiniensis* is *Streptomyces bottropensis.* However, the species boundary between these two strains is validated, with an average nucleotide identity of 90.7%, which is below the 95% threshold required for strains to be considered the

same species (Figueras *et al.*, 2014). The evolutionary relatedness of *S. goldiniensis* and *S. bottropensis* will be discussed in detail in *Chapter 4.* It is somewhat surprising however, despite producing almost identical and unique molecules, the kirromycin producer, *Streptomyces collinus* can be found on a separate clade from *S. goldiniensis*. These results suggest that the species are more distantly related, with an ANI of 84.0%.

3.2.6.6 Scaffold based assembly of *S. goldiniensis* using MeDuSa.

To optimise the assembly of the *S. goldiniensis* genome, MeDuSa, a Multi-Draft Based Scaffolder was used. This approach relies on the use of a closed genome to order and orientate contigs, as well as bridge contig gaps. The most closely related strain- as determined by AutoMLST- *Streptomyces bottropensis* was uploaded as the comparison genome. This returned an assembly with an improved contig number of ten. When the quality of this assembly was determined using QUAST, one contig was found to have significantly lower GC content than the others. Results from a BLASTn search revealed that this contig corresponded to the 18S rRNA sequence of a *Penicillium sp* mould and as a contaminant could therefore be excluded from further analysis. As a result, the *S. goldiniensis* assembly used for further analysis in this thesis comprises of nine contigs. A graphical representation of the GC content, contig number, N50 and L50 values for each assembly can be found in *Figure 3-10.* A tabular summary of the RAST and BLAST results from each stage of the assembly process can be found in *Table 3-B*. In *Figure 3-11,* the contigs which comprise this final assembly are depicted as a bandage diagram, and a summary of the QUAST results from this assembly can be found.

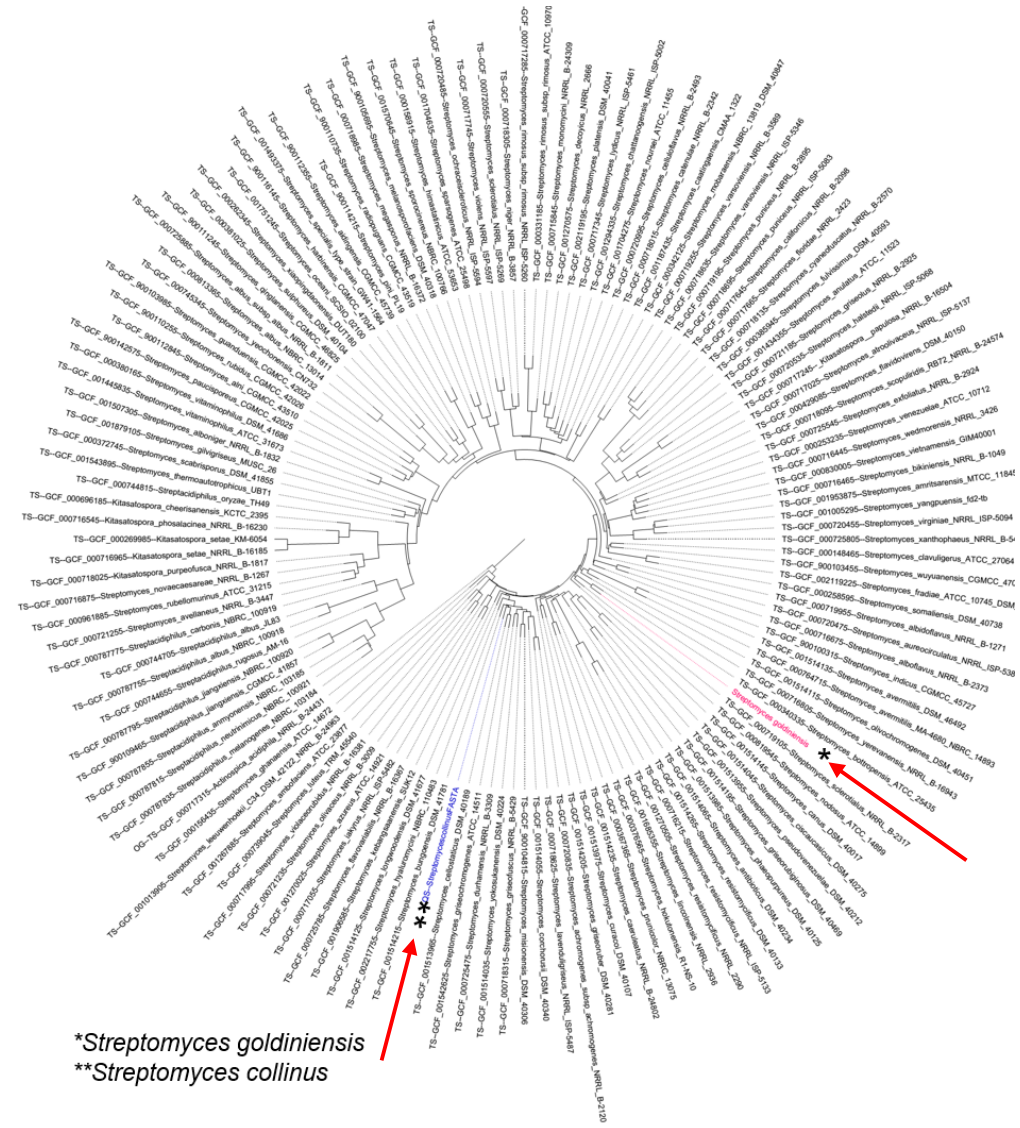


Figure 3-9: **Phylogenetic tree output from the results of AutoMLST** (Alanjary *et al.,* 2017)**.** Tree was edited in FigTree V3 and is depicted in circular mode. * Indicates the position of *S. goldiniensis* (pink) and ** indicates the position of *S. collinus* (blue).

Table 3-B: **Summary of QUAST and RAST results from sequential *S. goldiniensis* assemblies**

	Illumina MiSeq	PacBio	Combined assembly	MeDuSa
Domain	Bacteria	Bacteria	Bacteria	Bacteria
Taxonomy	*Streptomyces*	*Streptomyces*	*Streptomyces*	*Streptomyces*
Size (bp)	1048400	9908880	9969327	10,009,569
GC Content (%)	70	71.1	71.04	71
N50	13064	193883	421525	658502
L50	223	16	7	5
Contigs	2854	117	77	10
Subsystems	452	336	336	328
Coding sequences	8700	9294	9294	9025
Neighbours	*Streptomyces scabei*	*Streptomyces scabei*	*Streptomyces avermitilis MA-4680*	*Streptomyces avermitilis MA-4680*

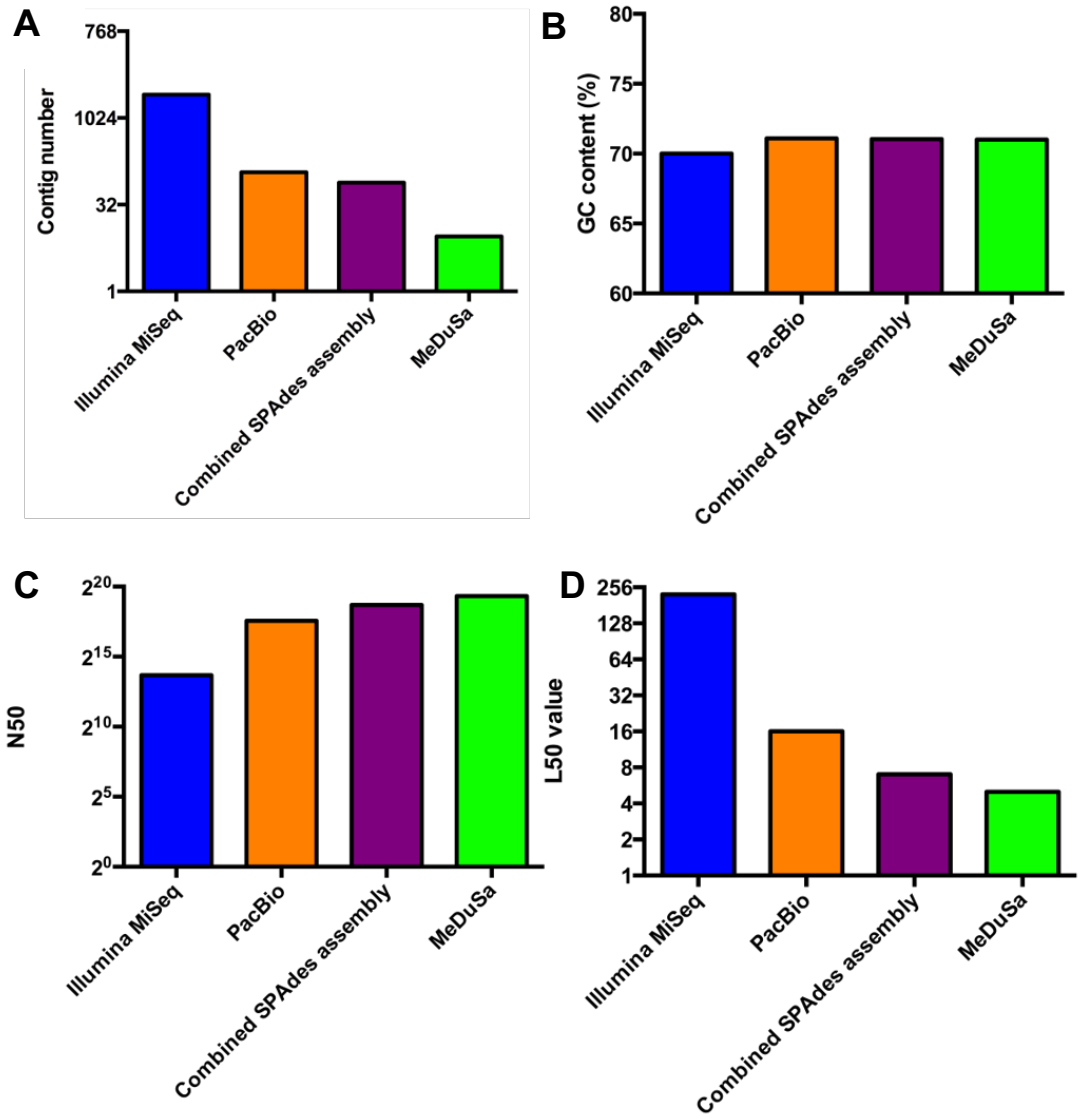


***Figure 3-10*: Histogram series depicting the change in *S. goldiniensis* assembly quality from sequential sequencing and assembly steps**. Values were generated from QUAST assembly quality analysis tool (A) Contig number (B) GC content (C) N50 value (D) L50 value.

.

A

B

Taxonomy	Streptomyces
Size (bp)	1005022
GC Content	71.0
N50	9950726
L50	1
Contigs	9
Subsystems	328
Coding sequences	9025
Neighbours	*Streptomyces bottropensis*
tRNA	81

Figure 3-11: **Summary of final *S. goldiniensis* genome assembly**. (A) Bandage graphical representation of the contigs contributing to the final assembly and their relative size. (B) QUAST summary from final S. goldiniensis.

3.2.6.7 *S. goldiniensis* does not possess plasmids.

The structure of *Streptomyces* genomes differs from the genomes of other bacteria (Hopwood, 1967, 2006), as typically, the majority of their genes can be found on a large linear chromosome which is characteristic of the genus. In addition to this, *Streptomyces* species can harbour circular plasmids, such as pIJ101 from *Streptomyces lividans* which exists in high copy number and replicates by the rolling-circle method which can be observed across the kingdom bacteria (Kieser *et al.*, 1982). Uniquely, some *Streptomyces* species also possess giant linear plasmids which contain genes which can encode for essential and non-essential processes. For example, the clavulanic acid producer *Streptomyces clavuligerus* harbors four large linear plasmids, which range in size from 12-450 kb (Kinashi, 2011). These encode for both genes involved in specialised metabolism as well as genes such as the Tap and Tpg proteins which are required for replication and maintenance of the *Streptomyces* telomeres (Bao and Cohen, 2003). In *S. coelicolor,* the broad-spectrum cyclopentanone antibiotic methylenomycin is encoded for by the 350 kb megaplasmid scp1(Hobbs *et al.*, 1992; Corre *et al.*, 2010; Kinashi, 2011). With the final *S. goldiniensis* assembly containing 8 contigs less than 15 kb in length which could not be mapped to the genome, it was hypothesised that one of more of these contigs may constitute a plasmid. To investigate this, blast was used to mine each contig for the presence of genes which encode for plasmid partitioning machinery - or *par* genes- in addition to other plasmid-associated replicons. The results from this analysis can be found in *Table 3-C* and show that there was no predicted *par* genes on any of the 8 minor contigs, with all eight minor contigs displaying homology to specialised metabolism-associated genes. To confirm this hypothesis physically, pulse field gel-electrophoresis (PFGE) was carried out to rule out the presence of these contigs representing independently replicating and physically separate sections of the genome. As shown in *Figure 3-12,* in the lane representing undigested *S. goldiniensis*

DNA, there are no bands corresponding to plasmids, whereas in the digested controls, bands can be observed. This data, in addition to the absence of plamid-associated bands in standard agarose gel-electrophoresis, confirms the absence of both small circular and giant linear plasmids in the *Streptomyces goldiniensis* genome. Despite the absence of the characteristic Terminal Inverted Repeat sequences which are typically associated with the telomores of *Streptomyces,* it is expected that contig one does contain the vast majority of the genome. Therefore, this genome was annotated by Prokka (Seemann, 2014) and the genbank files from this output were used to create a genomic map with CGviewer, which details the coding sequences, mapped to the contigs of the *S. goldiniensis* genome. This map can be found in *Figure 3-13.*

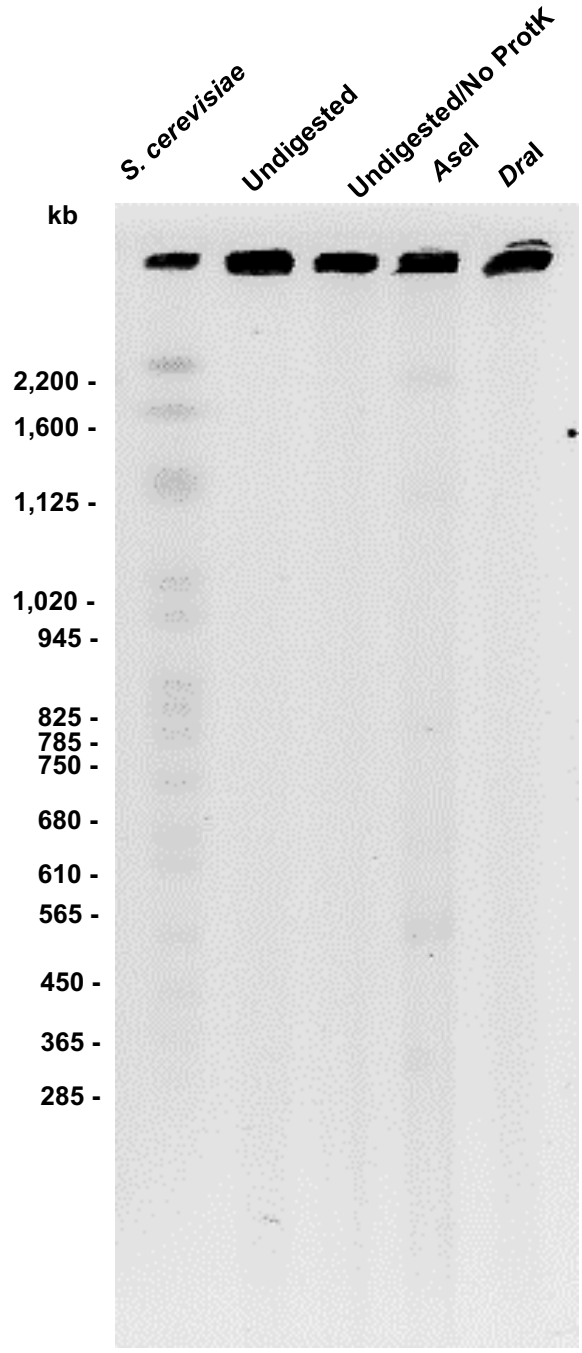


Figure 3-12: **Pulse field gel electrophoreses of genomic DNA from *S. goldiniensis.*** = DNA from *S. cerevisiae* is used as the marker. Lanes show undigested *S. goldiniensis* DNA, *S. goldiniensis* DNA digested with *AseI* and *S. goldiniensis* DNA digested with *DraI.*

Table 3-C: **Summary of BLAST results from *S. goldiniensis* contigs.**

Contig	Size	Blast result	Percentage Identity
1	9.8 mb	*Streptomyces bottropensis* chromosome	93
2	3.6 kb	*Streptomyces sp.* WAC 01438 chromosome, complete genome	86
3	13.5 kb	Streptomyces	93
4	3.5 kb	DQ149987.1 *Streptomyces neyagawaensis* concanamycin A biosynthetic gene cluster, complete sequence	91
5	2.3	CP022161.1 *Streptomyces* sp. 1H-SSA4 genome	95.7
6	5.2	Select seq DQ149987.1 *Streptomyces neyagawaensis* concanamycin A biosynthetic gene cluster, complete sequence	81
7	6.2	*Streptomyces* sp. Z022 chromosome, complete genome	83
8	3.2 kb	*Streptomyces* sp. CdTB01, complete genome	94
9	3.5 kb	*Streptomyces tendae strain* 139 chromosome, complete genome	81

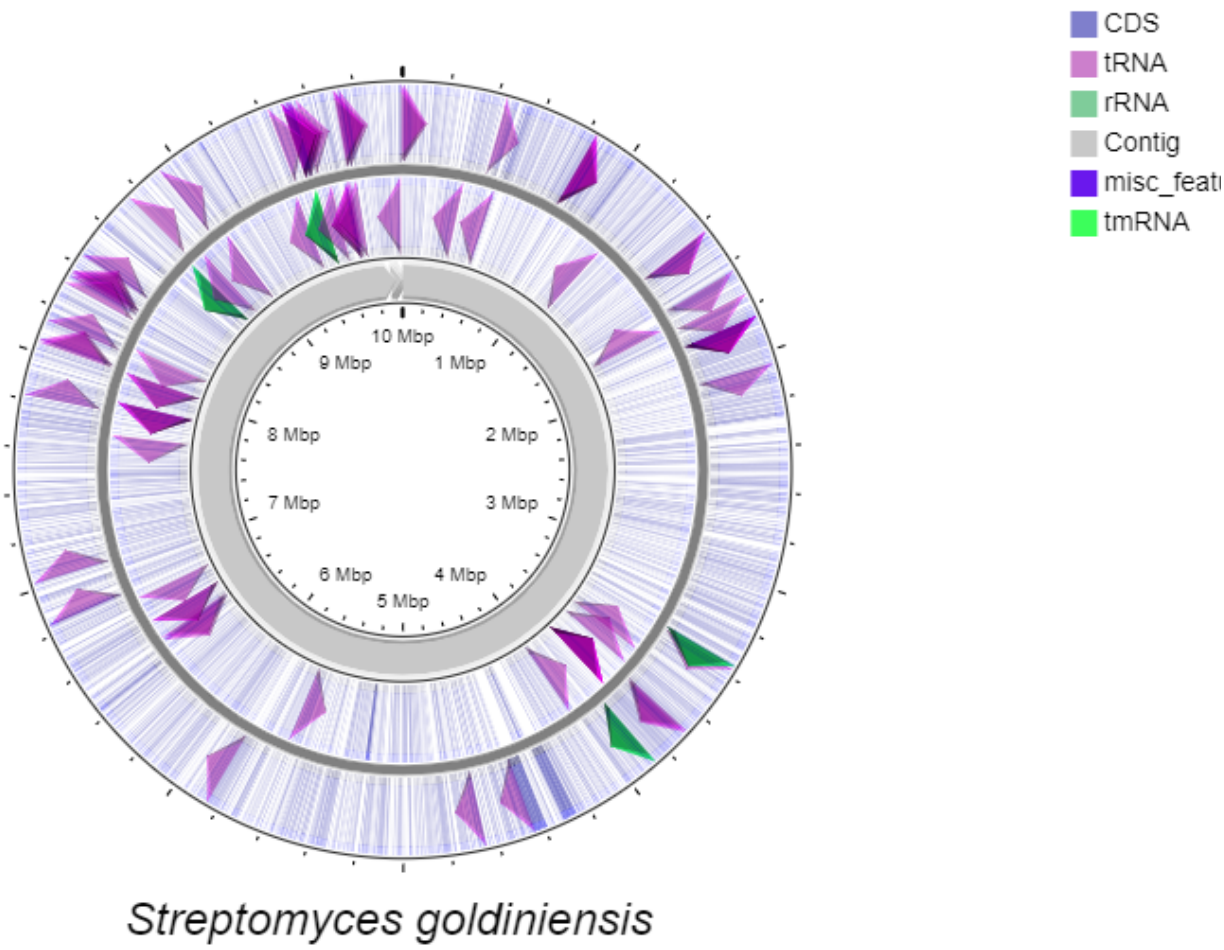


Figure 3-13: Map of *S. goldiniensis* genome constructed by CGviewer. Outer circle represents coding sequences on the major strand, the inner circle represents coding sequences on the minor strand.

3.2.7 The *S. goldiniensis* genome encodes 36 putative biosynthetic gene clusters

Streptomyces are renowned for their ability to biosynthesise a diverse range of molecules with a plethora of structures, functions and pharmaceutical properties. However, it was not until the publication of the first *Streptomyces* genome in 2002 (Bentley *et al.*, 2002) - that of the model actinomycete *Streptomyces coelicolor* - that the true biosynthetic potential of these organisms was understood. As expected, gene clusters encoding for the previously characterised molecules actinorhodin and undecylprodigiosin were identified. However, surprisingly, an additional 18 'cryptic' biosynthetic gene clusters were identified which were proposed to encode for molecules such as siderophores and terpenes. A large proportion of these gene clusters remain 'inactive' or 'silent' under laboratory conditions, and therefore, to understand the true biosynthetic capabilities of these organisms, a whole genome sequence is required (Ikeda, Shin-Ya and Omura, 2014). Therefore, to understand the biosynthetic capacity of *S. goldiniensis* and to preliminarily identify a putative aurodox gene cluster, the whole genome of *S. goldiniensis* was mined for specialised metabolite gene clusters using antiSMASH software. A summary of the results from antiSMASH can be found in *Table 3-D.*. From this analysis it was confirmed that *S. goldiniensis* putatively encodes for a total of 36 specialised metabolites which were inspected manually through the use of BLASTp. As for many *Streptomyces* species- there is diversity across the *S. goldiniensis* BGCs with Type I, Type II and Type III Polyketide Synthase-based gene clusters being identified, as well as Non-Ribosomal Peptide Synthetases. Analysis with antiSMASH revealed the presence of conserved *Streptomyces* gene clusters such as geosmin, desferrioxamine and melanin. In addition, potentially completely novel gene clusters were identified with no predicted

homology to clusters in the antiSMASH database such as a putative bacteriocin which can be found from position 2,705,875 - 2,716,826 in the *S. goldiniensis* genome.

Table 3-D: **Summary of *S. goldiniensis* Biosynthetic Gene Clusters as predicted by antiSMASH.**

Region	Type	Region	Most similar	% similarity
1	T2PKS	32,278 - 104,355	Spore pigment	83
3	Lanthipeptide, lassopeptide	486,431 - 514,192	Citrulassin E (ripp)	100
4	Siderophore	2,159,777 - 2,170,184	NA	0
5	Terpene	2,487,747 - 2,508,882	Lipopeptide	4
6	NRPS, tfuA related	2,543,021 - 2,606,769	Diisonitrile antibiotic SF2768	66
7	Bacteriocin	2,705,875 - 2,716,826	NA	0
8	Terpene	2,777,350 - 2,799,499	Geosmin	100
9	NRPS-like	2,818,277 - 2,858,728	S56P1	11
10	NRPS-like, siderophore	3,085,338 - 3,126,119	Paulomycin	13
11	NRPS-like	3,541,716 - 3,582,164	Octacosamicin	12
12	Terpene	3,625,827 - 3,643,980	NA	0
13	Indole	4,158,688 - 4,181,228	Terfestatin	28
14	bacteriocin,bottromycin,NRPS,transAT-PKS,T1PKS,hglE-KS,lassopeptide	4,213,370 - 4,484,508	Kirromycin	40
15	NRPS-like	4,599,349 - 4,657,342	Echosides	17
16	Terpene	4,783,720 - 4,802,511	NA	0
17	Linardin	5,046,397 - 5,067,017	Pentostatine	17
18	Siderophore	5,085,773 - 5,100,346	NA	0
19	NRPS, melanin	5,167,798 - 5,232,264	Scabichelin	100
20	Terpene	5,257,095 - 5,277,644	NA	0
21	NRPS, T1PKS	5,284,180 - 5,346,580	Herboxidiene	4
22	Terpene	5,480,336 - 5,505,711	Isorenieratene	100
23	T2PKS	5,770,343 - 5,856,353	Fluostatins	60
24	Lanthipeptide, bacteriocin	6,153,412 - 6,179,099	Informatipeptin	100
25	NRPS-like, T1PKS	6,410,852 - 6,457,973	Marineosin	45
26	Terpene	6,507,863 - 6,532,462	Hopene	100
27	Ectoine	6,852,064 - 6,861,735	Ectoine	100
28	Ladderane	6,937,157 - 6,978,500	Colabomycin	11
29	T3PKS, Terpene	7,764,486 - 7,805,661	Merochlorin	12
30	Melanin	8,189,939 - 8,200,406	Melanin	80
31	Siderophore	8,344,177 - 8,355,949	Desferrioxamine A	83
32	Ectoine	8,471,171 - 8,481,599	Ectoine	75
33	Terpene	9,170,031 - 9,191,797	NA	0
34	Bacteriocin	9,838,044 - 9,848,331	NA	0
35	Terpene	9,968,421 - 10,005,017	NA	0

3.2.8 Identification of the putative aurodox biosynthetic gene cluster

From the antiSMASH output, a gene cluster was identified in position 4,213,370 - 4,484,508 on the *S. goldiniensis* genome with 40% similarity to the gene cluster which encodes for kirromycin, the non-methylated aurodox derivative, and hence, it was proposed that this gene cluster putatively encodes for aurodox. A diagram of this supercluster, which is over 270 kb in length, can be found in *Figure 3-14.* The region of this supercluster which putatively encodes for aurodox can be found highlighted in green. This 87 kb region was identified through comparison to the kirromycin gene cluster (Weber *et al.*, 2008). In depth BLASTp analysis has shown that 23 of the 25 genes in the putative aurodox region of the cluster show some level of functional homology to genes in the kirromycin clyster. The role of the aurodox gene cluster will be dissected in more detail in the chapters to follow. Moreover, one of the two genes lacking a homolog in the kirromycin cluster has been putatively identified as a SAM-dependent *O*-methyl transferase, predicted to catalyse the methylation of the nitrogen on position one of the pyrridone group of kirromycin, thus, converting kirromycin to aurodox. This function of this gene will be discussed in more detail in *Chapter 4* and *Chapter 5.* The final gene encodes for a hypothetical protein with no predicted conserved domains.

At 270 kb, the aurodox 'supercluster' is one of the largest biosynthetic gene clusters ever reported (Hashimoto *et al.*, 2018), and is predicted to encode for multiple, independently regulated products. To rule out the notion that the presence of this 'supercluster' is an artefact of genome assembly, primer pairs were designed to amplify the overlap between the putative aurodox region and upstream and downstream genes. The PCR products from this cluster position check can be observed from the agarose gel in *Figure 3-14.* PCR products were confirmed by sequencing, and therefore, it was determined that this supercluster is a true

representation of the position of the aurodox cluster on the genome. Genes encoding Bottromycin A2 – an anti-MRSA compound- can be found upstream of the aurodox region of the supercluster, and can also be found in *S. bottropensis,* a closely related strain to *S. goldiniensis.* The bottromycin A2 gene cluster, which was recently redefined, and is limited to just twelve genes (Crone, Leeper and Truman, 2012), is found in its entirety within the supercluster. Furthermore, an additional *ycaO*-type protein is predicted to be encoded within these bottromycin genes. This has no homology to the core bottromycin A2 genes and is predicted to catalyse post-translational activation of amides (Burkhart *et al.*, 2017). It is predicted that the expression of this gene is developmentally regulated, due to the presence of a TTA codon , indicated by grey triangles in *Figure 3-14.* Downstream of the putative aurodox gene cluster, a group of genes are present which have 100% similarity to the gene cluster for the macrolide vacuolar ATP-ase inhibitor concanamycin A. Furthermore, glycolipid synthase and lassopeptide-encoding genes can be found downstream of the aurodox region of the supercluster.

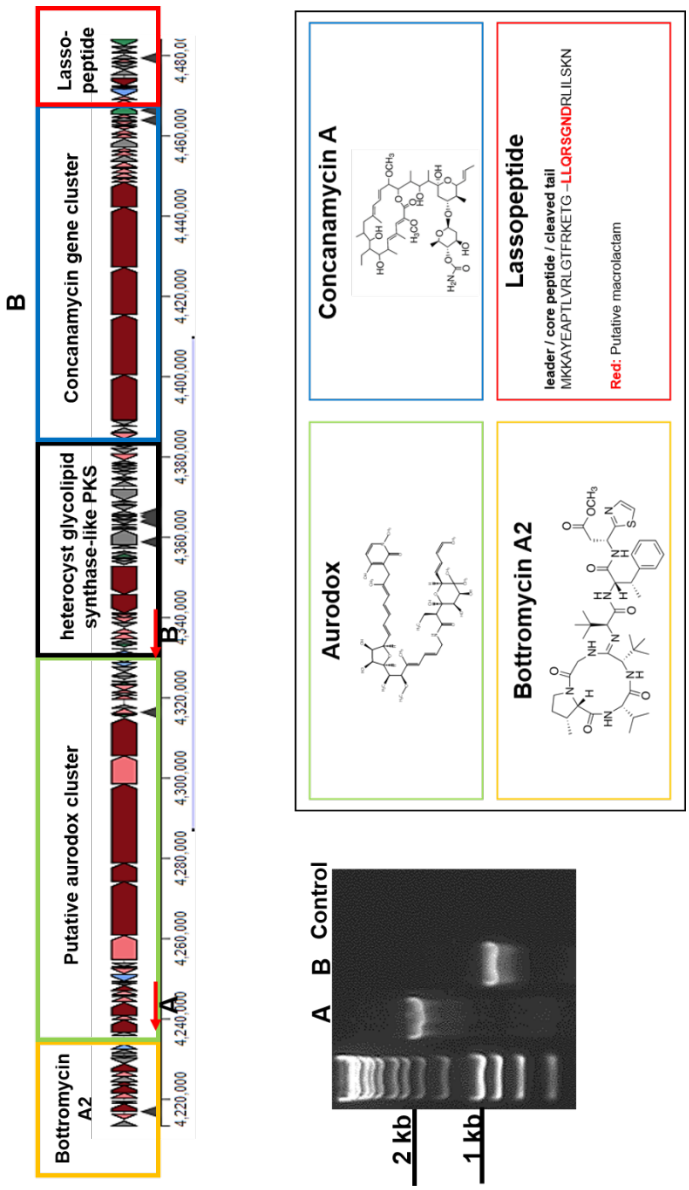


Figure 3-14: **Representation of the putative aurodox supercluster.** Schematic diagram taken directly from antiSMASH. Genes coloured red represent core biosynthetic genes, pink genes represent other biosynthetic genes, green genes represent regulatory genes, blue genes represent transport-associated genes and grey colour indicates other genes involved in biosynthesis. TTA codons are indicated by dark grey triangles under genes. Primer pair binding sites are indicated by 'A' and 'B' on the diagram and their corresponding PCR products are shown on the gel with the lanes labelled accordingly.

Figure 3-15: **List of similar gene clusters to the aurodox region of the predicted supercluster, according to antiSMASH.**

3.3 Summary

The work in this chapter has primarily focussed on improving the understanding of the physical and genomic characteristics of the aurodox producer, *S. goldiniensis.* Firstly, stereo and epifluorescence microscopy have allowed phenotypic characterisation of *S. goldiniensis* in response to differing environmental conditions and to visualise the stages of the *Streptomyces* life cycle. Furthermore, conditions for aurodox fermentation in the laboratory have been optimised, and the aurodox titre has been improved. In agreement with Berger *et al,* aurodox production could be observed in fermentations using the patented aurodox production media, however, an 80-fold improvement in aurodox yield was detected when the fermentation media was changed to GYM. As a result, there is now a robust aurodox fermentation method in place, as well as a strategy for the approximation of aurodox yield in extracts, for use in the following works of this thesis.

Sequencing of the *S. goldiniensis* genome in this study, using a combination of Illumina, PacBio and Oxford Nanopore minION techniques and applying a k-mer dependent assembly algorithm facilitated by SPades (Bankevich *et al.*, 2012) allowed for all three sets of data to be utilised in the genome assembly. Furthermore, scaffold-based assembly via MeDuSa resulted in a final assembly of the genome, which is comprised of 9 contigs, with an N50 value of 9950726 and an L50 value of one. Thus, with more than 99% of the genome on one contig, the quality of the assembly was verified and hence, it was carried forward for use in standard annotation and analysis pipelines.

In order to comprehend the biosynthetic capabilities of *S. goldiniensis* and identify a putative aurodox gene cluster, antiSMASH software was used. From this analysis, 36 putative biosynthetic gene clusters were identified from the *S. goldiniensis* genome

which theoretically encode for both conserved and novel metabolites. Significantly, a 270 kb supercluster was identified, of which 87 kb displays a high level of functional homology to the kirromycin gene cluster. It is therefore hypothesised that this stretch of DNA encodes for the aurodox gene cluster. Further characterisation of this gene cluster will be carried out in future chapters.

Chapter 4 : Genomic characterisation of the biosynthesis and evolution of aurodox in *Streptomyces goldiniensis.*

4.1 Introduction

Specialised metabolites from bacteria are typically encoded for by groups of genes known as biosynthetic gene clusters (Bentley *et al.*, 2002). These clusters comprise of genes which are necessary for the total assembly of small molecules, as well as pathway-specific regulatory, transport and immunity genes (de Lima Procópio *et al.*, 2012). To expand the possibilities of aurodox being used as an anti-virulence therapy optimisation of the aurodox fermentation procedure is required. Improvement and diversification of the aurodox fermentation process can be facilitated by gaining an understanding of the steps involved aurodox biosynthesis through characterisation of the biosynthetic gene cluster (Gomez-Escribano and Bibb, 2012). Until this project, the whole genome of *S. goldiniensis* remained unsequenced and the biosynthetic gene cluster encoding for aurodox production had not been identified. In *Chapter 3*, a putative aurodox biosynthetic gene cluster was identified – a hybrid Polyketide Synthase (PKS)/Non Ribosomal Peptide Synthetase (NPRS)- which showed, on initial analysis to have significant homology to the kirromycin biosynthetic gene cluster (Weber *et al.*, 2008). In the following chapter, an in depth, gene by gene comparison of the aurodox and kirromycin gene cluster has been conducted, as well as whole-cluster homology analysis between these gene clusters. Going forward, a central aim of this work is the characterisation of the catalytic steps involved in aurodox biosynthesis. In addition, bioinformatic analysis tools including antiSMASH have been used to detail the catalytic domains of these large Type I PKS and NRPS genes in an

attempt to characterise the biosynthetic process. Furthermore, the tailoring reactions involved in construction of the final aurodox molecule have been investigated, and importantly, the final step in aurodox biosynthesis – the conversion of the kirromycin precursor molecule to aurodox via methylation of N1- has been characterised bioinformatically. Furthermore, the regulatory mechanisms underpinning aurodox expression have been identified as well as aurodox resistance mechanisms.

In addition, this chapter has worked towards the understanding of the evolutionary events which have resulted in the emergence of the putative aurodox biosynthetic gene cluster. With the genes which putatively encode for aurodox contributing to a larger 'supercluster', it was postulated that homologous recombination between similar clusters in related strains such as *Streptomyces bottropensis* (the bottromycin A2 producer) and *Streptomyces collinus* (the kirromycin producer) has resulted in the establishment of the daughter 'supercluster'. Hence, a combination of antiSMASH and BiGSCAPE/CORASON (Navarro-Muñoz *et al.*, 2018) have been used to characterise the overlap in biosynthetic capacity of these three strains. Additionally, recombination detection programming (RDP; (Martin and Rybicki, 2000) has been used to apply algorithms to predict homologous recombination events between the aurodox, kirromycin and bottromycin A2 cluster which may have resulted in the formation of the aurodox cluster.

4.2 Results

4.2.1 The putative aurodox biosynthetic gene cluster shares functional homology with the kirromycin gene cluster despite gene rearrangements.

Kirromycin- a non-methylated derivative of aurodox- is encoded by an 82 kb gene cluster consisting of trans and *cis*-PKS genes alongside NRPS genes (*kirAI-kirAVI;* Weber *et al.*, 2008). These modular 'megasynthase' enzymes work in synchrony to synthesise the kirromycin backbone before tailoring enzymes such as the methyltransferase KirM and the Crotonyl-CoA reductase/carboxylase KirN decorate the molecule to completion before it is exported by the Major Facilitator Superfamily-like exporter encoded for *kirTI/kirTII* (Weber *et al.*, 2008; Robertsen *et al.*, 2018). Although the possibility that aurodox may have arisen independently from entirely unique sources as observed in the broad range of iron siderophores in bacteria (Fischbach, Walsh and Clardy, 2008) and fosfomycin (Kim *et al.*, 2012) was considered, this was not identified as the most likely evolutionary source of the molecule. Therefore, it was postulated that the similarity in molecular structure between aurodox and kirromycin (*Figure 4.1)* would infer a similarity in the gene products is responsible for biosynthesis and the enzymatic assembly pathway, and hence, it was predicted that homology between their retrospective gene clusters would exist. From preliminary analysis shown in *Figure 4.1,* at a protein level, similarity between the aurodox and kirromycin biosynthetic gene cluster (BGC) is clear. As expected, an orthologous PKS region consisting of both *cis* acetyl-transferases and *trans*-malonyl and ethylmalonyl transferase docking domains is predicted in the aurodox gene cluster. In the kirromycin gene cluster, these have been assigned gene names *kirAI-kirAVI.* Consistent with the kirromycin cluster nomenclature, in the aurodox cluster, these genes have been named (*aurAI-aurAVI)*. A homologous set of core genes encoding for these multi-domain PKSs can also be found in the factumycin

clusters of *Kitasatospora setae* and *Streptomyces spp.* WAC5292 (*facAI-facAVI)* again spread across six independent genes, encoding for the same intermediate molecule (Gui *et al.*, 2015). Hence, evidence suggests that the biosythesis of the aurodox backbone is a conserved process facilitated by these core, modular enzymes. In *Table 4-A,* a description of each aurodox gene can be found, ordered from left to right in accordance with *Figure 4.1,* and its kirromycin homolog - or absence thereof- is detailed. In *Table 4-B,* a description of the putative aurodox genes and their function, as determined by BLASTp is detailed. The individual functions of these genes will be discussed throughout this chapter- with particular attention paid to the catalytic roles of the PKS and NRPS genes.

Of the 25 genes present in the putative aurodox biosynthetic gene cluster, 23 gene homologs are shared with the kirromycin gene cluster, with homology confirmed by BLAST (*Table 4-A)*. There are two genes in the aurodox BGC which do not share homology with the kirromycin BGC. Firstly, a hypothetical protein without conserved domains or predicted functions (*aurHI)*. Secondly, somewhat unsurprisingly, a SAM-dependent *O*-methyl transferase *(aurM*)* was identified and is located between the conserved aspartate-1-decarboxylase (*aurD)* and ethylmalonyl transferase (*aurCII)* genes. It is therefore hypothesised that AurM* catalyses the conversion of kirromycin to aurodox.

Comparison of the aurodox and kirromycin gene clusters at a nucleotide level using EasyFig 2.2.2 (Sullivan, Petty and Beatson, 2011), indicated that the genes within the aurodox and kirromycin BGCs are ordered differently (*Figure 4.2*) One feature of interest is the apparent inversion of the tailoring genes, which appear to flip from upstream to downstream of the core PKS units. In the factumycin clusters of *K. setae* and *S. globoscus,* the genes maintain the order observed in the kirromycin cluster (Thaker *et al.*, 2012) . Several evolutionary and recombination events may have

contributed to this alteration to the order of these clusters. In summary, these data suggest that several sequential horizontal gene transfer events have occurred, resulting in the aurodox supercluster.

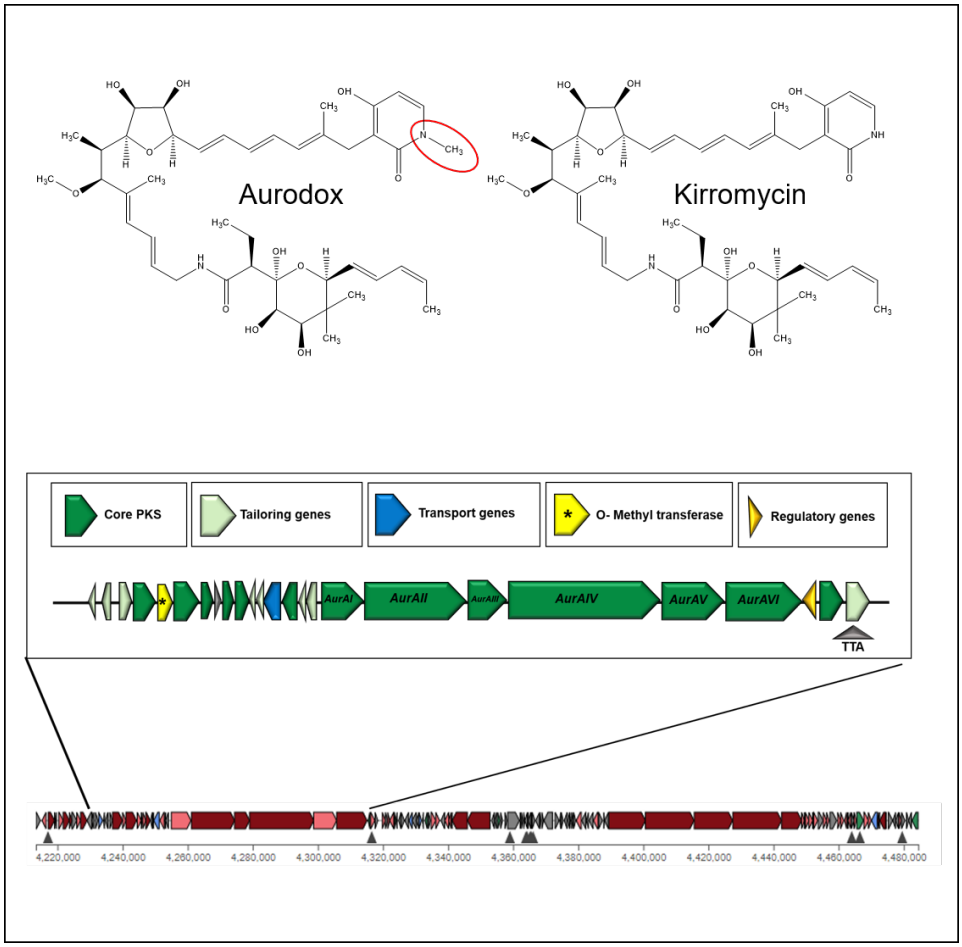


Figure 4-1: **Schematic diagram detailing putative aurodox biosynthetic genes as part of the larger aurodox 'supercluster'**. Upper panel depicts direct cluster prediction output from antiSMASH (278 kb), lower panel highlights genes putatively involved in aurodox biosynthesis (89 kb). Dark green/burgundy arrows represent core PKS genes int the upper and lower panels respectively, light green/pink arrows show tailoring genes, blue arrows show transport genes in both panels, and regulatory genes are shown in orange. The yellow arrow (*) represents a gene predicted to encode an additional SAM-dependent *O*-methyltransferase.

Table 4-A: **List of aurodox genes (ordered left to right according to Figure 4.1) and their kirromycin homologs.**

Gene	Kirromycin Homolog	Function
aurQ	N/A	Hypothetical protein
aurHV	*kirHV*	Hypothetical protein
aurHVI	*kirHVI*	Hypothetical protein
aurB	*kirB*	Non-Ribosomal Peptide Synthetase
aurM*	N/A	Type I SAM-dependent O- Methyltransferase
aurCI	*kirCI*	S-malonyltransferase
aurOII	*kirOII*	Cytochrome P450 hydroxylase
aurD	*kirD*	Aspartate 1-decarboxylase
aurM	*kirM*	Class I SAM-dependent O-methyltransferase
aurCII	*kirCII*	Ethylmalonyl-transferase
aurX	*kirX*	Dieckmann cyclase
aurHII	*kirHIII*	DUF3037 domain-containing protein
aurN	*kirN*	Crotonyl-CoA carboxylase/reductase
aurT	*kirTI/kirTII*	Major Facilitator SuperfamilyTransporter
aurIVI	*kirHIV*	Magnesium ATP-ase
aurAI	*kirAI*	Type 1 Polyketide Synthase
aurAII	*kirAII*	Trans-AT PKS
aurAIII	*kirAIII*	Non-ribosomal peptide synthetase
aurAIV	*kirAIV*	Trans-AT PKS
aurAV	*kirAV*	SDR family NAD(P)-dependent oxidoreductase
aurAVI	*kirAVI*	Type I PKS
aurR	*kirRI/RII*	TetR/AcrR family transcriptional regulator
aurOI	*kirOI*	Cytochrome p450 hydroxylase
aurHI	*kirHI*	Phytanoyl-CoA Dioxygenase

Table 4-B: Description of putative aurodox genes as determined by BLASTp.

Gene	Type	Function	Closest BLAST hit	AA similarity
aurQ	Other Gene	Hypothetical protein	MULTISPECIES *Streptomyces*	94/107 (87%)
aurHII	Other gene	Hypothetical protein	Streptomyces globosus	225/259 (86%)
aurHIII	Other Gene	Hypothetical protein	*Streptomyces sp.* SM11	242/288 (84%)
aurAVII	Core Biosynthetic Gene	Non-ribosomal peptide synthetase	*Streptomyces roseoverticillatus*	887/1023 (86%)
aurM*	Additional Biosynthetic Gene	SAM dependent- O-methyltransferase	*Streptomyces roseoverticillatus*	296/333 (88%)
aurCI	Core Biosynthetic Gene	S-malonyltransferase	*Streptomyces roseoverticillatus*	973/1076(90%)
aurOII	Additional Biosynthetic Gene	Cytochrome P450 hydroxylase	*Streptomyces sp.* CB00455	344/401 (85%)
aurD	Other Gene	aspartate 1-decarboxylase	*Streptomyces roseoverticillatus*	129/137 (94%)
aurM	Additional Biosynthetic Gene	class I SAM-dependent O- methyltransferase	*Streptomyces klenkii*	301/319 (94%)
aurCII	Core Biosynthetic Gene	Acyltransferase domain-containing protein	*Streptomyces sp. CB00455*	383/440(87%)
aurX	Other Gene	Dieckmann Cyclase	*Streptomyces globosus*	231/259 (89%)
aurHIV	Other gene	Uncharacterized protein, DUF2087 family	*Streptomyces sp. SM11*	149/178 (83%)
aurN	Additional Biosynthetic Gene	crotonyl-CoA carboxylase/reductase	*Streptomyces azureus*	427/443 (96%)
aurT	Transport related gene	Major Facilitator SuperfamilyTransporter	*Streptomyces klenkii*	509/557 (91%)
aurHVI	Other Gene	Magnesium ATP-ase	*Streptomyces sp. SM11*	97/109 (88%)
aurAI	Core Biosynthetic Gene	Type 1 Polyketide Synthase	*Streptomyces roseoverticillatus*	2513/3129 (80%)

aurAII	Additional Biosynthetic Gene	SDR family NAD(P)-dependent oxidoreductase	*Streptomyces roseoverticillatus*	1922/2361 (81%)
aurAIII	Additional Biosynthetic Gene	SDR family NAD(P)-dependent oxidoreductase	*Streptomyces sp. AVP053U2*	5416/6729 (80%)
aurAIV	Core Biosynthetic Gene	non-ribosomal peptide synthetase	*Streptomyces roseoverticillatus*	1350/1551 (87%)
aurAV	Core Biosynthetic Gene	SDR family NAD(P)-dependent oxidoreductase	*Streptomyces klenkii*	3627/4506 (80%)
aurAVI	Additional Biosynthetic Gene	SDR family NAD(P)-dependent oxidoreductase	*Streptomyces klenkii*	1707/2018 (84%)
aurR	Regulatory Gene	TetR/AcrR family transcriptional regulator	*MULTISPECIES Streptomyces*	178/182 (97%)
aurOI	Additional Biosynthetic Gene	Cytochrome p450 hydroxylase	*Streptomyces viridosporus*	390/406 (96%)
aurHVI	Other gene	Phytanoyl-coA Dioxygenase	*Streptomyces viridosporus*	267/278 (96%)

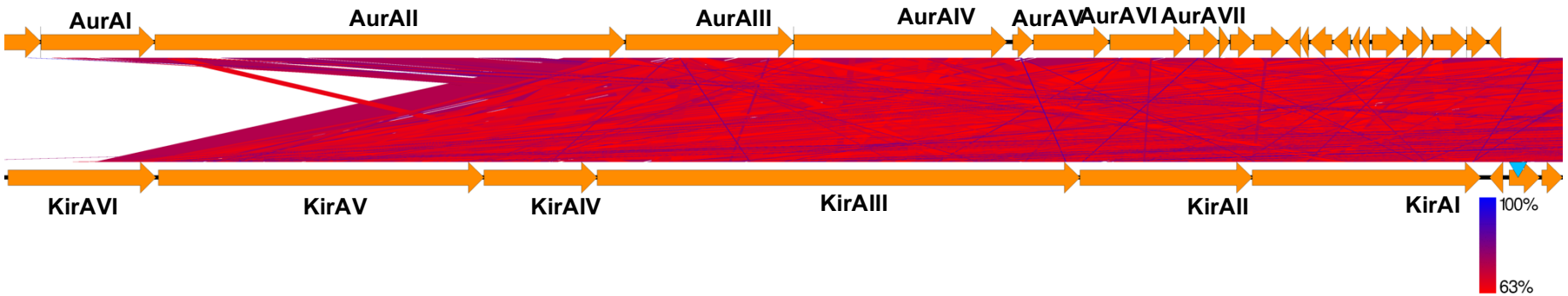


Figure 4-2: **Diagram displaying nucleotide similarity between the PKS regions of the putative aurodox gene cluster (upper) and kirromycin gene cluster (lower).**

Figure was generated by EasyFig 2.2, Scale represents conditionally formatted nucleotide sequence similarity. Similarity threshold was set at 60% (red)

4.2.2 The Putative aurodox Cluster lacks a canonical phosphopantetheinyl transferase (PPTase)

In PKS and NRPS containing BGCs, genes encoding phosphopantetheinyl transferases (PPTases) are often present (Owen, Copp and Ackerley, 2011). These genes encode enzymes responsible for the post-translational activation of acyl carrier proteins (ACPs) in PKSs and peptidyl carrier proteins (PCPs) in NRPSs. In the kirromycin gene cluster, *kirP* encodes for an Sfp-type PPTase which catalyses the activation of the ACPs and PCPs during kirromycin biosynthesis via transfer of the phosphopantetheinyl arm, which is required for the covalent attachment of biosynthetic intermediates (Pavlidou *et al.*, 2011). There is an absence of such a PPTase in the putative aurodox gene cluster and, remarkably, there are no predicted PPTase genes throughout the entire aurodox 'supercluster' (*Table 4-A)*. Despite the expression of PKSs/NRPSs being entirely possible without a cluster-encoded PPTase, expression of these gene clusters is often low (Pavlidou *et al.*, 2011; Thaker *et al.*, 2012). Therefore, given that aurodox expression occurs in *S. goldiniensis* to detectable levels, evidence suggests that a PPTase must be present in the genome to catalyse activation of the ACPs and PCPs within the cluster. This led to two hypotheses: firstly, that a *trans*-acting PPTase from elsewhere in the genome acts upon the aurodox cluster. This is probable, given the presence of two PPTases at distant loci in the *S. goldiniensis* genome. One of these PPTases, is located directly upstream of the aurodox supercluster (4464768…4465448) and shares 48% amino acid identity with *kirP.* The second hypothesis is that the hypothetical protein that lacks a homolog in the kirromycin cluster (*aurHI)* is an uncharacterised PPTase. However, with no conserved domains or homology to canonical *Streptomyces* PPTases, this finding would be significant.

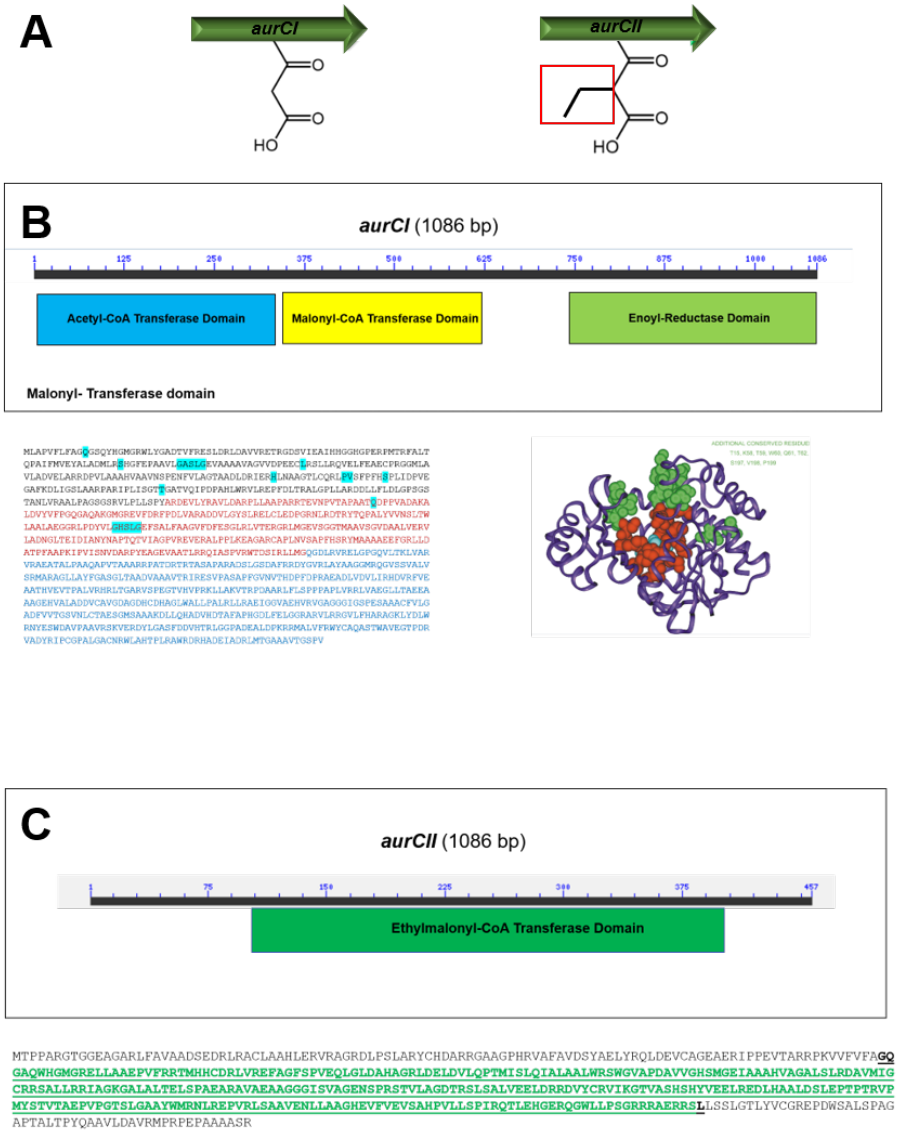
4.2.3 Understanding the role of the malonyl-CoA and ethymalonyl-CoA transferases *aurCI* and *aurCII* in the supply of aurodox extender units.

It is important to decipher the mechanisms behind the incorporation of precursors to the aurodox biosynthetic pathway. An understanding of the starter and extender units which are complicit in aurodox synthesis can enable the design of isotopically-labelled carbon feeding experiments (Weber *et al.*, 2008). In turn, these experiments can result in the synthesis of a labelled molecule for utilisation in biosynthesis and mechanism of action studies (Deshpande, Ambedkar and Shewale, 1988). Furthermore, understanding precursor supply can ultimately facilitate synthetic biology and mutasynthesis approaches to aurodox derivative production, with alterations to extender units leading to the formation of novel molecules. The available evidence from analysis by BLASTp and 'clusterblast' algorithms from antiSMASH lead to the hypothesis that precursor supply is conserved between the aurodox and kirromycin pathways. In nature, the presence of *trans*-AT PKSs is uncommon, and the combined use of *trans* and *cis*-ATs is unique, with examples limited to enzymes homologous to *kirAI-kirAVII* (Robertsen *et al.*, 2018). However, in agreement with the enzymatic pathway proposed by Weber *et al* for kirromycin*,* the data suggest that a combination of *cis*-AT domains within AurAVI and aurVII) and *trans*-acting malonyl and ethylmalonyl transferase enzymes (*aurCI* and *aurCII)* supply extender units successively.

In the putative aurodox biosynthetic pathway, it is hypothesised that *aurCI* (1086 aa) encodes a homolog of the malonyl-coA transferase KirCI (1086 aa, *Table 4-B),* with BLASTp conserved domain analysis predicting that both genes share an unusual AT1-AT2-ER domain pattern, previously observed and characterised in *kirCI* (Musiol *et al.*, 2013)*.* As is the case for *kirCI,* these data suggest that the initial AT domain of *aurCI* (AT1)*,* is inactive as an acyl-transferase, with mutations in conserved, essential

residues. Significantly, the histidine to alanine mutation in the GHSxG conserved motif required for acyl-transferase activity (Yadav, Gokhale and Mohanty, 2003) is conserved between KirCI and AurCI. In KirCI, the absence of AT activity by this domain was demonstrated using autoradiographic techniques (Musiol *et al.*, 2013), however it is hypothesised that this domain encodes PKS-proofreading activity due to homology with *pedC* which is involved in the proofreading of pederin PKS activity (Jensen *et al.*, 2012). The second, central AT domain of AurCI shows greater than 90% amino acid homology with the *kirCI*_AT2 domain and hence, with the presence of the conserved GHSxG catalytic domain, this gene is proposed to encode trans-malonyl activity, transporting malonyl-CoA to the ACP of domain four (Musiol *et al.*, 2013). The terminal ER domain is predicted to encode trans-acting ER activity (*Figure 4-3)*.

The second putative *trans*-AT enzyme, *aurCII* is a predicted homolog of the ethylmalonyl-CoA transferase *kirCII*. Studies have shown that during kirromycin biosynthesis, *kirCII* specifically interacts with the ACP of module five to incorporate ethylmalonyl_CoA into the kirromycin backbone, hence, providing a mechanism to extend polyketide diversity (Musiol *et al.*, 2011). With a 66% amino acid similarity to *kirCII,* and the presence of a conserved ethylmalyonyl-transferase domain, evidence suggests that this gene encodes for an enzyme which catalyses the loading of an ethylmalonyl-coA extender unit to the ACP of module five during aurodox biosynthesis (*Figure 4-3,4-4).*



MTPPARGTGGEAGARLFAVAADSEDRLRACLAAHLERVRAGRDLPSLARYCHDARRGAAGPHRVAFAVDSYAELYRQLDEVCAGEAERIPPEVTARRPKVVFVFAGQ
GAQWHGMGRELLAAEPVFRRTMHHCDRLVREFAGFSPVEQLGLDAHAGRLDELDVLQPTMISLQIALAALWRSWGVAPDAVVGHSMGEIAAAHVAGALSLRDAVMIG
CRRSALLRRIAGKGALALTELSPAEARAVAEAAGGGISVAGENSPRSTVLAGDTRSLSALVEELDRRDVYCRVIKGTVASHSHYVEELREDLHAALDSLEPTPTRVP
MYSTVTAEPVPGTSLGAAYWMRNLREPVRLSAAVENLLAAGHEVFVEVSAHPVLLSPIRQTLEHGERQGWLLPSGRRRAERRSLLSSLGTLYVCGREPDWSALSPAG
APTALTPYQAAVLDAVRMPRPEPAAAASR

Figure 4-3 **Summary of the role of malonyl-CoA transferase *aurCI* and the ethylmalonyl transferase *aurCII*.** (A) Schematic diagram of the delivery of aurodox extender units by AurCI and AurCII. (B) Description of modular domains in AurCI. In the amino acid sequence, black text represents AT1 domain, red font represents AT2 domain and blue font corresponds to ER domain. Residues highlighted in cyan are essential for catalysis. Protein confirmation is adapted from Yamav *et al*, 2003. (C) Description of AurCI. Green text represents ethylmalonyl transferase domain.

4.2.4 **The structural complexity of aurodox is facilitated by the unusual PKS/NRPS organisation in AurI-AurVI.**

The modular catalytic domains of AurI to AurVI are hypothesised to perform the enzymatic extension process which largely follows those of KirI to KirVI (*Figure 4.4*). In AurAI and AurAII, acetyl Co-A is extended by Claisen condensation reactions, however (Reid *et al.*, 2003) the unusually arranged PKSs do not follow the typical KS-AT-ACP organisation. In these unusual PKS arrangements there are two additional dehydratase domains present which were not identified in the predicted kirromycin pathway (Weber *et al.*, 2008). It is hypothesised that these domains are not unique to the aurodox pathway and are also present in the kirromycin pathway, with the clustertools algorithm predicting the presence of this domain, and the same dehydration reaction and double bond formation is conserved. Again, it is thought that advances in pathway prediction software has facilitated the detection of these domains.

In the putative aurodox pathway, *aurAIII* encodes a hybrid NRPS/PKS enzyme consisting of consecutive condensation and adenylation domains which catalyse the condensation of glycine and the incorporation of the amide bond. This process appears to be conserved between the aurodox and kirromycin pathway. The roles of *aurAIV-AVII* appear to mimic the actions of *kirAIV-AVII* with identical domain organisation predicted and greater than 75% amino acid similarity in all cases. These enzymes are predicted to extend the aurodox intermediate chain before AurB catalyses the incorporation β-alanine, which is synthesised by the aspartate-1-decarboxylase, AurD. The specificity of AurB for these specific residues is indicated by the DTLQLGVIWK consensus sequence which is indicative of an active site which accepts alanine (Rausch *et al.*, 2005).

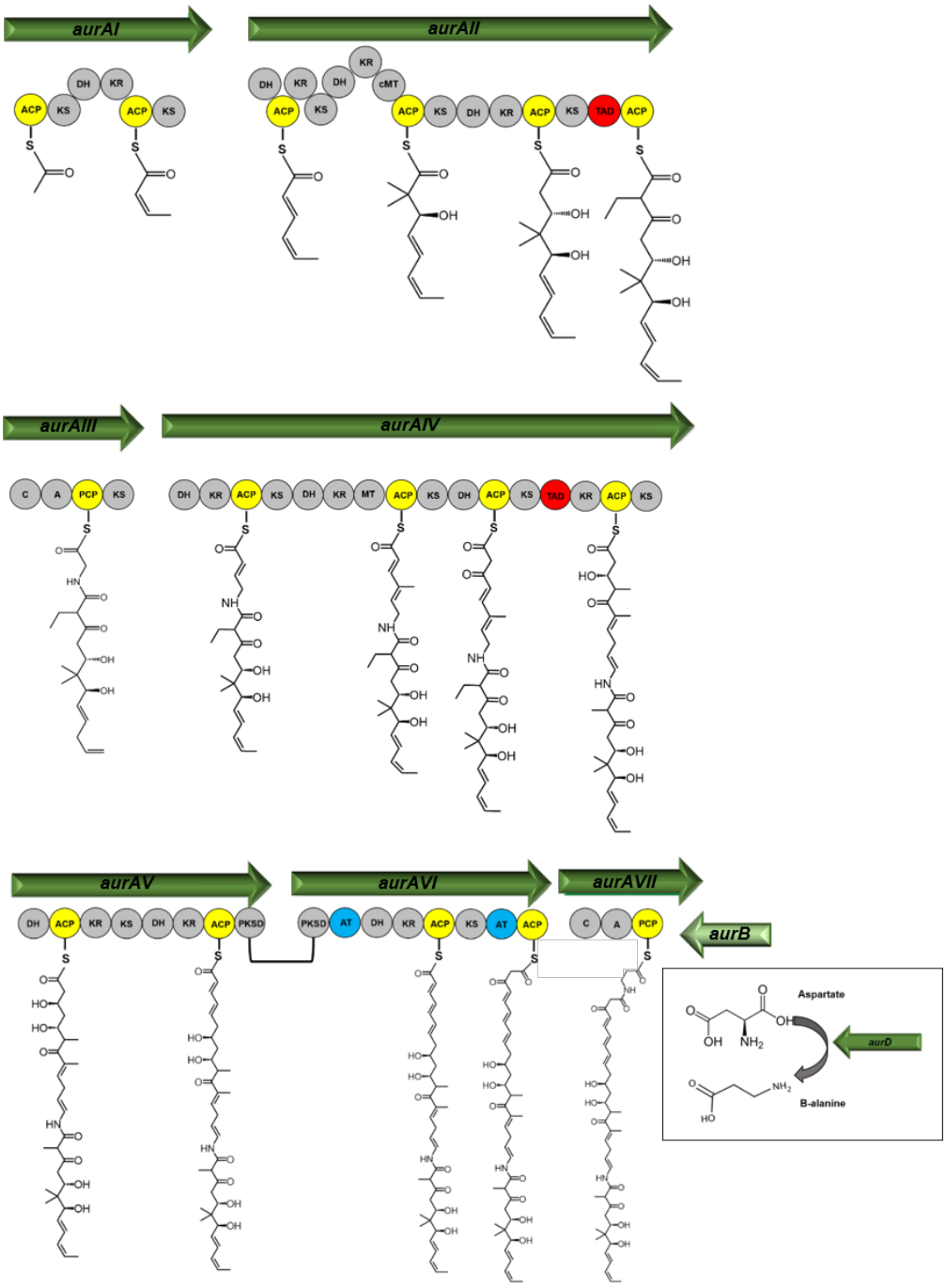


Figure 4-4: **Summary of PKS reactions involved in aurodox backbone synthesis**. ACP: acyl carrier protein; AT: acyltransferase domain; C: condensation domain; DH: dehydratase domain ER: enoyl reductase domain; KR: keto reductase domain; KS: keto synthase domain; MT: methyl transferase domain; PCP: peptidyl carrier protein; TAD: trans-AT docking domain

4.2.5 AurX is a putative Dieckmann cyclase responsible for simultaneous cyclisation and cleavage of the aurodox intermediate from AurB.

One of many unique features of the aurodox/kirromycin biosynthetic pathway is the absence of a classical thioesterase (Robertsen *et al.*, 2018). Typically, thioesterases are responsible for the removal of PKS/NRPS intermediates from the ACP/PCP, as they catalyse the cleavage of the ester bond via the removal of water (Kotowska *et al.*, 2002). Therefore, as this cleavage is essential for release of the PK-NRP, it was hypothesised that *aurX* encodes this function. BLASTp results indicate that *aurX* encodes a member of the Dieckmann Cyclase (DC) enzymes, with 89% amino acid similarity to the DC enzyme from *Streptomyces globoscus (Table 4-B).* These recently discovered enzymes catalyse the simultaneous cyclisation and cleavage of PK-NRPs (Gui *et al.*, 2015) and it therefore appears that the AurX functions in this role in aurodox biosynthesis, putatively catalysing ring closure and the formation of the goldinoic acid moiety. Again, the data suggests that this mechanism is conserved between both pathways.

4.2.6 The initial decorative reactions are conserved between the aurodox and kirromycin pathways.

Following the cleavage of the aurodox intermediate from AurB by AurX, tailoring enzymes facilitate the maturation of molecule. It is predicted that the SAM-dependent *O*-methyltransferase encoded by *aurM* catalyses the reduction and methylation of O-43 (Robertsen *et al.*, 2018). It is postulated that the Cytochrome P450 hydroxylases encoded by *aurOI* and *aurOII* catalyse the hydroxylation of C-30 and C-16 respectively. Recent work has also demonstrated that KirOI can be assigned to the formation of the double bond between C5 and C6 on the pyridone ring (Robertsen *et al.*, 2018), with gene disruption studies resulting in 5,6-dihydro-kirromycin. With 93.3% amino acid similarity and the formation of these bonds in both structures, it is

very likely that AurOI catalyses identical reactions to KirOI. In the aurodox pathway, it is predicted that AurOII simultaneously catalyses the hydroxylation of C-16 and the closure of the tetrahydrofuran ring, which was characterised biochemically by Robertsen *et al*. It is hypothesised these reactions lead to the formation of the final precursor molecule during aurodox biosynthesis, kirromycin. A schematic summary of these reactions can be found in *Figure 4-5.*

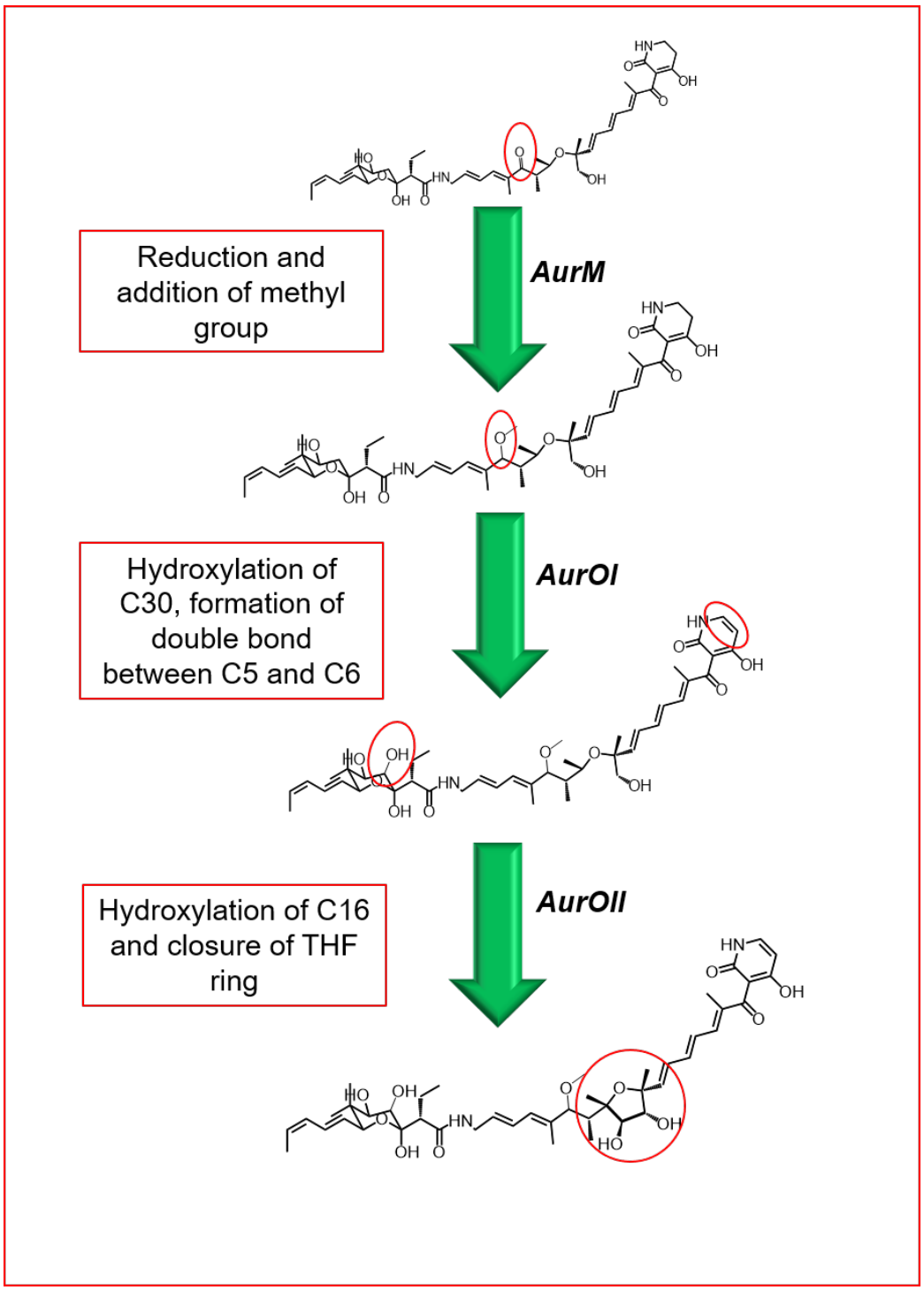


Figure 4-5: **Schematic representation of the aurodox decorating enzymes and their putative reactions**. AurM catalyse they reduction and methylation of O-14, AurOI is a cytochrome P450 and hydroxylates C30 and catalyses double bond formation between C5 and C6, AurOII hydroxylates C30 and closes THF ring.

4.2.7 The putative final step in aurodox biosynthesis is the conversion of kirromycin to aurodox by the SAM-dependent *O*-methyltransferase, AurM*.

In the putative aurodox BGC, *aurM*,* encodes a SAM-dependent *O*-methyltransferase (O-MTase, *Figure 4-2*) which does not display homology to the genes of the kirromycin biosynthetic gene cluster. It is predicted that this gene catalyses the conversion of the final aurodox precursor, kirromycin, to the mature aurodox molecule, via methylation of the pyridone ring. Given that aurodox, but not kirromycin, can inhibit T3S in EPEC without effecting growth, this methylation is predicted to modulate T3SS activity and therefore should be investigated. Phylogenetically, *aurM** forms a distinct clade within the SAM-dependent *O*-MTases of *Kitasatospora setae* and *Streptomyces globosus* which produce factumycin-like molecules (*Figure 4-6)*. Although these reactions have not yet been characterised biochemically, it is likely that these two *O*-MTases also catalyse the methylation of the pyridone rings situated N1 position of factumycin-like molecules (Thaker *et al.*, 2012).

To enhance the hypothesis, experiments were designed which allowed for the detection of aurodox precursors in fermentation extracts. *S. goldiniensis* was grown in GYM media, chloroform was used to extract aurodox and aurodox intermediates, and HPLC-MS analysis was carried out. These data support the hypothesised model, with the typical 7.22 minute elfamycin peak corresponding to m/Z ratios of 795.9 and 810.9, which can be assigned to kirromycin and aurodox respectively *(Figure 4-6,*). This was a consistent phenomenon in LC-MS analysis of aurodox extractions from *S. goldiniensis* and hence further evidence supports the idea that kirromycin is final aurodox precursor molecule

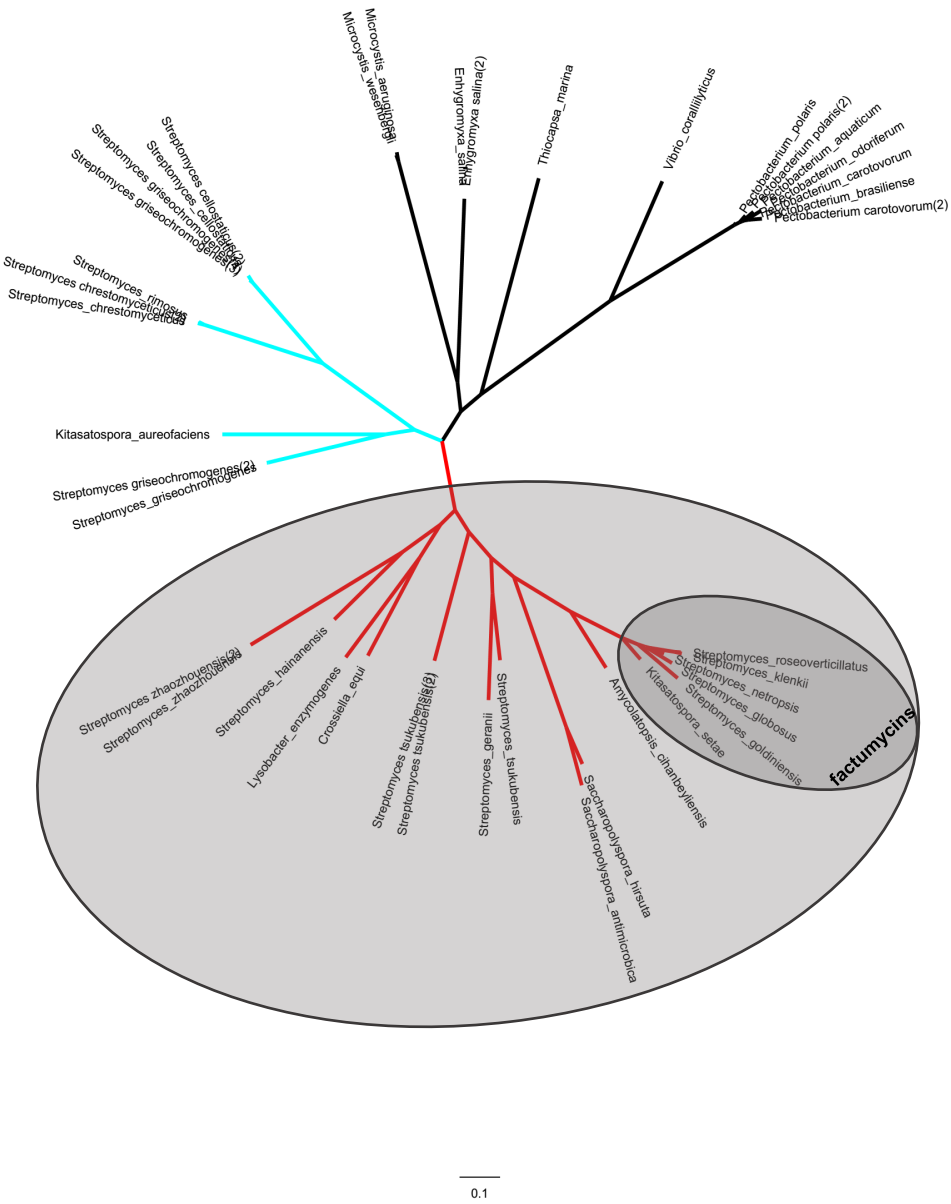


Figure 4-6: **Phylogenetic tree of *O*-methyltransferase genes of specialised metabolite BSGs from selected bacterial species**. Phylogenetic analysis of SAM-dependent *O*-methyltransferases species by Maximum Likelihood method (MEGA7). Evolutionary history was inferred by using the Maximum Likelihood using 500 bootstraps method based on the Tamura-Neil model . The tree with the highest log likelihood is shown. The tree is drawn to scale, with branch lengths correlating to the number of substitutions per site (scale bar). The analysis involved 41 amino acid sequences, which were taken from top 100 most similar proteins from BLASTp, with designated species. All positions containing gaps and missing data were eliminated.. The clade in which AurM* clusters is highlighted, with *O*-methyltransferase from 'factumycin' and related genes highlighted in dark grey

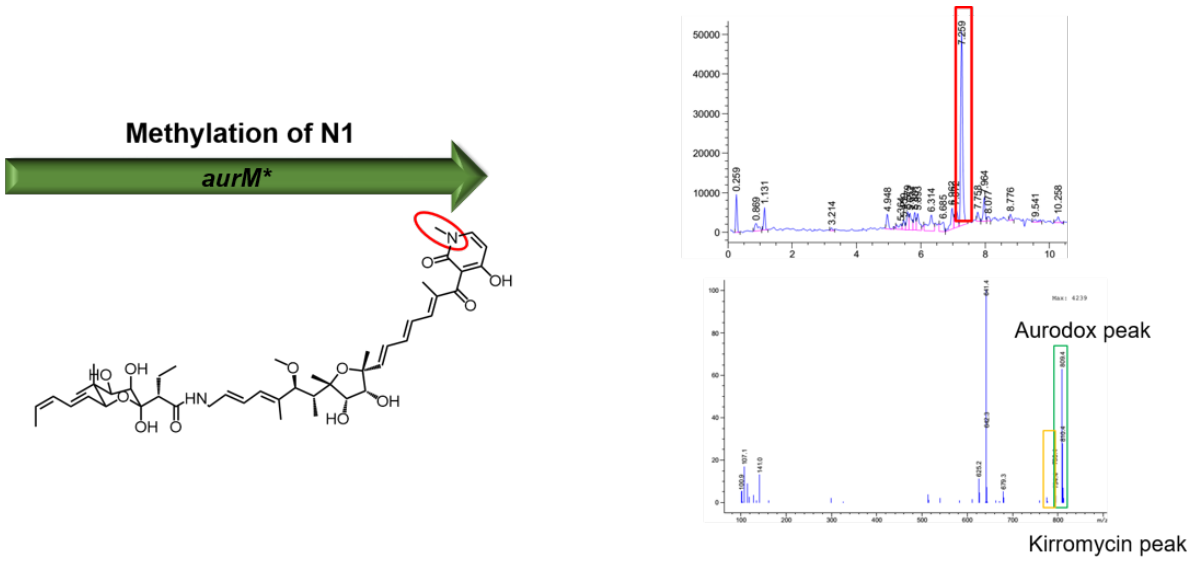


Figure 4-7: **Representation of Methylation Reaction Putatively Catalysed by *aurM*.*** Schematic indicates methylation position. LCMS analysis shown describes the peaks which indicate the presence of aurodox and kirromycin in *S. goldiniensis* fermentation extracts

.

4.2.8 The TetR- type transcriptional regulator, AurR is predicted to regulate control the expression of the aurodox gene cluster in *S. goldiniensis.*

Regulation of specialised metabolism is important as the expression of BGCs is costly in terms of energy and resources. Therefore, it is vital that these BGCs are expressed only when required in order to optimise the cost to benefit ratio in bacterial ecological warfare. In bacteria, the most common type of regulatory system is the 'one-component regulatory system', where the regulatory protein has a sensing domain and a DNA binding domain located on the same polypeptide (Cuthbertson and Nodwell, 2013). These proteins which are almost exclusively formed of alpha-helices that will bind to ligands which in turn alter the conformation of the protein in order to modify its affinity for DNA binding (Thaker, Spanogiannopoulos and Wright, 2010). In the putative aurodox cluster, *aurR* encodes a classical TetR-type regulator with 98% amino acid similarity to the TetR/AcrR family of regulators from *Streptomyces.* In the phylogenetic tree shown in *Figure 4-8,* it is demonstrated that AurR is most closely related to the 'autoregulatory' TetR-type proteins from *S. viridosporus* and *S. spongicola.* These act as repressors in antibiotic-encoding biosynthetic gene clusters, and when bound to the correct ligand, will inhibit the expression of the biosynthetic gene cluster in question, due to reduced affinity for promoter binding. It is known from previous studies that the presence of aurodox in *S. goldiniensis* media results in the downregulation of aurodox production via feedback inhibition, with a concentration of 400 µg/ml completely inhibiting aurodox biosynthesis (Liu, Hermann and Miller, 1977). In this study, the reversibility of feedback inhibition lead to the hypothesis that a likely mechanism of action is allosteric inhibition of aurodox biosynthetic enzymes by aurodox itself. However, at this point, the genetic nature of regulation of biosynthetic gene clusters in bacteria was very poorly understood. Consequently, the results from this study, notably the relatedness to common TetR repressors, in combination with the results of the feedback inhibition study by *Liu* et al, (1977)

suggest that the protein encoded by *aurR* binds to aurodox which results in a conformational change that prevents AurR from interacting with the promoter region and hence, inhibits expression of the aurodox biosynthetic gene cluster.

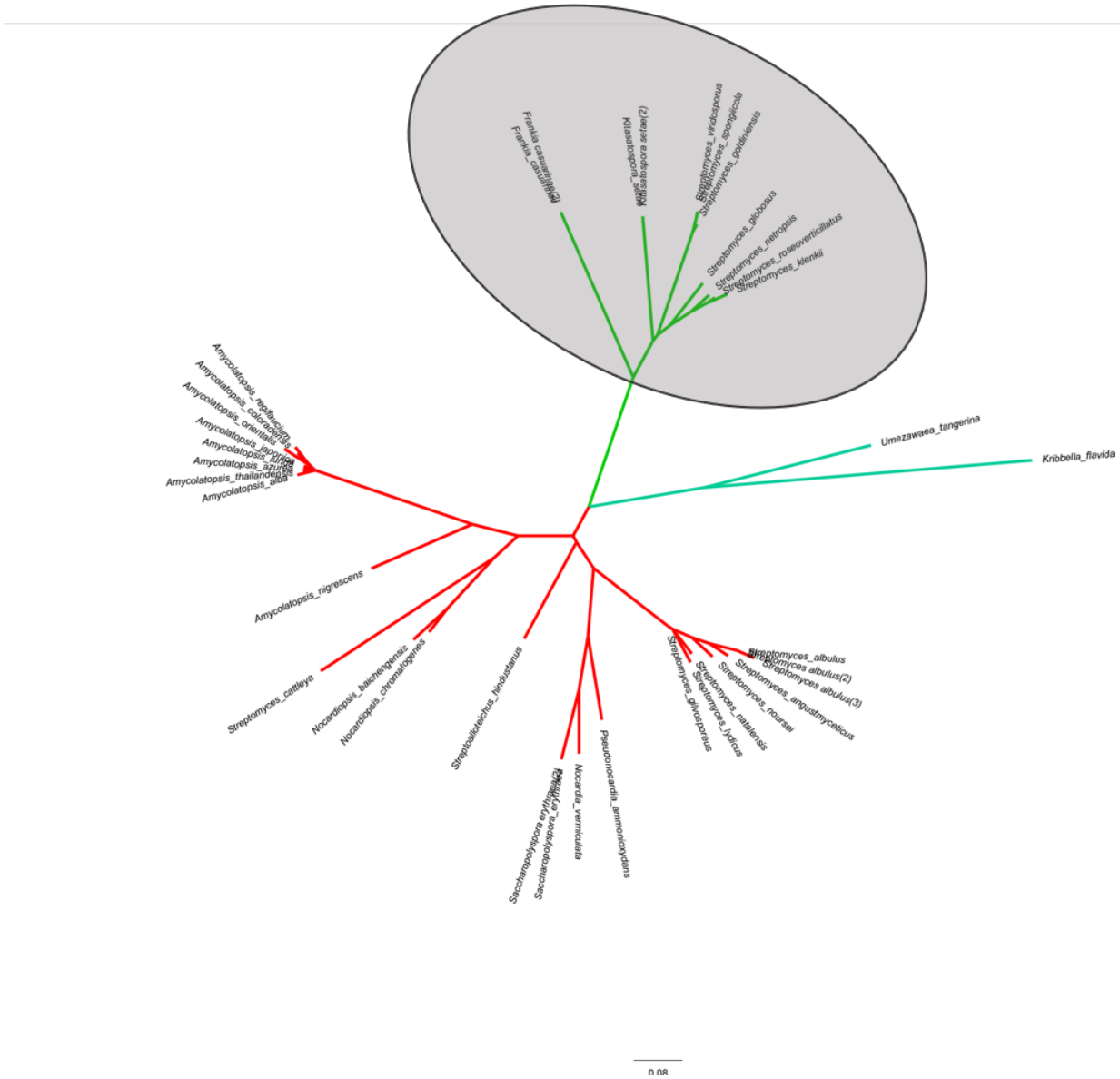


***Figure 4-8*: Phylogenetic tree of the TetR/AcrR Family Regulator AurR and related proteins from BSGs from selected bacterial species.** Phylogenetic analysis of TetR/AcrR family by Maximum Likelihood method (MEGA7)**.** Evolutionary history was inferred by using the Maximum Likelihood method using 500 bootstraps method based on the Tamura-Neil model. The tree with the highest log likelihood is shown. The tree is drawn to scale, with branch lengths correlating to the number of substitutions per site (scale bar). The analysis involved 38 amino acid sequences, which were taken from top 100 most similar proteins from BLASTp, with designated species. All positions containing gaps and missing data were eliminated.. The clade in which AurR clusters is highlighted.

4.2.9 Aurodox is putatively exported by the Major-facilitator Superfamily-Type Exporter, AurT.

Specialised metabolite transporters are essential for the translocation of small molecules into the extracellular space (Zhou *et al.*, 2016). Additionally, they play a vital role in resistance in organisms producing compounds with a conserved, intracellular target, as they require an efflux mechanism for immunity, as to prevent 'suicidal' effects of the compound. The diversification of target proteins can also influence resistance. Hence, an additional copy of EF-Tu (*tuf2)* is carried by many elfamycin producers, and this is the case for *S. goldiniensis* (Thaker *et al.*, 2012), with a *tuf2* gene positioned 21 kb upstream of the aurodox supercluster, it is hypothesised that this gene plays an essential role in aurodox resistance. Moreover, in the case of the elfamycins which target EF-Tu (Rückert *et al.*, 2013), there is often a cluster encoded resistance mechanism present. In the putative aurodox biosynthetic gene cluster, a Major-Facilitator Superfamily protein is encoded for by *aurT* (1677 bp). This is a homolog of the kirromycin transporter/exporter which is encoded for by *kirTI* (903 bp) and *kirTII* (747 bp), with the N and C-terminus being encoded by two separate ORFs based on the original kirromycin gene cluster annotation (Weber *et al.*, 2008). It should however be noted that this protein may be encoded by a single ORF, as the obsolete sequencing technologies used (Roche 454) were more likely to introduce artificial stop codons that modern sequencing methods (Trimble *et al.*, 2012). Despite the authors admittedly attempting to counteract this by the use of GC FramePlot by ARTEMIS to predict ORFs (Ishikawa and Hotta, 1999) it is possible that a stop codon has been artificially introduced between the two proteins which are typically encoded by a single ORF, with *aurT* almost exactly totally the combined length of the *kirTI* and *kirTII* ORFs (Weber *et al.*, 2008; Robertsen *et al.*, 2018).

Phylogenetic analysis of MFS proteins from BGCs again demonstrated clustering of AurT with exporters of other elfamycins such as the factumycin producers *Streptomyces globoscus* and *Streptomyces klenkii* (*Figure 4-9*). It is hypothesised that this protein exports aurodox and therefore, plays an important role in aurodox resistance in addition to export of the molecule.

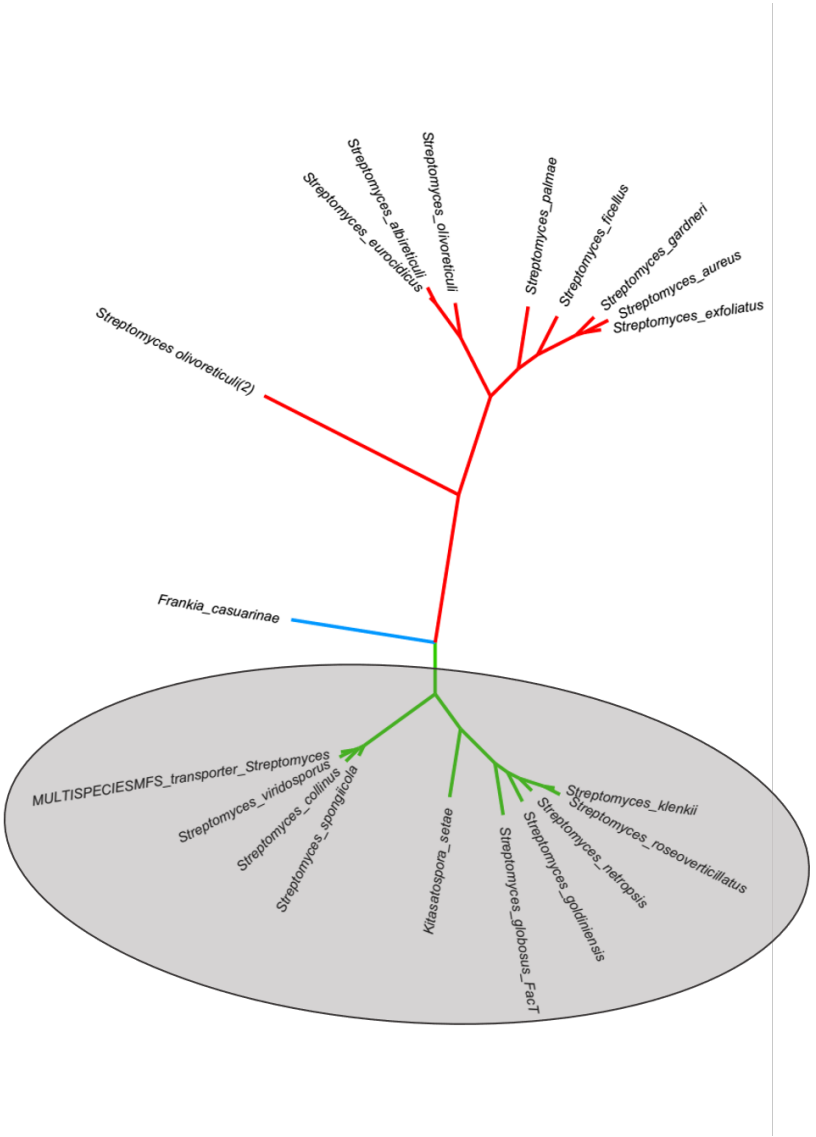


Figure 4-9: **Phylogenetic Tree of the MFS Family Transporter AurT and related proteins.** Phylogenetic analysis of AurT family by Maximum Likelihood method (MEGA7)**.** Evolutionary history was inferred by using the Maximum Likelihood method using 500 bootstraps method based on the Tamura-Neil model . The tree with the highest log likelihood is shown. The tree is drawn to scale, with branch lengths correlating to the number of substitutions per site (scale bar). The analysis involved 20 amino acid sequences, which were taken from the top 100 most similar proteins from blastP, with designated species. All positions containing gaps and missing data were eliminated.. The clade in which AurT clusters is highlighted.

4.2.10 Understanding the Evolution of the Aurodox Biosynthetic Gene Cluster: Comparing the biosynthetic potential of *S. goldiniensis* to the kirromycin producer *S. collinus* and the botrromycin A2 producer, *S. bottropensis.*

The aurodox 'supercluster' is 278 kb in length and is predicted by antiSMASH to encode for several independent specialised metabolites including bottromycin, aurodox, concanamycin A and a novel lassopeptide (*Chapter 3*). It is unsurprising that *S. goldiniensis* encodes these, given its phylogenetic relatedness to *S. bottropensis* (*Chapter 3*), a bottromycin A2 producer, as phylogenetic relatedness has been shown to correlate with biosynthetic output (Jensen and Fenical, 2005; Machado, Eva C Sonnenschein, *et al.*, 2015; van der Meij *et al.*, 2017). A direct comparison of the *S. goldiniensis, S.collinus* and *S. bottropensis* genomes demonstrated significant homology between *S. goldiniensis* and *S. bottropensis*, with much lower levels of homology between *S. goldiniensis* and *S. collinus* (*Figure 4-10).*

Given these observations it was hypothesised that there is a significant commonality between the biosynthetic capabilities of *S. goldiniensis* and *S. bottropensis*, in addition to the obvious overlap which exists between *S. goldiniensis* and *S. collinus* in the aurodox/kirromycin clusters. Furthermore, a thorough understanding of BGCs from each species at a genetic level may yield further information regarding the evolutionary events which lead to the formation of these BGCs, specifically, the aurodox supercluster. For example, the presence of repetitive regions may have facilitated recombination events. To test this idea, a range of computational analysis methods were applied to each genome sequence to characterise, primarily, each BGC within each of the strains, but also, how these clusters relate to each other. Firstly, comparative genome mining using antiSMASH was carried out to generate a comparison of the classes of BGC found in *S. goldiniensis, S. collinus* and *S. bottropensis.* From this analysis, a clear overlap between BGCs and the types of

molecules produced by each of the strains could observed (*Figure 4-11*). Unsurprisingly, several molecules were encoded for by all three strains including conserved molecules such as geosmin, desferrioxamine A and melanin (Hwang *et al.*, 2014) .

Interestingly, common biosynthetic gene clusters between *S. collinus* and *S. goldiniensis* were limited to those encoding common *Streptomyces* metabolites, with the exception of the aurodox and kirromycin clusters. On sequencing of the *S.collinus* genome, it was concluded that two copies of the kirromycin gene cluster were present on each arm of the chromosome (Rückert *et al.*, 2013), with the homology indicated to the aurodox cluster in *S. goldiniensis,* cluster 13 in *Figure 4-12.* The distinct specialised metabolisms of *S. collinus* and *S. goldiniensis* validates their distant relationship, demonstrated by AutoMLST analysis (*Chapter 3*). Similarly, the overlap in biosynthetic gene clusters between *S. collinus* and *S. bottropensis* extended only to conserved *Streptomyces* molecules. In *S. collinus,* the kirromycin cluster is not flanked by bottromycin encoding genes as is the case for the aurodox BGC, but with genes encoding enzymes involved in primary metabolic processes and therefore, there was no predicted homology between the kirromycin and bottromycin A2 cluster of *S. bottropensis.* In addition, this analysis revealed that of the 34 BGCs in the *S. goldiniensis* genome, 26 were shared with *S. bottropensis,* hence confirming the relatedness of the strains extended to biosynthetic capabilities.

To further characterise the shared BGCs between *S.goldiniensis, S. collinus* and *S. bottropensis,* BiGSCAPE (Navarro-Muñoz *et al.*, 2020) was used to construct BGC similarity networks. BiGSCAPE analysis enables the grouping of BGCs together in terms of their relatedness at an amino acid level, using the MiBIG database (Medema *et al.*, 2015) to infer molecular output, therefore, clusters encoding similar molecules are clustered together. The output from this analysis further demonstrates the

similarity between the specialised metabolism of *S. goldiniensis* and *S. bottropensis* (*Figure 4-12).* The hybrid nature of the bottromycin A2/aurodox region, in addition to the overlap in biosynthetic capabilities between *S. goldiniensis* and *S. bottropensis* contribute significantly to sequence homology. Therefore, these analyses have led to the hypothesis that inter and intragenomic recombination between repetitive regions of the *S. collinus, S.bottropensis* and *S. goldiniensis* genomes have resulted in the formation of the aurodox/kirromycin cluster, and therefore, the evolution of this cluster should and can be further investigated.

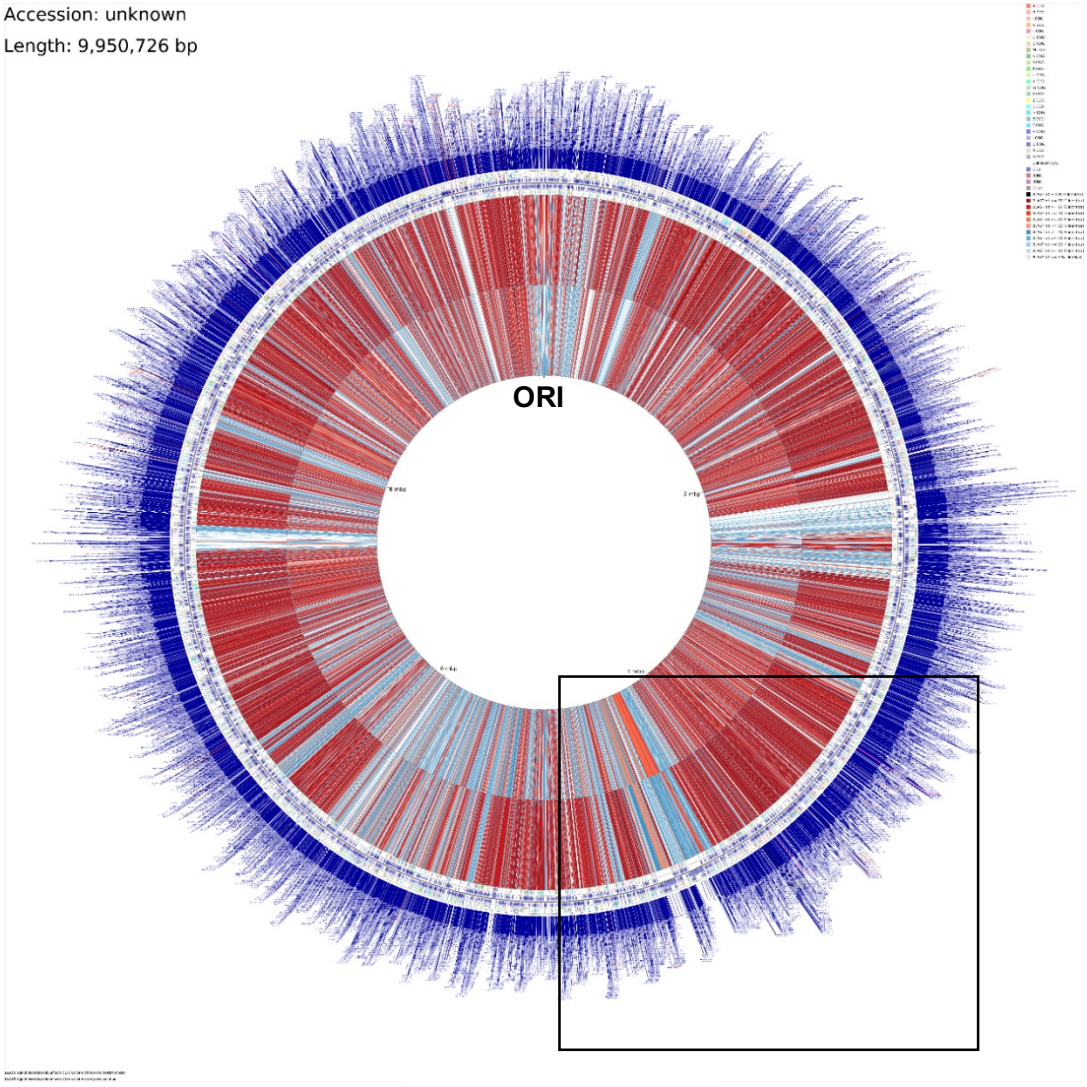


Figure 4-10: **Comparison of the *S. goldiniensis*, *S. collinus* and *S. bottropensis* Genomes.** Map was generated using CG View. Outer ring represents *S. goldiniensis* as annotated reference genome, with *S. collinus* on the inside and S. *bottropensis* in the centre. Key describes colours relating to amino acid similarity. The region of the genome which encodes the aurodox supercluster is highlighted. The homology between the aurodox encoding cluster and kirromycin cluster is clear, indicated by the red colouring, here, *S. bottropensis* homology is low. However, the homology switches to bottromycin A2 at the 5' end of the cluster, where bottromycin homology exists. It is the high level of homology in this area that could facilitate homologous recombination events which may have led to the formation of the aurodox gene cluster.

.

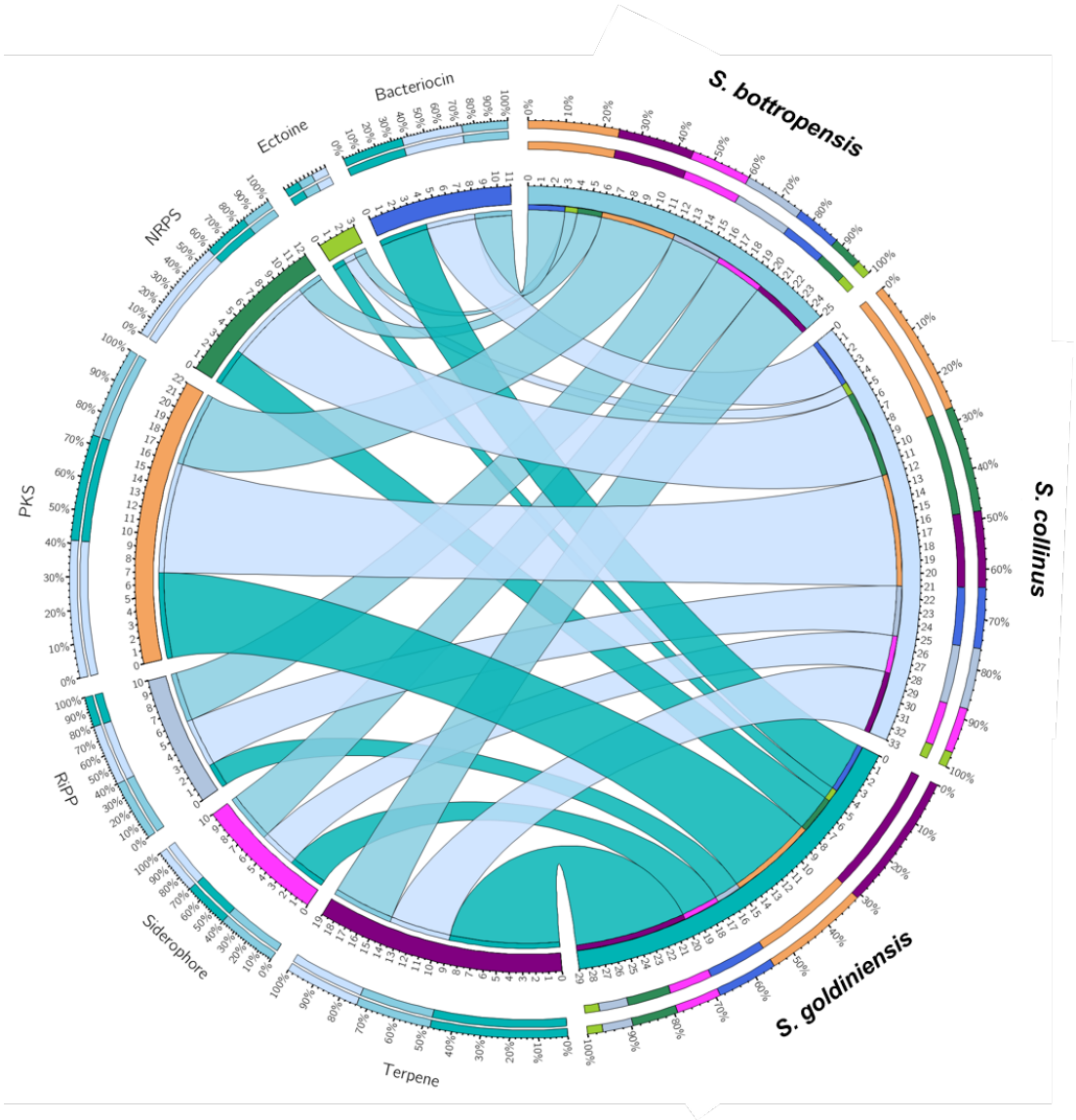


Figure 4-11: **Circos Plot Describing Gene Clusters Encoded for by *S. goldiniensis, S. collinus* and *S. bottropensis*.** Clusters predicted by antiSMASH. Percentages describe the ratio of each class of cluster encoded by each strain. Chord diagram was built using 'Circos Plot Viewer' function provided by http://mkweb.bcgsc.ca/tableviewer/.

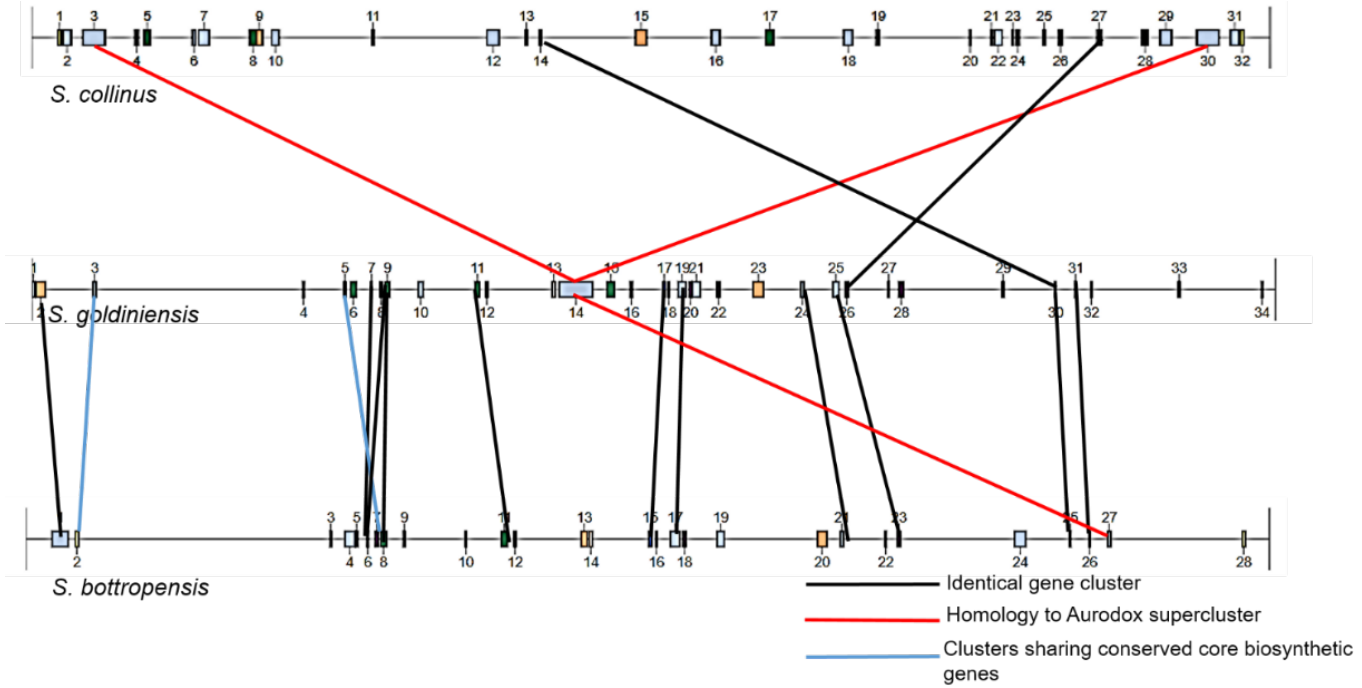


Figure 4-12: **Comparison of Biosynthetic Gene Clusters from *S. goldiniensis, S.collinus* and *S. botrropensis.*** Kirromycin clusters from *S. collinus* are 2 and 30 respectively. Cluster 14 in *S. goldiniensis* genome represents the aurodox 'supercluster'. Cluster 27 represents Bottromycin A2 cluster in *S. bottropensis.* List of gene clusters predicted by antiSMASH to be present in each strain can be found in Appendix B.

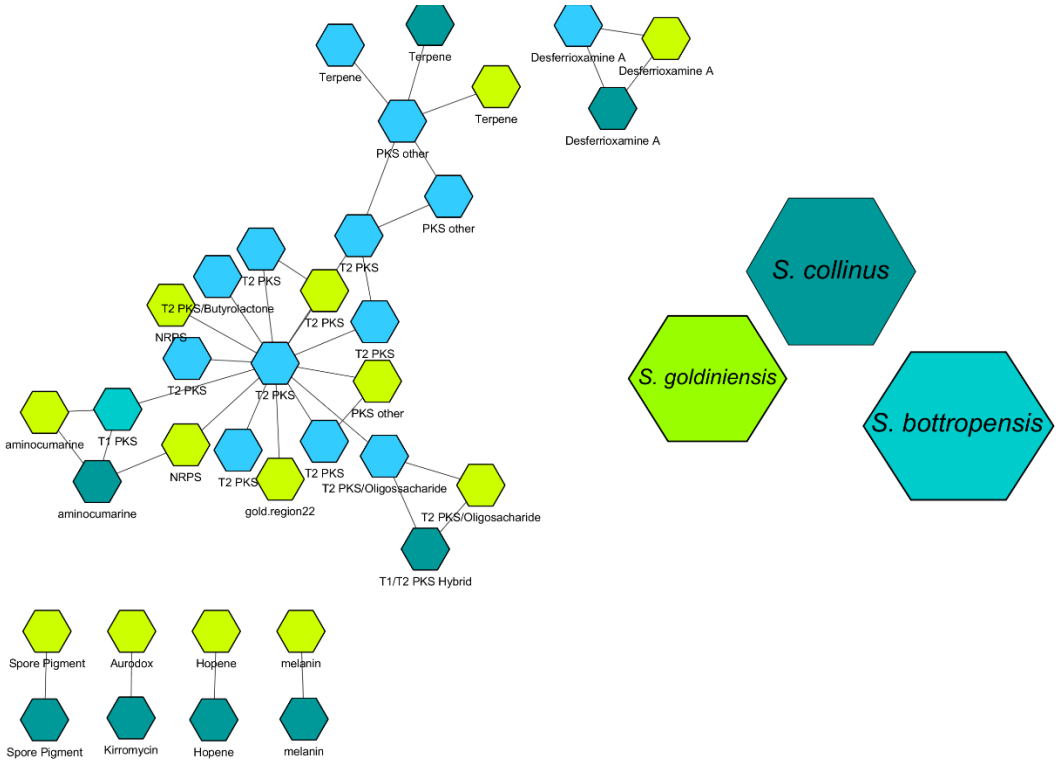


Figure 4-13: **Similarity Network from BGCs of *S. collinus, S. goldiniensis* and *S. collinus.*** Green nodes represent *S. goldiniensis* clusters, navy represents *S. collinus* and light blue represents *S. bottropensis.* Network generated by BiGSCAPE and Cytoscape.

4.2.11 Recombination Detection Programme (RDP) analysis predicts several recombination events between the kirromycin and bottromycin A2 gene clusters.

Homologous Recombination (HR) has played an important role in interspecies and intraspecies horizontal gene transfer in *Streptomyces.* It is widespread and is thought to occur at a higher rate than in other families of bacteria, and has facilitated speciation (Doroghazi and Buckley, 2010). Recent research has shown that spatial interactions between *Streptomyces* species can facilitate massive gene flux between species, with indels preferentially occurring on the chromosome arms (Tidjani *et al.*, 2019). Polyketide Synthases are particularly open to HR events due to their concurrent domain arrangements. Therefore, with the presence of several regions of repetitive, homologous sequences in the aurodox/kirromycin/bottromycin genes clusters, and two copies of the cluster kirromycin and bottromycin BGCs on the chromosome arms, it was postulated that the aurodox/kirromycin cluster is a 'daughter cluster', existing as a result of a recombinatory event between the kirromycin/aurodox BGCs and bottromycin A2 gene clusters. Furthermore, this opens the idea that *S. goldiniensis* and *S. bottropensis* share a close common ancestor and it is these recombination events that have resulted in speciation (Nouioui *et al.*, 2018)

To test this hypothesis, Recombination Detection Programming was applied (Martin and Rybicki, 2000) . This software facilitates the detection of historic recombination events between genetic sequences. Six specific recombination events were detected between the gene clusters, with Bottromycin A2 consistently contributing as a minor parent. The daughter cluster and major parent however, varied between aurodox and kirromycin (*Figure 4-14*). Although, these results do suggest that the aurodox/kirromycin clusters did arise as a result of HR, it does not conclusively determine which arose first. With *S. collinus* encoding two kirromycin BGCs on the

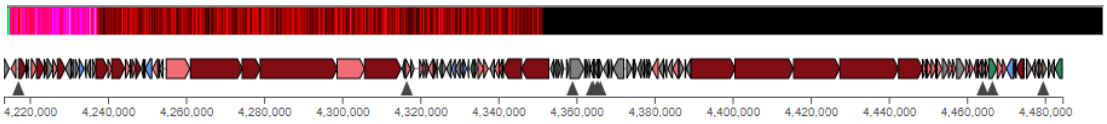
chromosome arms, and the aurodox supercluster occurring in the centre of the chromosome, in a similar position to bottromycin A2 in *S. bottropensis,* interactions between DNA on each arm facilitated between Tap and Tpg interactions could have allowed for the physical interactions required for HR, however this hypothesis is difficult to test. Moreover, the information gained in these experiments leads to the question of whether the methylation to the kirromycin molecule was gained or lost in these clusters and what the ecological role of this methylation event plays. Understanding this may give clues as to gene cluster arose first: aurodox or kirromycin?

The overlap of genes between the aurodox and kirromycin gene clusters is clear, despite the rearrangement of the BGC (*Figure 4-3,)* and therefore, HR appears to have contributed aurodox/kirromycin formation. However, other HGT events are capable of influencing the formation of BGCs and evidence of these can be found within the aurodox supercluster. Reordering events such as those that are found between the aurodox and kirromycin cluster are rare and are often the result of the transfer of core biosynthetic genes via transposition. For example, genes encoding the cladoniamides in *Streptomyces uncialis* are orthologous to genes encoding similar bis-indole alkaloids such as BE-54017, an anti-tumour molecule of unknown origin discovered from fragments of environmental DNA and rebeccamycin from *Saccharothrix aerocolonigenes* (Sánchez *et al.*, 2002; Ryan, 2011). However, these genes are distributing differently within each cluster, with transposons playing a role in these rearrangements. However, few other examples of this exist. With multiple transposases present in the aurodox 'supercluster' (*Table 4-C)*, including an IS5-type transposase 10 kb upstream of the aurodox encoding region, the contribution of transposition events in the formation of the aurodox cluster cannot be overlooked, particularly as these are not found in the regions which flank the kirromycin cluster.

***Table 4-C*: Summary of transposases within aurodox supercluster and their location**

Transposase	Location	Coding region
SMCOG1097	150212..151045	Between aurodox and concanamycin A
SMCOG1097	151326..152060	Between aurodox and concanamycin A
SMCOG1130	153527..154765	Between aurodox and concanamycin A
SMCOG1026	164686..165201	concanamycin A

A



B

Event Number	Recombinant	Major parent	Minor parent
1	Kirromycin	Aurodox	Bottromycin A2
2	Aurodox	Kirromycin	Bottromycin A2
3	Aurodox	Kirromycin	Bottromycin A2
4	Kirromycin	Aurodox	Bottromycin A2
5	Kirromycin	Unknown	Bottromycin A2
6	Aurodox	Kirromycin	Bottromycin A2

Figure 4-14: **Recombination Detection Programme Output.** Panel A highlights similarity between clusters. Pink represents bottromycin A2, red represents kirromycin, black represents aurodox supercluster. Areas susceptible to recombination (determined by homology/repetition) can be observed between bottromycin A2 and aurodox clusters followed by kirromycin and aurodox clusters. Panel B summaries specific recombination events detected by RDP algoirithm.

4.3 Summary

In this work, the putative aurodox gene cluster previously identified from the whole genome of the aurodox producer *S. goldiniensis (Chapter 3)* was subject to detailed analysis using computational techniques. This 88 kb cluster of genes can be found within a larger 278 kb cluster which is predicted to encode for several individual molecules including bottromycin A2 and concanamycin A. Using antiSMASH, BLASTp and EasyFig 2.2 it was demonstrated that the aurodox gene cluster displays homology with the kirromycin gene cluster with 23 genes of the 25 genes in the putative aurodox BGC found the kirromycin cluster. Many of the unusual features found in the kirromycin BGC were conserved in the putative aurodox encoding region, including PKS domains split across multiple genes and both *trans* and *cis* acyl transferases involved in the assembly (Weber *et al.*, 2008). Significantly, an abnormal rearrangement of genes between the aurodox and kirromycin clusters can be observed.

Given the methylated structure of aurodox, it was unsurprising that the putative aurodox cluster encodes a SAM-depenendant *O*-methyltransferase, which these analyses suggest is responsible for the conversion of kirromycin- shown biochemically to be the final aurodox intermediate- to the mature aurodox molecule. In summary, the bioinformatic analyses carried out in the chapter has indicated that the putative aurodox BGC identified in *Chapter 3* is very likely to encode for aurodox production. However, to confirm this hypothesis, the role of the cluster must be investigated using genetic and biochemical techniques.

A further aim of this chapter was to gain an understanding of the evolutionary events which resulted in the formation of the aurodox gene cluster. With *S. goldiniensis* proving to be closely related to the botrromycin A2 producer, *S. bottropensis,* the common BGCs between these two species were characterised and found to be

significant, suggesting that the two strains share a recent common ancestor. Therefore, it was postulated the vast areas of sequence similarity in the BGCs of the two genomes provides suitable conditions for HR events, as PKS units of BGCs are particularly susceptible to recombination events due to their highly repetitive domain organisation. Hence, the putative aurodox, kirromycin and bottromycin A2 BGCs were subject to RDP analysis. Here, specific historic recombination events between the clusters were identified, with bottromycin A2 consistently contributing as a minor parent, and aurodox and kirromycin clusters stochastically changing between parent and recombinant BGCs. Despite the presence of transposases within the aurodox supercluster, there was an absence of transposases in the predicted aurodox encoding region. Therefore, this work suggests that it is likely that HR has contributed significantly to and the formation of the aurodox/kirromycin clusters and to the chemical diversity in the specialised metabolome of *S. goldiniensis.* In addition, the diversity of BGCs in the environment is profound, and there could be many other BGCs in existence which may have contributed to aurodox/kirromycin formation through HR throughout their evolution, and the methods available cannot currently account for this. In addition, this analysis was unable to decipher which cluster, aurodox or kirromycin, arose first. Furthermore, these analyses could not determine whether *aurM** was derived from an intra or inter genomic source. These are questions which can only be answered through physical, experimental tests such as competition assays to determine what ecological advantages- if any- methylation to kirromycin has. Through understanding this, the order in which the aurodox and kirromycin occurred may be deciphered.

Chapter 5 : A biochemical characterisation of the aurodox Biosynthetic Gene Cluster and the mechanism of self-resistance.

5.1 Introduction

In the previous chapter the putative aurodox biosynthetic gene cluster was identified from the genome of *S. goldiniensis.* This BGC is an 88 kb hybrid Type I PKS/NRPS with significant homology to the kirromycin cluster of *S. collinus* (Weber *et al.*, 2008) and the factumycin clusters of *K. setae* and *S. globoscus* (Thaker *et al.*, 2012)*).* However, to confirm the role of this cluster in aurodox biosynthesis, molecular and biochemical analysis must be carried out. Therefore, a central aim in the following chapter was to heterologously express aurodox biosynthesis into a *Streptomyces* host in an attempt to simultaneously confirm the that discrete stretch of DNA identified is responsible for aurodox production. To facilitate this, the putative aurodox cluster was cloned into a Phage Artifical Chromosome (PAC) for expression in a heterologous host strain. Obtaining this gene cluster on a PAC also enabled targeted gene deletions using well established methodologies to investigate the role of individual genes in aurodox biosynthesis.

In *Streptomyces* species which produce elfamycin compounds, additional, elfamycin-resistant copies of EF-Tu genes (*tuf2*) are often encoded within the genome. These genes have specific mutations within the elfamycin and GTP binding sites which not only prevent aurodox from binding but prevent the inhibition of GTP hydrolysis by these molecules (Thaker *et al.*, 2012). However, some elfamycin producers such as *Streptomyces cattleya* do not encode *tuf2* genes yet appear to display immunity. With most elfamycin gene clusters also encoding immunity genes such as the MFS- type exporter found in the aurodox and kirromycin gene clusters, the role of *tuf2* genes in

aurodox resistance is not clear. Therefore, the role of this protein in resistance and biosynthesis was explored.

Moreover, these studies aim to characterise the relationship between the enzymatic steps involved in aurodox and kirromycin biosynthesis. Computational analysis detailed in *Chapter Four* has revealed significant homology between the kirromycin and putative aurodox BGCs. Unsurprisingly, the aurodox cluster was predicted to encode the additional SAM-dependent *O*-methyltransferase, AurM*, which did not display homology to the genes of the kirromycin cluster. On analysis of the aurodox biosynthetic pathway, it was hypothesised that this enzyme is responsible for the conversion of the final precursor in the pipeline, kirromycin, to the mature aurodox molecule. Therefore, to validate these results the role of AurM* in aurodox biosynthesis was investigated through the expression of this protein in the kirromycin producer *S. collinus.*

5.2 Results

5.2.1 Cloning of the aurodox biosynthetic gene cluster into the Phage Artifical Chromosome pESAC-13A.

To confirm the role of the putative aurodox gene cluster in aurodox biosynthesis, a heterologous expression approach was taken. This required the cloning of the aurodox cluster into a suitable vector for expression in heterologous *Streptomyces* hosts. Although cloning methods such as linear plus linear homologous recombination (LLHR) and transformation associated recombination (TAR-cloning) have been described (Bian *et al.*, 2012; Y. Li *et al.*, 2015), the aurodox gene cluster is significantly larger in size than the largest gene clusters cloned by these methods (55 kb). Therefore, a more traditional approach was taken which involved the preparation of a PAC library from the genomic DNA of *S. goldiniensis.* During this process, genomic DNA was digested sequentially with BamHI to yield fragments of 150-200 kb, which were ligated in to the pESAC-13A vector, encoding an apramycin resistance cassette and a phage-derived ϕC31 integrase gene (Sosio, 2001, *Figure 5-1*). This service, provided by Bio S&T (Saint-Laurent, Canada) enabled the construction of a PAC library with 20 x genome coverage, stored in *E. coli* DH10-β. These PACs could then be screened for the presence of the aurodox cluster using PCR tests with primers designed to amplify regions of the cluster. Firstly, the 1920 clones generated were screened using a primer designed to amplify a DNA fragment from the centre of the aurodox cluster. PACs which tested positive for this region were then subjected to additional screens which identified the presence of the cluster boundary regions on the PAC. The oligonucleotides for this screen were designed, validated and sent to Bio S&T with *S. goldiniensis* gDNA. The regions that they have been designed to amplify are indicated in *Figure 5-2A/B.*

From the PCR screening of the PACs carried out by Bio S and T, two triple positive clones were identified (*Figure 5-2C*) – i.e. clones that exhibit an amplicon for all three PCR products. Before further analysis was carried out using these PACs *in house*, the presence of the entire putative cluster within these PACs was re-confirmed using PCR (*Figure 5-2D).* Experiments at Bio S and T also used PFGE to determine insert size (*Figure 5-3*).

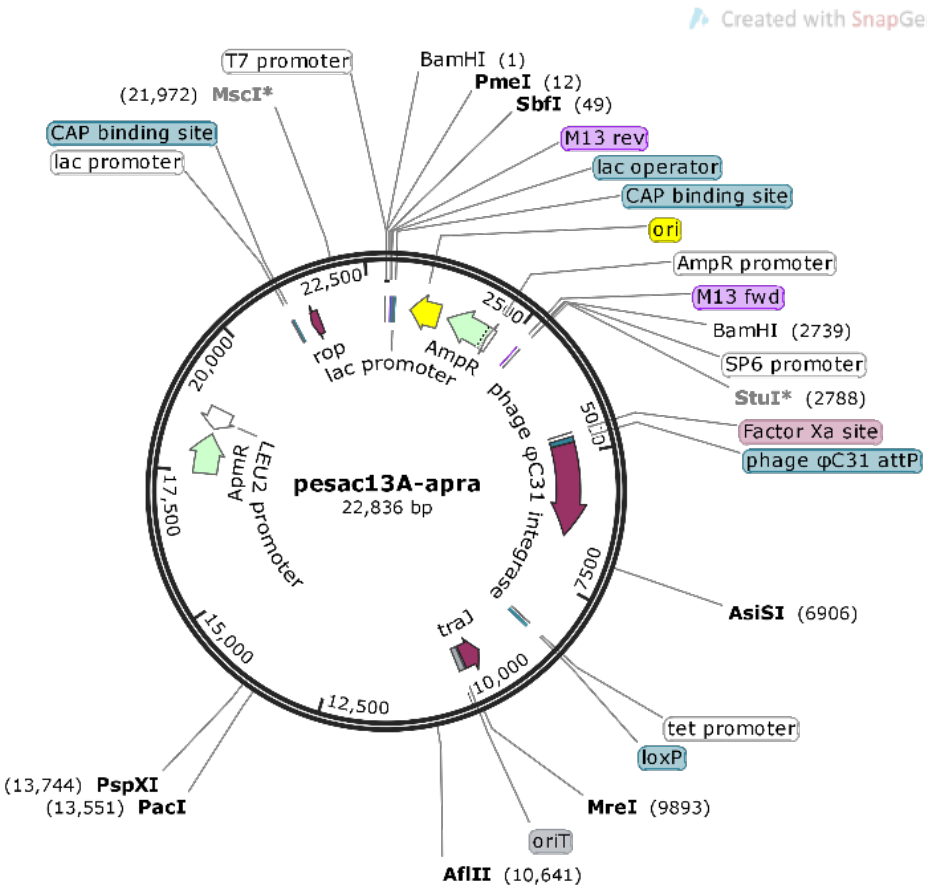


Figure 5-1: Plasmid map of pESAC-13A (Sosio, 2001). Vector encodes ϕC31 integrase, ampicillin resistance gene (replaced by gDNA at BamHI sites during PAC construction) and apramycin resistance gene.

.

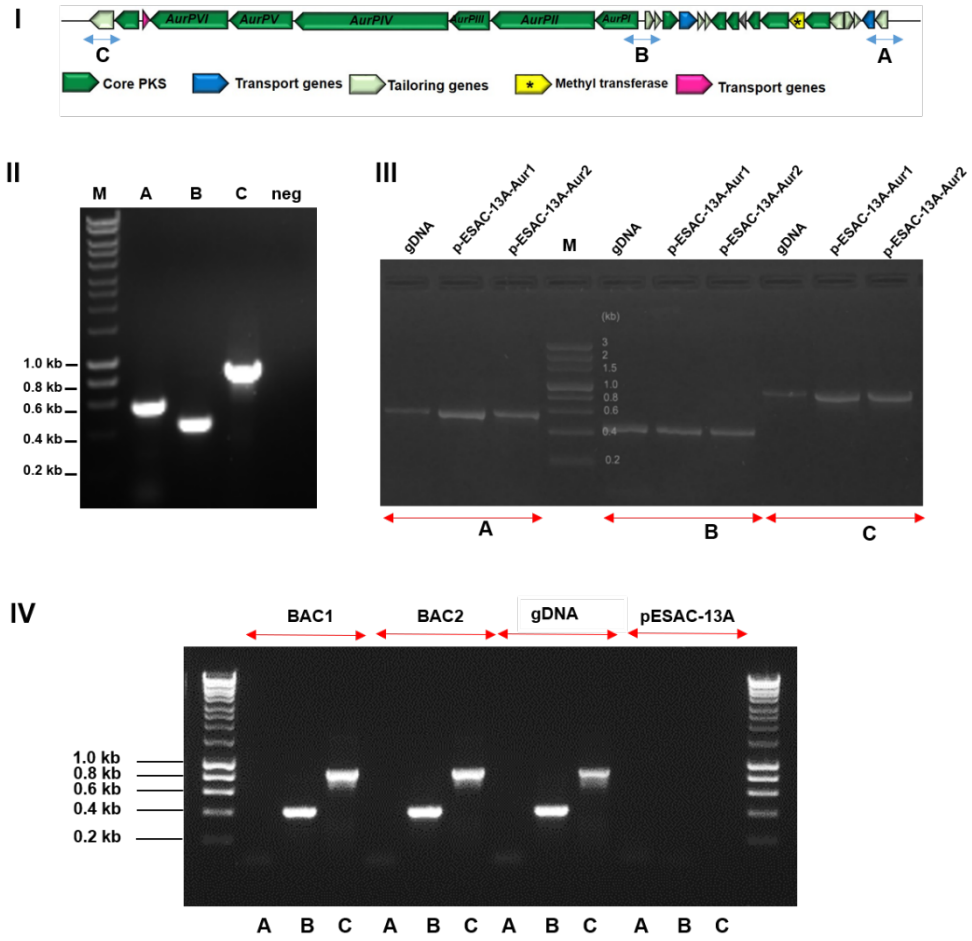


Figure 5-2: **Screening of pESAC-13A PAC library for triple positive aurodox clones.** (I) Map of aurodox gene cluster indicating target regions for PCR amplification during PAC screening. II) Agarose gel electrophoresis showing design of primers for Bio S and T screening and validation of their results. PCR products are 580, 512 and 863 bp respectively. Sequencing of amplicons from eurofins was used to confirm PCR products matched the putative aurodox cluster (III) Identification of two triple positive clones for putative aurodox encoding cluster, agarose gel electrophoresis carried out by Bio S and T (IV) PCR and agarose gel electrophoresis confirmation that PACs derived from Bio S and T screening were positive for the aurodox cluster.

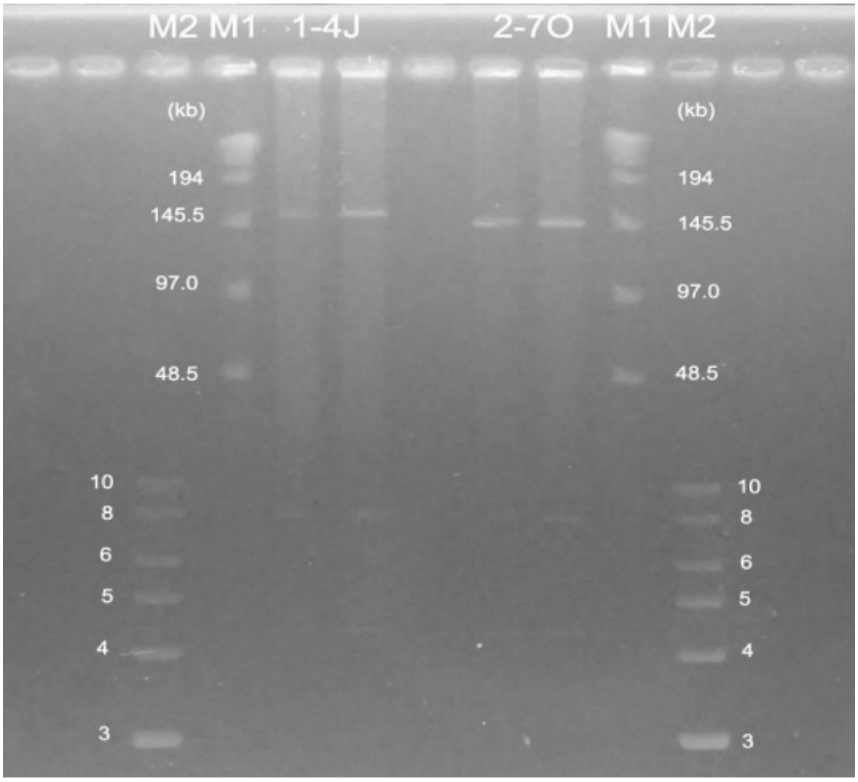


Figure 5-3: **Pulse field gel electrophoresis for confirmation of PAC insert size**. PAC1 (1-4J) and PAC2 (2-7O) were digested with BamHI and subjected to PFGE. Insert sizes were 160 kb and 155 kb respectively. Image was supplied and PFGE was carried out by Bio S and T.

5.2.2 Selection of suitable heterologous hosts for aurodox expression.

Heterologous expression (HE) is a robust method for the confirmation of the involvement of specific BGCs in specialised metabolite biosynthesis. In recent years, this method has not only allowed for the characterisation of BGCs of known metabolites but has facilitated the activation of 'silent' or 'cryptic' biosynthetic gene clusters (Huo *et al.*, 2019). A wide range of heterologous hosts have been used to enable the expression of natural product gene clusters, with varying success (Baltz, 2010). Historically, heterologous host strains were selected for characteristics which allowed for efficient fermentations and high natural product yields. *Streptomyces albus* J1074 has historically been used for BGC expression, as it reaches high levels of biomass yet does not produce significant concentrations of native small molecules under standard fermentation conditions (Sánchez *et al.*, 2002). The practicality of *S. albus* as a heterologous host was effectively demonstrated when it was used for the HE of steffimycin biosynthesis (Gullón *et al.*, 2006). In addition, *S. albus* has proved valuable for the HE of some non-*Streptomyces* BGCs such as the thiocoraline BGC from *Micromonospora spp* (Lombó *et al.*, 2006). Several other wild type strains such as *S. lividans* and *S. avermitilis* have also been used for HE of BGCs (Baltz, 2010).

In recent years, a more targeted approach has been taken to the generation of strains for HE, with several strains engineered specifically for the expression of heterologous natural product gene clusters (Baltz, 2010). One such host is *Streptomyces coelicolor* M1152. This strain, derived from *S. coelicolor* M145 has a reduced genome due to the deletion of four endogenous BGCs encoding for actinorhodin, prodiginine, the Type I PKS coelimycin C1 and Calcium-dependent antibiotic (CDA). In addition, point mutations were made in *rpoB* and *rpsL* to pleiotropically increase specialised metabolite yield (Gomez-Escribano and Bibb, 2012). This strain has been used extensively for HE, with the common occurrence of yield improvement cementing the

use of the term ‘superhost’ to describe this strain (Iftime *et al.*, 2015; Q. Li *et al.*, 2015; Braesel, Tran and Eustáquio, 2019).

For the majority of antibiotic natural products, the resistance mechanism is located within the associated BGC (Alanjary *et al.*, 2017). However, with elfamycin compounds, multiple immunity genes encoded at distinct loci can contribute to resistance. Primarily, an efflux pump encoded within the BGC such as KirT/AurT, and also an additional copy of EF-Tu (*tuf2*) located elsewhere (Thaker *et al.*, 2012). The presence of these *tuf2* genes is widespread in *Streptomyces* and can be found in over 100 *Streptomyces* genomes. *Therefore*, the elfamycin sensitivity/resistance phenotype of potential heterologous hosts had to be investigated to ensure that aurodox expression could occur without effects on the hosts that could limit production. To account for this, several strains which have previously been used for HE of BGCs including *S. albus, S. venezuelae, S. collinus* and *S. coelicolor* M1152 were analysed at a whole genome sequence level to identify EF-Tu copy number and amino acid sequences. These were subsequently subjected to phylogenetic analysis to identify strains with related EF-Tu genes to the *tuf2/tuf3* genes of the elfamycin producers which are involved in elfamycin resistance.

From this analysis, it was identified that the kirromycin-type elfamycin producers *S. collinus, S. goldiniensis* and *S. ramocissimus* encode copies of EF-Tu which form a phylogenetic clade, separate from the conserved Tuf1 proteins of *Streptomyces*. In addition, *S. coelicolor and S. venezuelae* encode *tuf2*-like EF-Tus which cluster with those of elfamycin producers (*Figure 5-4*). Two hypotheses were proposed for the reason behind the presence of these genes in *S. coelicolor* and *S. venezuelae* which do not produce elfamycins. Firstly, these genes could have evolved or been acquired via HGT as a defence mechanism for ecological encounters with elfamycin-producing bacteria. In addition, it is possible that a cryptic BGC of *S. coelicolor* or *S. venezuelae*

encodes or is associated with a cluster which encodes an elfamycin or has evolved from an elfamycin cluster that would have required *tuf2* for immunity .

Following on from the computational identification of *tuf2* carrying strains, four strains were carried forward to the next stage of screening for a suitable heterologous host for aurodox production. These were *S. collinus, S. albus* and *S. venezuelae* chosen for their *tuf2* genotype, and *S. albus* due to its sound characteristics for use as a heterologous host. Consequently, these strains, as well as *S. goldiniensis* found to encode *tuf2* 20 kb upstream of the aurodox supercluster, were subject to aurodox sensitivity assays. In the kirromycin producer *Streptomyces ramocissimus* which encodes three copies of EF-Tu, the resistant protein *tuf3* was found to be expressed specifically in the stationary stages of growth, coinciding with kirromycin biosynthesis (Olsthoorn-Tieleman *et al.*, 2007). To account for this, the sensitivity of both spores and mycelium (after four days culture in GYM) to aurodox was analysed (*Figure 5-5*).

From this analysis it would appear that aurodox resistance in *S. goldiniensis* is growth phase dependent, with spores unable to grow in the presence of 1 mg/ml aurodox. This was also the case for *S. albus,* however this strain also showed a delayed sporulation phenotype when mycelia were exposed to aurodox. In the case of *S. goldiniensis,* it was hypothesised that the expression of *tuf2* coincided with the expression of the aurodox gene cluster. Hence reverse transcriptase-PCR experiments were carried out to determine the onset of *tuf2* expression. However, with *tuf1* and *tuf2* sequences highly similar at a nucleotide level, discriminative primers for each gene could not be designed. In addition, *S. coelicolor* M1152, *S. venezuelae* and *S. collinus* displayed an aurodox resistance phenotype throughout their life cycle. Therefore, *S. albus, S. collinus, S. coelicolor* M1152 and *S. venezuelae* were tested as heterologous hosts for expression of the aurodox encoding PACs as the aurodox

resistant phenotype shown is likely increase the capacity of the strains for aurodox biosythesis.

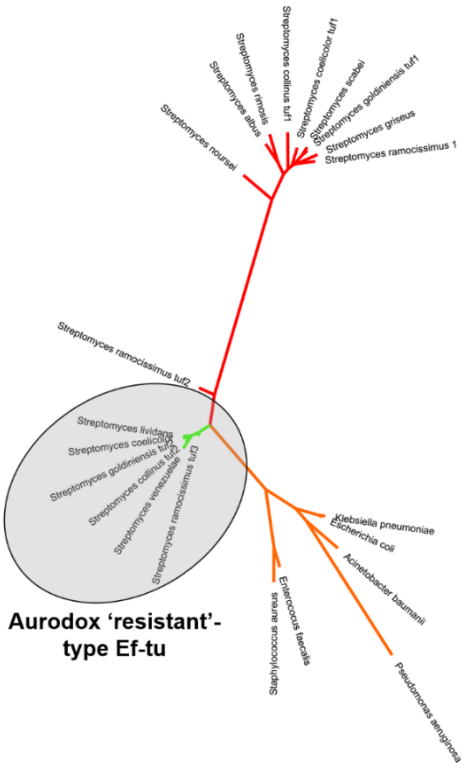


Figure 5-4: **Phylogenetic analysis of Elongation Factor Thermo-unstable amino acid sequenced.** The evolutionary history was inferred by using the Maximum Likelihood method based on the JTT matrix-based model using 500 bootstraps. Tree with the highest log likelihood (-4310.3386) is shown. The percentage of trees in which the associated taxa clustered together is shown next to the branches. Initial tree(s) for the heuristic search were obtained automatically by applying Neighbor-Join and BioNJ algorithms to a matrix of pairwise distances estimated using a JTT model, and then selecting the topology with superior log likelihood value. The tree is drawn to scale, with branch lengths measured in the number of substitutions per site. The analysis involved 22 amino acid sequences. All positions containing gaps and missing data were eliminated. There was a total of 327 positions in the final dataset.

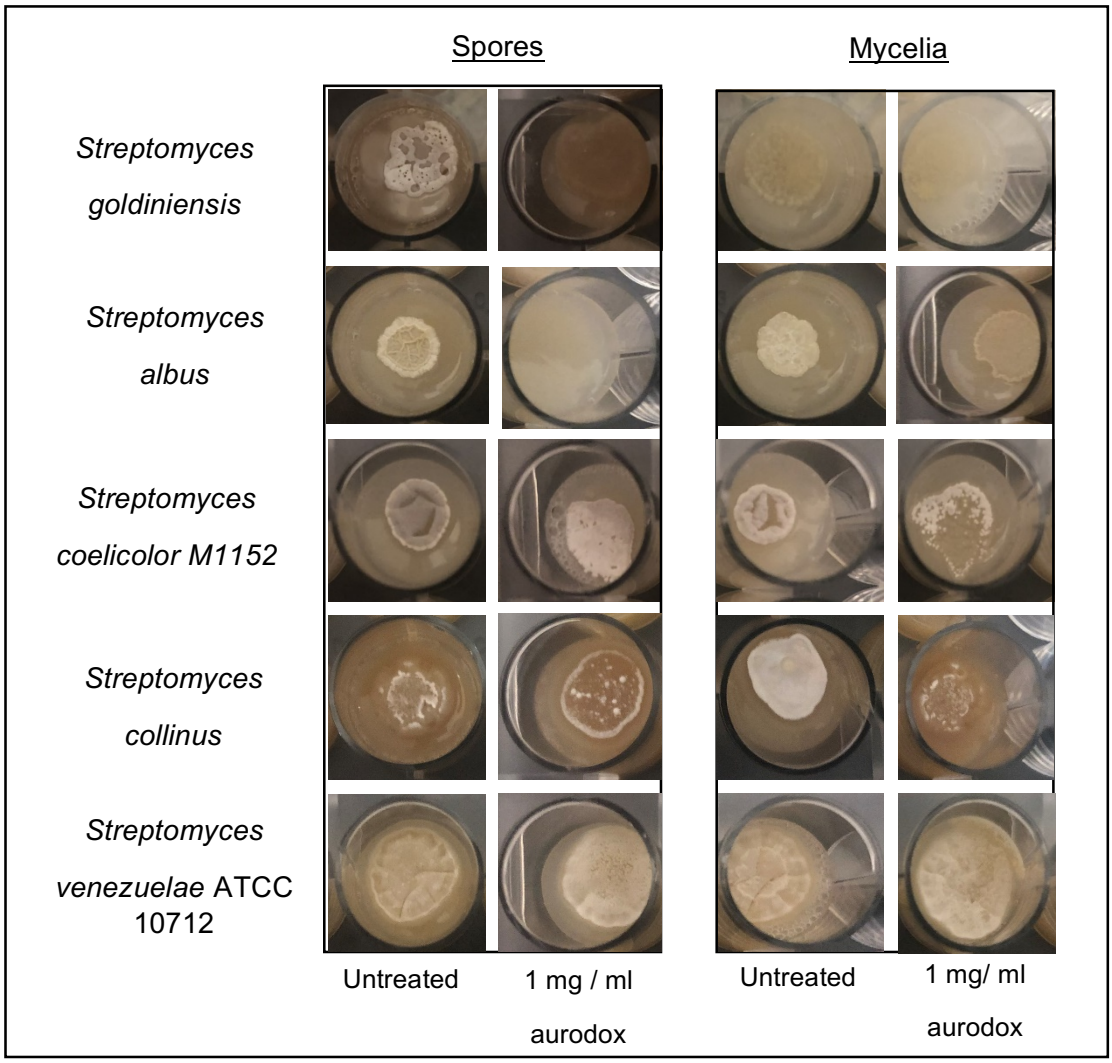


Figure 5-5: **Succeptibility of *Streptomyces* species to aurodox.** Strains were cultivated in 12-well plates on MS agar for 5 days, with or without 1 mg/ml aurodox in DMSO. Untreated samples were grown in the same concentration of DMSO. For the spore succeptibility tests, approximately 10^6 *Streptomyces* spores were spotted on to agar. Mycelia was harvested after four days in liquid GYM media and spotted on to agar.

5.2.3 Heterologous expression of pESAC-13A-AurI in multiple *Streptomyces* hosts.

To facilitate the transfer of pEASAC-13A-AurI into the chosen *Streptomyces* hosts, a conjugative approach was taken. This typically involves the transformation of the the non-methylating *E. coli* strain ET12567/pUZ8002, with the vector of interest, whilst also harbouring the pUZ8002 plasmid which encodes the machinery required for the transfer of vectors from *E.coli* to *Streptomyces* (Kieser *et al.*, 2000a). However, due to the large size of this PAC (~180 kb), transformation of ET12567/pUZ8002 was not possible and therefore tri-parental mating was used to facilitate the transfer of pESAC-13A-AurI in to ET12567. This approach involves the use of *E. coli* DH10-β containing the vector of interest, in addition to *E. coli* DH10-β containing the driver plasmid pR9604 (*bla*) and ET12567 (*cat*). By culturing the strains together on antibiotic-free agar, horizontal gene transfer is encouraged (Qin *et al.*, 2017). The resulting colonies can then be streaked on to selective agar to obtain ET12567 + vector of interest + pR9604. In the case of pESAC-13A-AurI, individual colonies were streaked on to LB agar with apramycin, carbenicillin and chloramphenicol and screened for the presence of the PAC by colony PCR (*Figure 5-6*).

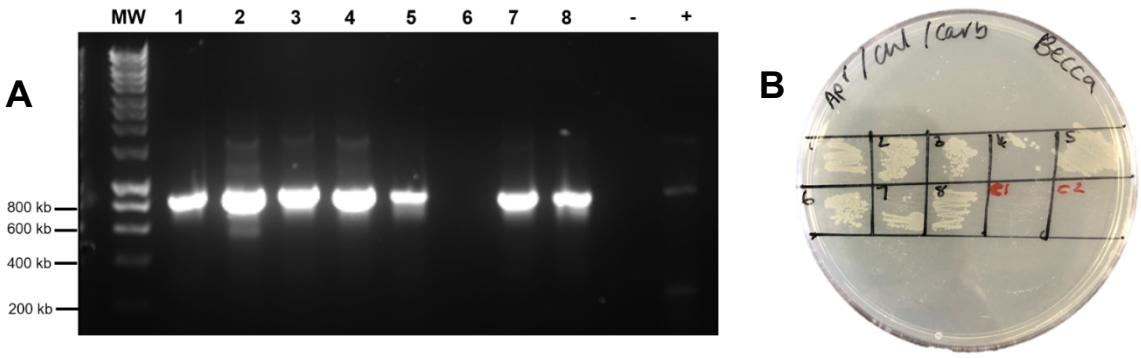


Figure 5-6: **Confirmation of pESAC-13A in ET12567.** The PAC used for subsequent studies was pESAC-13A-Aur1 (1-5J, *Figure 5-3*) (A) Colonies were picked post tri-parental mating from selection on apramycin, chloramphenicol and carbenicillin and patched-streaked on to selective antibiotics. Colony PCR was carried out using primer set C (from Bio S and T screening, *Figure 5-2*). Negative control lane (-) represents pESAC-13A empty vector, and positive control (+) represents pESAC-13A-aurI. (B) Patch-streaked colonies are represented by lanes 1-8. Controls are ET12567 and DH10-β + pESAC-13A-AurI (C1 and C2 respectively).

Conjugation of pESAC-13A-Aur could then be executed using the ET12567 + pR9604 + pESAC-13A-AurI strain generated from the tri-parental mating experiments. Primary exconjugants were patched on to MS agar containing apramycin and nalidixic acid, with the wild type as a control (*Figure 5-7A*). For each host, one exconjugant lineage was carried forward, with exception of *S. collinus* where two were carried forward, to investigate the effect of the heterogeneous bald and sporulating phenotypes observed in the exconjugant population. *Streptomyces* BGCs, particularly those derived from HGT events, are found to be highly unstable and prone to mutation (Chen *et al.*, 2002; Komatsu *et al.*, 2010). Therefore, the presence of the entire putative aurodox cluster was confirmed in each of the exconjugants by colony PCR (*Figure 5-7B).*

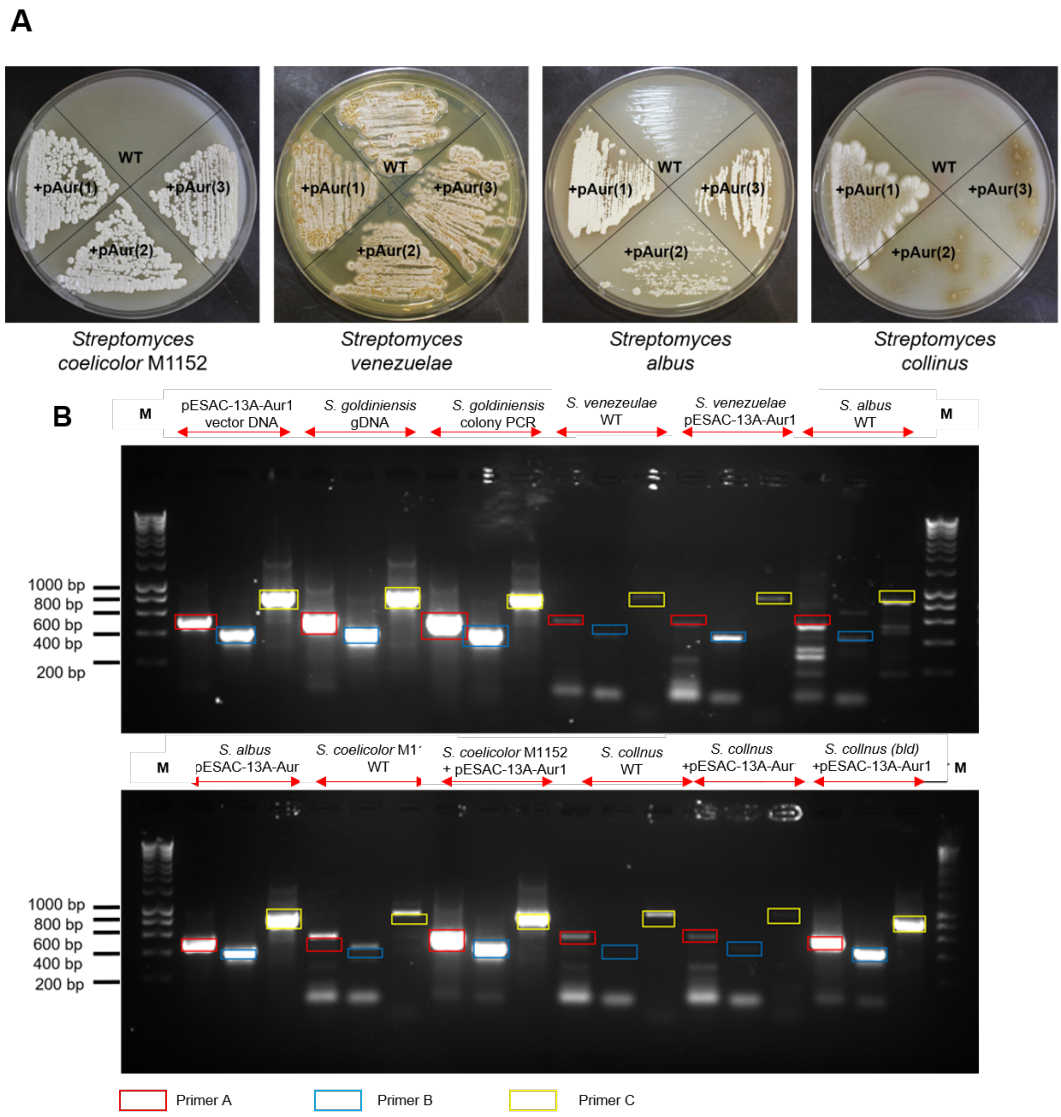


Figure 5-7: **Conjugation of pESAC-13A-Aurl into *S. coelicolor* M1152, *S. venezuelae*, *S. albus* and *S. collinus***. (A) Primary exconjugants streaked on to MS + apramycin + nalidixic acid, with the wild type shown as a control (B). The presence of the aurodox region of the PAC was confirmed using colony PCR with primer sets A, B and C from Bio S and T screening.

5.2.4 Heterologous aurodox biosynthesis can be achieved in *S. collinus and S. coelicolor* M1152

To characterise the ability of the of the engineered host strains to produce aurodox, they were cultured under the standard fermentation conditions described in *Chapter 3.* Subsequently, chloroform extracts were prepared and subjected to LCMS analysis to facilitate the detection of aurodox. In each case, the wild type strains were also analysed to enable recognition of distinct peaks in the HPLC trace specific to the strains expressing the putative aurodox BGC. In addition, an *S. goldiniensis* WT extract and 1 mg/ ml of authentic aurodox standard in DMSO were analysed for the determination and validation of the retention time associated with aurodox (7.2 minutes, *Figure 5-8*).

On preliminary analysis, a peak with a retention time of 7.2 minutes was identified in the HPLC trace from *S. coelicolor* M1152 + pESAC-13A-AurIthat was not present in the parental strain (*Figure 5-9*). When the molecules under this peak were analysed using negative mode MS, a peak corresponding to the aurodox negative-ion at a m/z ratio of 793 could be identified, indicating the presence of aurodox in the *S. coelicolor* M1152 +pESAC-13A-AurI extract. Furthermore, in addition to the kirromycin peak (6.9 minutes) found in the extracts from *S. collinus* WT, in *S. collinus* +pESAC-13A-AurI, a peak corresponding to the retention time of 7.2 minutes could also be identified, with ions present with an m/z of 793 in negative mode. These data provide evidence for aurodox biosynthesis in the *S. coelicolor* M1152 and *S. collinus* heterologous hosts and hence, demonstrate the role of the putative aurodox biosynthetic gene cluster identified from the genome sequence in the biosynthesis of the compound.

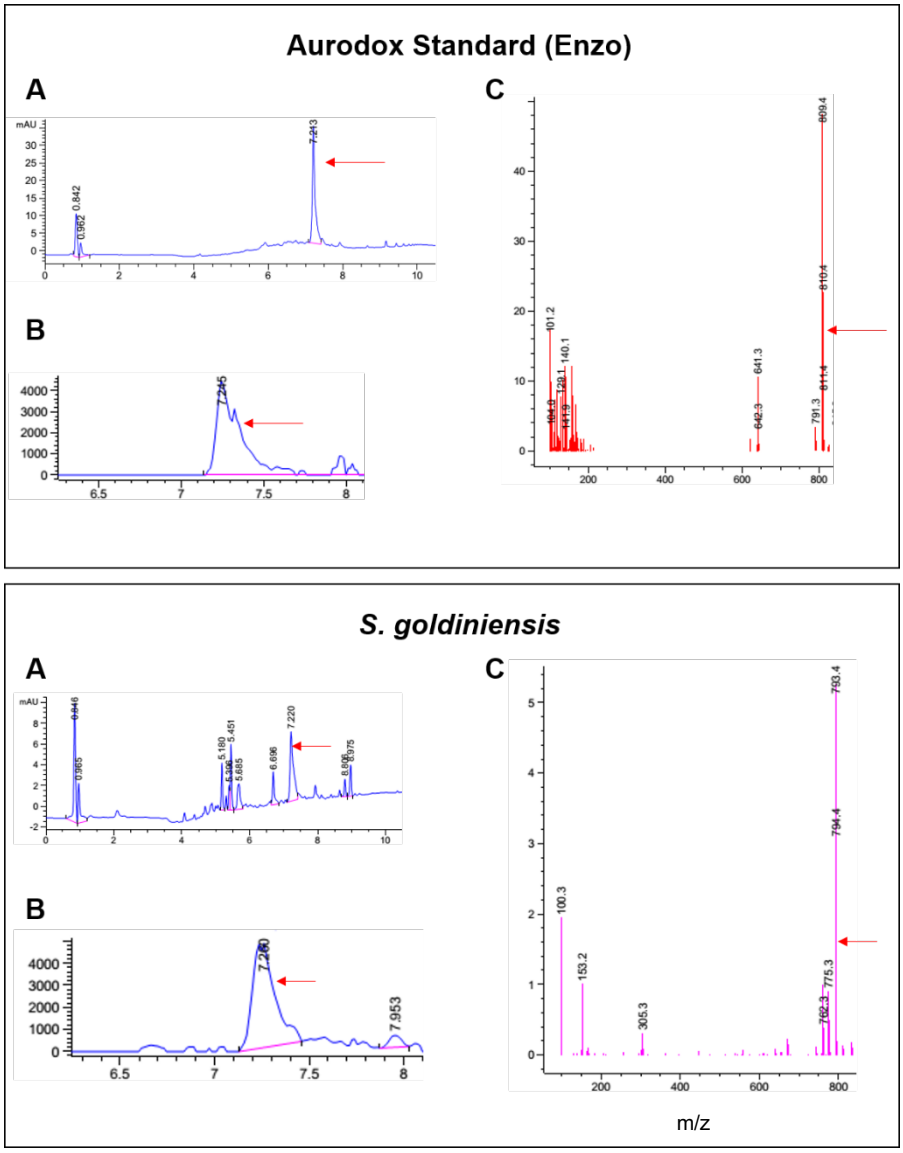


Figure 5-8: LCMS analysis of Aurodox Standard and *S. goldiniensis* extracts. Aurodox standard trace was generated using 1 mg/ml aurodox (Enzo) in DMSO. Extracts were generated according to standard aurodox fermentation and extraction protocols (Chapter 3). (A) Total ion chromatograms from HPLC analysis. Red arrows indicate retention time peak corresponding to aurodox. (B) HPLC trace showing signal intensity from aurodox associated peaks (~ 7.2 minutes). (C) Mass Spectrometry analysis of molecules under aurodox-associated peak. In standard, Positive Scan is shown and m/z of 811 shows aurodox cation. In *S. goldiniensis* negative scan mode peaks are shown, with 793 m/z peak corresponding to aurodox anion

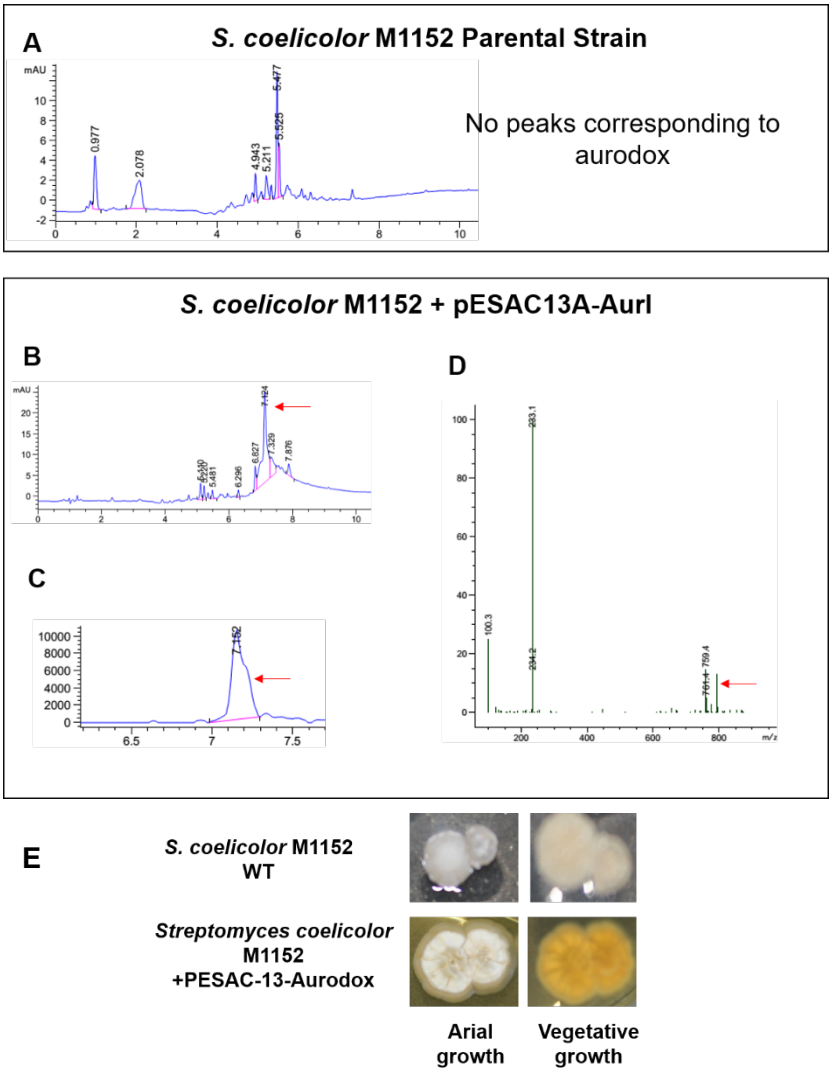
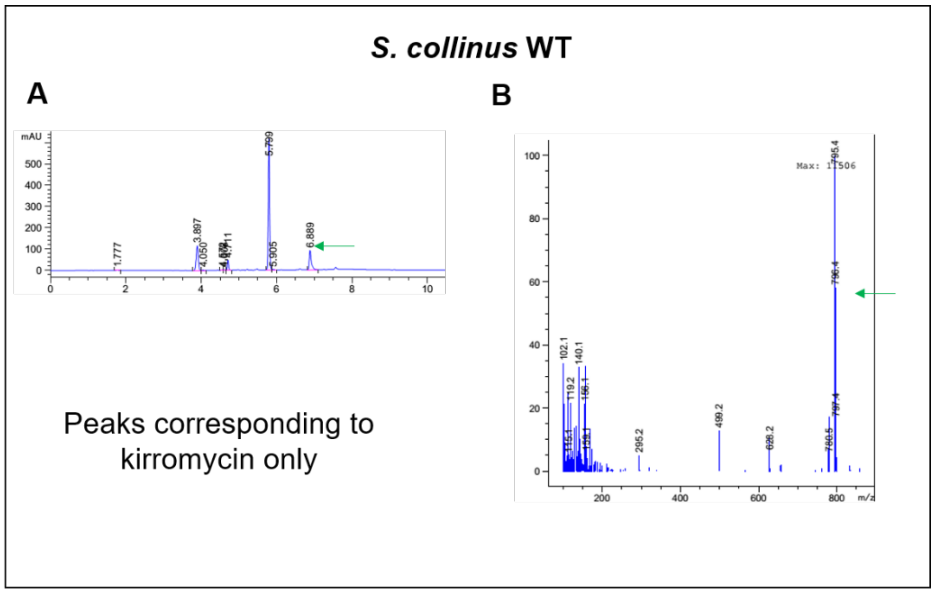


Figure 5-9: **LCMS analysis of fermentation extracts from *S. coelicolor* M1152 WT and *S. coelicolor* M1152 + pESAC-13A-AurI**. Extracts were generated according to standard aurodox fermentation and extraction protocols (Chapter 3). (A) Total ion chromatogram from HPLC analysis of S. coelicolor M1152, with no significant peaks corresponding to aurodox. (B) Total ion chromatogram from HPLC analysis of S. coelicolor M1152 + pESAC-13A-AurI, peak corresponding to aurodox (~ 7.2 minutes) is indicated by red arrow. (C) HPLC trace showing signal intensity from aurodox associated peak(~ 7.2 minutes) (D) Mass Spectrometry analysis of molecules under aurodox-associated peak. Negative scan mode peaks are shown, with 793 m/z peak



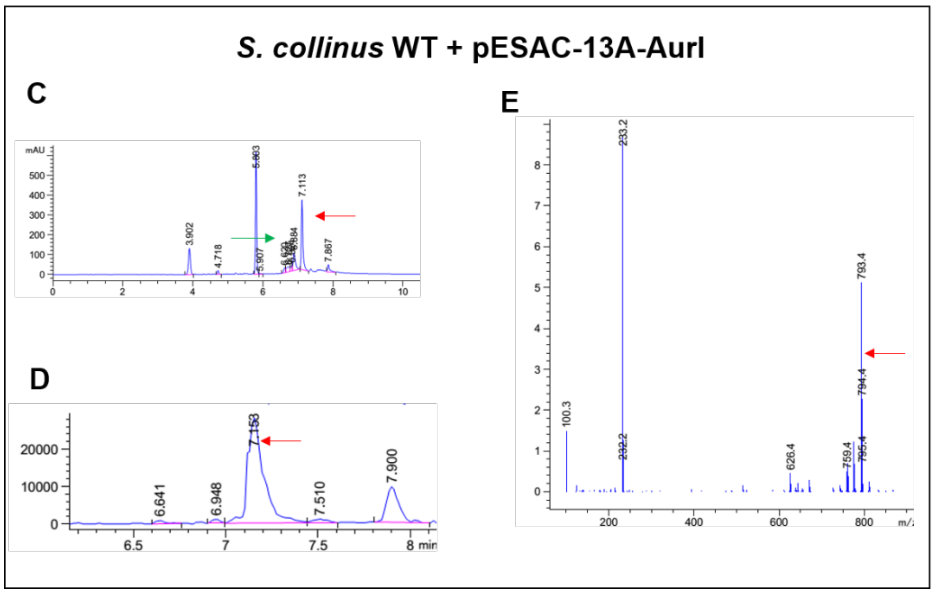


Figure 5-10: **LCMS analysis of fermentation extracts from *S. collinus* WT and *S. collinus* + pESAC-13A-AurI**. Extracts were generated according to standard aurodox fermentation and extraction protocols (*Chapter 3*). (A) Total ion chromatogram from HPLC analysis of *S. collinus*, with no significant peaks corresponding to aurodox. Peaks associated with kirromycin are indicated with green arrows (B) Mass Spectrometry analysis of molecules under kirromycin-associated peak. Positive scan mode peaks are shown, with 795 m/z peak corresponding to kirromycin anion. (C) Total ion chromatogram from HPLC analysis of *S. collinus* M1152 + pESAC-13A-AurI, peak corresponding to kirromycin (~6.9 minutes) is indicated by green arrow peak corresponding to aurodox (~ 7.2 minutes) is indicated by red arrow. (D) HPLC trace showing signal intensity from aurodox associated peak(~ 7.2 minutes) (E) Mass Spectrometry analysis of molecules under aurodox-associated peak. Negative scan

5.2.5 Aurodox biosynthesis could not be achieved in *S. albus* or *S. venezuelae* through integration of the pESAC-13A-AurI alone.

On examination of extracts from *S. venezuelae* ATCC 10712 and *S. venezuelae* +pESAC-13A-AurI, there was an absence of peaks at the appropriate aurodox retention time within the HPLC trace. Given the consistency of the fermentation and extraction protocols, it is unlikely that the absence of this compound within the extract is a result of insufficient extraction protocols or sample degradation. Therefore, this suggests the integration of the aurodox cluster on pESAC-13A-AurI alone was not sufficient to refactor aurodox biosynthesis *S. venezuelae.* Furthermore, in *S. albus*, despite the presence of a broad peak between 7 and 7.5 minutes, an aurodox-associated m/z peak could not be identified from the MS data. In addition, this peak could also be found in the wild type extract, and therefore, an association with aurodox could be disregarded. To confirm the presence or absence of aurodox production in these extracts, *K. pneumoniae* bioassays were carried out (*Figure 5-13*). The results from these growth inhibition assays confirmed that *S. coelicolor* +pESAC-13A-AurI and *S. collinus* pESAC-13A-AurI displayed greater activity against *K. pneumoniae* than the extracts from the parental strains and hence provides further evidence of aurodox biosynthesis in these strains. In *S. venezuelae* +pESAC-13A-AurI, bioactivity was diminished, suggesting that the integration of the PAC may have had negative effects on the expression of native antibiotic BGCs. For *S. albus* +pESAC-13A-Aur, bioactivity was increased from the parental strain, however, this could not be conclusively attributed to aurodox production, as disruption to the genome may have led to the upregulation of a native antibiotic BGC (Devine *et al.*, 2019). Another possibility is the presence of aurodox or its precursors in the extracts at a undetectable concentrations using LCMS which are sufficient to induce bioactivity. In summary, the integration of pESAC-13A-AurI alone was not sufficient to enable detectable aurodox biosynthesis in *S. venezuelae* and *S. albus.*

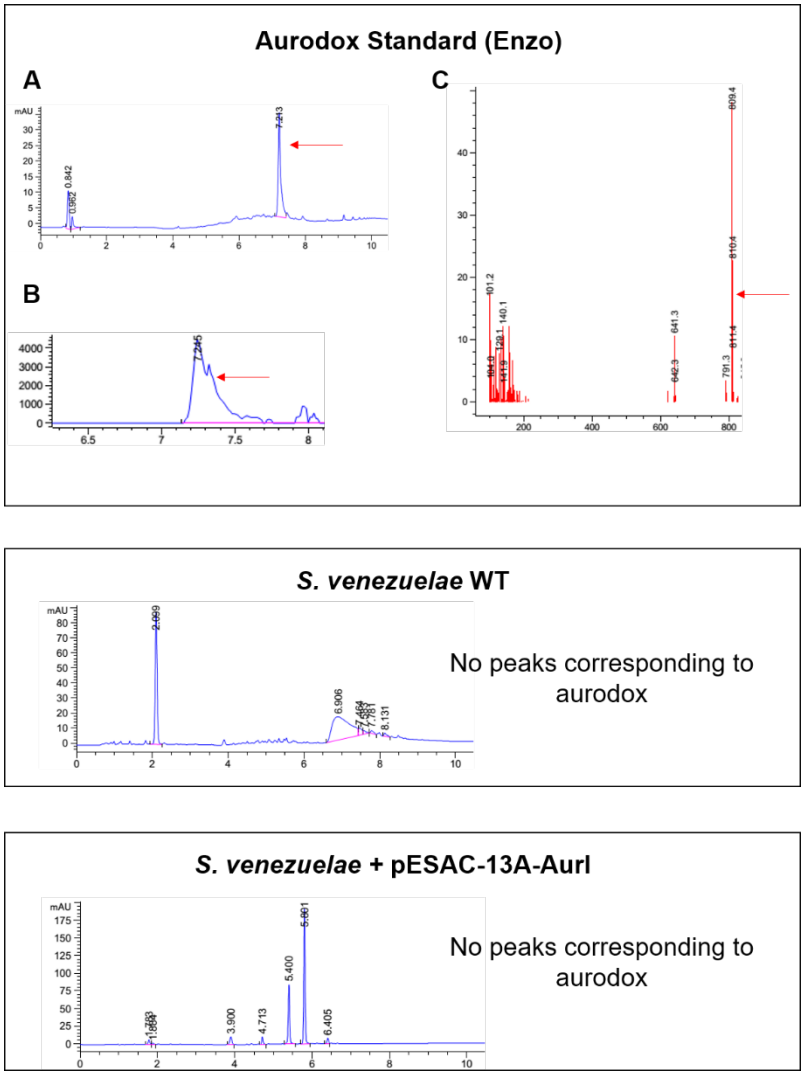


***Figure 5-11*: LCMS analysis of fermentation extracts from *S. venezuelae* WT and *S. venezuelae* + pESAC-13A-Aurl extracts**. (A) Total ion chromatograms from HPLC analysis. Red arrows indicate retention time peak corresponding to aurodox. (B) HPLC trace showing signal intensity from aurodox associated peaks (~ 7.2 minutes). (C) Mass Spectrometry analysis of molecules under aurodox-associated peak. In standard, Positive Scan is shown and m/z of 811 shows aurodox cation Extracts were generated according to standard aurodox fermentation and extraction protocols (*Chapter 3*). Total ion chromatogram from HPLC analysis of both strains are shown, with no significant peaks corresponding to aurodox.

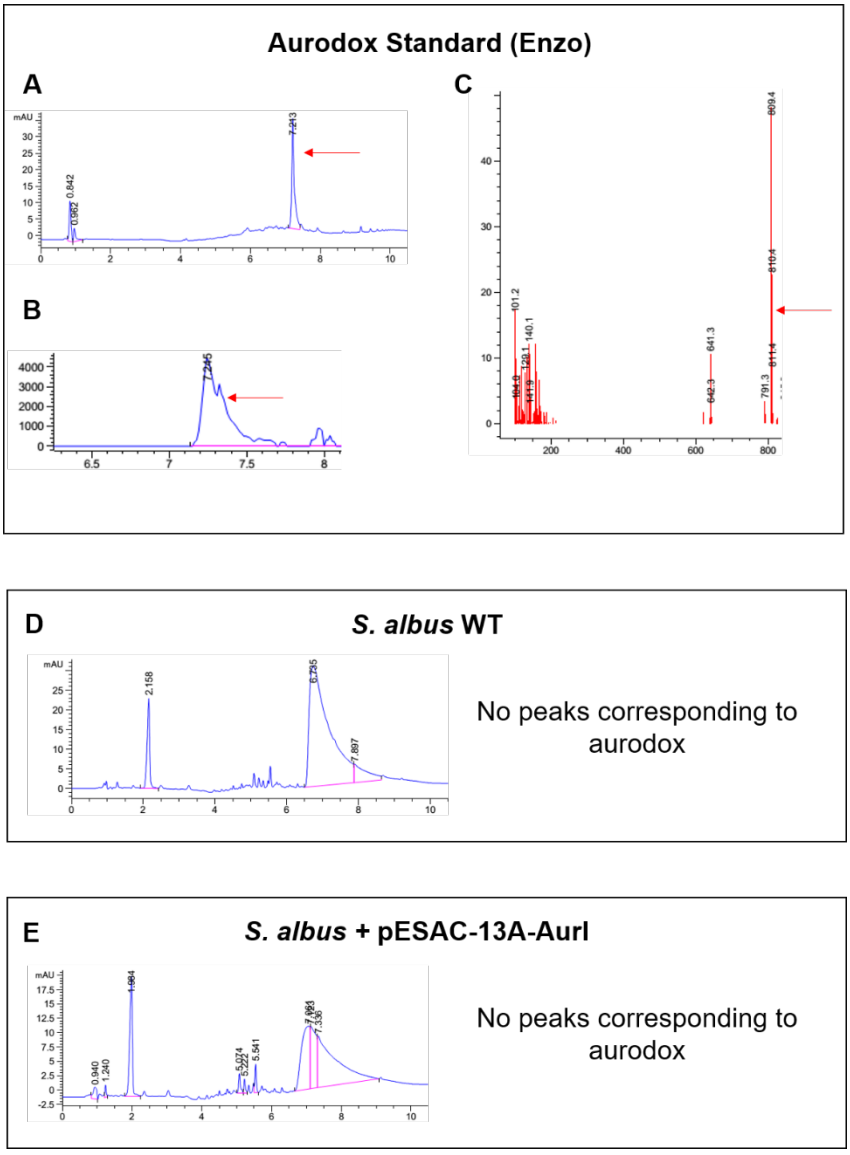


Figure 5-12: **LCMS analysis of fermentation extracts from *S. albus* WT and *S. albus* + pESAC-13A-AurI extracts**. Extracts were generated according to standard aurodox fermentation and extraction protocols (*Chapter 3*). (A) Total ion chromatograms from HPLC analysis. Red arrows indicate retention time peak corresponding to aurodox. (B) HPLC trace showing signal intensity from aurodox associated peaks (~ 7.2 minutes). (C) Mass Spectrometry analysis of molecules under aurodox-associated peak. In standard, Positive Scan is shown and m/z of 811 shows aurodox cation. In *S. goldiniensis* negative scan mode peaks are shown, with 793 m/z peak corresponding to aurodox anion. **(D/E)** Total ion chromatogram from HPLC analysis of both WT/ +pESAC-13A-AurI are shown, with no significant peaks corresponding to aurodox.

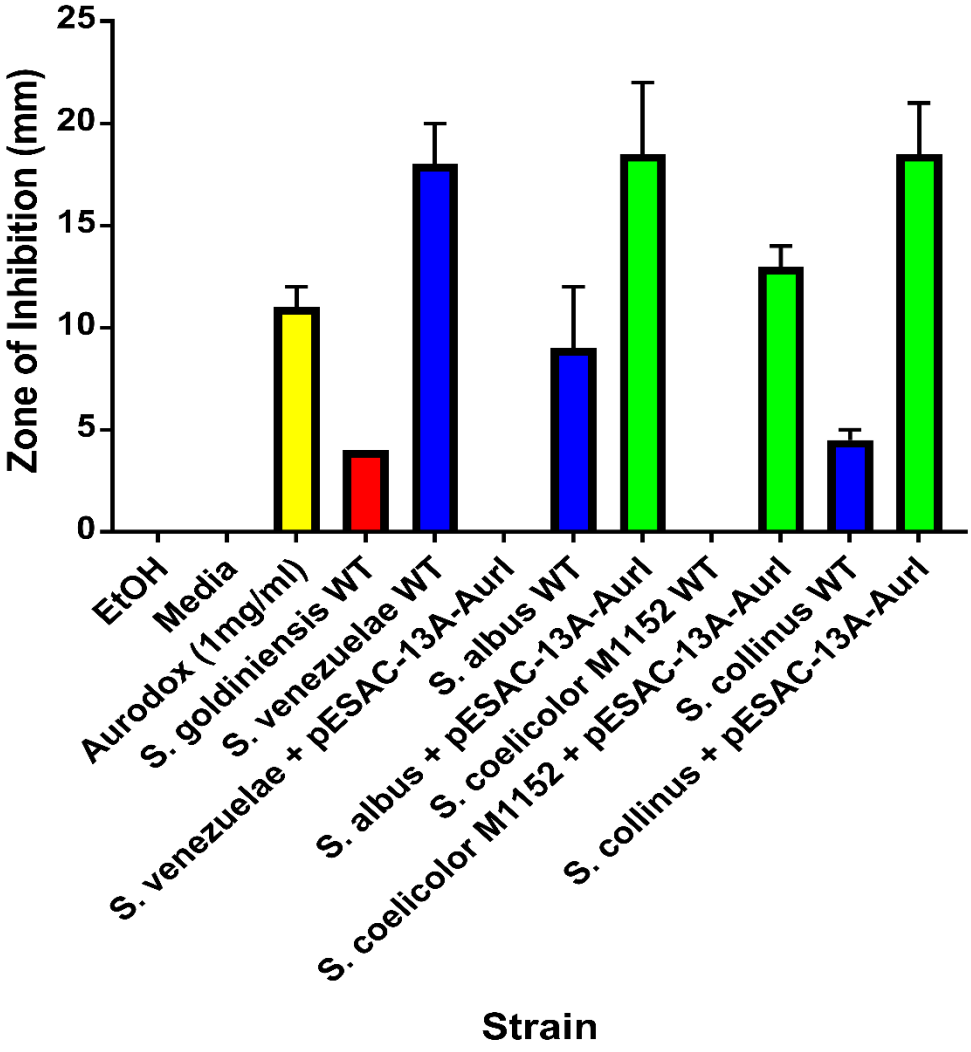


***Figure 5-13*: Activity of extracts from heterologous aurodox producers against *K. pneumoniae*.** Disk diffusion assays were used for bioactivity measurements. Assays were carried out in triplicate, with the error bars representing the standard deviation from the mean. Ethanol and Media blanks were used as negative controls with aurodox standard (Enzo) and *S. goldiniensis* extract used as positive controls

.

5.2.6 Understanding the role of *tuf2* in aurodox resistance and biosynthesis.

The function of the aurodox BGC in biosynthesis was confirmed by heterologous expression in *S. coelicolor* M1152 and *S. collinus.* Despite being able to attribute aurodox production to a specific DNA sequence, *S. albus* and *S. venezuelae* were unable to synthesise aurodox to detectable levels when harbouring this DNA sequence. In both cases, growth of the strains was limited whilst harbouring the PAC, suggesting that aurodox may have been expressed at some level, but was associated with negative consequences for the cells. Phylogenetic analysis of EF-Tu sequences from *Streptomyces* revealed that EF-Tu from *S. albus* can be found on a separate clade from the resistant *tuf2* genes found among the elfamycin producers (*Figure 5-4*). Hence this confers an aurodox-sensitive phenotype in *S. albus*. It was therefore hypothesised that the absence of a resistant copy of EF-Tu in *S. albus* was negatively impacting the ability of these strains to produce the compound whilst possessing the genetic information required to do so. In the case of *S. venezuelae* which does possess a *tuf2*-like gene and is resistant to aurodox (*Figure 5-5*), the likely reason behind the lack of aurodox biosynthesis was hypothesised to be a result of low EF-Tu copy number. Furthermore, it was hypothesised that the MFS-type exporter AurT alone is insufficient to confer resistance to aurodox.

To test this hypothesis, *tuf2* from *S goldiniensis* was cloned into the integrating vector, pMS82 (*hyg,* ϕBT integrase, *Figure 5-14*) and was expressed in each heterologous host, alongside pESAC-13A-AurI. To account for the impact of integration events on gene expression, strains were also constructed with pESAC-13A and pMS82 empty vectors. Therefore, for each host, seven strains were constructed. These are described in *Figure 5-15*.

In order to characterise the relative advantages and disadvantages of aurodox expression during competitive growth, strains were cultured adjacently to each other,

and phenotypic changes were observed between several of the constructed strains and the wild type. In *S. albus,* it appears that growth in the strain expressing *tuf2* alone is inhibited when grown in competition with *S. albus* +pESAC-13A-AurI + *tuf2.* This suggests that firstly, *S. albus* + pESAC-13A-AurI is producing aurodox leading to inhibition of *S. albus* + pMS82_tuf2 growth. Consequently, if this strain is producing aurodox, it implies that *tuf2* alone is not enough to confer aurodox resistance in *S. albus,* and that the Mfs-type protein AurT is also required. Furthermore, growth of *S. albus* + pESAC13-AurI is limited when compared to the parent. This suggests that the integration of pESAC-13A-AurI alone decreases the fitness of *S. albus,* evidenced by the appearance of suppressor mutants throughout the lawn of growth (*Figure 5-15A*). In *S. coelicolor* M1152*,* a similar pattern is observed, however, *S. coelicolor* M1152 + pESAC-13A-AurI + pMS82_tuf2 appears to limit the transition of aerial hyphae to mature spores in *S. coelicolor* + pMS82_tuf2 (Figure *5-15D*) rather than general growth. This provides further evidence that both *tuf2* and a transporter such as AurT/KirT are both required for aurodox resistance, however, the deletion of *tuf2* in *S. goldiniensis* could confirm this hypothesis which is true for *S. ramocissimus..* Observations from co-cultures of *S. collinus* constructed strains demonstrate that *S. collinus* M1152 +pESAC-13A-AurI is inhibiting aerial growth in *S. collinus* WT (*Figure 5-15B*). It does appears that the heterologously expressed aurodox from *S. collinus* inhibits the development of the *S. collinus*, which displays a consistently bald phenotype. However, with this strain possessing both KirT and *tuf2,* and resistance to extracellular aurodox previously demonstrated, the mechanism by which aurodox overcomes the native elfamycin resistance mechanisms of *S. collinus* is unclear (*Figure 5-5*). This could however, be due to the production of higher overall antibiotic yields with *S. collinus* + pESAC-13A+AurI possessing an additional BGC compared to the wild type. Therefore, to investigate this further *in vivo* conversion of kirromycin

to aurodox in *S. collinus* must be characterised. In *S. venezuelae,* each engineered strain displayed wild type morphology (*Figure 5-15B*).

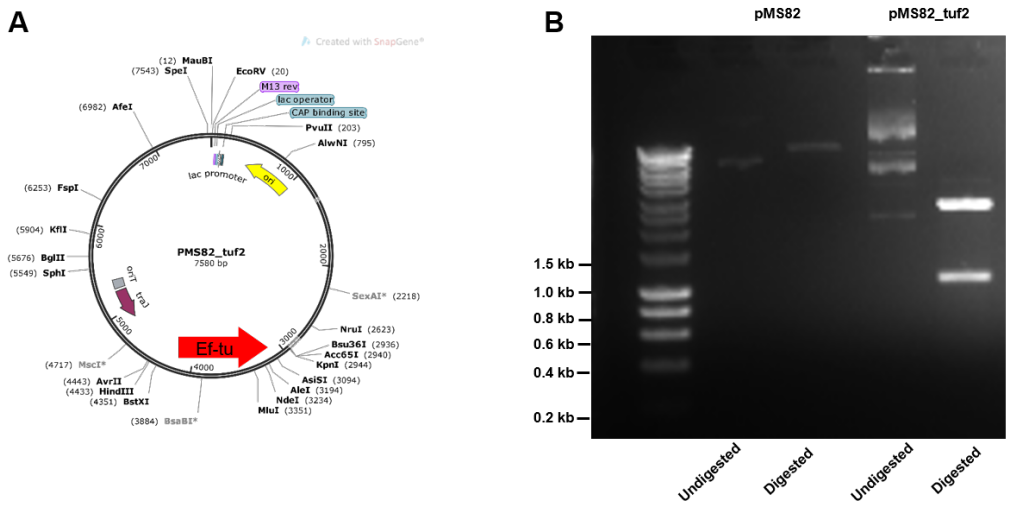


Figure 5-14: **Construction of PMS82_tuf2 by Genscript**. (A) Plasmid map if PMS_tuf2 generated by Snapgene TM. (B) Plasmid Digest of PMS82_tuf2 with BamHI and NdeI to confirm presence of 1197 bp *tuf2 gene.*

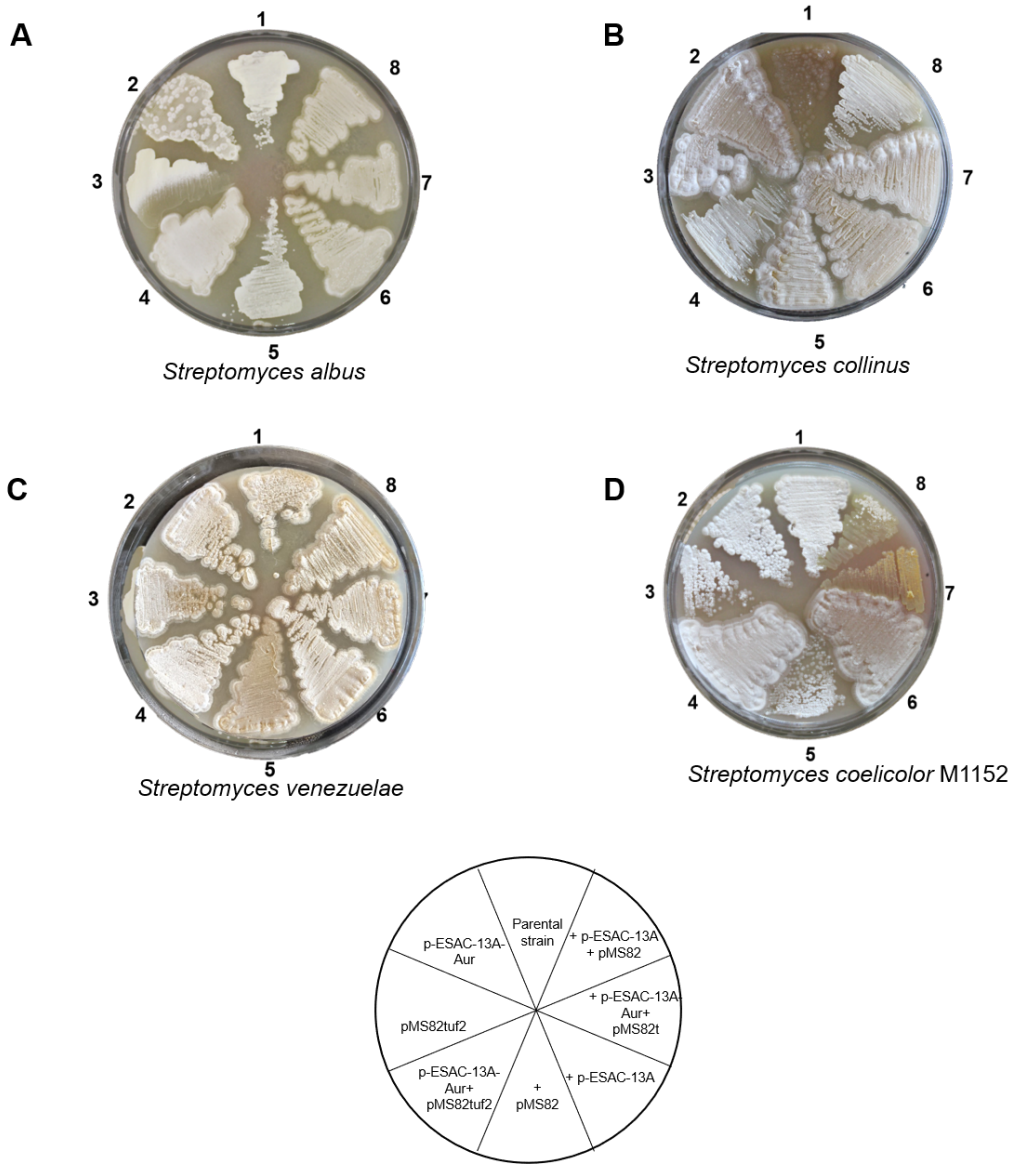


***Figure 5-15*: Images of heterologous aurodox hosts and empty vector controls.** Position of strains on MS plates are described in the legends. (A) *S. albus* (B) *S. collinus* (C) *S. venezuelae* (D) *S. coelicolor* M1152.

5.2.7 Expression of *tuf2* enables aurodox production in *S. venezuelae.*

Upon initial analysis, it was revealed that *S. venezuelae* possessed distinct properties which suggested it may be a suitable heterologous host for aurodox expression. These include a *tuf2*-like EF-Tu gene and constitutive aurodox resistance. Despite this, integration of pESAC-13A-AurI alone was insufficient for HE of the aurodox BGC

in this strain. Therefore, an additional copy of *tuf2* from *S. goldiniensis* was introduced in to *S. venezuelae* alongside pESAC-13A-AurI. LCMS analysis of the fermentation extracts from this strain revealed the presence of a sharp peak at the aurodox retention time of 7.2 minutes in the HPLC trace which was not present in the trace from the empty vector control (*Figure 5-16*). Furthermore, an MS signal was detected with an m/z ration of 793, which corresponds to the m/z of the aurodox anion (negative scan). These data confirm the presence of aurodox in the *S. venezuelae* +pESAC-13A-AurI + pMS82_tuf2 extract and therefore confirm that an additional copy of *tuf2* is required for aurodox biosynthesis in *S. venezuelae.* Furthermore, the signal intensity of the aurodox peak is 16-fold higher than the peak associated with aurodox in the *S. goldiniensis* extract, indicating that this sample can produce a significantly improved aurodox titre.

In *S. albus,* despite evidence for aurodox production in response to *tuf2* expression including the retrieval of golden extract resembling aurodox and a small peak in the HPLC at 7.2 minutes, this could not be confirmed by MS and therefore, the aurodox production status of *S. albus* + pESAC-13A-AurI + pMS82_tuf2 remains inconclusive (*Figure 5-17*).

Several hypotheses may explain the essentiality of this additional copy of EF-Tu. Firstly, it is possible that there is specificity in the *tuf2* sequence required to confer aurodox resistance. However, with no amino acid changes between the aurodox or GDP binding regions of EF-Tu from *S. venezuelae* and *tuf2* from *S. goldiniensis,* this is thought to be unlikely. However, it is possible that amino acid alterations in other regions of EF-Tu may influence aurodox sensitivity (Berchtold *et al.*, 1993). Furthermore, with *S. venezuelae* encoding only one EF-Tu gene, it is not possible that a lack of *tuf2* expression could lead to aurodox sensitivity, as this protein serves an essential process. In addition, the number of copies of EF-Tu could influence aurodox

sensitivity (*Figure 5-18*). It could conspire that this additional copy of EF-Tu encoded on pMS82_tuf2 simply facilitates an increase in abundance of the protein, leading to a change in binding stoichiometry and leading to sufficient immunity in *S. venezuelae* for aurodox biosynthesis to occur.

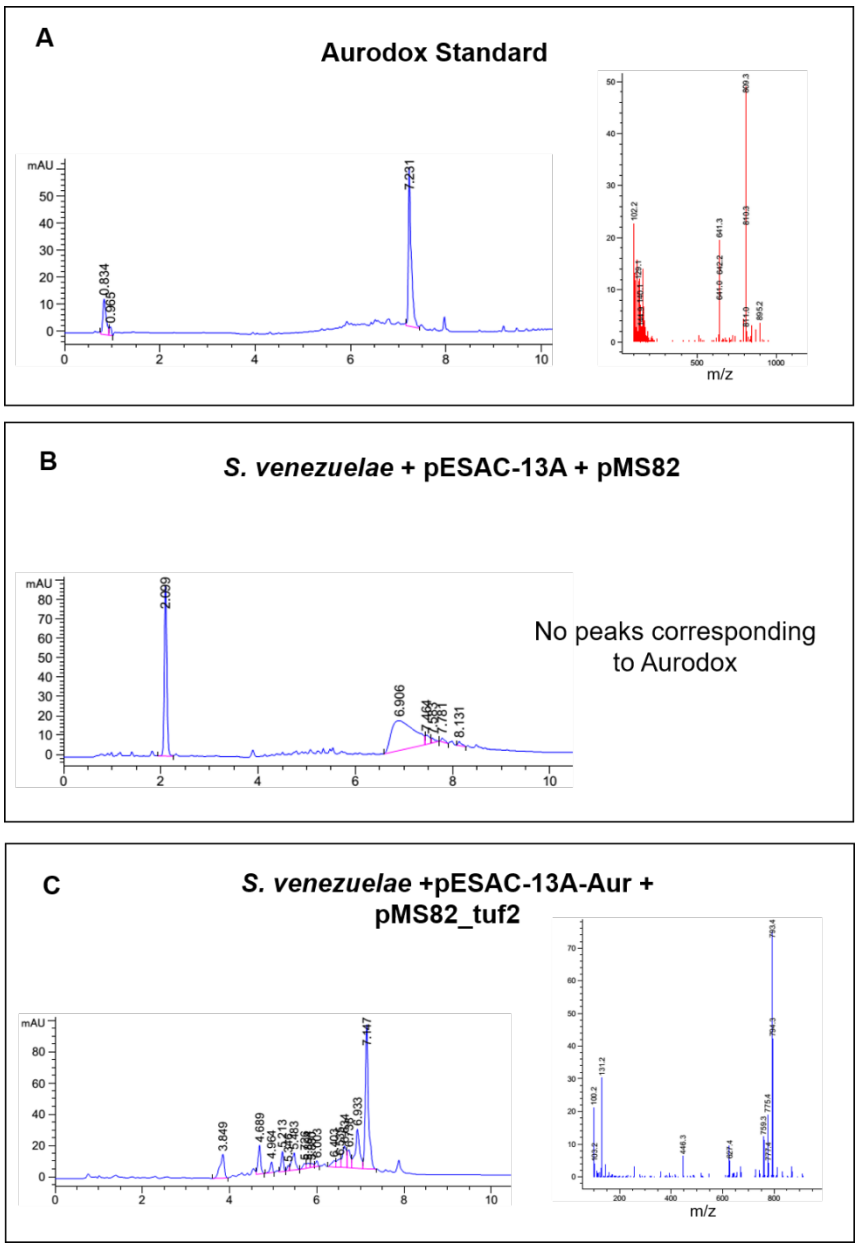


Figure 5-16: **LCMS analysis of fermentation extracts from and *S. venezuelae* + pESAC-13A + pMS82 and *S. venezuelae* + pESAC-13A-AurI + pMS82_tuf2.** Extracts were generated according to standard aurodox fermentation and extraction protocols (*Chapter 3*). (A) Total ion chromatogram from HPLC analysis of aurodox standard (1 mg/ml in DMSO, Enzo) and MS signals under aurodox retention time of 7.2 minutes. (B) Total ion chromatogram from HPLC analysis of extracts from *S. venezuelae* + pESAC-13A + pMS82. (C) Total ion chromatogram from HPLC analysis of *S. venezuelae* + pESAC-13A-AurI + pMS82_tuf2, MS data from 7.2 minute aurodox peak is shown. MS data shows ions detected in negative scan mode.

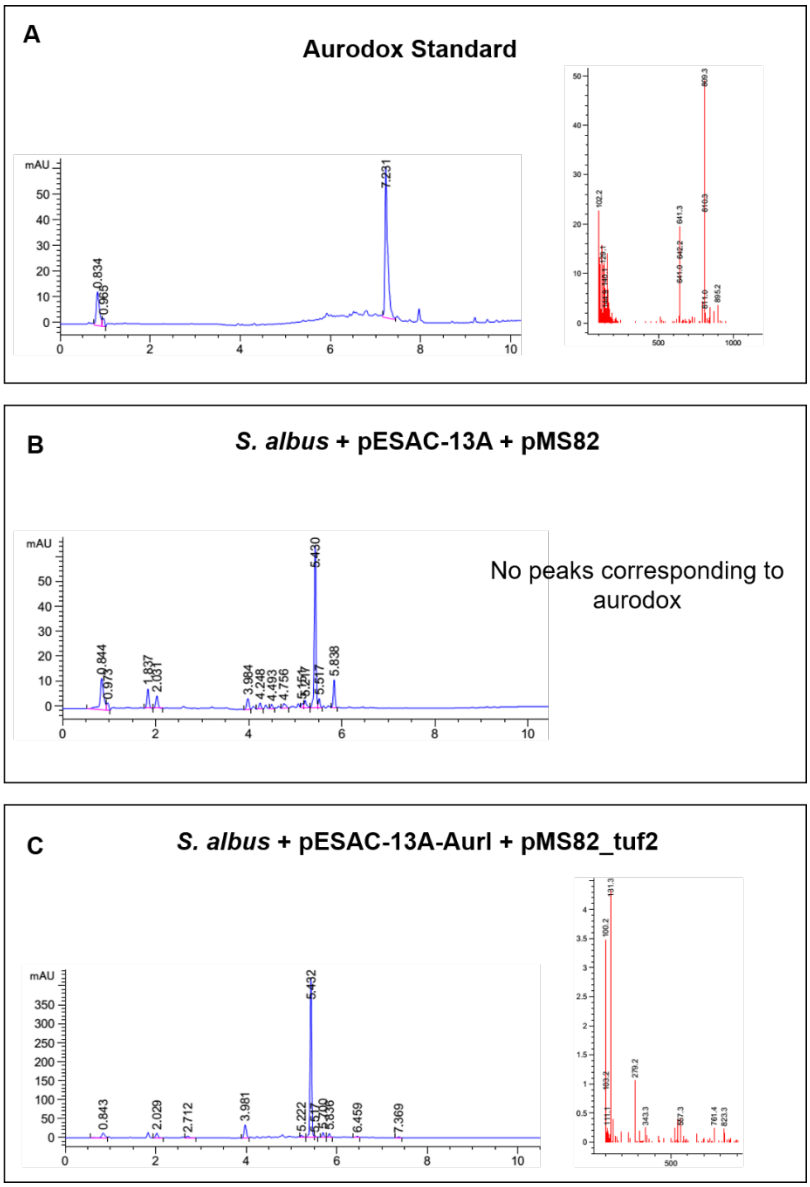


***Figure 5-17*: LCMS analysis of fermentation extracts from and *S. albus* + pESAC-13A + pMS82 and *S. albus* + pESAC-13A-Aurl + pMS82_tuf2.** Extracts were generated according to standard aurodox fermentation and extraction protocols (*Chapter 3*). (A) Total ion chromatogram from HPLC analysis of aurodox standard (1 mg/ml in DMSO, Enzo) and MS signals under aurodox retention time of 7.2 minutes. (B) Total ion chromatogram from HPLC analysis of extracts from *S. albus*+ pESAC-13A + pMS82. (C) Total ion chromatogram from HPLC analysis of *S. albus* + pESAC-13A-Aurl + pMS82_tuf2, MS data from 7.2 minute aurodox peak is shown. MS data shows ions detected in negative scan mode.

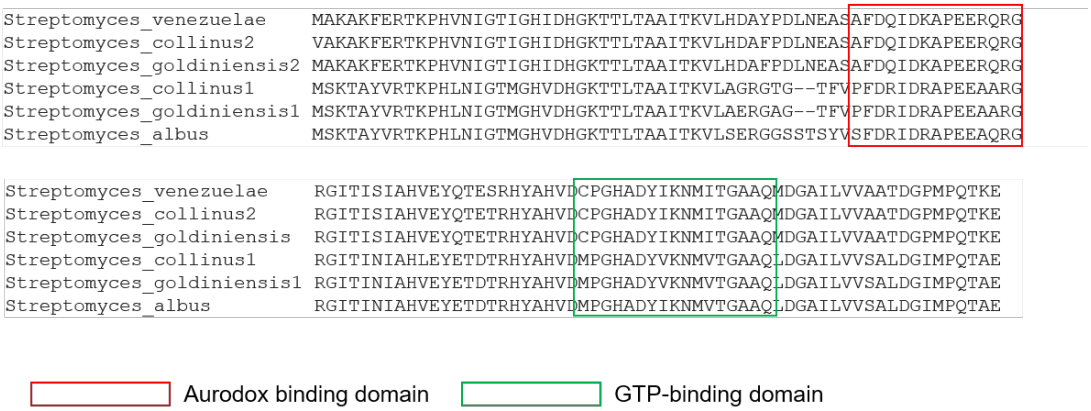


Figure 5-18*: Alignment of EF-Tu sequences from *S. goldiniensis* and *S. venezuelae. Alignments carried out by Clustal Omega on C-terminus sequences. Aurodox binding region (red) and GTP binding domain (green) are highlighted.

5.2.8 Disruption of the *aurAI* gene to confirm the role of the aurodox biosynthetic gene cluster in aurodox biosynthesis.

Heterologous aurodox biosynthesis can be achieved in multiple hosts through expression of pESAC-13A-AurI. This PAC has been confirmed by PCR to encode the entire cluster of genes with a proposed role in aurodox biosynthesis (*Figure 5-2*). However, with a total of 160 kb of the *S. goldiniensis* genome ligated in to pESAC-13A-AurI, it is likely that the predicted aurodox encoding region of pESAC-13A-AurI is flanked by PKS genes such as those from the concanamycin A or bottromycin A2 BGCs found directly upstream and downstream of the aurodox encoding region. Therefore, further evidence is required to specifically associate the PKS genes of the aurodox cluster (*aurAI-aurVII*) with the catalytic roles in aurodox biosynthesis that had been computationally characterised. To enable this, a modified version of the redirect PCR targeting of *Streptomyces* method (Gust *et al.*, 2003) was used to construct pESAC-13A-AurIΔaurAI, a mutant version of the PAC. By using this λ-*red* recombinase-based system, heterologous recombination enabled the replacement of *aurAI* with the hygromycin resistance gene (*Figure 5-19*). This was carried out by introducing pIJ10790 (*cat*) in to *E. coli* Dh10-*β* containing pESAC-13A-aurI. The hygromycin resistance gene was amplified from pIJ10700 with PCR-mediated overhangs with sequence homology to the regions directly flanking *aurAI* before arabinose induction of the λ red recombinase was carried out to facilitate heterologous recombination. The resulting mutated vector was then introduced into *S. coelicolor* M1152 via conjugation for comparison of metabolites with LCMS. Unfortunately, due to the research shutdown across the UK during the SARS-COVID2 pandemic, it was not possible to carry out this metabolite analysis, nor sequence the pESAC-13A-AurI to identify the genes flanking the aurodox cluster for the absolute determination of the aurodox BGC. However, this work will be carried out for any publications emerging from these works.

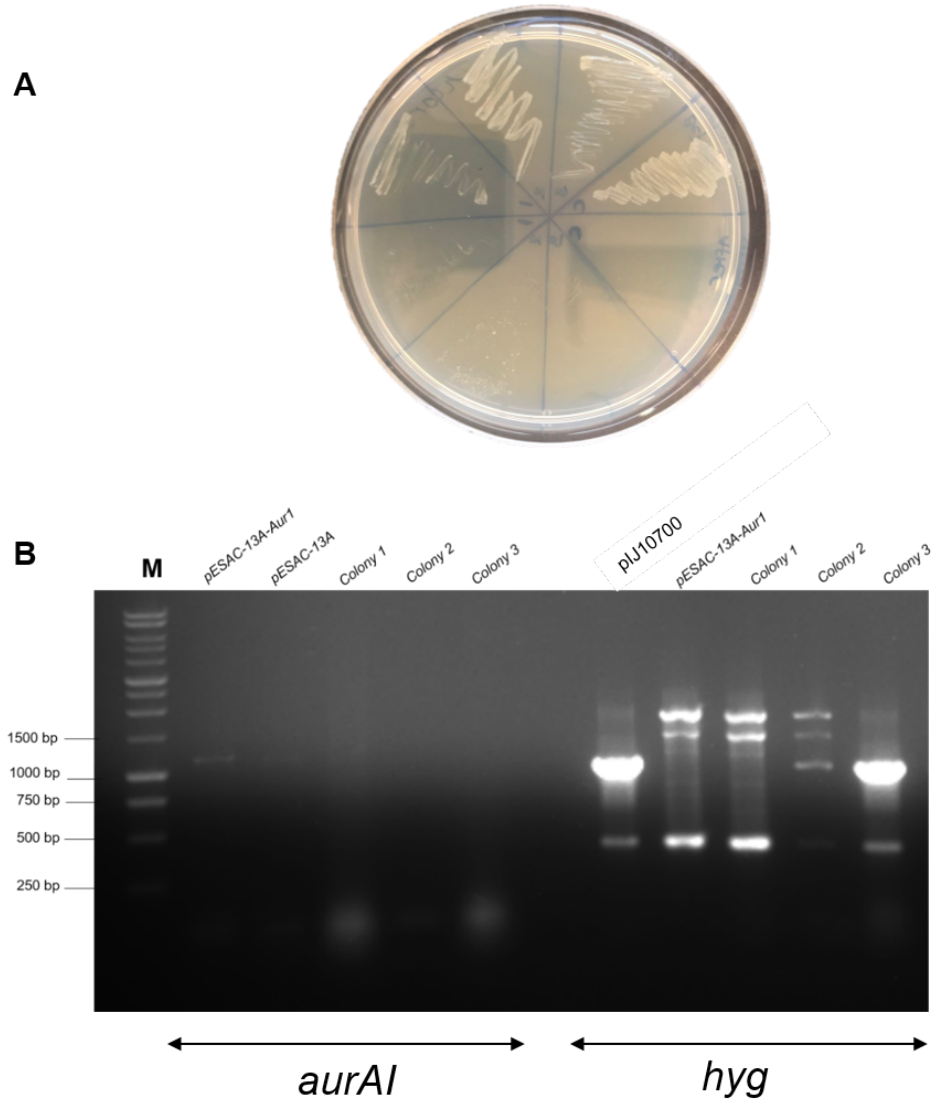


***Figure 5-19*: Construction of pESAC-13A-AurI*ΔaurAI* using redirect PCR targeting of *Streptomyces* genes.** (A) Curing of pIJ10790 (cat) from *E. coli* DH10-β + pESAC-13A-AurI-ΔaurAI. Cultures were grown in LB overnight in the absence of chloramphenicol at 37 ^{o}C to cure temperature-sensitive (30 ^{o}C) plasmid. (B) Confirmation of deletion of *aurAI* and presence of hygromycin resistance genes. PCR was used to amplify *aurAI*, with bands only present in the positive control (1120 bp), and hygromycin resistance (1200 bp), present within positive control and colony three. Hence, colony three was confirmed to contain DH10-β + pESAC-13A-AurI-ΔaurAI (*apr/hyg*).

5.2.9 Investigating the role of *aurM** in aurodox biosynthesis through heterologous expression of the gene in the kirromycin producer *S. collinus.*

In *Chapter 4,* the role of the SAM-dependent *O*-methyltransferase, AurM* was investigated. Given the absence of significant homology to kirromycin genes and the phylogenetic relatedness to methyl-transferases of the factumycin gene cluster, it was proposed that this gene was responsible for catalysing the addition of the methyl group to the pyridone moiety of the aurodox molecule. In addition, LCMS analysis of *S. goldiniensis* extracts revealed that kirromycin is likely to be the final aurodox intermediate. This led to the hypothesis that HE of *aurM** in *S. collinus* could catalyse the *in vivo* conversion of kirromycin to aurodox.

To test this hypothesis, an integrating vector was constructed to enable the thiostrepton inducible expression of *aurM** in *S. collinus.* Firstly, *aurM** was synthesised and cloned as a BamHI/NdeI fragment into a pUC57 vector (*bla*) by Eurofins, before Gibson Assembly was used to insert *aurM** in to pIJ6902 (*Figure 5-20*). The presence of *aurM** within pIJ6902 was confirmed through PCR and restriction digest. Subsequently, pIJ6902aurM* was introduced into *S. collinus* via conjugation and the presence of the integrated vector in the genome was confirmed by PCR (*Figure 5-21*).

From the initial competition analyses carried out between *S. collinus* + pIJ6902aurM* and the empty vector control strain *S. collinus* + pIJ6902, there was a clear inhibitory effect on the growth of *S. collinus* + pIJ6902 when in competition with *S. collinus* + pIJ6902aurM* (*Figure 5-21*). This is preliminary evidence of not only aurodox production in *S. collinus* but the methylation to kirromycin is sufficient to overcome the native elfamycin resistance mechanisms of *S. collinus*, *tuf2* and KirT. To further investigate this theory, both strains were subject to standard fermentation and extraction protocols and LCMS analysis was carried out on the fermentation extracts.

Initially, extract yields from *S. collinus* + pIJ6902aurM* were very low in the presence of Thiostrepton in concentrations as little at 10 ng/ml due to very limited growth. This could reflect the effect of aurodox on the producer, with thiostrepton targeting the 70S components of the ribosome (Murakami, Holt and Thompson, 1989), the double inhibition of translation may be a factor in this inhibition of growth. Therefore, cultures were allowed to reach stationary phase of growth before thiostrepton was added. In addition, gene expression under the control of the thiostrepton promoter *tipA* has been shown to be leaky (Murakami, Holt and Thompson, 1989) and therefore some expression of *aurM** in the absence of thiostrepton was expected.

Upon molecular analysis, signals corresponding to an m/z ratio of 793/811 were detected in the extracts from *S. collinus* +pIJ6902AurM* that were not present in the extracts from the empty vector controls, indicating the presence of aurodox in the extract from *S. collinus* +pIJ6902AurM* (*Figure 5-22*). The signals corresponding to aurodox are significantly weaker than those corresponding to kirromycin however, it is thought that further optimisation of the thiostrepton concentration in the media may alter this molecular. This data therefore biochemically confirms the role of AurM* as the SAM-dependent *O*-methyltransferase responsible for catalysing the methylation of kirromycin to become aurodox, whilst also providing further evidence that kirromycin is the final aurodox intermediate. Given the ability to aurodox to circumvent the native elfamycin resistance mechanisms of *S. collinus*, the acquisition of AurM* has significantly altered the activity the molecule. Therefore, these results provide an example of the profound effects that simple chemical modifications to specialised metabolites can induce.

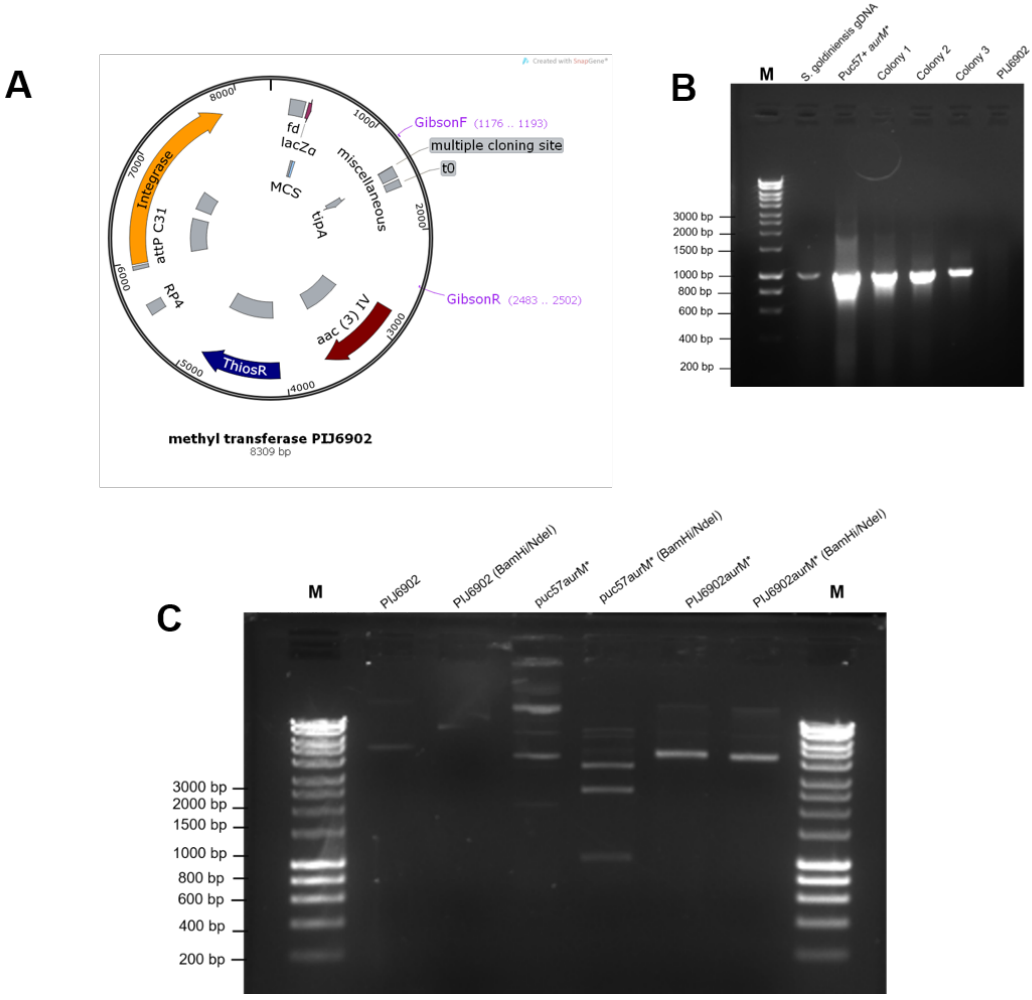


***Figure 5-20*: Construction of pIJ6902aurM* using Gibson Assembly.** (A) Plasmid map of pIJ6902aurM* generated by SnapGene TM. (B) Colony PCR to confirm the presence of aurM* (1002 bp) in PIJ6902_aurM*. *S. goldiniensis* gDNA was used as a positive control with pIJ6902 used as negative control. Presence of *aurM** was confirmed in all transformant colonies tested. (C) Restriction digest of pIJ6902/puc57AurM*/pIJ6902AurM* with BamHI and NdeI. No digestion by BamHI and NdeI in PIJ6902_aurM* demonstrates destruction of restriction sights by Gibson Assembly.

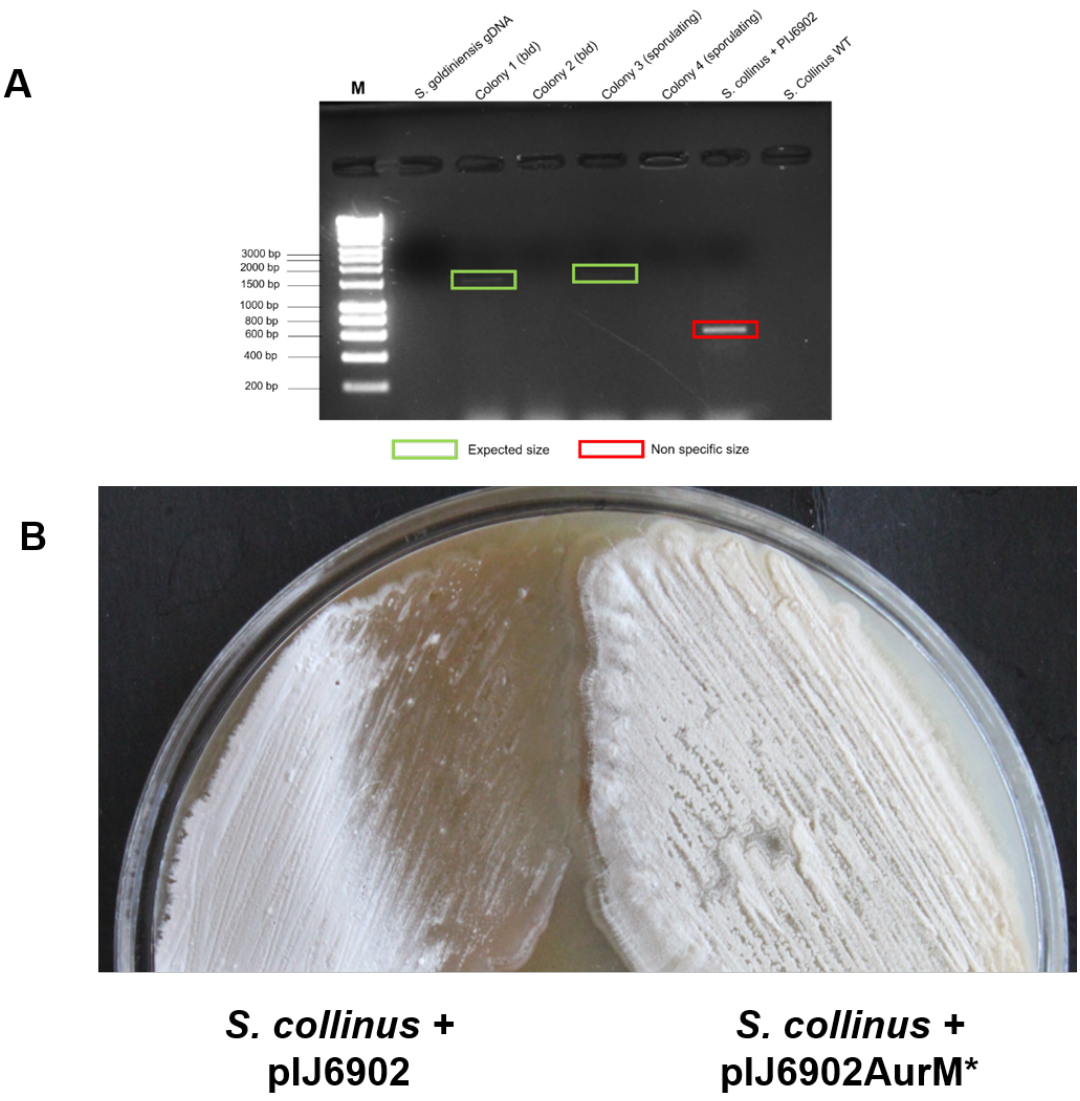


***Figure 5-21*: Confirmation of pIJ6902AurM* in *S. collinus*.** (A) PCR confirmation of pIJ6902AurM* integration into *S. collinus.* Primer pairs were designed to bind to pIJ6902 backbone resulting in a 650 bp product from the empty vector exconjugants and a 1652 bp product from pIJ6902AurM*. Wild type *S. goldiniensis* and *S. collinus* were used as negative controls. (B) Image of *S. collinus* + pIJ6902 and *S. collinus* + pIJ6902AurM* grown on MS agar + apramycin + Thiostrepton.

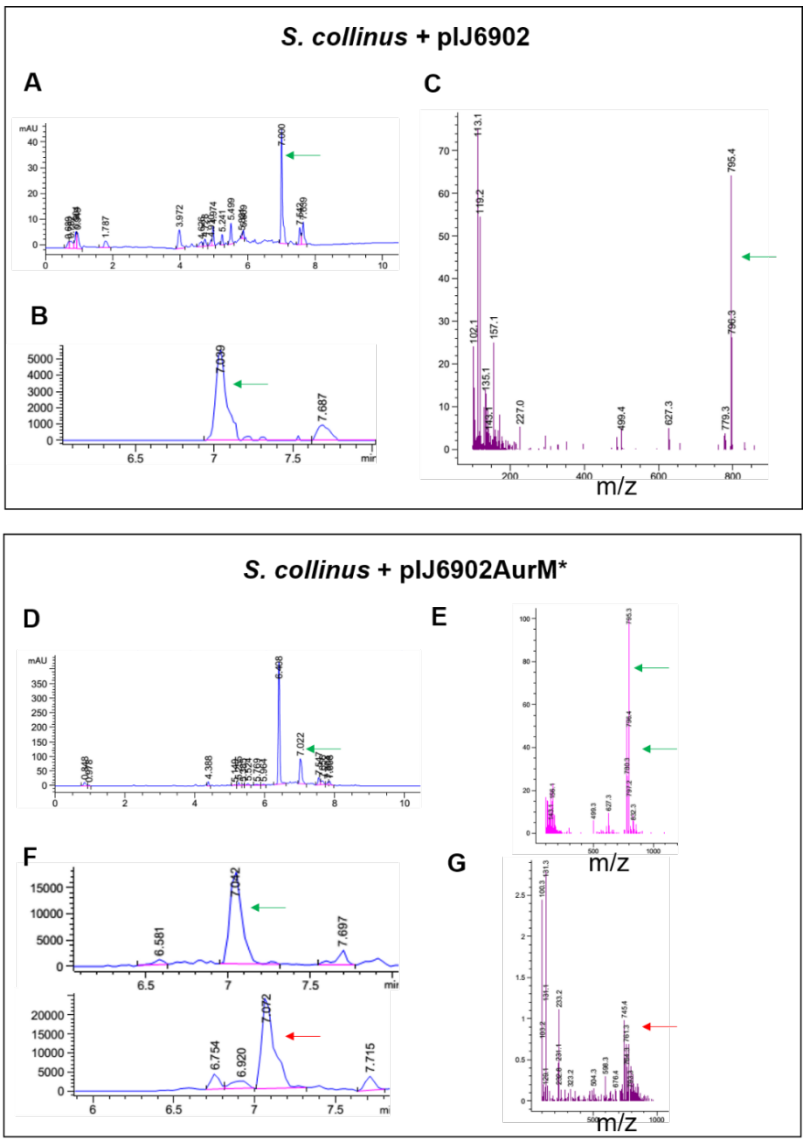


Figure 5-22:: LCMS analysis of *S. collinus* + pIJ6902AurM* versus empty vector control. Extracts were generated through standard fermentation and extraction protocol (*Chapter 3*) (A) Total Ion Chromatogram of extracts from *S. collinus* + pIJ6902, kirromycin peak (retention time 7.0 minutes) is indicated with a green arrow. (B) HPLC trace showing signal intensity from kirromycin associated peak. (C) Mass Spectrometry analysis of 7.0 minute retention time peak, peak with m/z ratio of 795 (kirromycin) is indicated with a green arrow. (D) Total Ion Chromatogram of extract from *S. collinus* + pIJ6902AurM with kirromycin peak indicated with a green arrow. (E) Mass Spectrometry analysis of 7.0 minute retention time peak, peak with m/z ratio of 795 (kirromycin) is indicated with a green arrow (F) HPLC trace showing signal intensity from kirromycin associated peak (green arrow) and aurodox associated peak (red arrow). (G) Mass Spectrometry analysis of 7.1 minute retention time peak (aurodox) signal with m/z ratio of 793 corresponding to aurodox is indicated with a red arrow.

5.3 Summary

In this chapter, the bioinformatic prediction of the putative aurodox BGC from previous chapters was validated using a range of molecular and biochemical techniques. Principally, the role of the putative aurodox BGC was confirmed through the heterologous expression of a pESAC-13A derived PAC (pESAC-13A-AurI), which was confirmed by PCR to encode the entirety of the putative BGC. To select a suitable host, multiple *Streptomyces* strains with a history of reliable use as heterologous hosts were screened for aurodox resistance. Furthermore, phylogenetic analysis of the elongation factor thermo-unstable genes (EF-Tu) was carried out to identify strains which demonstrated significant relatedness to the elfamycin-resistant *tuf2* genes of *S. goldiniensis* and *S. collinus.* Therefore, *S. collinus, S. coelicolor* M1152, *S. venezuelae* and *S. albus* were chosen for HE of the aurodox BGC, which was initially successful in *S. collinus* and *S. coelicolor* M1152. These results partially demonstrate the role of the putative aurodox gene cluster in aurodox biosynthesis however, it is possible for additional experiments to be carried out to further confirm this. Primarily, these should include the sequencing of pESAC-13A-AurI to determine the genes which flank the aurodox gene cluster, and following this, the disruption or removal of these genes from the PAC to demonstrate the catalytic role of the aurodox cluster in aurodox biosynthesis and conclusively determine the cluster boundaries. Further, irrefutable evidence for this can be provided through the analysis of extracts from *S. coelicolor* M1152 + pESAC-13A-AurIΔaurAI to test for the hypothesised absence of aurodox initiated by the disruption to the primary PKS unit of the cluster.

Supplementary EF-Tu (*tuf2)* genes in elfamycin producers are thought to contribute to immunity during the biosynthesis (Lutz Vogeley *et al.*, 2001; Olsthoorn-Tieleman *et al.*, 2007). However, with elfamycin BGCs such as the aurodox and kirromycin gene clusters encoding Mfs-type exporters (AurT/KirT; Weber *et al.*, 2008) the essentiality

of these genes for elfamycin immunity in producers has not been confirmed. Given the inability of *S. albus* and *S. venezuelae* to biosynthesise aurodox without *tuf2*, followed by aurodox production when the strains expressed these genes, (confirmed in *S. venezuelae,* suspected in *S. albus*), the work carried out in this chapter has enhanced the theory that both an exporter and a *tuf2* gene must be present to enable aurodox production. This contradicts the findings of Thaker *et al.*, (2012) who suggest that some elfamycin producers do not encode additional or resistant elongation factors. However, this analysis was based on the presence of an elfamycin-type BGC, without determining the expression status of the cluster. In addition, the quality of the genomes of the elfamycin producers used in the study were poor which may have resulted in additional copies of EF-Tu remaining undetected.

A central aim of this chapter was to confirm the catalytic role of the predicted SAM-dependent *O*-methyltransferase AurM* in the aurodox biosynthetic pathway. Expression of this gene in the kirromycin producer, *S. collinus* enabled the production methylated kirromycin- presumed to be aurodox- and hence simultaneously provided evidence that kirromycin is the final aurodox intermediate and that AurM* acts as a SAM-dependent *O*-methyl transferase, catalysing the addition of the methyl group to the pyridone moiety of kirromycin. Although retention time and m/z signals associated with aurodox could be detected in the LCMS analysis from the kirromycin producer expressing AurM*, the position of the additional methyl group cannot be determined by LCMS data alone. Therefore, to complete this data and conclusively confirm the hypothesis, LCMS/MS, NMR or C13 experiments should be carried out to confirm that the pyridone moiety is methylated (Musiol *et al.*, 2013).

The work described in this chapter has positive implications for the future study of aurodox. The increased aurodox yields provided by *S. coelicolor* M1152 +pESAC-13A-AurI and *S. venezuelae* + pESAC-13A-AurI + pMS82_tuf2 may enable increased

efficiency in production and purification of the compound for use in future studies. With aurodox recently being identified as an inhibitor of the EPEC T3SS (Kyota Kimura *et al.*, 2011), these engineered production strains can facilitate aurodox production for in depth studies including future work which aims to identify the mechanism by which aurodox inhibits T3S in enteric pathogens.

Chapter 6 : Characterisation of the mechanism of action of aurodox, A Type III Secretion System Inhibitor from *Streptomyces goldiniensis.*

6.1 Introduction

Enterohaemorrhagic *Escherichia coli* (EHEC) is a zoonotic pathogen responsible for foodborne outbreaks of diarrhoea which can escalate to haemorrhagic colitis and Haemolytic Uraemic Syndrome (HUS) (Kaper, Nataro and Mobley, 2004). Infections associated with the related pathotype Enteropathogenic *Escherichia coli* (EPEC) are typically less severe; however this pathotype remains the leading cause of diarrhoeal deaths in the developing world (Liu, Johnson and Cousens, 2012). As with many other Gram-negative pathogens, EHEC and EPEC utilise the Type Three Secretion System (T3SS) to facilitate colonisation within the host (Hueck, 1998). The T3SS is a needle-like apparatus which facilitates the translocation of effector proteins to the epithelial cells of the gut. Through the injection of these proteins which include the Translocated Intimin Receptor (Tir) and the Mitochondria Associated Protein (Map), EHEC and EPEC are able to polymerise actin filaments and form intimate junctions with epithelial cells, where several other virulence factors are expressed to destabilize cellular processes (Goosney, Gruenheid and Finlay, 2000; Pacheco and Sperandio, 2012). In the case of EHEC infection, symptoms can extend beyond the intestine as strains typically carry phage-derived Shiga toxins (Stx) which target organs such as the kidney and brain (Tarr, Gordon and Chandler, 2005). This often leads to the life-threatening condition HUS.

The T3SS of EHEC and EPEC is encoded by the highly-conserved Locus of Enterocyte Effacement (LEE) Island (Schmidt, 2010). This pathogenicity island encodes 42 genes on five polycistronic operons and is regulated by the master regulator, Ler. In turn, *ler* is regulated by specific regulators such as GrlA and GrlR, in addition to global regulators which mediate LEE expression in response to environmental stimuli (Iyoda *et al.*, 2006).

The reliance of EHEC and EPEC on the T3SS to initiate infection has identified it as a target for novel therapies to fight infection. Typically, these are part of a wider anti-virulence approach in which the aim is to prevent infection by the inhibition of a single virulence factor without inducing a reduction in growth (Allen *et al.*, 2014). Currently, treatment of EHEC infections with traditional antibiotics is not recommended due to stimulation of the bacterial SOS response (Pacheco and Sperandio, 2012). In response to DNA damage caused by antibiotics, the SOS response protein, RecA is overexpressed which results in activation of the Stx-encoding phage (Huerta-Uribe *et al.*, 2016). Hence, Stx production is upregulated and symptom severity increases. Additionally, the disruption to the native gut microbiome by broad-spectrum antibiotics can have negative consequences for the patient (Modi, Collins and Relman, 2014). As EHEC and EPEC infections are typically cleared naturally, anti-virulence approaches to treatment of EHEC and EPEC represent an exciting new strategy for the treatment of these infections. In addition, compounds that do not affect bacterial growth or survival reduce the evolutionary selective pressure on strains resistant to the treatment (Allen *et al.*, 2014), enhancing the long-term viability of the therapy.

Small compound inhibitors of the EPEC and EHEC T3SS have previously been identified (Veenendaal, Sundin and Blocker, 2009; Gauthier *et al.*, 2005). Notably, the salicylidene acylhydrazide (SA) family have been shown to inhibit T3S in a range of

enteric pathogens including EPEC, EHEC and *Salmonella enterica* (Zambelloni, Marquez and Roe, 2015). However, these compounds were found to bind to several bacterial protein targets and their mode of action has been shown to result from synergistic effects arising from a perturbation of the function of several conserved metabolic proteins (Wang *et al.*, 2011). Therefore, the conclusion was that, although effective, the SAs were rather promiscuous (Wang *et al.*, 2011).

Unsurprisingly, specialised metabolites of bacteria and fungi have been investigated for their anti-virulence properties. In nature, there are examples of natural product virulence-inhibitors being deployed in synchrony with expression of antibacterial compounds. For example, *Streptomyces clavuligerus* is known to co-express the beta-lactam antibiotic, cephamycin C with the beta-lactamase inhibitor clavulanic acid, to enhance the activity of cephamycin C by blocking the resistance mechanisms (Alexander and Jensen, 1998). Compounds such as these have evolved to confer competitive advantages in ecological warfare and therefore, the properties of these metabolites can vary. Hence, discovery of lead compounds has been successful from targeted natural product library screening for anti-virulence compounds, specifically for those which inhibit the T3SS (Zambelloni, Marquez and Roe, 2015).

aurodox, was identified from a large-scale compound screen by Kimura *et al* (2011) to inhibit the translocation of T3SS encoded effectors in EPEC. Low concentrations (1.5 μM) were shown to inhibit secretion and abolition of detectable effector proteins was observed at 6 μM. In this study, the findings of Kimura and his co-workers were replicated, and this works shows for the first time that the compound was also effective against the human pathogen EHEC. Moreover, Kimura and his co-workers tested the effect of the compound *in vivo* through the use of a *Citrobacter rodentium* murine infection model. Here, it was shown that mice treated with the compound survived lethal infections with limited effects on the intestinal tract. Although the effects of the

compound on T3S in EPEC were characterised by Kimura *et al*, the wider effects and the mechanism of action of the compound were not elucidated (Kyota Kimura *et al.*, 2011). Therefore, there is a need to gain a better understanding of the mechanism of action of aurodox for the development of the compound in to an anti-virulence drug.

As previously mentioned, aurodox was originally identified as an antibiotic compound with antibacterial effects upon Gram-positive pathogens such as *Streptococcus pyogenes* and *Staphylococcus aureus* (Berger *et al.*, 1972). Aurodox has since been well-characterised in terms of its bactericidal mechanism, with a mild effect upon *E. coli* growth reported using concentrations greater than 1 mg/ml, 200 times higher than used in T3S assays (5 µg/ml; Maehr *et al.*, 1978). The limited spectrum of antibiotic activity of aurodox has resulted in it being rejected as an antibiotic to treat infections in humans, yet the discovery of the novel T3SS inhibitory properties through screening for the inhibition of EPEC-mediated haemolysis suggests that aurodox may be a candidate drug for repurposing.

In this chapter, the effect of aurodox on T3S in EHEC O157:H7 (TUV93-0), EPEC O127:H6 (E2348/69) and *Citrobacter rodentium* (ICC168) has been characterised and it was demonstrated that the effects of aurodox are independent of growth. Furthermore, transcriptomic analysis has been used to show that aurodox inhibits the T3SS at the level of transcription by repression of the LEE master regulator, *ler*. Hence, the data suggest a model in which aurodox acts upstream of Ler and not directly on the T3SS itself. Finally, it is shown that aurodox does not induce expression of *recA*, which is essential for the production of Stx or induce shiga toxin production in *Citrobacter rodentium* carrying the Shiga toxin phage. It is proposed that these properties make aurodox a candidate anti-virulence therapy for the treatment of EPEC and EHEC infections.

6.2 Results

6.2.1 Aurodox inhibits the translocation of T3SS-associated effector proteins without affecting growth.

In a previous study, Kimura *et al* (2011) demonstrated that aurodox reduced the translocation of EPEC effector proteins in a concentration-dependent manner. Hence, the affect of the compound on other LEE-encoding pathogens, primarily EHEC (TUV93-0) and *C. rodentium* (ICC168) were explored with an aim to determine whether the mechanism of T3SS inhibition was independent of an inhibition in growth.

Each strain was cultured in media appropriate for the expression of the T3SS (Roe *et al.*, 2003). Aurodox was added to the cultures at the point of inoculation at increasing concentrations ranging from 1.5 µg/ml to 5 µg/ml (1.2-6 µM) and bacteria were grown through 4 generations to an OD_{600} of 0.7-0.9. Supernatant proteins were precipitated and whole cell lysates prepared as a comparator. The fractions were separated with SDS-PAGE and the gels were stained with Coomassie Brilliant Blue.

For the supernatant fractions, a concentration dependent reduction in T3SS-associated effector proteins was observed (*Figure 6-1A*). The dominant bands from the gel were excised to permit in-gel trypsin digestion and analysis by tandem mass spectrometry. Proteomic Mass Spectrometry was used in this case to eliminate the effects of variations in antibody binding which can be affected by the differences in serotype. This confirmed that, for all three pathogens, the two most dominant bands were comprised of three well-known effector proteins: Tir, EspB and EspD (*Table 6A-C*). In both EPEC and EHEC, EspB/EspD form a pore in the host cell membrane, whilst Tir, the trans-intimin receptor, is translocated to the surface of the target cell to act as a receptor for intimin, which facilitates the attachment and effacement of host cells. In contrast, there was no change in the profile of the cellular proteins (*Figure 6-*

1B), indicating that the mechanism of T3SS inhibition was not result of generic block in protein synthesis. Moreover, at the highest concentration used (5 µg/ml) aurodox does not inhibit growth of EPEC, EHEC or *C. rodentium* (*Figure 6-1B*).

Importantly, no statistically significant decrease in cell viability was observed (*Figure 6-1B*). Therefore, at the concentration of aurodox used, the mechanism of T3S inhibition is independent of any defect in growth or viability.

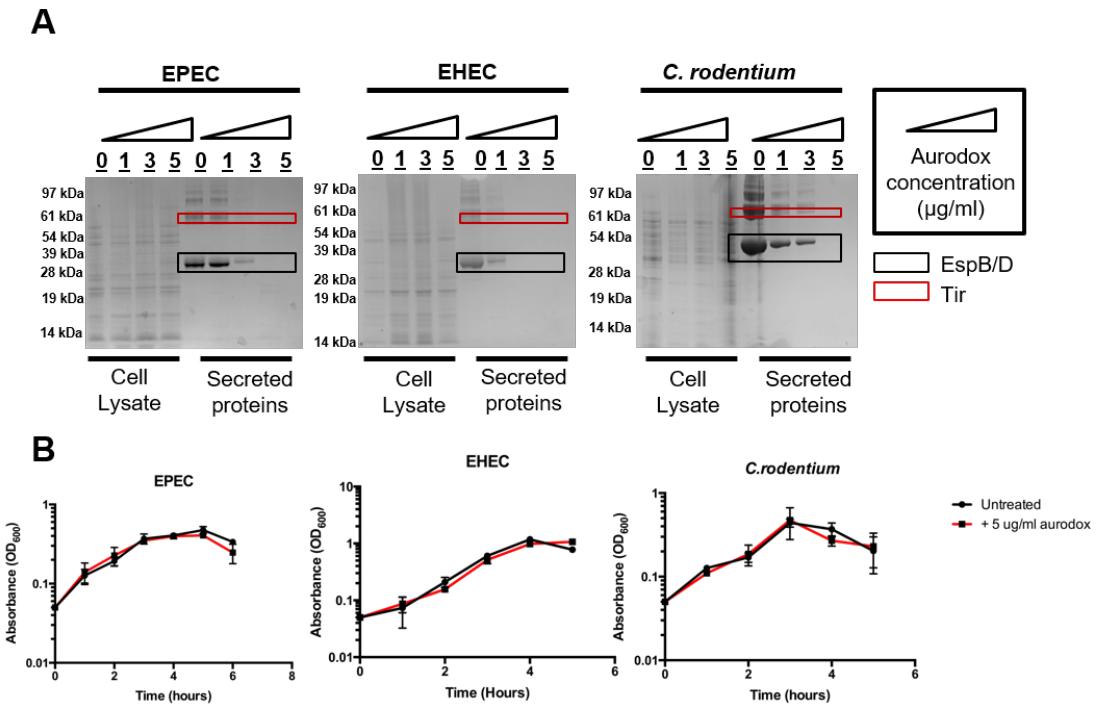


Figure 6-1: **Aurodox inhibits secretion of T3SS-associated effector proteins in EPEC, EHEC and *Citrobacter rodentium* without effecting bacterial growth.** (A) Secreted protein fractions were prepared by culturing strains in T3S-inducing media and proteins were precipitated using Trichloroacetic Acid. Whole cell fractions were prepared by lysis using BugBuster ™ solution. Samples were resolved using SDS gel-electrophoresis followed by Coomassie staining. Marked protein bands were excised and trypsin-digested for Mass Spectrometry identification of proteins. (B) Aurodox (5 µg/ml) does not inhibit the growth of EPEC, EHEC or *C. rodentium*. Strains were grown in Lysogeny Broth and growth rates were determined spectrophotometrically (OD_{600}), n=3. Error bars plotted are Standard Deviation from the mean. Changes in growth were not statistically significant.

Table 6-A: MASCOT result for EPEC secreted protein bands

Protein	Mass	MASCOT score	Matches
EspB	31564	648	11/11
EspD	39033	874	14/14
Tir	57954	2336	50/50
EspP	135881	2655	47/47

Table 6-B: **MASCOT result for EPEC secreted protein bands**

Protein	Mass	MASCOT score	Matches
EspB	31564	-	-
EspD	39033	-	-
Tir	57954	377	6/6

Table 6-C: **MASCOT result for EPEC secreted protein bands**

Protein	Mass	MASCOT score	Matches
EspB	31564	-	-
EspD	39033	-	-
Tir	57954	377	6/6

6.2.2 Aurodox inhibits the ability of EHEC to attach and efface epithelial cells

Previous studies using a murine infection model tested the inhibitory effects of aurodox on *C. rodentium* pathogenicity, demonstrating a marked improvement in the survival of infected mice and a reduction in colon damage when treated with aurodox (Kyota Kimura *et al.*, 2011) however, the effects were not demonstrated at the cellular level.

To investigate the effects of aurodox on the attachment of EHEC to host cells, which is known to be driven via the T3SS an *in vitro* infection assay was used. The effect of aurodox (5 µg/ml) on uninfected HeLa cells was tested to confirm there was no overt cytotoxicity of the compound (*Figure 6-2*). HeLa cells were infected with 10^7 EHEC bacteria constitutively expressing GFP. Host cell actin cytoskeleon was stained with Phalloidin-Alexa555. In addition, cell lysis was carried out to quantify the attached bacteria as a percentage of the initial innoculum. Epifluorescence microscopy images were used to quantify the number of infected cells and to investigate any effects on cell morphology (*Figure 6-3A)*.

At a mutliplicity of infection (MOI) of 150, untreated EHEC produced consistent infections with all HeLa cells showing > 50 adhered bacteria per cell. The HeLa cells displayed morphological changes associated with bacterial infection including actin condensation caused by lesion formation and cell rounding (*Figure 6-3A*). Addition of aurodox (5 µg/ml) substantially reduced both phenotypes, with only 36 % of cells becoming infected and a marked reduction in attaching and effacing lesions. Indeed, the infections by EHEC were typically restricted to a small number of bacteria per cell (< 5). Quantification of bacterial adherence efficiency by colony forming unit counts showed that EHEC infection was reduced by more than 7000-fold when treated with aurodox (*Figure 6-3B*), demonstrating the potent anti-virulence capacity of aurodox.

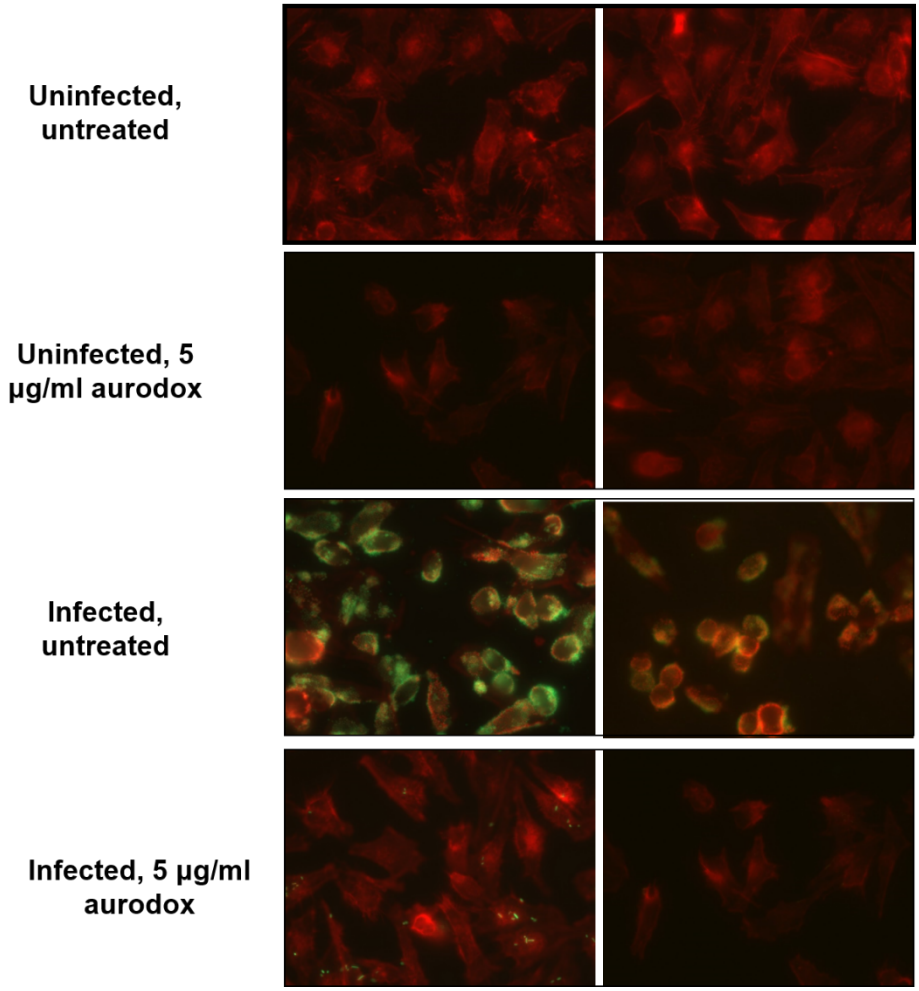


Figure 6-2: **Effect of aurodox on EHEC infection of epithelial cells and A/E lesion formation**. (A) Representative epifluorescence microscopy images from EHEC cell infection assay. Bacteria were transformed with 10^7 EHEC (TUV-930) transformed with p*rpsM-gfp* encoding GFP to facilitate quantification of infected cells and imaging. HeLa cells were actin-stained with Alexa-Fluor-Phaloidin (488).

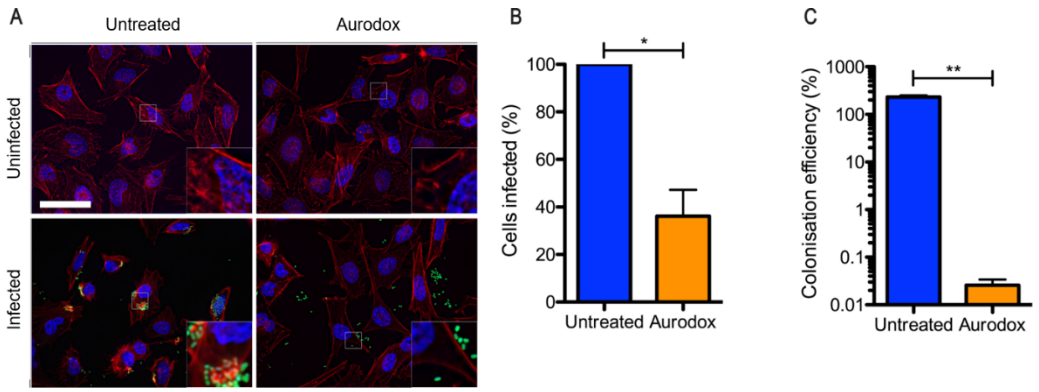


Figure 6-3: **Effect of aurodox on EHEC infection of epithelial cells and A/E lesion formation.** (A) Representative epifluorescence microscopy images from EHEC cell infection assay. Cells were infected with 10^7 EHEC cells transformed with p*RpsM-gfp* to facilitate quantification and imaging. HeLa cells were actin-stained with Alexa-Fluor-Phaloidin. Scale bar represents 50 µm. Insets contain a ×4 magnification of the indicated area. (B) Quantification of adherence efficiency and percentage of infected HeLa cells. Percentage of infected cells was determined microscopically by taking the average number of infected cells from ten images from three coverslips per experiment, p=0.014. Adherence percentage of aurodox-treated cells is the number of cells retrieved in the cell lysate expressed as a percentage of the number of cells retrieved in the CFU values were determined in triplicate from three assays p=0.0016. Asterisks indicate statistical significance.

6.2.3 Aurodox inhibits expression of the EHEC Locus of Enterocyte Effacement (LEE)

To gain insights into the possible molecular mechanism underlying the inhibition of the T3SS whole transcriptome analysis was used. EHEC was grown in MEM-HEPES ± 5 µg/ml aurodox and RNA was extracted from triplicate cultures. Transcripts were mapped to the EHEC EDL933 reference genome and the mean fold change was calculated. In total, 84 chromosomal genes and four pO157 genes were significantly downregulated, using the standard fold change of -1.5, consistent with the cutoff used to define the *ler* regulon (Bingle *et al.*, 2014), and an EDGE test *P* value of < 0.05. Consistent with the secretion data, analysis of LEE transcription revealed downregulation of the island after treatment with aurodox (*Figure 6-4)*, with 25 of the 42 LEE genes significantly downregulated, including the master regulators *ler* and *grlA* (Iyoda *et al.*, 2006). Additionally, the expression of a further 16 genes in the LEE were downregulated to some extent but did not meet the statistical requirements for significance. Therefore, 41 of 42 genes in the LEE showed some level of downregulation in the presence of aurodox. The only gene in the LEE pathogenicity island which was upregulated was *escS*, a gene encoding for a T3SS structural membrane protein. However, the upregulation of this gene can be counteracted by the downregulation of the expression of the rest of the LEE, as its function is dependent on trimeric proteins EscR/T, whose coding genes are both downregulated. It therefore appears that the marginal increase is an artefact in the data. A summary of the upregulated and downregulated EHEC genes in response to aurodox can be found in *Table 6D* and *6E.*

Table 6-D: **Summary of EHEC genes which are downregulated in response to aurodox treatment.**

Gene	Fold Change	P Value
espH	-1.51	0.04
metE	-1.51	0.01
ydgA	-1.52	0.00827
Z2386	-1.53	0.0021
escJ	-1.53	0.03
rpoE	-1.54	0.000591
Z2387	-1.55	0.00607
adhE	-1.56	0.000144
yeaD	-1.6	0.000375
sepZ	-1.6	0.0019
espF	-1.61	0.00231
yhcK	-1.61	0.05
spfA	-1.61	0.00155
xasA	-1.63	0.04
sepL	-1.63	0.00117
rrsE	-1.63	0.01
ptsG	-1.66	0.000000407
escI	-1.66	0.01
mokP	-1.67	0.01
abgT	-1.68	0.04
escN	-1.69	0.0000174
cesD	-1.7	0.000775
espJ	-1.71	0.000591
rrsD	-1.71	0.006
sepQ	-1.72	0.0000902
escF	-1.72	0.00782
rrsB	-1.74	0.00216
map	-1.74	0.000582
tktB	-1.76	0.00121
lifA	-1.77	0.03
tir	-1.78	0.00000692
osmC	-1.79	0.0000687
escC	-1.8	0.0000001
espX	-1.81	0.00149
pyrI	-1.82	0.00268

cesF	-1.82	0.00867
ysgA	-1.83	0.000126
poxB	-1.84	0.04
espL2	-1.87	0.00881
nleB	-1.89	0.02
orf29	-1.89	0.000106
rrfF	-1.93	0.00827
katE	-1.93	0.04
espW	-1.99	4.08E-09
rrfG	-2	0.00206
cesT	-2.02	4.63E-08
Z2063	-2.02	0.0000147
escE	-2.03	0.000533
Z5138	-2.04	0.0000436
yjcO	-2.13	0.00135
cesD2	-2.13	1.01E-11
manZ	-2.16	8.56E-13
espD	-2.18	5.84E-08
nleA	-2.2	0.0000411
manY	-2.21	1.69E-14
manX	-2.29	4.59E-19
espA	-2.32	4.26E-10
eae	-2.34	0
rorf1	-2.34	0.02
espB	-2.36	5.06E-10
fruA	-2.44	0.00000123
grlA	-2.45	0.00000008
rcsA	-2.47	1.67E-08
rrlG	-2.53	0.00527
yneB	-2.54	0.0047
ssrS	-2.57	0.000106
cpsB	-2.74	0.00827
ffs	-2.8	0.03
wzc	-2.97	0.03
fruB	-3.07	0.0000147
yghX	-3.51	0.00607
fruK	-3.59	0.000538
wcaH	-3.8	0.04
wcaG	-3.83	0.000000108
yjbE	-3.93	0.04
wcaD	-3.94	0.000563
wcaE	-3.98	0.00704

gmd	-5.04	5.22E-16
yebV	-5.89	0.02
wcaI	-6.07	1.64E-08
wcaC	-10.12	0.01
cpsG	-12.12	0.00000259
yjbF	-43.7	0.00554

Table 6-E: : **Summary of EHEC genes which are upregulated in response to aurodox treatment.**

Gene	Fold Change	P value
yheN	1.51	0.00000166
Z2005	1.52	0.00000151
yafH	1.53	0.00000191
pyrF	1.53	0.00000211
yqiB	1.53	0.00000142
trmA	1.57	0.00000173
yegQ	1.58	0.00000195
Z4370	1.58	0.00000284
sapA	1.59	0.00000169
hemA	1.59	0.00000443
yfhJ	1.6	0.00000159
folX	1.6	0.00000145
gabP	1.61	0.0000134
yfcI	1.61	0.00000143
mdaA	1.61	0.0000014
yliG	1.61	0.00000227
stpA	1.62	0.00000146
yhiW	1.62	0.00000157
ycaD	1.62	0.00000115
bcr	1.63	0.00000205
terC	1.63	0.00000125
ydeA	1.65	0.00000243
yafS	1.67	0.00000161
yccJ	1.67	0.00000161
abc	1.67	0.00000269
hypF	1.68	0.00000127
pheA	1.69	0.00000154
ygaF	1.69	0.0000126
yfbS	1.7	0.00000188
yfhE	1.71	0.0000069
leuC	1.71	0.00000764
pheP	1.72	0.0000026
yedV	1.72	0.00000183

leuB	1.73	0.00000588
ycgN	1.73	0.00000112
ydcP	1.75	0.00000323
fadB	1.75	0.00000154
nrdD	1.76	0.000000974
ygaT	1.77	0.0000152
sodB	1.78	0.00000114
fimB	1.78	0.00000164
trpC	1.78	0.00000712
bigA	1.8	0.00000249
prpE	1.81	0.00000134
yieP	1.81	0.00000165
leuD	1.81	0.00000392
Z2806	1.82	0.0000014
yfcD	1.83	0.00000279
yciA	1.84	0.00000161
narQ	1.85	0.00000125
acrD	1.85	0.00000109
ubiX	1.85	0.00000217
leuA	1.89	0.00000923
Z4832	1.91	0.00000209
Z2491	1.96	0.0000022
espR	1.99	0.000000894
yjhS	2	0.00000103
ycdZ	2.02	0.00000163
ycfS	2.03	0.00000321
Z4833	2.11	0.00000233
yadB	2.12	0.00000143
yhfC	2.18	0.000005
trpB	2.25	0.0000161
trpA	2.27	0.0000104
phoE	2.37	0.000000808
Z0701	2.39	0.000000646
ykgF	2.4	0.000000647
frdB	2.43	0.000000755
metF	2.59	0.00000174
yjcZ	2.63	0.000000653
torR	2.65	0.000000871
ebgR	2.9	0.00000124
yjiY	2.96	0.00000328
aldH	3	0.0000013
yafO	3	0.00000055

yliF	3	0.00000055
ymdF	3.14	0.000000991
hcaR	3.2	0.000000882
Z3235	3.22	0.000000609
Z1644	3.27	0.00000105
yfhH	3.29	0.000000774
bssR	3.39	0.00000172
mtr	3.66	0.0000141
CII	4.01	0.000000447
narJ	4.19	0.000000679
ybiH	4.93	0.00000108
trpE_1	5.23	0.00000223
dicB	5.62	0.000000395
prpR	5.62	0.000000395
Z2986	5.63	0.000000397
trpD	6.05	0.00000744
hycA	6.27	0.000000452
yaeJ	6.28	0.00000111
Z4398	6.28	0.000000452
yjgZ	8.26	0.00000062
napC	8.97	0.000000682
ydfZ	9.64	0.000000739
torD	10.09	0.0000025
torA	13.4	0.0000104
sorA	16.27	0.000000344
torC	18.24	0.00000365
Z6041	18.81	0.000000401
cmtA	18.84	0.000000402
ecpD	33.8	0.000000739

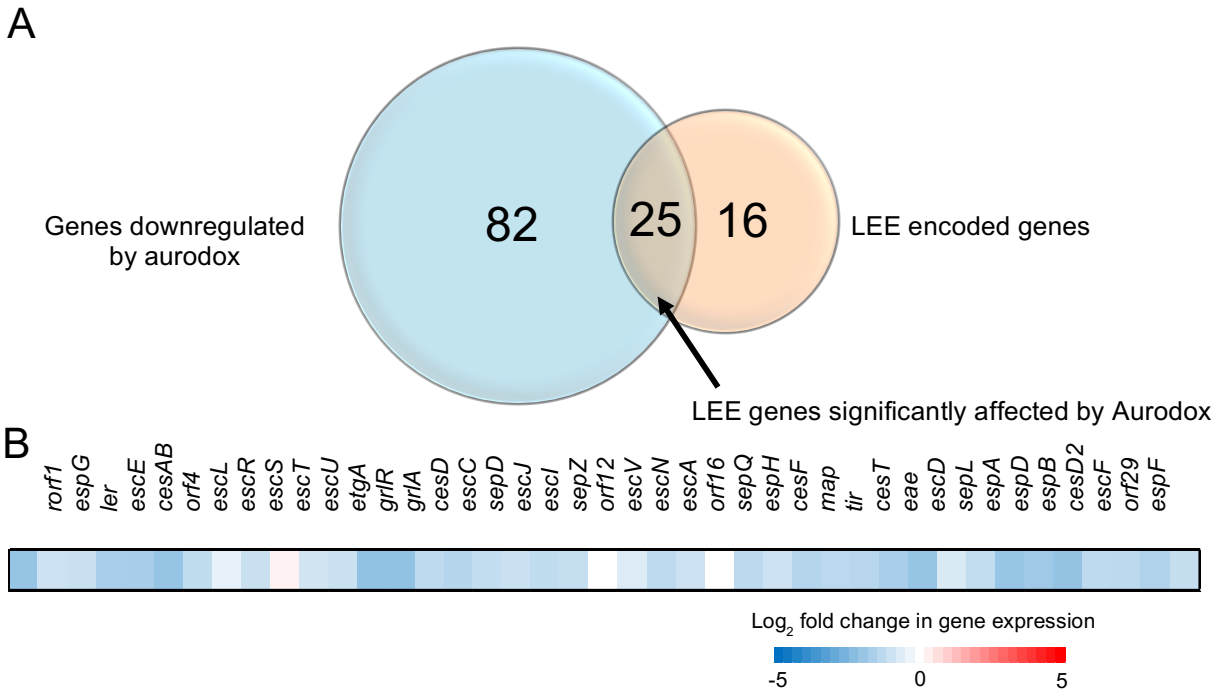


***Figure 6-4**:* **Representation of transcriptional changed induced by Aurodox in EHEC.** (A) Venn diagram representing overlap between genes significantly downregulated by aurodox (>1.5 fold, p <0.05) and LEE genes. (B) Heatmap representation of Log_2 fold change in LEE gene expression after treatment with Aurodox.

6.2.4 Aurodox inhibits the expression of other, Ler-regulated virulence genes outside of the Locus of Enterocyte Effacement

The data also revealed that many genes in the colanic acid biosynthesis operon were downregulated in the presence of aurodox. The initial results show that 20 of the 21 genes in this operon are downregulated in response to aurodox with 15 significant changes (P = 0.05; Table S6). Notably, expression of *rcsA,* a key regulator of colanic acid biosynthesis that is known to form part of the Ler dependent regulon (Bingle *et al.*, 2014) was downregulated 2.5-fold (*Figure 6-4*). These data suggest that suppression of the colanic acid operon is likely due to expression changes in the key regulators *rcsA* and *ler*. These changes could indicate further potential of aurodox for the treatment of EHEC infections, as colanic acid production is required *in vivo* during infection (Mao, Doyle and Chen, 2006). Furthermore, the extracellular polysaccharides encoded for by this operon play a role in the formation of biofilms in the environment. However, as the EHEC strain used in this study does not readily form biofilms in the laboratory environment, the effect of aurodox on biofilm phenotype could not be analysed.

In addition, 103 genes were upregulated in the presence of aurodox which include various metabolic genes (*prpR, trpA, trpB),* and genes encoding sensory proteins (e.g. *narQ)*. Importantly, there was no upregulation of SOS response- associated genes or virulence genes. The most upregulated gene was *ecpD,* a putative fimbrial chaperone protein, with a $\log_2$ fold change of 5. A complete list of these genes can be found in Appendix A.

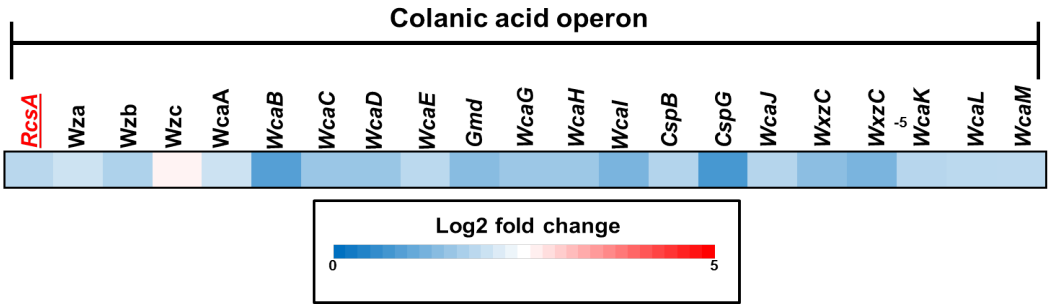


Figure 6-5: **Representation of transcriptional changes induced by aurodox in colanic acid biosynthesis operon**. Heatmap representation of Log_2 fold change in colonic acid operon gene expression after treatment with Aurodox.

6.2.4 Aurodox inhibits T3SS expression through downregulation of the master regulator *ler.*

To confirm the virulence-related aurodox targets identified by RNA-seq, GFP gene reporter assays (transcriptional fusions) were used to validate observed changes in gene expression. GFP reporter plasmids for *ler* and the housekeeping gene *rpsM* were transformed into EHEC and the relative fluorescence was calculated during exponential phase. This showed that expression of *ler* was significantly reduced in the presence of aurodox (*Figure 6-6A*). There was no significant change in expression of the housekeeping gene, *rpsM* (*Figure 6-6A*), consistent with the RNA-seq data. A reduction in *ler* expression was also observed for EPEC and *C. rodentium* (*Figure 6-6*) indicating a common mechanism by which aurodox causes a reduction of T3SS expression across these pathogens.

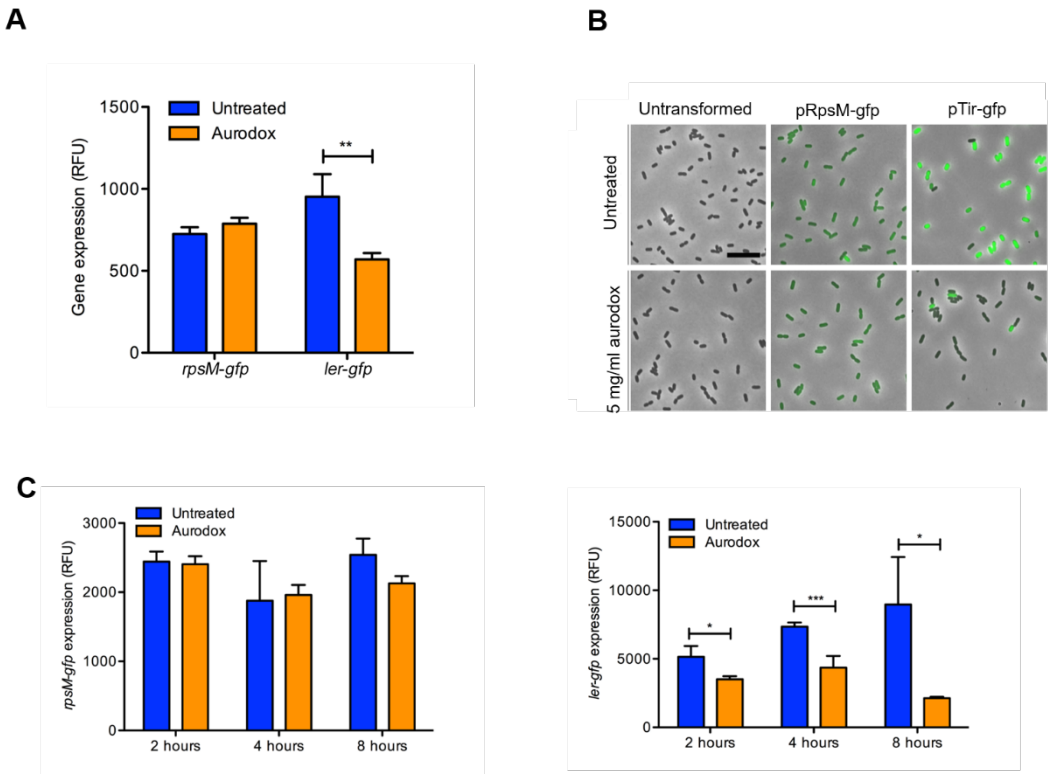


Figure 6-6: **GFP Transcriptional-fusion reporter assay analysis of Type III Secretion expression in EHEC and EPEC**. (A) Activity of housekeeping gene, *rpsM* promoter in response to aurodox. Changes in GFP levels were not significantly significant (p>0.05), Activity of *ler* promoter in response to aurodox treatment, change in GFP expression was significant (p= 0.010). (B) Widefield Epifluorescence microscopy images of EHEC treated with aurodox, scale bars represent 10 µm. (C) Activity of EPEC housekeeping gene, *rpsM* promoter in response to aurodox. Changes in GFP levels were not significantly significant (p>0.05), Activity of *ler* (EPEC) promoter in response to aurodox treatment, changes in GFP levels were significant at all time points (p=0.013, 0.0046 and 0.0246)

6.2.5 Overexpression of Ler allows EHEC to overcome the aurodox-induced T3S knockdown phenotype

Inhibition of the T3SS by aurodox could be explained by a number of mechanisms. These include: (1) directly inhibiting *ler* expression or function, (2) binding to a component of the T3SS resulting in negative feedback of *ler* or (3) affecting an upstream regulator of *ler*. To test if aurodox might be directly binding a structural component of the T3SS, a series of EPEC deletion mutants in various apparatus protein-encoding genes were evaluated and their effect on the transcription of *ler* were studied. The rationale being that if a structural component were a target of aurodox, then this could result in downregulation of *ler*, and therefore the other genes in the T3SS and the wider *ler* regulon. However, analysis of strains lacking four idependant structural components of the T3SS apparatus proteins showed that there was no effect on *ler* promoter activity when assayed using a GFP reporter (*Figure 6-7A*). Moreover, addition of aurodox resulted in reduced expression of *ler* in both WT EHEC and an *escC* deletion mutant (*Figure 6-7B*). These data show that there is no apparent feedback from structural components on the expression of the master regulator of the T3SS itself and that aurodox results in downregulation of *ler* expression, irrespective of whether the T3SS is fully assembled or not.

Having excluded T3SS apparatus proteins as the likely target of aurodox it was postulated that a regulator upstream or downstream of *ler* was the target. To address this, EHEC was transformed with the pVS45 plasmid encoding *ler* under transcriptional control of an arabinose-inducible promoter (Sperandio *et al.*, 2000a). By selectively overproducing Ler the sensitivity of EHEC to aurodox-mediated T3SS inhibition could be tested. The transformed strain was grown in media to induce LEE expression both in the presence and absence of 5 μg/ml aurodox and 2% arabinose to control overexpression of *ler*. The fractions were analysed using immunoblotting

with antibodies against the T3SS protein Tir. In the absence of arabinose (uninduced *ler*) the addition of aurodox results in T3SS inhibition as seen in previous experiments (*Figure 6-7C*). However, overexpression of *ler* by addition of arabinose completely overcomes the phenotype of T3SS-mediated repression normally associated with addition of aurodox (*Figure 6-7C*). This result supports the hypothesis that aurodox acts by either directly acting on Ler or by a mechanism involving an upstream regulator of *ler.*

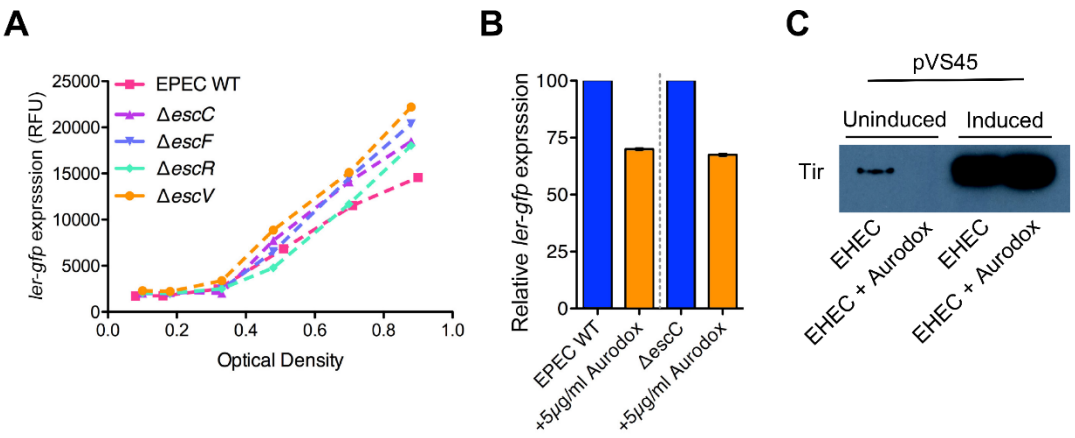


Figure 6-7: **(A) Analysis of *ler* expression in T3SS-defective EPEC mutants.** Mutant strains were transformed with p*ler-gfp.* Optical Density and fluorescence were measured at hourly intervals and expression was quantified in Relative Fluorescence Units. (B) Relative *ler* expression during exponential phase in WT EPEC vs EPEC ΔescC. *Ler* expression in treated samples is expressed as a percentage of untreated samples. (C) Immunoblot analysis of secreted Tir protein in uninduced vs induced EHEC + pVS45 for *ler* overexpression, when untreated and treated with aurodox.

6.2.6 Aurodox does not induce the SOS response in EHEC or *recA*-mediated shiga-toxin (Stx) expression in *Citrobacter rodentium*

Antimicrobial compounds that induce DNA damage are known to stimulate the bacterial SOS response. A key protein in this regulon is RecA, a co-protease that functions in the autocatalytic cleavage of the LexA and λ repressors resulting in depression of approximately 40 genes involved in the SOS response and importantly, an increased expression of bacteriophage-encoded Stx (Imamovic and Muniesa, 2012). As a potential treatment for EHEC infection, we therefore aimed to investigate if aurodox had any undesirable effects on Stx expression. RNA-seq transcriptome data did not reveal differential expression of SOS regulated genes (as defined by Fernández et al (Fernández De Henestrosa *et al.*, 2000) in response to aurodox. To validate this, EHEC was transformed with a plasmid containing *a* p*recA-gfp* fusion and promoter activity measured over a time course (*Figure 6-8A*). Ciprofloxacin was used a positive control of *recA* expression and provided log-fold stimulation compared with the uninduced control (*Figure 6-8A*). In contrast, the data show that aurodox does not induce *recA* transcription in EHEC at four and eight-hour time points (*Figure 6-8A*). The data imply that addition of aurodox does not induce the SOS response. Therefore, by implication, we postulated that Stx expression would remain unaffected.

To test this hypothesis, we directly examined the effect of aurodox on Stx expression *in vitro.* In the UK, Shiga toxin-expressing Enterohaemorrhagic *Escherichia coli* constitutes a Hazard Group 3 pathogen, and as a result, it requires enhanced containment facilities that are not readily available. Therefore, *Citrobacter rodentium* DBS771 (Hazzard Group 2) was used. This strain has been engineered to express Stx via the λ phage system and is ultimately regulated by the same *recA* system as EHEC. In addition to an untreated sample, this strain was grown in the presence of aurodox, ciprofloxacin and aurodox, and ciprofloxacin. The secreted proteins were

then harvested and probed for the presence of Stx using an antibody for the beta subunit of the shiga toxin (*Figure 6-8B*). The resulted bands have shown that aurodox does not induce the expression of Stx, in contrast to ciprofloxacin, for which a band corresponding to Stx can clearly be observed. These findings therefore enhance the potential of the compound to be used as a treatment for EHEC infections, as in contrast to traditional antibiotics, shiga-toxin production is not induced via the *recA* pathway.

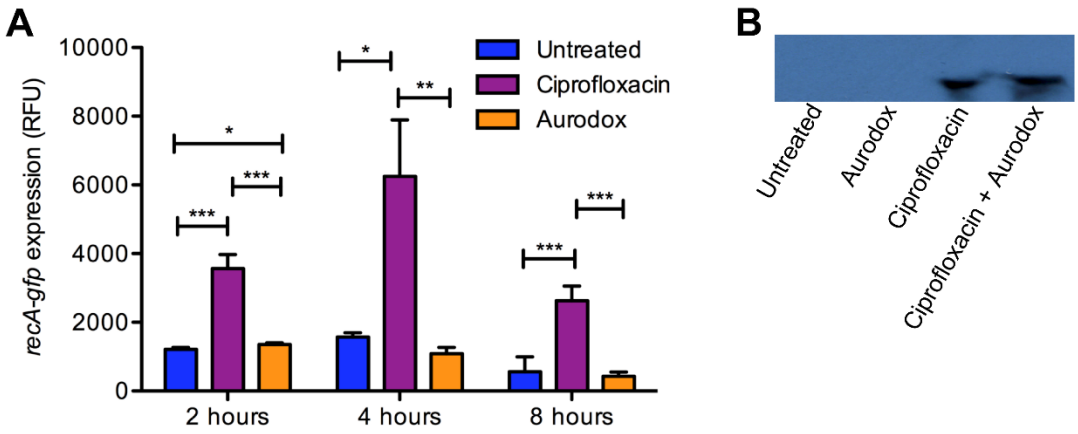


Figure 6-8: **Analysis of effect of aurodox on *recA*- mediated Stx expression. (A) Analysis of in vitro *recA-gfp* expression in aurodox-treated EHEC**. Aurodox and ciprofloxacin were added at equal molar concentration (6 μM). N=4. Error bars correspond to standard deviation from the mean. P -values for aurodox treated versus ciprofloxacin treated were 0.0004, 0.008 and 0.000141 respectively. (B) Immunoblot analysis of Stx expression in *Citrobacter rodentium* DBS100. Primary antibody was used to detect the Stx1 Beta Subunit (6 KDa).

6.2.7 Aurodox heterologously produced by *Streptomyces coelicolor* M1152 pESAC-13A-AurI/ pMS82_tuf2) inhibits *ler* expression in EHEC.

In order to understand the role of the aurodox biosynthetic gene cluster, heterologous expression in *Streptomyces coelicolor* M115 was carried out. During these experiments (Chapter 5) it was demonstrated that *Streptomyces coelicolor* M1152 (+ pESAC-13A-AurI/ pMS82_tuf2) can be engineered to produce aurodox to higher yields than the native producer, *Streptomyces goldiniensis.* In addition, the clean background of *S. coelicolor* M1152 and its superhost properties have facilitated streamlined aurodox purification.

In order to confirm that the aurodox expressed in *Streptomyces coelicolor* M1152 (+ pESAC-13A-AurI/ pMS82_tuf2) was produced in an active form and maintained Type III Secretion Inhibitory activity, agar plugs of *S. coelicolor* M1152 expressing aurodox, in addition to the empty vector control *Streptomyces coelicolor* M1152 (pESAC-13A/ pMS82) and the native aurodox producer *S. goldiniensis* were used in bioassays to visualise the effect of aurodox production on growth and Type III Secretion in EHEC. Zones of inhibition in both EHEC *rpsM*-gfp and *ler*-gfp were measured under visible light and UV light in order to observe the killing effect, characterised by a zone of clearance, in parallel with the T3S downregulatory effect, confirmed by an absence of GFP expression (*Figure 6-9*). The results from these experiments endorse several of the hypotheses of this project. Firstly, zones of inhibition observed around the agar plugs containing *S. coelicolor* M1152 (pESAC-13A-AurI/ pMS82_tuf2) were significantly larger than those associated with *S. goldiniensis* ($p=0.0036$) and the empty vector control strain, which did not show any bioactivity against EHEC. Thus, supporting the results from Chapter 5 which demonstrate that *S. coelicolor* pESAC-13A-AurI/ pMS82_tuf2 produces aurodox to higher levels than the native producer.

In addition, the results show that under UV light, the zone of inhibition associated with *S. coelicolor* pESAC-13A-AurI/ pMS82_tuf2 is significantly increased in diameter from the zone of clearance observed by visible light (p=0.003). These results suggest that as the compound diffuses through the agar, aurodox concentrations decrease until there are no effects of the compound on growth, however, the downregulatory effects on *ler* expression are maintained at these low concentrations. For EHEC expressing *RpsM:gfp*, there is no significant change in zone diameter under UV light, again, confirming the targeted nature of the effects of aurodox on *ler* gene expression and corroborating the results in *Figure 6-1* which show that aurodox downregulates T3S without effecting growth.

The effect of purified aurodox extract on *ler* activity in EHEC was examined using GFP gene fusion reporter assays. In these experiments, semi-quantitative LCMS analysis and *Klebsiella pneumoniae* bioassays were carried out to approximate the aurodox concentration of the extract. Consequently, the aurodox was diluted to approximately 6 μM, the aurodox concentration used throughout these experiments which is has been shown to downregulate Type III Secretion without effecting growth. In these experiments, it can be observed that the aurodox produced by *Streptomyces coelicolor* M1152 pESAC-13A-AurI/ pMS82_tuf2 downregulates *ler* expression in EHEC significantly (p= 0.043) compared to the untreated control, whilst *rpsM* expression remains unaffected (p=0.47). Although the aurodox standard used as a control in this assay did downregulate *ler* expression in EHEC, this change was not statistically significant (p=0.053), which may be explained by the limited shelf-life of the compound which had been in storage for several years. In conclusion, alongside LCMS data, these studies confirm that *S. coelicolor* M1152 pESAC-13A-AurI/ pMS82_tuf2 can biosynthesise aurodox in an active form which maintains its T3S-inhibitory properties.

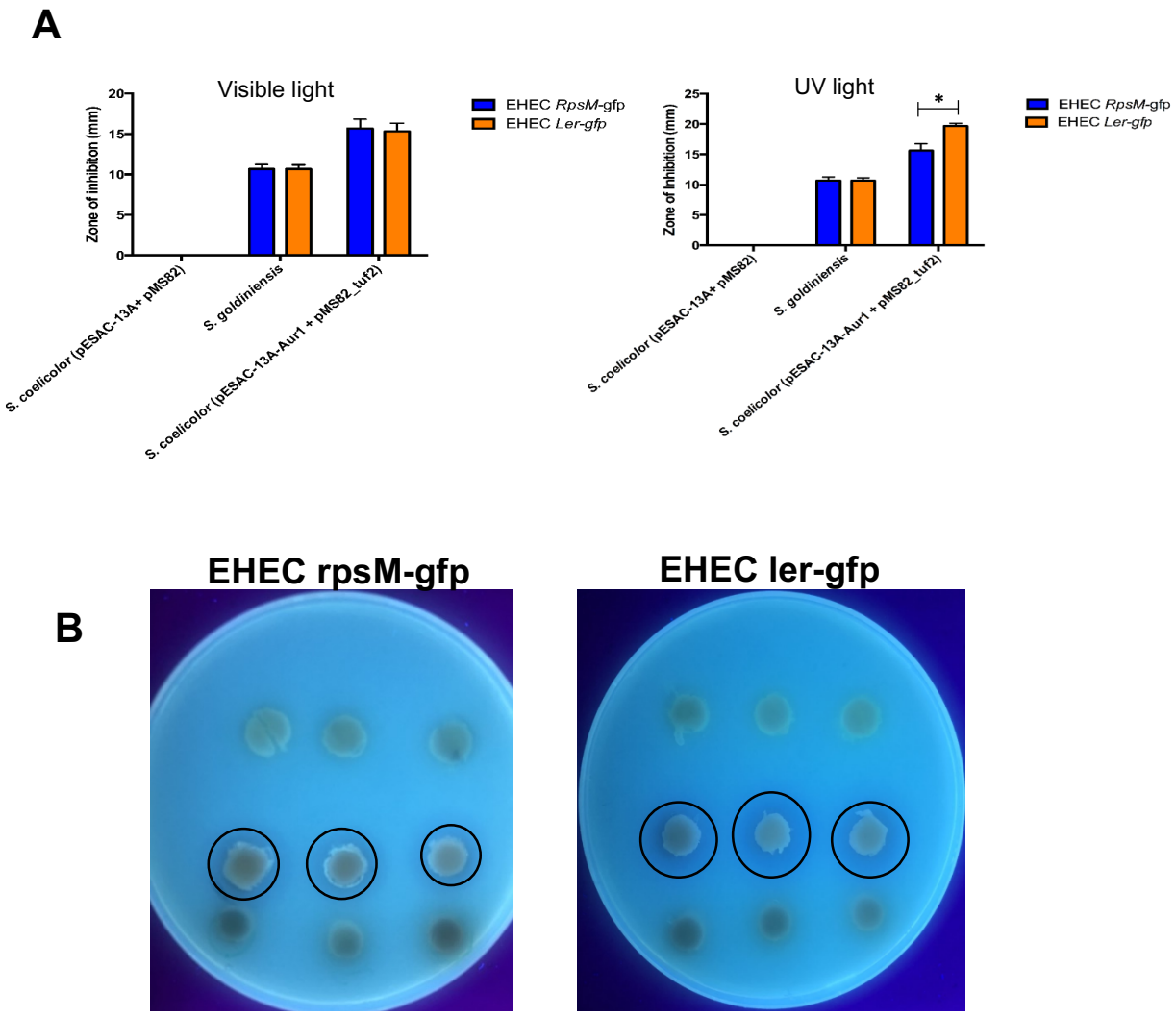


Figure 6-9: **Confirmation of activity of aurodox from heterologous expression in *Streptomyces coelicolor* M1152**. (A) Measurement of zones of inhibition under visible light. Zones of EHEC clearance were measured for EHEC + *rpsM-gfp* and *ler-gfp*. Changes in growth between EHEC + *rpsM-gfp* and *ler-gfp* were not statistically significant for all plugs ($p>0.05$) (B) Measurement of zones of inhibition under UV light. Change in zone of inhibition between EHEC *ler-gfp* and EHEC *rpsM-gfp* is were statistically significant ($p=0.033$). (C) Visual representation of zones of inhibition in EHEC.

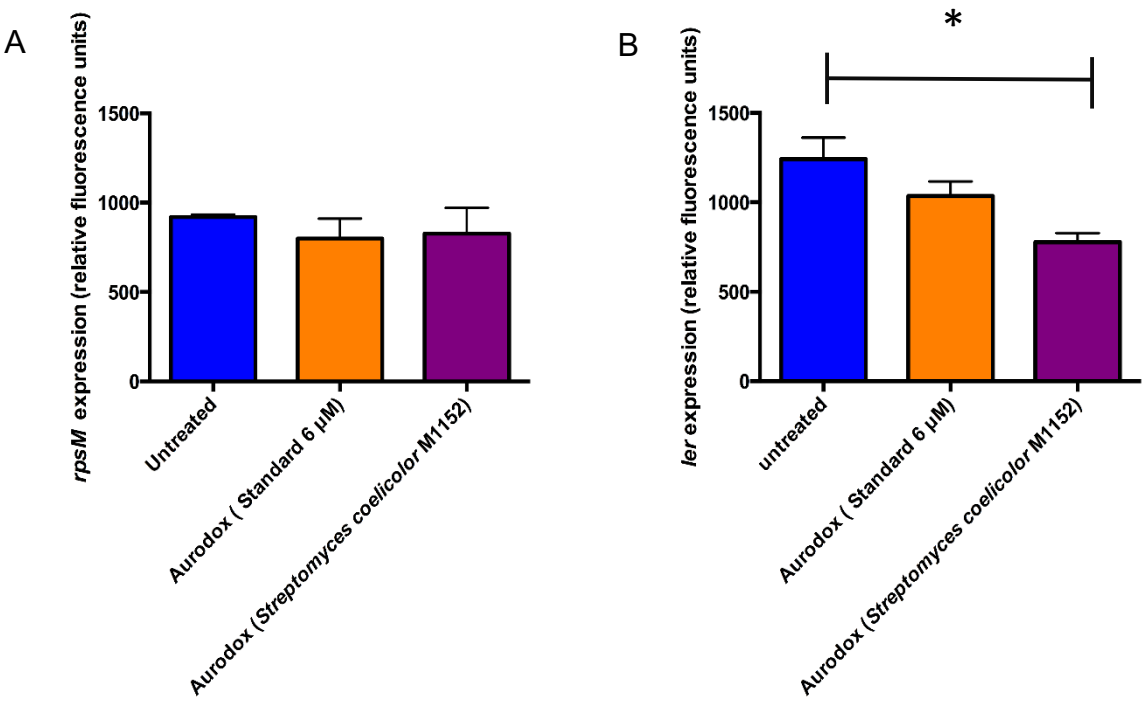


Figure 6-10: **GFP Transcriptional-fusion reporter assay analysis of *ler* expression in EHEC treated with aurodox (standard) and aurodox expressed in *Streptomyces coelicolor* M1152** (A) Expression of housekeeping gene*, rpsM* in response to aurodox treatment. Changes in gene expression were not statistically significant (p>0.05). (B) Expression of *ler* in response to aurodox treatment, change in gene expression between untreated and aurodox from Stx M1152 was statistically significant (p=0.043)

6.3 Summary

This work in this chapter has focussed on several main questions. First, can Aurodox block the function of the T3SS in EHEC O157:H7, as well as EPEC and *C. rodentium*? This is important to address because there are very limited treatment options for EHEC infections due to complications associated with Shiga toxin expression. Second, are there inhibitor-specific transcriptional responses that may help us understand the mode of action? Thirdly, does aurodox alter the bacterial the SOS response affecting expression of bacteriophage-encoded Stx? Finally, is aurodox produced via heterologous expression in *S. coelicolor* M1142 still bioactive and capable of inhibiting the T3SS?

In agreement with Kimura et al, (2011) we observed that aurodox resulted in a concentration dependent inhibition of T3S in EPEC and *C. rodentium.* The same result was found in EHEC, a result that shows a likely common mechanism of inhibition. Importantly, at the concentrations used, there were no effects on bacterial growth rate or viability in any of the species tested. This finding suggests that the mechanism by which aurodox inhibits the T3SS is distinct from the reported target of bactericidal activity elongation factor Tu (EF-Tu) (L Vogeley *et al.*, 2001), a vital component of the protein biosynthesis machinery. It is notable that aurodox is an extremely poor antibiotic against *E. coli*, with prior work in Chapter 3 showing limited inhibition zones at 1 mg/ml; some 200× the concentrations used to inhibit the T3SS (Berger *et al.*, 1973a). Furthermore, the distinct nature of the mechanisms of T3S inhibition and bactericidal activity have positive implications for resistance. Although resistance to aurodox as an antibiotic may arise through mutations in EF-Tu, these mutations would not confer resistance to aurodox as an inhibitor of the T3SS or restore virulence. Although it is possible mutations in *ler* or its regulators may result in resistance to

aurodox as a T3S inhibitor, these would not be strongly selected for in the same way as conventional antibiotics.

Our transcriptomic data provides the first insight into how aurodox affects global gene expression. Remarkably, only 3.24% of the genome showed a significant up or downshift in transcription. This shows excellent specificity and selectivity of inhibition, both very desirable qualities for an anti-virulence compound (Allen *et al.*, 2014). Perhaps predictably, much of the LEE was seen to be downregulated, including transcription of the master regulators *ler* and *grlA* (Iyoda *et al.*, 2006). Further, many of the genes affected were part of the *ler* regulon suggesting a mode of action centred around this protein. Critically, overexpression of *ler*, by an inducible expression system, completely over-rides the effects of aurodox on secretion of Tir via the T3SS. The data from this chapter also confirms that deletion of components of the T3SS system, such as structural proteins, does not lead to transcriptional feedback on *Lee1*. In conjunction with the transcriptomics data, the proposed working model is that aurodox binds a target that affects transcription of *ler*. With several other regulatory pathways and environmental signals known to regulate the expression of *ler,* there are many potential targets (Sperandio *et al.*, 2000a). The complexity and high molecular weight of aurodox suggests that the compound may have the ability to compromise the integrity of the outer membrane on cell entry, and hence activate the expression of the membrane stress response sigma factor, *rpoE.* Increased expression of this sigma factor has been reported to reduce T3S via downregulation of *ler*, particularly in the presence of Zinc (Xue *et al.*, 2015). However, our transcriptional analysis has shown 1.54-fold downregulation of the sigma factor, and therefore, it is believed this pathway is unlinked.

It has also been demonstrated, for the first time, that aurodox does not stimulate expression of *recA*, a key protein that helps mediate the SOS response in *E. coli* (Kim

and Little, 1993). The SOS response is known to be a key regulator of lambdoid bacteriophage transcription by stimulating autocleavage of the phage repressor, cI, in response to DNA damage, triggering a cascade that leads to transcription of the Q anti-terminator transcript and the lytic cycle (Fuchs *et al.*, 1999). As both the *StxA* and *StxB* genes in EHEC strains are located on the genomes of resident prophages (Nguyen and Sperandio, 2012), it is not surprising that they are upregulated by the SOS response that can induced by treatments that damage bacterial DNA, including UV light, selected antibiotics, and mitomycin C. This response to traditional broad-spectrum antibiotics has made treatment of EHEC problematic and highlights an important advantage of aurodox: it can suppress virulence without the “sting in the tail” associated with Shiga toxin expression and damage.

Finally, the work in this chapter has demonstrated that aurodox which has been purified through heterologous expression of the aurodox gene cluster in *S. coelicolor* M1152 is active and is able to inhibit the expression of the Type III Secretion System master regulator, *ler.* This result is important, significant, and the validation of this aurodox activity has positive implications for the future of aurodox production via fermentation as the improved production strains can facilitate the upscaling of aurodox production for the drug discovery pipeline.

Chapter 7 : Discussion and Future work

There is an urgent need for novel strategies to combat bacterial infections (Jackson, Czaplewski and Piddock, 2018). With the instance of antibiotic resistance increasing and a downturn in the discovery of suitable antibiotics for clinical use in humans, novel approaches must be taken. This thesis aimed to investigate one such approach: an antivirulence strategy for the treatment of diarrheagenic *E. coli* infections of the gut. There are currently no antibiotic therapies recommended for the treatment of EHEC and EPEC, due to the severe side effects associated with their use. Therefore, this thesis has focussed on the understanding of the properties of aurodox, primarily its biosynthesis and mechanism of action, with a view to repurposing this shelved antibiotic as antivirulence therapy.

7.1 Whole genome sequencing of *Streptomyces goldiniensis*.

To fulfil the objectives of this project, a robust whole genome sequence of *S. goldiniensis* was required to enable the characterisation of aurodox biosynthesis. It was postulated that a high quality genome sequence could provide essential information regarding the genes involved in aurodox biosynthesis with enough resolution to decipher distinct similarities and differences between the aurodox BGC and the gene clusters of other kirromycin-type elfamycins (Weber *et al.*, 2008; Thaker *et al.*, 2012; Robertsen *et al.*, 2018). Therefore, a combination of PacBio and Illumina MiSeq sequencing technologies were applied to enable the complementation of the short reads provided by Illumina MiSeq (2 x 300 bp), with long reads (20 kb) provided by PacBio. Using this method, which was previously validated by Gomez-Escribano *et al.*, (2016) using enhanced sequencing depth and coverage were acquired which facilitated the resolving of the repetitive sequence regions in the modular PKS and NRPS domains that could not be resolved using a single technology. It was hypothesised that the use of the Oxford Nanopore MinION to provide ultra-long reads

would further enhance this sequencing depth to allow for the accurate *de novo* assembly of the *S. goldiniensis* genome (Kruasuwan *et al.*, 2017). The MinION sequencer is compact, with a short library preparation procedure to enable use of the sequencer in field studies (Naômé *et al.*, 2018). However, these advantages were contrasted by poor data output, with raw reads amounting to a 6-fold coverage of the genome (*Figure 3-8*). This is a poor return, and although some ultra-long individual reads of up to 121 kb did aid in the reduction of the contig number, the return was relatively poor for the large investment required to sequence with MinION. One issue that has been identified with Nanopore sequencing of *Streptomyces* is the identification of long G and C repeat sequences as errors, resulting in the deactivation of the sequencing channels. This was observed during the MinION sequencing procedures of this study and resulted in the systematic reduction of the pores available for sequencing and therefore led to a reduction in data output. Despite a previous attempt being made to assess the suitability of Nanopore sequencing for high GC organisms, *Bordetella pertussis* (60 % GC) was used for validation (Jain *et al.*, 2015). With the majority of *Streptomyces* species harbouring genomes with over 70% GC content, novel procedures and algorithms consequently been implemented which prevent the deterioration of the pores and consequently, the quality of *Streptomyces* sequence has improved (Laver *et al.*, 2015). Several *Streptomyces* genomes have now been sequenced using MinION sequencing (Naômé *et al.*, 2018; Yusof *et al.*, 2020), and therefore, this could be used in future if in attempt to close the genome of *S. goldiniensis* by assigning the whole genome to one contig.

The sequencing data provided by this triple-sequencing approach of *S. goldiniensis* was ultimately assembled using the *S. bottropensis* genome as a basis for scaffold based assembly (*Table 3-B*). This resulted in a genome assembly consisting of nine contigs contributing to the large linear chromosome, with the presence of plasmids

ruled out by PFGE. Although the genome could not be closed, the final genome assembly generated in this study was sufficient for detailed annotation, comparative genomic approaches and importantly, genome mining for the presence of BGCs. Significantly, this has led to the identification of 33 BGCs, including the putative aurodox BGC which formed the basis of subsequent studies (*Figure 3-14*)

7.2 Aurodox and kirromycin: can comparing the BGCs provide clues as to how the clusters arose?

The identification of the aurodox BGC has enabled the characterisation of the putative enzymatic steps involved in aurodox biosynthesis. Moreover, profound homology between the aurodox and kirromycin BGCs has been demonstrated, with common genes including the core PKS/NRPS units encoded by *aur/kirAI-AVII.* In addition, recombination detection software has predicted several recombination events between the two gene clusters which emphasise their evolutionary relatedness.

Despite the similarities in the structure, function and genetic basis of aurodox and kirromycin, this study has recognised multiple differences between the BGCs, at both a sequence and functional level. One key difference is the distinguishing presence of the SAM-dependent *O*-methyltransferase (AurM*) within the aurodox cluster, responsible for catalysing the conversion of the final aurodox intermediate, kirromycin, to aurodox. With aurodox and kirromycin displaying such similar target and activity properties, the effects of this methylation were previously poorly understood. This therefore provoked questions as to how these clusters evolved, their origin and accordingly, has the capacity to methylate kirromycin been lost or gained? To address this question, the relative advantages of the methylation to kirromycin were assessed through expression of AurM* in the kirromycin producer *S. collinus.* Growing the strains adjacently on agar plates indicates that aurodox can inhibit the growth of the empty vector control, suggesting that the native immunity mechanisms of *S. collinus*

are insufficient to confer resistance to aurodox. This suggests that there must have been co-evolution of these self-resistance mechanisms (AurT and Tuf2) with the acquisition of AurM*. It also provides evidence that the methylation of aurodox was gained rather than lost, indicating that kirromycin is an evolutionary predecessor of aurodox. Yet with both strains encoding the AurT/KirTI (65% aa similarity) transport mechanisms and the *tuf2* immunity genes, it appears extraordinary that a simple methylation to the compound could result in the circumvention of the immunity mechanisms in *S. collinus* to cause the profound effects on growth which are demonstrated in *Figure 5-21* (Reddy *et al.*, 2012). However, this is an effect that can be observed across many organisms, as methylation is a chemically efficient method of modulating biological function. The ability of this simple modification- magic methylation- to change or improve the potency of a compound has resulted in the development of many synthetic methodologies to facilitate this reaction as well as more complex alkylation, for use in medicinal chemistry (McKean, Hoskisson and Burley, 2020) . Moreover, it has been shown previously that alterations as seemingly innocuous as a change in molecular charge can result in impaired ability of an exporter/importer pump to transport molecules across the membrane (Mirandela *et al.*, 2019). Therefore, it is possible that this methylation is sufficient for aurodox to avoid recognition by these transporters in *S. collinus*, allowing aurodox to accumulate within the cytoplasm and enabling a fatal interaction between aurodox the EF-Tu of *S. collinus*. However, given the ability of *S. collinus* to biosynthesise and export aurodox when heterologously expressing AurM*, it appears that KirT is able to export aurodox to some extent, but with less efficiency. Further evidence to support this can be found in *Chapter 5,* where the ability of *S. collinus* to resist aurodox at 1 mg/ml throughout its lifecycle is demonstrated. One potential explanation for this could be that KirT has a greater affinity for kirromycin, which is biosynthesised during stationary phase, requiring export. In this case, the export of kirromycin is prioritised over

aurodox. In addition, the presence of excess kirromycin or aurodox is likely to prevent the expression of *kirT*, which is controlled by the TetR-type regulator, KirR which is inactivated when bound to kirromycin (Weber *et al.*, 2008). Although it is possible that the confirmed presence of two kirromycin clusters within the *S. collinus* genome could counteract this, the relative expression of each gene cluster has not yet been investigated (Rückert *et al.*, 2013). To address these questions, the relative abilities of KirT to export aurodox can be explored via determination of MIC in response to the expression of these proteins (Thaker *et al.*, 2012). It should be noted that adjacent growth is not a robust competition method, and fails to take in to account the diversity of the ecological niches which the strains could encounter natively during ecological warfare. However, this snapshot of the interaction between the strains does neatly demonstrate the advantages to enzymatic modification of specialised metabolites for avoidance of native resistance mechanisms (Zhang *et al.*, 2015).

With experimental evidence supporting a model whereby the addition of a methyl group to kirromycin to form aurodox facilitates the avoidance of the native elfamycin resistance mechanisms of *S. collinus,* it appears that kirromycin is not only a direct precursor of aurodox but is the evolutionary predecessor of the molecule. If this is the case, the cluster may have acquired an early version of AurM* from an *inter* or *intra* genomic source. There are two additional *O*-methyl transferases encoded within the *S. goldiniensis* genome, annotated as mitomycin C *O*-methyl transferases, similarly to AurM*. However, amino acid sequence similarity remains low at 35% and hence it is possible that this *aurM** was acquired through horizontal gene transfer (Theobald *et al.*, 2018). Furthermore, preliminary studies have found that the T3SS inhibitory properties displayed by aurodox are not shared by kirromycin. This is further evidence of the profound functional effects that simple chemical modifications to natural products can induce. Therefore, it is possible that functionally, aurodox may benefit

from further derivatisation which this study has formed the basis for. Fundamentally, the presence of the aurodox cluster on pESAC-13A-AurI can facilitate genetic manipulation such as gene disruptions or modifications through the addition of enzymes such as halogenases (Sánchez *et al.*, 2002; Kittilä *et al.*, 2017; Qin *et al.*, 2017). The development of *ler* GFP reporter assays (*Chapter* 6) in this study as a method of gauging levels of T3SS can therefore be applied as a screen for derivatives with improved T3SS inhibitory properties to prioritise those for further analysis with improved pharmacokinetic properties.

7.3 Application of modern computational analyses methods to update the kirromycin BGC annotation

This study has applied contemporary computational analysis methods to both the aurodox and kirromycin BGCs which has provided not only a detailed, modular annotation of the aurodox BGC, but has facilitated advances in the annotation of the kirromycin BGC. This comparison has provided some evidence to support the hypotheses of Weber et al ., who had previously assigned putative domains to the PKS/NRPS genes (Robertsen *et al.*, 2018). Moreover, this analysis has highlighted previously unknown aspects of the kirromycin/aurodox gene clusters and identified potential discrepancies in the annotation of the kirromycin encoding genes.

In the published annotation of the kirromycin gene cluster (Weber *et al.*, 2008), both the Mfs-type exporter KirT protein is predicted to have the N and C terminal polypeptide sequences encoded separately (*kir*TI/*kir*TII, *kirRI/kirRII*). However, the historic Roche 454 sequencing technology used to sequence the kirromycin cluster is prone to introducing premature stop codons (Trimble *et al.*, 2012). Therefore, it is likely that these proteins are encoded by one ORF, as they are in the aurodox BGC, with single nucleotide errors sufficient to induce these mis-annotations. In addition, there are few examples in the literature where these exporters are encoded by

independent genes and the benefit to doing so appears unclear (Cuthbertson and Nodwell, 2013; Zhou *et al.*, 2016).

Furthermore, the updates to the clustertools algorithm from antiSMASH has provided further evidence to support the presence of the dehydratase domains in Kir/AurAI-III that were annotated purely by predicted function, with the genetic and structural basis of these domains uncharacterised at the time of initial publication (Akey *et al.*, 2010). Computational analysis using antiSMASH has identified these DH domains within both the aurodox and kirromycin gene clusters,and therefore, identification of these DH domains by two different algorithms corroborated the evidence and augments the role of the domains in biosynthesis whilst exemplifying the functional homology shared between the clusters (Weber *et al.*, 2008). However, biochemical studies using DH-disruption mutants should be used to provide conclusive evidence of the role of these domains aurodox and kirromycin biosynthesis.

7.4 Engineering of a heterologous aurodox production strain for high efficiency purification for use in further mechanism of action studies.

The heterologous expression of aurodox performed in *Chapter 5* served two purposes: 1) Confirmation of the role of the putative aurodox BGC in its biosynthesis, and 2) to provide candidate strains for upscaling of aurodox production for mechanism of action studies. Namely, *S. coelicolor* M1152 + pESAC-13A-AurI and *S. venezuelae* + pESAC-13A-AurI +pMS82_tuf2 provided significantly improved aurodox titres compared to *S. goldiniensis* under consistent fermentation conditions (*Figure 5-9, 5-16*). The improvement in aurodox production by *S. coelicolor* M1152 expressing pESAC-13A-AurI was also demonstrated by the relative increase in GFP inhibition when cultured with EHEC + *ler::GFP* (*Chapter 5*). Therefore, this study has formed a

basis for additional improvements to ultimately construct an aurodox production strain for use in scaled-up fermentations, Yet, there are further avenues which should be explored for genetic modification of these strains to further improve aurodox yields before this production strain is finalised for use in fermentations.

One such avenue that could be explored is the expression of a *cis*-acting PPTase in the production strain. Unlike the kirromycin biosynthetic gene cluster, the aurodox BGC does not encode a canonical PPTase. Although, as discussed in *Chapter 4,* it is possible that the hypothetical gene *aurHI* which is not found in the kirromycin cluster could encode an uncharacterised PPTase, it is thought that the PPTase present 10 kb upstream of the aurodox supercluster is likely to act upon the ACPs and PCPs of the aurodox NRPS and PKSs. Given that the genomic insert encoding aurodox in pESAC-13A is only 155 kb, the presence of a PPTase on pESAC-13A-AurI can be ruled out, even without sequencing of the PAC. Currently, the activation of the ACPs/PCPs involved in aurodox production relies on the native PPTases of the heterologous hosts. Therefore, it is hypothesised that cloning a PPTase on to this PAC could lead to increases in aurodox titre (Thaker *et al.*, 2012). Ultimately, when a final production strain is developed it will be tested in an 8 L fermentation system in aim to produce substantial quantities of aurodox for further mechanism of action studies.

7.5 Clinical applications of aurodox: development of aurodox as an anti-virulence therapy.

The ability of aurodox to inhibit the Type III Secretion system via the downregulation of the EHEC virulence gene regulator *ler* is remarkable. The ecological role of the compound is clear: by binding to the EF-Tu of competing organisms, aurodox can inhibit translation and hence confer a competitive advantage to *S. goldiniensis.* However, the purpose of producing a compound which inhibits virulence in human

enteric pathogens is less clear, with the organisms unlikely to interact in the environment (Porter, 1971; Ferens and Hovde, 2011). However, many natural product inhibitors of the T3SS have been identified from a range of sources including the caminosides from the marine sponge *Caminus sphaeroconia* and Cytosporone B from the *Dothiorella* fungus (Linington *et al.*, 2002; Zhan *et al.*, 2008), with most of these metabolites also possessing additional properties such as insecticidal effects. Originally, it was hypothesised that perhaps the inhibition of T3S in EHEC and EPEC was an indirect result of aurodox binding to EF-Tu. However, with a complete inhibition of T3S at 5 μg/ml, with no effect on growth rate or changes in the expression of the translation machinery genes, this hypothesis was rejected. Therefore, the dual properties of aurodox can be attributed to a 'molecular coincidence' that has the potential for exploitation for human, and potentially animal, health.

Despite the previous identification of multiple EHEC T3SS inhibitors (Hudson *et al.*, 2007; Pendergrass and May, 2019), aurodox has distinct advantages over these compounds which nominate it as a primary candidate for further development for medicinal use. Firstly, where most other Type III Secretion systems inhibitors target the structure of the secretion system or the translocation of a single effector protein involved in colonisation, aurodox inhibits the expression of the entire pathogenicity island and hence, the effects of the compound are more profound. In *Chapter 6,* a 7000-fold reduction in colonisation efficiency of EHEC induced by aurodox exemplifies these effects. Moreover, aurodox is the first natural product T3SS inhibitor which has been proven not to negatively impact Stx production in EHEC. Furthermore, aurodox is a compound of bacterial origin, meaning its biosynthetic genes are clustered, which is not the case for the fungi or marine sponges which have been a source of other natural product T3SS inhibitors. This innate property opens the compound to the prospect of synthetic biology approaches to the production of

derivatives. Again, this study has provided the foundations for these derivatisation experiments through the characterisation of the enzymatic steps involved in biosynthesis and the cloning of the aurodox gene cluster on to pESAC-13A, which can enable the targeted genetic manipulation which was near impossible in *S. goldiniensis*. Moreover, bacterial hosts enable fermentation procedures which can provide large quantities of compounds for industrial or medicinal use, such as those used to produce clavulanic acid (Li and Townsend, 2006)

There are many contributing factors which can influence the decision to move forward with the process of drug development. Primarily, there must be a clinical need (Zambelloni, Marquez and Roe, 2015; Totsika, 2016; Calvert, Jumde and Titz, 2018; Jackson, Czaplewski and Piddock, 2018). In the case of EHEC, this clinical need is great, with the negative consequences of antibiotic treatment resulting in current treatment strategies relying on simple interventions such as rehydration therapy (Goldwater and Bettelheim, 2012). However, there are fundamental aspects of aurodox that must be investigated before it can be developed for the clinically. Firstly, a more definitive target of aurodox in EHEC should be identified at the protein level. Moreover, the ability of aurodox to prevent EHEC infections is profound, and studies by Kimura *et al,* (2011) have shown that the compound can inhibit the pathogenesis of *C. rodentium* in mice, leading to their survival. However, the ability of aurodox to reverse a previously established EHEC infection must be analysed to establish a suitable clinical role for the compound. If the compound can reverse EHEC infections to clear them, its use as treatment for EHEC patients is more likely. In the case that aurodox can only prevent the establishment of infection, it may yet have a role in prophylaxis (community EHEC outbreaks) or in agriculture, by preventing cows from becoming colonised with EHEC initially to prevent its infiltration of the food chain (Ferens and Hovde, 2011). There is a precedent for the use of aurodox in agriculture,

as it was previously used as a growth promoter in poultry (Berger *et al.*, 1973a). The compound has also been listed in multiple patents for clinical trials. Despite the results of these studies having limited accessibility, the pharmacokinetic properties of the compound in humans is available, demonstrating the historic trialling of the compound in humans. The fact that this compound has been used repeatedly in trials for various ailments is indicative that the issues with aurodox as a treatment were a result of poor efficacy a opposed to negative side effects (*Drugbank*, 2020). Hence, these properties further enhance the potential of aurodox to be used as an antivirulence therapy for EHEC and EPEC infections.

7.6 Conclusions and Future Work

The works encompassed in this thesis have advanced our understanding of aurodox in several respects. Initially, fundamental research into the previously poorly characterised aurodox production strain *S. goldiniensis* was carried out, with the morphology and physiology characterised, leading to the development of a robust small-scale aurodox fermentation procedure in the laboratory. These results enhanced the subsequent whole genome sequencing of *S. goldiniensis,* with the final assembly consisting of nine contigs. In future, PCR experiments may be used to bridge the gaps between the contigs, however the genome in its current form has been accepted to the NCBI genome database, passing stringent quality checks in the process.

A central aim of these works was the identification and elucidation of the aurodox biosynthetic gene cluster and biosynthetic pathway. This objective has been met through the use of a combination of whole genome sequencing and genome mining followed by heterologous expression and subsequent biochemical validation of aurodox production. In addition, an *aurAI* gene knockout version of pESAC-13A-AurI has been constructed and expressed in *S. coelicolor* M1152, however, the

metabolites produced by the strain could not be analysed within the timeframe of this thesis due to restrictions implement as a result of the global SARS-COVID2 pandemic. Therefore, the analysis of these extracts should be prioritised to reinforce the role of the predicted aurodox PKS units in biosynthesis. Moreover, this study, for the first time, has identified the 25-gene, 87 kb aurodox BGC with significant homology to the kirromycin gene cluster, and the enzymatic steps involved in aurodox biosynthesis have been computationally characterised. The similarities and differences between these two gene clusters have also been extensively analysed, resulting in the identification of targets for strain improvement and derivatisation and the confirmation of the role of AurM* in the conversion of kirromycin to aurodox. In conclusion, the data support the idea that kirromycin is both an evolutionary and biosynthetic precursor to aurodox, with several HGT and recombination events identified between the aurodox, kirromycin, bottromycin and unknown BGCs resulting in the formation of the aurodox cluster. The results of this study have also implicated Tuf2 as an essential gene for aurodox immunity.

This work also provides a framework for future experimental characterisation of specific genes in aurodox biosynthesis, using the works of Weber *et al,* 2008 and Robertsen *et al,* 2018 as basis for the prioritisation of genes for disruption. In addition, further tailoring enzymes such as halogenases can be used to produce novel aurodox derivatives (Qin *et al.*, 2017). There is also potential for the application of a synthetic chemistry approach to the production of aurodox derivatives. In recent studies, the use of SAM analogues to introduce diverse chemical groups in place of a standard methylation has been demonstrated (McKean *et al.*, 2019). These methods represent efficient approaches for the generation of a bank of potentially patentable structures, which can subsequently be tested for bioactivity through the use of the T3SS, *ler::gfp* reporter assay system described in *Chapter 6.*

Finally, this study has shown conclusively that aurodox inhibits the expression of the EHEC T3SS via the downregulation of the master virulence gene regulator *ler.* It is significant that aurodox is able to inhibit EHEC colonisation of the human gut without inducing the *recA*-mediated Stx response, which has deadly consequences in patients (McHugh *et al.*, 2019). In order for aurodox to progress into the clinic, there are several immediate questions that must be addressed. As a priority, a definitive mechanism of action must be elucidated through the identification of a definitive protein or structural target. The data described in *Chapter 5* supports a model whereby aurodox acts upstream of *ler* in the regulatory cascade to inhibit its expression, however this must be proven experimentally. One such method to achieve this is the identification of aurodox insensitive mutants, namely, those which maintain *ler* expression after treatment with aurodox. Several passages of aurodox treatment followed by separation of fluorescing strains using a cell sorter will facilitate the identification of these mutants. The resistant phenotype can be subsequently confirmed using SDS-PAGE analysis of secreting proteins before WGS is used to identify 'resistance-conferring' mutations. Furthermore, preparation of a photoreactive group for UV light-induced covalent trapping of aurodox binding proteins may be used to allow protein pull-down assays to identify a specific protein target of aurodox, however, this may be complicated by the abundance of EF-Tu in the cell. There are also many wider questions to address regarding the clinical role of anti-virulence therapies and their wider roll in the microbiome. Given the increase in symptom severity observed in EHEC patients as a result of the dysbiosis of the gut microflora, it is important that the effects of the use of sub-bactericidal aurodox concentrations on the gut microbiome is assessed (Allen *et al.*, 2014; Langdon, Crook and Dantas, 2016; Calvert, Jumde and Titz, 2018) that may require vast bodies of research to understand, and are likely outside of the scope of this project.

In summary, this thesis has addressed two main aspects of the aurodox problem: its biosynthesis and its mechanism of action. Significant progress has been made in the understanding of both and as result of the conclusions detailed above, this body of work provides a basis for the future development of aurodox as antivirulence strategy for the treatment of Enteropathogenic and Enterohaemorhagic *E. coli* infections of the human gut.

Chapter 8 References

AbdelKhalek, A., Abutaleb, N. S., Mohammad, H. and Seleem, M. N. (2019) 'Antibacterial and antivirulence activities of auranofin against *Clostridium difficile'*, *International Journal of Antimicrobial Agents*, 53(1), pp. 54–62. doi: 10.1016/j.ijantimicag.2018.09.018.

Akey, D. L., Razelun, J. R., Tehranisa, J., Sherman, D. H., Gerwick, W. H. and Smith, J. L. (2010) 'Crystal Structures of Dehydratase Domains from the Curacin Polyketide Biosynthetic Pathway', *Structure*. doi: 10.1016/j.str.2009.10.018.

Alam, M. T., Merlo, M. E., Takano, E. and Breitling, R. (2010) 'Genome-based phylogenetic analysis of Streptomyces and its relatives', *Molecular Phylogenetics and Evolution*. doi: 10.1016/j.ympev.2009.11.019.

Alanjary, M., Kronmiller, B., Adamek, M., Blin, K., Weber, T., Huson, D., Philmus, B. and Ziemert, N. (2017) 'The Antibiotic Resistant Target Seeker (ARTS), an exploration engine for antibiotic cluster prioritization and novel drug target discovery', *Nucleic Acids Research*. doi: 10.1093/nar/gkx360.

Alanjary, M., Steinke, K. and Ziemert, N. (2019) 'AutoMLST: an automated web server for generating multi-locus species trees highlighting natural product potential', *Nucleic acids research*, 47(W1), pp. W276–W282. doi: 10.1093/nar/gkz282.

Alexander, D. C. and Jensen, S. E. (1998) 'Investigation of the Streptomyces clavuligerus cephamycin C gene cluster and its regulation by the CcaR protein', *Journal of Bacteriology*.

Allen, R. C., Popat, R., Diggle, S. P. and Brown, S. P. (2014) 'Targeting virulence: can we make evolution-proof drugs?', *Nature reviews. Microbiology*. Nature Publishing Group, 12(4), pp. 300–308. doi: 10.1038/nrmicro3232.

Amara, A., Takano, E. and Breitling, R. (2018) 'Development and validation of an updated computational model of Streptomyces coelicolor primary and secondary metabolism', *BMC Genomics*. doi: 10.1186/s12864-018-4905-5.

Aminov, R. I. (2010) ‘A brief history of the antibiotic era: Lessons learned and challenges for the future’, *Frontiers in Microbiology*, 1(DEC). doi: 10.3389/fmicb.2010.00134.

Antony-Babu, S., Stien, D., Eparvier, V., Parrot, D., Tomasi, S. and Suzuki, M. T. (2017) ‘Multiple Streptomyces species with distinct secondary metabolomes have identical 16S rRNA gene sequences’, *Scientific Reports*. doi: 10.1038/s41598-017-11363-1.

Baba, T. *et al.* (2006) ‘Construction of *Escherichia coli* K-12 in-frame, single-gene knockout mutants: The Keio collection’, *Molecular Systems Biology*. doi: 10.1038/msb4100050.

Balsells, E., Shi, T., Leese, C., Lyell, I., Burrows, J., Wiuff, C., Campbell, H., Kyaw, M. H. and Nair, H. (2019) ‘Global burden of *Clostridium difficile* infections: A systematic review and meta-analysis’, *Journal of Global Health*, 9(1). doi: 10.7189/jogh.09.010407.

Baltz, R. H. (2010) ‘Streptomyces and Saccharopolyspora hosts for heterologous expression of secondary metabolite gene clusters’, *Journal of Industrial Microbiology and Biotechnology*. doi: 10.1007/s10295-010-0730-9.

Banjo, M., Iguchi, A., Seto, K., Kikuchi, T., Harada, T., Scheutz, F. and Iyoda, S. (2018) ‘*Escherichia coli* H-Genotyping PCR: A complete and practical platform for molecular h typing’, *Journal of Clinical Microbiology*. doi: 10.1128/JCM.00190-18.

Bankevich, A. *et al.* (2012) ‘SPAdes: A New Genome Assembly Algorithm and Its Applications to Single-Cell Sequencing’, *Journal of Computational Biology*, 19(5), pp. 455–477. doi: 10.1089/cmb.2012.0021.

Bao, K. and Cohen, S. N. (2003) ‘Recruitment of terminal protein to the ends of Streptomyces linear plasmids and chromosomes by a novel telomere-binding protein essential for linear DNA replication’, *Genes and Development*. doi: 10.1101/gad.1060303.

Beckham, K. S. H. and Roe, A. J. (2014) ‘From screen to target: insights and approaches for the development of anti-virulence compounds’, *Frontiers in Cellular and Infection Microbiology*, 4(September), pp. 1–8. doi: 10.3389/fcimb.2014.00139.

Beilhartz, G. L., Tam, J., Zhang, Z. and Melnyk, R. A. (2016) 'Comment on A small-molecule antivirulence agent for treating *Clostridium difficile* infection', *Science Translational Medicine*. doi: 10.1126/scitranslmed.aad8926.

Bentley, R. and Meganathan, R. (1982) 'Biosynthesis of vitamin K (menaquinone) in bacteria', *Microbiological Reviews*. doi: 10.1128/mmbr.46.3.241-280.1982.

Bentley, S. D. *et al.* (2002) 'Complete genome sequence of the model actinomycete Streptomyces coelicolor A3(2)', *Nature*. doi: 10.1038/417141a.

Berchtold, H., Reshetnikova, L., Reiser, C. O., Schirmer, N. K., Sprinzl, M. and Hilgenfeld, R. (1993) 'Crystal structure of active elongation factor Tu reveals major domain rearrangements.', *Nature*, 365(6442), pp. 126–132. doi: 10.1038/365126a0.

Berger, julius, Lehr, H. H., Teitel, S., Maehr, H. and Grunberg, E. (1972) 'A New Antibiotic X-5108 of Streptomyces origin I. Production, isolation and Properties', *The Journal of antibiotics*, XXVI(I), pp. 15–22.

Berger, J., Lehr, H., Teitel, S., Maehr, H. and Grunberg, E. (1973a) 'A new antibiotic X-5108 OF Streptomyces origin. I. Production, isolation and properties', *Journal of Antibiotics*, 26(1), pp. 15–22. Available at: http://www.ncbi.nlm.nih.gov/pubmed/4781274.

Berger, J., Lehr, H., Teitel, S., Maehr, H. and Grunberg, E. (1973b) 'A new antibiotic X-5108 OF Streptomyces origin. I. Production, isolation and properties', *Journal of Antibiotics*, 26(1), pp. 15–22. doi: 10.7164/antibiotics.26.15.

Bian, X., Huang, F., Stewart, F. A., Xia, L., Zhang, Y. and Müller, R. (2012) 'Direct Cloning, Genetic Engineering, and Heterologous Expression of the Syringolin Biosynthetic Gene Cluster in *E. coli t*hrough Red/ET Recombineering', *ChemBioChem*. doi: 10.1002/cbic.201200310.

Bibb, M. J. (2005) 'Regulation of secondary metabolism in streptomycetes', *Current Opinion in Microbiology*, pp. 208–215. doi: 10.1016/j.mib.2005.02.016.

Bibb, Mervyn J., Fernández-Martínez, L. T., Borsetto, C., Gomez-Escribano, J. P., Bibb, Maureen J., Al-Bassam, M. M. and Chandra, G. (2014) 'New insights into chloramphenicol biosynthesis in Streptomyces venezuelae ATCC 10712', *Antimicrobial Agents and Chemotherapy*, 58(12), pp. 7441–7450. doi: 10.1128/AAC.04272-14.

Bingle, L. E. H. *et al.* (2014) 'Microarray analysis of the Ler regulon in enteropathogenic and enterohaemorrhagic *Escherichia coli* strains', *PLoS ONE*, 9(1), pp. 1–12. doi: 10.1371/journal.pone.0080160.

Blanco, P., Hernando-Amado, S., Reales-Calderon, J., Corona, F., Lira, F., Alcalde-Rico, M., Bernardini, A., Sanchez, M. and Martinez, J. (2016) 'Bacterial Multidrug Efflux Pumps: Much More Than Antibiotic Resistance Determinants', *Microorganisms*. doi: 10.3390/microorganisms4010014.

Blattner, F. R. *et al.* (1997) 'The complete genome sequence of *Escherichia coli* K-12', *Science*. doi: 10.1126/science.277.5331.1453.

Blin, K., Medema, M. H., Kottmann, R., Lee, S. Y. and Weber, T. (2016) 'The antiSMASH database, a comprehensive database of microbial secondary metabolite biosynthetic gene clusters', *Nucleic Acids Research*. doi: 10.1093/nar/gkw960.

Bolger, A. M., Lohse, M. and Usadel, B. (2014) 'Trimmomatic: A flexible trimmer for Illumina sequence data', *Bioinformatics*. doi: 10.1093/bioinformatics/btu170.

Bonacorsi, S. and Bingen, E. (2005) 'Molecular epidemiology of *Escherichia coli* causing neonatal meningitis', *International Journal of Medical Microbiology*. doi: 10.1016/j.ijmm.2005.07.011.

Bosi, E., Donati, B., Galardini, M., Brunetti, S., Sagot, M. F., Lió, P., Crescenzi, P., Fani, R. and Fondi, M. (2015) 'MeDuSa: A multi-draft based scaffolder', *Bioinformatics*, 31(15), pp. 2443–2451. doi: 10.1093/bioinformatics/btv171.

Braesel, J., Tran, T. A. and Eustáquio, A. S. (2019) 'Heterologous expression of the

diazaquinomycin biosynthetic gene cluster', *Journal of Industrial Microbiology and Biotechnology*. doi: 10.1007/s10295-019-02187-1.

Burkhart, B. J., Schwalen, C. J., Mann, G., Naismith, J. H. and Mitchell, D. A. (2017) 'YcaO-Dependent Posttranslational Amide Activation: Biosynthesis, Structure, and Function', *Chemical Reviews*. doi: 10.1021/acs.chemrev.6b00623.

Bush, M. J., Tschowri, N., Schlimpert, S., Flärdh, K. and Buttner, M. J. (2015) 'c-di-GMP signalling and the regulation of developmental transitions in streptomycetes', *Nature Reviews Microbiology*, 13(12), pp. 749–760. doi: 10.1038/nrmicro3546.

Calvert, M. B., Jumde, V. R. and Titz, A. (2018) 'Pathoblockers or antivirulence drugs as a new option for the treatment of bacterial infections', *Beilstein Journal of Organic Chemistry*, pp. 2607–2617. doi: 10.3762/bjoc.14.239.

Campellone, K. G., Robbins, D. and Leong, J. M. (2004) 'EspFU is a translocated EHEC effector that interacts with Tir and N-WASP and promotes Nck-independent actin assembly', *Developmental Cell*, 7(2), pp. 217–228. doi: 10.1016/j.devcel.2004.07.004.

Carver, T. J., Rutherford, K. M., Berriman, M., Rajandream, M. A., Barrell, B. G. and Parkhill, J. (2005) 'ACT: The Artemis comparison tool', *Bioinformatics*. doi: 10.1093/bioinformatics/bti553.

CDC (2018) 'Multistate Outbreak of *E. coli* O157:H7 Infections Linked to Romaine Lettuce (Final Update)', *2018 Outbreaks*.

Cegelski, L., Marshall, G. R., Eldridge, G. R. and Hultgren, S. J. (2008) 'The biology and future prospects of antivirulence therapies', *Nature Reviews Microbiology*, pp. 17–27. doi: 10.1038/nrmicro1818.

Chakraburtty, R. and Bibb, M. (1997) 'The ppGpp synthetase gene (relA) of Streptomyces coelicolor A3(2) plays a conditional role in antibiotic production and morphological differentiation', *Journal of Bacteriology*. doi: 10.1128/jb.179.18.5854-5861.1997.

Champness, W. C. (1988) ‘New loci required for Streptomyces coelicolor morphological and physiological differentiation.’, *Journal of bacteriology*. doi: 10.1128/jb.170.3.1168-1174.1988.

Chase-Topping, M. E. *et al.* (2007) ‘Risk factors for the presence of high-level shedders of Escherichia coli O157 on Scottish farms’, *Journal of Clinical Microbiology*. doi: 10.1128/JCM.01690-06.

Chater, K. F. (1972) ‘A morphological and genetic mapping study of white colony mutants of Streptomyces coelicolor.’, *Journal of general microbiology*. doi: 10.1099/00221287-72-1-9.

Chater, K. F. and Chandra, G. (2006) ‘The evolution of development in Streptomyces analysed by genome comparisons’, *FEMS Microbiology Reviews*. doi: 10.1111/j.1574-6976.2006.00033.x.

Chen, C. W., Huang, C. H., Lee, H. H., Tsai, H. H. and Kirby, R. (2002) ‘Once the circle has been broken: Dynamics and evolution of Streptomyces chromosomes’, *Trends in Genetics*. doi: 10.1016/S0168-9525(02)02752-X.

Chevrette, M. G., Gutiérrez-García, K., Selem-Mojica, N., Aguilar-Martínez, C., Yañez-Olvera, A., Ramos-Aboites, H. E., Hoskisson, P. A. and Barona-Gómez, F. (2020) ‘Evolutionary dynamics of natural product biosynthesis in bacteria’, *Natural Product Reports*. doi: 10.1039/c9np00048h.

Chng, C., Lum, A. M., Vroom, J. A. and Kao, C. M. (2008) ‘A key developmental regulator controls the synthesis of the antibiotic erythromycin in Saccharopolyspora erythraea’, *Proceedings of the National Academy of Sciences of the United States of America*. doi: 10.1073/pnas.0803622105.

Claessen, D., Rink, R., De Jong, W., Siebring, J., De Vreugd, P., Boersma, F. G. H., Dijkhuizen, L. and Wösten, H. A. B. (2003) ‘A novel class of secreted hydrophobic proteins is involved in aerial hyphae formation in Streptomyces coelicolor by forming amyloid-like fibrils’, *Genes and Development*, 17(14), pp. 1714–1726. doi: 10.1101/gad.264303.

Clements, A., Young, J. C., Constantinou, N. and Frankel, G. (2012) 'Infection strategies of enteric pathogenic *Escherichia coli*', *Gut Microbes*. doi: 10.4161/gmic.19182.

Cornelis, G. R. (2006) 'The type III secretion injectisome.', *Nature reviews. Microbiology*, 4(11), pp. 811–825. doi: 10.1038/nrmicro1526.

Corre, C., Haynes, S. W., Malet, N., Song, L. and Challis, G. L. (2010) 'A butenolide intermediate in methylenomycin furan biosynthesis is implied by incorporation of stereospecifically 13C-labelled glycerols', *Chemical Communications*. doi: 10.1039/c000496k.

Costa, T. R. D., Felisberto-Rodrigues, C., Meir, A., Prevost, M. S., Redzej, A., Trokter, M. and Waksman, G. (2015) 'Secretion systems in Gram-negative bacteria: Structural and mechanistic insights', *Nature Reviews Microbiology*. doi: 10.1038/nrmicro3456.

Crone, W. J. K., Leeper, F. J. and Truman, A. W. (2012) 'Identification and characterisation of the gene cluster for the anti-MRSA antibiotic bottromycin: Expanding the biosynthetic diversity of ribosomal peptides', *Chemical Science*. doi: 10.1039/c2sc21190d.

Cuthbertson, L. and Nodwell, J. R. (2013) 'The TetR Family of Regulators', *Microbiology and Molecular Biology Reviews*. doi: 10.1128/mmbr.00018-13.

Datsenko, K. A. and Wanner, B. L. (2000) 'One-step inactivation of chromosomal genes in.pdf', *Proceedings of the National Academy of Sciences*, 97(12), pp. 6640–6645. doi: 10.1073/pnas.120163297.

Davies, J. (2013) 'Specialized microbial metabolites: Functions and origins', *Journal of Antibiotics*. doi: 10.1038/ja.2013.61.

Deng, W. *et al.* (2004) 'Dissecting virulence: Systematic and functional analyses of a pathogenicity island', *Proceedings of the National Academy of Sciences of the United States of America*, 101(10), pp. 3597–3602. doi: 10.1073/pnas.0400326101.

Deshpande, B. S., Ambedkar, S. S. and Shewale, J. G. (1988) 'Biologically active secondary

metabolites from Streptomyces', *Enzyme and Microbial Technology*, pp. 455–473. doi: 10.1016/0141-0229(88)90023-3.

Devine, R., Qin, Z., Wilkinson, B. and Hutchings, M. (2019) 'Discovering novel antimicrobials from Streptomyces formicae, a symbiont of fungus farming plant ants, using CRISPR/Cas9 genome editing', *Access Microbiology*. doi: 10.1099/acmi.ac2019.po0067.

Donnenberg, M. S. (2013) *Escherichia coli: Pathotypes and Principles of Pathogenesis: Second Edition*, *Escherichia coli: Pathotypes and Principles of Pathogenesis: Second Edition*. doi: 10.1016/C2011-0-07692-5.

Doroghazi, J. R. and Buckley, D. H. (2010) 'Widespread homologous recombination within and between Streptomyces species', *ISME Journal*. doi: 10.1038/ismej.2010.45.

Drugbank (2020).

Du, L., Liu, R. H., Ying, L. and Zhao, G. R. (2012) 'An efficient intergeneric conjugation of DNA from *escherichia coli* to mycelia of the lincomycin-producer *Streptomyces lincolnensis*', *International Journal of Molecular Sciences*, 13(4), pp. 4797–4806. doi: 10.3390/ijms13044797.

Dunne, K. A. *et al.* (2017) 'Sequencing a piece of history: Complete genome sequence of the original *Escherichia coli* strain', *Microbial Genomics*. doi: 10.1099/mgen.0.000106.

Elliot, M. A., Bibb, M. J., Buttner, M. J. and Leskiw, B. K. (2001) 'BldD is a direct regulator of key developmental genes in Streptomyces coelicolor A3(2)', *Molecular Microbiology*. doi: 10.1046/j.1365-2958.2001.02387.x.

Elliot, M. A., Karoonuthaisiri, N., Huang, J., Bibb, M. J., Cohen, S. N., Kao, C. M. and Buttner, M. J. (2003) 'The chaplins: A family of hydrophobic cell-surface proteins involved in aerial mycelium formation in Streptomyces coelicolor', *Genes and Development*, 17(14), pp. 1727–1740. doi: 10.1101/gad.264403.

Elliot, M., Damji, F., Passantino, R., Chater, K. and Leskiw, B. (1998) 'The bldD gene of

Streptomyces coelicolor A3(2): A regulatory gene involved in morphogenesis and antibiotic production', *Journal of Bacteriology*. doi: 10.1128/jb.180.6.1549-1555.1998.

Famelis, N. *et al.* (2019) 'Architecture of the mycobacterial type VII secretion system', *Nature*. doi: 10.1038/s41586-019-1633-1.

Ferens, W. A. and Hovde, C. J. (2011) '*Escherichia coli* O157:H7: Animal reservoir and sources of human infection', *Foodborne Pathogens and Disease*. doi: 10.1089/fpd.2010.0673.

Fernández De Henestrosa, A. R., Ogi, T., Aoyagi, S., Chafin, D., Hayes, J. J., Ohmori, H. and Woodgate, R. (2000) 'Identification of additional genes belonging to the LexA regulon in *Escherichia coli*', *Molecular Microbiology*, 35(6), pp. 1560–1572. doi: 10.1046/j.1365-2958.2000.01826.x.

Figueras, M. J., Beaz-Hidalgo, R., Hossain, M. J. and Liles, M. R. (2014) 'Taxonomic affiliation of new genomes should be verified using average nucleotide identity and multilocus phylogenetic analysis', *Genome Announcements*. doi: 10.1128/genomeA.00927-14.

Fischbach, M. A., Walsh, C. T. and Clardy, J. (2008) 'The evolution of gene collectives: How natural selection drives chemical innovation', *Proceedings of the National Academy of Sciences of the United States of America*. doi: 10.1073/pnas.0709132105.

Flärdh, K. and Buttner, M. J. (2009) 'Streptomyces morphogenetics: dissecting differentiation in a filamentous bacterium.', *Nature Reviews Microbiology*, 7(1), pp. 36–49. doi: 10.1038/nrmicro1968.

Fordtran, J. S. (2006) ' Colitis Due to *Clostridium difficile* Toxins: Underdiagnosed, Highly Virulent, and Nosocomial ', *Baylor University Medical Center Proceedings*, 19(1), pp. 3–12. doi: 10.1080/08998280.2006.11928114.

Fratamico, P. M., DebRoy, C., Liu, Y., Needleman, D. S., Baranzoni, G. M. and Feng, P.

(2016) ‘Advances in molecular serotyping and subtyping of Escherichia coli’, *Frontiers in Microbiology*. doi: 10.3389/fmicb.2016.00644.

Friedmann, H. C. (2006) ‘Escherich and Escherichia’, *Advances in Applied Microbiology*. doi: 10.1016/S0065-2164(06)60005-1.

Fuchs, S., Mühldorfer, I., Donohue-Rolfe, A., Kerényi, M., Emödy, L., Alexiev, R., Nenkov, P. and Hacker, J. (1999) ‘Influence of RecA on in vivo virulence and Shiga toxin 2 production in *Escherichia coli* pathogens’, *Microbial Pathogenesis*, 27(1), pp. 13–23. doi: 10.1006/mpat.1999.0279.

Gauthier, A., Robertson, M. L., Lowden, M., Ibarra, J. A., Puente, L. and Finlay, B. B. (2005) ‘Transcriptional Inhibitor of Virulence Factors in Enteropathogenic *Escherichia coli’*, 49(10), pp. 4101–4109. doi: 10.1128/AAC.49.10.4101.

Gerlach, R. G. and Hensel, M. (2007) ‘Protein secretion systems and adhesins: The molecular armory of Gram-negative pathogens’, *International Journal of Medical Microbiology*. doi: 10.1016/j.ijmm.2007.03.017.

Gibson, D. G. (2011) ‘Enzymatic assembly of overlapping DNA fragments’, in *Methods in Enzymology*. doi: 10.1016/B978-0-12-385120-8.00015-2.

Girard, G., Traag, B. A., Sangal, V., Mascini, N., Hoskisson, P. A., Goodfellow, M. and van Wezel, G. P. (2013) ‘A novel taxonomic marker that discriminates between morphologically complex actinomycetes.’, *Open biology*. doi: 10.1098/rsob.130073.

Goldwater, P. N. and Bettelheim, K. a (2012) ‘Treatment of enterohemorrhagic *Escherichia co*li (EHEC) infection and hemolytic uremic syndrome (HUS)’, *BMC Medicine*, 10(1), p. 12. doi: 10.1186/1741-7015-10-12.

Gomez-Escribano, J. P., Alt, S. and Bibb, M. J. (2016) ‘Next generation sequencing of actinobacteria for the discovery of novel natural products’, *Marine Drugs*. doi: 10.3390/md14040078.

Gomez-Escribano, J. P. and Bibb, M. J. (2012) 'Streptomyces coelicolor as an expression host for heterologous gene clusters', in *Methods in Enzymology*, pp. 279–300. doi: 10.1016/B978-0-12-404634-4.00014-0.

Gomez-Escribano, J. P., Castro, J. F., Razmilic, V., Chandra, G., Andrews, B., Asenjo, J. A. and Bibb, M. J. (2015) 'The Streptomyces leeuwenhoekii genome: De novo sequencing and assembly in single contigs of the chromosome, circular plasmid pSLE1 and linear plasmid pSLE2', *BMC Genomics*. doi: 10.1186/s12864-015-1652-8.

Goosney, D. L., Gruenheid, S. and Finlay, B. B. (2000) 'Gut feelings: enteropathogenic *E. coli* (EPEC) interactions with the host.', *Annual review of cell and developmental biology*, 16, pp. 173–189. doi: 10.1146/annurev.cellbio.16.1.173.

Grant, J. R., Arantes, A. S. and Stothard, P. (2012) 'Comparing thousands of circular genomes using the CGView Comparison Tool', *BMC Genomics*. doi: 10.1186/1471-2164-13-202.

Grant, S. G. N., Jesseet, J., Bloomt, F. R. and Hanahan, D. (1990) 'Differential plasmid rescue from transgenic mouse DNAs into *Escherichia coli* methylation-restriction mutants (bacterial restriction/DNA methylation/cloning mammalian DNA/heterogeneous transgene expression/insulin gene regulation)', *Genetics*, 87, pp. 4645–4649.

Gregory, M. A., Till, R. and Smith, M. C. M. (2003) 'Integration site for Streptomyces phage φBT1 and development of site-specific integrating vectors', *Journal of Bacteriology*, 185(17), pp. 5320–5323. doi: 10.1128/JB.185.17.5320-5323.2003.

Gui, C., Li, Q., Mo, X., Qin, X., Ma, J. and Ju, J. (2015) 'Discovery of a new family of dieckmann cyclases essential to tetramic acid and pyridone-based natural products biosynthesis', *Organic Letters*. doi: 10.1021/ol5036497.

Gullón, S., Olano, C., Abdelfattah, M. S., Braña, A. F., Rohr, J., Méndez, C. and Salas, J. A. (2006) 'Isolation, characterization, and heterologous expression of the biosynthesis gene cluster for the antitumor anthracycline steffimycin', *Applied and Environmental Microbiology*.

doi: 10.1128/AEM.00734-06.

Gurevich, A., Saveliev, V., Vyahhi, N. and Tesler, G. (2013) 'QUAST: Quality assessment tool for genome assemblies', *Bioinformatics*. doi: 10.1093/bioinformatics/btt086.

Gust, B., Challis, G. L., Fowler, K., Kieser, T. and Chater, K. F. (2003) 'PCR-targeted Streptomyces gene replacement identifies a protein domain needed for biosynthesis of the sesquiterpene soil odor geosmin', *Proceedings of the National Academy of Sciences of the United States of America*, 100(4), pp. 1541–1546. doi: 10.1073/pnas.0337542100.

Hao, C., Huang, S., Deng, Z., Zhao, C. and Yu, Y. (2014) 'Mining of the pyrrolamide antibiotics analogs in Streptomyces netropsis reveals the amidohydrolase-dependent "iterative strategy" underlying the pyrrole polymerization', *PLoS ONE*, 9(6). doi: 10.1371/journal.pone.0099077.

Hashimoto, T., Hashimoto, J., Kozone, I., Amagai, K., Kawahara, T., Takahashi, S., Ikeda, H. and Shin-Ya, K. (2018) 'Biosynthesis of Quinolidomicin, the Largest Known Macrolide of Terrestrial Origin: Identification and Heterologous Expression of a Biosynthetic Gene Cluster over 200 kb', *Organic Letters*. doi: 10.1021/acs.orglett.8b03570.

Hempel, A. M. *et al.* (2012) 'The Ser/Thr protein kinase AfsK regulates polar growth and hyphal branching in the filamentous bacteria Streptomyces', *Proceedings of the National Academy of Sciences of the United States of America*. doi: 10.1073/pnas.1207409109.

Herbert, L. J. *et al.* (2014) '*E. coli* O157 on Scottish cattle farms: Evidence of local spread and persistence using repeat cross-sectional data', *BMC Veterinary Research*. doi: 10.1186/1746-6148-10-95.

Hernandes, R. T., Elias, W. P., Vieira, M. A. M. and Gomes, T. A. T. (2009) 'An overview of atypical enteropathogenic *Escherichia coli*', *FEMS Microbiology Letters*. doi: 10.1111/j.1574-6968.2009.01664.x.

Hobbs, G. *et al.* (1992) 'An integrated approach to studying regulation of production of the

antibiotic methylenomycin by Streptomyces coelicolor A3(2)', *Journal of Bacteriology*. doi: 10.1128/jb.174.5.1487-1494.1992.

Hobbs, G., Frazer, C. M., Gardner, D. C. J., Cullum, J. A. and Oliver, S. G. (1989) 'Dispersed growth of Streptomyces in liquid culture', *Applied Microbiology and Biotechnology*. doi: 10.1007/BF00258408.

Hodgson, D. A. (2000) 'Primary metabolism and its control in Streptomycetes: A most unusual group of bacteria', *Advances in Microbial Physiology*. doi: 10.1016/s0065-2911(00)42003-5.

Homma, T., Nuxoll, A., Brown Gandt, A., Ebner, P., Engels, I., Schneider, T., Götz, F., Lewis, K. and Conlon, B. P. (2016) 'Dual targeting of cell wall precursors by teixobactin leads to cell lysis', *Antimicrobial Agents and Chemotherapy*, (August), p. AAC.01050-16. doi: 10.1128/AAC.01050-16.

Hopwood, D. A. (1967) 'Genetic analysis and genome structure in Streptomyces coelicolor.', *Bacteriological reviews*. doi: 10.1128/mmbr.31.4.373-403.1967.

Hopwood, D. A. (2006) ' Soil To Genomics: The Streptomyces Chromosome ', *Annual Review of Genetics*. doi: 10.1146/annurev.genet.40.110405.090639.

Hoskisson, P. A., Hobbs, G. and Sharples, G. P. (2001) 'Antibiotic production, accumulation of intracellular carbon reserves, and sporulation in Micromonospora echinospora (ATCC 15837)', *Canadian Journal of Microbiology*. doi: 10.1139/cjm-47-2-148.

Hudson, D. L., Layton, A. N., Field, T. R., Bowen, A. J., Wolf-Watz, H., Elofsson, M., Stevens, M. P. and Galyov, E. E. (2007) 'Inhibition of type III secretion in Salmonella enterica serovar typhimurium by small-molecule inhibitors', *Antimicrobial Agents and Chemotherapy*, 51(7), pp. 2631–2635. doi: 10.1128/AAC.01492-06.

Hueck, C. J. (1998) 'Type III Protein Secretion Systems in Bacterial Pathogens of Animals and Plants', *Microbiology and Molecular Biology Reviews*, 62(2), pp. 379–433.

Huerta-Uribe, A. *et al.* (2016) 'Identification and characterization of novel compounds blocking Shiga toxin expression in *Escherichia coli* O157:H7', *Frontiers in Microbiology*, 7(NOV), pp. 1–9. doi: 10.3389/fmicb.2016.01930.

Huo, L., Hug, J. J., Fu, C., Bian, X., Zhang, Y. and Müller, R. (2019) 'Heterologous expression of bacterial natural product biosynthetic pathways', *Natural Product Reports*. doi: 10.1039/c8np00091c.

Hwang, K. S., Kim, H. U., Charusanti, P., Palsson, B. T. and Lee, S. Y. (2014) 'Systems biology and biotechnology of Streptomyces species for the production of secondary metabolites', *Biotechnology Advances*. doi: 10.1016/j.biotechadv.2013.10.008.

Iftime, D., Jasyk, M., Kulik, A., Imhoff, J. F., Stegmann, E., Wohlleben, W., Süssmuth, R. D. and Weber, T. (2015) 'Streptocollin, a Type IV Lanthipeptide Produced by Streptomyces collinus Tü 365', *ChemBioChem*. doi: 10.1002/cbic.201500377.

Iguchi, A. *et al.* (2009) 'Complete genome sequence and comparative genome analysis of enteropathogenic *Escherichia col*i O127:H6 strain E2348/69', *Journal of Bacteriology*. doi: 10.1128/JB.01238-08.

Ikeda, H., Shin-Ya, K. and Omura, S. (2014) 'Genome mining of the Streptomyces avermitilis genome and development of genome-minimized hosts for heterologous expression of biosynthetic gene clusters', *Journal of Industrial Microbiology and Biotechnology*, pp. 233–250. doi: 10.1007/s10295-013-1327-x.

Imamovic, L. and Muniesa, M. (2012) 'Characterizing RecA-independent induction of Shiga toxin2-encoding phages by EDTA treatment.', *PloS one*, 7(2), p. e32393. doi: 10.1371/journal.pone.0032393.

Ishikawa, J. and Hotta, K. (1999) 'FramePlot: a new implementation of the Frame analysis for predicting protein-coding regions in bacterial DNA with a high G+C content', *FEMS Microbiology Letters*. doi: 10.1111/j.1574-6968.1999.tb13576.x.

Iyoda, S., Koizumi, N., Satou, H., Lu, Y., Saitoh, T., Ohnishi, M. and Watanabe, H. (2006) 'The GrlR-GrlA regulatory system coordinately controls the expression of flagellar and LEE-encoded type III protein secretion systems in enterohemorrhagic *Escherichia col*i.', *Journal of bacteriology*, 188(16), pp. 5682–5692. doi: 10.1128/JB.00352-06.

J., R. A. *et al.* (2004) 'Co-ordinate single-cell expression of LEE4- and LEE5-encoded proteins of *Escherichia coli* O157:H7', *Molecular Microbiology*. Wiley/Blackwell (10.1111), 54(2), pp. 337–352. doi: 10.1111/j.1365-2958.2004.04277.x.

Jackson, N., Czaplewski, L. and Piddock, L. J. V. (2018) 'Discovery and development of new antibacterial drugs: Learning from experience?', *Journal of Antimicrobial Chemotherapy*, pp. 1452–1459. doi: 10.1093/jac/dky019.

Jain, M., Fiddes, I. T., Miga, K. H., Olsen, H. E., Paten, B. and Akeson, M. (2015) 'Improved data analysis for the MinION nanopore sequencer', *Nature Methods*, 12(4), pp. 351–356. doi: 10.1038/nmeth.3290.

Jarvis, K. G., Girón, J. a, Jerse, a E., McDaniel, T. K., Donnenberg, M. S. and Kaper, J. B. (1995) 'Enteropathogenic *Escherichia col*i contains a putative type III secretion system necessary for the export of proteins involved in attaching and effacing lesion formation.', *Proceedings of the National Academy of Sciences of the United States of America*, 92(August), pp. 7996–8000. doi: 10.1073/pnas.92.17.7996.

Jensen, K. *et al.* (2012) 'Polyketide proofreading by an acyltransferase-like enzyme', *Chemistry and Biology*. doi: 10.1016/j.chembiol.2012.01.005.

Jensen, P. R. and Fenical, W. (2005) 'New natural-product diversity from marine actinomycetes', in *Natural Products: Drug Discovery and Therapeutic Medicine*. doi: 10.1007/978-1-59259-976-9_14.

Jones, A. C., Gust, B., Kulik, A., Heide, L., Buttner, M. J. and Bibb, M. J. (2013) 'Phage P1-Derived Artificial Chromosomes Facilitate Heterologous Expression of the FK506 Gene Cluster', *PLoS ONE*. doi: 10.1371/journal.pone.0069319.

Kaper, J. B., Nataro, J. P. and Mobley, H. L. (2004) 'Pathogenic Escherichia coli.', *Nature reviews. Microbiology*, 2(2), pp. 123–140. doi: 10.1038/nrmicro818.

Karmali, M. A. (1989) 'Infection by verocytotoxin-producing Escherichia coli', *Clinical Microbiology Reviews*. doi: 10.1128/CMR.2.1.15.

Karmali, M. A., Petric, M., Steele, B. T. and Lim, C. (1983) 'SPORADIC CASES OF HAEMOLYTIC-URAEMIC SYNDROME ASSOCIATED WITH FAECAL CYTOTOXIN AND CYTOTOXIN-PRODUCING ESCHERICHIA COLI IN STOOLS', *The Lancet*. doi: 10.1016/S0140-6736(83)91795-6.

Karp, P. D. *et al.* (2007) 'Multidimensional annotation of the Escherichia coli K-12 genome', *Nucleic Acids Research*. doi: 10.1093/nar/gkm740.

Katz, L. and Baltz, R. H. (2016) 'Natural product discovery: past, present, and future', *Journal of Industrial Microbiology {&} Biotechnology*, 43(2), pp. 155–176. doi: 10.1007/s10295-015-1723-5.

Kieser, T., Bibb, M. J., Buttner, M. J., Chater, K. F. and Hopwood, D. A. (2000a) 'Practical Streptomyces Genetics', *John Innes Centre Ltd.*, p. 529. doi: 10.4016/28481.01.

Kieser, T., Bibb, M. J., Buttner, M. J., Chater, K. F. and Hopwood, D. A. (2000b) 'Practical Streptomyces Genetics', *John Innes Centre Ltd.*, p. 529. doi: 10.4016/28481.01.

Kieser, T., Hopwood, D. A., Wright, H. M. and Thompson, C. J. (1982) 'pIJ101, a multi-copy broad host-range Streptomyces plasmid: Functional analysis and development of DNA cloning vectors', *MGG Molecular & General Genetics*. doi: 10.1007/BF00330791.

Kim, B. and Little, J. W. (1993) 'LexA and lambda Cl repressors as enzymes: specific cleavage in an intermolecular reaction.', *Cell*, 73(6), pp. 1165–1173.

Kim, S. Y., Ju, K. S., Metcalf, W. W., Evans, B. S., Kuzuyama, T. and Van Der Donk, W. A. (2012) 'Different biosynthetic pathways to fosfomycin in Pseudomonas syringae and Streptomyces species', *Antimicrobial Agents and Chemotherapy*. doi: 10.1128/AAC.06478-

11.

Kimura, Kyota, Iwatsuki, M., Nagai, T., Matsumoto, A., Takahashi, Y., Shiomi, K., Omura, S. and Abe, A. (2011) 'A small-molecule inhibitor of the bacterial type III secretion system protects against in vivo infection with Citrobacter rodentium.', *The Journal of antibiotics*, 64(2), pp. 197–203. doi: 10.1038/ja.2010.155.

Kimura, K, Iwatsuki, M., Nagai, T., Matsumoto, A., Takahashi, Y., Shiomi, K., Omura, S. and Abe, A. (2011) 'A small-molecule inhibitor of the bacterial type III secretion system protects against in vivo infection with *Citrobacter rodentium*', *The Journal of antibiotics*, 64(2), pp. 197–203. doi: 10.1038/ja.2010.155.

Kinashi, H. (2011) 'Giant linear plasmids in Streptomyces: A treasure trove of antibiotic biosynthetic clusters', *Journal of Antibiotics*. doi: 10.1038/ja.2010.146.

Kitov, P. I., Sadowska, J. M., Mulvey, G., Armstrong, G. D., Ling, H., Pannu, N. S., Read, R. J. and Bundle, D. R. (2000) 'Shiga-like toxins are neutralized by tailored multivalent carbohydrate ligands', *Nature*, 403(6770), pp. 669–672. doi: 10.1038/35001095.

Kittilä, T. *et al.* (2017) 'Halogenation of glycopeptide antibiotics occurs at the amino acid level during non-ribosomal peptide synthesis', *Chemical Science*. doi: 10.1039/c7sc00460e.

Kodani, S., Hudson, M. E., Durrant, M. C., Buttner, M. J., Nodwell, J. R. and Willey, J. M. (2004) 'The SapB morphogen is a lantibiotic-like peptide derived from the product of the developmental gene ramS in Streptomyces coelicolor', *Proceedings of the National Academy of Sciences of the United States of America*. doi: 10.1073/pnas.0404220101.

Kodani, S., Lodato, M. A., Durrant, M. C., Picart, F. and Willey, J. M. (2005) 'SapT, a lanthionine-containing peptide involved in aerial hyphae formation in the streptomycetes', *Molecular Microbiology*. doi: 10.1111/j.1365-2958.2005.04921.x.

Komatsu, M., Uchiyama, T., Omura, S., Cane, D. E. and Ikeda, H. (2010) 'Genome-minimized Streptomyces host for the heterologous expression of secondary metabolism.',

Proceedings of the National Academy of Sciences of the United States of America, 107(6), pp. 2646–51. doi: 10.1073/pnas.0914833107.

Konstantinidis, K. T., Ramette, A. and Tiedje, J. M. (2006) 'The bacterial species definition in the genomic era', in *Philosophical Transactions of the Royal Society B: Biological Sciences*. doi: 10.1098/rstb.2006.1920.

Koren, S., Walenz, B. P., Berlin, K., Miller, J. R. and Phillippy, A. M. (2016) 'Canu: scalable and accurate long-read assembly via adaptive k-mer weighting and repeat separation', *bioRxiv*, p. 071282. doi: 10.1101/071282.

Kotowska, M., Pawlik, K., Butler, A. R., Cundliffe, E., Takano, E. and Kuczek, K. (2002) 'Type II thioesterase from Streptomyces coelicolor A3(2)', *Microbiology*. doi: 10.1099/00221287-148-6-1777.

Kruasuwan, W., Salih, T. S., Brozio, S., Hoskisson, P. A. and Thamchaipenet, A. (2017) 'Draft genome sequence of plant growth-promoting endophytic Streptomyces sp. GKU 895 isolated from the roots of sugarcane', *Genome Announcements*. doi: 10.1128/genomeA.00358-17.

Kuehne, S. A., Cartman, S. T., Heap, J. T., Kelly, M. L., Cockayne, A. and Minton, N. P. (2010) 'The role of toxin A and toxin B in *Clostridium difficile* infection', *Nature*, 467(7316), pp. 711–713. doi: 10.1038/nature09397.

Kumar, S., Stecher, G. and Tamura, K. (2016) 'MEGA7: Molecular Evolutionary Genetics Analysis version 7.0 for bigger datasets.', *Molecular biology and evolution*, p. msw054. doi: 10.1093/molbev/msw054.

Kuščer, E., Coates, N., Challis, I., Gregory, M., Wilkinson, B., Sheridan, R. and Petković, H. (2007) 'Roles of rapH and rapG in positive regulation of rapamycin biosynthesis in Streptomyces hygroscopicus', *Journal of Bacteriology*. doi: 10.1128/JB.00129-07.

Labeda, D. P. *et al.* (2012) 'Phylogenetic study of the species within the family

Streptomycetaceae', in *Antonie van Leeuwenhoek, International Journal of General and Molecular Microbiology*. doi: 10.1007/s10482-011-9656-0.

Langdon, A., Crook, N. and Dantas, G. (2016) 'The effects of antibiotics on the microbiome throughout development and alternative approaches for therapeutic modulation', *Genome Medicine*. doi: 10.1186/s13073-016-0294-z.

Latoscha, A., Drexler, D. J., Al-Bassam, M. M., Bandera, A. M., Kaever, V., Findlay, K. C., Witte, G. and Tschowri, N. (2020) 'c-di-AMP hydrolysis by the phosphodiesterase AtaC promotes differentiation of multicellular bacteria', *Proceedings of the National Academy of Sciences*. doi: 10.1073/pnas.1917080117.

Laver, T., Harrison, J., O'Neill, P. A., Moore, K., Farbos, A., Paszkiewicz, K. and Studholme, D. J. (2015) 'Assessing the performance of the Oxford Nanopore Technologies MinION', *Biomolecular Detection and Quantification*, 3, pp. 1–8. doi: 10.1016/j.bdq.2015.02.001.

Lee, J. H., Kim, Y. G., Cho, H. S., Ryu, S. Y., Cho, M. H. and Lee, J. (2014) 'Coumarins reduce biofilm formation and the virulence of *Escherichia coli* O157:H7', *Phytomedicine*, 21(8–9), pp. 1037–1042. doi: 10.1016/j.phymed.2014.04.008.

Li, Q., Song, Y., Qin, X., Zhang, X., Sun, A. and Ju, J. (2015) 'Identification of the Biosynthetic Gene Cluster for the Anti-infective Desotamides and Production of a New Analogue in a Heterologous Host', *Journal of Natural Products*. doi: 10.1021/acs.jnatprod.5b00009.

Li, R. and Townsend, C. A. (2006) 'Rational strain improvement for enhanced clavulanic acid production by genetic engineering of the glycolytic pathway in Streptomyces clavuligerus', *Metabolic Engineering*. doi: 10.1016/j.ymben.2006.01.003.

Li, Y., Li, Z., Yamanaka, K., Xu, Y., Zhang, W., Vlamakis, H., Kolter, R., Moore, B. S. and Qian, P. Y. (2015) 'Directed natural product biosynthesis gene cluster capture and expression in the model bacterium Bacillus subtilis', *Scientific Reports*. doi: 10.1038/srep09383.

de Lima Procópio, R. E., da Silva, I. R., Martins, M. K., de Azevedo, J. L. and de Araújo, J. M. (2012) 'Antibiotics produced by Streptomyces', *Brazilian Journal of Infectious Diseases*, pp. 466–471. doi: 10.1016/j.bjid.2012.08.014.

Ling, L. L. *et al.* (2015) 'A new antibiotic kills pathogens without detectable resistance', *Nature*. doi: 10.1038/nature14098.

Linington, R. G., Robertson, M., Gauthier, A., Finlay, B. B., Van Soest, R. and Andersen, R. J. (2002) 'Caminoside A, an antimicrobial glycolipid isolated from the marine sponge Caminus sphaeroconia', *Organic Letters*. doi: 10.1021/ol0268337.

Liu, C. min, Hermann, T. and Miller, P. A. (1977) 'Feedback inhibition of the synthesis of an antibiotic: Aurodox (X-5108)', *The Journal of Antibiotics*. doi: 10.7164/antibiotics.30.244.

Liu, L. *et al.* (2012) 'Global, regional, and national causes of child mortality: An updated systematic analysis for 2010 with time trends since 2000', *The Lancet*. doi: 10.1016/S0140-6736(12)60560-1.

Liu, L., Johnson, H. and Cousens, S. (2012) 'Global, regional, and national causes of child mortality: an updated systematic analysis for 2010 with time trends since 2000', *Lancet, The*, 379(9832), pp. 2151–2161.

Lombó, F., Velasco, A., Castro, A., De La Calle, F., Braña, A. F., Sánchez-Puelles, J. M., Méndez, C. and Salas, J. A. (2006) 'Deciphering the biosynthesis pathway of the antitumor thiocoraline from a marine actinomycete and its expression in two Streptomyces species', *ChemBioChem*. doi: 10.1002/cbic.200500325.

Lyras, D. *et al.* (2009) 'Toxin B is essential for virulence of *Clostridium difficile'*, *Nature*, 458(7242), pp. 1176–1179. doi: 10.1038/nature07822.

Machado, H., Sonnenschein, Eva C, *et al.* (2015) 'Genome mining reveals unlocked bioactive potential of marine Gram-negative bacteria', *BMC Genomics*, 16(1), p. 158. doi: 10.1186/s12864-015-1365-z.

Machado, H., Sonnenschein, Eva C., Melchiorsen, J. and Gram, L. (2015) 'Genome mining reveals unlocked bioactive potential of marine Gram-negative bacteria', *BMC Genomics*, 16(1). doi: 10.1186/s12864-015-1365-z.

MacNeil, D. J., Gewain, K. M., Ruby, C. L., Dezeny, G., Gibbons, P. H. and MacNeil, T. (1992) 'Analysis of Streptomyces avermitilis genes required for avermectin biosynthesis utilizing a novel integration vector', *Gene*, 111(1), pp. 61–68. doi: 10.1016/0378-1119(92)90603-M.

Maehr, H., Leach, M., Williams, T. H. and Blount, J. F. (1980) 'The chemistry of aurodox and related antibiotics', *Canadian Journal of Chemistry*, 58, pp. 501–526. doi: 10.1139/v80-080.

Maehr, H., Leach, M., Yarmchuk, L. and Mitrivic, M. (1978) 'Chemical Conversion of Mocimycin to Aurodox and derivatives of Aurodox, Goldinamine and Mocimycin', *The Journal of antibiotics*, XXXII(4), pp. 361–367.

Maehr, H., Leach, M., Yarmchuk, L. and Mitrovic, M. (1979) 'Antibiotic X-5108. IX. Chemical conversion of mocimycin to aurodox and derivatives of aurodox, goldinamine and mocimycin', *J Antibiot (Tokyo)*, 32(4), pp. 361–367. Available at: http://www.ncbi.nlm.nih.gov/entrez/query.fcgi?cmd=Retrieve&db=PubMed&dopt=Citation&list_uids=381270.

Mallick, E. M. *et al.* (2012) 'A novel murine infection model for Shiga toxin-producing *Escherichia coli*', *Journal of Clinical Investigation*, 122(11), pp. 4012–4024. doi: 10.1172/JCI62746.

Mao, Y., Doyle, M. P. and Chen, J. (2006) 'Role of colanic acid exopolysaccharide in the survival of enterohaemorrhagic *Escherichia coli* O157:H7 in simulated gastrointestinal fluids', *Letters in Applied Microbiology*, 42(6), pp. 642–647. doi: 10.1111/j.1472-765X.2006.01875.x.

Maplestone, R. A., Stone, M. J. and Williams, D. H. (1992) 'The evolutionary role of secondary metabolites - a review', *Gene*. doi: 10.1016/0378-1119(92)90553-2.

Marinelli, F. (2009) 'Antibiotics and Streptomyces: The future of antibiotic discovery', *Microbiology Today*.

Martin, D. and Rybicki, E. (2000) 'RDP: Detection of recombination amongst aligned sequences', *Bioinformatics*. doi: 10.1093/bioinformatics/16.6.562.

McHugh, R. E., O'Boyle, N., Connolly, J. P. R., Hoskisson, P. A. and Roe, A. J. (2019) 'Characterization of the mode of action of aurodox, a type III secretion system inhibitor from *Streptomyces goldiniensis*', *Infection and Immunity*. doi: 10.1128/IAI.00595-18.

McKean, I. J. W., Hoskisson, P. A. and Burley, G. A. (2020) 'Biocatalytic Alkylation Cascades: Recent Advances and Future Opportunities for Late-Stage Functionalization', *ChemBioChem*.

McKean, I. J. W., Sadler, J. C., Cuetos, A., Frese, A., Humphreys, L. D., Grogan, G., Hoskisson, P. A. and Burley, G. A. (2019) 'S-Adenosyl Methionine Cofactor Modifications Enhance the Biocatalytic Repertoire of Small Molecule C-Alkylation', *Angewandte Chemie - International Edition*. doi: 10.1002/anie.201908681.

McLean, T. C., Lo, R., Tschowri, N., Hoskisson, P. A., Al Bassam, M. M., Hutchings, M. I. and Som, N. F. (2019) 'Sensing and responding to diverse extracellular signals: An updated analysis of the sensor kinases and response regulators of *Streptomyces* species', *Microbiology (United Kingdom)*. doi: 10.1099/mic.0.000817.

McLean, T. C., Wilkinson, B., Hutchings, M. I. and Devine, R. (2019) 'Dissolution of the disparate: Co-ordinate regulation in antibiotic biosynthesis', *Antibiotics*. doi: 10.3390/antibiotics8020083.

Medema, M. H. *et al.* (2015) 'Minimum Information about a Biosynthetic Gene cluster', *Nature Chemical Biology*. doi: 10.1038/nchembio.1890.

van der Meij, A., Worsley, S. F., Hutchings, M. I. and van Wezel, G. P. (2017) 'Chemical ecology of antibiotic production by actinomycetes', *FEMS Microbiology Reviews*. doi:

10.1093/femsre/fux005.

Mellmann, A. *et al.* (2011) 'Prospective genomic characterization of the german enterohemorrhagic *Escherichia coli* O104:H4 outbreak by rapid next generation sequencing technology', *PLoS ONE*. doi: 10.1371/journal.pone.0022751.

Mikheyev, A. S. and Tin, M. M. Y. (2014) 'A first look at the Oxford Nanopore MinION sequencer', *Molecular Ecology Resources*, 14(6), pp. 1097–1102. doi: 10.1111/1755-0998.12324.

Mills, E., Baruch, K., Charpentier, X., Kobi, S. and Rosenshine, I. (2008) 'Real-Time Analysis of Effector Translocation by the Type III Secretion System of Enteropathogenic *Escherichia coli'*, *Cell Host & Microbe*, 3(2), pp. 104–113. doi: 10.1016/j.chom.2007.11.007.

Mirandela, G. D., Tamburrino, G., Hoskisson, P. A., Zachariae, U. and Javelle, A. (2019) ' The lipid environment determines the activity of the *Escherichia coli* ammonium transporter AmtB ', *The FASEB Journal*. doi: 10.1096/fj.201800782r.

Modi, S. R., Collins, J. J. and Relman, D. A. (2014) 'Antibiotics and the gut microbiota', *Journal of Clinical Investigation*, 124(10), pp. 4212–4218. doi: 10.1172/JCI72333.

Mühlen, S. and Dersch, P. (2016) 'Anti-virulence strategies to target bacterial infections', *Current Topics in Microbiology and Immunology*, 398, pp. 147–183. doi: 10.1007/82_2015_490.

Murakami, T., Holt, T. G. and Thompson, C. J. (1989) 'Thiostrepton-induced gene expression in Streptomyces lividans', *Journal of Bacteriology*. doi: 10.1128/jb.171.3.1459-1466.1989.

Musiol, E. M., Greule, A., Härtner, T., Kulik, A., Wohlleben, W. and Weber, T. (2013) 'The AT2 domain of KirCI loads malonyl extender units to the ACPs of the kirromycin PKS', *ChemBioChem*. doi: 10.1002/cbic.201300211.

Musiol, E. M., Härtner, T., Kulik, A., Moldenhauer, J., Piel, J., Wohlleben, W. and Weber, T.

(2011) ‘Supramolecular templating in kirromycin biosynthesis: The acyltransferase KirCII loads ethylmalonyl-CoA extender onto a specific ACP of the trans-AT PKS’, *Chemistry and Biology*. doi: 10.1016/j.chembiol.2011.02.007.

Naômé, A. *et al.* (2018) ‘Complete genome sequence of Streptomyces lunaelactis MM109T, isolated from cave moonmilk deposits’, *Genome Announcements*. doi: 10.1128/genomeA.00435-18.

Navalkele, B. D. and Chopra, T. (2018) ‘Bezlotoxumab: An emerging monoclonal antibody therapy for prevention of recurrent *Clostridium difficile* infection’, *Biologics: Targets and Therapy*, pp. 11–21. doi: 10.2147/BTT.S127099.

Navarro-Muñoz, J. *et al.* (2018) ‘A computational framework for systematic exploration of biosynthetic diversity from large-scale genomic data’, *bioRxiv*. doi: 10.1101/445270.

Navarro-Muñoz, J. C. *et al.* (2020) ‘A computational framework to explore large-scale biosynthetic diversity’, *Nature Chemical Biology*. doi: 10.1038/s41589-019-0400-9.

NETER, E., WESTPHAL, O., LUDERITZ, O., GINO, R. M. and GORZYNSKI, E. A. (1955) ‘Demonstration of antibodies against enteropathogenic *Escherichia coli* in sera of children of various ages.’, *Pediatrics*.

Nguyen, Y. and Sperandio, V. (2012) ‘Enterohemorrhagic *E. coli* (EHEC) pathogenesis.’, *Frontiers in cellular and infection microbiology*, 2(July), p. 90. doi: 10.3389/fcimb.2012.00090.

Nouioui, I. *et al.* (2018) ‘Genome-based taxonomic classification of the phylum actinobacteria’, *Frontiers in Microbiology*. doi: 10.3389/fmicb.2018.02007.

Ochoa, T. J. and Contreras, C. A. (2011) ‘Enteropathogenic *Escherichia coli i*nfection in children’, *Current Opinion in Infectious Diseases*. doi: 10.1097/QCO.0b013e32834a8b8b.

Ohnishi, Y., Seo, J. W. and Horinouchi, S. (2002) ‘Deprogrammed sporulation in Streptomyces’, *FEMS Microbiology Letters*, pp. 1–7. doi: 10.1016/S0378-1097(02)00996-5.

Okamoto, S., Taguchi, T., Ochi, K. and Ichinose, K. (2009) 'Biosynthesis of Actinorhodin and Related Antibiotics: Discovery of Alternative Routes for Quinone Formation Encoded in the act Gene Cluster', *Chemistry and Biology*. doi: 10.1016/j.chembiol.2009.01.015.

Olsthoorn-Tieleman, L. N., Palstra, R. J. T. S., Van Wezel, G. P., Bibb, M. J. and Pleij, C. W. A. (2007) 'Elongation factor Tu3 (EF-Tu3) from the kirromycin producer Streptomyces ramocissimus is resistant to three classes of EF-Tu-specific inhibitors', *Journal of Bacteriology*. doi: 10.1128/JB.01810-06.

Owen, J. G., Copp, J. N. and Ackerley, D. F. (2011) 'Rapid and flexible biochemical assays for evaluating 4′- phosphopantetheinyl transferase activity', *Biochemical Journal*. doi: 10.1042/BJ20110321.

Pacheco, A. R. and Sperandio, V. (2012) 'Shiga toxin in enterohemorrhagic *E.coli:* regulation and novel anti-virulence strategies.', *Frontiers in cellular and infection microbiology*, 2(June), p. 81. doi: 10.3389/fcimb.2012.00081.

Patzer, S. I. and Braun, V. (2010) 'Gene cluster involved in the biosynthesis of griseobactin, a catechol-peptide siderophore of Streptomyces sp. ATCC 700974', *Journal of Bacteriology*. doi: 10.1128/JB.01250-09.

Pavlidou, M., Pross, E. K., Musiol, E. M., Kulik, A., Wohlleben, W. and Weber, T. (2011) 'The phosphopantetheinyl transferase KirP activates the ACP and PCP domains of the kirromycin NRPS/PKS of Streptomyces collinus Tü 365', *FEMS Microbiology Letters*, pp. 26–33. doi: 10.1111/j.1574-6968.2011.02263.x.

Pendergrass, H. A. and May, A. E. (2019) 'Natural Product Type III Secretion System Inhibitors', *Antibiotics*, 8(4), p. 162. doi: 10.3390/antibiotics8040162.

Porter, J. N. (1971) 'Prevalence and Distribution of Antibiotic-Producing Actinomycetes', *Advances in Applied Microbiology*. doi: 10.1016/S0065-2164(08)70540-9.

Qin, Z., Munnoch, J. T., Devine, R., Holmes, N. A., Seipke, R. F., Wilkinson, K. A.,

Wilkinson, B. and Hutchings, M. I. (2017) 'Formicamycins, antibacterial polyketides produced by Streptomyces formicae isolated from African Tetraponera plant-ants', *Chemical Science*. doi: 10.1039/c6sc04265a.

Rausch, C., Weber, T., Kohlbacher, O., Wohlleben, W. and Huson, D. H. (2005) 'Specificity prediction of adenylation domains in nonribosomal peptide synthetases (NRPS) using transductive support vector machines (TSVMs)', *Nucleic Acids Research*. doi: 10.1093/nar/gki885.

Reddy, V. S., Shlykov, M. A., Castillo, R., Sun, E. I. and Saier, M. H. (2012) 'The major facilitator superfamily (MFS) revisited', *FEBS Journal*. doi: 10.1111/j.1742-4658.2012.08588.x.

Reid, R., Piagentini, M., Rodriguez, E., Ashley, G., Viswanathan, N., Carney, J., Santi, D. V., Richard Hutchinson, C. and McDaniel, R. (2003) 'A model of structure and catalysis for ketoreductase domains in modular polyketide synthases', *Biochemistry*. doi: 10.1021/bi0268706.

Rhoads, A. and Au, K. F. (2015) 'PacBio Sequencing and Its Applications', *Genomics, Proteomics and Bioinformatics*. doi: 10.1016/j.gpb.2015.08.002.

Richter, M. and Rosselló-Móra, R. (2009) 'Shifting the genomic gold standard for the prokaryotic species definition', *Proceedings of the National Academy of Sciences of the United States of America*. doi: 10.1073/pnas.0906412106.

Robertsen, H. L., Musiol-Kroll, E. M., Ding, L., Laiple, K. J., Hofeditz, T., Wohlleben, W., Lee, S. Y., Grond, S. and Weber, T. (2018) 'Filling the Gaps in the Kirromycin Biosynthesis: Deciphering the Role of Genes Involved in Ethylmalonyl-CoA Supply and Tailoring Reactions', *Scientific Reports*. doi: 10.1038/s41598-018-21507-6.

Roe, A. J., Yull, H., Naylor, S. W., Martin, J., Smith, D. G. E., Gally, D. L. and Woodward, M. J. (2003) 'Heterogeneous Surface Expression of EspA Translocon Filaments by *Escherichia coli* O157 : H7 Is Controlled at the Posttranscriptional Level', *Infection and immunity*, 71(10),

pp. 5900–5909. doi: 10.1128/IAI.71.10.5900.

Ronald , M., A. (2006) *Microbiological Media for the*, *Yeast*.

Rückert, C. *et al.* (2013) 'Complete genome sequence of the kirromycin producer Streptomyces collinus Tü 365 consisting of a linear chromosome and two linear plasmids', *Journal of Biotechnology*, 168(4), pp. 739–740. doi: 10.1016/j.jbiotec.2013.10.004.

Ryan, K. S. (2011) 'Biosynthetic gene cluster for the cladoniamides, bis-indoles with a rearranged scaffold', *PLoS ONE*. doi: 10.1371/journal.pone.0023694.

Saidijam, M. *et al.* (2006) 'Microbial Drug Efflux Proteins of the Major Facilitator Superfamily', *Current Drug Targets*. doi: 10.2174/138945006777709575.

Salerno, P., Persson, J., Bucca, G., Laing, E., Ausmees, N., Smith, C. P. and Flärdh, K. (2013) 'Identification of new developmentally regulated genes involved in Streptomyces coelicolor sporulation', *BMC Microbiology*. doi: 10.1186/1471-2180-13-281.

Sambrook, J., Fritsch, E. F. and Maniatis, T. (1989) *Molecular Cloning: A Laboratory Manual. Cold Spring Harbor laboratory press*, *New York*. doi: 574.873224 1/1989.

Sánchez, C., Butovich, I. A., Braña, A. F., Rohr, J., Méndez, C. and Salas, J. A. (2002) 'The Biosynthetic gene cluster for the antitumor rebeccamycin: Characterization and generation of indolocarbazole derivatives', *Chemistry and Biology*. doi: 10.1016/S1074-5521(02)00126-6.

Sangal, V., Goodfellow, M., Jones, A. L., Schwalbe, E. C., Blom, J., Hoskisson, P. A. and Sutcliffe, I. C. (2016) 'Next-generation systematics: An innovative approach to resolve the structure of complex prokaryotic taxa', *Scientific Reports*. doi: 10.1038/srep38392.

Schatz, A., Bugle, E. and Waksman, S. A. (1944) 'Streptomycin, a Substance Exhibiting Antibiotic Activity Against Gram-Positive and Gram-Negative Bacteria', *Proceedings of the Society for Experimental Biology and Medicine*. doi: 10.3181/00379727-55-14461.

Schmidt, M. A. (2010) 'LEEways: tales of EPEC, ATEC and EHEC.', *Cellular microbiology*,

12(11), pp. 1544–1552. doi: 10.1111/j.1462-5822.2010.01518.x.

Schniete, J. K., Cruz-Morales, P., Selem-Mojica, N., Fernández-Martínez, L. T., Hunter, I. S., Barona-Gómez, F. and Hoskisson, P. A. (2018) 'Expanding primary metabolism helps generate the metabolic robustness to facilitate antibiotic biosynthesis in Streptomyces', *mBio*. doi: 10.1128/mBio.02283-17.

Schwedock, J., McCormick, J. R., Angert, E. R., Nodwell, J. R. and Losick, R. (1997) ' Assembly of the cell division protein FtsZ into ladder-like structures in the aerial hyphae of Streptomyces coelicolor ', *Molecular Microbiology*. doi: 10.1111/j.1365-2958.1997.mmi507.x.

Seemann, T. (2014) 'Prokka: Rapid prokaryotic genome annotation', *Bioinformatics*. doi: 10.1093/bioinformatics/btu153.

Shirling, E. B. and Gottlieb, D. (1966) 'Methods for characterization of Streptomyces species', *International Journal of Systematic Bacteriology*. doi: 10.1099/00207713-16-3-313.

Sosio, M. (2001) 'Assembly of large genomic segments in artificial chromosomes by homologous recombination in *Escherichia coli*', *Nucleic Acids Research*, 29(7), pp. 37e – 37. doi: 10.1093/nar/29.7.e37.

Sperandio, V., Mellies, J. L., Delahay, R. M., Frankel, G., Adam Crawford, J., Nguyen, W. and Kaper, J. B. (2000a) 'Activation of enteropathogenic *Escherichia coli* (EPEC) LEE2 and LEE3 operons by Ler', *Molecular Microbiology*, 38(4), pp. 781–793. doi: 10.1046/j.1365-2958.2000.02168.x.

Sperandio, V., Mellies, J. L., Delahay, R. M., Frankel, G., Adam Crawford, J., Nguyen, W. and Kaper, J. B. (2000b) 'Activation of enteropathogenic *Escherichia coli* (EPEC) LEE2 and LEE3 operons by Ler', *Molecular Microbiology*. doi: 10.1046/j.1365-2958.2000.02168.x.

Stackebrandt, E. and Goebel, B. M. (1994) 'Taxonomic note: A place for DNA-DNA reassociation and 16S rRNA sequence analysis in the present species definition in bacteriology', *International Journal of Systematic Bacteriology*. doi: 10.1099/00207713-44-4-

846.

Sullivan, M. J., Petty, N. K. and Beatson, S. A. (2011) 'Easyfig: A genome comparison visualizer', *Bioinformatics*. doi: 10.1093/bioinformatics/btr039.

Tam, J., Hamza, T., Ma, B., Chen, K., Beilhartz, G. L., Ravel, J., Feng, H. and Melnyk, R. A. (2018) 'Host-targeted niclosamide inhibits C. difficile virulence and prevents disease in mice without disrupting the gut microbiota', *Nature Communications*, 9(1). doi: 10.1038/s41467-018-07705-w.

Tarr, P. I., Gordon, C. a and Chandler, W. L. (2005) 'Shiga-toxin-producing *Escherichia coli* and haemolytic uraemic syndrome.', *Lancet*, 365(9464), pp. 1073–1086. doi: 10.1016/S0140-6736(05)71144-2.

Thaker, M. N., García, M., Koteva, K., Waglechner, N., Sorensen, D., Medina, R. and Wright, G. D. (2012) 'Biosynthetic gene cluster and antimicrobial activity of the elfamycin antibiotic factumycin', *MedChemComm*. doi: 10.1039/c2md20038d.

Thaker, M., Spanogiannopoulos, P. and Wright, G. D. (2010) 'The tetracycline resistome', *Cellular and Molecular Life Sciences*, pp. 419–431. doi: 10.1007/s00018-009-0172-6.

Theobald, S. *et al.* (2018) 'Uncovering secondary metabolite evolution and biosynthesis using gene cluster networks and genetic dereplication', *Scientific Reports*. doi: 10.1038/s41598-018-36561-3.

Tidjani, A. R., Lorenzi, J. N., Toussaint, M., Van Dijk, E., Naquin, D., Lespinet, O., Bontemps, C. and Leblond, P. (2019) 'Massive gene flux drives genome diversity between sympatric *Streptomyces* conspecifics', *mBio*. doi: 10.1128/mBio.01533-19.

Totsika, M. (2016) 'Benefits and Challenges of Antivirulence Antimicrobials at the Dawn of the Post-Antibiotic Era', *Drug Delivery Letters*, 6(1), pp. 30–37. doi: 10.2174/2210303106666160506120057.

Tozzoli, R. and Scheutz, F. (2014) 'Diarrhoeagenic *Escherichia coli* Infections in Humans', in

Pathogenic Escherichia coli. doi: 10.2215/cjn.00550113.

Trimble, W. L., Keegan, K. P., D'Souza, M., Wilke, A., Wilkening, J., Gilbert, J. and Meyer, F. (2012) 'Short-read reading-frame predictors are not created equal: Sequence error causes loss of signal', *BMC Bioinformatics*. doi: 10.1186/1471-2105-13-183.

Tringe, S. G. and Hugenholtz, P. (2008) 'A renaissance for the pioneering 16S rRNA gene', *Current Opinion in Microbiology*, 11(5), pp. 442–446. doi: 10.1016/j.mib.2008.09.011.

Tschowri, N. (2016) 'Cyclic dinucleotide-controlled regulatory pathways in Streptomyces species', *Journal of Bacteriology*. doi: 10.1128/JB.00423-15.

Veenendaal, A. K. J., Sundin, C. and Blocker, A. J. (2009) 'Small-molecule type III secretion system inhibitors block assembly of the Shigella type III secreton.', *Journal of bacteriology*, 191(2), pp. 563–570. doi: 10.1128/JB.01004-08.

Vogeley, Lutz, Palm, G. J., Mesters, J. R. and Hilgenfeld, R. (2001) 'Conformational change of elongation factor Tu (EF-Tu) induced by antibiotic binding. Crystal structure of the complex between EF-Tu??GDP and aurodox', *Journal of Biological Chemistry*, 276(20), pp. 17149–17155. doi: 10.1074/jbc.M100017200.

Vogeley, L, Palm, G. J., Mesters, J. R. and Hilgenfeld, R. (2001) 'Conformational change of elongation factor Tu (EF-Tu) induced by antibiotic binding. Crystal structure of the complex between EF-Tu.GDP and aurodox.', *The Journal of biological chemistry*, 276(20), pp. 17149–17155. doi: 10.1074/jbc.M100017200.

Waksman, S. A. and Henrici, A. T. (1943) 'The Nomenclature and Classification of the Actinomycetes.', *Journal of bacteriology*. doi: 10.1128/jb.46.4.337-341.1943.

Wang, D. *et al.* (2011) 'Identification of bacterial target proteins for the salicylidene acylhydrazide class of virulence-blocking compounds.', *The Journal of biological chemistry*, 286(34), pp. 29922–29931. doi: 10.1074/jbc.M111.233858.

Wasteson, Y. (2001) 'Zoonotic *Escherichia col*i.', *Acta veterinaria Scandinavica.*

Supplementum.

Weber, T., Laiple, K. J., Pross, E. K., Textor, A., Grond, S., Welzel, K., Pelzer, S., Vente, A. and Wohlleben, W. (2008) 'Molecular Analysis of the Kirromycin Biosynthetic Gene Cluster Revealed β-Alanine as Precursor of the Pyridone Moiety', *Chemistry and Biology*, 15(2), pp. 175–188. doi: 10.1016/j.chembiol.2007.12.009.

van Wezel, G. P. and McDowall, K. J. (2011) 'The regulation of the secondary metabolism of *Streptomyces*: new links and experimental advances', *Natural Product Reports*, 28(7), p. 1311. doi: 10.1039/c1np00003a.

Van Wezel, G. P. and McDowall, K. J. (2011) 'The regulation of the secondary metabolism of *Streptomyces*: New links and experimental advances', *Natural Product Reports*. doi: 10.1039/c1np00003a.

Wilcox, M. H. *et al.* (2017) 'Bezlotoxumab for Prevention of Recurrent *Clostridium difficile* Infection', *New England Journal of Medicine*, 376(4), pp. 305–317. doi: 10.1056/NEJMoa1602615.

Wiles, T. J., Kulesus, R. R. and Mulvey, M. A. (2008) 'Origins and virulence mechanisms of uropathogenic *Escherichia coli*', *Experimental and Molecular Pathology*. doi: 10.1016/j.yexmp.2008.03.007.

Willey, J. M. and Gaskell, A. A. (2011) 'Morphogenetic signaling molecules of the streptomycetes', *Chemical Reviews*. doi: 10.1021/cr1000404.

Woese, C. R., Blanz, P. and Hahn, C. M. (1984) 'What isn't a Pseudomonad: The Importance of Nomenclature in Bacterial Classification', *Systematic and Applied Microbiology*. doi: 10.1016/S0723-2020(84)80019-3.

Xu, L. *et al.* (2019) 'OrthoVenn2: a web server for whole-genome comparison and annotation of orthologous clusters across multiple species', *Nucleic acids research*. doi: 10.1093/nar/gkz333.

Xue, Y., Osborn, J., Panchal, A. and Mellies, J. L. (2015) 'The RpoE stress response pathway mediates reduction of the virulence of enteropathogenic *Escherichia coli* by zinc', *Applied and Environmental Microbiology*, 81(11), pp. 3766–3774. doi: 10.1128/AEM.00507-15.

Yadav, G., Gokhale, R. S. and Mohanty, D. (2003) 'Computational approach for prediction of domain organization and substrate specificity of modular polyketide synthases', *Journal of Molecular Biology*. doi: 10.1016/S0022-2836(03)00232-8.

Yagüe, P. *et al.* (2016) 'Subcompartmentalization by cross-membranes during early growth of *Streptomyces* hyphae', *Nature Communications*. doi: 10.1038/ncomms12467.

Yusof, N. *et al.* (2020) ' Complete Genome Sequence of Lignin-Degrading *Streptomyces* sp. Strain S6, Isolated from an Oil Palm Plantation in Malaysia ', *Microbiology Resource Announcements*. doi: 10.1128/mra.01332-19.

Zambelloni, R., Marquez, R. and Roe, A. J. (2015) 'Development of antivirulence compounds: A biochemical review', *Chemical Biology and Drug Design*, 85(1), pp. 43–55. doi: 10.1111/cbdd.12430.

Zaslaver, A., Bren, A., Ronen, M., Itzkovitz, S., Kikoin, I., Shavit, S., Liebermeister, W., Surette, M. G. and Alon, U. (2006) 'A comprehensive library of fluorescent transcriptional reporters for *Escherichia coli*', *Nature Methods*. Nature Publishing Group, 3, pp. 623–628.

Zhan, Y. *et al.* (2008) 'Cytosporone B is an agonist for nuclear orphan receptor Nur77', *Nature Chemical Biology*. doi: 10.1038/nchembio.106.

Zhang, G., Li, Y., Fang, L. and Pfeifer, B. A. (2015) 'Tailoring pathway modularity in the biosynthesis of erythromycin analogs heterologously engineered in *E. coli*', *Science Advances*, (May), pp. 1–8. doi: 10.1126/sciadv.1500077.

Zhou, Z., Sun, N., Wu, S., Li, Y. Q. and Wang, Y. (2016) 'Genomic data mining reveals a rich repertoire of transport proteins in *Streptomyces*', *BMC Genomics*. doi: 10.1186/s12864-016-

2899-4.

Appendix A

1) Biosynthetic gene clusters encoded by *S. bottropensis* (antiSMASH)

Region	Type	Most similar	Percentage
1	T1PKS,NRPS, T2PKS	Spore pigment	83
2	Lanthipeptide	NA	0
3	Siderophore	NA	0
4	T2PKS, PKS-like	Rishirilide B	100
5	Terpene	FD-594	6
6	Bacteriocin	NA	0
7	Terpene	Geosmin	100
8	NRPS-like	S56-p1	0
9	Siderophore	NA	0
10	Ectoine	Ectoine	100
11	NRPS-like	Octacosamicin	12
12	Terpene	NA	0
13	T1PKS	Maduropeptin	3
14	Indole	Terfestatin	28
15	Linardin	Pentostatine	9
16	Siderophore	NA	0
17	NRPS, Melanin	Scabicelin	100
18	Terpene	Hopene	92
19	NRPS-like, T1PKS	Polyoxin	5
20	T1PKS	Auroromycin	11
21	Lanthipeptide, bacteriocin	Informatipeptin	100
22	Butyrolactone	Griseoviridin	8
23	Terpene	Hopene	92
24	PKS-like, LAP,butyrolactone, T1PKS	RK-682	45
25	Melanin	Melanin	60
26	Siderophore	Desferrioxamine B	83
27	Bacteriocin, bottromycin	Bottromycin A2	51
28	Lanthipeptide	NA	0

2) Biosynthetic gene clusters encoded by *S. collinus*

Region	Type	Most similar known cluster	Similarity
Region 1	lanthipeptide	Streptocollin	75%
Region 2	lassopeptide,NRPS	Azicemicin	11%
Region 3	NRPS,transAT-PKS,T1PKS	Kirromycin	62%
Region 4	terpene	Isorenieratene	100%
Region 5	NRPS-like	Lasalocid	11%
Region 6	terpene,melanin	Melanin	71%
Region 7	NRPS,T1PKS,bacteriocin	Foxicins A–D	12%
Region 8	NRPS-like	Stenothricin	13%
Region 9	T3PKS	Herboxidiene	6%
Region 10	T1PKS,hglE-KS	Cinnamycin	14%
Region 11	ectoine	Ectoine	100%
Region 12	thiopeptide,T2PKS,NRPS,PKS-like,oligosaccharide	Rabelomycin	29%
Region 13	melanin	Istamycin	5%
Region 14	siderophore	Desferrioxamine B	83%
Region 15	T2PKS	Spore pigment	83%
Region 16	T1PKS,NRPS		
Region 17	NRPS	PM100117 / PM100118	8%
Region 18	ladderane,NRPS	Calcium-dependent antibiotic	25%
Region 19	terpene	Albaflavenone	100%
Region 20	siderophore		
Region 21	bacteriocin	Lipopeptide 8D1-1 / lipopeptide 8D1-2	4%
Region 22	other,T1PKS	Tetronasin	3%
Region 23	bacteriocin		
Region 24	terpene	Geosmin	100%
Region 25	siderophore		
Region 26	terpene	Pentalenolactone	58%
Region 27	terpene	Hopene	92%
Region 28	other	A-503083	7%
Region 29	NRPS,bacteriocin	Informatipeptin	57%
Region 30	NRPS,transAT-PKS,T1PKS	Kirromycin	62%
Region 31	NRPS,lassopeptide	Azicemicin	11%
Region 32	lanthipeptide	Streptocollin	75%

Appendix B: