Exploitation of underused Streptomyces through a combined metabolomics-genomics workflow to enhance natural product diversity.

BURNS, J.

2020

Exploitation of underused Streptomyces through a combined metabolomics-genomics workflow to enhance natural product diversity

Joshua Burns

A thesis submitted in partial fulfilment of the requirements of the

Robert Gordon University

for the degree of Doctor of Philosophy

This research programme was carried out

in collaboration with NCIMB Ltd.

May 2020

ACKNOWLEDGEMENTS

Firstly, I would like to thank my project supervisors Christine Edwards and Linda Lawton at RGU and Sam Law at NCIMB. Their guidance and support at all stages of my PhD, and the Knowledge Transfer Partnership on which it was based, have been essential for the entirety of the project and I am extremely grateful for it.

The other members of CyanoSol in RGU were hugely helpful so thanks go to Len Montgomery, Aakash Welgama, Carlos Pestana, Calum McNerney, Julia Waack, Declan Maxwell, and Joe Palmer for all their help and advice at every stage of the process. I am also indebted to Joanna Reed, Philippa Hulme, and Coline Buanic for their help in data generation for Chapter 5, and to Dan Swan at NCIMB for all his help with *Streptomyces* genome sequencing and input on Chapter 4. Next are all the other RGU students who made my stay in Aberdeen enjoyable through a mix of their friendship, willingness to sit through complaints, and (usually) tolerating my jokes. Amongst others, I'd most like to thank Calum McNerney, Matteo Scipioni, Franzi Pohl, Hazel Ramage, Mhairi Paul, Zoi Papadatou, Teo Stoyanova, Martin Corsie, Quirin Werthner, Tesnime Jebara, Ahmed Salaheldin, PJ Barron and Qas Ali.

I would also like to thank the staff at RGU, especially Andrea Macmillan and Dorothy McDonald for all their help with admin and large amounts of free tea. I also need to thank the other staff at NCIMB, in particular Carol Phillips for her contributions to the project and advice on the industry and business aspects.

I am deeply grateful to Innovate UK for funding the project through KTP grant 10124, and for the guidance of the KTP advisors Mark Abbs and Ian Heywood.

Finally, I want to thank my family for the constant love and support that they have given me over the last few years and at every other time. To my Mum, Dad, Caroline, Rebecca, Ofra, Maya, Eldad and Lottie, thank you, and I wouldn't have been able to do any of it without you.

LIST OF ABBREVIATIONS

A	Adenylation
ACP	Acyl carrier protein
ACT	Actinorhodin
AMR	Antimicrobial resistance
ANI	Average nucleotide identity
antiSMASH	Antibiotics and secondary metabolite analysis shell
AT	Acyl transferase
autoMLST	Auto multi locus sequence typing
BA	Bennett's agar
BEH	Ethylene bridged hybrid
BGC	Biosynthetic gene cluster
BLAST	Basic local alignment search tool
BLASTN	Basic local alignment search tool-nucleotide
BLASTP	Basic local alignment search tool-protein

BUSCO	Benchmarking universal single copy orthologs
C	Condensation
CCR	Carbon catabolite repression
CDA	Calcium dependent antibiotic
CDS	Coding sequences
CS	Corn steep liquor-soya flour
DSMZ	Deutsche sammlung von mikroorganismen und zellkulturen
ESBL	Extended spectrum beta-lactamase
ENV+	Polystyrene-divinylbenzene
EP	Extended phenotype
FASTA	Fast-all
GlcNAc	N-acetyl glucosamine
GNPS	Global natural products social molecular networking
GYM	Glucose yeast-malt
HP	High polarity pool

ISP5	International S*treptomyces* project medium 5
KS	Ketosynthase
LAP	Linear azoline containing peptide
LC-MS	Liquid chromatography-mass spectrometry
LP	Lower polarity pool
M19	Medium 19
M400	Medium 400
MHA	Muller-Hinton agar
MHB	Muller-Hinton broth
MiBIG	Minimum information about a biosynthetic gene cluster
MM	Molasses-meat extract
MMM	Minimal mannitol medium
MP	Medium polarity pool
MRSA	Methicillin-resistant *Staphylococcus aureus*
MS/MS	Mass spectrometry/mass spectrometry

MYM	Maltose yeast-malt
NaBu	Sodium butyrate
NCIMB	National collection of industrial and marine bacteria
NGS	Next generation sequencing
NRPS	Non-ribosomal polyketide synthetase
NUT	Nutrient agar
OM	Oatmeal medium
OSMAC	One strain many compounds
PBP	Penicillin-binding protein
PCA	Principal component analysis
PCP	Peptidyl carrier protein
PE	Paired end
PepS	Peptone-sucrose medium
PKS	Polyketide synthetase
PLS-DA	Partial least squares-discriminant analysis

ProYM	Proline yeast-malt
QToF	Quadrupole time of flight
QUAST	Quality assessment tool for genome assemblies
RAST	Rapid annotation using subsystem technology
RiPP	Ribosomally synthesized and post-translationally modified peptides
RED	Undecylprodigiosin
SCC*mec*	Staphylococcal cassette chromosome *mec*
SE	Single end
SEA	Soil extract Agar
SGT	Soybean-glucose-tryptone
SMet	Specialised metabolite
SM	Sucrose-meat extract
SPNT	Supernatant
TSA	Tryptone soya agar
UPLC	Ultra-performance liquid chromatography

WHO	World health organisation
WGS	Whole genome sequencing
XTT	2,3-bis (2-methoxy-4-nitro-5-sulfophenyl)-5-[(phenylamino) carbonyl]-2H-tetrazolium hydroxide

ABSTRACT

The genus *Streptomyces* is the source of approximately two thirds of all clinically used antibiotics. Despite being the source of so many specialised metabolites, genomic analysis indicates that most *Streptomyces* strains have the potential to produce around 25 bioactive metabolites, some of which may be the basis of novel therapies. This makes culture collections of *Streptomyces* spp. an easily accessible but underused resource to mine for genomic and metabolomic variety. Therefore, the main aim of this project was to initiate exploitation of the culture collection at NCIMB Ltd. by expanding the available chemical space from underutilised *Streptomyces* for the production of novel antibiotics. This primarily used a mixture of metabolomic and genomic methods.

A high-throughput culture parameter screen was designed around multiple carbon sources, nitrogen sources, and extraction sample times. This was tested on the model species *S. coelicolor* A3(2) to compare differences in the production of known specialised metabolites, using UPLC-MS to analyse crude extracts from growth on agar. Data was analysed using MZmine and putative metabolites identified using freely available MS/MS databases, primarily GNPS. This showed clear variation in production of 9 identified metabolites including deferoxamines, germicidins, undecylprodigiosin and coelichelin as a result of different culture parameters. Therefore, the screen successfully expanded the available chemical space, so was applied to non-model *Streptomyces* strains.

The screen was used to compare the total metabolomic variety produced by 3 *Streptomyces* isolated from different environments in order to select a strain for further investigation. Comparing metabolomic features using principal component analysis showed the Costa Rican soil isolate *S. costaricanus* to produce the most variety versus the other 2 *Streptomyces* strains. The metabolite family most responsible for principal component separation was identified as the actinomycins. Scale-up of both agar and broth culture was used for metabolite dereplication and bioassays against multidrug resistant *Acinetobacter baumannii*, one of the bacteria on the World Health Organisation's list of pathogens most urgently requiring new therapies. Fractions were derived from broth culture supernatant and agar crude extract by flash chromatography, resulting in semi-purified fractions. The predominant metabolite families in

fractions were actinomycins and deferoxamines, which were further split by polarity into separate fractions. This resulted in rapid purification of metabolites, with 1 fraction comprising 80% deferoxamine B by weight. Fractions were tested against *A. baumannii* using the 2,3-bis (2-methoxy-4-nitro-5-sulfophenyl)-5-[(phenylamino) carbonyl]-2H-tetrazolium hydroxide (XTT) assay, which showed partial inhibition of growth at 50 µg/ml. Examining the bioactive fractions showed potentially novel minor peaks which could be responsible for bioactivity.

A high-quality full genome of *S. costaricanus* was obtained using a combination of MiSeq and MinION sequences. This was analysed with RAST and antiSMASH to determine the specialised metabolite potential of *S. costaricanus*. AntiSMASH detected 33 biosynthetic gene clusters (BGCs), above the mean for *Streptomyces*. Thus, the confirmed genomic potential also suggested a wider metabolite variety, as indicated by the metabolomic screen. Some of the 33 BGC products had been previously detected by UPLC-MS, like actinomycin D and deferoxamine B. Other BGCs had 0% homology to known BGCs, including a terpene BGC which only showed core gene homology to 2 other *Streptomyces*. One of these strains shared all of the BGCs with *S. costaricanus*, including their sequential order and closely approximated genomic locations. Comparison of marker genes with autoMLST gave preliminary evidence for the taxonomic reclassification of *S. costaricanus* as a strain of *S. griseofuscus.*

Starting from a large collection of unexploited *Streptomyces*, this project catalogued the metabolomic and genomic diversity of a single strain and its bioactive potential. Together, the project stages formed a workflow for further exploitation of NCIMB *Streptomyces* and other microbes.

CHAPTER 1

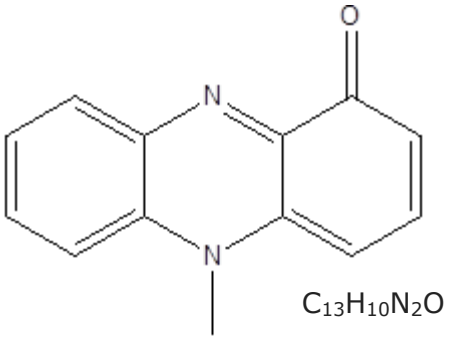
INTRODUCTION

1 Introduction

1.1 An overview of modern microbial antibiotic discovery

Modern microbial derived antibiotic discovery can be traced back to 2 main events. The first of these was the accidental, but hugely significant, discovery of penicillin by Alexander Fleming in 1928 (Fleming, 1929). While penicillin is sometimes referred to as the first microbially-derived antibiotic (Gaynes, 2017), this is not strictly correct. Pyocyanin (Figure 1.1), produced by *Pseudomonas aeruginosa*, was shown to inhibit microbial growth including that of *Bacillus anthracis* in infected rabbits (Woodhead & Wood, 1890; Gould, 2016). Pyocyanin has likely been overlooked as it had limited applicability compared to penicillin, along with the fact that most of the original publications were not in English.

Figure 1.1 Structure of pyocyanin.



The second event was the development of a systematic antibiotic discovery platform by Selman Waksman, beginning what was later termed the "golden age" of antibiotic discovery running for approximately 20 years from 1940 – 1960 (Katz & Baltz, 2016). This platform led to the discovery of many of the antibiotics currently used in the clinic (Brown & Wright, 2016). The origins, microbes, and derived antibiotics are outlined in the following sections, demonstrating the early successes of microbial specialised metabolites (SMets) as medical therapies.

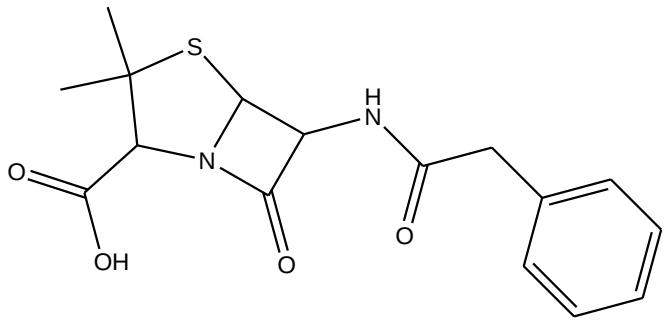
1.1.1 Origins of antibiotic discovery: penicillin

Fleming's original research that gave rise to the discovery of penicillin was on colony variation in "pyogenic *Staphylococcus*", or *Staphylococcus aureus* (Fleming, 1944). This involved routine microscopy, which required removing the lid from a culture plate. Following the exposure of one plate to the air, and after several days of incubation at room temperature, a mould contaminant – described by Fleming as a "felted mass", indicating a filamentous species of fungi – was seen on the plate, along with the inhibition and lysis of *S. aureus* colonies.

After isolation and nutrient broth culture of the contaminant, the culture supernatant was found to inhibit multiple pathogenic microbes including *Streptococcus pyogenes*, *Streptococcus pneumoniae*, *Neisseria gonorrhoeae*, and diptheroid *Corynebacterium* spp. No toxic effects were seen when intravenously injected into rabbits or mice, after hourly application to human conjunctiva, and human leukocyte function was not impaired (Fleming, 1929). Antimicrobial bioactivity remained even when diluted up to 800-fold, which was greater than the maximum useful dilution factor for the commonly used antiseptic phenol. This indicated production of a potent antibacterial compound with clear therapeutic applications.

The mould itself was originally identified as a *Penicillium* species, hence the name penicillin for the newly discovered antibiotic. Multiple natural, or non-synthetic, penicillins are now known; penicillin G, shown in Figure 1.2 is the major product and the one referred to in this section, with other minor metabolites like penicillins V, F, and K also present (Torres *et al.*, 1999).

Figure 1.2 Structure of penicillin G.



$C_{16}H_{18}N_2O_4S$

Taxonomic characterisation of the organism producing penicillin has undergone multiple revisions. It was putatively classified as being closest to *P. rubrum* (Fleming, 1929), but later reidentified by Thom (1945) as a strain of *P. notatum*, which itself was then merged into the species *P. chrysogeneum* (Samson *et al.*, 1977). Though still commonly referred to as *P. chrysogenum*, Houbraken *et al.* (2011) used more recent metabolic and molecular methods to further separate *P. chrysogeneum* and what they propose as Fleming's original strain, *P. rubens*. Despite this, *P. chrysogenum* is still sometimes used in the literature as the first discovered producer of penicillin (Ziemons *et al.*, 2017).

Production of penicillin by strains of *P. rubens* was demonstrated to vary across a wide range of concentrations, from 0 to 129 µg/ml (Raper *et al.*, 1944). *P. rubens* is known to undergo high levels of mutation, which led Thom (1945) to hypothesise that this was to support SMet variety. Thom (1945) therefore suggested that the strain was present at low levels within Fleming's lab until it mutated to produce enough penicillin to outcompete other microbiota. Thom (1945) also reported that his collection continued to receive requests for *P. rubrum* for several years after its renaming to *P. notatum*, which were replied to by sending *P. notatum* with an explanation of the then current taxonomy. This emphasises that, while taxonomic processes may lead to re-classification of industrially and medically relevant strains, culture collections are important for both maintaining the original strain and ensuring that requests for a SMet producer are correctly documented.

Penicillin was introduced to the clinic in the early 1940s with good results, though not all patients survived as problems with scale-up and isolation led to a limited supply of pure drug (Abraham *et al.*, 1941); treatment was also hampered by penicillin's low stability, which continues to be a problem (Miller, 2002). Following successful production trials in Oxford (Chain *et al.*, 1940), production was extended across the United States in 1941, eventually providing enough to treat allied forces in World War II (Gaynes, 2017). While Fleming's discovery of penicillin was not intentional, its importance was realised and successfully transformed into a successful medical product. It also hinted at other similar microbial SMet with similar activities, ready to be discovered and exploited through a systematic antibiotic discovery programme. This would be

confirmed through the research of Selman Waksman, investigating soil microbes and their specialised metabolites, predominantly from the genus *Streptomyces*.

1.1.2 Origins of antibiotic discovery: The Waksman platform

1.1.2.1 Pre-antibiotic era research

The Waksman platform was an antibiotic discovery system focused on using soil bacteria to produce bacteriostatic or bactericidal compounds active against a range of pathogenic microbes. Waksman's research originally concentrated on the interactions of different soil microbes with the environment and aspects of primary metabolism. As examples, Waksman (1918) focused on proteolytic enzymes of soil fungi and "actinomycetes", which refers to the now famous natural product producing order *Actinomycetales* and particularly the genus *Streptomyces* (Hua Zhu *et al.*, 2014). Waksman & Joffe (1922) concentrated on soil sulphur oxidation, and Waksman & Skinner (1926) on cellulose decomposition. In reference to *Streptomyces*, many of the strains were originally assigned to the *Actinomyces* genus when their antibiotics were discovered; one example is *Actinomyces griseus*, the producer of streptomycin (Schatz *et al.*, 1944). These were later reclassified into *Streptomyces*, which is how they are referred to here unless in a direct quote from an article.

There was a brief mention of the *Actinomycetales* as producers of "substances toxic to bacteria" resulting in a zone of inhibition "free from fungus and bacteria" on agar plate culture in Waksman & Starkey (1923), though the main focus of the paper was on soil sterilisation. However, a link was not made between these toxic substances and their potential effects on other microbes including human pathogens (Woodruff, 2014). So, while Waksman was familiar with *Actinomycetales* that had the ability to inhibit the growth of other microbes, this was more in an ecological context rather than as sources of medicines. The first article specifically on soil microbes and their potential for medicinal use was published in 1940 (Waksman & Woodruff, 1940), in which 4 microbial groups of special interest were identified.

The first, *Pseudomonas* spp. has already been mentioned in section 1.1 as the source of pyocyanin, the first antibiotic derived from a microbe. The second group was *Bacillus* spp., including *Bacillus brevis* which had recently been found

by a past student of Waksman to produce a substance active against Gram-positive bacteria (Dubos, 1939). This substance, named tyrothricin, was actually a mixture of the antibiotics gramicidin and tyrocidine (Dubos & Hotchkiss, 1941). The mixture and the overall components, especially tyrocidine, are unsuitable for clinical use as they induced haemolysis (Rammelkamp & Weinstein, 1942). The remaining 2 groups are *Actinomycetales* and then fungi, with *Penicillium* and *Trichoderma* noted as 2 genera of interest. *Actinomycetales*, especially *Streptomyces* spp., were the main workhorse of the platform and so will be explored in more detail.

1.1.2.2 Discovery of actinomycin

Dubos' identification of a new use for *B. brevis* was responsible for Waksman's refocusing on antibiotic discovery, especially those active against Gram-negatives as penicillin, gramicidin, and tyrocidine are mostly limited to Gram-positive bacteria (Woodruff, 2014). Initial tests used the same method as Dubos (1939): supplementing soils with a target pathogen over multiple months, selecting for any microbes able to destroy the pathogens and use their nutrients for metabolism. The major difference was the use of Gram-negatives to supplement the soil pre-plating, with the logic that any enriched soil microbes should be able to kill Gram-negatives. Whether this actually affected results is unknown since no negative control was done.

By enriching with *Escherichia coli* and *Klebsiella aerogenes* then plating out serial dilutions of the soil, one *Actinomycetales* species (Waksman & Woodruff, 1941) was found to consistently inhibit the target pathogens. *K. aerogenes* is referred to as *Aerobacter aerogenes* in Waksman & Woodruff (1941); its reclassification to *Entereobacter aerogenes* and then *K. aerogenes* is described in Tindall *et al.* (2017). Similar to the above confusion over *Penicillium* names for culture collection requests, this highlights the importance of maintaining a collection aware of these differences to drug discovery projects, as an inconsistent target test pathogen would prevent reproducible experiments.

The antagonist microbe was determined to be a new species now known as *S. antibioticus* (Katz *et al.*, 1956), and the antibiotic named actinomycin. An early version of the One Strain Many Compounds, or OSMAC, (Bode *et al.*, 2002) method was applied to culture *S. antibioticus* and assay bioactivity, varying

nitrogen sources with starch as the carbon source. This showed it to be primarily active against Gram-positives: *K. aerogenes* was not inhibited at all by 200 µg/ml crude extract, though *E. coli* was inhibited at the same concentration from multiple media extracts. In comparison, *B. subtilis* was totally inhibited at 1 µg/ml.

It should be noted that as this experiment took place before the current antibiotic crisis, these are unlikely to be multi-drug resistant strains so actinomycin is now even less effective than in Waksman & Woodruff (1941). Alongside this, actinomycin was also found to be toxic to animals. Mice were killed by a dosage of 10 µg per 20 g body weight, and was lethal whether injected into the body, subcutaneously, or taken orally (Waksman & Tishler, 1941). The bioactivity of actinomycin against both prokaryotes and eukaryotes is due to its mechanism of action, as it forms a complex with DNA by inserting itself at CpG sites, prevent RNA Polymerase accessing the DNA and so stopping transcription. While it is not used as an antibiotic, the cytotoxic effects of actinomycin do make it suitable for anticancer chemotherapy (Lo *et al.*, 2013). Actinomycin has historically been used to treat Wilms tumour, a renal cancer (Fernbach & Martyn, 1966), but more recently as an enhancer of immunotoxins which target cancerous cells (Liu *et al.*, 2016)

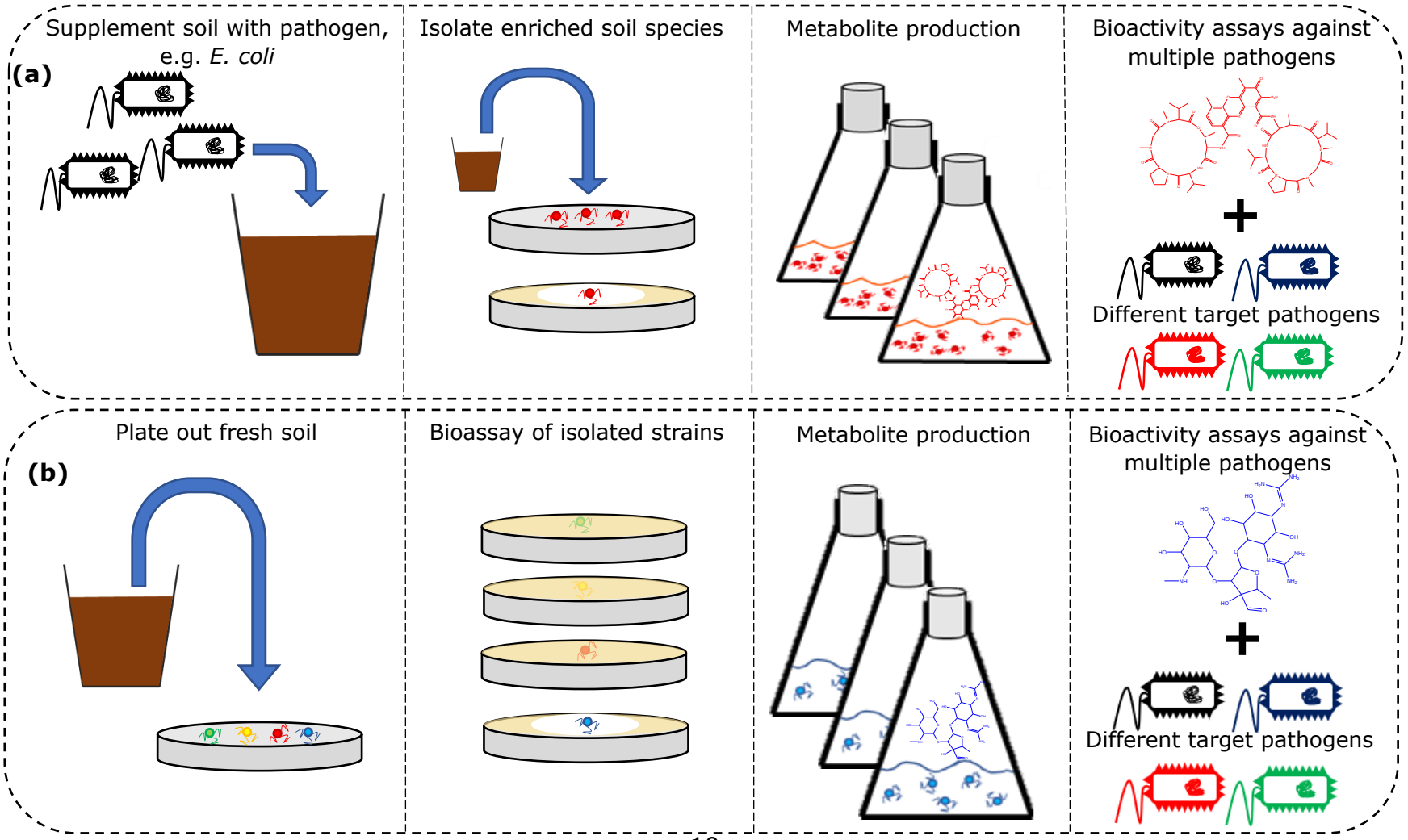
Actinomycin was the first antibiotic to be discovered by the Waksman platform, and while it was not the ideal Gram-negative antibiotic with no toxic side effects, it was proof of the concept that a targeted screen of *Actinomycetales* would produce novel bioactive metabolites.

1.1.2.3 Platform expansion and discovery of streptomycin

The screening platform was modified to directly plate soil suspensions onto agar plates and test the colonies for bioactivity; the two methods are summarised in Figure 1.3. This led to the discovery of antibiotics from fungi: fumigatin and fumigacin from *Aspergillus fumigatus*, and clavacin from *Aspergillus clavatus* (Waksman *et al.*, 1943). Continuing with the trend of Gram-negatives being more difficult to target, fumigatin and fumigacin showed antibiotic activity against Gram-positives but not Gram-negatives with the exception of some against *E. coli*. Clavacin showed strong activity against both types at similar levels, but all 3 of these antibiotics were toxic to eukaryotic cells and therefore

not suitable for widescale usage. The first platform antibiotic to show potent activity towards Gram-negative and Gram-positive bacteria with apparently minimal toxicity was streptothricin from *S. lavendulae* (Waksman & Woodruff, 1942). Unfortunately streptothricin was later found to have delayed kidney toxicity, halting clinical trials (Inamori *et al.*, 1979).

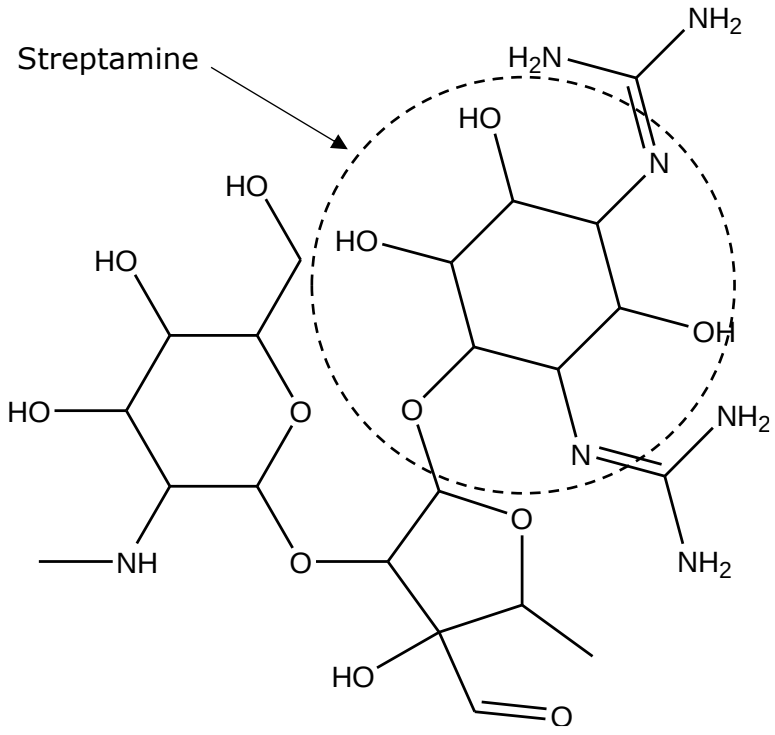
Figure 1.3 Waksman Platform antibiotic discovery methods **(a)** initial method based on soil supplementation with target pathogen, and **(b)** revised method using isolation then bioactivity screening.



Screening soil isolates for bioactive SMet producers had led to the discovery of novel compounds, but none of them exhibited inhibition of Gram-negative strains while causing minimal toxic side effects to animal models. The first metabolite discovered by the platform to fulfil this was streptomycin, produced by *S. griseus* (Schatz *et al.*, 1944). Streptomycin is effective against both Gram-positive and Gram-negative bacteria with no lethally toxic effects to eukaryotes since, unlike actinomycin, it acts on a prokaryote-specific target. Streptomycin disrupts protein synthesis by binding to the 30S ribosomal subunit (Luzzatto *et al.*, 1968), warping the structure at the 16S codon recognition site (Demirci *et al.*, 2013) resulting in an incorrect translational complex which misreads codons and produces random sequence proteins. Some neurotoxic effects were seen, but the reduced derivative dihydrostreptomycin was found to alleviate these (Waksman & Lechevalier, 1949); later experiments did find a nephrotoxic effect to the broader class of antibiotics, but these are avoided through correct body mass to dosage level calculations (Pedersen *et al.*, 1987).

Initially streptomycin was used to treat tuberculosis, caused by another actinobacterium *Mycobacterium tuberculosis*. Streptomycin was the first drug to inhibit tuberculosis progression in the commonly used animal model, guinea pigs (Hinshaw *et al.*, 1946). It was also the first of the aminoglycoside antibiotic group to be discovered, which are characterised by an inositol derivative – streptamine for streptomycin – bonded to an aminosugar group, with 2 or more further amino groups and free hydroxyls (Becker & Cooper, 2013) shown in Figure 1.4.

Figure 1.4 Structure of streptomycin with inositol derivative segment noted.



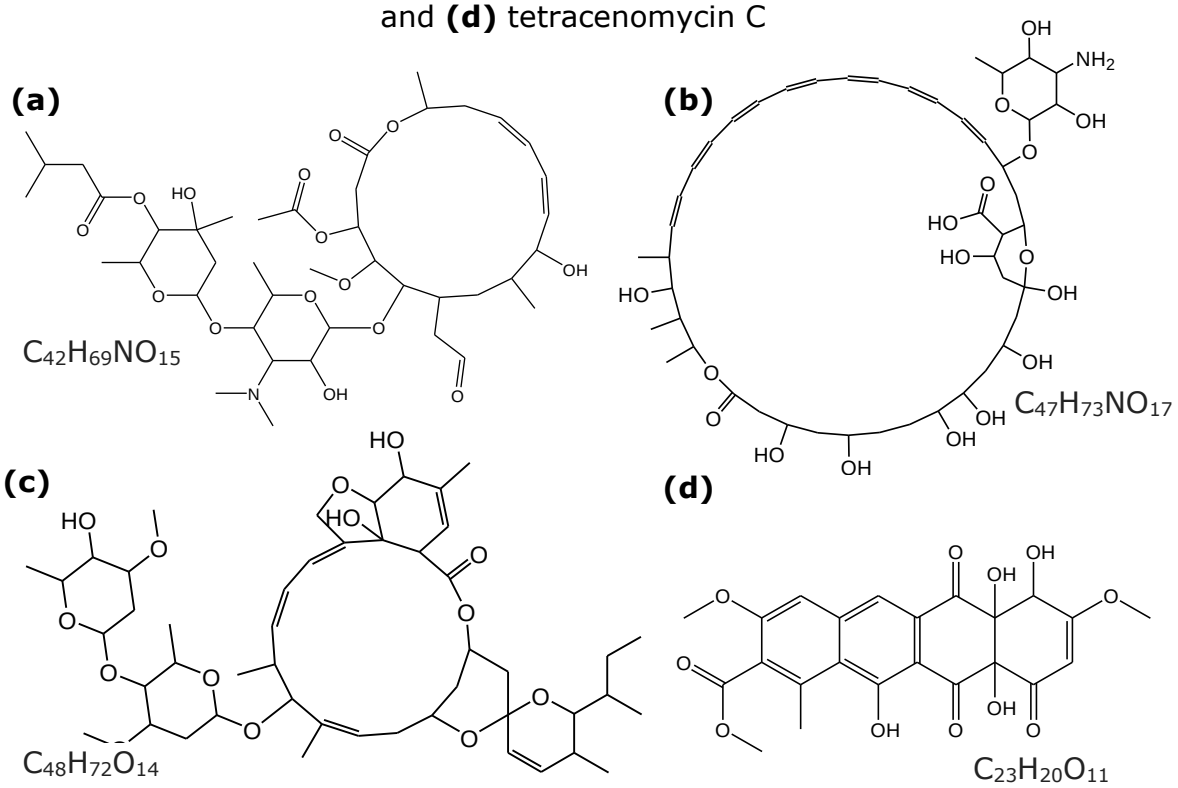
Another aminoglycoside, neomycin, would later be discovered in Waksman's lab (Waksman & Lechevalier, 1949). Produced by *S. fradiae*, it was – like streptomycin – active against *M. tuberculosis*, but crucially against strains which were already displaying resistance to streptomycin. It also displayed similar if not better levels of Gram-negative antibiotic activity, with 5 µg/ml being enough to totally prevent growth of streptomycin resistant *E. coli*.

1.1.2.4 Common classes of bioactive specialised metabolites

Streptomycin belongs to the polyketide metabolite class, which are produced by polyketide synthases (PKS) (Gomes *et al.*, 2013). Biosynthetic gene clusters (BGCs) for PKS are common throughout *Streptomyces* spp. and other *Actinomycetales* genera (Risdian *et al.*, 2019), and synthesise a wide range of metabolites with a variety of structures and bioactivities. These include antibacterials like josamycin from *S. narbonensis*, the antifungal amphotericin B from *S. nodosus*, antihelminith avermectin from *S. avermitilis*, or the antitumour tetracenomycin produced by *S. glaucescens* (Caffrey *et al.*, 2001; Thompson *et*

al., 2004; Petkovic *et al.*, 2006; Katz & Baltz, 2016). Structures are shown in Figure 1.5.

Figure 1.5 Structures of **(a)** josamycin, **(b)** amphotericin B, **(c)** avermectin, and **(d)** tetracenomycin C



PKS operate as an assembly line, with each segment performing a different enzymatic function to construct the metabolite. There are 3 types of PKS in *Streptomyces*, types I, II, and III, which differ in their mechanisms and product classes, and may also be combined to build hybrid metabolites. The two most common types in *Streptomyces,* types I and II, are briefly outlined below. An acyl-CoA unit is generally used as the starting building block for both types.

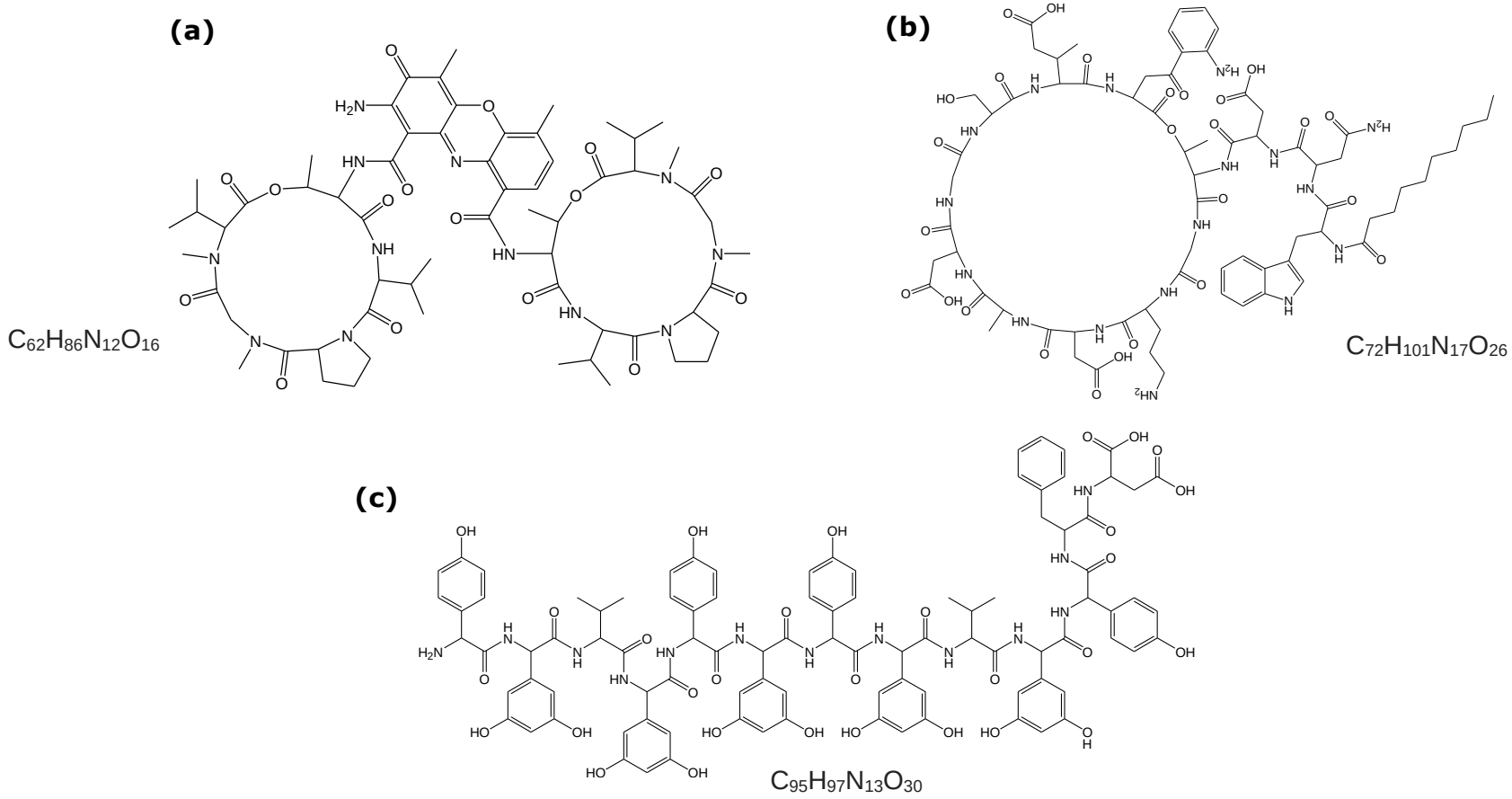
In type I PKS, the acyl-CoA is processed by multiple modules, with each module containing a set of functional domains. A module must include a ketosynthase (KS), an acyltransferase (AT), and end on an acyl carrier protein domain (ACP) as a minimum. The starter unit is built into a metabolite through addition of extender units like methylmalonyl-CoA, which are joined through the KS catalysing C-C bond formation. The extender unit is selected by the AT domain, and the product transferred to the next module by the ACP. Further modifications are done by other domains, like keto reductases or dehydrateases

(Risdian *et al.*, 2019). This results in PKS megastructures which must be associated to synthesise the product, which unidirectionally passes through each module, until release from the C terminus of the PKS is catalysed by a thioesterase (Musiol-Kroll & Wohlleben, 2018)

In comparison, type II PKS are discrete, iterative enzymes, which include a KS, chain-length factor, and ACP. Additional downstream tailoring enzymes like ketoreductases, aromatses, or cyclases can be included to enhance the minimal PKS, followed by further tailoring to the scaffold from methyltransferases, oxygenases, and glycosyltransferases. Type II PKS produce polycyclic aromatic compounds, with most of the structural variety and bioactivity coming from the downstream tailoring (Zhang & Tang, 2009).

A 2nd common class of biosynthetic enzymes are the non-ribosomal peptide synthases (NRPS). Similar to PKS, *Streptomyces* NRPS construct a variety of metabolites with multiple bioactivites, but unlike PKS, directly incorporates amino acids into the structure without using ribosomal products. These include the antitumour antibiotic actinomycin from *S. antibioticus,* antibacterial daptomycin from *S. roseosporus*, and antiviral feglymycin produced by *S.* sp. DSM 11171 (Férir *et al.*, 2012; Katz & Baltz, 2016); structures are shown in Figure 1.6. Similar to type I PKS, most NRPS operate as sequence of modular multifunctional enzymes, specifically condensation (C), adenylation (A), and peptidyl carrier protein (PCP) domains. The C domain forms C-N bonds between amino acids, which are recruited by the A domain, and transferred to the PCP domain to be recognised by the C domain on the next module, or released from the NRPS by the hydrolysing or cyclysing effects of a thioesterase (Nikolouli & Mossialos, 2012). NRPS and PKS clusters can form combined hybrid products for even more structural diversity, and correspondingly varied bioactivity, such as oxazolomycin from *S. albus* JA3453 (Zhao *et al.*, 2010).

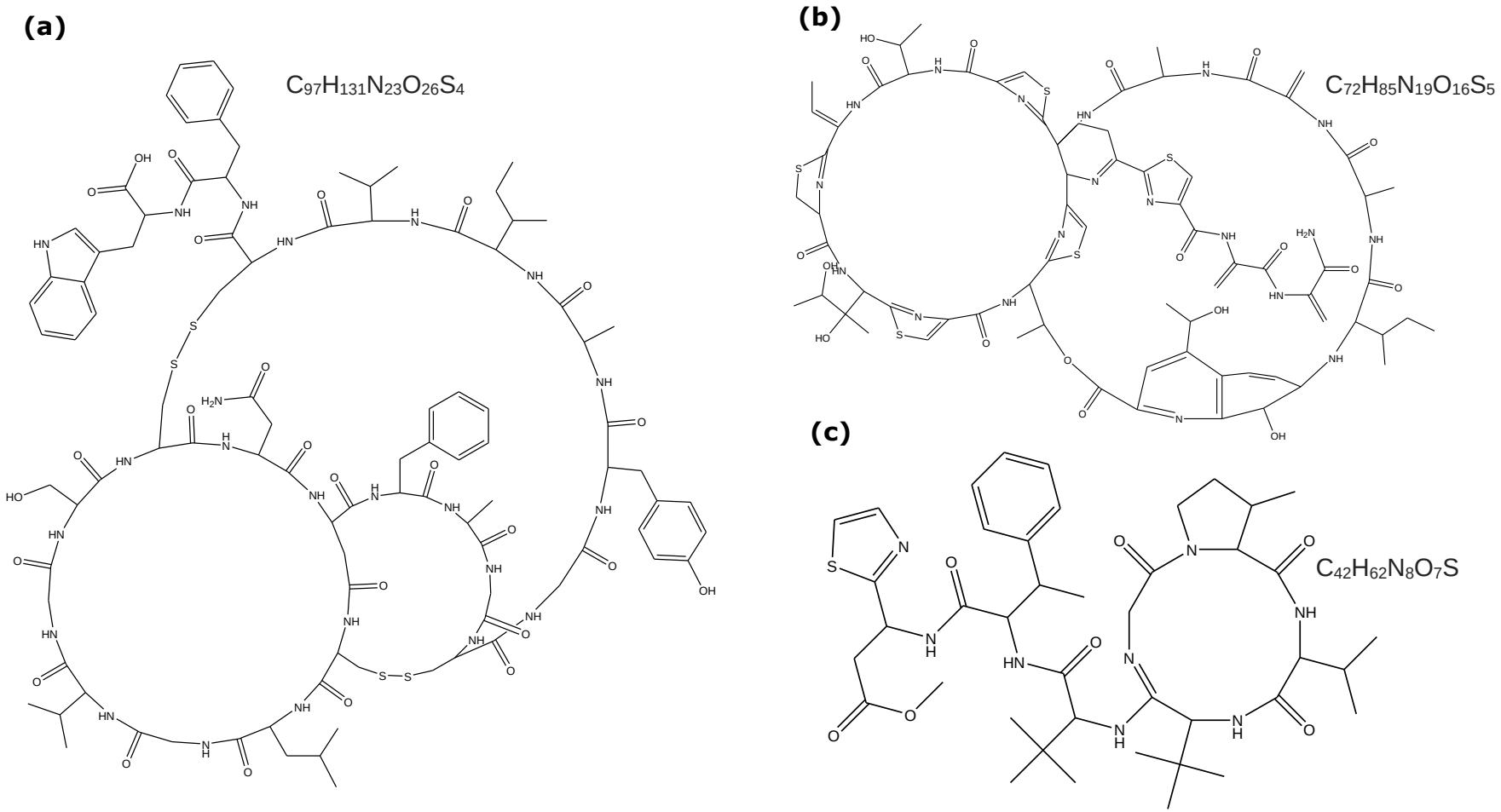
Figure 1.6 Examples of non-ribosomal peptides produced by *Streptomyces* spp. **(a)** actinomycin D, **(b)** daptomycin, **(c)** feglymycin.



A final class of metabolites are the ribosomally synthesised and post-translationally modified peptides, or RiPPs. RiPPs undergo multiple rounds of post-translational processing from the precursor peptide to form the mature product. In *Streptomyces*, starting at the N terminus, the precursor product is generally made of a leader peptide, followed by the core peptide, then a recognition sequence. The leader peptide is involved in assigning post-translational tailoring to the core peptide and export, while the recognition sequence can cyclise or enable proteolytic cleavage and produce the mature RiPP (Arnison *et al.*, 2013).

As with polyketides and non-ribosomal peptides, RiPPs are present in *Streptomyces* in numerous sub-types, including lassopeptides, thiopeptides, and bottromycins. As RiPPs are less well studied than polyketide and non-ribosomal peptides, there are few current commercial or medical applications for them, with the food additive nisin from *Lactotoccus lactis* as an exception (Tan *et al.*, 2019). Despite this, multiple bioactivites are still exhibited by RiPPs. *S.* sp. AA6537 produces the lasso peptide siamycin I, which inhibits human immunodeficiency virus envelope fusion (Lin *et al.*, 1996). *S. azureus*, *S. laurentii* and other species synthesise the antibacterial thiostrepton (Mocek *et al.*, 1993), and the MRSA-inhibiting bottromycins (Hou *et al.*, 2012; Arnison *et al.*, 2013). Structures for these metabolites are shown in Figure 1.7.

Figure 1.7 Structures of the RiPPs **(a)** siamycin I, **(b)** thiostrepton, and **(c)** bottromycin A2



1.1.2.5 Adoption by other researchers

As a result of the success of Waksman's research into *Actinomycetales*, particularly *Streptomyces* there was widespread stimulation of similar research in both academia and industry (Woodruff, 2014). These resulted in the rapid discovery of other new antibiotics through to the 1960s. Some of these are shown in Table 1.1, which has been limited to just antibacterial metabolites produced by *Streptomyces*, and corresponding structures in Figure 1.8. All of the listed metabolites attack the ribosome through various different mechanisms of action. This was not an intentional selection but does reflect the importance of having a prokaryote-specific target in chemotherapy which is more likely to lead to a successful drug with minimal side effects in eukaryotes.

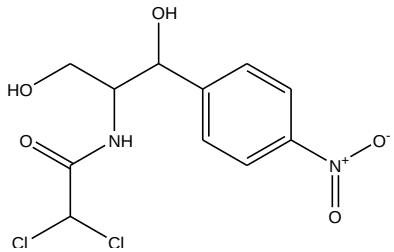
As a result of this drive towards *Streptomyces* based antibiotic discovery, over half of known antibiotics are produced by *Streptomyces* (D'Costa *et al.*, 2006) as well as approximately two thirds of commonly used clinical and veterinary drugs (Hopwood, 2013).

Table 1.1 *Streptomyces* derived antibiotics from the golden age of discovery.

Metabolite	Species	Mechanism	Activity	Reference
Chloramphenicol	*Streptomyces venezualae*	Binds to 50S ribosomal subunit, halting peptide chain elongation	Gram + Gram -	Ehrlich *et al.* (1947)
Oxytetracycline	*Streptomyces rimosus*	Binds to 30S subunit, preventing tRNA binding	Gram + Gram -	(Duggar, 1948)
Oleandomycin	*Streptomyces antibioticus*	Binds to 50S subunit, preventing complex formation	Gram + Gram - (weak)	Garrod (1957)
Lincomycin	*Streptomyces lincolnensis*	Binds to 23S RNA, stopping peptide bonding	Gram + Gram – (weak)	Macleod *et al.* (1964)
Kanamycin	*Streptomyces kanamyceticus*	Binds to 30S subunit, causing mRNA misreads and incorrect amino acid recruitment	Gram + Gram -	Maeda *et al.* (1957)

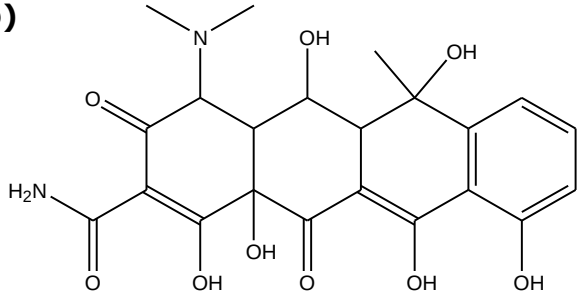
Figure 1.8 Structures of **(a)** chloramphenicol, **(b)** oxytetracycline, **(c)** lincomycin, **(d)** oleandomycin, and **(e)** kanamycin.

(a)



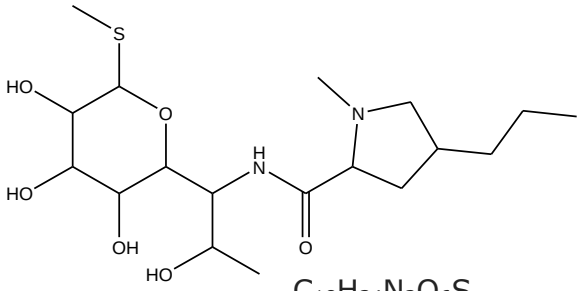
$C_{11}H_{12}Cl_2N_2O_5$

(b)



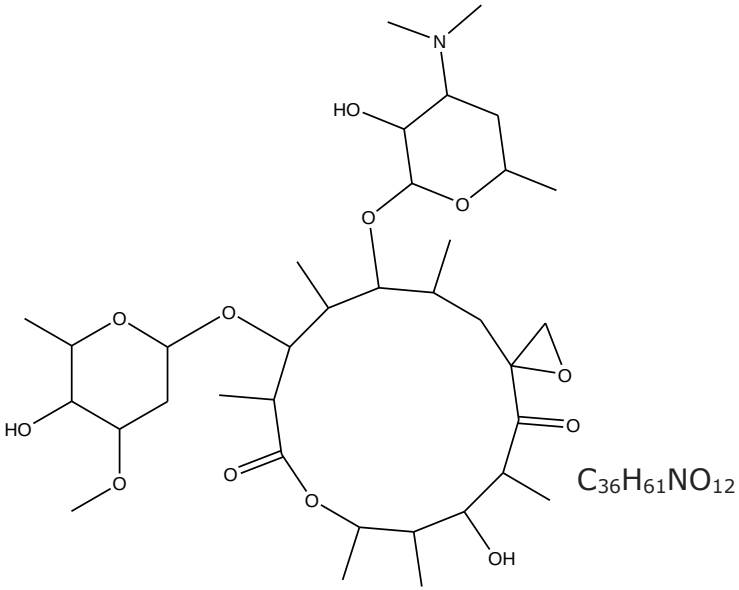
$C_{22}H_{24}N_2O_9$

(c)



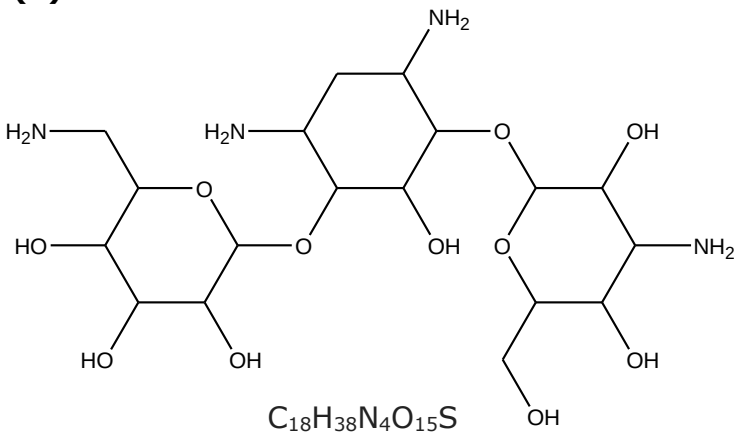
$C_{18}H_{34}N_2O_6S$

(d)



$C_{36}H_{61}NO_{12}$

(e)



$C_{18}H_{38}N_4O_{15}S$

Some of the novel metabolites like neomycin were explicitly noted to be of use as resistance was increasing to a previously effective drug. Unfortunately, the rate of SMet discovery was declining: approximately 12000 SMet discovered up to the 1960s – of which 160 were used as drugs or as the chemical scaffold for development – but only a further 10000 SMet were found by 2002 (Marinelli, 2009).

Since the end of the Golden Age in the 1960s, only 2 new classes of antibiotics have been discovered, both of which are only active against Gram-positive species: daptomycin, a lipopeptide produced by *S. roseosporus* (Debono *et al.*, 1987), and teixobactin from *Eleftheria terrae* (Ling *et al.*, 2015), for which no mechanisms of resistance have yet been found. The likelihood of discovering a new class is estimated to be low: based on the frequency of commonly found BGCs, 10% of species produce streptothricin, 1% streptomycin, and 0.1% tetracycline; daptomycin has a discovery probability of 2×10^{-7} (Lewis, 2013).

Shortly after the widespread production of penicillin, Fleming predicted in 1945 that the increased prescription of penicillin would lead to selection of resistant strains (Rosenblatt-Farrell, 2009). This had already been confirmed 5 years earlier (Abraham & Chain, 1940) with the discovery of a β-lactamase from *E. coli*. Antibiotic usage has increased along with population and intensive agriculture or farming. The majority of antibiotic usage in the US has been on livestock, with over 13,000,000 kg – about 80% of total US consumption – being used to prevent disease (Chang *et al.*, 2015).

The selective pressure is further increased by routine administration, increasing the chance of antimicrobial resistance (AMR) becoming a stable trait, leading to AMR related diseases and further transfer of AMR genes to different environments. Along with an increase in global travel, AMR has become not only more difficult to treat but also to contain. This is expanded in the next section, which briefly describes the progression of AMR. It also emphasises why novel antimicrobials are desperately needed, especially for nosocomial infections in immunocompromised patients (Terra *et al.*, 2018).

1.2 Rise of antimicrobial resistance

1.2.1 Origins of AMR genes

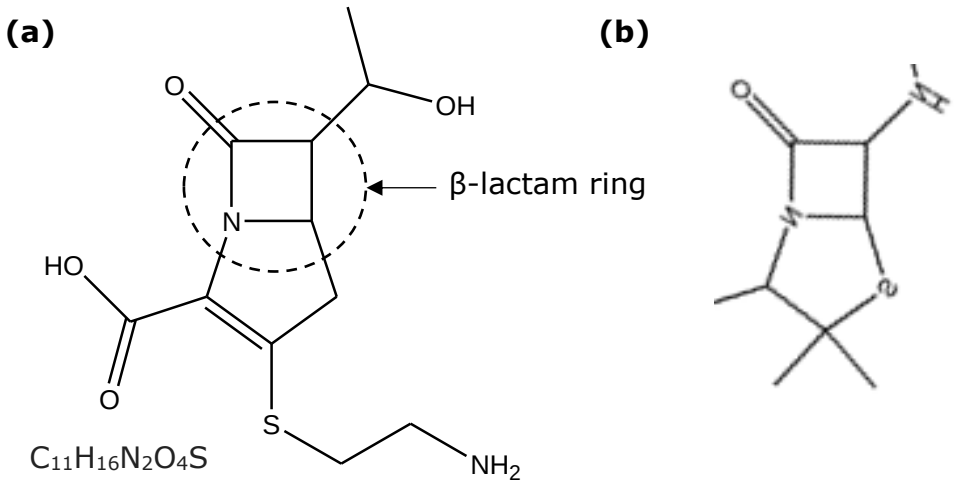
AMR is not a new phenomenon. Some form of resistance must have existed as long as antibiotics themselves have been produced, since microbes producing bioactive SMet must have been able to prevent any damage being done to their own cells. Genes for streptomycin resistance have been found in Siberian permafrost approximately 3,000,000 years old (Mindlin *et al.*, 2008). These genes (*strA, strB*, and *aadA*) were not found in permafrost-preserved *S. griseus* or any other *Streptomyces*, but in a mixed population of different genera; the Gram-negative species included *Pseudomonas* sp. and *Acinetobacter* sp., both of which are on the World Health Organisation's (WHO) list of pathogens most needing new treatments (Tacconelli *et al.*, 2018). In *S. griseus strA* and *strB* produce self-resistance by inactivating intracellular streptomycin through the enzyme streptomycin 6-phosphotransferase. This accounts for 1 of 4 main mechanisms for AMR: drug inactivation, altering the target structure, blocking drug uptake, and efflux pumps.

β-lactams, such as penicillin, are inactivated by the β-lactamases. β-lactams are generally associated with fungi, and were originally thought to not require any self-resistance mechanisms since they are not active against fungi (Ogawara, 1981). However, *Streptomyces* produces β-lactam antibiotics such as the carbapenem thienamycin, isolated from *S. cattyleya* (Ogawara, 2016), or the cephamycin family (Stapley *et al.*, 1972). Carbapenems differ from the subclass penicillins belong to (penams) through the substitution of a sulphur to carbon, and an unsaturated 5-carbon ring, hence "**carb**apen**e**ms" (Figure 1.9).

Though β-lactams are primarily active against Gram-positive bacteria, *Streptomyces* spp. are highly resistant to β-lactams due to penicillin-binding proteins which facilitate β-lactam capture inside the cell for release outside it (Ogawara, 2016). In comparison, Gram-negative pathogens use β-lactamases to destroy the lactam ring and bioactivity, which would seem to be totally separate mechanisms of action. However, β-lactamases are thought to have evolved from penicillin-binding proteins (PBP) and have some sequence homology (Meroueh *et*

al., 2003), presenting a 2nd class of AMR which may have its origins in *Streptomyces.*

Figure 1.9 Structure of **(a)** thienamycin with **(b)** core penam structure of penicillin for comparison.



Streptomyces species are generally not pathogenic to humans, though the genus does include notable plant pathogens such as the cause of potato scab, *S. scabies* (Seipke & Loria, 2009). *S. somaliensis* and *S. sudanensis* (Seipke *et al.*, 2012) are the main exceptions in humans, with recent unprecedented instances of *S. griseus* mycetoma (Chander *et al.*, 2015) and *S. atratus* pneumonia (Ariza-Prota *et al.*, 2015). Therefore, the problem is not so much the AMR-gene reservoir with *Streptomyces* but the mishandling of antibiotics leading to the selection of cassettes for multidrug resistance, and their inevitable spread.

1.2.2 Global Spread of AMR

1.2.2.1 ESKAPE pathogens

The most dangerous group of MDR-microbes are termed the "ESKAPE" pathogens: *Enterococcus faecium*, *Staphylococcus aureus, Klebsiella pneumoniae, Acinetobacter baumannii, Pseudomonas aeruginosa*, and *Enterobacter* spp. (H. Zhu *et al.*, 2014), especially dangerous when acquired as nosocomial infections. The ESKAPE list correlates with the WHO designations of

pathogens most urgently requiring new medicines, which includes all of the ESKAPE Gram-negative strains (Tacconelli *et al.*, 2018). From the 2 Gram-positive members, methicillin-resistant *Staphylococcus aureus* (MRSA) became an important nosocomial and eventually community-associated pathogen (Mediavilla *et al.*, 2012). Resistance in MRSA is due to the expression of another PBP, PBP 2a, derived from the *mecA* gene carried within the staphylococcal cassette chromosome *mec* (SCC*mec*) (Kwon *et al.*, 2005). SCC*mec* confers resistance to methicillin, a 2nd generation β-lactam, giving MRSA its name, along with resistance to all other β-lactams (Ender *et al.*, 2004). Resistance to other antibiotics can be acquired through the incorporation of different genes into SCC*mec* such as Tn*554* (Wielders *et al.*, 2002) which confers resistance to erythromycin and spectinomycin (Murphy *et al.*, 1985).

Resistance can also occur through a non-specific efflux pump or preventing diffusion through a porin. However, much of resistance is due to the inactivation of the frontline drug class, the carbapenems, by β-lactamases. Biofilms are also a major contributor to AMR for the ESKAPE pathogens, but this section will be restricted to the genetic basis of it (Eze *et al.*, 2018)

The major set of recent resistance enzyme genes in *K. pneumoniae*, and the *Enterobacteriacae* as a whole, confer resistance to carbapenems along with other drugs. MDR-Carbapenem resistance in *K. pneumoniae* can be acquired through a variety of genes, recently including the ^{bla}NDM-1 gene, standing for New Delhi metallo-β-lactamase after its discovery location (Kumarasamy *et al.*, 2010). Its product, NDM-1, hydrolyses all β-lactams apart from aztreonam. ^{bla}NDM-1 is horizontally transferred with 2 other genomic regions on the plasmid in conjugation, which together provide resistance against all antibiotics aside from fluoroquinolones and colistin (Yong *et al.*, 2009).

Since its discovery, NDM-1 producing strains have been detected across the world, including the UK. This strain was only treatable using colistin and tigecycline, a tetracycline derivative (Kumarasamy *et al.*, 2010). Colistin resistance, acquired through a mutated *mgrB* gene which modifies the lipopolysaccharides in the outer membrane, were recently reported in India (Pragasam *et al.*, 2016), indicating that a microbe resistant to all the antibiotics of last resort is a definite possibility.

A. baumanii infections are typically associated with already highly immunocompromised hospitalised patients in intensive care, surgical patients, or those exposed to contaminated medical devices (Kumburu *et al.*, 2019). As a predominantly nosocomial infection, MDR strains are even more deadly to these groups, especially if general hygiene procedures are not followed. It is also naturally competent, readily taking up exogenous DNA, increasing the likelihood of novel virulence or resistance mechanisms, of which β-lactamases are the most common mechanism to date (Lee *et al.*, 2017)

A. baumannii resists carbapenems through a class of enzymes termed oxacillin-hydrolysing, or "OXA", primarily through OXA-23, OXA-40, and OXA-58 though many other OXA-variants exist; the number refers to the order in which they were discovered and not any specific similarity (Evans & Amyes, 2014). Carbapenemases have significant structural homology with oxacillinases, so are classed as a subset of oxacillinase. Despite this, *A. baumannii* carbapenemases are technically inefficient at hydrolysing oxacillins (Santillana *et al.*, 2007).

MDR genes and cassettes are known to spread throughout the ESKAPE pathogens – for example, between *Pseudomonas aeruginosa* and *A. baumanii* (DaSilva & Domingues, 2016) – increasing the dangers of widespread AMR.

1.2.2.2 Factors Contributing to Widespread AMR

The rise of AMR is intrinsically linked to the misuse of antibiotics over a range of issues, with each country having its own set of issues that ultimately make AMR a global problem. These can be a complex mixture of unnecessary prescriptions or over the counter antibiotic purchases, large volume usage in agriculture, failure to complete a course of medication, and waste drugs leaching into the environment. Examples of some of these are given below, but not all the literature is necessarily negative: Finland lowered the incidence of erythromycin resistant *Streptococcus pyogenes* infections from 9.2% to 7.5% over 1997 to 2000 with a decrease in overall macrolide consumption (Bergman *et al.*, 2004), so positive change is realistic.

Adopting a One Health approach using combined policy across multiple areas is an essential strategy (Zinsstag *et al.*, 2011), of which new antibiotics are a vital part. Antimicrobial stewardship in medicine is often the focus of public

engagement or prescription regulation efforts, but medical uses are not the main usage of antibiotics in the US. An estimated 70% of antibiotics are bought by the agricultural sector as preventatives to promote growth (O'Neill, 2016). In 1996, with just streptomycin and oxytetracycline, this amounted to 136,000 kg of antibiotic (Fair & Tor, 2014). This further leaches into the soil, expanding the microbial resistome beyond the already present AMR-genes.

The implications of unnecessary prescriptions and over the counter purchases can be illustrated by examples from the United States and South Asia. Each year in the US, 40 million people are diagnosed with respiratory diseases. Of those, 27 million are given antibiotics unnecessarily (O'Neill, 2016), either selecting for multidrug resistant (MDR) strains within the patient's microbiome or in nearby waterways as they are discarded; this is at least a decrease from 34 million inappropriate antibiotic prescriptions in 2010-11 (Fleming-Dutra *et al.*, 2016). More broadly, across all emergency care visits, Fleming-Dutra *et al*. (2016) estimate that up to 30% of oral antibiotics were incorrectly prescribed, totalling 46 million cases, likely contributing to the 23,000 deaths a year in the US due to MDR pathogens.

Many countries in South Asia have minimal regulation on access to antibiotics, leading to it becoming a reservoir for MDR strains. Chung The *et al.* (2016) studied the prevalence of resistance to ciprofloxacin in treatment of infections by *Shigella sonnei*, which causes diarrhoea through damaging the gastrointestinal mucosa. They found an MDR strain of *S. sonnei* distinct from other resistant strains which had achieved global dominance over other strains. This was clearly linked to rates of over the counter purchasing of antibiotics from pharmacies in south Asia: Nepal & Bhatta (2018) show that ciprofloxacin was the most frequently purchased fluoroquinolone in South East Asian pharmacies, and is used by over 40% of the self-medicating population.

In contrast, rates in Europe range from 1% to 4%, partially due to stricter restrictions on what pharmacies can sell. While rates may vary between countries, the global nature of travel means that AMR genes can easily be spread across the world. ESBL-containing *Enterobacteriaceae* are often found in popular tourist areas, with up to 67% of visitors to Southeast Asia and 88% to India becoming colonised (if not infected) by these strains (Frost *et al.*, 2019).

Fortunately colonisation is not permanent and will often disappear after 3 months (Arcilla *et al.*, 2017), but this may be long enough for a hospital or other vulnerable community outbreak or to transfer new resistance genes to the established microbiota.

The prevalence of resistance genes in the environment has been illustrated by multiple studies. An examination of 480 soil strains by D'Costa *et al.* (2006) found widespread resistance genes for a set of 11 clinically important natural and synthetic antibiotics. Daptomycin resistance was the most common, with 100% of the strains being resistant and 80% of a further screen of 80 strains able to inactivate daptomycin on top of their resistance to it. Resistance was not only found to natural antibiotics, as 294 of the 480 were resistant to synercid, a semisynthetic combination of two *Streptomyces* metabolites which target different sections of the 50S ribosomal subunit.

Wastewater treatment plants are common reservoirs for AMR as they provide a link between human antibiotic waste in the influent and the wider environment in the effluent, facilitating the dispersal of resistant strains (Marti *et al.*, 2013). Marti *et al.* (2013) studied one plant on the Ter river in Spain, recording antibiotic concentrations and AMR gene prevalence upstream and downstream of the plant. Their results showed 0.91 μg/ml of ciprofloxacin and 0.18 μg/ml clarithromycin in the influent, down to 0.23 μg/ml ciprofloxacin and 0.16 μg/ml clarithromycin in the effluent. Comparing AMR gene levels in biofilms upstream and downstream of the plant showed lower levels of the β-lactamase genes ^{bla}TEM and ^{bla}SHV downstream of the plant, with minimal effect on AMR gene prevalence in sediment.

Chen *et al.* (2018) studied the prevalence of AMR genes in response to β-lactam and aminoglycoside overusage in Chinese aquaculture. The majority of resistance genes – 77% – were for resistance against a single class of antibiotics, but approximately 85% of resistance plasmids included at least 3 types of AMR genes, resulting in strains resistant to multiple antibiotic classes. Since aquaculture farm dosages can use multiple classes of antibiotics at a time, this likely contributes to the formation of multidrug resistance cassettes on single plasmids.

A solution to AMR will have to involve both the healthcare and industrial sectors. Initiatives to increase good antimicrobial stewardship across hospitals, nursing homes (Barlam *et al.*, 2016) and pharmacies (Wickens *et al.*, 2013) are of undeniable importance, but do not help develop new therapies. Although there was initial substantial financial investment into antimicrobial SMet discovery, much of this lessened following the decrease in discovery rates (Woodruff, 2014). Though the global market for antibiotic sales is $40 billion/year, $35.3 billion is from unpatented antibiotics (O'Neill, 2016), removing much incentive for investing in early stage discovery of novel unique metabolites.

New antibiotics are clearly needed to help combat AMR. *Streptomyces*, and other *Actinomycetales* genera, are the most prolific producers of antimicrobial SMet, with a wide variety of biosynthetic mechanisms. Understanding which *Streptomyces* characteristics contribute to its SMet-producing capabilities is essential for the discovery of novel SMet and the development of new therapies.

1.3 *Streptomyces* and specialised metabolism

1.3.1 Characteristic *Streptomyces* biology

The family *Streptomycetaceae* was proposed by Waksman & Henrici (1943) originally containing two genera, *Streptomyces* and *Micromonospora*. *Streptomyces* are Gram-positive filamentous microbes, typically with a mycelia diameter of 0.7 µm (Manteca, Fernandez, *et al.*, 2005). *Streptomyces* spp. are primarily known for their SMet production and chemotherapeutic potential, but this is underpinned by a variety of biological traits rarely found in other bacteria, such as multicellularity, morphological differentiation, and even cannibalism through programmed cell death (PCD) (González-Pastor *et al.*, 2003; Manteca, Fernandez, *et al.*, 2005; van Dissel *et al.*, 2014). These characteristics, and some of their relevance to SMet, are summarised in Table 1.2.

Table 1.2 *Streptomyces* characteristics and their contribution to SMet.

***Streptomyces* characteristic**	**Contribution to SMet**	**Reference**
Large genome size	Genome usually 8 to 12 MB, provides raw material for SMet evolution	Harrison & Studholme (2014)
Linear genome	Core stable region flanked by genetically diverse arms which contain most SMet genes	van Keulen (2014)
Multicellular	Different cells produce different compounds within one colony	Yagüe *et al.* (2012)
Multistage differentiation	SM production varies through stages, so culture metabolome varies over time	Yagüe *et al.* (2012)
Niche diversity	Different niches select for production of different metabolites	Rateb *et al.* (2011); Niu *et al.* (2016)

Some of these characteristics are shared with other *Actinomycetales* genera such as *Nocardia*, *Micromonospora* or *Amycolatopsis*, which are also known for their SMet (Tanaka *et al.*, 2013; Barka *et al.*, 2016; Holmes *et al.*, 2016). However, as stated, historically *Streptomyces* have been identified as the most prolific producers of SMet. While this may be due to their relative overabundance from environmental isolates (Tiwari & Gupta, 2013), screening over 700 *Actinomycetales* genomes for their BGC content showed *Streptomyces* to have the largest amount of BGCs/genome at a mean of 21.6 BGCs (Doroghazi *et al.*, 2014). Therefore, within *Actinomycetales*, *Streptomyces* retains the most potential for antibiotic discovery projects.

While characteristic of *Streptomyces*, the individual traits are not necessarily relevant for SMet. A linear genome is not a predictor of bioactive metabolite

production, as the Gram-negative order *Myxococcales* contains multiple SMet producers – notably *Sorangium cellulosum* and *Myxococcus* spp. (Gerth *et al.*, 2003) – which have a conventional circular genome (Garcia *et al.*, 2009). The shared traits between the two orders (multicellular, predominantly soil dwelling, multistage differentiation, large GC rich genome) do reinforce that the other characteristics are important for SMet producers (Zaburannyi *et al.*, 2016). The linear architecture may be important within the phylum *Actinobacteria*: essential primary metabolism gene synteny exists across both SMet producers and non-producers, such as the pathogen *M. tuberculosis*, and so the genomic architecture could potentially lead to more rapid BGC evolution on the genome terminal regions (Hopwood, 2006).

The list of characteristics most relevant to SMet production are multicellularity, multistage differentiation, and niche variety, since all of those will require different compounds by necessity to cope with extreme growth conditions. Niche variety is not an essential part of the *Streptomyces* lifecycle. The ability to quickly mutate and evolve novel chemistry aids *Streptomyces* survival in a range of standard soil to extreme environments (Pathom-aree *et al.*, 2006; Rateb *et al.*, 2011; Terra *et al.*, 2018), but is more a consequence of its characteristics not a reason for its SMet variety. In contrast, differentiation expands SMet production by requiring a different set of SMet at each lifecycle stage, and multicellularity enables differential expression of SMet over different cells. As these characteristics are more closely linked to SMet production, they will be focused on in the next section.

1.3.2 *Streptomyces* lifecycle

1.3.2.1 Lifecycle on solid media

The majority of work on *Streptomyces* lifecycle and differentiation has been achieved on solid agar media. This is outlined in Figure 1.10, and is broadly defined by 4 main stages: spore gemination, growth of primary compartmentalised mycelium ("MI") through the substrate, multinucleate secondary compartmentalised mycelium ("MII") which are the main SMet producers, and spore formation (Yagüe *et al.*, 2012).

After contact with a suitable location with enough resources to fuel germination, a spore sends out a germ tube, potentially 1 from each tip of the spore (Čihák *et al.*, 2017). The tube grows by tip extension, incorporating peptidoglycan into the apex (Sigle *et al.*, 2016), eventually forming branches and becoming the substrate MI hyphae (Bush *et al.*, 2015).

The MI hyphae forms cross-membranes producing compartments approximately 1 μm wide, giving an early uninucleate MI hyphae which is more frequently protoplasmic than later stages (Yagüe *et al.*, 2016). A round of PCD then occurs, producing a variegated dead/alive sequence of cells along the hyphae with the dead cells acting as fuel for the remaining viable segments to transition to multinucleate MII hyphae (Manteca, Fernandez, *et al.*, 2005).

The MII hyphae are the main producer of SMet. Transcriptomic analysis of *S. coelicolor* A3(2) MI and MII phases showed a distinct correlation of SMet transcripts to MII. Transcripts for the blue pigment actinorhodin (ACT), undecylprodigiosin (RED), calcium dependent antibiotic, geosmin, and the cryptic polyketide cpk were all associated with the MII hyphae (Yagüe *et al.*, 2013). This stage can be further separated into two sub-categories: the early substrate MII, which lacks a hydrophobic layer, and the late aerial hyphae, which expresses the hydrophobic rodlins and chaplins (Manteca *et al.*, 2007). The 8 chaplins (ChpA – H) form amyloid-like fibrils woven into the cell walls, which are sorted into the mature rodlet layer by rodlins RdlA and B, both of which are necessary to fully complete the lifecycle (Claessen *et al.*, 2004). Production of these, and the aerial hyphae, is fuelled by a 2nd round of PCD, after which the multigenomic aerial hyphae is compartmentalised into unigenomic pre-spores (Yagüe *et al.*, 2013).

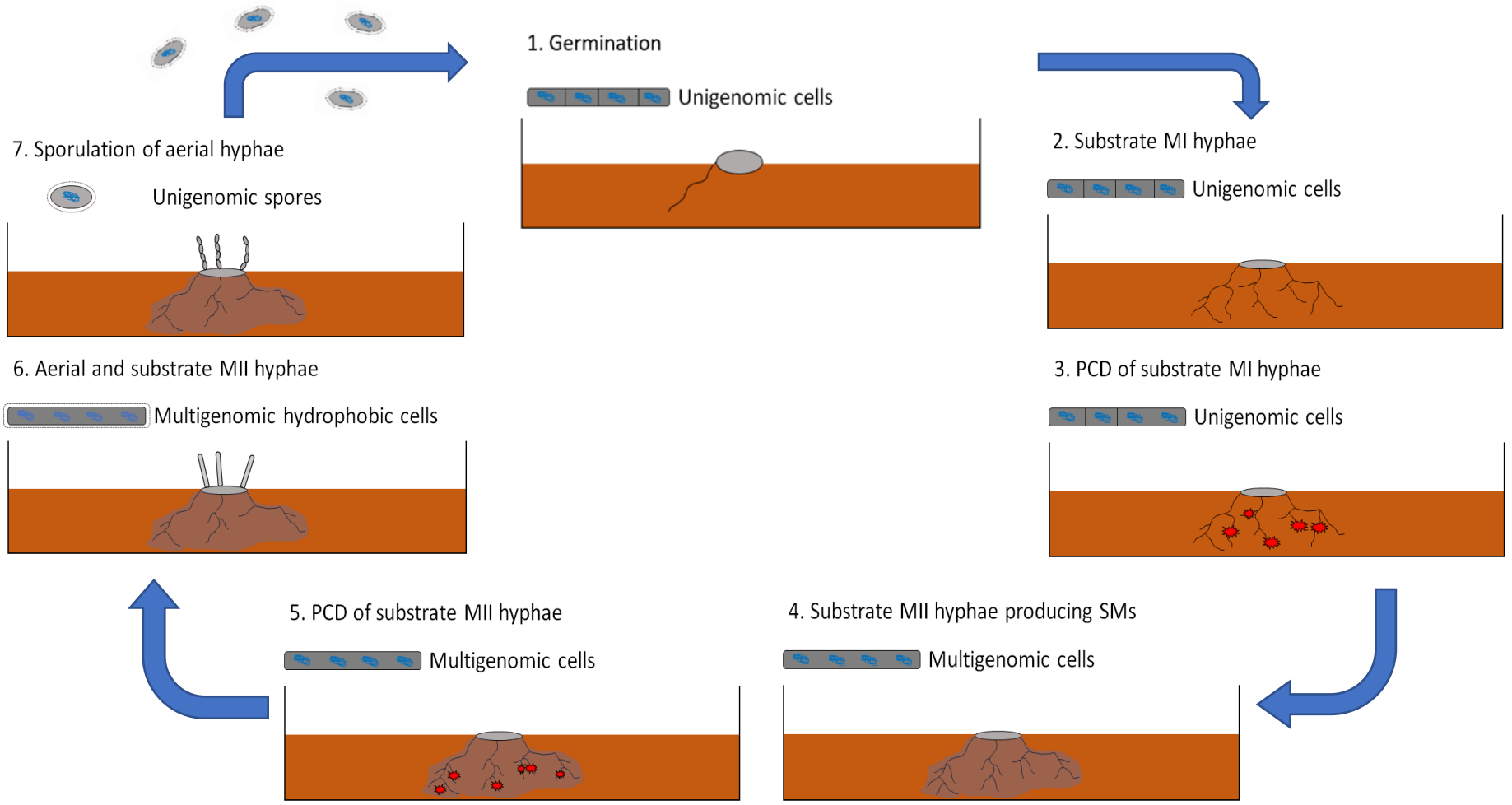
After maturation, the spores are spread by an external force to other suitable media, where the cycle restarts. In case an appropriate location is not immediately available, the spores are able to go into a dormant state with low to no levels of metabolic activity (Bobek *et al.*, 2017) until conditions are more favourable for successful germination. Spore germination may not totally depend on external forces like nutrition levels as levels of the cell wall protein NepA can also affect germination levels. De Jong *et al.* (2009) knocked out the gene for this putative dormancy regulator increasing germination from 6% to

19% of total spores after 3 h, reaching 100% germination 4 h before the wild type (10 h against 14 h). Spores also contain elevated levels of the sugar trehalose, which makes up to 25% of the dry weight of *S. griseus* spores (McBride & Ensign, 1987) and is potentially used as an initial energy source in suboptimal conditions.

In total, *Streptomyces* undergo a hugely complex set of multicellular differentiation and programmed cell death processes, though it is unclear if these are actually necessary for its SMet productivity. One potential solution is the tradeoff in resources required between maintaining biomass and SMet production. If sporulation occurs too early in a strategy to escape adverse conditions, less total biomass may be produced resulting in the strain being outcompeted anyway; on the other hand, overproducing SMet to defend its resources may not leave enough for sporulation.

PCD could allow relative overproduction and defence, followed by PCD to free up more resources for sporulation. This is reinforced by *Streptomyces* moving relatively slowly compared to other species, except in rare cases (Jones *et al.*, 2017), which limits its access to nutrients. Preventing other organisms from using the available resources is therefore more important, hence why overproduction of antimicrobials is desirable. While uncommon across known prokaryotes, PCD has been recorded in other soil species. Once stressful conditions trigger sporulation, a *Bacillus subtilis* cell will secrete sporulation-blocking metabolites to block development of other *B. subtilis* cells around it, eventually lysing them to use the additional nutrients to fuel sporulation (González-Pastor *et al.*, 2003).

Figure 1.10 *Streptomyces* lifecyle showing main stages of differentiation on agar: germination, MI and MII PCD, aerial hyphae and sporulation.



1.3.2.2 Genetic regulation of morphology and metabolism

Expression of SMet by *Streptomyces* is tightly regulated and can correlate with morphological development. Most of the research into *Streptomyces* lifecycle and genetics has been done using the model species *S. coelicolor* A3(2), from which 2 main mutant phenotypes have been generated: bald mutants, which lack aerial hyphae and affect the *bld* genes, and white mutants, which affect the *whi* gene family and cannot produce mature spores (Yagüe *et al.*, 2012; Hackl & Bechthold, 2015). Each of these *S. coelicolor* A3(2) mutant types has different specialised metabolomes, reflecting different stages of morphological development. Both gene families are also known to code for sigma factors which redirect RNA polymerase to specific genes, or act as pleiotropic and specific regulators of SMet (Manteca *et al.*, 2010; Bush *et al.*, 2015).

Both *whi* and *bld* genes are conserved in all known *Streptomyces* as of 2015 (Kim *et al.*, 2015). Therefore, results obtained with *S. coelicolor* A3(2) are unlikely to be limited to just *S. coelicolor* A3(2), so examples of mutants will be given for both genes.

1.3.2.3 Effect of *whi* mutants on specialised metabolism

Mutagenesis of *whiA* forces hyphae to take a tightly coiled spring-like form lacking the characteristic grey polyketide SMet associated with *S. coelicolor* A3(2) mature spores. Normally, wild-type spores are compartmentalised into 1 µm segments by septa which split the multigenomic aerial hyphae into unigenomic spores (Sigle *et al.*, 2016). These septa are not seen in *whiA* mutants, producing hyphae longer than in the wild-type at 100 µm long (Findlay *et al.*, 1999). Similar phenotypes are seen from other *whi* mutants (*whiG, whiB, whiH)*, all unpigmented, and varying in the intensity of hyphal coiling and septa loss.

Production of the grey pigment in *S. coelicolor* A3(2) is controlled by *whiE* (Kelemen *et al.*, 1998). This gene is conserved in other *Streptomyces*, though the pigment colour can vary – *S. halstedii* mature spores are green, not grey. *whiE* is under the regulation of multiple other genes and their products. One of

these is *sigF*, which codes for the sigma factor σF (Kelemen *et al.*, 1998). Negative *sigF* mutants are not totally devoid of pigment, as some parts of the *whi* cluster remain active, specifically the *whiEP1* promoter. *whiEP2* is made inactive, preventing expression of *whiE*-ORFVII which results in a brown spore pigment. σF therefore likely regulates part of modification of the pigment's chemical structure instead of its production. Which genes are ultimately responsible for the full regulation of *whiE* is unknown, as well as the ecological function of the pigment (Salerno *et al.*, 2013). One suggestion by Chater *et al.* (2010) is that it provides protection against a range of stresses, whether abiotic like heat and radiation, or biological ones like digestive enzymes from predatory microbes.

1.3.2.4 Effect of *bld* mutants on specialised metabolism

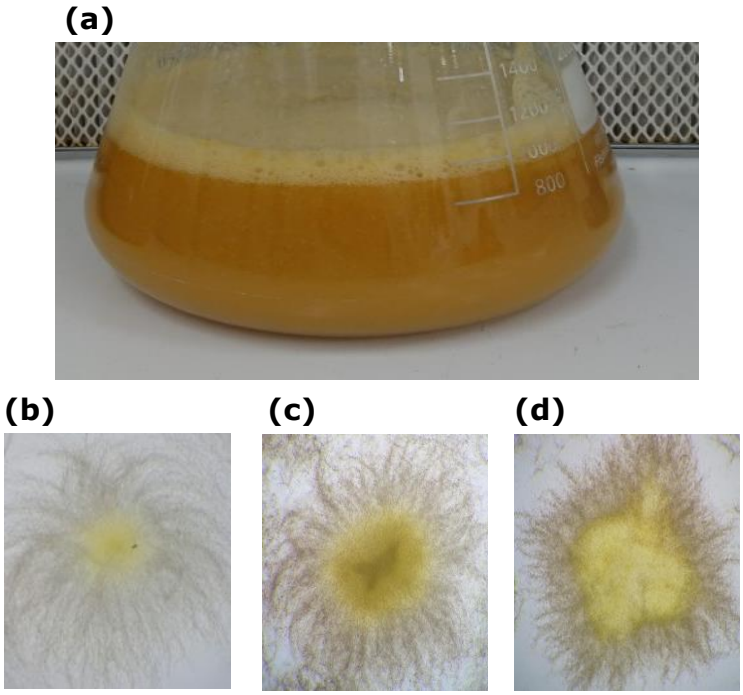
bld mutants are affected earlier in the morphological cycle than *whi*, so it could be assumed that the effect on SMet would be more inhibitory. This is the case for the *bldD* global regulator, which if mutagenised, pleiotropically affects SMet such that none of the characteristic antibiotics expected to be produced by *S. coliecolor* A3(2) are detected (Elliot *et al.*, 2003). As a global regulator, *bldD* affects the expression of many genes: its product, the 18 kDa DNA binding protein BldD, was found to control around 170 transcriptional genes (den Hengst *et al.*, 2010).

Within those, 42 genes coded for DNA-binding regulators, with known targets including genes for morphological differentiation and antibiotic production. These included *cvn9* – ACT regulation - and *nsdA*. In the wild type, BldD binds upstream of *nsdA*, preventing it from being transcribed. If *nsdA* is knocked out then *S. coelicolor* displays increased sporulation and overproduction of ACT and 2 other antibiotics, methylenomycin and the Calcium Dependent Antibiotic (Li *et al.*, 2006). One other gene found to be controlled by BldD was *bldC*, which codes for a DNA binding protein which also controls morphology and production of ACT (Hunt *et al.*, 2005).

1.3.2.5 Differentiation in liquid media and soils

A broth lifecycle is not discussed in detail as, unlike on agar, most species do not undergo full multicellular differentiation in broth culture (Yagüe *et al.*, 2013; Manteca & Yagüe, 2018). The MI hyphae do differentiate to MII as evidenced by PCD and SMet production at the centre of bacterial pellets. Examples of these dense pellets with central SMet production after 72 h in *S. costaricanus* (pellet diameter approximately 0.1 mm) are shown in Figure 1.11. There are exceptions, notably *S. venezuelae* which does fully sporulate in liquid media, but the majority of species do not display the full lifecycle occurring on agar (Glazebrook *et al.*, 1990).

Figure 1.11 **(a)** *S. costaricanus* 1 L culture at 72 h displaying pellet formation and **(b, c, d)** pigment production within pellets at 40x magnification.



Less work has also been done on *Streptomyces* development in more soil-like, or natural, conditions. The literature indicates that the MI hyphae are the dominant hyphal type, unlike in the lab where MII hyphae lasts the longest, and no sporulation was seen after 30 days (Manteca & Sanchez, 2009). As no PCD nor sporulation was observed, it is possible that other nutritional or biological stimuli are needed to induce these in the wild.

The slow growth would also make this unsuitable for a generalised SMet discovery project unless genetic modification of natural condition-specific SMet genes was incorporated. Many methods exist for SMet discovery and choosing the correct one for a project is essential to its success.

1.4 Antimicrobial discovery using *Streptomyces*

1.4.1 Production potential of unexploited *Streptomyces*

As the golden age of antibiotic discovery ended and synthetic drug libraries began to receive more attention, it was assumed that most *Streptomyces* metabolites had already been discovered. It had been known for many decades that culture parameters affect SMet production (Waksman & Tishler, 1941; Schatz *et al.*, 1944), but there was no evidence to suggest that *Streptomyces* was definitively able to produce other undiscovered antibiotics. This changed with the full genome sequence of *S. coelicolor* A3(2), which showed many more potential antibiotic BGCs within the genome than had been seen in the lab (Bentley *et al.*, 2002); the same proved to be true for other *Streptomyces*, with typically about 30 BGCs in total (Baltz, 2016). The vast majority of these BGCs are either unexpressed or weakly expressed in the lab, and so termed "cryptic" or "silent" BGCs (Moore *et al.*, 2012). Finding the right conditions to elicit expression of all of these BGCs is crucial for microbial drug discovery.

This means that all the previously examined *Streptomyces* used in the initial surge of research, plus all those isolated since, have huge unexploited potential to contribute novel bioactive compounds and scaffolds in the fight against AMR. Following their initial analysis, these predominantly soil *Streptomyces, Actinomycetales* and other microbes have been deposited in culture collections. This corresponds to a huge amount of potential metabolomic diversity in a regulated and accessible collection. Therefore, the simplest way to access a huge amount of niche diversity and the corresponding chemical space is to use the less researched strains in culture collections. One collection is NCIMB Ltd., which contains approximately 500 strains of *Streptomyces* in its collection (NCIMB Ltd, 2017).

1.4.2 Strategies to discover new metabolites from *Streptomyces*

There is a huge diversity of potential methods in SMet discovery from *Streptomyces*, each suited to specific projects. The major ones are listed below in Table 1.3 (genetic methods) and Table 1.4 (culture dependent). Other methods do exist, such as engineering the biosynthetic machinery by mutating the ribosome and RNA Polymerase (Ochi *et al.*, 2004; Shentu *et al.*, 2016), but this is arguably more suited to product optimisation (Kawai *et al.*, 2007; Wang *et al.*, 2008) than discovery so will not be discussed further. Metagenomic methods are also not included although they are valuable for discovering novel BGCs without having to rely on culture (Reen *et al.*, 2015) as this is unsuited to a culture collection based project.

Table 1.3 Genetic manipulation strategies for metabolite discovery, primarily in *Streptomyces*.

Method	Advantages	Disadvantages	Source
Alters a BGC's regulatory elements to control expression	Single product with known regulation	Identification of regulatory elements needed, just one product	Bunet *et al.* (2011); Luo *et al.* (2013)
Modifies pleiotropic regulators to affect multiple BGCs	Expression of multiple BGCs using genus-conserved regulators	Relies on known regulators, little predictive ability	McKenzie *et al.* (2010); Becerril *et al.* (2018)
Links BGC to a reporter, indicating activation	Simplifies detection, high-throughput mutagenesis	BGC/reporter may not combine properly	Guo *et al.* (2015); Xiang *et al.* (2009)
Cloning and heterologous expression of specific BGC	Chosen production host, simplifies detection	BGC may not clone and express, size limits	Lin *et al.* (2006); Jiang *et al.* (2012)

Table 1.4 Culture dependent strategies for metabolite discovery, primarily in *Streptomyces*.

Method	Advantages	Disadvantages	Source
Pleiotropic regulation via altering nutrition	Applicable to all species, many growth conditions	No predictive power, complex results	Bode *et al.* (2002); Rateb *et al.* (2011)
Trigger BGC production with stringent response	Variety of stresses, applicable to all species	May damage organism too much, little initial predictive power	Sevcikova *et al.* (2004); Jakeman *et al.* (2006)
Co-culture to elicit ecology-based reactions to invasion or mutualism	Potentially specific activity against eliciting microbe, can use co-isolated strains	Other microbes may increase extract complexity or have no effect on SMet production	Onaka *et al.* (2011); Shin *et al.* (2018)

As the *Streptomyces* strains used will be previously underexploited, they are unlikely to have any significant genetic manipulation tools developed nor a full genome sequence to identify BGCs. This makes the genetic tools initially unsuitable, though a full genome sequence is desirable to combine SMet genomics and metabolomics. This leaves the culture dependent methods, which are more applicable to a library of ecologically diverse strains, with the potential to elicit a different set of SMet with each culture condition.

This set of methods is typified through the One Strain Many Compounds method or OSMAC, briefly mentioned in section 1.1.2.2. Varying culture parameters has long been known to produce different levels of SMet as evidenced by its usage in early optimisation of actinomycin, and it was not fully codified as a method for cryptic BGC activation until Bode *et al.* (2002). The authors identified 20 SMet from a single strain, covering many types of chemical scaffolds and natural product class. Different culture conditions included temperature, pH, type of vessel, oxygenation, pressure, and minerals; ideally, being able to mimic the ecological conditions a strain was isolated from would be a good application of OSMAC to identify any ecological cues that drive SMet production. Of these

conditions, changing the nutrients in media is the easiest with potentially pleiotropic effects (Rateb *et al.*, 2011), and avoids the complications of co-culture.

1.4.2.1 Carbon source regulation of specialised metabolism

The simplest way to vary culture conditions is through the carbon and nitrogen sources. Carbon catabolite repression (CCR) is a well-studied feature of SMet, where one carbon source can lead to repression of some BGCs and expression of others until that carbon source has been exhausted (Romero-Rodríguez *et al.*, 2016). Glucose is normally repressive, since it is an excellent energy source which indicates that nutrient levels are high, making spending resources on SMet an uncompetitive strategy as any bioactive compounds may be unable to repress growth levels of other microbes while restricting growth of the producer (Ruiz *et al.*, 2010). Other less preferred carbon sources will, since they are not as immediately energetically rich, be either neutral or upregulate expression (Sánchez *et al.*, 2010).

As one example, wild-type *S. lividans* is known to contain the BGC for ACT but does not normally express it under standard lab conditions (Kim *et al.*, 2001). Both rich and minimal glucose containing media do not allow production of ACT but using glycerol as a carbon source induced expression through upregulation of *afsR2* mRNA. Its product, AfsR2, is a pleiotropic regulator of specialised metabolism through its upregulation of the enzyme guanosine pentaphosphate synthetase/polyribonucleotide nucleotidyltransferase, which is involved in the synthesis of the alarmone (p)ppGpp (Kim *et al.*, 2005). (p)ppGpp is partially responsible for stress response in bacteria, and correlates with the induction of SMet and differentiation (Hesketh *et al.*, 2007; Wang *et al.*, 2008).

SMet can also be upregulated through the addition of the carbon and nitrogen source N-acetyl glucosamine (GlcNAc), though this depends on whether it is added to rich or poor media (Rigali *et al.*, 2008). As a component of both bacterial and fungal cell walls (Konopka, 2012), its presence can indicate either breakdown of neighbouring hyphae and other microbes, or fungi invading the *Streptomyces'* space. Both of these are hostile conditions, but under rich media, GLcNAc represses SMet; under growth on poor media, SMet production is pleiotropically upregulated.

GlcNAc is transported into the cell via the sugar transferase system, which also phosphorylates it to GlcNAc-6-P. This is then acetylated by NagA to glucosamine-6-P (Rigali *et al.*, 2008), which is an allosteric effector of DasR, a global regulator. The binding of glucosamine-6-P to DasR releases it from the DNA, where it is located upstream of multiple antibiotic-linked genes. In *S. coelicolor* A3(2), expression of *act*II-ORF4 and *redZ*, which are involved in ACT and RED synthesis, are upregulated (Rigali *et al.*, 2008). The difference with rich media is thought to stem from the origin of GlcNAc, since it can be both from fungal (chitin degradation) or bacterial (peptidoglycan degradation) cells. It is suggested by Rigali *et al.* (2008) that under "feast" conditions where external chitin is abundant and useful as a food source, this implies sufficient external nutrients to continue growth without needing to produce SMet. Under "famine" conditions where cell wall hydrolysis is taking place, conditions are more stressful, so SMet production is upregulated.

The full mechanism of CCR is not fully known (Romero-Rodríguez *et al.*, 2016) across the genus, likely differing between species and carbon sources. In *S. coelicolor* A3(2), high levels of glucose represses at transcription enzymes for the alternate carbon sources galactose, fructose, arabinose and glycerol (Hodgson, 1982), potentially mediated through glucose kinase which converts glucose to glucose-6-phosphage. Glucose kinase mutants do not show glucose catabolite repression (McCormick & Flärdh, 2012), indicating an essential part in CCR, but the mechanism is also unknown. Glucose kinase does not contain a DNA binding region, so may interact with other regulators for its effects like AfsR2 (Liu *et al.*, 2013).

1.4.2.2 Phosphate and nitrogen regulation of specialised metabolism

Phosphate availability is monitored through the two-component system PhoR-PhoP. PhoR, a histidine kinase, is triggered by a lack of phosphate to phosphorylate PhoP (Santos-Beneit *et al.*, 2009; van Wezel & McDowall, 2011). PhoR is normally prevented from phosophorylating PhoP by the Pst phosphate transport system, which under phosphate rich conditions is bound to PhoP (van Wezel & McDowall, 2011). The phosphorylated PhoP then binds to areas of DNA known as PHO boxes, regulating transcription of those regions. High phosphate

is known to reduce production of many antibiotics, including jadomycin B in *S. venezuelae*, rapamycin in *S. hygroscopicus,* and nikkomycin in *S. tendae* (Hege-Treskatis *et al.*, 1992; Cheng *et al.*, 1995; Jakeman *et al.*, 2006). Phosphate limitation is therefore an important part of regulating SMet across *Streptomyces*.

Similar to carbon source regulation, any nitrogen source which is too effective at promoting growth is likely to repress SMet production since it will not promote a stress response. A well characterised nitrogen-reliant regulatory mechanism is the AfsQ two-component system, which is comprised of AfsQ1 and AfsQ2 (R Wang *et al.*, 2013). Upon growth in glutamate-containing media, the membrane bound AfsQ2 phosphorylates AfsQ1 which is then able to perform a variety of functions, including upregulation of antibiotic genes when bound to a conserved pair of 5 base pair sequences separated by a 6 base pair interval. The exact mechanism which triggers AsfQ2 to phosphorylate AsfQ1 in the presence of glucose is still unclear (Liu *et al.*, 2013).Upregulated antibiotic genes include *act*II-ORF4, *redZ* and *cdaR*, the last of which is involved in calcium dependent antibiotic synthesis.

AsfQ1 shares binding sites with the nitrogen regulator GlnR (R Wang *et al.*, 2013). The GlnR regulon includes at least 15 genes, mostly for enzymes, which facilitate nitrogen uptake and alternative source usage. Chromatin immunoprecipitation-microarray analysis (ChIP-chip) showed 36 GlnR binding sites, mostly upstream of other genes involved in nitrogen metabolism (He *et al.*, 2016). GlnR not only affects SMet by regulating primary metabolic precursors, but in some cases regulates SMet directly though fewer examples are known. The gene *jadR1,* part of the jadomycin BGC in *S. venezuelae*, contains a GlnR binding site predicted to activate transcription of the gene; GlnR was also found to bind to the commonly assayed ACT and RED activator genes *act*II-ORF4 and *redZ* (He *et al.*, 2016).

1.4.2.3 Stressors and elicitors

Nutrient variety is not the only abiotic factor *Streptomyces* must respond to, as all environments will present fluctuations in salinity, pH (Jiang *et al.*, 2018), metals (Morgenstern *et al.*, 2015), temperature, antibiotics from other microbes, and other challenges. These, and the survival responses *Streptomyces* exhibit, can function as signals which lead to the expression of cryptic BGCs that require

specific ecological conditions. This can lead to unexpected new functions for the signalling molecule, such as sub-inhibitory levels of antibiotics acting as transcriptional triggers of otherwise cryptic BGCs (Yim *et al.*, 2006).

This method has the advantage of being very easy to carry out: the addition of 3% dimethyl sulfoxide to media increased production of chloramphenicol and tetracenomycin C in *S. venezuelae,* and thiostrepton from *S. azureus* (Chen *et al.*, 2000). As these are all different classes of metabolite, this implies a global method of action for solvent-based induction, likely as a result of triggering the stringent response and increasing chaperone activity to protect from the denaturing effects of DMSO (Yoon & Nodwell, 2014). However, this approach is wholly reliant on luck so should not be the sole strategy.

Jadomycin B is also upregulated by adding a solvent, in this case 6% ethanol, but also through heat shock. Doull *et al.* (1993) found that raising the culture temperature from 27°C to 42°C increased jadomycin B from essentially 0 to 25 µg/ml. The jadomycin B BGC contains 2 well characterised pathway transcription factors, the activator *jadR1* and repressor *jadR2*, which negatively regulates *jadR1* through its product JadR2 (Yoon & Nodwell, 2014). Exactly how heat shock upregulates SMet in *Streptomyces* in general is unclear. It has been suggested that the chaperone DnaK, which would normally stabilise binding of the transcription factor HspR, would be removed by increased numbers of misfolded proteins after heat shock and allow HspR to initiate transcription of its target genes (Bucca *et al.*, 1997; Yoon & Nodwell, 2014). Unfortunately, there is no HspR recognition site in the jadomycin BGC, nor in another ethanol and heat shock activated BGC, validamycin A (Liao *et al.*, 2009).

Cryptic BGCs have also been elicited in response to the rare earth metal scandium, usually as scandium chloride. After initial proof of concept demonstration by upregulating ACT in *S. coelicolor* A3(2), scandium was found to increase expression of otherwise cryptic or very low expression clusters by 2.5 to 12 fold (Kawai *et al.*, 2007; Tanaka *et al.*, 2010). Similar to solvents, this is a simple method which is applicable to most strains but will require optimisation to avoid over or under stressing the culture. Unfortunately, the mechanism by which scandium activates *Streptomyces* is not known, and there are few other examples in the literature despite it being suggested as generally applicable.

Antibiotics, as microbial products, can also elicit a reaction without harmful inhibitory effects. The addition of lincomycin, produced by *S. lincolnensis*, to *S. lividans* culture resulted in multiple novel SMet which were not seen without lincomycin (Imai *et al.*, 2015). Levels of ACT, though not a cryptic BGC, were also increased in *S. coelicolor* by adding erythromycin. Whether antibiotics are actually produced in large enough concentrations to kill other microbes in the soil, and if they function both as messengers and antibiotics, is unfortunately outside the scope of this thesis. Despite that, this is a novel way to trigger SMet using an aspect of microbial ecology which could potentially be very effective when co-culturing microbes found from the same niche.

There are further methods, such as epigenetic modification (Kumar *et al.*, 2016) to allow RNA Polymerase access to the genome, salt shock (Kol *et al.*, 2010), pH shock (Kontro *et al.*, 2005; Jiang *et al.*, 2018) and messenger molecules like γ-Butyrolactones which both stimulate and inhibit a wide range of SMet (Demain, 1998; Liu *et al.*, 2013). While the large number of culture-dependent methods does give a huge amount of choice when working with an unexploited *Streptomyces*, they are much less useful if the metabolomic output cannot be precisely examined. Analytical natural product chemistry is a vital part of the drug discovery process, especially as samples have become faster to analyse in greater detail.

1.4.3 Analytical chemistry enhances specialised metabolite discovery

An ideal untargeted natural product metabolomics experiment will be able to immediately identify the known metabolites within an extract, assign putative structural analysis to unknowns, link together families of metabolites, and discard any irrelevant media ingredients. While not possible yet, this is becoming more likely through more advanced ultra-high pressure liquid chromatography-mass spectrometry (LC-MS) and nuclear magnetic resonance technologies, which are the primary tools in microbial SMet analysis (Rao *et al.*, 2013; Wu *et al.*, 2015; Wu *et al.*, 2016; Covington *et al.*, 2017).

LC-MS is a combined technique, using polarity-based column separation to untangle the mix of compounds within a sample. These sequentially elute into the mass spectrometer, where due to the separation they are ideally analysed

one at a time, producing an LC-MS chromatogram of all the metabolites within an extract (that were able to ionise). The whole process can be separated into 3 main blocks: extract preparation, chromatographic separation, and MS detection (Walker, 2013). These must remain consistent in order to compare any new results against previous ones, and so obtain a proper metabolite profile. An element of bias will be introduced based on the extraction, chromatography, and MS methods chosen so an experiment will not be truly untargeted (Krug & Müller, 2014), but this can be said for most analytical techniques.

Analytical chemistry is arguably the key difference in a modern untargeted metabolomics project to the older OSMAC efforts, since it facilitates the important task of preventing work on already known compounds, or dereplication (Zhang, 2005). If the aim is novel metabolites and an extract contains mostly known compounds then it can easily be discarded, and focus shifted to another culture condition or microbe. While this was also possible through large databases of natural products, these often require a subscription or lack mass spectral data. Recent attention has been on freely available mass spectral databases, designed to facilitate the metabolite discovery process in light of AMR challenges, increased computational power, and LC-MS advances.

The primary LC-MS/MS database, Global Natural Products Social molecular networking (GNPS) uses both donated user data and larger scale government spectral databases to dereplicate samples. While a metabolite not appearing in the database does not mean that it is novel, it can be used in combination with other tools to focus prioritisation efforts on specific extracts or metabolomic features.

1.5 Thesis aim

1.5.1 Biosynthetic potential of unexploited *Streptomyces*

Streptomyces has been the main workhorse for drug discovery projects. Continued improvements in genomic and metabolomic analysis have highlighted the SMet variety that many *Streptomyces* strains can potentially produce, so long as the correct culture parameters are used. Therefore, there is still much unexploited potential for *Streptomyces* to produce novel bioactive SMet, especially when selecting from a geographically and ecologically diverse culture

collection. Using an OSMAC screen is a rapid method to access this potential and is applicable to any of the strains in the collection.

SM diversity can be robustly assessed using UPLC-MS based metabolomics, as there are various bioinformatics tools which allow comparison of large metabolomics data sets to explore the variety within them. This can be used to choose scale-up conditions, link metabolites to specific elicitors, and calculate concentration of a metabolite. The combination of OSMAC and UPLC-MS should therefore provide information on the pleiotropic regulation of SMet, single elicitor-SMet links, and initial structural determination through mass spectra.

1.5.2 Selection of target for bioactivity assays

Despite the surge in Gram-negative AMR, as of 2017 only 31% of drugs in development are active against the Gram-negative ESKAPE pathogens, and many of these will not necessarily be active against the most resistant strains identified by the WHO (Simpkin *et al.*, 2017). Resolving this will need urgent widespread action on fundamental research combating the critical priority pathogens through multiple strategies, one of which is novel antibiotics. A recent assessment of the antibiotic pipeline found only 1 ongoing clinical trial against MDR *A. baumannii* (Hesterkamp, 2016). Therefore, *A. baumannii* and potentially the other ESKAPE Gram-negatives are excellent initial targets.

1.5.3 Primary aim

The hypothesis of this thesis is that *Streptomyces* culture collections are an underexploited source of novel specialised metabolites. The primary aim will therefore be to enhance specialised metabolite diversity by expanding the available chemical space from an underexploited *Streptomyces* strain, which can be used as a basis for bioactivity assays against *A. baumannii*. This will be achieved through 4 main objectives.

1) Development of a robust culture parameter-based screen to elicit production of multiple SMet, verified by testing on the model species *S. coelicolor* A3(2).

S. coelicolor A3(2) is the most widely researched *Streptomyces* strain for genomics and metabolomics studies (Bentley *et al.*, 2002; Liu *et al.*, 2013; Hoskisson & van Wezel, 2019). Therefore, there is a correspondingly large amount of published literature on its SMet, aiding identification of any detected compounds and so validation of the screen for non-model strains.

2) Application of the screen to underexploited *Streptomyces* strains to select a main strain and culture parameters.

The majority within NCIMB are soil isolates, so these will be studied first and then compared with non-soil isolate strains to ensure rational selection of a final underexploited strain. SMet identification will be more difficult than with *S. coelicolor* A3(2) due to the lack of published literature, but the usage of metabolomics software will aid separation of metabolites from media compounds.

3) Full genome sequencing of the strain to determine its SMet biosynthetic potential.

As an underexploited strain, there will likely not be a full genome sequence available at the start of the project. Therefore, acquisition and analysis of a novel high-quality sequence will allow prediction of its BGCs and comparison to other *Streptomyces* strains.

4) Culture scale-up, bioactivity assays and metabolite dereplication to inform isolation of bioactive SMet.

The final objective will use the cumulative strain and media selection data to produce enough culture for fraction-based purification and bioactivity assays. This will enable focused analysis on specific fractions and ideally isolation of novel bioactive SMet.

Together, these objectives link genomic and metabolomic data for a previously unexploited *Streptomyces* strain, developing a framework for the exploitation of further strains.

CHAPTER 2

METABOLOMIC SCREENING OF *S. COELICOLOR* A3(2)

2 Metabolomic screening of *S. coelicolor* A3(2)

2.1 Introduction

2.1.1 Specialised metabolism of the model species *S. coelicolor*

S. coelicolor A3(2) has been the most well studied species of *Streptomyces*, at least from a genetic perspective, for over 50 years (Kieser *et al.*, 1992; Rutledge & Challis, 2015; Hoskisson & van Wezel, 2019). This has contributed to its common use in the lab, eventually becoming the first *Streptomyces* to be fully sequenced (Bentley *et al.*, 2002). The specific strain sequenced was *S. coelicolor* A3(2) M145, which lacks the giant linear plasmid SCP1 and circular plasmid SCP2.

S. coelicolor is taxonomically more closely related to *S. violaceoruber* ISP5049 than the *S. coelicolor* Müller type strain. The difference between the two *S. coelicolor* strains has been noted for over 50 years (Kutzner & Waksman, 1959), but the name *S. coelicolor* has been maintained due to its use as a model species. Therefore, for this thesis, "*S. coelicolor*" refers to the model species *S. coelicolor* A3(2) and the type strain will be given as *S. coelicolor* Müller.

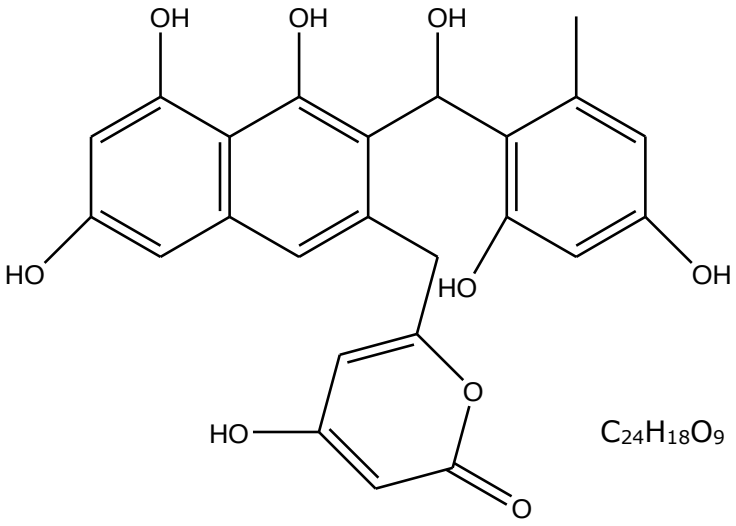
Much of the research on *S. coelicolor* has focused on unravelling the genetics of differentiation, sporulation, and how these processes affect specialised metabolism (SMet). Most *Streptomyces* will undergo the characteristic differentiation cycle of germination, vegetative growth, aerial hyphae production and ultimately sporulation in the lab. These can be easily distinguished from each other and as such are used as markers to connect the effect of a gene to a developmental stage, which many *Streptomyces* species could be used for. What has made *S. coelicolor* so widely used is its production of two main pigments: the blue actinorhodin (ACT) and red undecylprodigiosin (RED), which can be used as proxies for the effect a gene has on SMet.

2.1.1.1 Regulation of spore development and pigmentation by *whi* genes

The *whi* genes, named for the deficient mutant's white aerial hyphae colour, regulate both SMet and differentiation. For example, if *whiA* is knocked out, *S. coelicolor* spores germinate, go through vegetative growth and produce aerial hyphae, indicating that standard cellular processes and programmed cell death are not affected (Findlay *et al.*, 1999). However, after developing aerial hyphae, *whiA* mutants do not further differentiate into spores at the hyphae tips. Instead, the hyphae take a tightly coiled spring-like form lacking the characteristic grey polyketide associated with *S. coelicolor* mature spores. Wild-type spores are separated by septa, splitting the unigenomic spores into approximately 1 μm segments (Sigle *et al.*, 2016). The *whiA* mutants generated by Findlay *et al.* (1999) lack septa, forming hyphae over 100 μm long if uncoiled. Other *whi* mutants (*whiG, whiB, whiH)* produced similarly white unpigmented phenotypes, varying in hyphal coiling intensity and septa numbers.

The gene cluster directly responsible for "grey pigment" production in *S. coelicolor* is *whiE* (Kelemen *et al.*, 1998), which is conserved in other *Streptomyces* even if the associated colour can change: *S. halstedii* produces a green spore pigment. The genes are regulated by multiple other elements, with partial regulation by *sigF* and its sigma factor product σF (Kelemen *et al.*, 1998). Knocking out *sigF* does not affect each gene within *whiE*, since the activity of the *whiEP1* promoter is not altered, but *whiEP2* is neutered. This prevents expression of *whiE-ORFVII*, resulting in a brown spore pigment, indicating that σF regulates the pigment's chemical structure rather than its whole production. The final structure of the pigment is not known, but the intermediate Tw95a has been characterised (van Keulen & Dyson, 2014) following heterologous expression of *whiE* and *whiE-ORFIV* to produce a functional PKS. Its structure is shown in Figure 2.1.

Figure 2.1 Structure of the grey spore pigment intermediate Tw95a.



The full regulation of *whiE*, and the actual function of the pigment, is unknown (Salerno *et al.*, 2013). Chater *et al.* (2010) speculate that the pigment is covalently linked to the cell wall to provide some level of stress protection, whether from abiotic factors like heat or radiation, or against predatory organisms and their digestive enzymes. Similarly, decaying plant matter is a common food source for *Streptomyces*, but may end up consumed by animals like earthworms also eating the plants. If the pigment provides some protection against digestion by one of the most numerous soil animals its genes will be heavily selected for (Chater & Chandra, 2006).

2.1.1.2 Production of undecylprodigiosin is linked to increased cell death and sporulation in *S. coelicolor*

The *whi* example shows how specialised metabolism and differentiation, in this case sporulation, are tightly linked in *S. coelicolor*. Specialised metabolites and other stages of the lifecycle are similarly connected. One of the most well studied *S. coelicolor* compounds, RED, was found by Tenconi *et al.* (2018) to contribute to programmed cell death, a trait characteristic of complex multicellular species. Expression of *redD*, the corresponding BGC, was specifically found where cell lysis was occurring, potentially to prevent other microbes from taking advantage of free nutrients through cytotoxic activity of RED. Removing *redD* reduces cell

death, and so lessens available fuel for sporulation. Since not all *Streptomyces* contain *redD* it is unknown if there are other metabolites involved in cell death and if they are also intracellular instead of being secreted like many other antibiotics.

2.1.2 Metabolites produced by *S. coelicolor*

Prior to genome sequencing, *S. coelicolor* was known to produce five specialised metabolites with some antibiotic activity. Most of these were easily identified due to being pigments : blue ACT, brown methylenomycin, *whiE* grey spore pigment, RED, and the non-pigmented calcium dependent antibiotic (CDA) (van Keulen & Dyson, 2014). The impact of culture conditions on pigment production, primarily ACT and RED, had been observed since the isolation of *S. coelicolor* (Kutzner & Waksman, 1959) hence their use as general indicators of SMet. Since sequencing a further 16 biosynthetic gene clusters (BGCs) have been identified in *S. coelicolor* (Rutledge & Challis, 2015).

Despite being the model *Streptomyces*, there are still clusters remaining which do not have an associated product. A list of metabolites with known BGCs is given inTable 2.1, using van Keulen & Dyson (2014) and Challis (2014) as guides. van Keulen & Dyson (2014) note 9 additional BGCs with unknown products: a type I PKS, type II fatty acid, an NRPS, 4 ribosomally synthesised polypeptide BGCs, and 2 unknown siderophores.

The previously mentioned ACT and grey spore pigment are both polyketides, as is the yellow pigmented coelimycin P1 *(cpk* cluster*)*. Coelicmycin was originally discovered by removing the clusters for ACT, RED, and CDA, thereby increasing the metabolic flux to the otherwise silent *cpk* cluster (Gomez-Escribano *et al.*, 2012). Therefore, it is unlikely to be detected in any metabolomics experiments which does not use genetic engineering given the common expression of those 3 BGCs. Structures of characteristic *S. coelicolor* metabolites are shown in Figure 2.2.

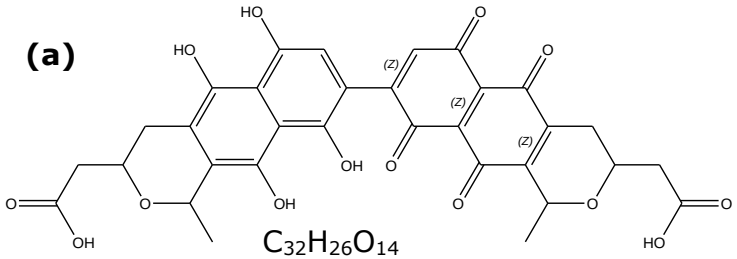
Table 2.1 *S. coelicolor* BGCs with known products. Some genes are labelled "SCO", indicating *S. coelicolor* nomenclature.

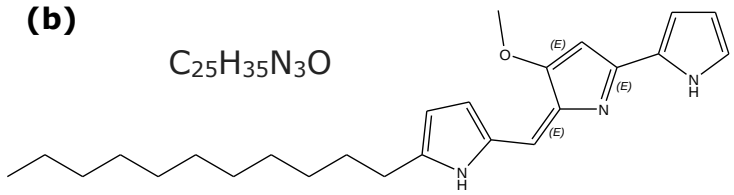
BGC class	Product	Gene or BGC	Bioactivity	Source
PKS				
TI PKS	Coelimycin	*cpk*	Antibiotic	Gomez-Escribano *et al.* 2012
TII PKS	Actinorhodin	*act*	Antibiotic	Mak & Nodwell (2017)
TII PKS	Tw95a	*whiE*	-	Lee *et al.*, (2012)
TIII PKS	Flaviolin	*thns*	-	Izumikawa *et al.*, (2003)
TIII PKS	Germicidin A	*gcs*	Anti-spore germination	Chemler *et al.*, (2012)
Peptide				
NRPS-lipopeptide	CDA	*cda*	Antibiotic	Wang *et al.*, (2013)
NRPS	Coelibactin	SCO *7676–7692*	Zincophore	
NRPS-hydroxamate	Coelichelin	*cch*	Siderophore	Lautru *et al.*, (2005)
RiPP	SapB	*ram*	Antibiotic	Kodani *et al.*, (2004)
Terpene				
Tetraterpene	Isorenieratene	*crtEIBV*	-	Takano *et al.*, (2005)
Triterpene	Hopene	SCO *6759–6771*	-	Seipke & Loria, (2009)

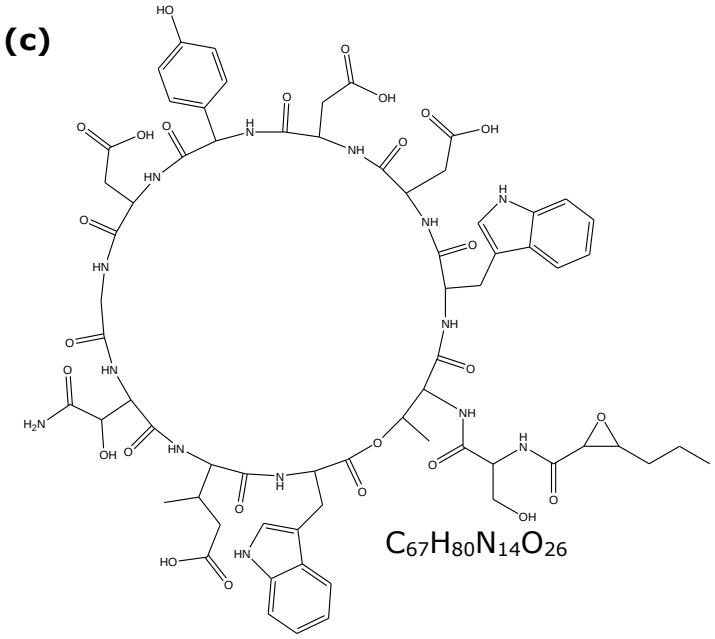
Sesquiterpene	Albaflavenone	*cyp*170A1	Antibiotic	Čihák *et al.*, (2017)
Sesquiterpene	Geosmin	*cyc2*	-	Yagüe *et al.* (2013)
Fatty acid				
PKS/TI fatty acid	Eicosapentaenoic acid	*pufA*	-	Shulse & Allen, (2011)
Miscellaneous				
Oligopyrrole	Undecylprodigiosin	*red*	Anticancer, antibiotic	Hu *et al.*, (2002)
Siderophore	Deferoxamine B	*des*	Siderophore	Tunca *et al.*, (2007)
γ-Butyrolactone	SCB1	*scbA*	Signalling	Hsiao *et al.*, (2007)
Cyclopentanone	Methylenomycin A	*mmy*	Antibiotic	Hobbs *et al.*, (1992)
Butenolide	Methylenomycin furans	*mmfRLHP*	Antibiotic	Corre *et al.*, (2010)

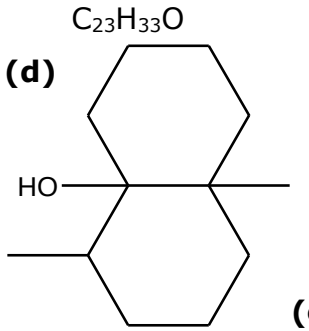
Terpenes are common *Streptomyces* BGCs, though are found throughout microbial and plant life (Marinelli, 2009). In *S. coelicolor* the most well-known is geosmin (Gust *et al.,* 2003), which is so far universal to *Streptomyces* (Kim *et al.*, 2015) and is responsible for the earthy smell of soil. While the smell makes it easy to superficially detect, geosmin's volatility makes it unsuitable for LC-MS analysis. Though it is a common metabolite, like many other compounds its function remains unknown.

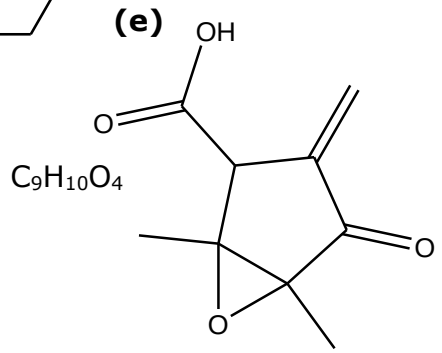
Figure 2.2 Structures and formulae of characteristic *S. coelicolor* metabolites **(a)** actinorhodin, **(b)** undecylprodigiosin, **(c)** calcium dependent antibiotic, **(d)** geosmin, **(e)** methylenomycin, and **(f)** albaflavenone.

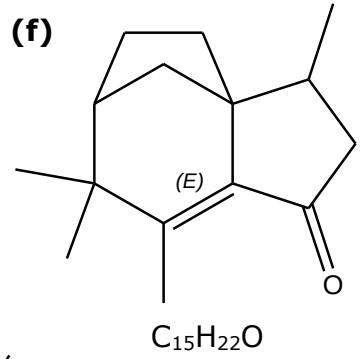












However, Gust *et al.* (2003) do note that mutants in *cyc2*, one of the genes responsible for geosmin biosynthesis, show "reduced" sporulation. As with RED and spore pigment, this again correlates a metabolite to a key part of the *Streptomyces* lifecycle. Geosmin is also produced by a range of other microbes within *Actinomycetales, Myxobacteria*, *Cyanobacteria*, and within fungi *Penicillium* spp. (Mattheis & Roberts, 1992; Giglio *et al.*, 2008). While not all of these are spore formers, indicating potential functions other than sporulation, all of these clades include prolific natural product producers.

Other metabolite classes include the oligopyrroles, like RED and streptorubin B, and the siderophore deferoxamines B, E, G, and D. Like geosmin, deferoxamines are highly conserved in *Streptomyces* (Kim *et al.*, 2015), but are not the only siderophores produced by *S. coelicolor*. The non-ribosomal polypeptide coelichelin is another siderophore, but it is not conserved to the same extent as deferoxamines. Following its identification in the partial genome structure of *S. coelicolor*, Lautru *et al.* (2005) detected coelichelin by comparing culture supernatants of coelichelin cluster (*ccH)* mutants with the wild type after growth in iron deficient media. Genome mining can lead to an overload of genomic information with little corresponding metabolomic data, so this is a good example of a targeted genomics/metabolomics strategy to isolate the product of a cluster of interest.

2.1.3 Chapter aim and objectives

The project requires a robust untargeted metabolomics system to reliably identify known metabolites over a variety of culture time points, highlight productive growth conditions of both known and unknown metabolites, all following removal of media associated compounds. *S. coelicolor* was the ideal choice to develop this method given the large amount of research on its metabolites, their regulation, and how this links to the phenotype. Therefore, the aim was to establish an untargeted metabolomics method which can assess the change in metabolites over different growth conditions. Within this, the objectives were to:

1) Culture *S. coelicolor* in a range of conditions to observe pigmentation and morphological variety.
2) Examine how sample time effects the chemotype within a growth condition.
3) Develop a combined microbiology and analytical chemistry method which can be applied to NCIMB *Streptomyces* and other microbes.

2.2 Materials and Methods

2.2.1 Chemicals

2.2.1.1 UPLC-MS solvents

Methanol, acetonitrile and formic acid (HPLC grade) were purchased from Fisher Scientific (Loughborough UK). Water was purified using a Purelab Flex system (18.2 MΩ, ELGA LabWater, High Wycombe UK).

Undecylprodigiosin (>93% purity) was purchased from Abcam (Cambridge UK).

2.2.1.2 Growth media

A full list of culture media and ingredients is given in Appendix table 1.

Media were selected based on their use in the International Streptomyces Project (Shirling & Gottlieb, 1966), general specialised metabolite production enhancement (Goodfellow & Fiedler, 2010) maintenance and sporulation effects (Shepherd *et al.*, 2010), minimal media (H. Zhu *et al.*, 2014), or made based on components in other media. The overall set included a mixture of rich and poor media (especially SEA and MMM) to compare metabolites upregulated by nutrient-deficient growth with ones potentially upregulated by large culture densities, and a mixture of carbon source alteration (maltose yeast-malt or MYM, proline yeast malt or ProYM) to potentially simplify the reason for any metabolomic or phenotypic differences.

2.2.2 *Streptomyces* selection from NCIMB collection

It was earlier stated that *S. coelicolor* A3(2) would always be referred to as *S. coelicolor.* An exception is made here as the culture collection uses the correct taxonomic name. *S. violaceoruber* (Waksman and Curtis) BAA-471 was purchased from the American Type Culture Collection (Manassas, US).

The NCIMB collection was searched for *Streptomyces* strains with minimal recent specialised metabolite research, and ideally those with minimal bioactivity research overall. More recent accessions with potentially ongoing research were also included based on the novelty of their native environment. *S. coelicolor*

Müller was also included to demonstrate its difference to *S. coelicolor*. The selected NCIMB strains are listed in Table 2.2.

Table 2.2 *Streptomyces* species selected from the NCIMB collection

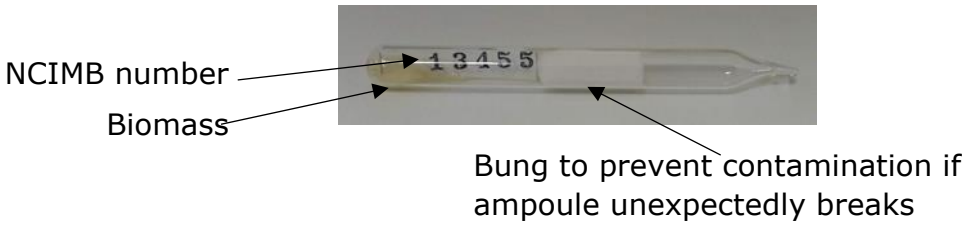
Strain	NCIMB Number	Isolated from	Year of accession
S. coelicolor Müller	9798	-	1965
S. halstedii	9839	"Deeper soil layers"	1966
S. pseudoechinosporeus	9918	Desert sand, Kazakhstan	1966
S. avellaneus	11000	Cave soil, Italy	1973
S. flaviscleroticus	11008	Soil, India	1973
S. purpureus	11313	Soil, Hawaii	1977
S. sp	12947	Lavender field soil, Hungary	1989
S. costaricanus	13455	Soil, Costa Rica	1997
S. rhizosphaericus	13674	Ectorhizosphere, Indonesia	1997
S. sanyensis	14820	Mangrove sediment, China	2012
S. aridus	14965	Atacama Desert, Chile	2014

2.2.3 Lyophilised strain revival and maintenance

2.2.3.1 Lyophilised strain resuscitation

NCIMB strains are stored in glass ampoules (Figure 2.3) which require snapping to access the freeze-dried culture for resuscitation. The ampoule was scored with a diamond tipped cutter and snapped inside a laminar flow hood to prevent contamination, and a sterile Pasteur pipette used to suspend the culture in approximately 1 ml glucose yeast-malt (GYM) broth. The suspended culture was spotted onto a GYM plate and spread to isolate single colonies, and into a 20 ml GYM vial, both of which were incubated at 30°C. This was repeated for all Strains in Table 2.2, and the American Type Culture Collection *S. coelicolor* ampoule.

Figure 2.3 NCIMB lyophilised culture ampoule.



2.2.3.2 *Streptomyces* strain maintenance

2.2.3.3 Spore plate culture

The GYM broth *Streptomyces* cultures (100 µl) were used to inoculate GYM and Soya Flour Mannitol plates, then spread for confluent growth. The plates were incubated at 30°C until the culture had sporulated.

2.2.3.4 Spore suspensions

Spores were harvested to create spore suspensions for later inoculations, modifying the method described in Shepherd *et al.* (2010).

All non-biological materials were sterilised either by autoclaving or by the manufacturer prior to use. A spore filter was constructed by placing a 20 ml syringe barrel into a 50 ml centrifuge tube (Fisher Scientific) and filling the bottom of the barrel with sterile cotton wool. The spores were harvested by pipetting 5 ml sterile reverse osmosis (RO) water onto the plate and gently rubbing the surface with a cotton bud to dislodge and suspend spores. The suspension was then pipetted into the cotton wool and filtered by replacing the syringe plunger. The syringe was discarded.

The tube was then centrifuged at 5,000 x g for 5 min, after which the supernatant was removed with a pipette and discarded. The spore pellet was resuspended in 1 ml 25% sterile aqueous glycerol, vortexed to mix, and stored at -80°C.

A 10 µl aliquot was incubated on a fresh Soya Flour Mannitol plate to check for contamination; 50 µl was used for further extract cultures on 9 cm agar plates, and 20 µl into 24 and 25-well plates.

2.2.4 *S. coelicolor* co-culture with NCIMB strains in 25-well plates

S. coelicolor was cultured in Sterilin square 25-well plates (Fisher Scientific) with each well containing 2 ml M19 media. M19 was selected as it induced relatively early and high production levels of the blue pigment ACT from *S. coelicolor* compared to the other media, which acted as a crude proxy for specialised metabolite production (Kim *et al.*, 2007; Wang *et al.*, 2008)

Spore suspension (20 µl) of *S. coelicolor* were spotted into the wells along with 20 µl suspension of each of the 11 soil NCIMB *Streptomyces* (Table 2.2). Aside from *S. coelicolor*, each of the 11 strains was inoculated into 2 wells: one for axenic growth, and another for co-culture with *S. coelicolor*. *S. coelicolor* had its own axenic well as a control.

2.2.5 Evaluation of *S. coelicolor* lifecycle and metabolites

2.2.5.1 Culture of *S. coelicolor* on agar for metabolite extraction

S. coelicolor spore suspension (50 µl) was spread to give confluent growth across 63 agar plates, which were incubated at 30°C. At each sample point, triplicate plates were sampled by cutting each into approximately 2 cm^3 pieces and crushing with a glass homogeniser (Fisher Scientific). Samples were taken at 3, 6, 9, 12, 24, 27, 30, 48, 54, 72, 96, 120, and 168 h then every following 24 h up to 15 d, requiring preparation of multiple spore suspensions.

2.2.5.2 Extraction of *S. coelicolor* metabolites

Samples were processed by transferring 3 x 1 g homogenised agar culture to 3 separate 15 ml centrifuge tubes per plate, which were then stored at -20°C until all samples had been collected. After defrosting, each tube was used for extraction with a different solvent: water, 50% aqueous methanol, and methanol, to cover a medium-high to high extract polarity range. Common less

polar solvents such as ethyl acetate or butanol were avoided to prevent extraction and rediscovery of known metabolites.

Extraction solvent (4 ml) was added to the respective tubes and shaken at room temperature in a rotary incubator (220 rpm) for 1 h with vortexting every 15 min. An aliquot (800 µl) of crude extract was transferred to a 1.5 ml microcentrifuge tube and centrifuged at 5000 x g for 5 min, and the supernatant filtered through a Whatman 0.5 µm PTFE filter into LC-MS autosampler snap top vials (ThermoFisher Scientific).

2.2.5.3 Metabolite analysis

Two sets of LC-MS conditions were used for *S. coelicolor* metabolite analysis. The first, Method A, used lab standard microcystic-optimised settings, and were used in the Lifecycle extraction trial.

Samples were analysed using a Waters Acquity UPLC coupled to a Waters Acquity PDA and Xevo G2 QToF (Manchester, UK) controlled by Masslynx v4.1. Separation was achieved using a BEH C18 column (2.1 x 50 mm, 1.7 µm particle size) with a target column temperature of 40°C. The mobile phase was ultrapure 18.2 MΩ water with 0.1% formic acid (solvent A) and acetonitrile with 0.1% formic acid (Solvent B) run on the gradient shown in Table 2.3. The auto sampler was maintained at 5°C and took 5 µL sample for injection into the LC-MS.

Table 2.3 LC settings for Method A.

Time/min	% A	% B
0	95	5
10	30	70
11	0	100
12	0	100
12.5	95	5
15	95	5

Centroid data was acquired using electrospray in positive ionisation mode with a mass scanning range of *m/z* 50 to 2000. Scan duration was 0.25 s and inter-scan delay 0.025 s. The capillary, sampling and extraction cones were set to 3.3

kV, 25 V and 3 V respectively, with source and desolvation temperatures at 80°C and 350°C respectively. Desolvation and cone gas flows were 500 l/h and 50 l/h. with cone gas at 50 l/h. Collision energy was 15 V. The reference calibration compound was Leu-enkephalin (Sztáray *et al.*, 2011) with a reference scan frequency of 10 s and reference cone voltage of 25 V.

Data was collected using MS^E. After an ion has been scanned it is subjected to a collision energy ramp to generate a second data set which approximates MS/MS fragmentation. The ramp settings were 20 eV start to 55 eV end.

2.2.5.4 Culture and staining of *S. coelicolor*

GYM plates were overlaid with sterile porous cellophane (Sigma-Aldrich) and inoculated to form a confluent lawn with 50 µl spore suspension. GYM was chosen because of its general use from being the 2nd International Streptomyces Project medium (Shirling & Gottlieb, 1966).

S. coelicolor cell time dependent morphology and viability was assessed with a Molecular Probes L-13152 Live/Dead BacLight Bacterial Viability Kit. At each sample time point cellophane was removed and a section cut out to place on a slide. This was stained following the method in Manteca *et al.* (2005). Briefly, the SYTO 9 and Propidium Iodide reagents were separately dissolved in 2.5 ml filter sterilised 18.2 MΩ water, then combined in a 1:1 ratio to create the dye master mix.

The mix was dropped over the sample to submerge it, a sterile cover slip placed over it, and left in the dark to stain for 10 min. Samples were observed using a Leica DMRB fluorescence microscope with 540 – 650 nm and 355 – 425 nm excitation to 645 – 675 nm (red) and 470 nm (green) emission filters respectively.

2.2.6 Evaluation of impact of culture media on *S. coelicolor* metabolism

The 8 media found in Appendix table 1 were used to fill 3 24-well plates, which were inoculated with 20 µl *S. coelicolor* spores/well. Plates were incubated at

30°C for 5, 10, and 15 d with one plate removed for processing at each time point.

A 1 cm diameter corer was used to remove a plug from the centre of each well, which was then placed into a 1.5 mm microcentrifuge tube. The corer was cleaned in fresh RO water between each well. Samples were stored at -80°C prior to extraction. After defrosting, the cores were crushed with a glass rod and metabolites extracted by adding 1 ml 50% aqueous methanol. The crude extract tubes were vortexed every 15 min for 1 h, after which a 500 µl sample from each tube was syringe filtered through a Millipore 0.2 µM cellulose filter (Merck, Southampton, UK) into a corresponding UPLC snap top vial.

A sterile media 24-well plate was processed in the same way to analyse blank media, ensuring data analysis would not focus on media compounds.

2.2.6.1 Metabolite analysis

The second set of UPLC-MS conditions, Method B, was based on literature examples of *Streptomyces* natural product metabolomics using the same UPLC-MS equipment (Rao *et al.*, 2011; Rao *et al.*, 2013)

Gradient settings changed from section 2.2.5.3 to the new method are listed in Table 2.4. The mobile phase solvents A and B are unchanged. Changes to mass spectra acquisition are shown in

Table 2.5. All other conditions remain the same as section 2.2.5.3.

Table 2.4 LC settings for Method B.

Time/min	% A	% B
Initial	95	5
0.2	95	5
1.1	75	25
2.25	70	30
3	60	40
6	5	95
7.5	5	95
7.75	95	5
10	95	5

Table 2.5 Method A and B MS settings.

	Method A	**Method B**
Sampling cone/V	25	35
Extraction cone/V	3	4
Source temperature/°C	80	120
Desolvation temperature/°C	350	500
Desolvation gas flow/l/h	500	1000
Collision energy/V	15	60

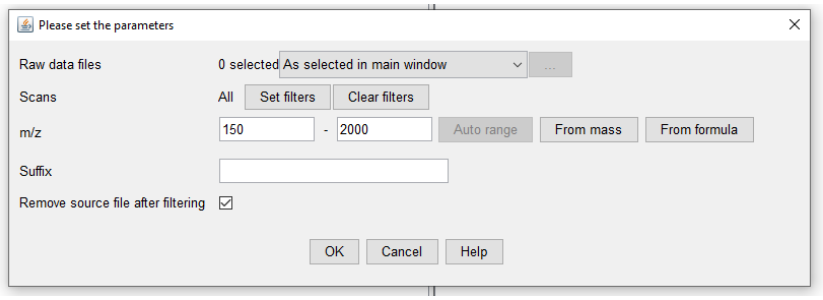
2.2.6.2 Data mining with MZmine

Waters RAW data files were converted to 32-bit .mzML files using MSConvert (Chambers *et al.*, 2012). The MS Level filter ScanEvent was set at 0 – 1 to remove UV data. The processed files were imported into MZmine v2.37 (Pluskal *et al.*, 2010; Myers *et al.*, 2017). The first stage, raw data methods, is displayed in shown in Figure 2.4.

2.2.6.3 Data visualisation

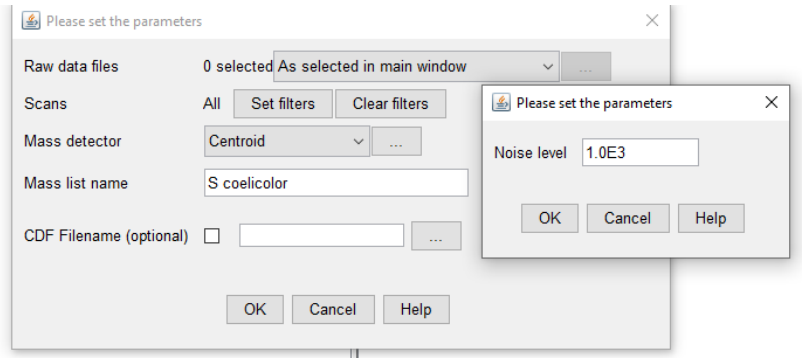
Structures were visualised with ChemDraw Ultra 7. Data was processed and graphs produced with GraphPad Prism 8.2.1.

Figure 2.4 Raw data methods used in MZmine after data import.



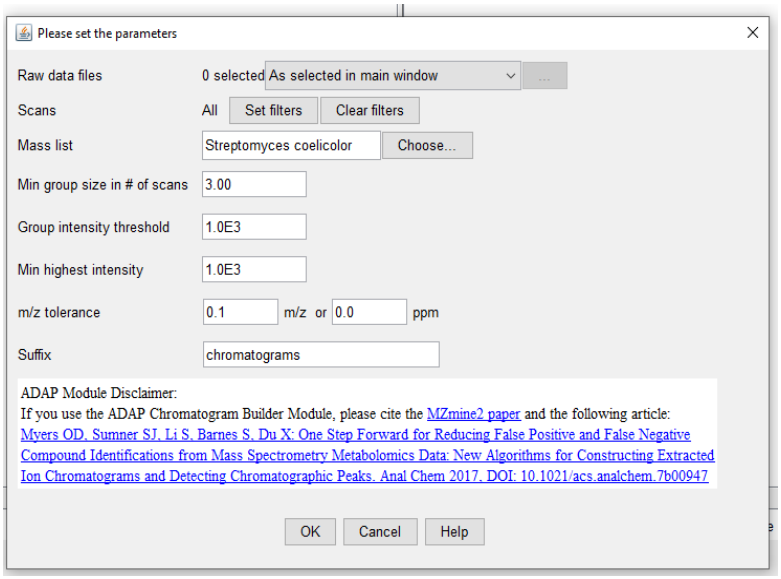
Crop filter:

Filter to only contain 150 – 2000 *m/z*, removing some common solvent peaks



Mass detection:

Set centroid mass detection lower limit to 1E3



ADAP chromatogram builder (Myers *et al.*, 2017):

Min. group size = 3

Intensity threshold = 1E3

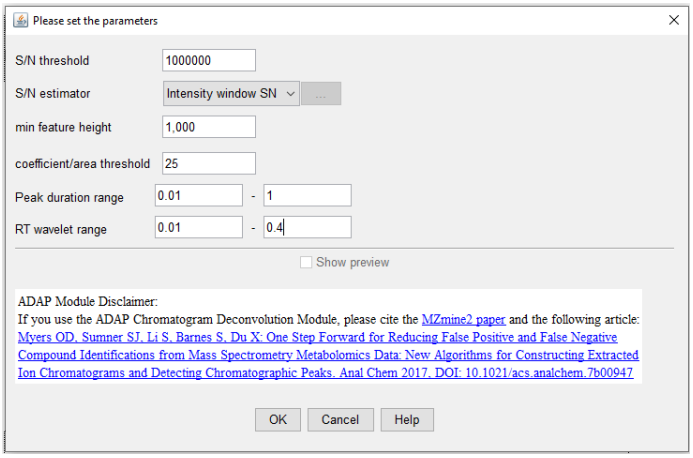
Min highest intensity: 1E3

m/z tolerance = 0.1 *m/z*

Prevents false positive chromatograms.

Having generated a chromatogram with the ADAP builder, these are converted to peak lists which are aligned into a single list, for example aligning triplicates (Figure 2.5).

Figure 2.5 Use of peak list methods to create peak lists from each file and align them into a master list.



Deconvolution (ADAP):

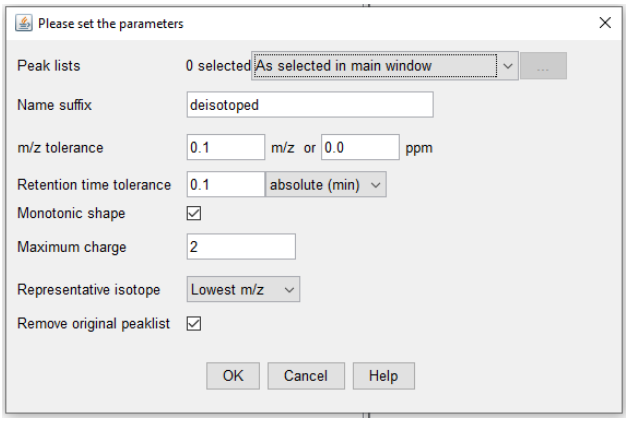
S/N = 1E6

Min. feature height = 1E3

Coefficient/area threshold = 25

Peak duration range = 0.01 – 1 min

RT wavelet range = 0.01 – 0.4 min



Isotopic peak grouper:

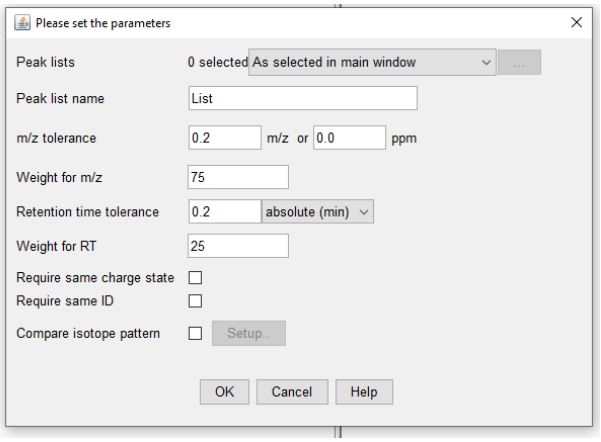
m/z tolerance = 0.1

RT tolerance = 0.1 min

Maximum charge = 2

Representative isotope = lowest *m/z*

Monotonic shape = yes



Alignment:

m/z tolerance = 0.1

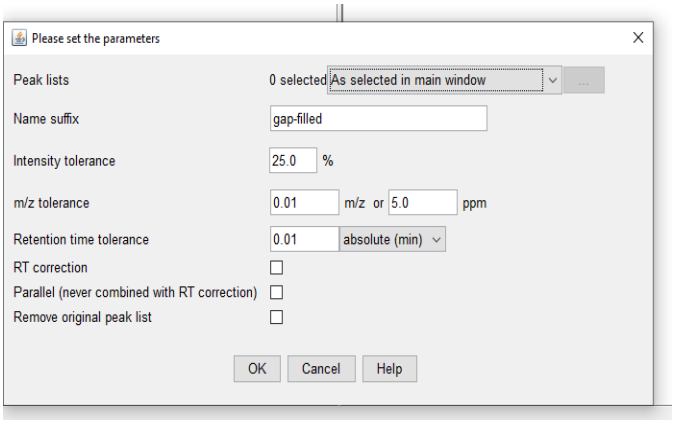
RT tolerance = 0.2 min

Weight for *m/z* = 75

Weight for RT = 25

Aligned peak lists need further quality control before being used for statistical analysis. Steps are shown in Figure 2.6. After processing, peak lists can be using for statistical analysis (principal component analysis, scatter plots) and exported to other metabolomics-related tools such as GNPS or MetaboAnalyst.

Figure 2.6 Processing of aligned peak lists.



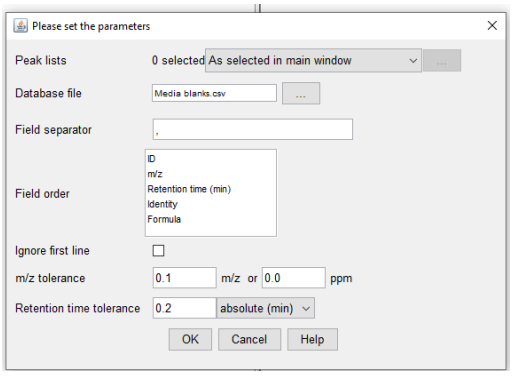
Gap filling:

Intensity tolerance = 25%

m/z tolerance = 0.01 *m/z*

RT tolerance = 0.01 min

Fills in any missing features in an aligned peak list using the constituent lists



Custom database identification:

m/z tolerance = 0.1 *m/z*

RT tolerance = 0.2 min

Allows for identification and removal of media blanks or other specific features.

2.3 Results and Discussion

2.3.1 *S. coelicolor* ACT production and co-culture

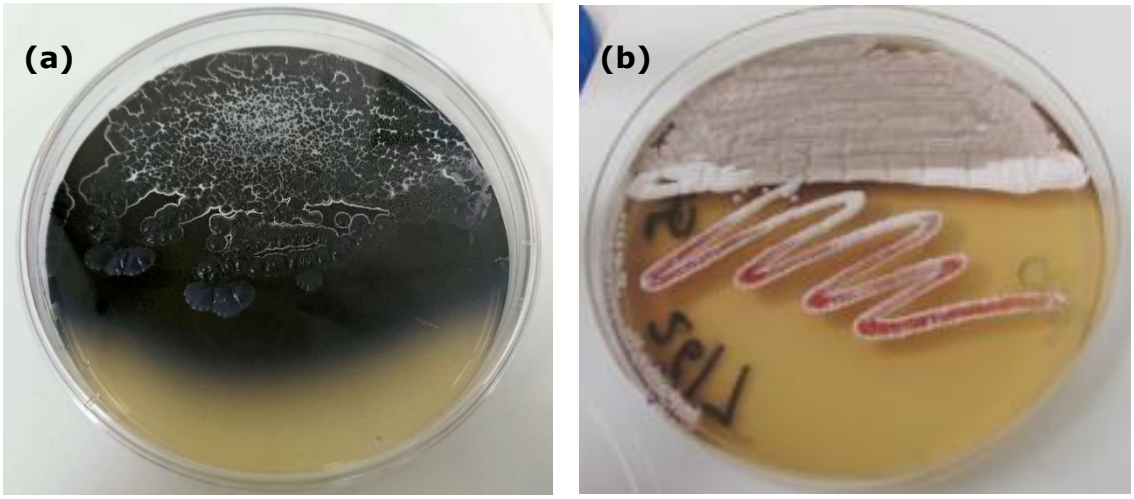
2.3.1.1 Effect of media on ACT production and differentiation

Preliminary culture of *S. coelicolor* on different media showed the mannitol-peptone medium M19 enabled relatively high and early production of the blue pigmented antibiotic ACT. Growth of *S. coelicolor* on other media is shown in Figure 2.14. However, differentiation was also delayed, with aerial hyphae only appearing in the denser regions of the colony.

Figure 2.7 emphasises the difference culture on M19 agar media makes with a comparison to GYM at 3 d, which is already developing mature spores and no ACT production, only RED. The presence of aerial hyphae on M19 culture is presumably a response to the lack of nutrients caused by high cell density in comparison to the outer less densely packed biomass. The identity of the pigment was confirmed by adjusting a culture plate extract's pH from neutral to alkaline and back, causing the characteristic blue to red (alkaline) then back to blue colour change (Coisne *et al.*, 1999).

Figure 2.7 **(a)** *S. coelicolor* A3(2) on M19 after 10 d showing actinorhodin production and sparse aerial hyphae, and **(b)** *S. coelicolor* on GYM after 3 d with grey mature spore pigment and prodigiosin.

.



On M19 the colonies lacked aerial hyphae for up to 5 d, the only culture condition to not produce any (Figure 2.14), while still activating specialised metabolism through ACT production. Growth on minimal media with mannitol as the carbon source has been reported (Rigali *et al.*, 2006) to restore aerial hyphae to *bld* strains. *bld*, or "bald", refers to an aerial hyphae-less phenotype caused by a variety of mutants; in this case, the mutation is in *dasR*, a master regulator controlling differentiation and specialised metabolism.

Rigali *et al.* (2006) also note that Soya Flour Mannitol medium, which is commonly used to induce sporulation in *Streptomyces* (Shepherd *et al.*, 2010), led to low sporulation and overproduction of ACT in their *S. coelicolor dasR* mutant. Strangely the opposite situation, upregulation of *dasR*, also produces a *bld* phenotype with high ACT levels but on multiple media. Both these descriptions match growth on M19, even though the strain used here is the wild type.

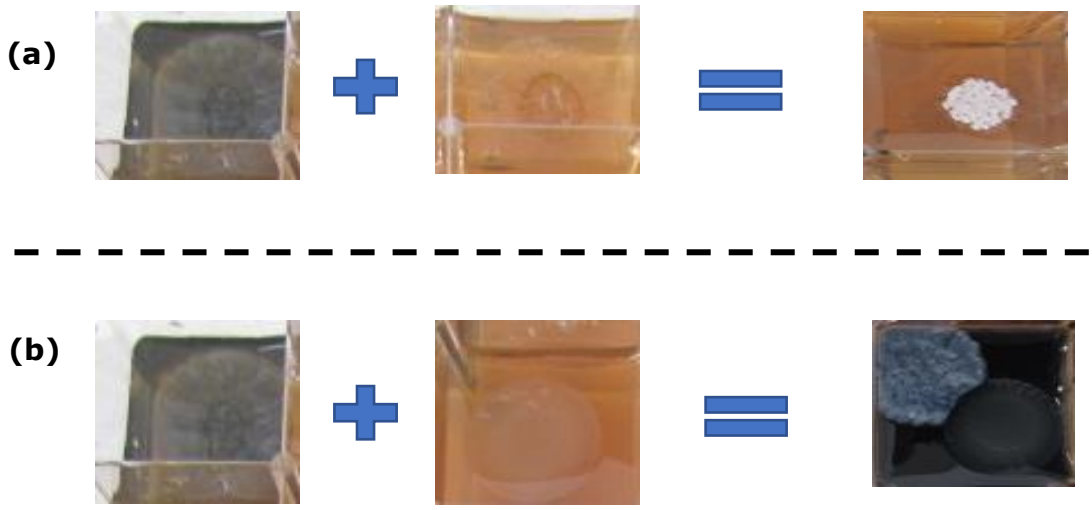
While there are many other genes that can affect production of aerial hyphae, such as *bldA* (Hackl & Bechthold, 2015) or *prsH* (Viollier *et al.*, 2003), these have different negative effects on *Streptomyces* pigment production and so specialised metabolism. Given the similar phenotypes of growth on M19 to that of *dasR* disruption on Soya Flour Mannitol medium or upregulation, it is possible that either of the components of M19 (mannitol and peptone) are affecting levels of DasR.

ACT production is commonly used as a proxy for specialised metabolism levels in *Streptomyces* since its secreted form – the lactone derivative γ-ACT (Coisne *et al.*, 1999) – can easily be visually approximated. The intracellular red pigment RED has also been used as an indicator of specialised metabolism (Sevcikova & Kormanec, 2004), but this was not detected. Like other metabolites, compound levels are affected by a mix of abiotic factors, for example salt or rare earth metals (Sevcikova & Kormanec, 2004; Tanaka *et al.*, 2010), and biotic factors like co-culture (Onaka *et al.*, 2011).

2.3.1.2 Co-culture both increases and decreases *S. coelicolor* ACT production

Selected NCIMB *Streptomyces* (Table 2.2) were co-cultured with *S. coelicolor* to screen for upregulation of specialised metabolism as indicated by pigment levels. The majority of strains either had no increase in ACT or lowered it. Figure 2.8 shows 2 co-cultures which either had morphological or upregulatory effects on ACT. Full co-culture is displayed in Appendix figure 1.

Figure 2.8 Co-culture of **(a)** *S. coelicolor* with *S. halstedii*, and **(b)** *S. coelicolor* with *S. purpureus.*



The majority of strains resulted in reduced pigment production. The extent of this varies, with some (*S. coelicolor*/*S. halstedii*) totally abolishing production and other *S. coelicolor* retaining pigment at the colony edge furthest from its competitor (*S. coelicolor*/*S. aridus*). Only one co-culture, *S. coelicolor*/*S. purpureus* positively affected ACT production, with both the intracellular and secreted pigment levels increasing. Presumably this is due to upregulating *S. coelicolor* ACT levels, as *S. purpureus* is not a known producer of any blue pigments. The *S. coelicolor*/*S. halstedii* co-culture induced aerial hyphae and possible sporulation in *S. halstedii*, not otherwise seen on this medium.

Why some strains decreased production is unknown. ACT requires a neutral to alkaline pH to be secreted as γ-ACT, which *Pseudomonas* species can prevent

through media acidification, but this does not stop intracellular pigmentation (Galet *et al.*, 2014). Traxler *et al.* (2013) grew *S. coelicolor* with multiple actinomycetes, leading to altered pigmentation and other metabolites, and determined that siderophore production was induced through iron scavenging competition.

Aside from ACT upregulation, this experiment did show the advantages of a multi-well screen compared to standard petri dish culture: more efficient in use of space, chemical and biological resources, and easier comparisons for strain under varying conditions.

2.3.2 Evaluation of *S. coelicolor* metabolites over time

As the model *Streptomyces*, *S. coelicolor* has the most characterised metabolites of any species in the genus. It was therefore chosen for a trial LC-MS metabolomics experiment to test one OSMAC parameter – sample time – along with the difference in metabolites extracted by 3 different high polarity solvents, and whether the LC-MS conditions used were suitable.

2.3.2.1 Fluorescence microscopy of *S. coelicolor* growth phases

Like other *Streptomyces*, *S. coelicolor* undergoes a process of differentiation over time which is linked to specialised metabolism. It was therefore hypothesised that specific differentiation stages may be correlated with the production of metabolites. Fluorescence microscopy showed clear development of individual spores, germination, and rapid growth of the MI hyphae in Figure 2.9. The first round of programmed cell death (PCD) is possibly also observed at 12 h, with some segments of hyphae stained red indicating damaged or dead cells. PCD fuels further growth to produce uneven concentrated islands of biomass across the plate (Manteca *et al.*, 2007), which is shown in Figure 2.10, until eventually aerial hyphae and spores were produced by 72 h (Figure 2.11).

The morphology during lifecycle was compared to previous fluorescence microscopy of *S. coelicolor*, *S. antibioticus*, and *S. lividans* (Manteca *et al.*, 2005; Manteca *et al.*, 2007) where spores were seen to follow the same differentiation pattern and partially similar approximate timings. These timings are compared with *S. antibioticus* and *S. lividans* (Table 2.6), as Manteca *et al.*, (2005) found

S. coelicolor and *S. lividans* timings to be identical. *Streptomyces* development is asynchronous so even within a strain these will vary (De Jong *et al.*, 2009).

Table 2.6 Fluorescent microscopy-based timings of *Streptomyces* growth phases

Species	Germination	MI	MII	Spores	Source
S. antibioticus	4 h	6 h	14 h	45 h	Manteca, *et al.*, (2005)
S. lividans	4 h	6 h	17 h	50 h	Manteca *et al.*, (2007)
S. coelicolor	**6 - 9 h**	**6 - 12 h**	18 h	60 h	This work

More sample points were used in Manteca *et al.,* (2005; 2007) which accounts for the slight gap between MII and spore formation. However, germination and appearance of MI did take over double the previously recorded time. Images at 6 h did show spores, and while there were potential stubs on some which could have been early germ tubes, it was not confirmed. Therefore, germination occurred between 6 and 9 h. This extends to the MI hyphae, as a germ tube can be considered MI hyphae so it may be most accurate to say MI appear around 9 h, but a typical mat of hyphae was only seen from 12 h onwards.

Figure 2.9 Live (green) dead (red) staining fluorescence microscopy of *S. coelicolor*. **(a)** Spores at 6 h, 200 x; **(b)** Germination tubes marked at 9 h, 400 x; **(c)** Early MI growth, 12 h 200 x; **(d)** Also 12 h, less crowded segment showing early clusters of growth and dead segments, marked; **(e)** Rapid growth around core clusters at 15 h, 100 x.

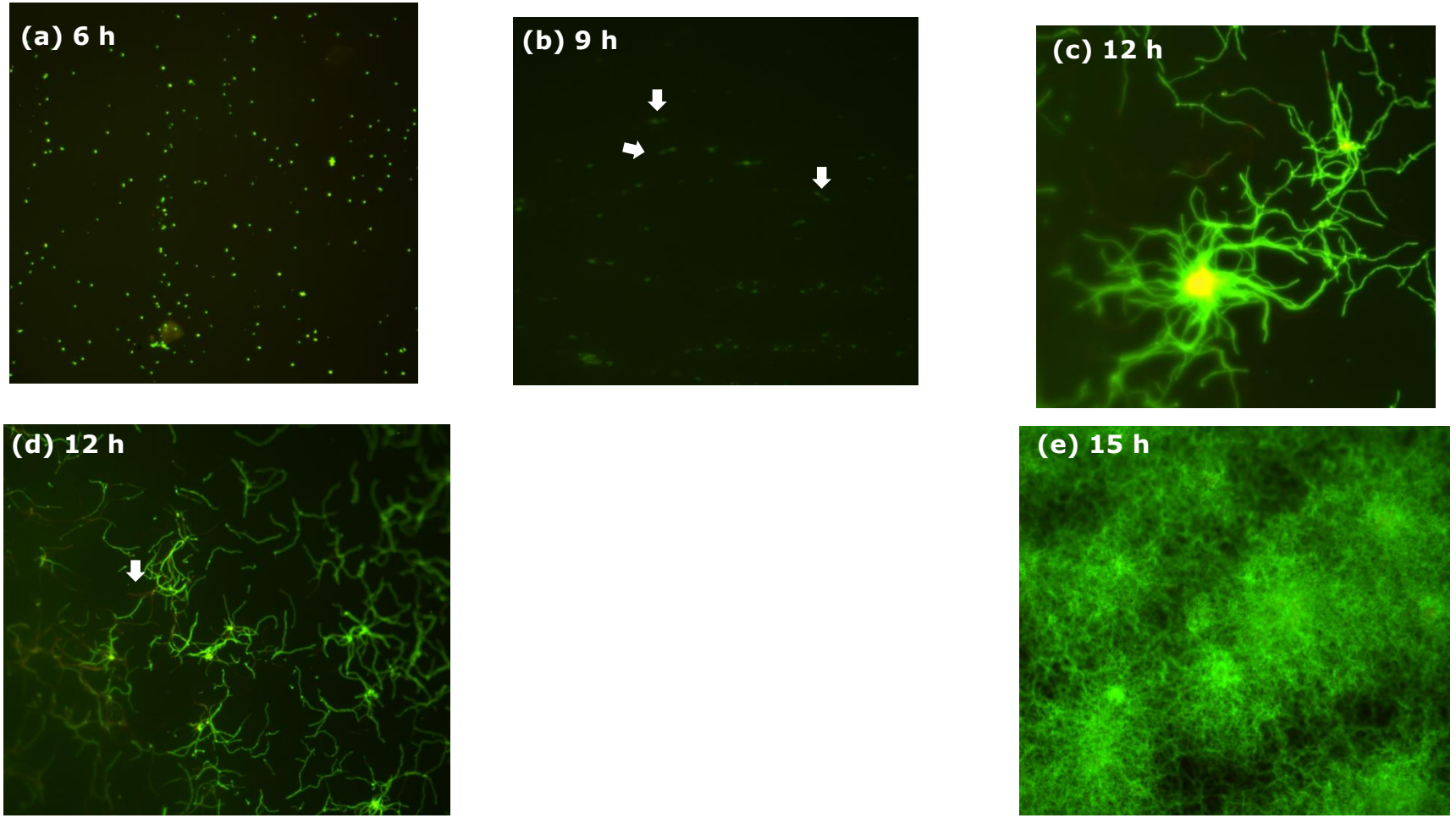


Figure 2.10 *S. coelicolor* growth island at 18 h with fluorescence microscopy showing **(a)** live and dead cells, **(b)** dead cells, and **(c)** light microscopy of the same island. All 200 x.

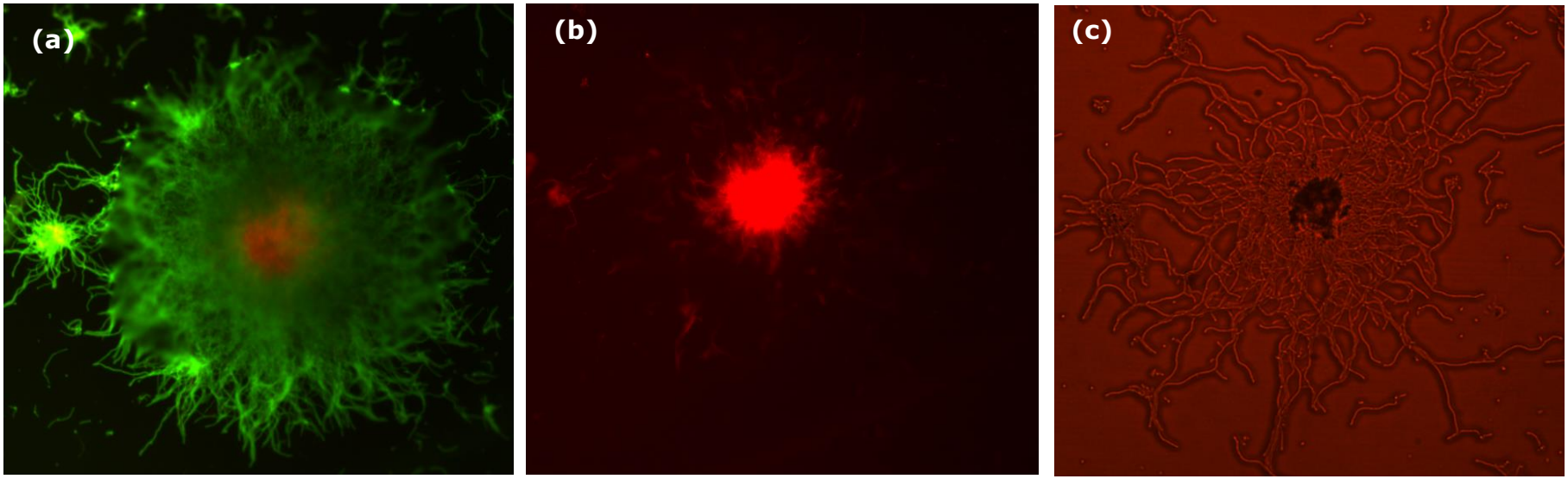
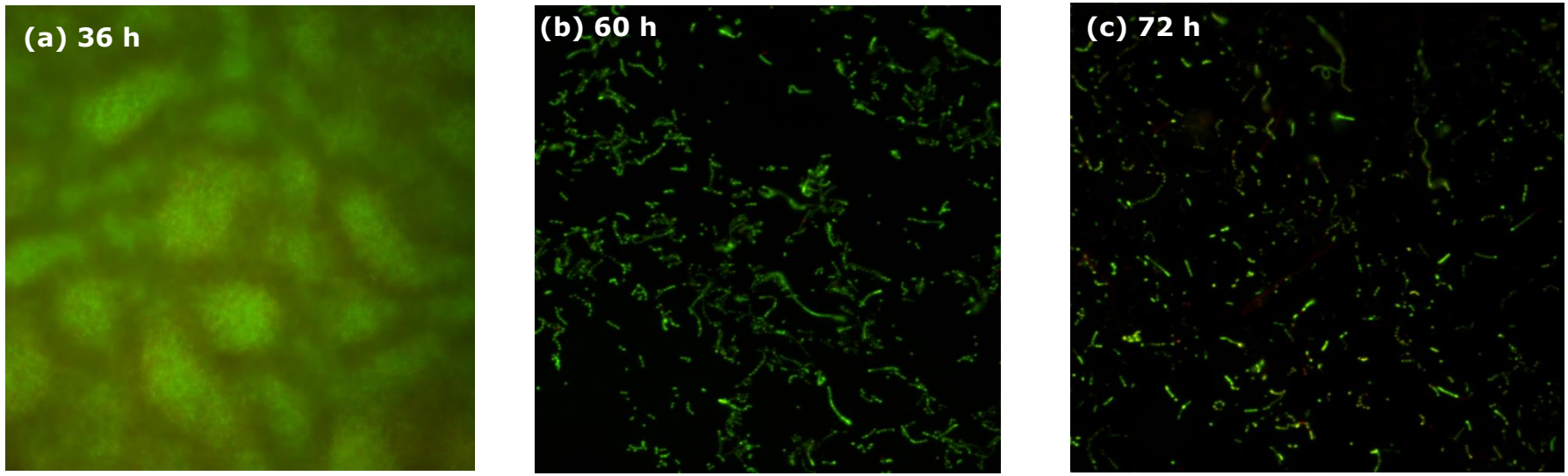


Figure 2.11 Late stages of *S. coelicolor* lifecycle. **(a)** At 36 h growth islands cover the surface with relatively empty channels between them, 100 x. **(b)** Cover slip print and stain of aerial hyphae at 60 h, 200 x, and **(c)** Cover slip and print stain of aerial hyphae and spores, 72 h, 200 x. Example spore chain marked with separated oval spores visible.



2.3.2.2 Alternation of metabolite analysis parameters

Unfortunately, there was no metabolomic data to correlate to the differentiation cycle. Though growth, differentiation, and pigment production were seen, LC-MS analysis showed no metabolite differences between the time points or the agar blank (Figure 2.12). The reason for the lack of detected metabolites was likely a combination of the inefficient extraction method and unoptimised LC-MS parameters. The nutrients and temperature (GYM agar, 30°C) were standard *Streptomyces* culture parameters (Shepherd *et al.*, 2010).

Therefore, the hypothesis of this section – that time and solvents will affect available metabolites – had not been properly tested. The experiment was redesigned to use the multi-well plate format of Section 2.3.1 which more effectively, if in a preliminary fashion, studied metabolite differences caused by altering the environment of *S. coelicolor*. Reuse of the multi-well plate format also required a new extraction method.

Three agar culture sample processing methods were compared: 1) cutting into pieces then leaving in solvent; 2) crushing then solvent; and 3) cutting, freeze/defrosting, then crushing and solvent. Each of the three did give detectable metabolites with 50% aqueous methanol extraction solvent unlike the previous experiment, specifically the siderophores deferoxamine B and coelichelin (Figure 2.13). Cutting up gave the fewest chromatogram metabolite peaks, so freeze/thawing then crushing was selected as the most suitable method for extraction. The LC-MS parameters were then changed to more *Streptomyces* optimised settings, specifically those in Rao *et al.* (2013) as the same model of UPLC-MS qToF was used.

Solvent selection can have a major impact on the extracted metabolites, as different solvent polarities will lead to a different set of metabolites being extracted. As part of a diverse OSMAC project using *Streptomyces* spp. and *Salinispora* spp., a predominantly marine genus within *Actinomycetales*, Crüsemann *et al.* (2017) sequentially extracted culture samples with the common solvents ethyl acetate, butanol, and methanol. After grouping the metabolomic features into related clusters, approximately 6000 unique clusters were found, with 3000 coming from ethyl acetate extracts, 1750 in butanol extracts, and 1250 from methanol extracts. When comparing overall molecular

families rather than feature clusters, methanol extracted the largest unique amount at 25 of 57 total. This translated into some known metabolites only being extracted by specific solvents, such as antimycin A1 by methanol, salinamide E by butanol, and arenimycin by ethyl acetate.

As these are commonly used solvents, it could be hypothesised that the use of uncommonly used solvents would also expand the chemical space. Barkal *et al.* (2016) extracted *Aspergillus nidulans* cultures with chloroform, and the high boiling point solvents 1-pentanol (138°C), and γ-caprolactone (219°C). Some known metabolites were preferentially extracted (for example, asperfuranone by γ-caprolactone), but there was no examination of whether novel metabolites were extracted.

Since the overall project was focused on multiple *Streptomcyes*, not just *S. coelicolor*, the multi-well plate screen was remade around growth conditions more likely to change global specialised metabolism regulation irrespective of the strain. Therefore the screen was designed around different nutritional conditions, since different minerals, metals, carbon and nitrogen sources are well documented as triggering specialised metabolite diversity (Bode *et al.*, 2002; Rateb *et al.*, 2011; Traxler *et al.*, 2013). The other difference from the first attempted screen was in sample points. Early time points are unlikely to give high enough levels of specialised metabolites, since early growth is more focused on biomass production. Therefore, sample points were changed to 5, 10, and 15 days to cover stationary and nutrient depletion stages.

Figure 2.12 Chromatograms of *S. coelicolor* culture on GYM plates extracted with water, 50% methanol, and 100% methanol after 15 d.

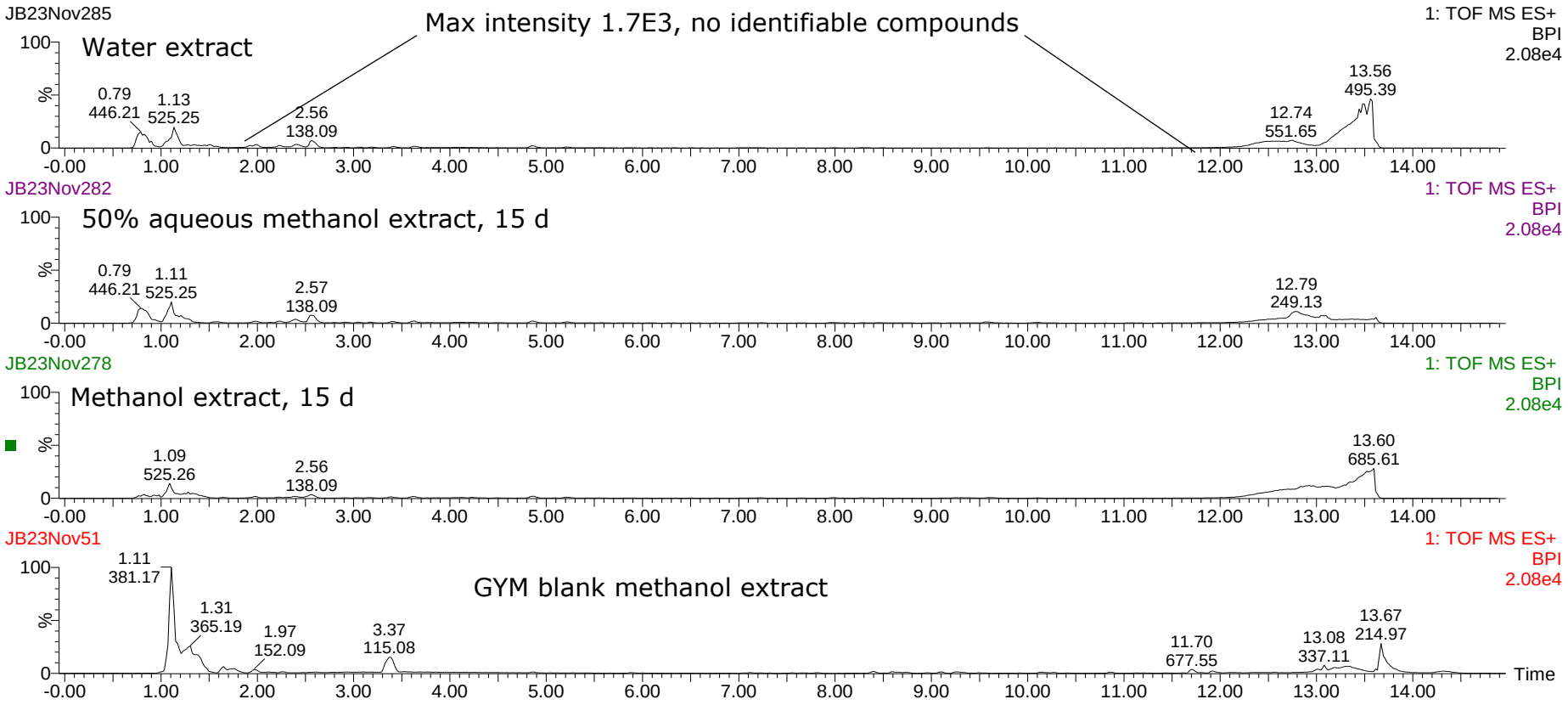
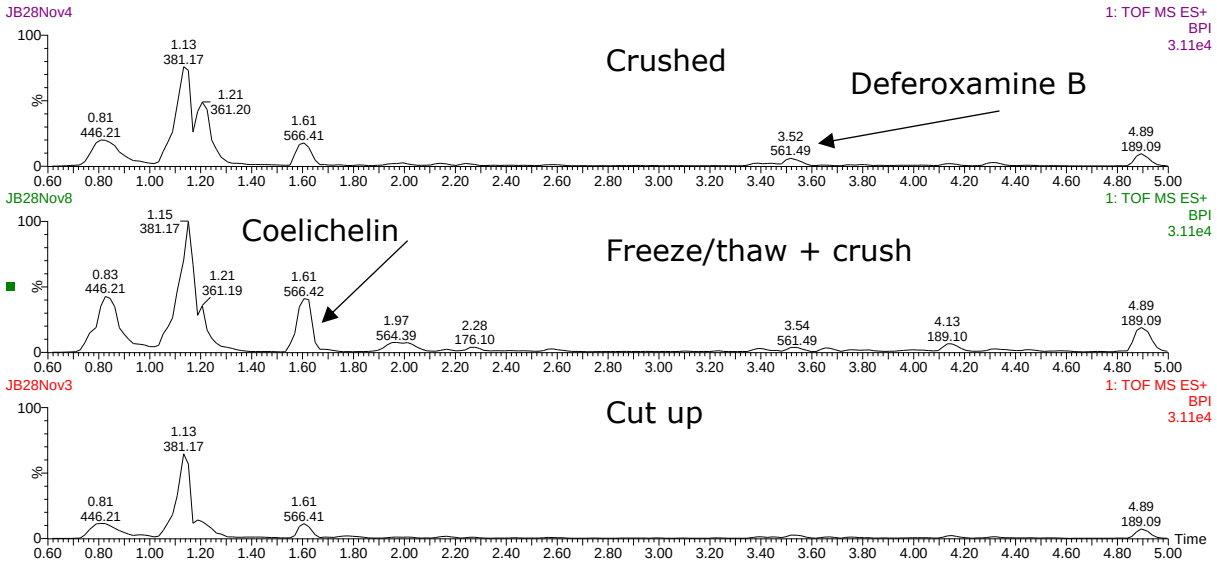


Figure 2.13 Comparison of 3 sample processing methods for extraction in 50% methanol. Chromatograms are representative of triplicate replicas.



2.3.3 OSMAC 24-well Plate Nutrition Screen

2.3.3.1 Morphological development of *S. coelicolor* in 24-well plates

The *S. coelicolor* nutritional treatment screen (Figure 2.14) contained a mix of rich and poor media with different carbon and nitrogen sources. At 5 d the only treatment with no aerial hyphae was M19. Grey spore pigment indicates mature spores in *S. coelicolor*, indicating that the differentiation and sporulation process was more rapid on the oatmeal (OM) and MYM, and 2 of 3 replicas with international *Streptomyces* project 5 medium (ISP5). The lack of consistent aerial hyphae in M19 continues through to 10 d, as though they were present by this time it was not across all replicas. Spore pigmentation differed in MYM and proline yeast-malt (ProYM), the latter of which still presented white aerial hyphae as opposed to MYM's dark grey pigment.

One of the minimal mannitol medium (MMM) cultures strongly expressed actinorhodin compared to the other two replicates, but at the expense of aerial hyphae in the centre of the colony. *Streptomyces* is known to rapidly mutagensise the antibiotic-BGC containing regions of its genome (Hopwood, 2006), which may be the reason for the difference to other MMM replicas at 10 d.

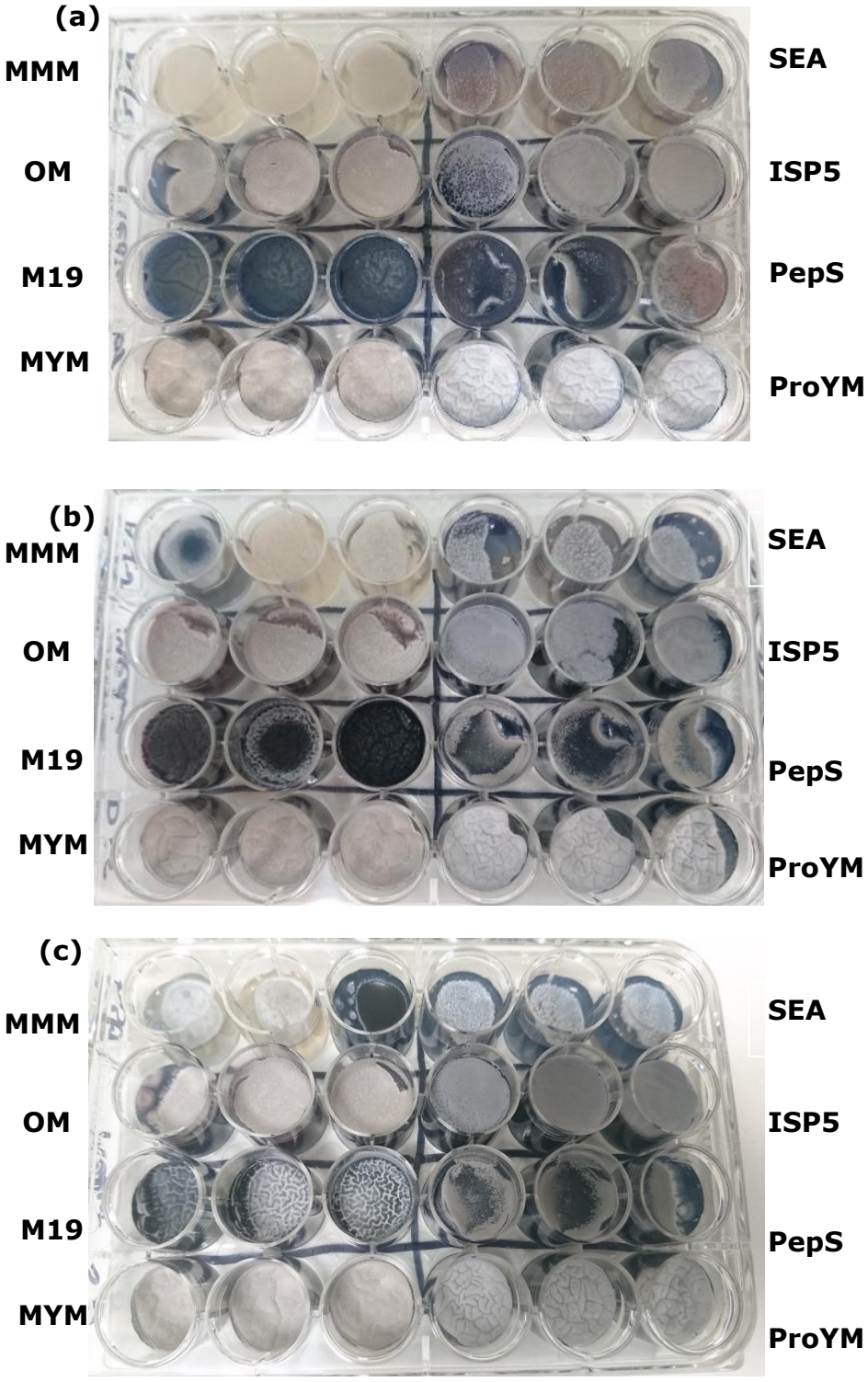
By 15 d ISP5 culture showed uniform dark grey spores, but some evidence of blue pigment droplets seen in previous time points remained in the leftmost well. These droplets, or extracellular vehicles, can contain a variety of proteins and peptides (Schrempf *et al.*, 2011) including antimicrobials. Aerial hyphae in M19 growth was much more consistent between replicas than at previous points and appears to be forming in-between the wrinkled colony surface peaks rather than uniformly across the whole colony surface. Like 10 d, another MMM replica produced much more ACT than the other 2 wells with less aerial hyphae, mimicking M19 culture in its productivity and differentiation. Plating out this colony would have been an interesting experiment to see if this was actually due to mutation and maintained across generations.

After extraction, samples were processed with MZmine (Pluskal *et al.*, 2010; Myers *et al.*, 2017) to remove media, non-triplicate, and <1E3 height features. An extracted ion chromatogram post-MZmine processing is shown in Figure 2.15, and a list of all detected *S. coelicolor* features in Figure 2.16. There are unhighlighted peaks in the chromatogram, which does not necessarily indicate an undiscovered metabolite from the model species, as this LC-MS method consistently results in a noisy chromatogram from 6 min onwards especially due to *m/z* 383 and 365, as well as peaks for removed media components.

The MZmine final peak list contains 122 features, of which 10 have putative identifications. Metabolites were identified using a custom database from global natural products social molecular networking (GNPS), other published data, and standard compound samples (Table 2.7). There were features with *m/*z values similar to previous albaflavenone UPLC-MS detection (Čihák *et al.*, 2017), but no fragment masses were provided so albaflavenone was not listed as identified. Comparison spectra are shown in Appendix figure 2 to Appendix figure 8.

Metabolites or adducts below the detection limit (>3 in a row 1E3+ height ions) are not listed but may still be present: MYM 10 contains CDA 4a, along with its Na and K adducts, and CDA 3a. Similarly, CDA is not listed as produced under PepS culture as it is present at <1E3 (mean 4.63E2 height).

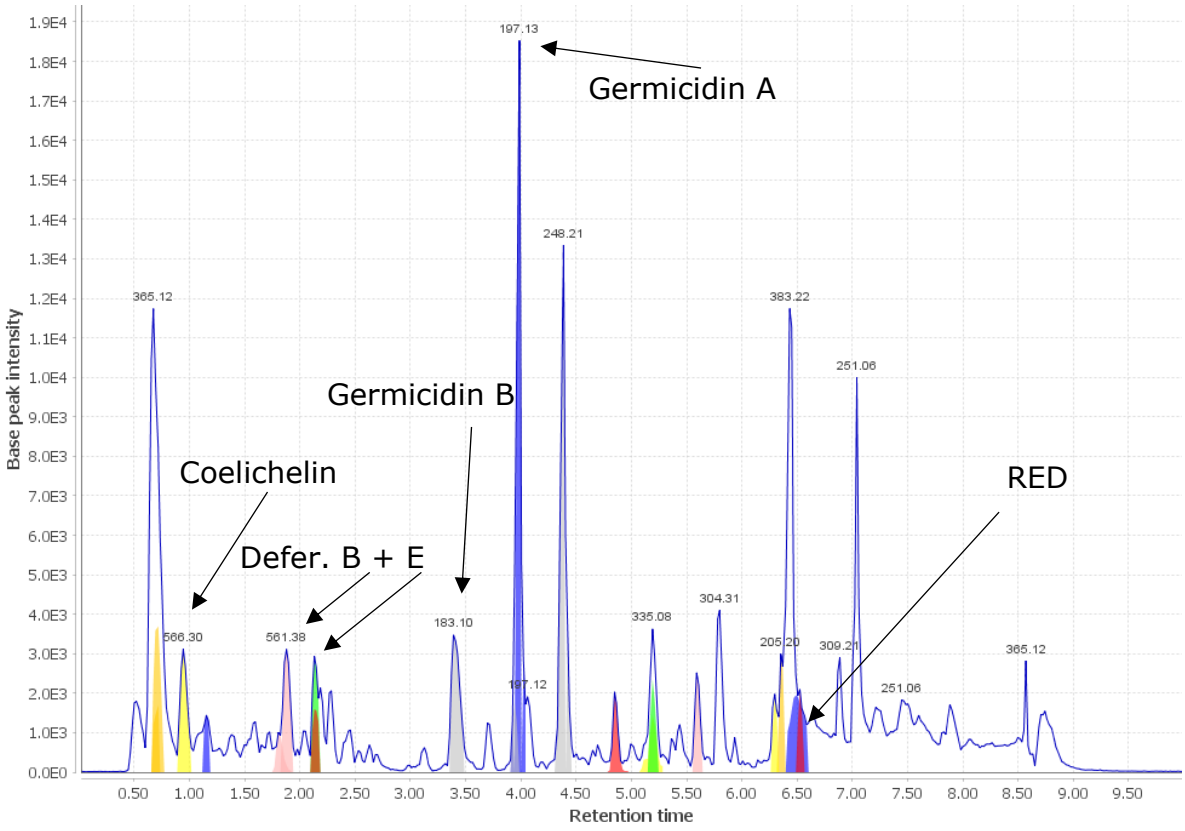
Figure 2.14 Culture of *S. coelicolor* in multiple media over **(a)** 5, **(b)** 10, and **(c)** 15 d. The plates are split into 8 media shown to the left or right of each triplicate set. Appendix table 1 lists media components. MMM = minimal mannitol medium; SEA = soil extract agar; OM = oatmeal medium; ISP5 = international *Streptomyces* project 5; M19 = medium 19; PepS = Peptone sucrose; MYM = maltose yeast malt; ProYM = proline yeast malt.



2.3.3.2 Analysis of metabolite variety

One hypothesis was that the rich media would lead to a large number of total features with fewer unique features as a percentage of the overall production, and the opposite for poor media as the starvation response or lack of carbon catabolite repression upregulates cryptic BGCs. This does not seem to be correct: the nutritionally poorest medium, soil extract agar (SEA), had the lowest total and 3rd lowest number of unique features, 27.27% of its overall total (Figure 2.16). However, the 2nd lowest total and lowest unique feature count (0) came from growth on PepS, which contains 20 g/l of starch as its carbon source, not a preferred carbon source for *S. coelicolor*. Three relatively rich media – MYM, ProYM, ISP5 – were 5th, 6th, and 7th lowest respectively by percentage uniqueness (10.00%, 9.52%, 6.45%).

Figure 2.15 MZmine extracted ion chromatogram for MYM 5 d culture

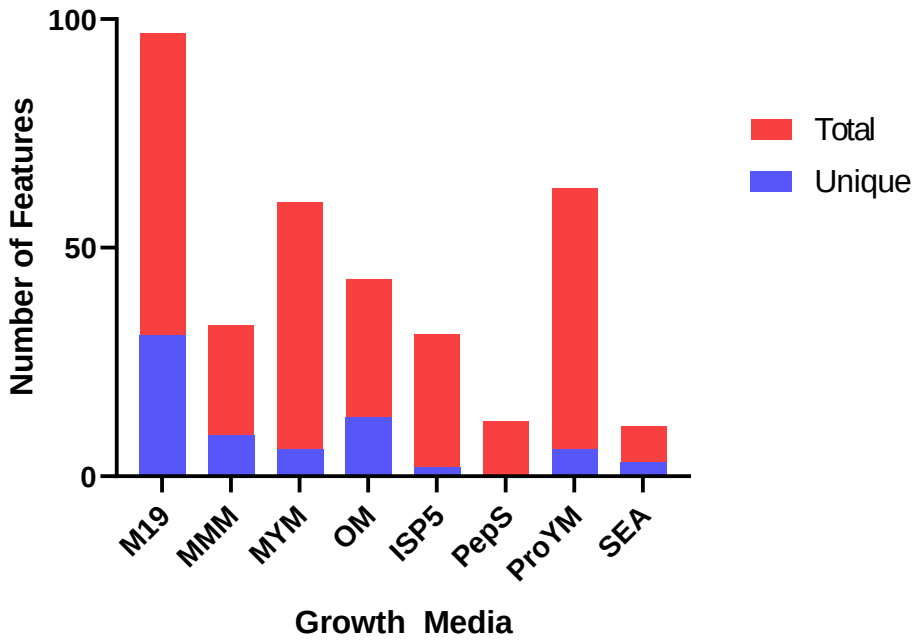


The largest number of unique features was obtained from growth on M19. It also produced the largest amount by relative peak area of germicidins A and B, deferoxamines E and B, RED, and 2nd largest amount of coelichelin. The relatively high numbers for 6 of the 9 identified metabolites compared to other media does fit with the hypothesis in 2.3.1 regarding DasR levels, since this would lead to an increase in specialised metabolism. Mannitol is not a preferred carbon source by *S. coelicolor*, with a doubling time of 6.1 h on mannitol media compared to 1.9 h with glucose (Hodgson, 1982).

2.3.3.3 Specific metabolite concentrations through peak area

ACT was visually noted as being produced in the majority of cultures. However, LC-MS analysis did not show actinorhodin at the levels it was assumed to be being produced at. Culture on SEA produced the most, but visually showed a far less intense pigment than other cultures which according to the LC-MS results contained no actinorhodin. Examining the UV absorbance for a sample with detected ACT (SEA 10 d) shows absorbance at 522 nm, which is replicated in M19 at the same retention time.

Figure 2.16 Total and Unique >1E3 intensity (height) features detected from OSMAC *S. coelicolor* culture.



A similar absorbance spectrum was recorded by Coisne *et al.* (1999) at 535 nm after extraction with acidic methanol. The equivalent peak in photodiode array data may be a more accurate way to determine ACT levels. Unfortunately, searching the sample assumed to have the most ACT – M19 culture 15 d – for 535 nm reveals a 2 min long additional peak which is clearly not related to a specific compound in a 10 min gradient elution. Two ACT *m/z* peaks with different retention times were detected in culture on SEA (4.69 min, 5.50 min). The earlier peak was designated as ε-ACT due to the presence of *m/z* 649.15 spectra, which is 0.02 *m/z* off the GNPS ε-ACT example. Examining a method in the literature reported to have detected ACT (Čihák *et al.*, 2017) does show some similarities, as both use the same solvents with a similar gradient in positive ionisation mode. The desolvation temperature was lower at 350°C, but even if this were the causative difference it would not explain why ACT was detected in some samples but not others.

RED was also not widely detected. The time sample points may not be suited to RED, as previous research found RED correlates with biomass accumulation, unlike a stereotypical specialised metabolite which will be found at stationary phase onwards (Hobbs *et al.*, 1990). Despite this, comparison to a 10 µg/ml standard showed similar peaks and spectra in the culture extracts at the two later sample times (M19 and MYM).

The detection time in later samples could fit with the previously mentioned work on RED and programmed cell death (Tenconi *et al.*, 2018), where RED is associated with the breakdown of the substrate hyphae to fuel differentiation into aerial hyphae. As aerial hyphae were not seen on M19 culture10 d, if RED is strongly associated with aerial hyphae then it would follow that RED would not appear until PCD is occurring to fuel aerial hyphae.

Each of the siderophores (coelichelin, deferoxamines B and E) were detected in the same culture media extract, with some difference between time points. Both B and E were produced at higher concentrations in M19, followed by ProYM and MYM. Relative amounts between time points were also similar across media for B and E, with most of the siderophores being expressed between 5 and 10 d, indicating a steady rate of expression. Coelichelin time point production differed between culture on different media. From *S. coelicolor* culture on M19,

approximately 12.5% is produced up to 5 d, after which a further approximately 62.5% of the total is synthesised, then the last 25% by 15 d. On MYM culture, the majority (~75%) is produced by 5 d, though the total amount by 15 d is just over half that of M19 culture. A graph showing total relative peak areas for all metabolites across all time points is presented in Figure 2.17, and individual metabolite graphs in Figure 2.18. ε-Actinorhodin is not shown in an individual graph as it was only detected in SEA at 15 d. MYM at 15 d could not be determined in triplicate as 2 files were unusable so has no error bars.

Table 2.7 Features with *m/z* matching previously reported *S. coelicolor* metabolites based on MZmine processing.

Potential Compound	**Observed *m/z* (M+H unless stated)**	**1E3+ Height Culture Parameters**	**Confirmed by**
CDA 4a	1495.60	MYM 5, 10; ProYM 5, 10	GNPS spectra
Deferoxamine B	561.38; 614.30 [M+Fe-2H]	M19 5 – 15; MYM 5, 10; ProYM 10, 15	GNPS spectra
Deferoxamine E	601.38	M19 5 – 15; MYM 5, 10; ProYM 5 - 15	GNPS spectra
Deferoxamine B [M-2H+Fe]			
Coelichelin	566.30	M19 5 – 15; MYM 5, 10; ProYM 5 - 10	GNPS spectra
ε-Actinorhodin	649.14	SEA 5 – 15	GNPS spectra; secreted pigment
γ-Actinorhodin	631.13	SEA 5 – 15 ISP5 10, 15	GNPS spectra; secreted pigment.
Germicidin A	197.12	**NOT** (MMM 5; PepS 5, 10; SEA 5 – 15); in all others.	MS^1 spectral comparison to Čihák *et al.*, 2017
Germicidin B	183.10	**NOT** (MMM 5; PepS 5; SEA 5 – 10); in all others.	Unconfirmed, likely given rt and spectral similarity to germicidin A
Undecylprodigiosin	394.22	M19 10, 15 MYM 5	Standard sample MSE spectra; GNPS

Figure 2.17 Total relative peak areas of detected metabolites from *S. coelicolor* culture over 15 d.

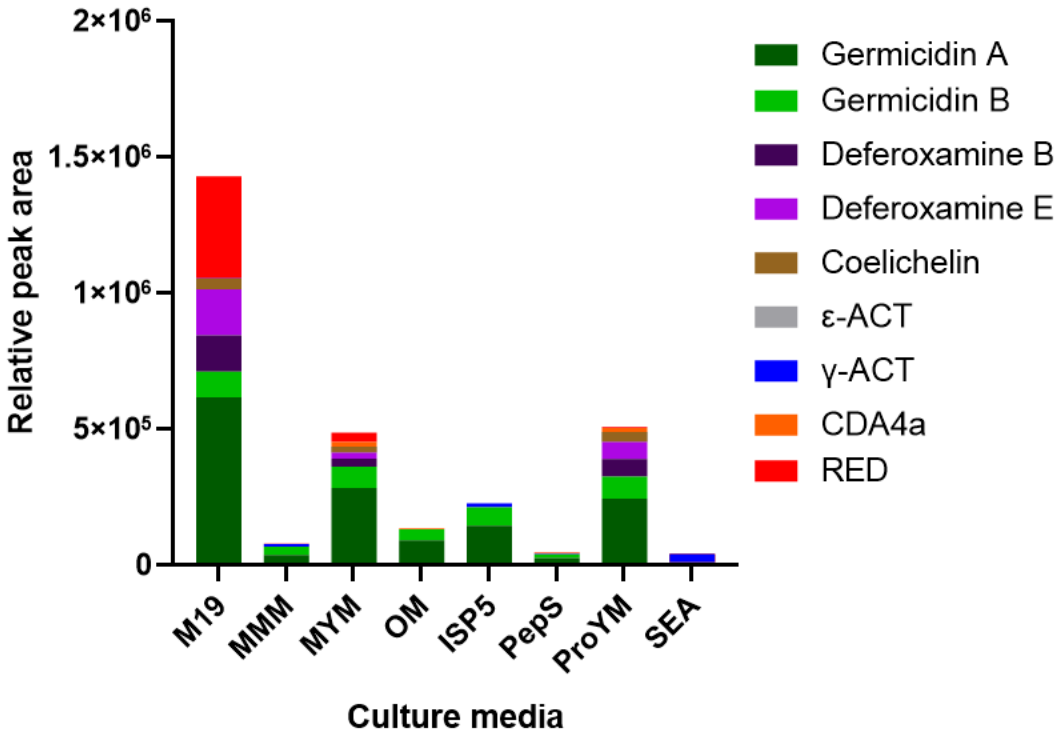
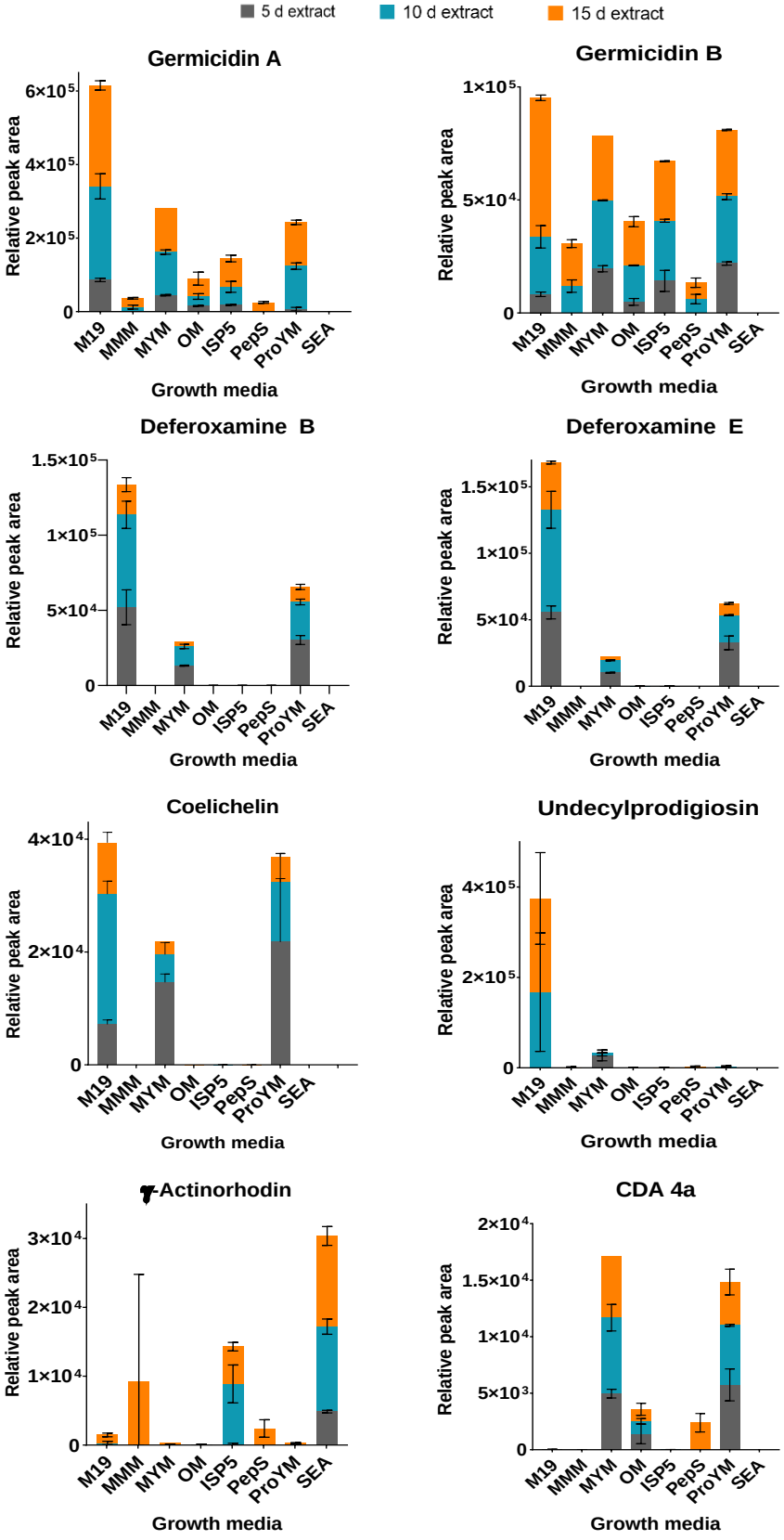


Figure 2.18 Relative peak areas of detected metabolites from OSMAC culture over 5, 10, and 15 d.



The OSMAC 24 multi-well plate screen showed clear advantages over standard petri dish culture and extraction. The initial failure of the time-based screening did lead to the more efficient well culture method so was useful in that respect, and the extra efficiency in use of space, chemicals, inoculation material, comparison between conditions, and easy extraction made the well screen a superior OSMAC method compared to the first experiment.

While it is impossible to optimise LC-MS conditions for all the metabolites in an untargeted metabolomics experiment, that at least 9 metabolites were detected from *S. coelicolor* across a range of retention times or polarities indicated that the method was applicable to other *Streptomyces*. Therefore, it was selected as the main method for OSMAC-based metabolomics for the NCIMB *Streptomyces.*

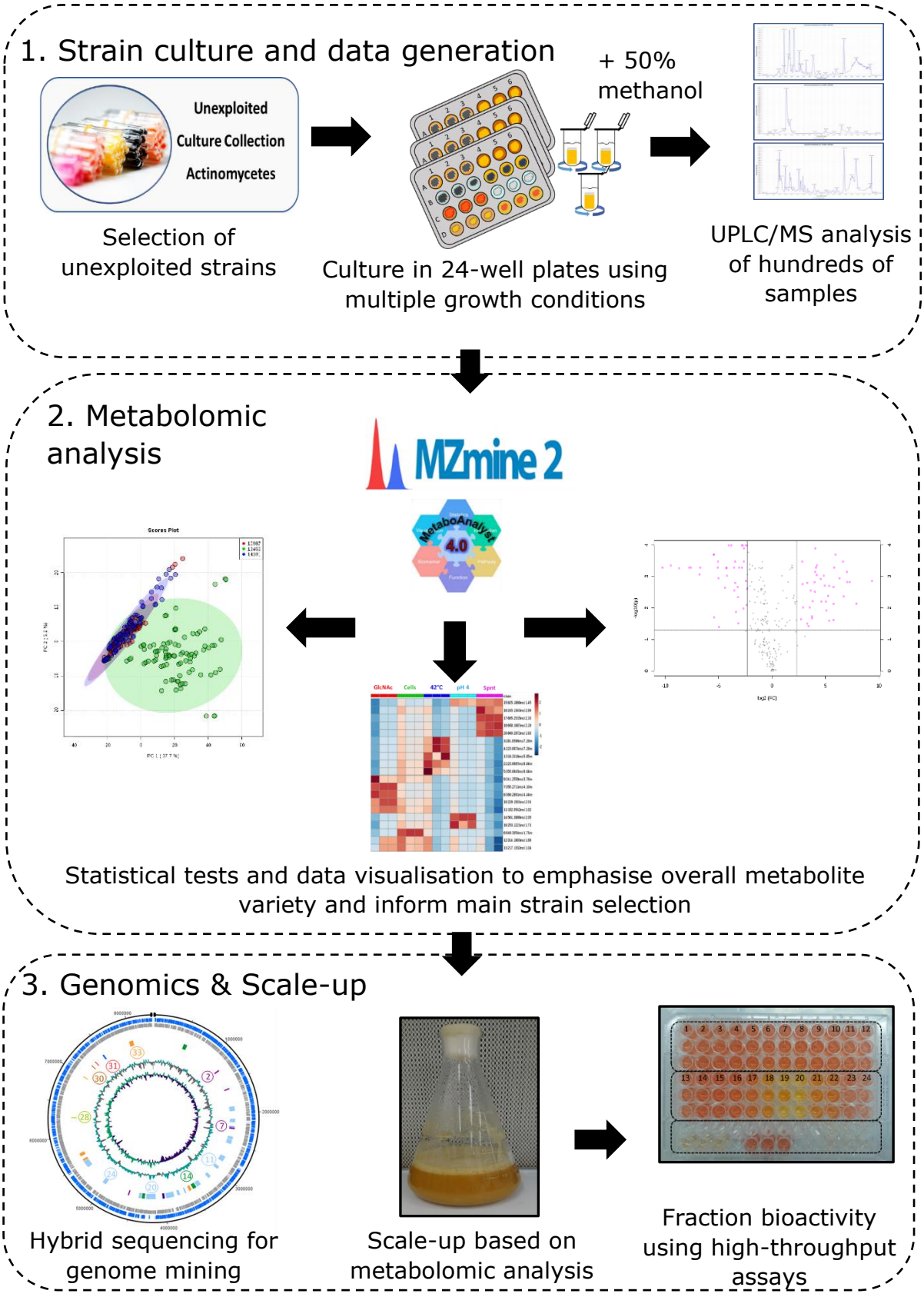
2.4 Conclusions

The aim for this chapter was to develop a metabolomic method which could efficiently expand the specialised metabolome of a strain, resulting in identifiable metabolites which could be compared across growth conditions. Using *S. coelicolor* to test the method was a logical choice given the amount of SMet research into it and its products. Production of different metabolites was clearly observed under different regimes, from initial visual identification of ACT in a range of biotic and predominantly abiotic growth treatments. The switch from standard agar plates to a more efficient well-plate culture was triggered by the co-culture experiments, as they demonstrated the applicability of a small-scale system testing multiple growth treatments while using non-model *Streptomyces* strains.

However, the LC-MS conditions used were not appropriate for each of the compounds as the difference in ACT levels showed when comparing visual intensity to its chromatogram intensity. Not being able to fully identify every class of compound is an inherent risk in an untargeted metabolomics project, which is compensated for by the detection of multiple other *S. coelicolor* metabolites from a range of different chemical classes. The metabolites present in the sample will been affected by the choice of extraction solvent, 50% aqueous methanol, which both targets the relatively unexplored higher polarity chemical space while providing more disruption to the cell for extraction than would be caused by pure water.

While metabolomics is an important tool for studying microbial natural products, it is only one of multiple approaches to elicit production of novel metabolites. Genomics-based methods are widely applied, using a full genome sequence to predict BGCs and so alter strain prioritisation or target expression of a BGC. Here, the metabolomic screening can be used to identify a specific non-model *Streptomyces* strain for further examination, including full genome sequencing. The combination of these methods, shown in an overall suggested workflow for metabolite discovery in Figure 2.19, presents a powerful multi-omics method for expanding metabolite variety and discovery of novel bioactive metabolites.

Figure 2.19 Proposed overall workflow for microbial natural products discovery.



CHAPTER 3

METABOLOMICS-BASED PROFILING AND SELECTION OF UNEXPLOITED *STREPTOMYCES* STRAINS

3 Metabolomics-based profiling and selection of unexploited *Streptomyces* strains

3.1 Introduction

3.1.1 Selection of unexploited *Streptomyces* strains

The NCIMB collection contains hundreds of *Streptomyces* from diverse ecological niches. Many of these strains have minimal genomic or specialised metabolite information. Therefore, rational strain selection for metabolomic screening and eventual culture scale-up required a mixture of methods to justify picking a small number of *Streptomyces* out of the whole collection. Previously used relevant non-genomic methods to examine *Streptomyces* metabolite variety include isolating from non-soil niches such as marine environments (Pereira *et al.*, 2014), the One Strain Many Compounds (OSMAC) method (Bode *et al.*, 2002), and classical bioactivity screening (Fedorenko *et al.*, 2015). These methods are directly applicable to guiding strain selection. Combining these methods for strain selection along with a high-throughput metabolomics screening procedure provides a pipeline for enhanced metabolite variety by exploiting unexamined strains in the NCIMB culture collection.

3.1.2 Untargeted metabolomics analysis of unexploited strains

The metabolomic profile of a strain can be an important tool in analysing the metabolomic output of single strain or a larger group of microbes. Some applications of profiling based on *Streptomyces* culture are given below.

3.1.2.1 Metabolomic profiling of symbiotic *Streptomyces* strains

A basic two-media OSMAC strategy was used by Viegelmann *et al.* (2014) to expand the accessible metabolome of *Streptomyces* sp. SM8, isolated from the marine sponge *Halicona simulans*. A combination of LC-MS and nuclear magnetic resonance (NMR) was used, reasoning that the two techniques are complimentary to one another: UPLC-MS can be much more sensitive, but the ionisation method may not pick up certain classes of metabolite. In comparison,

NMR is less sensitive but can convey a more complete metabolome as it is not biased towards metabolite classes.

After identifying antifungal compounds that were upregulated by one medium, Viegelmann *et al.* (2014) used principal component analysis (PCA) to test whether metabolites produced by *Streptomyces* sp. SM8 were also present in their host. Displaying the metabolites through scatter plot analysis showed two mostly separate profiles, with investigated antifungals unambiguously associated with *S.* sp. SM8 (reinforced by *Streptomyces* sp. SM8. having the corresponding biosynthetic gene cluster/BGC). There were some shared metabolites in both profiles. While this could indicate that some metabolites are produced by both species, it is also possible that the sponge is using bioactive metabolites produced by the *Streptomyces* symbiont.

Another sponge symbiont study examined the metabolomic profiles of *Streptomyces* and other microbes isolated from Mediterranean sponges (Cheng *et al.*, 2015). PCA of extracts was linked to anti-trypanosomal activity, with the assumption that bioactive outliers were more likely to produce novel metabolites. This led to the selection of *Streptomyces* sp. SBT348 for further study. MS/MS analysis could not identify the responsible bioactive compounds, indicating the production of an undiscovered anti-trypanosomal metabolite.

As part of a wide-ranging study on insect *Streptomyces* symbionts, Chevrette *et al.* (2019) studied the metabolome of a set of 120 strains including 69 insect symbionts. After culture on yeast peptone mannitol agar, metabolomic profiling showed significant metabolite diversity which corresponded to the diverse ecology and phylogeny of the strains. From the total metabolomic features produced, 59% were only found in a single strain. Using PCA to identify outliers led to the selection of *Streptomyces* sp. ISID311, isolated from a species of fungus-growing ant within the *Cyphomyrmex* genus. Examining which components contributed to profile separation led to the discovery of a novel antifungal metabolite, cyphomycin.

3.1.2.2 Stressors and elicitors to enhance metabolite production

An alternative approach to enhance discovery of novel metabolites is through the addition of a stressor or elicitor, producing a parameter-specific metabolomic

response. This can range from broadly applicable treatments like heat shocks through to addition of specific compounds that trigger upregulation of a certain pathway (Yoon & Nodwell, 2014).

Scandium chloride was originally used to enhance actinorhodin levels by up to 25 times from *S. coelicolor* by Kawai *et al.* (2007), though the mechanism of action of scandium and other rare earth metals to stimulate specialised metabolism is not known (Tanaka *et al.*, 2010). Applying scandium to growth media of other *Streptomyces* enhanced their metabolite production: actinomycin levels approximately doubled in *S. antibioticus*, and streptomycin from *S. griseus* tripled. Unfortunately, as with any OSMAC project, not every condition will stimulate specialised metabolism. Scandium did not have the same effect on each strain, so any effects on *S. costaricanus* metabolite production cannot be predicted.

Other culture parameters like pH can also affect metabolite production, whether positively through upregulation of novel compounds, or negatively by causing too much stress for a microbe. Kontro *et al.* (2005) also found that optimum pH is heavily dependent on the growth media used for an individual species, so as with scandium, it may be that use of a non-minimal medium facilitates growth at pH 10. As an alternative to altering the initial pH, acid and alkali shock have both been used to stimulate specialised metabolism (Kim *et al.*, 2007; Jiang *et al.*, 2018), but again this would probably use a non-minimal media starter culture. pH information at isolation would be useful to mimic ecological conditions but given environmental variability this may not reflect optimum conditions for novel metabolite production.

3.1.3 Chapter aim and objectives

Chapter 2 focused on the development of an untargeted metabolomics workflow that assessed metabolite diversity in *S. coelicolor*. The aim of this chapter was to apply this methodology to unexploited *Streptomyces* strains from the NCIMB culture collection to identify a suitable strain for further investigation. This was achieved with the following objectives:

1) Selection of a soil-isolated *Streptomyces* strain using initial bioactivity assays against Gram-positive and Gram-negative bacteria.

2) Selection of 2 non-soil *Streptomyces* strains from distinct niches.
3) Comparison of total metabolomic diversity from the 3 strains when cultured using the 24-well plate screen.
4) Identification of the strain producing the most different metabolomic profile compared to the other 2 *Streptomyces*, selecting this strain for further metabolomic screening and bioactivity assays.
5) Analysis of metabolomic data for this single strain to identify which culture parameters most enabled metabolite diversity.

3.2 Materials and Methods

3.2.1 Selection of NCIMB *Streptomyces* strains

Streptomyces strains from the NCIMB collection without any isolation information were excluded, along with any model or industrially common species such as *Streptomyces rimosus* (NCIMB 9980), isolated from bee intestines, as it is widely used to produce tetracyclines (Petkovic *et al.*, 2006).

Strains selected in section 2.2.2, soil strains with minimal published specialised metabolite information and different soil niches, were reused. Non-soil strains were chosen using similar criteria, though the selection pool was much smaller (Figure 3.2). The final 2 selected non-soil strains along with the previously selected soil strains are shown in Table 3.1.

Table 3.1 Soil *Streptomyces* strains used in antimicrobial activity assays, and non-soil comparison strains.

Strain	NCIMB Number	Isolated from	Year of accession
Soil isolates			
S. coelicolor Müller	9798	-	1965
S. halstedii	9839	"Deeper soil layers"	1966
S. pseudoechinosporeus	9918	Desert sand, Kazakhstan	1966
S. avellaneus	11000	Cave soil, Italy	1973
S. flaviscleroticus	11008	Soil, India	1973
S. purpureus	11313	Soil, Hawaii	1977
S. sp	12947	Lavender field soil, Hungary	1989
S. costaricanus	13455	Soil, Costa Rica	1997
S. rhizosphaericus	13674	Ectorhizosphere, Indonesia	1997
S. sanyensis	14820	Mangrove sediment, China	2012
S. aridus	14965	Atacama Desert, Chile	2014
Non-soil isolates			
Streptomyces sp.	14101	Marine sediment, Mariana Trench	2005
Streptomyces sp.	10987	PVC-coated wallpaper	1973

3.2.2 Antimicrobial activity screen of *Streptomyces* strains

The bioactivity of 12 NCIMB *Streptomyces* strains was measured against *E. coli* NCIMB 10236, *B. subtilis* NCIMB 10113, and Methicillin Resistant *S. aureus* NCIMB 50143 using an agar overlay assay. Spore suspension (20 µl) of each

strain were spotted onto agar plates of Tryptone Soya Agar (TSA), Nutrient Agar (NUT), OM, Bennett's Agar (BA), medium 19 (M19), medium 400 (M400), international streptomyces project medium 4 (ISP4), ISP5, Deutsche Sammlung von Mikroorganismen und Zellkulturen medium 5006 (DSMZ 5006), glucose yeast malt, and maltose yeast malt (GYM, MYM). Additional media recipes are found in Appendix section 8.2. Each plate was divided into quadrants to culture 4 *Streptomyces* at 30°C for 7 d. A flask of 50 ml nutrient broth was inoculated with a loop of each of the 3 test strains, and grown overnight at 37°C, 175 rpm. This preculture was added to 2 x 450 ml freshly autoclaved NUT agar once it was under 50°C, mixed, and poured over *Streptomyces* cultures to produce an indicator lawn. Zone of inhibition diameters were recorded after 24 h culture at 37°C. Only zones greater than 1 cm were included.

OM, M19, ISP5, GYM, and MYM ingredients are listed in Section 2.5.2. Other media are shown in Appendix table 2, except NUT and TSA which used the premade Sigma-Aldrich TSA and NUT mixes. All components were obtained from Sigma-Aldrich (Poole, UK)

3.2.3 24-well plate culture and extraction

3.2.3.1 24-well plate nutrition screen to assess metabolomic diversity

S. costaricanus NCIMB 13455 (*S. costaricanus*), *Streptomyces* sp. NCIMB 14101 (*Streptomyces* sp. MT5), and NCIMB 10987 (*Streptomyces* sp.) were selected based on initial bioactivity tests and isolate niche variety. Strains were grown as in section 2.2.6 with the following modifications:

1) Comparison plate with liquid media for each time point.
2) Time points reduced from 5/10/15 d to 5/10 d for agar and broth.
3) Soil extract agar, minimal mannitol medium, proline yeast malt and peptone sucrose media replaced with Sucrose-Meat extract (SM), Soybean-Glucose-Tryptone (SGT), Molasses-Meat extract (MM), and Corn steep liquor-Soya flour (CS) media. All solid media contained 20 g/l agar.

Additional media were formed from ingredients reported to upregulate specialised metabolism (Junker *et al.*, 1998; Tormo *et al.*, 2003; Fedorenko *et al.*, 2015). Compositions are found in Appendix table 3.

3.2.3.2 Extraction of metabolites from 24-well plate culture

Broth culture metabolites were extracted by adding 1 ml 50% aqueous methanol to each well, gently shaking to mix then leaving to sit for 5 min. A syringe pipette was then used to transfer 1 ml from each well to a respective 1.5 ml microcentrifuge tube. Agar cultures were extracted as section 2.2.6 up until the filtering stage. Both agar and broth extracts were then filtered by adding 500 µl of each to a corresponding Whatman Mini-UniPrep G2 Syringeless Filter UPLC Vial (GE Healthcare, Little Chalfont UK) and filtered using a Whatman Multicompressor. Samples were stored at -20°C until UPLC-MS analysis.

3.2.3.3 24-well plate culture using minimal media and elicitors

For minimal media and elicitor experiments, 24-well plates were used to culture *S. costaricanus* in the conditions given in Table 3.2 in triplicate. The minimal medium used is referred to as MMM and is distinct from the nutrition screen maltose-meat extract medium (MM).

Table 3.2 Minimal media growth conditions for *S. costaricanus*

	Condition	Elicitor purchased from
MMM+	No elicitor (just MMM culture)	-
	25 µM Scandium chloride	Sigma-Aldrich
	25 mM Sodium butyrate (NaBu)	Sigma-Aldrich
	10 mM N-acetyl glucosamine (GlcNAc)	Sigma-Aldrich
	MMM pH 4	-
	MMM pH 10	-
	Dead *A. baumannii* (100 µl)*	-
	A. baumannii culture supernatant (100 µl)	-
	MMM culture at 42°C	-

*Dead *A. baumannii* cells were prepared from *A. baumannii* NCTC 13301, a clinical β-lactam resistant strain, which was purchased from Public Health England and maintained on Muller-Hinton agar plates at 4°C.

A loop of *A. baumannii* NCTC 13301 colony was used to incubate in a 50 ml overnight culture of Muller-Hinton broth. The culture was transferred to a 50 ml centrifuge tube and centrifuged for 15 min, at 5000 x g. The supernatant was filtered through a 0.2 µm sterile filter into sterile 1.5 ml microcentrifuge tubes before adding to wells. Pellet was resuspended in 10 ml sterile 18.2 MΩ water and placed in a beaker of boiling water for 30 min. Heat killed cell suspension was then pipetted into wells. Each well was inoculated and extracted as the nutrition screen wells above.

3.2.4 Metabolite analysis by UPLC-MS

Samples were analysed using the *Streptomyces* specific set of parameters found in Section 2.2.8. The same MZmine (Pluskal *et al.*, 2010; Myers *et al.*, 2017) settings as Section 2.2.9 were also used. For the minimal media screen, samples of Muller-Hinton broth and *A. baumannii* supernatant were run with LC-MS to remove media or unwanted metabolites in MZmine.

3.2.4.1 Actinomycin quantification

A 10 µg/ml actinomycin D standard was analysed by UPLC-MS. This was used to quantify other actinomycins through comparisons to the extracted ion peak area of actinomycin D.

3.2.5 Statistical analysis

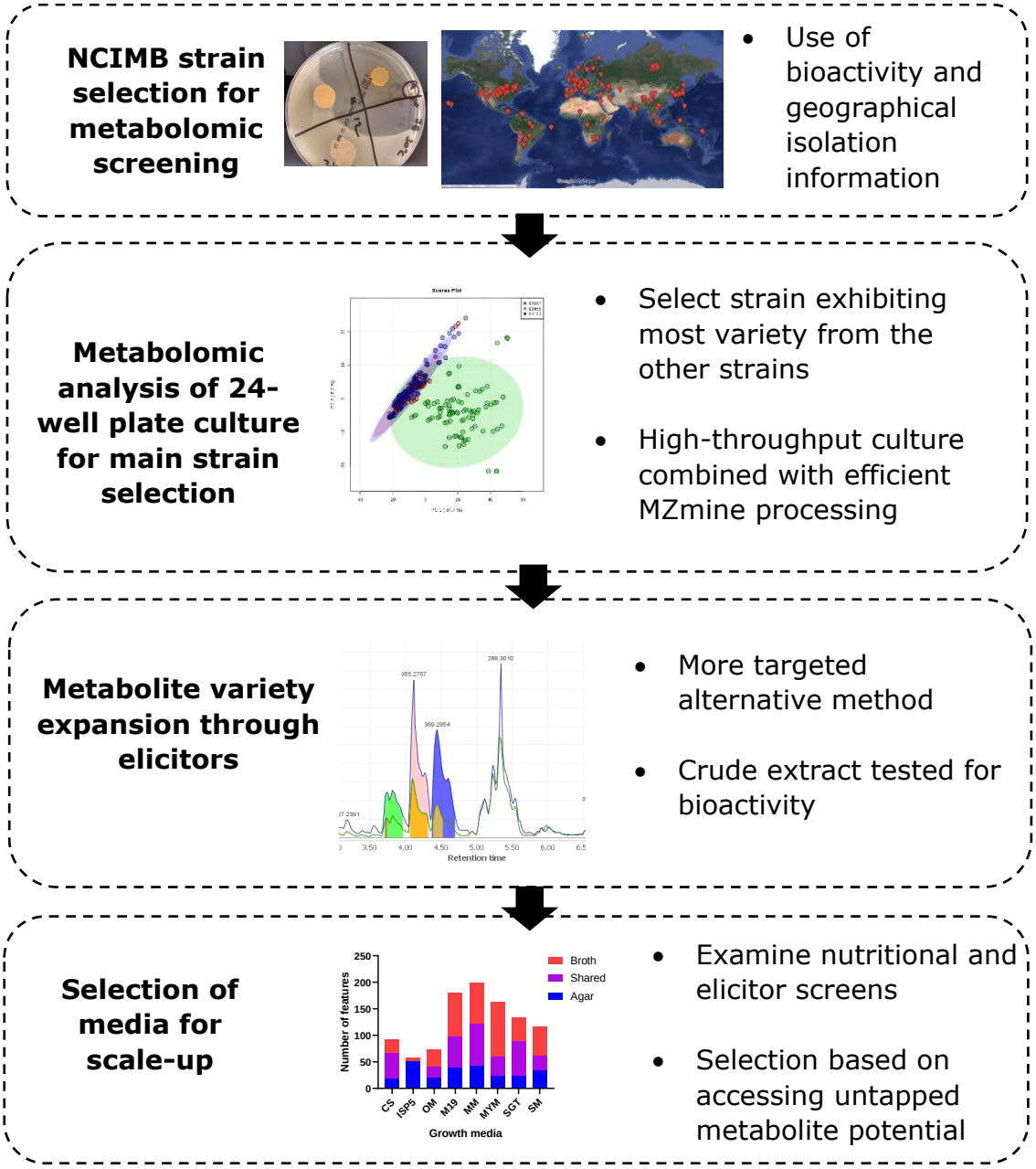
Initial PCA was achieved using the MZmine PCA setting on the processed combined peak list for all 3 strains, using feature Peak Area for the peak measurement setting. Further statistical analysis on *S. costaricanus* was performed by exporting the MZmine peak list to MetaboAnalyst v4.0 (Chong *et al.*, 2018). Settings used were default missing value settings, features filtered by Interquartile Range, normalised by sum, and data transformed using glog,

adapting the recommended or well performing untargeted metabolomics processing methods found by Di Guida *et al.* (2016).

3.2.6 Overview of proposed methods

This chapter builds on the 24-well plate culture and extraction procedure developed through chapter 2, applying it to 3 non-model *Streptomyces* to catalogue their metabolite variety and use it to inform further strain and culture parameter selection. The methods are summarised in Figure 3.1.

Figure 3.1 Summary of methods used to screen non-model *Streptomyces* strains.



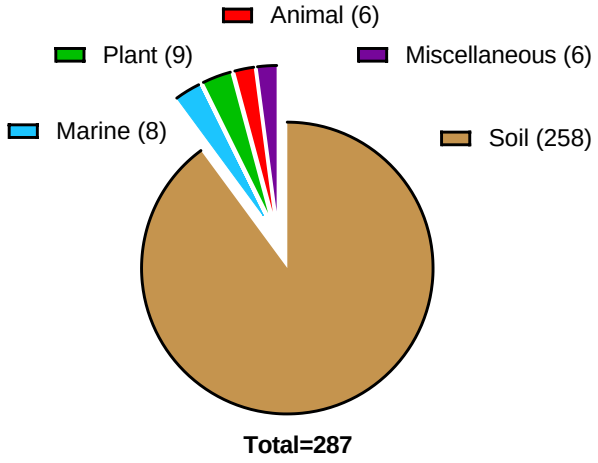
3.3 Results and Discussion

3.3.1 Selection of NCIMB *Streptomyces* strains

3.3.1.1 Geographical variety of isolates

At the time of writing there were 469 *Streptomyces* strains in the NCIMB collection. Of these, 287 have some level of isolation information, from just "soil" through to a specific country, elevation, and building (NCIMB 11762, "Soil…from Atacama Large Millimeter Array observatory site 4000 m above sea level, Atacama Desert, Chile"). A basic overview of niche type numbers is shown in Figure 3.2.

Figure 3.2 Breakdown of NCIMB *Streptomyces* strains with niche isolation data.



The older strains, generally deposited towards the end of the 1960s, frequently lack geographical information which biased the selection towards more recent isolates. The majority of continents have isolates taken from them, with Antarctica the only one missing from the collection. *Streptomyces* have been isolated from Antarctica, but only recently and in small numbers (Núñez-Montero *et al.*, 2019).

3.3.1.2 Antibacterial activity screen of selected soil isolates

A preliminary bioactivity screen was used to inform which soil *Streptomyces* would be selected for metabolomic screening. The soil strains were assayed against 3 test species: *Escherichia coli, Bacillus subtilis,* and methicillin resistant *Staphylococcus aureus*. Media and strains showing bioactivity against each of the test bacteria are shown from Figure 3.3 to Figure 3.5. Only strains showing bioactivity are included in the figures.

Bioactivity was more common against the two Gram-positive species: of the 121 conditions (11 strains, 11 media) tested, 10 (4 species) produced activity against *E. coli*, 18 (6 strains) against *B. subtilis*, and 27 (7 strains) against MRSA. Not all strains were active against both Gram-positive species even if growth of 1 of the 2 was inhibited. As an example, *S. costaricanus* inhibited MRSA after growth on MYM and GYM but did not inhibit *B. subtilis* under the same culture conditions. This implies some level of specificity in the antibiotics secreted.

Figure 3.3 Strains with bioactivity against *E. coli*.

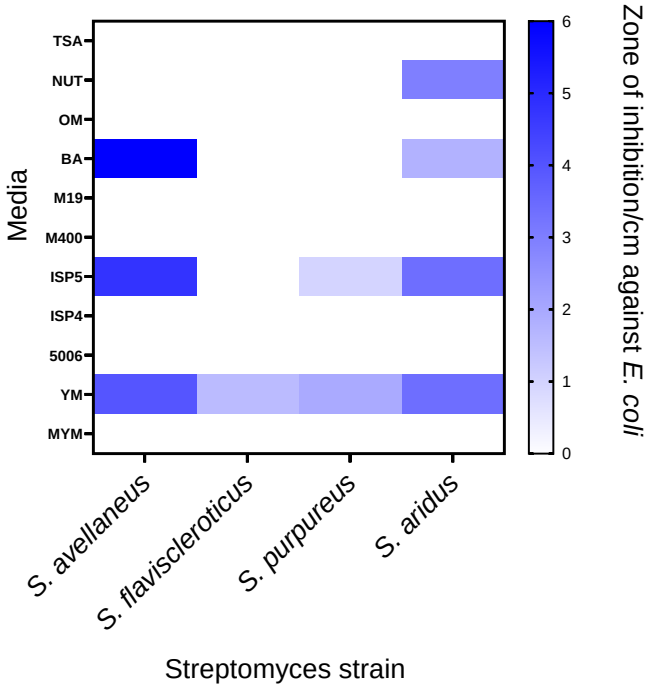


Figure 3.4 Strains with activity against *B. subtilis*

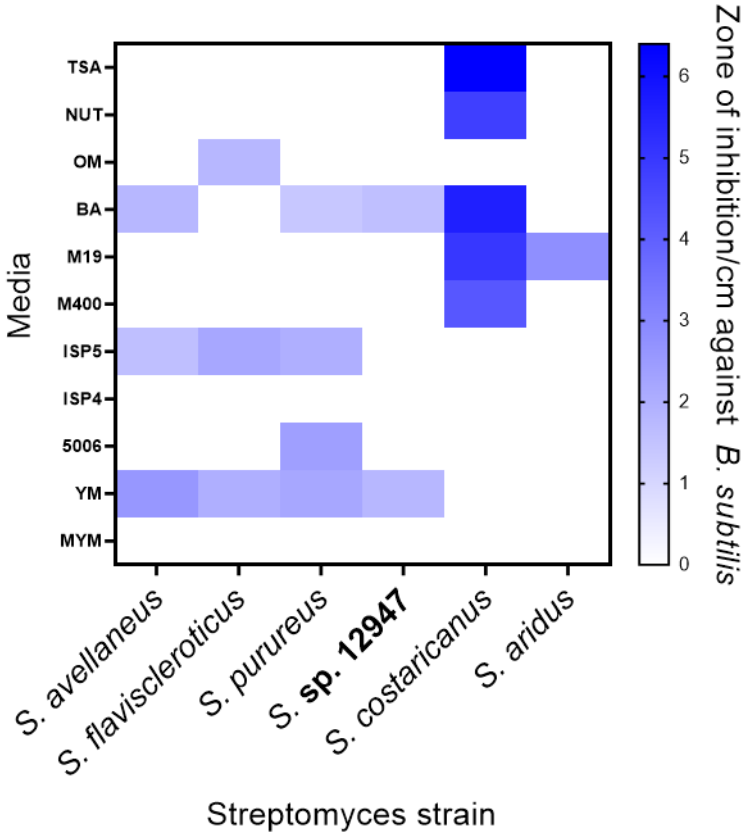
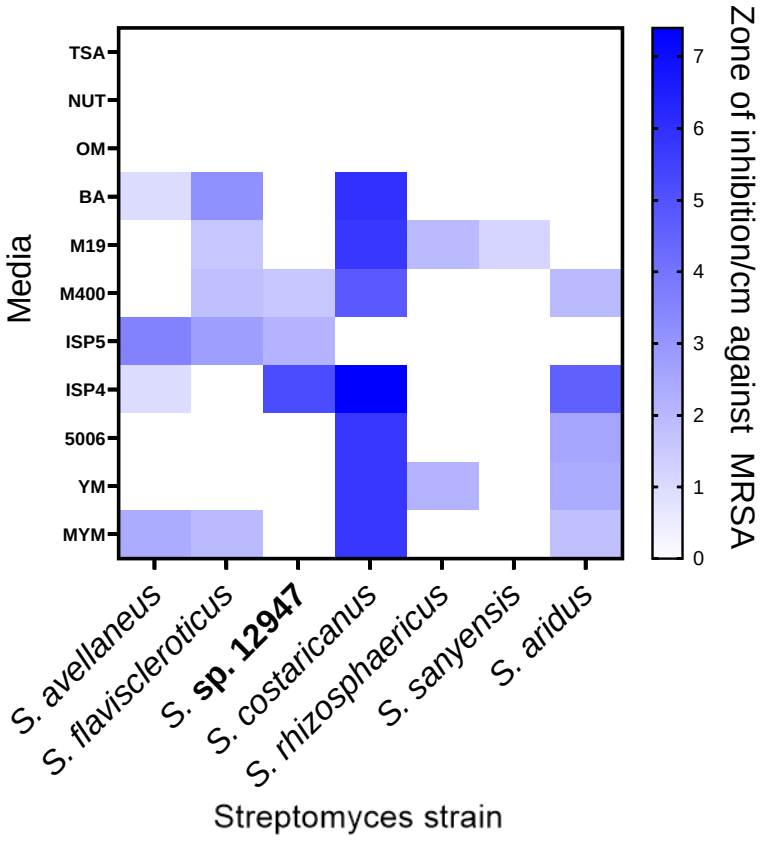


Figure 3.5 Strains with bioactivity against MRSA



3.3.1.3 Antibacterial metabolites produced by bioactive strains

Streptomyces aridus, isolated from the ultra-arid Atacama Desert, was most varied on media in activity against *E. coli,* producing 4 zones of which 3 were 3 cm or larger in diameter. The largest zone was produced by *S. avellaneus* (NCIMB 11000) on Bennett's Agar. This strain is a known chlorotetracycline producer (Ward & Allenby, 2018), a broad spectrum antibiotic active against Gram-positives and negatives.

S. avellaneus also produces chromomycins (Kawano *et al.*, 1990), which are active against tumour cells and Gram-positives. Gram-negative activity is not reported, potentially due to difficulties crossing the Gram-negative membrane (Kamiyama & Kaziro, 1966), so is unlikely to be the active compound against *E. coli* unless combined with an unknown emulsifier . Running the equivalent strain *S. avellaneus* NRRL B-3447 (accession ID NZ_JOFK00000000) through antiSMASH v5 (Blin *et al.*, 2019) to predict other potentially bioactive compounds was not useful as the shotgun genome sequence was too fragmented for robust predictions. Activity from *S. avellaneus* was seen from growth on BA, ISP5, and YM. Both BA (6 cm inhibition) and ISP5 (4.8 cm inhibition) have glycerol as a carbon source, potentially indicating antimicrobial upregulation with glycerol growth media.

Both Gram-positive species were inhibited by a wider set of *Streptomyces*. *S. costaricanus* had the largest zone of inhibition against MRSA (7.4 cm on ISP4) and *B. subtilis* (6.4 cm on TSA), the second of which is shown in Figure 3.6. *S. costaricanus* also consistently produced a zone on 7 of 11 media, with the lowest potency on M400 (4.8 cm). No other strains produced activity on as many media as *S. costaricanus*.

At the time of the experiment, no specific antimicrobial compound information was known about this species past its association with "nematode-suppressive soil" (Esnard *et al.*, 1995), and inhibition of the plant pathogens *Rhizoctonia solani* and *Phytium aphanidermatum* (Goodfellow *et al.*, 2012). Liu *et al.* (2019) later published on a marine strain of *S. costaricanus* which was found to produce actinomycins D, X0β, and X2 which could explain the bioactivity against other Gram-positive species and orange pigmentation seen in Figure 3.6.

Gram-negative species are more resistant to antibiotics than Gram-positives due to the presence of an outer membrane in Gram-negative species (Tsuchido & Takano, 1988). The outer membrane is comprised of a mix of phospholipids and lipopolysaccharides forming the main bilayer, with proteins performing various functions like porins and receptors placed throughout the layer (Delcour, 2009). This acts as a barrier to antibiotics that are active only when intracellular.

The outer lipopolysaccharide layer is highly impermeable to most molecules, and diffusion is mostly through porin channels which provide a clear resistance target for natural selection. Though 11 strains and media against 1 Gram-negative and 2 Gram-positive species is not a large enough sample to properly represent of intrinsic resistance, the fewer positive results against *E. coli* fits the expected trend. Other potentially bioactive metabolites produced by the inhibitive species are listed in Table 3.3. The DSMZ equivalent strain page for *S. flavisclerotícus* list (DSMZ, 2019) lists it as producing aburamycin and "antibiotic M5-18903", but no corresponding literature could be found to confirm this.

Figure 3.6 Inhibition of *B. subtilis* caused by *S. costaricanus* on TSA, showing orange pigmentation. Actinomycin is a red/dark orange colour.

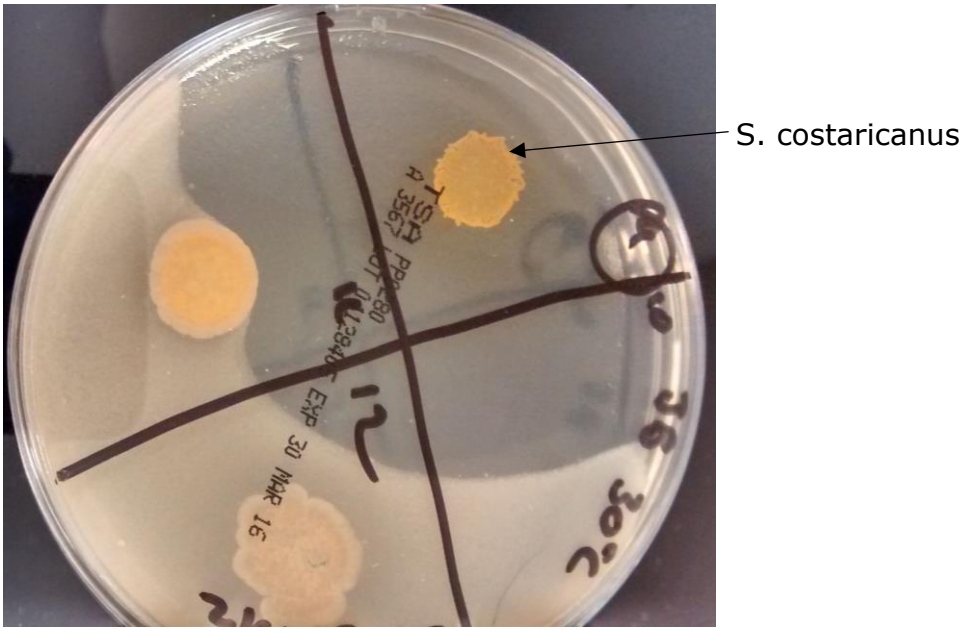


Table 3.3 Compounds produced by inhibitive strains against Gram-negatives and Gram-positive bacteria.

Species	Metabolites inhibiting Gram-negative growth
Streptomyces avellaneus	Chlorotetracycline
*Streptomyces flavisclerotic*us	-
Streptomyces pururеus	Chloramphenicol (Goodfellow *et al.*, 2012)
	Metabolites inhibiting Gram-positive growth
Streptomyces avellaneus	Chlorotetracycline, chromomycins
Streptomyces costaricanus	Actinomycins (Liu *et al.*, 2019)
*Streptomyces flavisclerotic*us	Chromomycins (Prajapati *et al.*, 2019)
Streptomyces pururеus	Chloramphenicol
Streptomyces rhizosphaericus	-
Streptomyces sanyensis	Potentially staurosporine (Sancelme *et al.*, 1994; Li *et al.*, 2013)

3.3.1.4 Selection of soil strain

S. costaricanus did not show any activity against *E. coli*. However, classical bioactivity-based screening is also not suitable for conveying the full range metabolites a species produces under even one growth condition. One bioactive molecule, likely one already discovered if using standard lab growth conditions, can hide the presence of any other similarly bioactive compounds, especially those expressed at low levels. A cryptic BGC coding for a highly bioactive metabolite may be expressed at a low level, rather than totally silent, requiring

scaleup and fractionation to isolate. Therefore *S. costaricanus* was chosen as the soil *Streptomyces* for specialised metabolite diversity evaluation.

S. costaricanus produced at least one bioactive antimicrobial in over half of the media against MRSA, and 5/11 against *B. subtilis*. Given the high rates of bioactivity it was likely that at least some of the compounds were previously identified. This also makes *S. costaricanus* a good choice for further analysis, since any identified compounds can be used to benchmark production of unknowns and help identify bioactive metabolites against Gram-negatives.

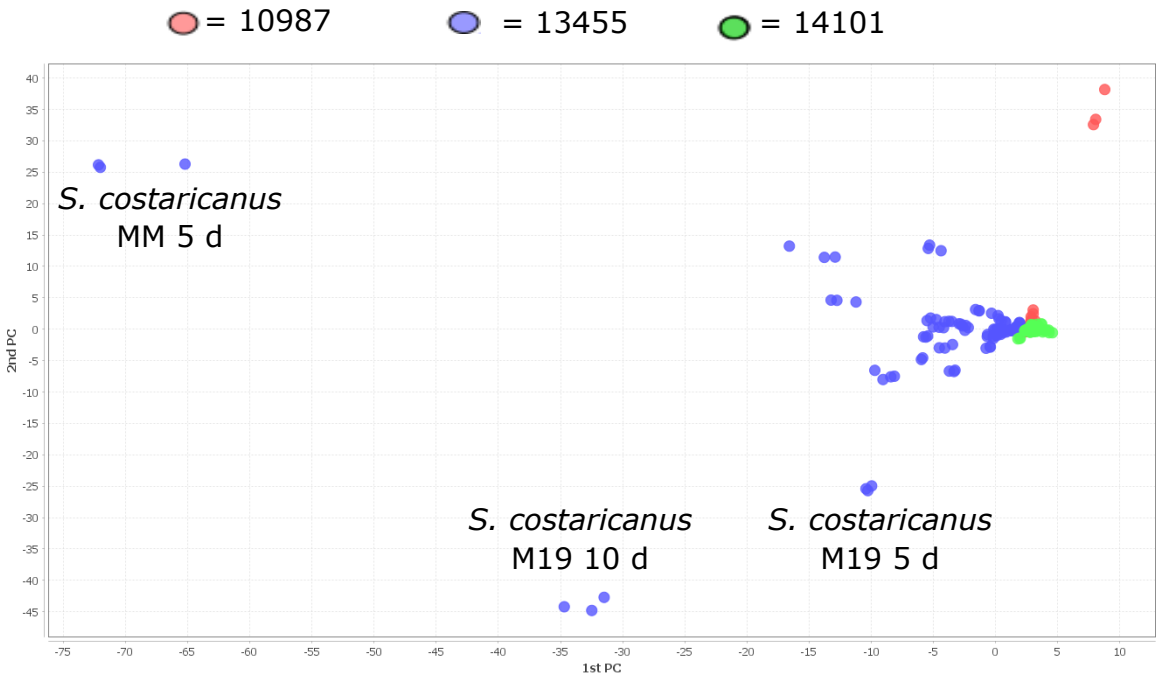
3.3.1.5 Selection of non-soil strains

The most geographically extreme isolate is from the Mariana Trench, *Streptomyces* sp. NCIMB 14101. Based on its totally different niche to *S. costaricanus* it was selected for 24-well plate screening. Of the remaining non-soil strains, though both listed as biosafety level 1, *S. gedanensis* and *S. willmorei* are both potential pathogens. The ATCC equivalent for *S. gedanensis*, ATCC 4880, is listed as a clinical isolate, and *S. willmorei* was isolated from a patient with streptotrichosis. While primary *Streptomyces* infections both in healthy and immunocompromised human hosts are rare (Kapadia *et al.*, 2007), to minimise risk neither were selected for further analysis. Of the remaining strains, *Streptomyces* sp. NCMIB 10987 had the most different isolation niche (growth on PVC wallpaper) so was also selected as the final strain.

3.3.2 Metabolomic analysis of *Streptomyces* strains

The three strains selected in 3.3.1 were each cultured using the 24-well plate system outlined in Chapter 2.2.6 then metabolites extracted for UPLC-MS analysis. After processing with MZmine, an initial PCA plot was used to reduce data complexity and show which strain produced the most metabolomic variety both within and between strains. Both *Streptomyces* sp. NCIMB 10987 and *Streptomyces* sp. NCIMB 14101 clustered close to each other, with the exception of *Streptomyces* sp. NCIMB 10987 MYM 10 d culture (Figure 3.7), and *S. costaricanus* samples mostly separate.

Figure 3.7 PCA of 3 NCIMB strains in MZmine. The three red circles in the upper right corner are *Streptomyces* sp. NCIMB 10987 MYM 10 d triplicate samples.



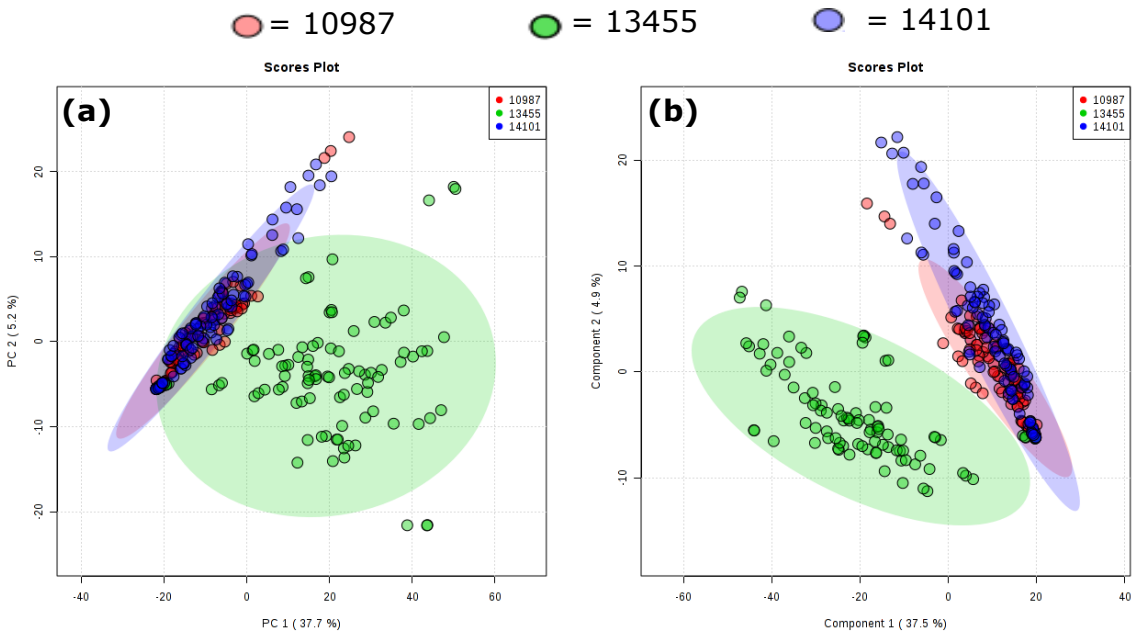
The MZmine peaklists were exported to MetaboAnalyst for more in-depth multivariate analysis. After processing the peak lists to produce a more accurate analysis (filtering, scaling, and transformation steps listed in methods), MetaboAnalyst PCA (Figure 3.8) also showed a clear difference between *S. costaricanus* and the other 2 strains, though the total variation explained using 2 principal components was under 50% (37.7 + 5.2%).

The majority of samples for each strain were grouped within the 95% confidence interval circles, indicating that analysed metabolites produced by each strain should significantly vary. The strain selection was therefore successful in that there was demonstrably significant metabolite variety between strains. There was some overlap between the intervals, potentially due to shared common *Streptomyces* compounds like deferoxamines. To reinforce the overall separation, and check if the shared features were significant between *S. costaricanus* and the other 2 strains, the data was re-run through a Partial Least Squares-Discriminant Analysis (PLS-DA).

PLS-DA is a supervised multivariate technique, so uses class identification information provided by the user in its analysis. This is in contrast to PCA which does not use any class information to group data. PCA is therefore considered a good initial exploratory tool, while PLS-DA is more suitable for separating data (Worley & Powers, 2012) so can act in tandem to confirm sample and class separation. Figure 3.8b shows a similar PLS-DA pattern to the PCA, in that *S. costaricanus* is separated from the other two strains except with no confidence interval overlap.

As a supervised method, PLS-DA can potentially group unrelated data if classes are randomly assigned, producing meaningless results (Szymańska *et al.*, 2012). Separation validity can be confirmed using a permutation analysis, which randomly assigns classes and checks separation similarity against the original output. Running the permutation analysis confirmed that the PLS-DA model was significantly different to one with random class assignments ($p < 0.001$), validating the model.

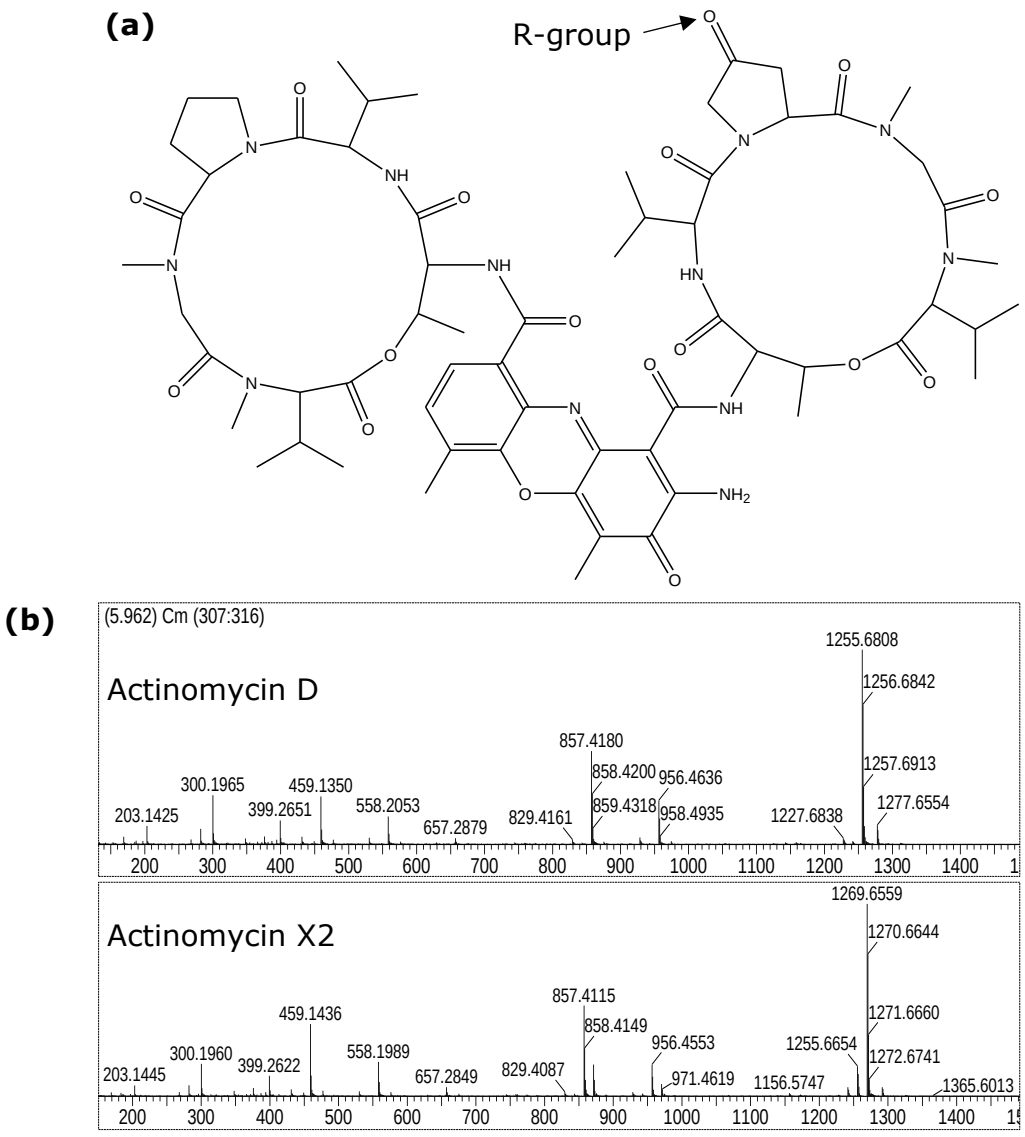
Figure 3.8 **(a)** PCA, and **(b)** PLS-DA of *S. costaricanus* (NCIMB 13455), *Streptomyces* sp. NCIMB 10987, and *Streptomyces* sp. NCIMB 14101.



S. costaricanus appeared to have the least overlap with the other 2 strains going by the PCA, indicating a greater metabolomic diversity was produced by *S. costaricanus* on these media than by *S.* sp. 10987 or *S.* sp. 14101.

Examining the PCA loadings plot, the three most important features for separation had m/z of 1269.67, 1293.67, and 1271.68 with rt of 6.80, 6.12, and 6.22 min respectively. Spectral analysis of each of the features gave a putative identity of actinomycins, a family of compounds produced by non-ribosomal polypeptide synthetases with antibiotic and anticancer activities depending on the family member (Katz *et al.*, 1956).

Figure 3.9 **(a)** Structure of actinomycin X2. **(b)** Comparison of actinomycin D standard to detected actinomycin X2.



The most abundant feature which also contributed most to separation – *m/z* 1269.67 – could be definitively identified as actinomycin X2 $[M+H]^+$ (shown with comparison to the most well studied actinomycin, actinomycin D in Figure 3.9), and *m/z* 1271.63 as actinomycin X0β $[M+H]^+$.

As other unidentified features also significantly contributed to separation, it was therefore hypothesised that focusing on *S. costaricanus* would give a higher chance of producing and identifying varied, novel metabolites rather than the other 2 *Streptomyces* species. As a further representation of metabolomic diversity, the total detected features across all 3 species were displayed using a heat map (Figure 3.10, names of specific media and feature *m/z* are not shown due to the large numbers of both), though it should be noted that as this shows features both metabolites and their MS-derived fragments are included. The heat map showed *S. costaricanus* metabolite intensities fluctuating more compared to the other 2 species, while also identifying which media produced the most variation and overall production within a strain. Therefore, data for *S. costaricanus* was re-run to produce the heat map of 569 features shown in Figure 3.11.

This heat map more clearly shows the compartmentalisation of features derived from *S. costaricanus* culture on different media. It also allowed basic overall comparison of the effect of sample time. As an example, CS agar and broth culture at 10 d share a run of higher intensity detected features which is not seen at 5 d. These features are not common to 10 d culture with every media, indicating that both the correct media and later culture time are required to produce the corresponding metabolites. Culture with ISP5, OM, and SGT produced the smallest blocks of uniquely higher detected features. This could have been due to a mix of poor growth, as in the case of ISP5 broth, unsuitable culture time with that media, or simply a non-optimal set of nutritional parameters for widespread novel metabolite production. In comparison, M19, MM, and MYM culture features all had larger runs of increased detection, as well as clear progressions from 5 d to 10 d over both agar and broth media, with M19 showing the longest run of features.

Figure 3.10 Heat map of 3 *Streptomyces* strain metabolomic features produced by growth on 8 agar and 8 broth media with an MS detection limit of 1E3 signal intensity. Left to right: *S.* sp. NCIMB 10987, *S. costaricanus*, *S.* sp. NCIMB 14101. Each row is a separate feature which retains its identity across the map for all the 3 strains. Blue indicates relatively lower feature detection, red higher feature detection. Blocks of higher intensity detection within a column does not indicate feature similarity, only that MZmine arranged higher detected features together from that sample, which are from multiple metabolites.

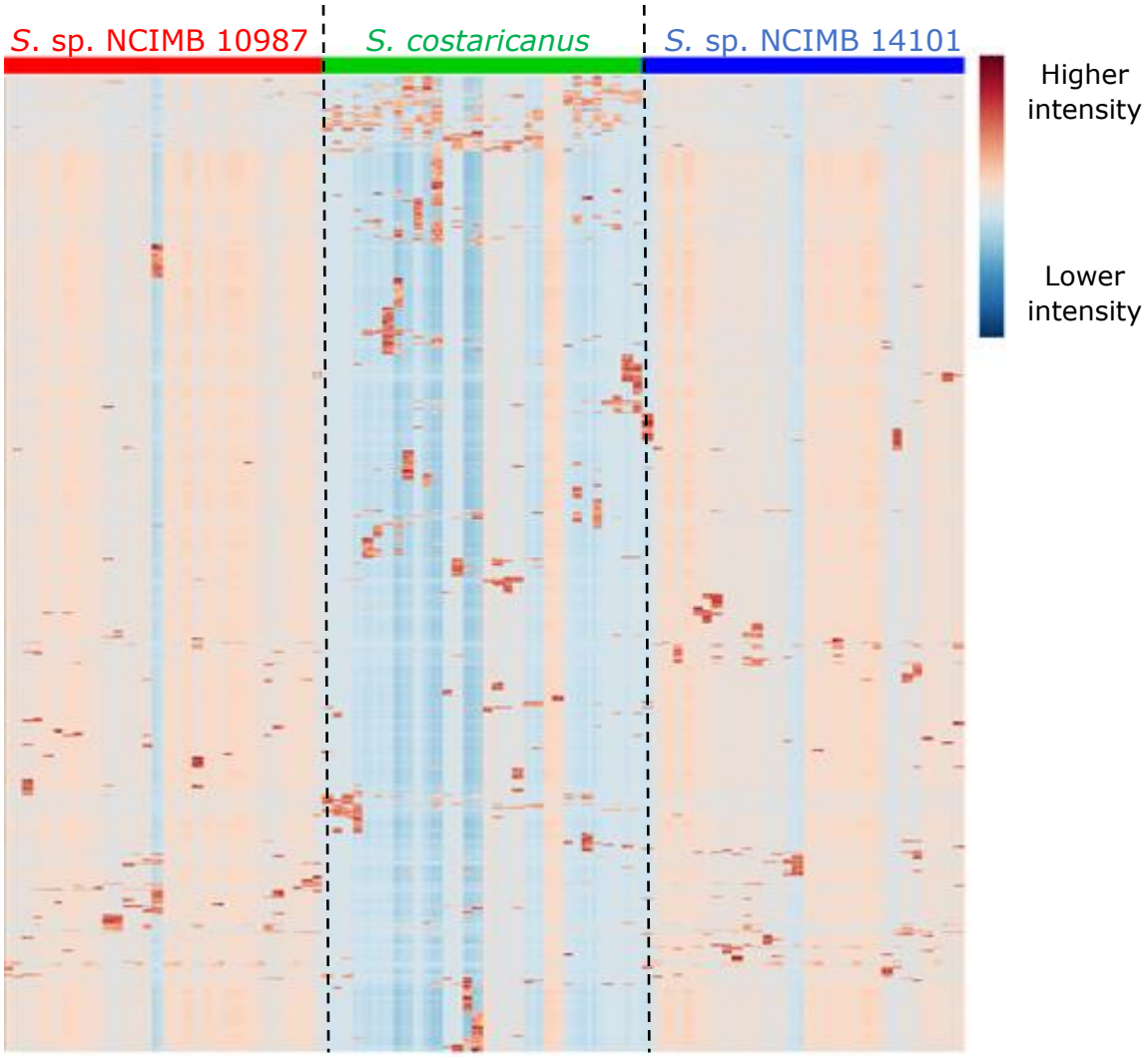
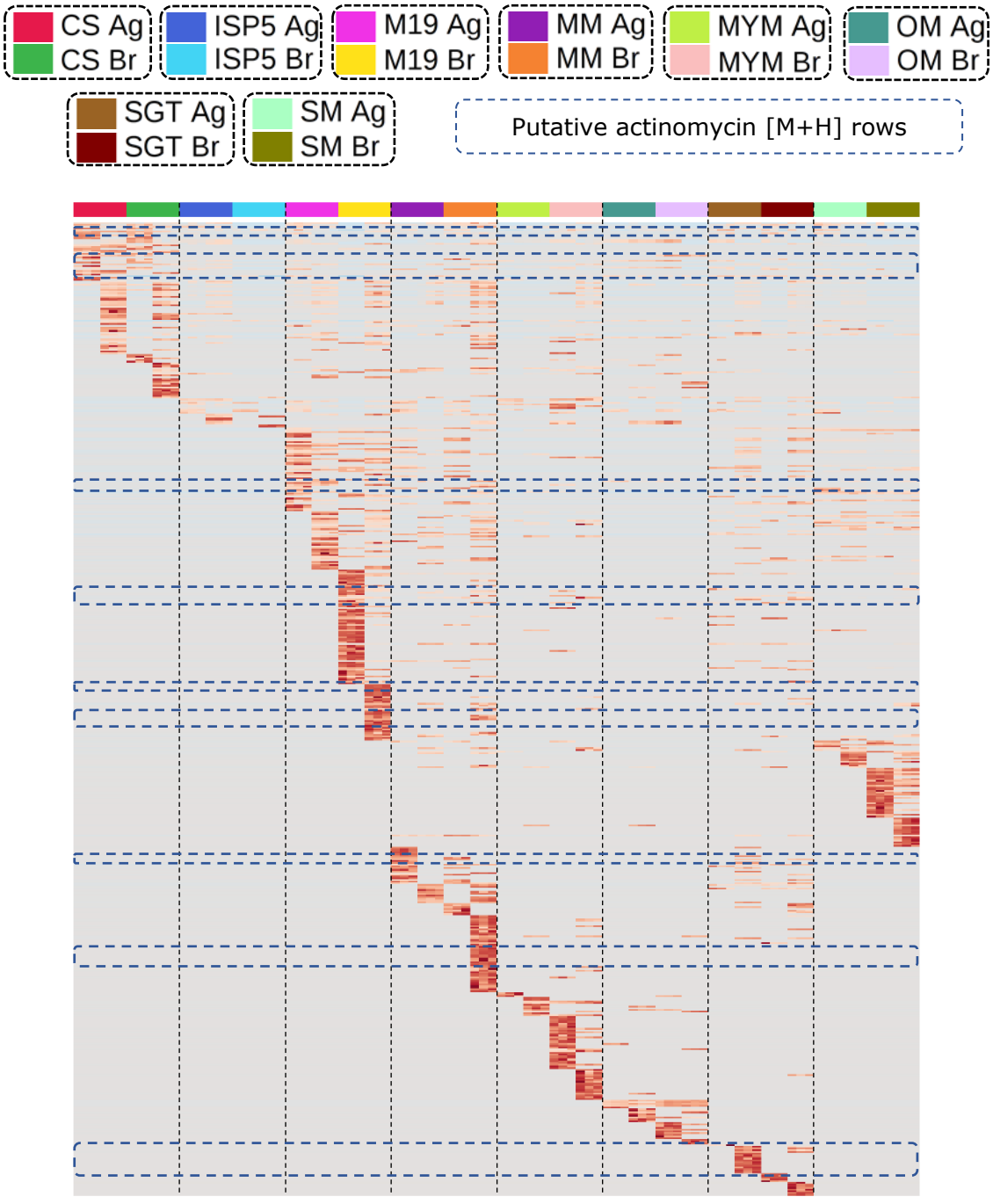


Figure 3.11 Heat map of features derived from *S. costaricanus* culture on agar and broth media. There are 2 columns present in each colour segment, which represent the 2 sample time points at 5 d then 10 d. Blue boxes indicate actinomycin feature containing rows, identified as the main cause of separation from the other 2 *Streptomyces* strains. Feature intentisty is indicated with grey (low intentsity) to red (high intensity) cell colour.



While the overall method did appear to have given useful information for strain selection, there are potential considerations to widespread application. One is if 2 of the 3 selected strains were closely related, for example two *Streptomyces* sp. strains that were actually *S. coelicolor* and *S. lividans* or even the same species. This would effectively skew the metabolomic results towards one strain by unintentionally decreasing the available comparative variety. While it cannot be confirmed without species classification, this is unlikely to apply here, as the non-soil strains were intentionally selected from different niches with vastly different survival pressures. Full genome sequences for the strains would also prevent this, but they were not available at the time of selection. It would also have enabled genome mining, giving an alternative – or complementary – strain selection process. Ideally this would be included earlier in any further strain comparison projects, especially given the decrease in long-read sequencing costs.

Secondly, as the overall aim was to create a generalised screen to expand specialised metabolism, the growth conditions may not be optimised for each individual strain. This issue is arguably shared by most OSMAC projects as without any prior knowledge about a strain's niche preferences or genetic regulation it is impossible to directly optimise an OSMAC screen. This can be partially resolved by using multiple growth conditions, but this needs to be weighed against the practical considerations of how much more data processing and analysis there will be for each strain. Similarly, though using a strain from an extreme environment may be a promising new metabolite source, its extreme environment may lead to different metabolite activation conditions than a typical soil *Streptomyces*. As the project was based on NCIMB *Streptomyces*, the majority of which are soil strains, this problem is another balance between a broad condition screen and tailoring to an individual isolate.

3.3.2.1 Impact of media on actinomycin production

As stated in section 3.3.2, the metabolites most contributing to separation in PCA were the actinomycins, specifically actinomycins X0β and X2. Actinomycin X2 differs to the most well studied family member, actinomycin D, by the addition of an oxygen R-group. The other detected actinomycin, X0β, contains an OH R-group at this location. Actinomycin D was produced but contributed less to PCA

separation. The main metabolite family data broadly agrees with a recently published paper on a marine strain of *S. costaricanus* where actinomycin D was found to be the main product; here on most media it was actinomycin X2 (Liu *et al.*, 2019).

General characteristics are given for the 3 detected actinomycins in Table 3.4. Antibacterial activity is strongest against Gram-positives as actinomycin is less able to penetrate the Gram-negative outer membrane to bind DNA and prevent transcription. Activity is still seen against some Gram-negatives for different family members, but the literature is not consistent about activity levels. Wang *et al.* (2017) detected the same actinomycins from a marine isolate of *S. heliomycini*, tested against *E. coli* and found it to be ineffective, whereas Zhang *et al.* (2016) report an MIC of 0.12 μM for actinomcyins D and X2.

Production of actinomycins by *S. costarcianus* grown on different media was compared as they were the most important PCA separation feature. Actinomycin X2, the predominant actinomycin, was produced by *S. costaricanus* at a maximum concentration of 15.14 μg/ml from M19 broth cultures after 10 d. The combined concentrations for all actinomycins over both broth and agar culture are shown in Figure 3.12, which displays significant variation in X2 production across *S. costaricanus* culture. Trends for actinomycin levels also varied from 5 to 10 d, with the majority of media leading to increasing production levels over time. Culture on other media like M19 or CS broth did not produce a significant increase over 5 d to 10 d.

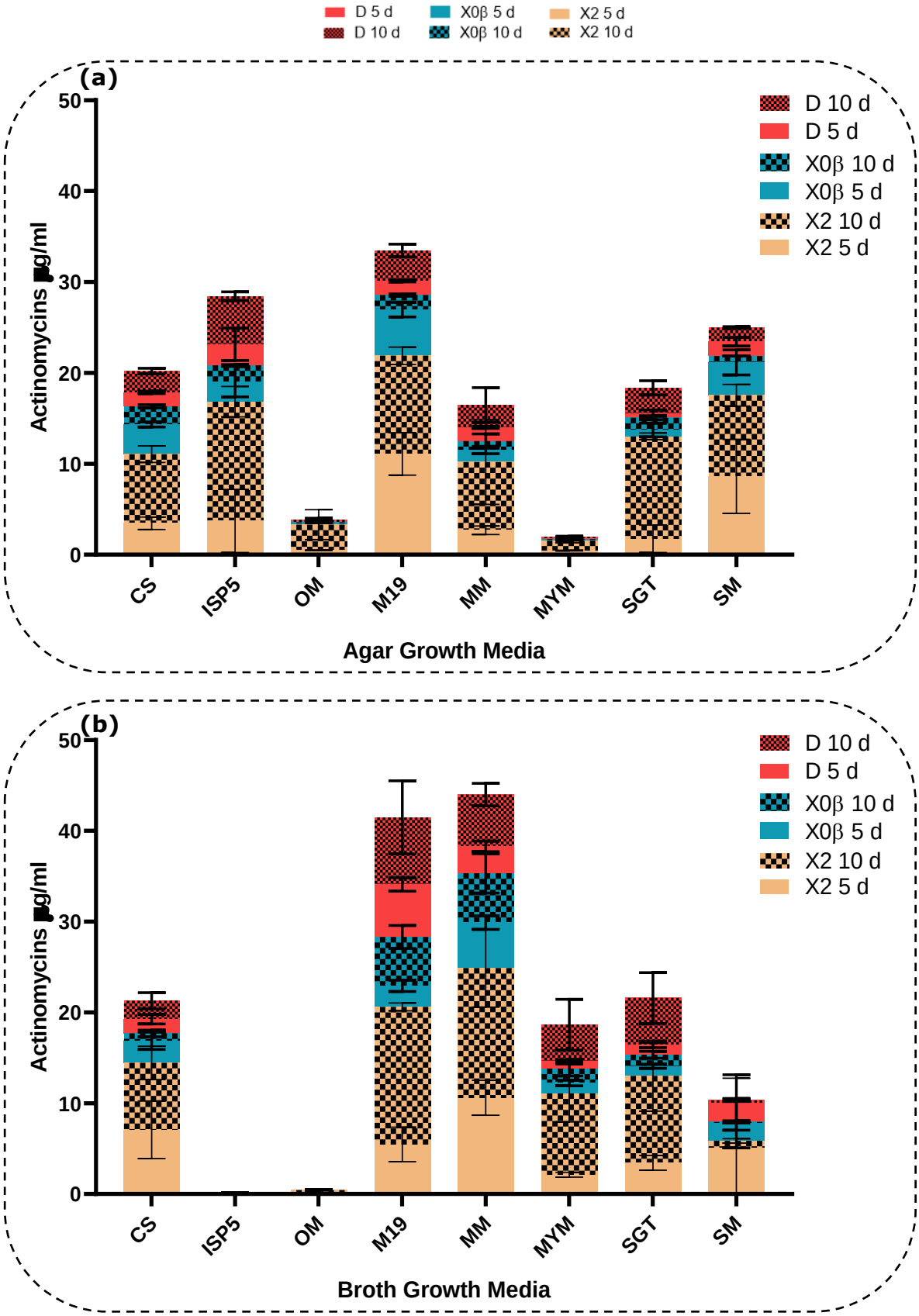
Table 3.4 Detected actinomycins and their characteristics

Actinomycin member	***m/z* of [M+H]$^+$**	**Formula of [M+H]$^+$**	**Bioactivity against**	**Source**
D	1255.64	$C_{62}H_{86}N_{12}O_{16}^+$	Tumours	Xiong *et al.* (2012); Liu *et al.* (2016); Zhang *et al.* (2016); Wang *et al.* (2017)
X2	1269.63	$C_{62}H_{84}N_{12}O_{17}^+$	Gram-positives	
X0β	1271.63	$C_{62}H_{86}N_{12}O_{17}^+$	Gram-negatives (weaker, less applicable)	

ISP5 agar culture gave significantly less compound at 5 compared to 10 d (3.71 µg/ml up to 13.12 µg/ml respectively), with an approximately 4-fold difference. Similarly, MYM broth culture produced 2.11 and 8.98 µg/ml over the time points. In comparison, CS broth culture did not significantly change. The medium leading to the most consistent production of actinomycin was M19 agar, which gave a slight decrease from 11.09 down to 10.08 µg/ml over the sample points. The M19 broth equivalent does not follow the same pattern, showing a 3-fold increase from 5.46 up to 15.14 µg/ml.

The greatest difference was seen between ISP5 agar and broth media, which gave one of the largest yields from agar but none from broth. This comparison is not wholly useful since growth was severely restricted, potentially absent, in ISP5 broth. Why *S. costaricanus* did not grow is unknown since ISP5 does not contain any abnormal components which would prevent growth without agar.

Figure 3.12 **(a)** Concentrations of 3 detected actinomycins derived from *S. costaricanus* agar culture, and **(b)** actinomycins derived from *S. costaricanus* broth culture. Key for actinomycin types and culture times:



3.3.3 Impact of minimal media culture supplemented with elicitors on metabolite production

The 8 media used in the nutrition screen contained a diverse range of carbon and nitrogen sources, and minerals. While growth on complex media can lead to a different metabolome in each condition, it also made it more difficult to link specific media components to specialised metabolites. Therefore, to further explore the metabolite diversity of *S. costaricanus*, it was grown in minimal mannitol media (MMM) with a single parameter altered over 9 other wells: 25 µM scandium chloride, 25 mM sodium butyrate (NaBu), 10 mM N-Acetyl Glucosamine (GlcNAc), pH 4, pH 10, *Acinetobacter baumannii* cells, *A. baumannii* supernatant, constant 42°C, and 42°C heat shock for 1 h after 5 d culture.

Actinomycins were not detected above the 1E3 MS^1 cutoff from minimal media culture. Analysis of MS/MS files of 24-well plate minimal media culture processed with MZmine with lower signal intensity inclusion (1E1) did detect actinomycins D and X2, along with deferoxamine B complexes [M+Al-2H; M+Fe-2H].

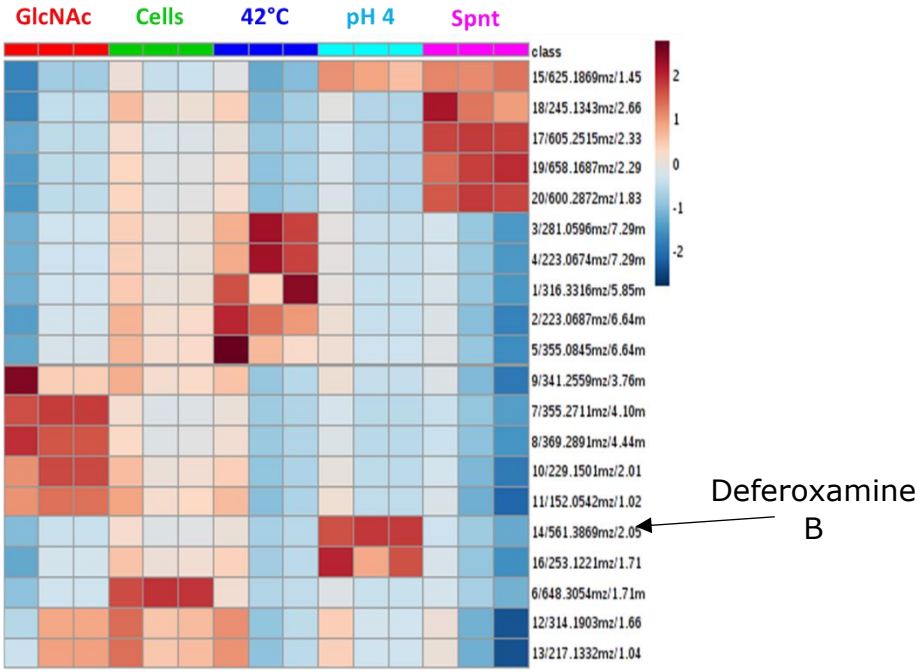
3.3.3.1 Elicitors upregulate production of *S. costaricanus* metabolites in minimal media culture

Comparing the other culture conditions to standard minimal media growth and removing those with no new metabolites eliminated scandium chloride and 42°C heat shock. This left 5 conditions, shown in Figure 3.13, producing 20 features not seen from minimal media culture. Features were predominantly specific to their elicitors, with only 1 being shared over multiple elicitors: *m/z* 625.19 rt 1.45 min, shared between pH 4 and 100 µl *A. baumannii* culture supernatant. Though elicited by culture supernatant this feature was not produced in response to heat-killed *A. baumanii* cells.

One feature was upregulated after addition of dead *A. baumannii* (*m/z* 648.31 rt 1.71) to minimal media. A feature with the same *m/z* and rt is present in the

original minimal media culture, but at an intensity approximately two orders of magnitude less than with *A. baumannii* added (2 – 3E1 against 2 – 4 E3).

Figure 3.13 Heat map of features found in MMM+elicitor *S. costaricanus* cultures but not in MMM-only culture. Dark blue through to white then dark red indicates relatively lower to higher LC-MS feature intensity as a proxy for metabolite production. Deferoxamine B, upregulated by culture at pH 4, was detected from rich media culture but not from minimal media.



Co-culture mediated induction of specialised metabolism can occur through a variety of methods, including secreted metabolites (D Wang *et al.*, 2013) or cell-cell contact (Onaka *et al.*, 2011). This study showed living cells along with physical contact were necessary for SMet upregulation; another study on *S. coelicolor* A3(2) used dead *Bacillus subtilis* and *Staphylococcus aureus* cells to increase production of undecylprodigiosin (Luti & Mavituna, 2011). Murphy *et al.* (2007) increased production of bacitracin A from *B. lichenformis* by adding oligosaccharides to the growth medium, indicating that the biochemical itself was sufficient for upregulation. *S. costaricanus* could be similarly reacting to *A. baumannii* thermostable cell wall components.

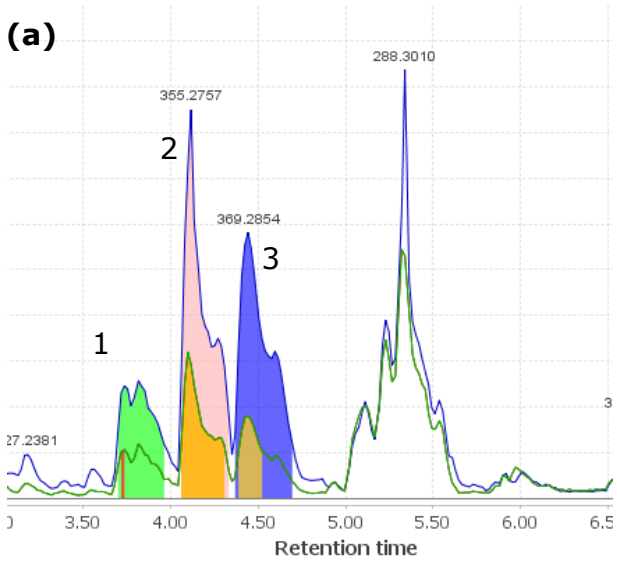
Similarly, higher temperature treatment elicited metabolites (*m/z* 316.33 rt 5.85; *m/z* 281.06 rt 7.29) that were produced under minimal media culture of *S. costaricanus* but below detection levels: *m/z* 316.33 goes from a mean intensity by height of 7.3E1 to 3.1E3, and *m/z* 281 1.4E2 to 3.3E3. Of the 2 pH4-specific metabolites, 1 (*m/z* 561.39, deferoxamine B) was already seen in the nutrition screen.

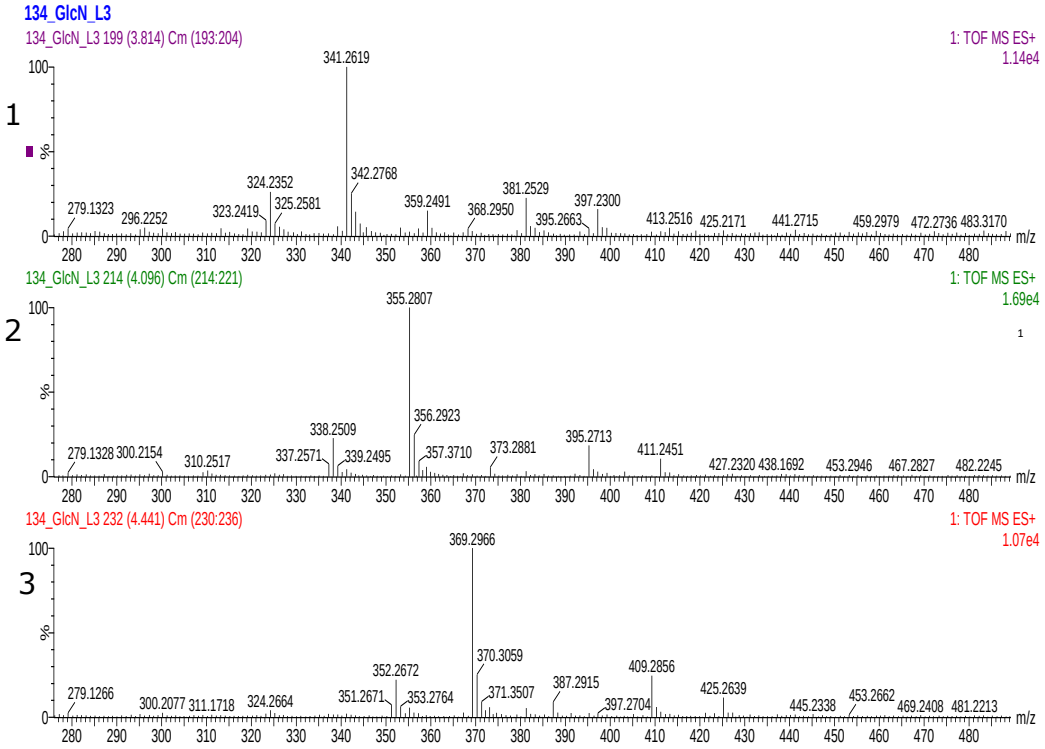
3.3.3.2 Putative metabolite family upregulated by addition of GlcNAc

GlcNAc has been shown to pleiotropically activate specialised metabolism in a range of *Streptomyces*, though unsurprisingly given the variety of the genus and complexity of genetic regulation, can also repress or have no effect on SMet (Rigali *et al.*, 2008). Of the 6 potentially novel metabolites produced by *S. costaricanus* under this condition, three (*m/z* 341.26, 355.27, 369.30, or 0, +14, +28) elute close together over 3.6 to 4.7 min. A similar fragmentation pattern is shown maintaining the 0/14/28 *m/z* difference, potentially indicating a molecular family separated by CH_2 and C_2H_4 fragments (Figure 3.14). Manually searching the GNPS database for these masses did not find any matches.

Figure 3.14 **(a)** LC/MS chromatograms of the upregulated compounds in response to 10 mM GlcNAc. **(b)** Mass spectra of compounds 1, 2 and 3, showing the consistent 0/14/28 mass fragment difference.

(b)



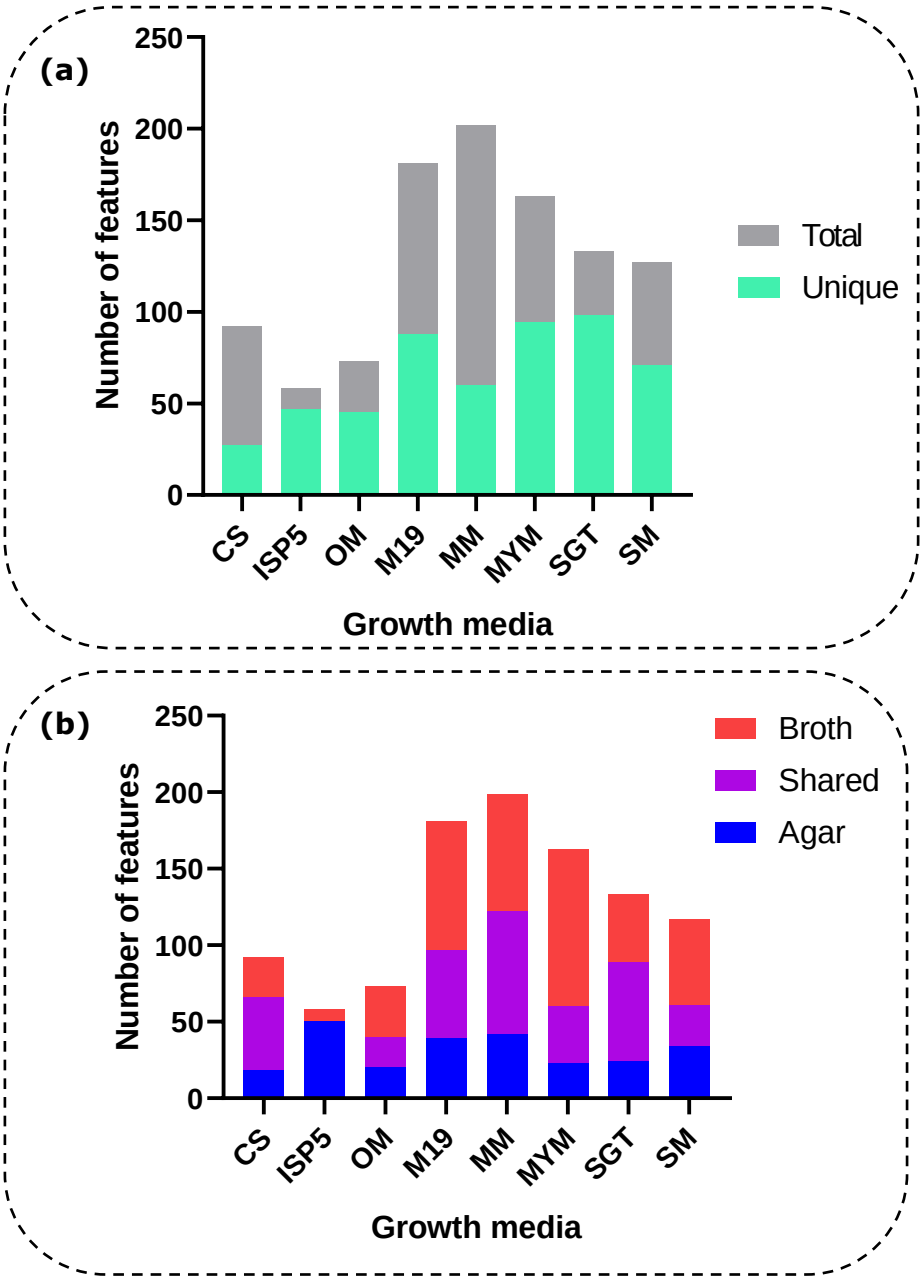


3.3.4 Selection of media for scale-up culture

Originally selection of the target strain for further investigation was based on PCA to show overall diversity in the *S. costaricanus* metabolome. However, while the PCA did separate the sample replicas into groups (Appendix figure 10), the variance explained by PC1+PC2+PC3 was not high (21.5%+9.1%+8.0%) and there was significant overlap between conditions. This made PCA unsuitable for deciding which condition should be used. Instead, total and unique feature numbers were used to show the most productive conditions. The overall effect of varying culture conditions on *S. costaricanus* is displayed in Figure 3.15, and was used along with carbon source repression likelihood and growth levels to make the scale-up selection.

Although MM culture gave the highest number of total detected features (202), there were a relatively low number of unique features (60), meaning those which were not detected from culture in any other growth media. As MM contains 3 potential carbon sources, 20 g/l molasses, 5 g/l glucose, 1 g/l yeast extract, the total number likely reflects high growth with no nutritional stress for *S. costaricanus*. A full comparison of features total and unique features for media is shown in Figure 3.15. This figure also shows the number of features found as a result of broth or agar culture from a specific medium, as well as any features shared between the two. For most conditions, broth led a greater number of features with the only exception being ISP5, in which *S. costaricanus* was not seen to grow.

Figure 3.15 **(a)** Total and unique features detected from *S. costaricanus* OSMAC culture, and **(b)** Features from broth and agar culture, or shared between them



The media leading to the highest number of unique features relative to its overall production was ISP5 (81.0%), followed by culture on SGT (73.7%), OM (61.6%), MYM (57.7%), SM (55.9%), M19 (48.6%), MM (29.7%), and CS (29.3%). This should not be taken as rankings of the "best" media for unique metabolite production by *S. costaricanus* – ISP5, though it is at the top of the list, gave no

growth in ISP5 broth and fewer unique features than all but CS and OM. CS and SGT included 1 g/l of glucose in case the potential main carbon source (Corn Steep Liquor and Soybean Oil respectively) could not be effectively utilised by *S. costaricanus*. Culture of *S. costaricanus* on both CS and SGT conditions gave good confluent growth, but SGT clearly outperforms CS with 133 to 92 total features and 71 to 27 unique features.

S. costaricanus produced consistently good growth on media without glucose (OM, M19, ISP5, MYM, M19, and MYM) with relatively high total/unique feature numbers. The least globally repressive carbon source used in the screen is potentially mannitol (Romero-Rodríguez *et al.*, 2016), which is the sole source of M19. Culture on M19 also gave the 2nd most overall (181) and 3rd most unique (88) features, and interestingly matched the bald phenotype shown by *S. coelicolor* on M19 in Chapter 2. Given its good growth and high-ranking feature production, M19 was selected as the scale-up medium. M19 broth gave more detectable features than Agar (148 against 70) so was selected for scale-up culture. Both MM and MYM were good candidates for scale-up, but the MM extracts contained fewer unique features than M19; MYM extracts had a similar number of unique features, but fewer overall. Maltose can also be a restrictive carbon source, especially as a glucose disaccharide (Sánchez *et al.*, 2010). Both of these points reinforce M19 as a suitable medium for further studies.

3.4 Conclusions

The NCIMB *Streptomyces* collection was surveyed for unexploited strains diverse range of isolation niches, using zone of inhibition-based bioactivity assays to select from the soil strains and niche variety for 2 other non-soil strains. This led to the selection of the strains NCIMB 13455 (*Streptomyces costaricanus*, Costa Rican soil), NCIMB 14101 (*Streptomyces* sp., Atacama Trench sediment), and NCIMB 10987 (*Streptomyces* sp., growth on PVC wallpaper). The 3 strains were screened using the 24-well plate culture screen developed in Chapter 2, leading to the identification of *S. costaricanus* as the most distinct from the other 2 strains.

The nutrition screen data for *S. costaricanus* showed the main produced metabolite family to be the actinomycins, followed by deferoxamines. Production levels of both of these families varied across media type and sample time, with M19 and ISP5 culture producing the most actinomycin on agar media, and M19 and MM culture producing the most in broth. Comparing the total features produced across all of the media showed M19, MYM, and MM culture gave the most features. Of those, M19 was selected for further culture scale-up as it enabled production of a relatively large number of unique features and contained a less repressive carbon source than other media.

In addition to rich media culture, *S. costaricanus* was also cultured in minimal media with multiple additives or stresses that could potentially have upregulated specialised metabolite production. While some of the produced metabolites were shared with the rich media screen, like deferoxamine B, addition of GlcNAc and heat-killed *A. baumannii* cells both elicited production of potentially novel metabolites. Unfortunately, production levels were too low to reasonably scale up and investigate the novel features. It was therefore decided that focusing on M19 culture would be more productive and so was chosen as the main culture scale-up parameter.

CHAPTER 4

CHARACTERISATION OF THE *S. COSTARICANUS* GENOME AND ITS BIOSYNTHETIC GENE CLUSTERS

4 Characterisation of the *S. costaricanus* genome and its biosynthetic gene clusters

4.1 Introduction

4.1.1 Genomic variation of specialised metabolism in *Streptomyces*

The linear genomes of *Streptomyces* contains a core region of approximately 2000 genes over 5 to 7.5 Mbp (Kim *et al.*, 2015; Wang *et al.*, 2015), most of which are essential for cell survival. Gene functions include vital processes like DNA replication, cell division, translation, and primary metabolism (Hopwood, 2006). This region displays considerable synteny not just in the genus but across the phylum, with analogues in both *S. coelicolor* and *M. tuberculosis* (Choulet *et al.*, 2006). The shared core region varies depending on which species are examined and how strict the criteria are for the core region: *S. coelicolor* and *S. avermitilis* are given as sharing around 4000 genes, gaining and losing them at approximately 10 genes every 1 million years since their last common ancestor (Chater & Chandra, 2006). In comparison, the flanking segments – or "arms" – are skewed towards genes which allow adaptation and survival in a changing environment, such as chitinases, other enzymes and biosynthetic gene clusters (BGCs) responsible for specialised metabolism.

The arms are also more genetically varied and malleable between strains, with a high rate of horizontal gene transfer aided by transposons, conjugative plasmid-associated genes, and a high rate of insertions and deletions (Hopwood, 2006; Choulet *et al.*, 2006). Localising this variation to the arms avoids damaging the essential core region, thereby increasing the potential of natural selection to act on specialised metabolism and other genes related to environmental adaption. Rearrangement on the arms can be rapid: Zhang *et al.* (2019) re-streaked 15 random *S. coelicolor* colonies, and found morphology varied from the original colonies, which increased again after a 2nd round of re-streaking. Sequencing 8 random colonies showed significant chromosome terminus rearrangement, with insertion sequences producing multi-300 kb amplifications mixed with deletions up to 878 kb long. Mutant variety correlated with altered specialised metabolite

profiles as well as increased inhibition of *B. subtilis*; colonies were essentially found to be split into mutants overproducing antibiotics, and non-producers dedicated to growth and expansion.

Exact examples of a specific gene, or BGC, as adaptions to a specific condition are rare due to the difficulty of determining previous gene acquisition events and if they correlate with a proposed environmental selective pressure. Larger scale genomic analysis can potentially be more informative. Tian *et al.* (2016) investigated 9 marine *Streptomyces* from across the South China Sea and compared them to publicly available sequences. Despite being from a broad range of locations across the sea, the marine strains shared increased levels of K^+ transporters, suggested to help with osmolarity regulation, as well as increased numbers of the transporter NhaA which exports Na^+.

A secreted metabolite which aids survival and so replication can be compared to an evolutionary biology concept usually associated with animal evolution, the "extended phenotype" (EP; Dawkins, 1982; Hunter, 2018), where a gene's impact on the animal and its corresponding evolutionary fitness is expressed separately from the animal. A beaver's dam or termite's nest are simple examples of the EP; the behaviour of a parasite-infected animal is a more complicated example of the parasite's EP.

Similarly, a secreted bioactive metabolite, whether a messenger molecule to its own strain or an antimicrobial weapon, or the breakdown of *Streptomyces* hyphae induced by programmed cell death to fuel its neighbour's differentiation can be viewed as benefiting the microbial host's own EP. Since the specialised metabolite regions of *Streptomyces* genomes change so frequently, this corresponds to accelerated selection of the microbe's genomic content and how the EP aids reproduction. Kamoun, (2007) uses EP in a similar sense to describe effector proteins secreted by microbes inside plants, as although they are produced by the microbe the protein itself functions on or inside a plant cell, thereby modifying the environment of the producing species.

The high genomic variation means that even closely related strains can differ in their specialised metabolome. Strain-level recombination has been found to exceed species-level recombination rates in the *S. flavogriseus* clade by over two orders of magnitude (Doroghazi & Buckley, 2010); this partially explains the

phenotypic and chemotypic diversity seen from a single species. Seipke (2015) compared the BGCs of *S. albus* strains, which contain between 25 and 30 clusters, and found a core set of 18 BGCs. Aside from those 18, each strain had at least 1 of 17 unique BGCs, with a range of 1 to 6 unique BGCs per genome. Of those 17, 3 had a precise predicted product: paulomycin, NRPS-oligosaccharide; kendomycin, type I-III PKS; fredericamycin, type II PKS. The remainder were a mix of NRPS, PKS, butyrolactone, bacteriocin, and lantipeptide BGCs.

This does not mean that specialised metabolism is completely variable across the genus: Kim *et al.* (2015) located 27 such genes conserved across 17 *Streptomyces*, including well known *Streptomcyes* metabolites geosmin, deferoxamine, and hopene. However, these genes are clearly in the minority compared to the overall fluidity of specialised metabolism. The falling price of Whole Genome Sequencing (WGS; Didelot *et al.*, 2012) has made routine genotyping more accessible, leading to a rapid increase in the amount of useful genomic data.

4.1.2 Sequencing *Streptomyces* genomes

As a complex multicellular genus, *Streptomyces*, especially non-model strains, can be challenging to work with. This is carried over to their genome, which has presented multiple challenges for bioinformaticians attempting to produce a robust complete genome. Two characteristics are particularly problematic: high GC content (Gomez-Escribano *et al.*, 2016), and similar repetitive sequences in BGCs (Xu & Wright, 2019).

High GC content is a key trait of *Streptomyces* and other actinomycetes with a typical minimum of 70% (Liu *et al.*, 2013). GC rich regions have a higher melting point than an evenly split sequence, and may form single-strand hairpin loops after the initial melting step which are not further amplified in PCR (Frey *et al.*, 2008). As a result, GC rich sequences can be under-amplified and so produce uneven to no sequencing coverage; as some assembly programmes treat uneven coverage as an error, this can lead to genuine sequences being ignored (Chen *et al.*, 2013).

Even if the uneven coverage regions are not ignored, other problems can arise in assembling repetitive regions as found by Scott & Ely (2015). These problems predominantly extend across short read sequencing technologies as the longer read methods do produce better coverage (Quail *et al.*, 2012). Both non-ribosomal peptide synthases (NRPS) and polyketide synthases (PKS) contain highly repetitive sequences which cannot be easily distinguished from each other (Gomez-Escribano *et al.*, 2016), especially if the assembly errors occur. This will lead to either fragmented BGCs spread across the genome, or two separate sequences being incorrectly assembled together.

However, as with GC richness, the longer read technologies are more suited to BGC sequencing as the longer read window prevents assembly errors over repetitive and GC rich sequences. This was reported by Goldstein *et al.* (2019), who in their comparison of MiSeq and MinION sequencing, found their Illumina *Pseudonocardia,* a GC rich *Actinomycetales* genus, generated a fragmented sequence of thousands of contigs. Analysing the full genome sequences in antiSMASH reinforced that the MiSeq sequence was biased away from GC and comprised fragmented BGCs. The addition of MinION data to the MiSeq reads using both SPAdes and Unicycler assembly programmes resulted in fewer and also larger contigs, though Unicycler was the better option for hybrid assembly.

The differences in *Streptomyces* genome quality between sequencers are well known; unfortunately, older database deposited sequences using Illumina, 454 or other shorter read technologies are likely to have both of the above problems. Studholme (2016) lists multiple examples with the then 653 *Streptomyces* assemblies in Genbank. Searching for 40 universally conserved single copy genes showed 239 of the 653 genomes to be missing genes, indicating their incompleteness.

It is further noted that the original sequence read data may not be published, making it impossible to check the assembly quality. These issues make it clear that the correct next generation technologies should be used for *Streptomyces* WGS, ideally using long reads as the structural basis with high coverage additional shorter reads to improve quality by minimalising frame shift and random base calling errors. Both long and short read technologies have their own advantages and disadvantages, so using a hybrid mixture of the 2 is a valid

option for a range for prokaryotic and eukaryotic species (Koren *et al.*, 2012; Zimin *et al.*, 2017; Wick *et al.*, 2017). In particular the longer, structurally correct reads which have a higher base calling error rate are complimented by the shorter reads, which call basses more accurately but would be difficult to correctly arrange without the longer read scaffolds.

A robust genome sequence is vital for predicting the full range of BGCs. Screening genomes to determine a microbe's BGC amount and so biosynthetic potential is a key part of genomics-based metabolite discovery. The huge increase in available, hopefully high quality, data on BGCs means that assessment of the BGC potential of a *Streptomyces* is achievable early on in a project and can provide useful information for metabolite dereplication and BGC targeting.

4.1.3 Genome mining for biosynthetic gene clusters

4.1.3.1 BGC rich bacterial clades

Genome mining is the process of searching a genome for a specific class of gene, which here are specialised metabolite BGCs. This can then be used to guide strain selection or as a base for genetic manipulation to elicit production of novel specialised metabolites (Lautru *et al.*, 2005; Corre & Challis, 2009). *Streptomyces* are not unique within bacteria as having high BGC numbers as other phyla contain high BGC clades (Baltz, 2017), some of which are briefly listed.

Within Firmicutes, a selection of *Paenibacillus* spp. and *Bacillus* spp. were found by Baltz (2016) to contain up to 17 BGCs and produce multiple antimicrobials such as surfactins and bacilysin from *B. subtilis* (Hamdache *et al.*, 2013). *Burkholderia* spp., a β-proteobacterium, are another source of natural products, with an average of 15 BGCs/species and a high of 27 in *B. pseudomallei*, (Liu & Cheng, 2014), a strain of which was exploited using a genome mining strategy by Biggins *et al.* (2011) to express a cryptic BGC. Finally, δ-proteobacteria contain the order *Myxococcales*, or myxobacteria, most notably *Sorangium cellulosum* (32 BGCs in *Sorangium cellulosum* so ce56, the most of any non-*Actinomycetales* member thus far) and the genus *Myxococcus*. Like *Streptomyces*, these are predominantly found in the soil, have a high GC%,

undergo a complex multicellular lifecycle ending in sporulation, and have a large genome (Garcia *et al.*, 2009).

While these are all valid candidates for genome mining and natural products research in general, *Streptomyces* remain the most BGC-dense genus, with a range of 3 – 43 BGCs and an average of 21.6 per genome from a set of 341 strains (Doroghazi *et al.,* 2014). This set did not include *S. rapamycinicus* NRRL 5491, which contains at least 48 BGCs (Baranasic *et al.*, 2013), so the actual average may be higher. There are reports of *Streptomyces* strains with over 50 BGCs (Jackson *et al.*, 2018), but given their rarity are worth confirming with multiple sequences to avoid fragmented BGCs artificially raising the total count.

Multiple methods and software exist to mine genomes for BGCs, some of which are still applicable to genomes. One example is the detection of *mbtH* genes, which code for small proteins, typically 60 – 70 amino acids long, which act as chaperones for NRPS products (Baltz, 2011). The original MbtH protein is found in *Mycobacterium tuberculosis*, where it helps solubilise the NRPS mycobactin (McMahon *et al.*, 2012). Homologues of *mbtH* are common across bacteria, but especially in *Actinobacteria*, for which NRPS are one of the most common BGC classes along with PKS and terpenes (Baltz, 2014). Searching for MbtH homologues will therefore indirectly assess NRPS numbers in a genome, with less homologous results potentially indicating a more novel NRPS.

The most widely used method for *Streptomyces* genome mining is the Antibiotics and Secondary Metabolite Analysis Shell, or antiSMASH (Medema *et al.*, 2011; Blin *et al.*, 2019). antiSMASH detects a wide range of BGCs, from the common NRPS/PKS/terpene clusters to rarer specific types like lassopeptides, cyanobactins, or bottromycins. Importantly it is also presented with an easy to use web interface making it accessible to anyone without needing to run the underlying command lines, vastly increasing adoption. It also supports multiple file formats (GenBank/FASTA/EMBL), will perform basic compound structure prediction where possible, and provides links to the antiSMASH database. The database provides a wide range of information on cluster type phylogeny, overall cluster type representation, and customisable cluster searches (Blin *et al.*, 2019).

antiSMASH detects BGCs within data by searching for core regions specific to cluster types then using the surrounding genomic context to build up the cluster,

identifying ancillary genes (transport, resistance, regulation, and others) potentially linked to the biosynthetic genes. These can be used to predict further information about the gene or its eventual metabolite product. This includes product location (intra or extracellular, based on transport genes e.g. multiple ABC transporters), regulatory elements like TTA codons in *Streptomyces* (Chater & Chandra, 2008) or cold/heat/other stress related sigma factors, and resistance genes which could identify a bioactive compound's mechanism of action. Resistance-based prediction is more comprehensively examined in the Antibiotic Resistance Target Seeker tool (Alanjary *et al.*, 2017).

4.1.4 Chapter aim and objectives

Integrating metabolomics and genomics data would result in a powerful tool to analyse metabolite production, regulation, and untapped potential of a microbe. The high variation of *Streptomyces* genomes would also make for an interesting comparison between the conserved likely core region genes and the more diverse specialised metabolite genes. Therefore, the aim of this chapter is to characterise the genome of *S. costaricanus* and its BGCs. This will be accomplished by:

1) Obtaining a high-quality full genome sequence for *S. costaricanus.*
2) Initial genome exploration and comparison to model *Streptomyces*
3) Genome mining to identify BGCs.
4) Determining if similar strains within the *S. costaricanus* clade share BGCs.

4.2 Materials and Methods

4.2.1 DNA extraction and whole genome sequencing of *S. costaricanus*

All DNA extraction and sequencing steps were done by the Identification department at NCIMB Ltd.

4.2.1.1 DNA extraction of *S. costaricanus* for Illumina library preparation and MiSeq sequencing

S. costaricanus was cultured for DNA extraction on yeast malt agar at 25°C for 7 d after which a loop of colony was used to inoculate a 20 ml vial yeast malt broth. The culture was incubated at 25°C, 150 rpm, for 5 d to generate biomass for DNA extraction.

DNA was extracted for Illumina MiSeq sequencing using a DNeasy UltraClean Microbial Kit (QIAGEN NV, Venlo NL) following the manufacturer's instructions. DNA amplification is required before the extracted DNA can be sequenced. MiSeq library preparation used an Illumina Nextera DNA Flex Library Prep Kit (Illumina, Cambridge UK). After generation, Protocol A: Standard Normalization was used to denature and dilute the library, after which samples were run in the MiSeq using a MiSeq v3 600 cycle cartridge. Sample documentation was created using Illumina Experimentation Manager.

4.2.1.2 DNA extraction of *S. costaricanus* for Oxford Nanopore library prep and MinION sequencing

The same *S. costaricanus* culture method was used for DNA extraction as was for Illumina sequencing. DNA was extracted using a Qiagen MagAttract HMW DNA kit, again following the manufacturer's instructions.

After extraction, the 1D Genomic DNA by Ligation protocol was followed using a SQK-LSK109 kit to repair fragmented DNA and attach sequencing adapters to the DNA ends. After repair and ligation to prepare the library, DNA was loaded into a FLO-MIN106 R9 flow cell and run in a MinION using MinKNOW.

4.2.2 Sequence assembly

All assembly and quality checking steps were done by the Next Generation Sequencing department at NCIMB Ltd.

WGS runs were assembled by Unicycler v0.4.7 (Wick *et al.*, 2017), SPAdes v3.13 (Bankevich *et al.*, 2012), and polished using Pilon v1.22. Sequence data was used to generate four assemblies for comparison: MinION + paired end (PE), MinION + combined PE/single end (SE) with no PE read pair linking, MinION + PE/SE with PE read pair linking, and MinION only. Sequence analytics were derived from QUAST v5.0.2 (Gurevich *et al.*, 2013) and genome completeness assessment by BUSCO v3.0.2 (Simão *et al.*, 2015) or CheckM (Parks *et al.*, 2015). Plasmids were searched for using the Gram Positive option of the PlasmidFinder tool (Carattoli *et al.*, 2014).

4.2.3 *Streptomyces* BGC Prediction with antiSMASH

Fast-All (FASTA) files were used as the input for antiSMASH v5. All Extra Features (KnownClusterBlast, ActiveSiteFinder, ClusterBlast, Cluster Pfam analysis, SubClusterBlast, and protein family-based gene ontology term annotation) were included. After running, GenBank cluster files were downloaded to create a DNAPlotter gene map in Artemis v17 (Carver *et al.*, 2012).

4.2.4 autoMLST cladistics

S. costaricanus and *Streptomyces* sp. endophyte N2 FASTA files were submitted to autoMLST in denovo mode with the default settings to produce a concatenated alignment. For the remainder of this chapter, "*Streptomyces* sp. endophyte N2" will be referred to as "*Streptomyces* sp. N2".

4.3 Results and Discussion

4.3.1 Whole genome assembly of *S. costaricanus*

All genome sequencing and assembly in this section was done by the Next Generation Sequencing department of NCIMB Ltd.

4.3.1.1 Initial single technology sequencing

Initial sequencing of *S. costaricanus* with an Illumina MiSeq produced 2 assemblies, both of which led to BGCs under 1 kb long when analysed antiSMASH v4 (Blin *et al.*, 2017) indicating a fragmented misaligned genome. Therefore, to compliment the short MiSeq reads, prevent misalignment, and bridge gaps in the short reads, a longer read Oxford Nanopore MinION sequence was generated. Characteristics for these sequences are listed in Table 4.1.

Table 4.1 Statistics for 3 initial full genome sequences

Sequencer	Coverage	Sequencing type	Contigs	Size/Mbp
Illumina MiSeq	20x	250 bp PE	1210	7.94
Illumina MiSeq	65x	300 bp SE	87	8.31
Nanopore MinION	140x	1 strand (1D)	1	8.25

4.3.1.2 Hybrid genome assembly

The 3 assemblies were processed into 4 combinations:

Assembly 1) Nanopore plus 250 PE.

Assembly 2) Nanopore plus a combined 300 SE and 250 PE library, but only using SE data from both.

Assembly 3) Nanopore plus 300 SE and 250 PE, using SE and PE data.

Assembly 4) Nanopore only.

Statistics for each hybrid assembly are shown in Table 4.2. This includes the N50, which refers to the length of the contig required to reach ≥50% of genome size when the contigs are ordered longest to shortest, and the L50, which is the number of contigs used to reach N50. A more complete sequence is likely to have a high N50 and low L50, but this on its own is not enough to confirm genomic quality. Assembly 4 is made of just 1 contig so has a low L50 and high N50 but is smaller than the other assemblies, indicating that there may be some missing data.

Assembly 1 is the largest sequence, but a large number of contigs – 19 out of 23 – under 5k bp, which indicates that the MinION data could not fully align all of the shorter read contigs produced from the MiSeq runs. Aligning assembly 1 to what should be the most structurally correct assembly (assembly 4, as it is 1 contig) using progressive MAUVE (Darling *et al.*, 2010) shows inconsistent genome mapping between the two assemblies; in comparison, aligning assemblies 2 and 4 shows very high synteny (Figure 4.1). Therefore, assembly 1 was not used for any further analysis as the comparisons of assemblies 1 and 4 versus 2 and 4 shows assembly 1 to be structurally incorrect. Furthermore, the low PE coverage will have led to incorrect base calling and likely amplified any low GC read bias. If this data were used to enhance the MinION sequence, the low short read quality would likely be insufficient to produce any improvement.

Assembly 4 has the lowest Benchmarking Universal Single-Copy Orthologs, given in in Table 4.3 (BUSCO; Simão *et al.*, 2015) score, despite its use as the structural reference. This indicates relative genomic incompleteness, and as the assembly also lacks any improvement from the MiSeq data assembly 4 will also not be used further. Although the long read length is advantageous for overall genomic structure and repetitive sequences, Nanopore – and other long read technologies, like Pacific Biosciences – assemblies can have a higher error rate per base than Illumina, at 5 to 15 against <1% (Wick *et al.*, 2017).

Nanopore technology operates through measurement of electric potential as a single strand of DNA passes through a nanopore, with each base giving a distinct alteration in potential and so allowing inference of the sequence (Rang *et al.*, 2018). This is also the first potential source of error, which can be induced through the DNA passing through at varying speeds, pushing multiple bases

through at once, especially if the bases have additional chemical modifications, or with a low signal to noise ratio. This can easily lead to fragmentation as errors build up over the sequence, resulting in a low BUSCO score.

Of the two most remaining assemblies – 2 and 3 – assembly 2 produced better BUSCO results with only 2 contigs. Assembly 2 was therefore chosen for further analysis as the best assembly and was re-processed to remove sequencing adapters and other contaminants followed by final polishing of the assembly using both sets of MiSeq data. This gave the final assembly sequence, for which statistics are shown in Table 4.4. Further genome quality checking with CheckM (Parks *et al.*, 2015), which like BUSCO uses a set of reference genes and their position within a genome to determine assembly quality, gave a completeness score of 99.89%, outperforming assembly 3.

Table 4.2 Statistics for the 4 hybrid assemblies.

Assembly Statistic	Assembly 1	Assembly 2	Assembly 3	Assembly 4
Genome size (bp)	8588679	8350565	8568038	8246559
N50 (bp)	7318482	5487367	3578644	8246559
L50	1	1	2	1
GC %	71.73	71.79	71.71	71.76
Total contigs	23	2	13	1
Contigs larger than 1 kb	23	2	13	1
Contigs larger than 5 kb	4	2	6	1
Contigs larger than 10 kb	4	2	5	1
Contigs larger than 50 kb	3	2	5	1
Largest Contig (bp)	7318482	5487367	3842139	8246559

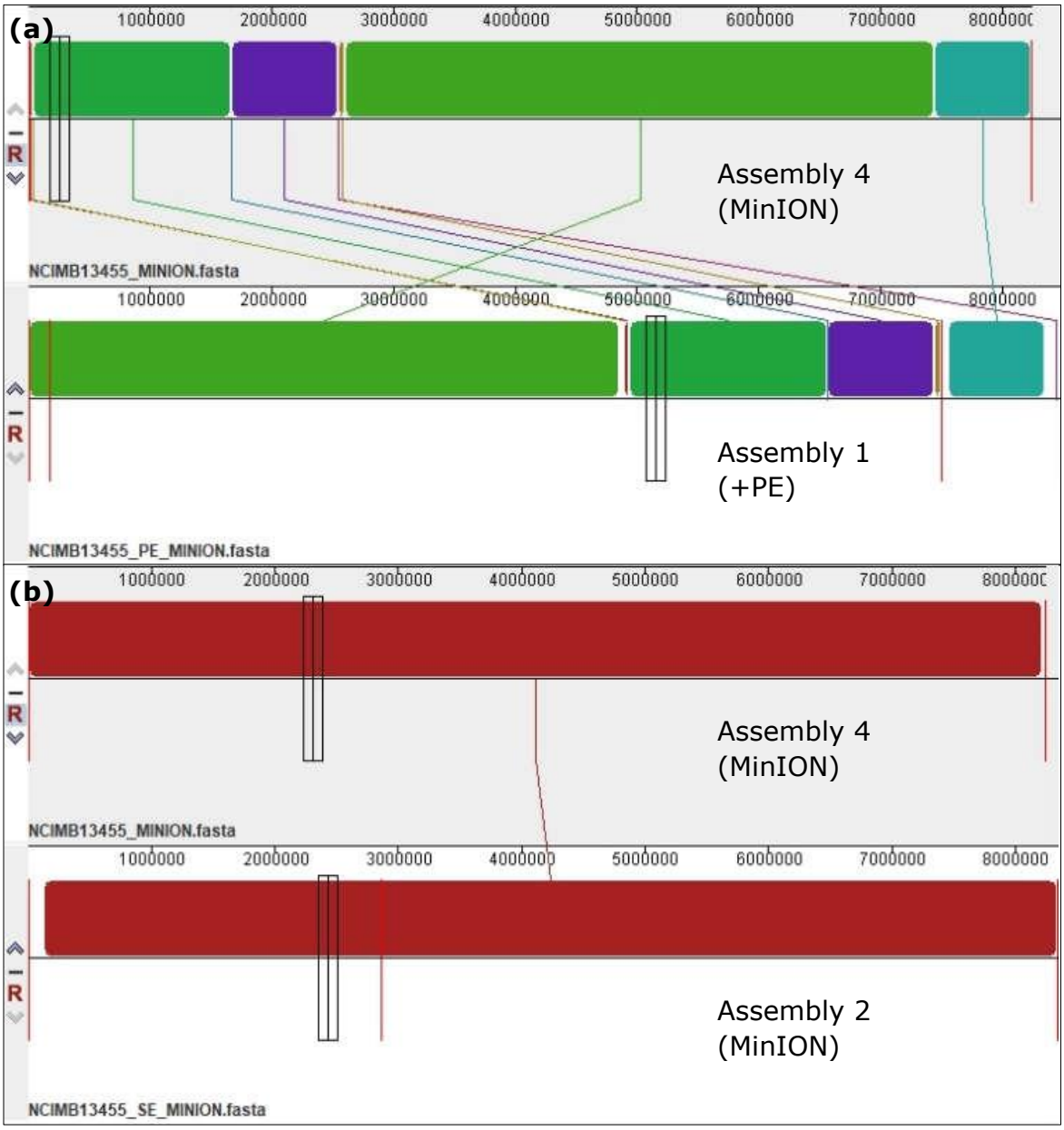
Table 4.3 Genome assembly completeness as determined by BUSCO.

Assembly	Complete Genes	Complete Single Copy Genes	Duplicate Genes	Fragmented Genes	Missing Genes	Completeness %
1	143	141	2	0	5	96.7
2	143	141	2	0	5	96.7
3	143	139	4	0	5	96.6
4	69	69	0	42	37	46.6

Table 4.4 Final assembly statistics.

Assembly Statistic	Value
Length (bp)	8349884
Contigs	1
GC %	71.79
Complete BUSCOs	143
Duplicated/Fragmented BUSCOs	0
Missing BUSCOs	5

Figure 4.1 Progressive MAUVE sequence alignment of Assemblies 1 vs 4, and 2 vs 4 to determine structural correctness.



4.3.2 Genomic exploration of *S. costaricanus* with RAST

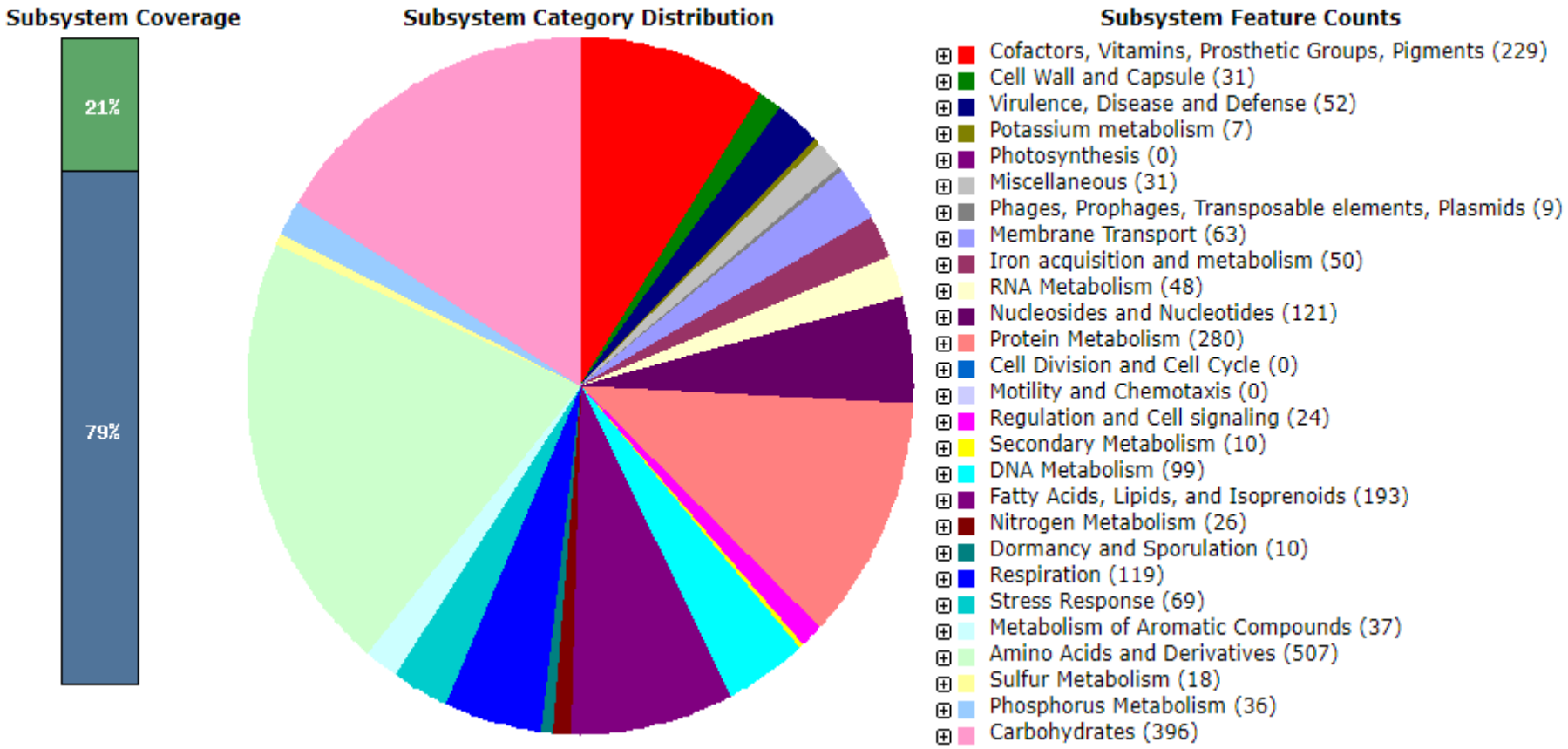
The *S. costaricanus* genome FASTA file was run through the Rapid Annotation using Subsystem Technology-SEED (RAST; Aziz *et al.*, 2008) server. RAST detected 87 RNAs and 7,449 coding sequences (CDS) across 341 subsystems; a subsystem is defined by Overbeek *et al.* (2005) as a set of functional roles which implement a biological process or structural complex. As an example, the process of glycolysis is comprised of enzymes carrying out different functional roles, and the ribosome complex is made of proteins performing functional roles.

Of the 7,449 CDS, 21% were located to a subsystem and 79% unallocated (Figure 4.2). This corresponds to 1546 proteins (1481 non-hypothetical, 65 hypothetical) and 5903 proteins (3172 non-hypothetical, 2731 hypothetical). The largest subsystem is "Amino Acid and Derivates" subsystem with 507 genes, followed by Carbohydrates at 396. Headings are not necessarily immediately reflective of gene function: the "Virulence, Disease and Defence" category contains *Mycobacterium* virulence operons, but these are given protein synthesis and DNA transcription functions. Subsystems may not include relevant genes which are known to be in the organism because there is no appropriate sub-category. For example, desferrioxamine E genes are detected, but another siderophore's genes are not as desferrioxamine E is the only microbial siderophore category.

Included in this category is the "Resistance to Antibiotics and Toxic Compounds" subsection. As a prolific bioactive specialised metabolite producer, *Streptomyces* species must be able to resist the toxic effects of their own antibiotics, e.g. with efflux pumps (Hopwood, 2007), so resistance genes should be expected even more than for a standard soil-inhabiting microbe. Resistance genes for streptothricin (an acetyltransferase), fluoroquinolones (2 topoisomerases and 2 DNA gyrases), and beta-lactams (class A beta-lactamase, metal-dependent hydrolase beta-lactamse) were detected. As stated in section 1.2.1, *Streptomyces* and other antimicrobial producers need resistance genes to avoid self-destruction and may be the source of some of these in the wider microbial population, so the presence of resistance genes is expected.

Specific resistance genes may also imply production of the associated antibiotic (Alanjary *et al.*, 2017). Efflux pumps may be more linked to self-resistance compared to destructive enzymatic methods since a pump will not destroy the metabolite, which would invalidate the energy cost used to make it. There are specific tools dedicated to detecting potential resistance genes associated with biosynthetic machinery (Alanjary *et al.*, 2017), but no immediately significant results were produced for *S. costaricanus.*

Figure 4.2 RAST Subsystem distribution for *S. costaricanus*.



Multiple subsystems relate to characteristic *Streptomyces* processes like specialised metabolism and sporulation, though the Secondary Metabolism subsystem itself contains categories more tailored towards plant metabolites and is not designed specifically for bacterial specialised metabolite detection. The "Iron Acquisition and Metabolism" subsystem contains genes for desferrioxamine E biosynthesis (*desA – E)* and secretion. Siderophores are widespread metabolites with a vital role in iron uptake, making them one of the few metabolites with known ecological functions (Wibberg *et al.*, 2018).

S. costaricanus was then compared to the model species *S. coelicolor* to assess homology between the two species. An overall protein sequence homology comparison between the two species is shown in Figure 4.3, along with any non-matching regions (white space in the circle), and a protein BLAST dot plot, which displays conservation and sequence order within the genome. Comparing the two species gives 24 proteins which retain 99%+ similarity most of which code for important ribosomal subunits transcriptional regulators or sporulation-associated proteins, the last of which are especially relevant for *Streptomyces*.

The dot plot shows some significant sequence inversions compared to *S. coelicolor*. As stated above, one *Streptomyces* characteristic is the variability of its genome to enhance BGC evolution, and as such inversions are seen when comparing between different species. As examples, *S. ambofaciens* is closely related to *S. coelicolor* but its genome contains 2 inversions around its origin of replication (Choulet *et al.*, 2006), *S. albulus* ZPM showed inversions to 11 other strains (Wang *et al.*, 2015), as did *S. albus* J1074 against *S. coelicolor* and *S. bingchenggensis* BCW 1 (Zaburannyi *et al.*, 2014).

Any of the 99%+ protein functions listed as hypothetical in both *S. coelicolor* and *S. costaricanus* were analysed with BLASTP to search for any further information. One protein remained hypothetical and could only be identified as a multispecies actinobacterial AURKAIP1/COX24 domain-containing protein; the DNA sequence s conserved in *Streptomyces* as shown by a BLASTN search, with 100% query cover in *Streptomyces* sp. N2 and *S. collinus* Tu 365 having the joint highest scores. AURKAIP1/COX24 has been identified as part of the S22 subunit in the *Mycobacterium smegmatis* ribosome (Hentschel *et al.*, 2017), which would

explain the level of conservation. The full list of ≥99% proteins is shown in Table 4.5, mostly comprising ribosomal subunits and transcription factors.

Figure 4.3 (**a**) Comparison of protein identity homologies in *S. coelicolor* to *S. costaricanus*. (**b**) Blast dot plot comparison of *S. coelicolor* to *S. costaricanus.* Diagonal lines show conserved regions between the two species. Perpendicular sections indicate inverted sequences that are otherwise conserved.

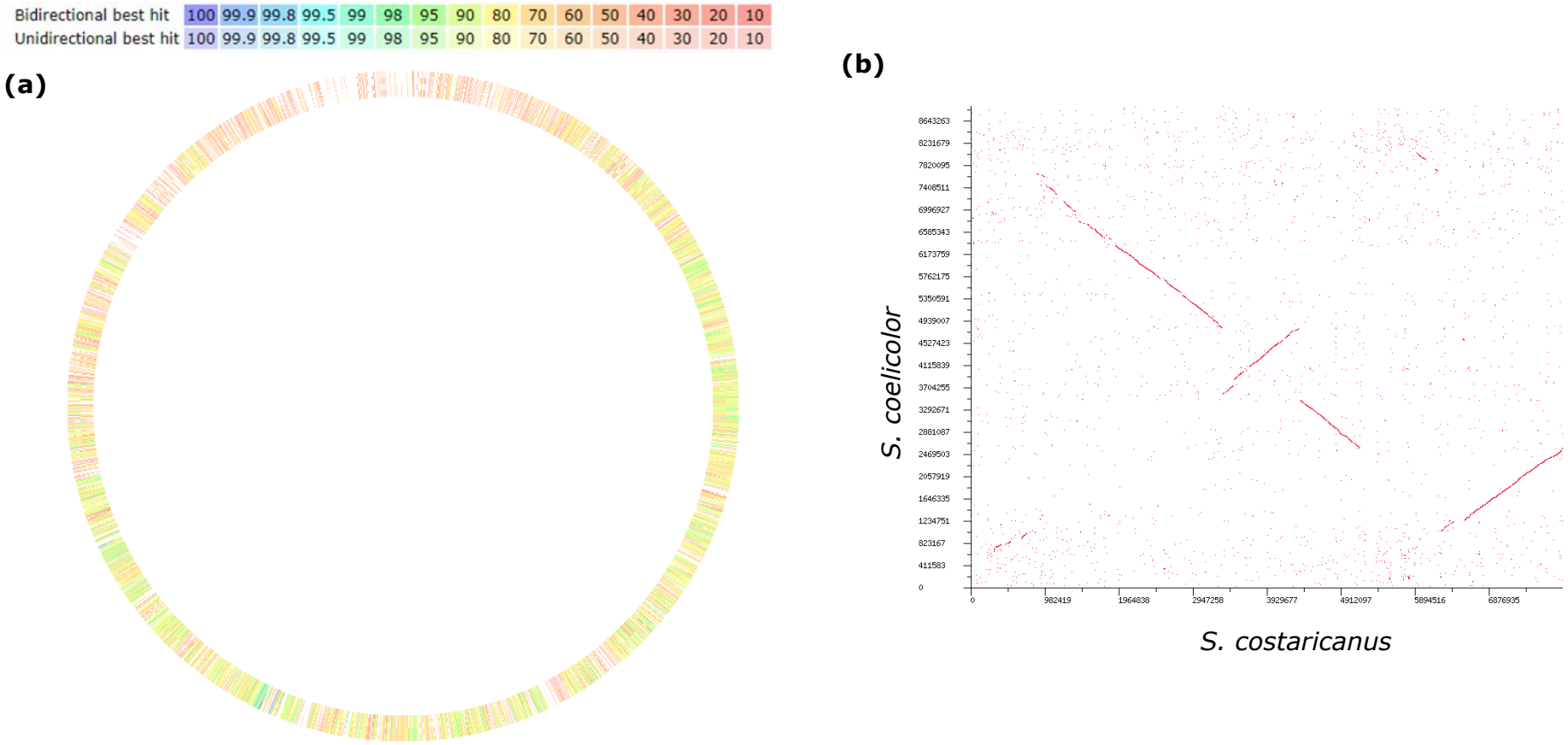


Table 4.5 List of sequences with >99% homology between *S. coelicolor* and *S. costaricanus*.

***S. costaricanus* Function**	**Homology % to *S. coelicolor***	***S. costaricanus* Function**	**Homology % to *S. coelicolor***
RbpA, RNAP-binding	100	Ribosomal Subunit L30p	100
Sporulation cell SsgB, sporulation and cell division	100	RNA Polymerase Alpha Subunit	100
WhiB, Sporulation regulator	100	LuxR family transcriptional regulator	100
Hypothetical	100	Peptidase/RNAse associated protein	100
Ribosomal Subunit L34p	100	LuxR family transcriptional regulator	99.54
Ribosomal Subunit S18p	100	Ribosomal Subunit S7p	99.36
BldC, DNA binding	100	GTP-ase activator	99.29
CarD-like regulator	100	Ribosomal Subunit S11p	99.25
Ribosomal Subunit L33p	100	Ribosomal Subunit L14p	99.18
Ribosomal Subunit S12p	100	Anti-sigma factor BldG	99.12
Ribosomal Subunit S10p	100	Putative electron transport protein	99.1
Ribosomal Subunit S14p	100	Putative integral membrane protein	99.09

4.3.3 Biosynthetic Gene Cluster Prediction with antiSMASH

The full genome sequence was comprised of 2 contigs. Prior to running in antiSMASH these were concatenated into a single contiguous sequence orientated to start at *dnaA*. BGCs in *S. costaricanus* were detected by running the single contig FASTA file through antiSMASH v5 (Blin *et al.*, 2019).

antiSMASH results includes searches to known clusters based on the Minimum Information about a Biosynthetic Gene Cluster database (Medema *et al.*, 2015; Blin *et al.*, 2019), providing a homology score for any matching clusters. This score does not necessarily reflect product identity, whether of the metabolite synthase or metabolite itself, as genes for transport, tailoring, resistance and other non-core functions are included in the score.

In total 33 BGCs were detected (Table 4.6), approximately 50% above the genus average of ~22 (21.6) reported by (Doroghazi *et al.*, 2014). Just under half (16) of these were designated as "Hybrids". This does not indicate a chemical hybrid product, or a true hybrid BGC like a PKS-NRPS. antiSMASH v5 first detects protoclusters using conserved core BGC sequence. After this, it assigns them to 1 of 4 types of candidate BGC based on the vicinity of the core region to the next nearest BGC: chemical hybrids, interleaved, neighbouring, and single. Cluster types, breakdown of all clusters into their constituent single protoclusters are shown in Figure 4.4, and their locations in Figure 4.5.

Figure 4.4 antiSMASH v detected BGCs. **(a)** Initial antiSMASH designations of the 33 clusters. **(b)** Breakdown of clusters into their protoclusters. *Streptomyces* typically contain mostly NRPS and PKS modules, so this distribution is expected. Clusters can contain multiple instances of a single class. LAP = linear azoline containing peptide

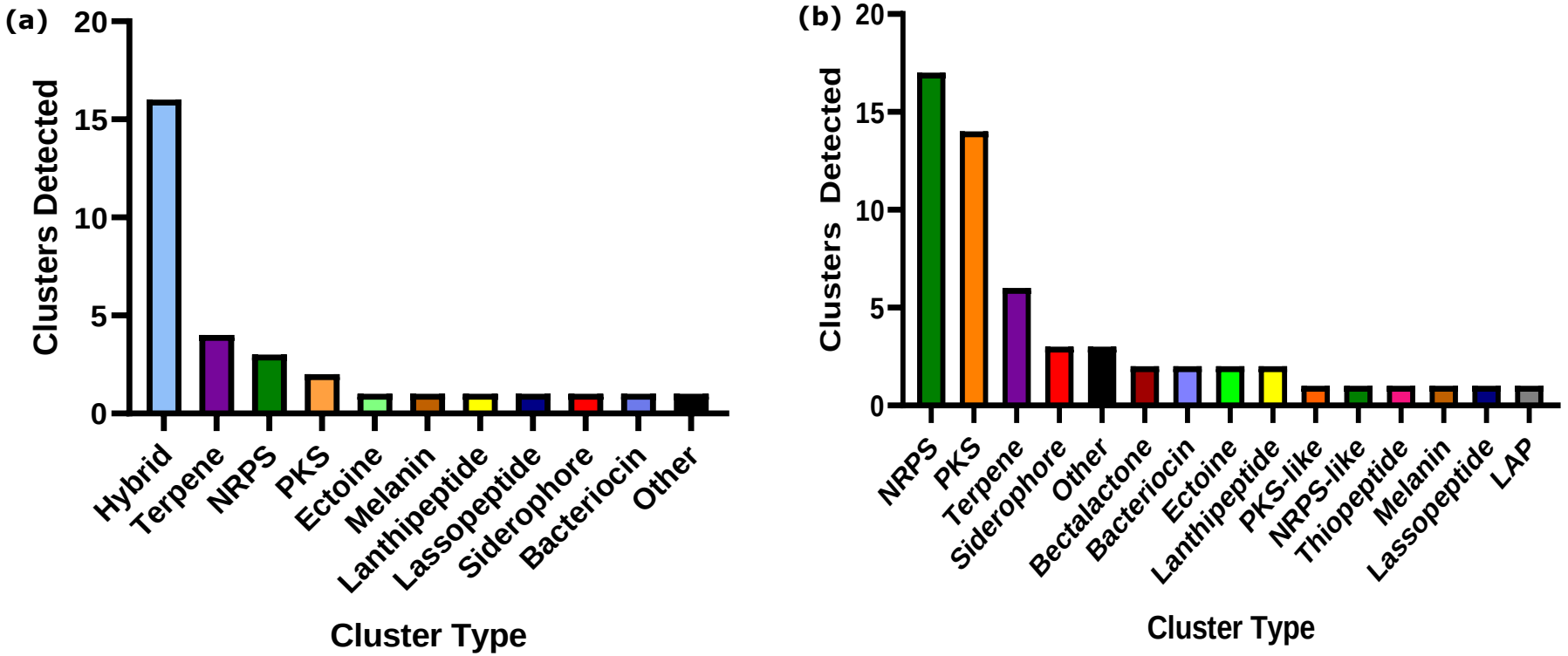


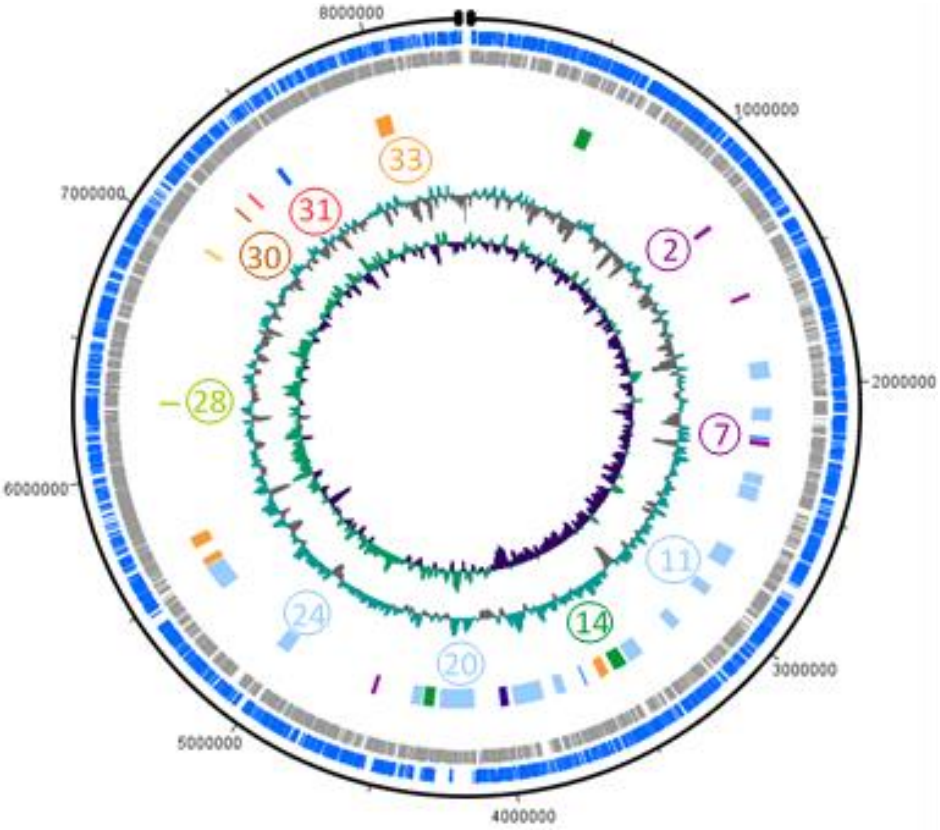
Table 4.6 Full list of all the BGCs detected by antiSMASH. Of the 33 clusters, 10 show 75%+ homology to known clusters, with all but one other under 30%. No homology at all is detected for 4 clusters.

BGC	Genomic location/bp	BGC Type	Homology % to known BGCs	Most similar known product
1	509431-570824	NRPS	8	PM100117 / PM100118
2	1227505-1247571	Terpene	100	Albaflavenone
3	1569701-1589303	Terpene	0	-
4	1878472-1956681	NRPS-Siderophore-T1PKS	22	Kinamycin
5	2090546-2148149	Other-T1PKS-PKSlike	6	Meilingmycin
6	2219984-2230770	Bacteriocin	0	-
7	2241446-2262833	Terpene	100	Geosmin
8	2390342-2440380	NRPS-T1PKS	2	Meilingmycin
9	2446874-2516868	NRPS-Siderophore	18	Friulimicin
10	2741819-2835191	NRPS-Betalactone	16	Kirromycin
11	2931217-2983913	Terpene-Thiopeptide-LAP	92	Hopene
12	3140929-3187387	NRPS-T1PKS	19	Salinosporamide A

13	3353513-3421730	NRPS-Lanthipeptide	12	Bleomycin
14	3436936-3497258	NRPS	78	Mirubactin
15	3525529-3566670	T1PKS	3	Tetronasin
16	3631967-3642182	Bacteriocin	28	Informatipeptin
17	3726161-3772157	NRPS-Betalactone	6	Formicamycins A-M
18	3827264-3959841	NRPS-T1PKS-Terpene	21	Conglobatin
19	3985861-4023927	Other	7	A-503083
20	4135970-4293147	T1PKS-NRPS	100	Pentamycin
21	4316655-4362268	NRPS	66	SF2768
22	4368966-4415961	NRPSlike-T1PKS	11	Borrelidin
23	4577293-4598966	Terpene	0	-
24	4994760-5063175	Other-NRPS	89	Actinomycin
25	5399527-5518637	NRPS-Ectoine	27	Stenothricin
26	5522190-5563254	T3PKS	7	Herboxidiene
27	5617099-5668844	T1PKS-hglEKS	19	Cinnamycin
28	6273774-6284184	Ectoine	100	Ectoine
29	6979189-7001879	Lanthipeptide	0	-

30	7207958-7218563	Melanin	100	Melanin
31	7292799-7304568	Siderophore	83	Deferoxamine B
32	7460750-7483303	Lassopeptide	12	Ikarugamycin
33	7945683-8018288	T2PKS-T1PKS	83	Spore pigment

Figure 4.5 Broken circle genome map of *S. costaricanus* with BGCs. Rings from outer to inner: Forward and Reverse coding sequences; BGC locations with 75%+ homology clusters numbered; GC% plot; G vs C skew with above the line skewed towards G.



BGC	BGC Type	Homology %	Product
2	Terpene	100	Albaflavenone
7	Terpene	100	Geosmin
11	Terpene-Thiopeptide-LAP	92	Hopene
14	NRPS	78	Mirubactin
20	T1PKS-NRPS	100	Filipin
24	Other-NRPS	89	Actinomycin
28	Ectoine	100	Ectoine
30	Melanin	100	Melanin
31	Siderophore	83	Desferrioxamine B
33	T2PKS-T1PKS	83	Spore pigment

Table 4.7 75%+ homology clusters and their potential products. Of these, Actinomycin and Desferrioxamine are the only ones identified in LC/MS spectra, though some (Geosmin) are too volatile for LC/MS.

4.3.4 Targeted cluster and product analysis

A minority (10/33) of clusters had 75%+ homology to known BGCs with designated products (Table 4.7). Of the remaining 23, 4 had 0% (BGCs 3, 6, 23, and 29), 7 under 10%, a further 7 under 20%, 4 under 30%, and 1 at 66%. Without genetic manipulation to induce production of each of the unknown clusters, which would likely be a difficult and time-consuming task in a non-model *Streptomyces* for a single BGC alone, further analysis will be limited to selected high (75%+) and low (0%) homology clusters, and where possible their diversity within the wider antiSMASH database. The selected BGCs coded for albaflavenone, actinomycin, and an unidentified terpene. These were chosen to include an identified BGC that was not detected from *S. costaricanus* culture, one that was both identified and detected in culture, and a totally unknown BGC as a core *Streptomyces* metabolite, known metabolite, and an unidentified product.

4.3.4.1 BGC 2: albaflavenone

The sequence and gene classification for BGC 2 is shown Figure 4.6, and is 100% predicted to code for an albaflavenone terpene synthase. "other genes" indicates genes that are not strictly relevant, such as transposases, some primary metabolism genes, and hypothetical or totally known genes.

Albaflavenone is a sesquiterpene antibiotic active against *B. subtilis* (Gürtler *et al.*, 1994) and is well known as a *Streptomyces* metabolite due to its presence in the *S. coelicolor* genome (Čihák *et al.*, 2017). Despite being a widespread BGC, it is not always expressed under standard culture conditions. Takamatsu *et al.* (2011) investigated a cryptic cluster in *S. avermitillis*, which is the industrial producer of the antihelminith avermectin. The cluster consists of two genes, *sav3031* and *sav3032*, with *sav3031* located downstream of *sav3032*. These genes are homologous to those for albaflavenone production in *S. coelicolor*, *sco5222* and *sco5223*.

Sav3031 codes for a cytochrome P450, which is maintained in *S. costaricanus* according to the predicted function of the equivalent gene. *Sav3032* has high homology to an *S. coelicolor* gene coding for an epi-isozizaene synthase, which catalyses conversion of epi-isozizaene to albaflavenone. The predicted *S. costaricanus* product is a "Terpene synthase/cyclase metal-binding domain"

product, which matches Takamatsu *et al.* (2011) who found that activity was dependent on Mg^{2+}. The metal ion binds to Sav3032 at Asp101 into a metal binding domain with the motif DDRHD, found in the translated amino acid sequence from *S. costaricanus*.

Searching the antiSMASH database for the corresponding MIBiG cluster gives 248 strains: 244 *Streptomyces*, *Actinospica acidiphila* – another actinomycete – and 3 unknowns. All of the results show 100% similarity, indicating a widely conserved cluster. Widespread conservation over *Streptomyces* and other genera could indicate an important ecological function for albaflavenone, but like the characteristic *Streptomyces* terpene geosmin, its function is unknown (Seipke, 2015).

4.3.4.2 BGC 24: actinomycin

The sequence and gene classification for BGC 24 is shown in Figure 4.7, and is 89% predicted to code for an actinomycin "NRPS-Other".

The cluster designated as a core biosynthetic "other" gene codes for LmbU, which regulates expression of the antibiotic lincomycin (Imai *et al.*, 2015). Therefore, the gene is likely better classified as a regulatory element, leaving the BGC as an NRPS. Other elements in the BGC include a cassette of 3 transport-related genes, which are predicted to code for 2 adenosine triphosphate-binding cassette (ABC) transporters, and 1 ABC-2 type transporter but running a BLASTP search further identifies it as a major facilitator superfamily (MFS) transporter. While this family does include many transporters, an MFS in multidrug resistant *Candida* spp. is able to use actinomycin as a substrate and transport it outside the cell (Redhu *et al.*, 2016), much as *S. costaricanus* will have to do to avoid self-toxicity.

Figure 4.6 BGC 2 (Albaflavenone) gene classifications

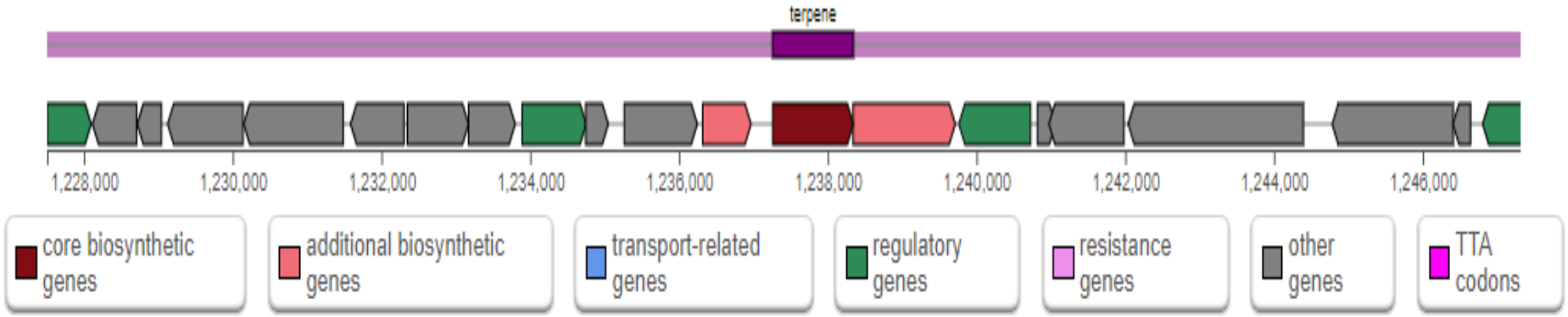


Figure 4.7 BGC 24 (actinomycin) gene classifications

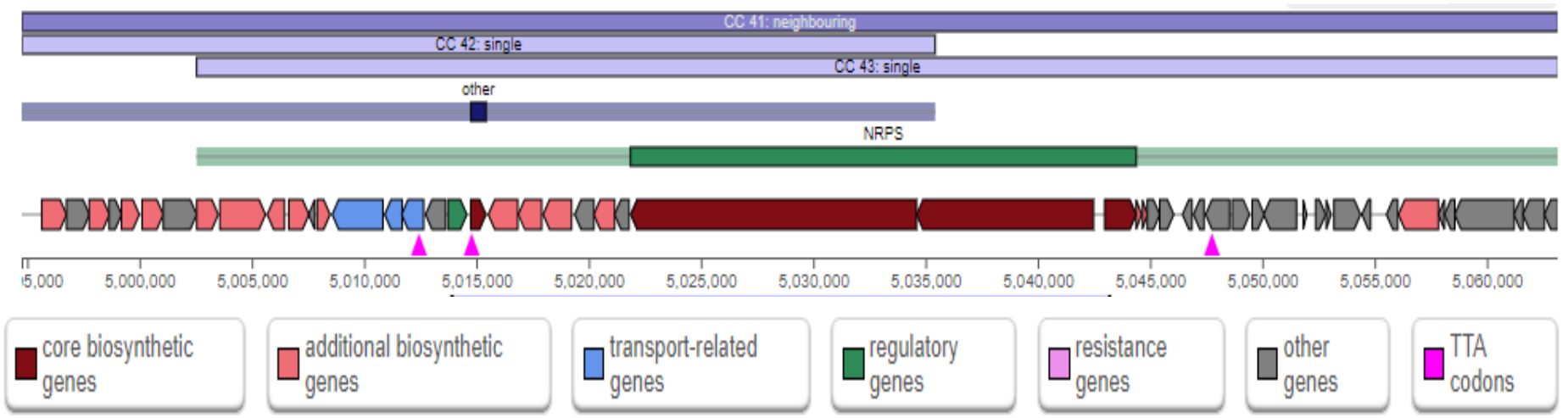
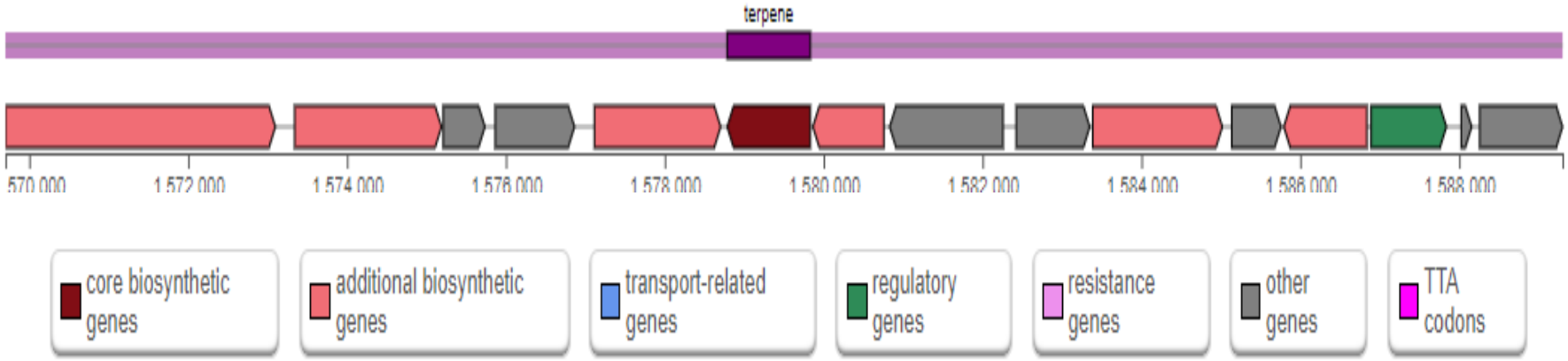


Figure 4.8. BGC 3 gene classifications.



The cluster contains 3 TTA (Leucine) codons, which are rarely seen in *Streptomyces* due to the GC richness: in *S. coelicolor*, there are only 147 instances of it, and they are associated with global scale regulation of antibiotic production and differentiation (Chater & Chandra, 2008; Liu *et al.*, 2013). This is due to their rarity as a codon, meaning that it can be used to designate genes which will only be properly transcribed and translated under TTA specific conditions.

This impacts on antibiotic production, sporulation and MI/MII differentiation, though *Streptomyces* can survive in axenic culture if its ability to use TTA codons is knocked out (Hackl & Bechthold, 2015). One of the codons appears within a putative transposase, identified via BLASTP ISL3 family transposase. A transposase-associated regulatory element would be a risky regulation method, so this is more likely to be by chance than selected regulatory function. However, the terminal ends of *Streptomyces* genomes are known for frequent transposon and plasmid mediated rearrangement, so this is standard *Streptomyces* genomic organisation.

4.3.4.3 BGC 3: unknown terpene

The sequence and gene classification for BGC 3 is shown in Figure 4.8, and is predicted to code for a terpene with 0% additional identification. There are multiple BGCs in *S. costaricanus*, and any non-model *Streptomyces*, with low levels of homology. Cluster 3 was selected for 2 reasons. Firstly, compared to many of the unknown "Hybrid" BGCs, it is a relatively simple 1-BGC type gene so is more likely to be reflective of its final product and would potentially make any later confirmatory genetic manipulation easier. Secondly, it displays very limited homology to strains outside of *Streptomyces*. Running a BLASTP search gives 2 *Streptomyces* with acceptable coverage and match levels, followed by *Frankia sapropphytica* – another *Actinomycetales* species – but low match % (Table 4.8).

Table 4.8 Highest BLASTP results for cluster 3 core synthase.

Strain	Coverage %	Match %
Streptomyces sp. N2	100	99
Streptomyces murinus	98	98
Frankia saprophytica	90	38

Searching the BGC as a whole in MiBIG also produces limited results, with only two results: 28% identification to a pentalenene synthase in the pristinol cluster of *S. pristinaespiralis*, and the same ID to PntA in the pentalenolactone cluster of *S. arenae*.

This implies that BGC 3 is very tightly conserved to a few species, or alternatively that other strains holding the BGC simply haven't been uploaded to the database yet. Assuming that the first choice is correct, one reason for the limited spread could be a very specific ecological niche requirement. If all 3 of the *Streptomyces*, as they had the most similar homology, are from niches with similar requirements then there would be pressure to maintain the terpene. The issue then is why it has not spread to other *Streptomyces*, given the tendency of specialised metabolite genes to recombine and spread throughout the genus. Knowing the ecological isolation information for *S. costaricanus* and the other strains would help identify any niche-specific pressures, but unfortunately there is no specific information past nematode suppressive Costa Rican soil.

4.3.5 Comparison to other members of the *S. costaricanus* clade

4.3.5.1 Difficulties with *Streptomyces* classification

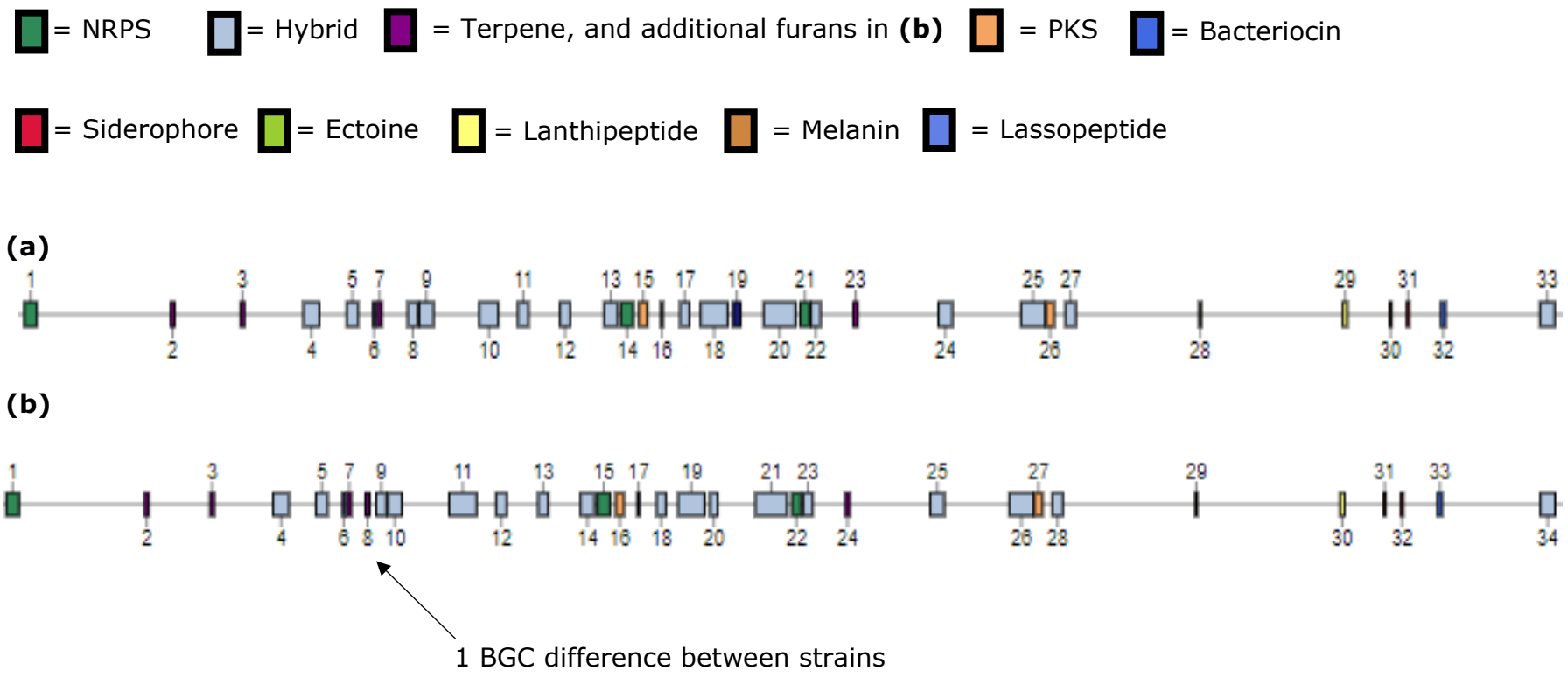
Streptomyces species are notoriously difficult to classify. Classical phenotype-based cladistics is unreliable given a strain's phenotypic variation even within a single colony; the commonly used 16S approach will show high similarity between species without reflecting the extreme levels of recombination in the genomic arm regions (Guo *et al.*, 2008). Unsurprisingly, this leads to species with identical 16S regions but different specialised metabolomes (Antony-Babu *et*

al., 2017), which is an issue if using phylogeny to determine which isolate should be screened for natural product production.

As noted in Section 4.3.4, BGC component BLAST searches produce the same set of strains at high (97%+) homologies. *Streptomyces* sp. N2 is the most frequent, followed by *S.* sp. LamerLS-31b, *S.* sp. DconLS, and various *S. griseofuscus* strains. Given the above difficulties with *Streptomyces* cladistics, it was decided to compare *Streptomyces* sp. N2's antiSMASH and classification outputs to *S. costaricanus* for any similarities.

Figure 4.9 shows antiSMASH outputs for the two strains. The BLAST results had already identified significant similarity between BGC identity between the two, but not their locations within the genome. Orientating both genomes to DnaA shows a near total identity in both BGC composition and also localisation. BGC 1 position differs by under 1 kb, with similar differences for all other clusters, with the exception of a *Streptomyces* sp. N2 specific furan BGC.

Figure 4.9 Comparison of the order of **(a)** *S. costaricanus* BGCs to **(b)** *Streptomyces* sp. N2 BGCs. Colours indicate antiSMASH BGC type:



Labeda *et al.* (2012) group *S. costaricanus* into a clade with *S. graminearus, S. griseofuscus, S. murinus*, and *S. phaeogriseichromatogenes* based on their 16S sequences. A separate previous study (Kurosawa *et al.*, 2006) forms a similar clade linking *S. murinus* to *S. padanus* and *S. misionensis*, which were not included in Labeda *et al*. (2012). While 16S has been the standard tool for microbial classification, it can be an unreliable guide if used in isolation (Alanjary *et al.*, 2019). A more comprehensive approach is to use multi locus sequence analysis, which spans a range of marker genes to form a Multi Locus Sequence Tree (MLST). One recent tool, autoMLST (Alanjary *et al.*, 2019), queries the full genome input against reference genomes plus BGC predictions to give a mix of classification and secondary metabolite information. autoMLST uses the average nucleotide identity (ANI) of two strains to determine if they are the same species, with 95%+ being sufficient.

4.3.5.2 autoMLST and BGC-based cladistics

S. costaricanus was run through autoMLST along with *Streptomyces* sp. N2 to determine their taxonomic relatedness and compare it to the noted BGC similarity. The full tree is shown in Figure 4.10. *Streptomyces* sp. N2 has been putatively identified by the depositors as *S. padanus* based on its secondary metabolite profile (Worsley & Hutchings, personal communication) so unless the inferred 16S matches between Labeda *et al.* (2012) and Kurosawa *et al.* (2016) are not correct the two strains should be in the same clade.

S. costaricanus and *Streptomyces* sp. N2 were mapped to the same clade and given an estimated ANI of 97.6% and 97.5% respectively to *S. griseofuscus*, P value <0.05. This indicates that both species are potentially strains of *S. griseofuscus*, which would take precedence over the type strain of *S. costaricanus* as *S. griseofuscus* is the older type strain (Nagaoka *et al.*, 1986). *S. costaricanus* was re-run through RAST to compare protein homology against *Streptomyces* sp. N2, *S. griseofuscus*, and *S. coelicolor* as a relative outgroup. The results (Figure 4.11) followed autoMLST cladistics: *Streptomyces* sp. N2 was the most similar with 2798 100% homologous proteins, and 2680 more at 99%+.

Given the high ANI between the two strains, and the exceptionally close antiSMASH results, it is suggested that microbial drug discovery projects use a

mixture of antiSMASH and autoMLST results to ensure a diverse set of *Streptomyces*.

The suggested overall workflow does include antiSMASH as a predictor of specialised metabolite potential but following metabolomic screening. As genome sequencing becomes more affordable and efficient, moving genomics to the front of a project with multiple candidate microbes is arguably more useful especially when combined with autoMLST and other tools. Using the BGC similarity alongside cladistics would ensure a wide taxonomic range, increasing the chances of novel BGC expression and discovery.

Using the autoMLST-selected strains to search for patterns in the spread of actinomycin production did not show a similar clade-limited pattern as with the terpene cluster 3, possibly indicating that more common BGCs are not useful for ensuring diversity. A literature search found actinomycin to be produced by 6 other *Streptomyces* species in the autoMLST list, outside of *S. costaricanus* and *S.* sp. N2: *S. avermitilis* (Chen *et al.*, 2012)*, S. antibioticus* (Katz *et al.*, 1956)*, S. griseofuscus* (Kurosawa *et al.*, 2006)*, S. griseoruber* (Praveen & Tripathi, 2009)*, S. iakyrus* (Bitzer *et al.*, 2006)*,* and *S. luteus* (Zeng *et al.*, 2019).

The 3 antiSMASH-designated core regions of the actinomycin BGCs in both *S. costaricanus* and *S.* sp. N2 were compared using BLASTP. The combined genes had an equal number of bases with an overall identity of 98.61% similarity after conversion to amino acids, corresponding to 7235 of 7337 matching residues. The 102 differences were not significantly located in any region of the sequences. The most common changes both involved alanine, with 15/102 differences from alanine to threonine (or vice versa) substitutions, and 10/102 from alanine to valine. Neither of these valine nor threonine share the exact same properties as alanine, though valine like alanine is aliphatic (Ud-Din *et al.*, 2020) possibly limiting any beneficial or harmful impacts on actinomycin synthesis. Comparing the sequence differences and synthesis efficiencies across the BGC-containing strains would be an interesting project, possibly using purified enzymes in a cell free system (Li *et al.*, 2018) to limit the effect of other components.

Figure 4.10 autoMLST of *S. costaricanus* and *Streptomyces* sp. N2. Colours represent ANI groups of 95%+. Arrow shows the location of *S. costaricanus* and *S.* sp. N2. Red stars mark production of actinomycins by either the given strain or its species

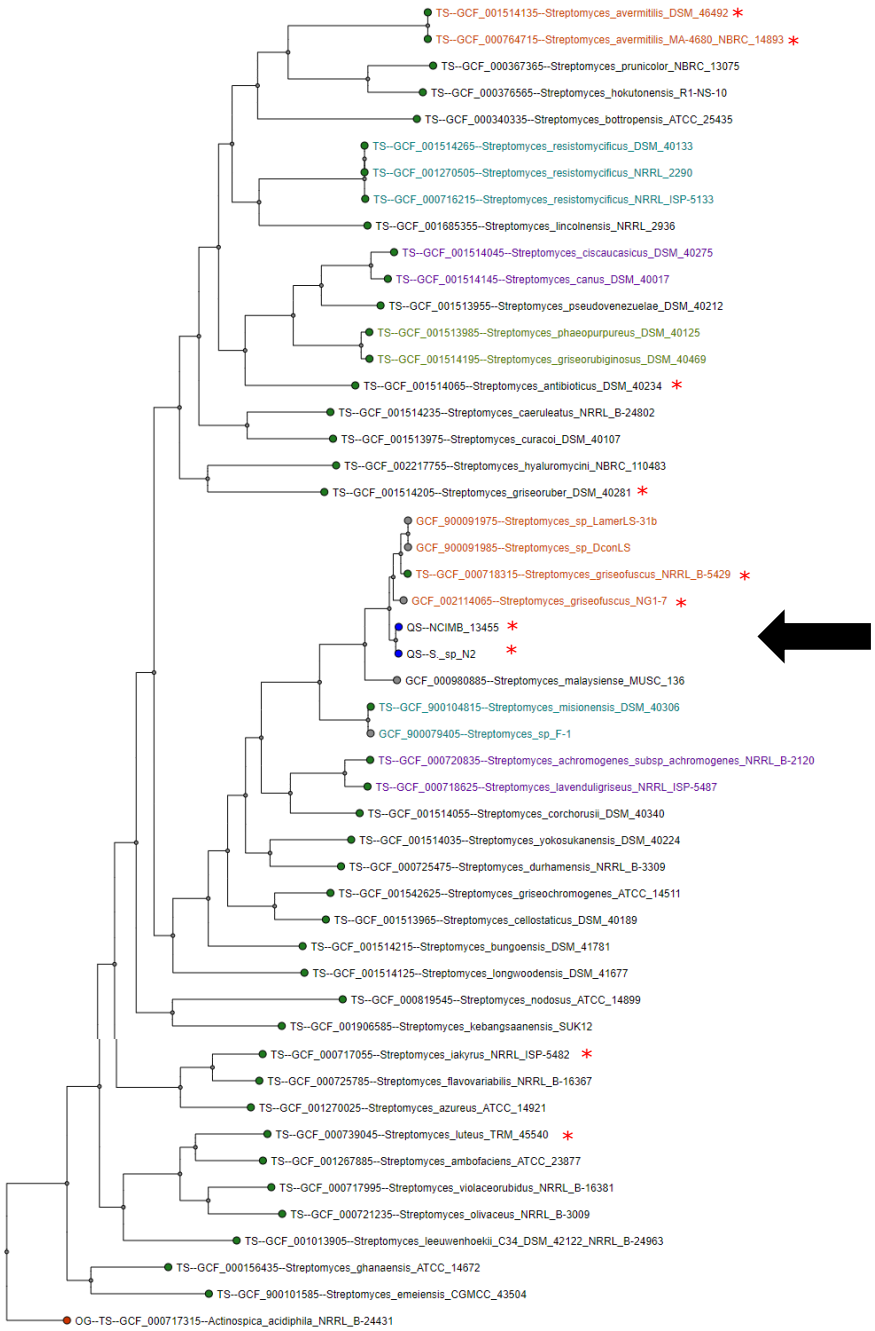
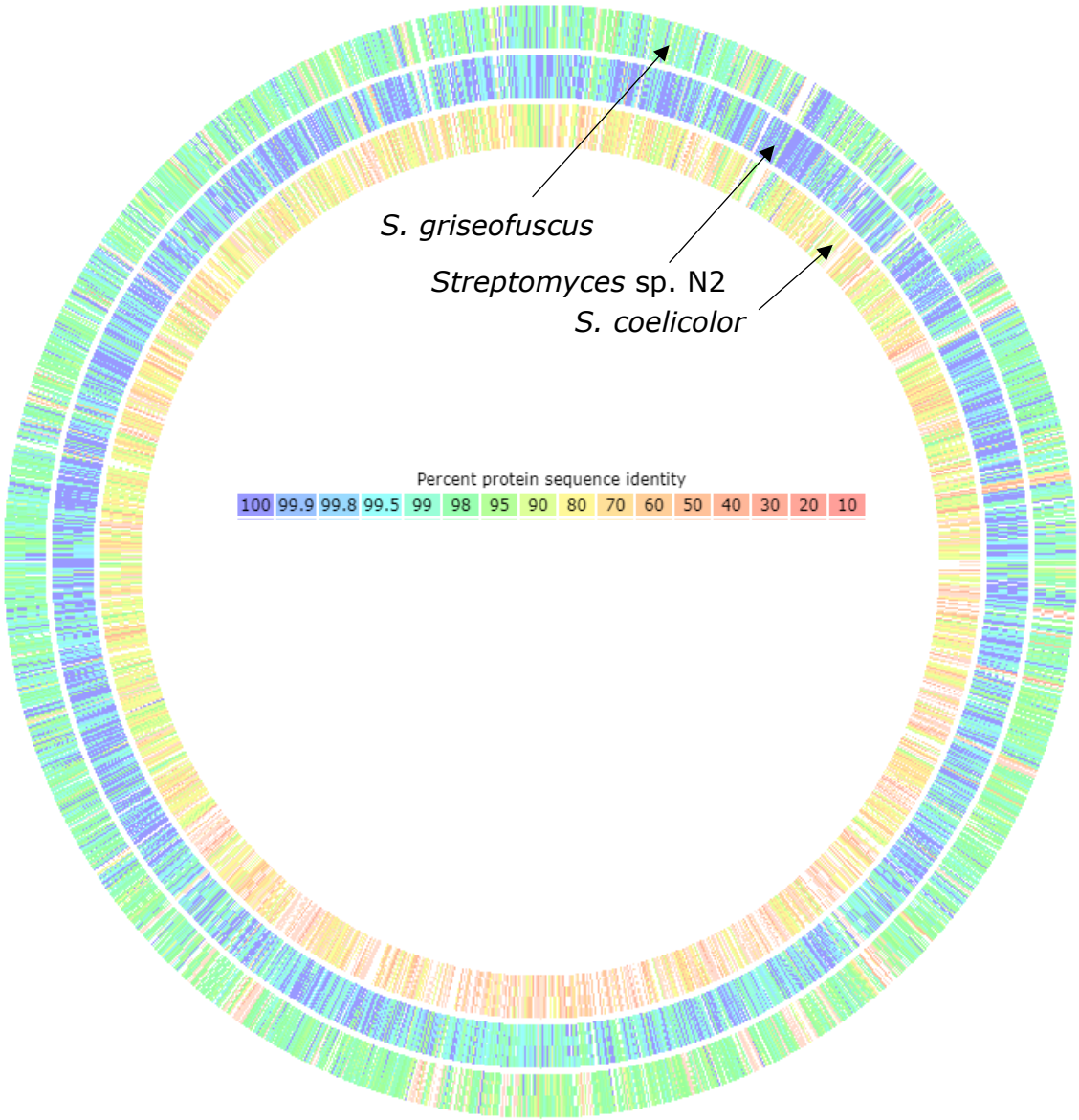


Figure 4.11. RAST protein homology comparison of *S. griseofuscus*, *Streptomyces* sp. N2, and *S. coelicolor* to *S. costaricanus*



4.4 Conclusions

The objectives for this chapter were to obtain a high-quality genome sequence of *S. costaricanus* and use this to explore BGCs along with general genome characteristics compared to other well-studied *Streptomyces*. Using standard genomic completeness and quality tools the sequence was found to be of good quality, which enabled further exploration.

Some results like the presence of the deferor4xamine BGC were expected, given its presence in the core *Streptomyces* genome region, and others – actinomycins – confirmed metabolomic results from previous chapters. The large number of unknown BGCs highlighted the novelty of the genome, as well as the deficiencies in linking available clusters and known metabolites. This emphasised that the metabolomic screening strategy developed over Chapters 2 and 3 correctly selected a strain with high novel BGC potential.

The final use of the sequence was to analyse *S. costaricanus* cladistics. That *Streptomyces* sp. N2 closely grouped to *S. costaricanus* was expected, given its BGC similarity and genomic orientation. Further species reassignment of *S. costaricanus* is outside the sc ope of this thesis, but MLST and BGC similarity does suggest a close familial link between *S. costaricanus* and *Streptomyces* sp. N2 if not also with *S. griseofuscus*.

CHAPTER 5

SCALE-UP CULTURE OF *S. COSTARICANUS* FOR EXTRACTION AND ISOLATION OF BIOACTIVE METABOLITES

5 Scale-up culture of *S. costaricanus* for extraction and isolation of bioactive metabolites

5.1 Introduction

5.1.1 Effect of culture scale-up on metabolite production

The previous chapters examined the impact of nutrition, stresses, elicitors, and solid versus liquid medium on metabolite production in *S. coelicolor* A3(2) and *S. costaricanus*. These studies were all performed on a relatively miniaturised scale to enable more efficient comparison of metabolomic and phenotypic data compared to standard flask and plate culture. While this did lead to generation of a wide range of metabolomic and genomic data, culture scale-up is essential for isolation and extraction of metabolites for use in bioactivity assays, chemical dereplication and characterisation. This adds potentially multiple variables to the project, as how scale-up proceeds can affect metabolite production. This could be impacted by, for example, the vessel shape, oxygenation levels, inoculum differences, messenger molecules, or pH change. Some examples are discussed below.

5.1.1.1 Vessel Depth and Diameter

Barkal *et al.* (2016) used their own miniaturised system to study specialised metabolism in *Aspergillus nidulans*, altering both physical and chemical parameters to study the resulting chemotypes. The physical parameters included both diameter and depth of the culture vessel, though only well diameter was found to influence metabolism. Using principal component analysis to visualise the separation, there was a significant difference between metabolites produced by *A. nidulans* grown in 2, 3.5, 5.5, and 10 mm diameter agar-containing wells, with no separation found from well depth.

The example is not perfect in that well depth will be less relevant using agar since it is a solid medium and so will restrict movement especially in a filamentous microbe, though this will depend on the organism growth speed and culture duration. As growth through the medium to unused nutrients still

happens, nutrient starvation in the smaller volumes is not the primary factor for the variation. The principal component analysis levels of variation explained were not particularly high (29% PC1, 18% PC2) limiting its predictive power and potentially giving a false impression of the diversity of metabolism.

Separation was still significant especially when comparing between the smallest and largest wells, then to the standard agar plate extract control, so the overall example is still valid. It is unknown if there is a linear relationship between specific metabolites produced and the physical characteristics of the culture vessel or if after a certain diameter there are no additional metabolite changes. The wells of the 24-well plate are 15.6 mm wide, larger than that reported by Barkal *et al.* (2016), so assuming the same regulation applies to *S. costaricanus* as *A. nidulans*, some of the differential impact of scale-up should be relatively low.

5.1.1.2 Oxygenation and Growth Rate

Physical characteristics can also affect oxygenation levels. Growth in baffled or spring-coiled flasks is common in culture of S*treptomyces* (Kieser *et al.*, 2000) since it increases oxygen availability and prevents some of the clumping exhibited in liquid media, leading to higher and more dispersed growth. As with nutrition levels, this does not mean that more oxygenation necessarily leads to improved metabolite production, since specialised metabolism is regulated by multiple factors.

This was demonstrated by growth of *S. lividans* in baffled and spring-coiled flasks where production of undecylprodigiosin (RED) was consistently lower than in standard shake flasks, in some cases 385 times less efficient (Gamboa-Suasnavart *et al.*, 2018). The lower oxygenation from the standard shake flasks was reflected in an anaerobic level respiratory quotient and slower growth rate. It was therefore proposed that a low oxygen transfer rate can trigger the onset of specialised metabolism and reducing primary metabolite production, lowering the growth rate but increasing specialised metabolite production. In contrast, expression of the antifungal metabolite pimaricin from *S. natalensis* is upregulated via quorum sensing, so will be positively affected by higher oxygenation and growth rate (Recio *et al.*, 2004)

The literature is not entirely in agreement over the positive impact of low oxygen transfer to specialised metabolism, especially given the complex regulation of the expression of different metabolites. *Streptomyces* species are obligate aerobes and are unable to grow under anaerobic conditions. *S. coelicolor* can survive anaerobic conditions for 24 d then immediately resume growth once oxygen is added, as do the majority of strains tested to date, although spores of *S. avermitilis* decline to 20% viability by 10 d and reach 0% by the end of the time course (van Keulen *et al.*, 2007). *S. coelicolor* can express a set of respiratory nitrate reductases normally found in facultative and obligate anaerobes. These may help anaerobiosis through reducing nitrate to maintain the proton motive force (Falke *et al.*, 2016).

Sohoni *et al.* (2012) used 24-square deep well plates to make their own miniaturised cultivation system for *S. coelicolor*, adding glass beads of various sizes to the wells to study the effect on chemotype and phenotype. Oxygen transfer was increased by 3 mm beads, as was production of the characteristic pigment actinorhodin by approximately 25 times. The two studies cannot be directly compared since they study different metabolites in different vessel types with different organisms, but it is a useful reminder of the potential variety in regulation between strains and culture types. Since this study has so far not investigated glass beads and used standard wells, scaling up to standard shake flasks without beads or baffles will ideally give a similar metabolite profile.

5.1.1.3 Inoculum Type

A standardised inoculum is a vital part of maintaining consistency across cultures. There are two common master stock storage methods: freezing mycelia (Charusanti *et al.*, 2012) and spore suspensions (Shepherd *et al.*, 2010), each of which has advantages in certain situations. When maintaining a large *Streptomyces* or other mycelial microbe library, frozen mycelia arranged on a plate are a simple way to quickly replicate stocks using a multi-pin replicator. For specialised metabolite experiments which rely on transitions between growth phases (Wentzel *et al.*, 2012), cell density and messenger molecules (Polkade *et al.*, 2016), a spore suspension is a more precise way of measuring out the same concentration of inoculum each time. All other factors remaining equal, this

guarantees identical growth rates for each experiment and so reliable metabolite production.

Using a pin replicator for the 24-well plate would have technically saved inoculation time, but the time saving is easily outweighed by the differences in inoculum densities. A spore suspension is not ideal for inoculating a larger culture volume. Even if it would standardise the inoculum between scale-up and the 24-well plate screen, the 1 ml spore suspension obtained per plate is inadequate to seed multiple larger culture volumes. Starter cultures are a frequent way to maintain inoculum equality between scale-up vessels (Rateb *et al.*, 2011; Thaker *et al.*, 2013; Xie *et al.*, 2014) while limiting master stock usage.

While a complete link cannot be maintained between the 24-well plate screen and scale-up culture, a logical choice of vessel type and culture techniques ensure that the difference between the small and large cultures can be minimised. This will in turn keep variation between scale-up replicas minimal, giving reliable metabolite production in each flask.

5.1.1.4 Scale-up of S. costaricanus

At the start of the project there were no reports of scale-up, or any recent experiments, using *S. costaricanus*. Since then, 1 study has appeared on the actinomycin D biosynthetic gene cluster in a marine strain of *S. costaricanus* (Liu *et al.*, 2019). The volumes of media used in the study are also smaller than the planned scale-up (50 ml, 250 ml flasks) and is focused on smaller scale extraction of actinomycin D to explore the effects of different mutants on biosynthesis.

Both agar and broth media were used for scale-up, but this is not a direct comparison since extracting from the agar and biomass will include both intracellular and secreted metabolites. While some *Streptomyces* metabolites are intracellular, such as RED, most antimicrobials are thought to be secreted into the environment (Senges *et al.*, 2018). This will be affected by the differences of agar versus submerged growth but explains why metabolomics experiments frequently use only the supernatant instead of the pellet.

5.1.2 Chapter Aim and Objectives

Chapters 2 and 3 developed an untargeted metabolomics workflow based on metabolite extraction from 24-well plates. The media volumes used in the plates are not sufficient for metabolite isolation from *S. costaricanus*, so scale-up was required. The growth condition selected was M19 media after 5 d growth, since these parameters produced relatively high numbers of both total features and those specific to culture in M19, and principal component analysis (Section 3.2.3) showed 5 d as significantly different from most other conditions. Therefore, scale-up using M19 broth would enable the final part of the early stage metabolite discovery workflow: metabolite purification through fractionation, bioactivity, and novelty assessments. The objectives for this chapter were therefore to:

1) Scale-up *S. costaricanus* using M19 agar and broth culture.
2) Extraction of agar culture metabolites and analysis with UPLC-MS.
3) UPLC-MS analysis of broth culture supernatant.
4) Flash fractionation of both supernatant and agar extract to purify metabolites.
5) Bioactivity guided investigation and further purification of metabolites in relevant fractions.

5.2 Materials and Methods

5.2.1 Scale-up of *S. costaricanus* culture, metabolite extraction, and fractionation

5.2.1.1 *S. costaricanus* broth culture and extract collection

S. costaricanus inoculum was prepared as in Chapter 3, Section 3.3.6. Inoculum (2 ml/flask) was used to inoculate 15 x 500 ml baffled shake flasks containing 200 ml M19 broth for a total of 3 L. Flasks were shaken at 200 rpm, 30°C, for 5 d. The supernatant was separated from biomass by centrifuging at 4,122 x g and for immediate fractionation to prevent enzymatic degradation.

Biomass for pellet metabolite extraction trials was extracted by suspending the pellet in 5 times its volume of 50% aqueous methanol and shaking at 200 rpm overnight, room temperature.

5.2.1.2 *S. costaricanus* agar culture and crude extract preparation

Inoculum was prepared as in Chapter 3, Section 3.3.6. Inoculum (250 µl/plate) was added to 117 plates with approximately 25 ml M19 agar in each, for an approximate total of 3 L agar across the plates. Plates were allowed to dry in a laminar flow hood then left to incubate at 30°C for 5 d.

After incubation, metabolites were extracted by removing the agar and biomass from the petri dish and submerged in 3 L of 50% aqueous methanol. The mixture was left at room temperature for 24 h, then the solvent drained and centrifuged at 4,122 x g for 10 min to obtain the crude extract.

5.2.1.3 Flash chromatography of broth and agar extracts

The agar crude extract was diluted down to 10% methanol using 18.2 MΩ water, after which both the supernatant and agar extract were processed with the same method.

An Isolera One Flash Purification System (Cardiff, UK) was used to precondition a Biotage 120 g C18 resin cartridge using 250 ml of methanol followed by 250 ml

18.2 MΩ water, both at a flow rate of 50 ml/min. The extract was then loaded onto the cartridge at a flow rate of 50 ml/min. A gradient using 10% increments of methanol (2 x column volume or 250 ml/step) was used to elute compounds from the C18 cartridge in 120 ml maximum fractions, monitoring the eluate at 240 nm.

Compounds unretained by the C18 resin were loading onto an ENV+ (Biotage) resin cartridge to capture more polar chemicals that would otherwise have been discarded. The ENV+ cartridge was preconditioned, loaded and culture fractions eluted in the same way as the C18 cartridge. Both sets of fractions were stored at -4°C for later UPLC-MS analysis. Dry weights were calculated by aliquoting 3 ml of each fraction in a pre-weighed vial, then evaporated to dryness in a Genevac EZ2 centrifugal evaporator (Genevac, Ipswich UK) and subtracting the dry vial weight from the vial plus fraction weight.

5.2.1.4 Polarity-based fraction pooling and purification

Fractions were pooled into 3 groups by their most dominant metabolite family (Actinomycins, Deferoxamines, or Other), diluted with 18.2 MΩ water to 10% methanol concentration, loaded onto respective 120 g C18 cartridges, and eluted over pool polarity specific gradients: actinomycins from 40 to 100% methanol, deferoxamines 25 to 75% methanol, and the remaining fractions 0 to 60% methanol. The maximum fraction volume was also set to 60 ml to further concentrate metabolites, and elution flow rate at 30 ml/min. Cartridges 2 and 3 were also flushed with 100% methanol following the gradient to elute any remaining metabolites.

Deferoxamines were quantified using a semi-purified deferoxamine fraction (80% by mass), comparing its peak areas to dry weight and applying that to other fractions.

5.2.2 Growth and inhibition of *A. baumannii*

5.2.2.1 Growth curve of *A. baumannii*

A colony of *Acinetobacter baumannii* NCTC 13301 growing on Muller-Hinton agar was used to inoculate a 50 ml Muller-Hinton broth (MHB) overnight culture (37°C, 175 rpm). After overnight culture 1 ml was taken to inoculate 100 ml fresh MHB, which was used to determine lag, log, and stationary phase at OD_{600} every 30 min in triplicate. Measurements were taken using a Spectronic Helios Epsilon spectrophotometer (Thermo Fisher Scientific, Loughborough, UK).

5.2.2.2 Assessing inhibition of *A. baumannii* growth

Bioactivity was tested by adapting the XTT procedure used by Pettit *et al.* (2005). The overnight culture and inoculation of 100 ml MHB was repeated. After 4 h (mid-log phase) a culture aliquot was adjusted to an OD_{600} of 0.05. This was transferred to a 96-well plate (100 µl/well) with an equal amount of stock fraction, except in the control wells (media only, *A. baumannii* culture only, *A. baumannii* and 1 µg/ml colistin). Plates were wrapped in foil and left to incubate for 20 h at 37°C. The sodium salt of XTT (Sigma-Aldrich, Poole UK) was dissolved to 1 mg/ml in PBS and combined in a 12.5:1 ratio with 1 mM Menadione (Sigma-Aldrich) dissolved in acetone. The mixture was added to wells (25 µl/well), shaken gently, and incubated for a further 1 h. The plate was then read at 490 nm, and inhibition calculated by the formula:

$$100 * (1 - \left(\frac{Mean\ well\ absorbance - mean\ MHB\ absorbance}{Mean\ negative\ control\ absorbance - mean\ MHB\ absorbance}\right))$$

Colistin MIC was determined by a separate XTT assay using serial dilutions of 0.125 µg/ml up to 32 µg/ml. The negative control was untreated cells in MHB.

5.2.3 UPLC-MS/MS analysis and metabolite identification

5.2.3.1 UPLC-MS/MS conditions

UPLC-MS conditions were as Chapter 2, section 2.2.6. MS/MS used the Data Directed Acquisition setting in Masslynx V4.1, with the following MS/MS specific

changes compared to previous chapter conditions: *m/z* range was 150 – 2000, MS/MS collision energy ramped from 20 – 40 V, and scan duration and interscan delay was 1 s and 0.025 s respectively.

5.2.3.2 GNPS dereplication

Files were converted to .mzML using MSConvert (Chambers *et al.*, 2012). The Subset filter was selected from the filter menu, and the scanEvent subfilter selected with parameters 1 – 2 to select UPLC-MS and UPLC-MS/MS functions from MS/MS RAW files. The converted files were uploaded to GNPS for dereplication and molecular networking. GNPS parameters for the METABOLOMICS-SNETS-V2 pipeline are shown in Table 5.1.

Table 5.1 GNPS settings used to search M19 culture fractions.

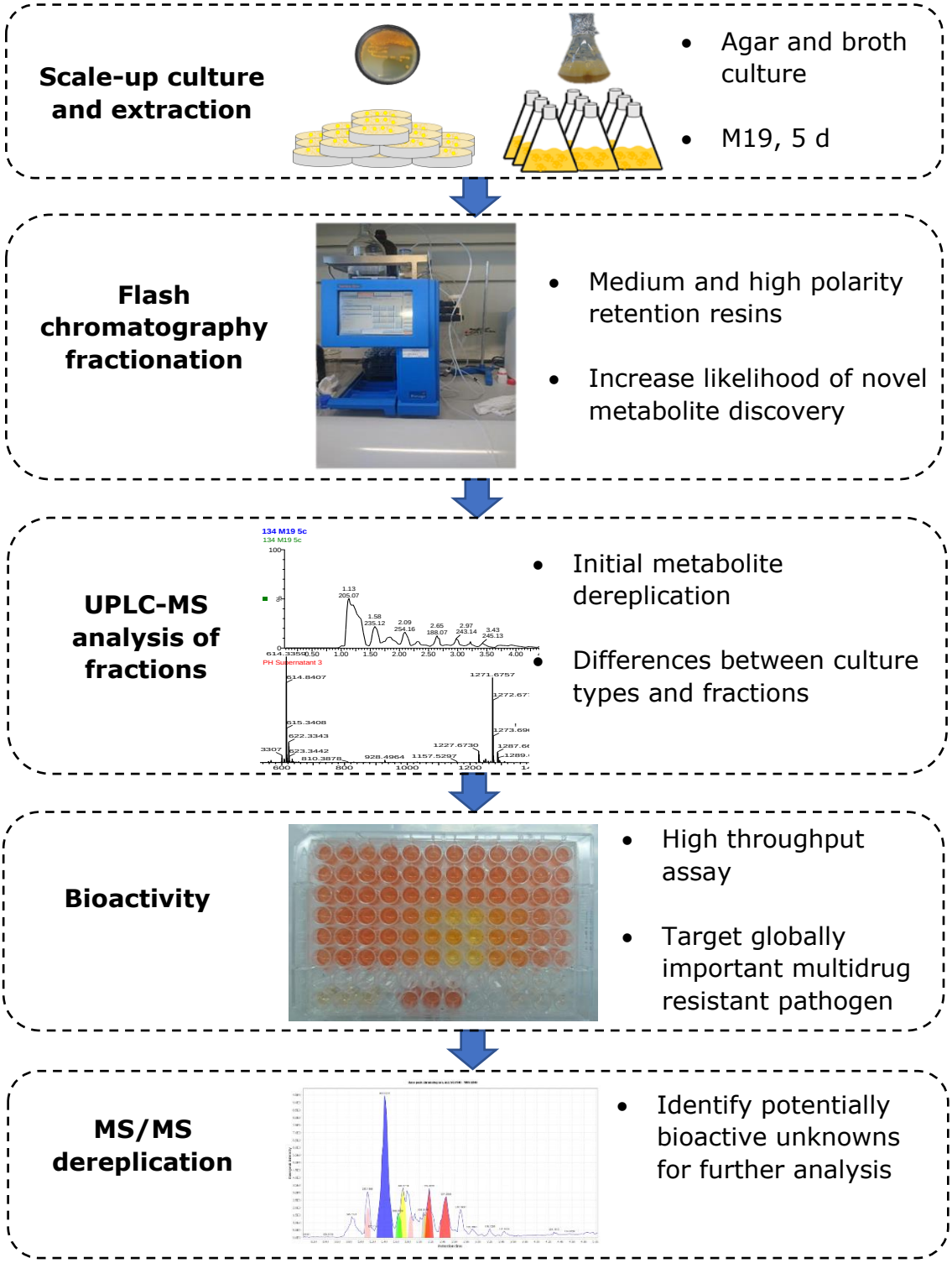
Parameter	Setting
Precursor Ion Mass Tolerance	0.2 Da
Fragment Ion Mass Tolerance	0.2 Da
Minimum Library Peak Match	6
Score Threshold	0.7

GNPS results were refined using MolNetEnhancer v1.2.5 (Ernst *et al.*, 2019) , and the graphml output used to create a molecular network in Cytoscape v3.1.7. The network was filtered to exclude blank compounds, and networks which were comprised of only one node.

5.2.4 Overview of proposed methods

The methods used in this chapter are designed to exploit an area of the *S. costaricanus* chemical space that is less likely to have previously been examined. The methods are summarised in Figure 5.1.

Figure 5.1 Summary of methods for culture scale-up, fraction bioactivity, and metabolite dereplication.



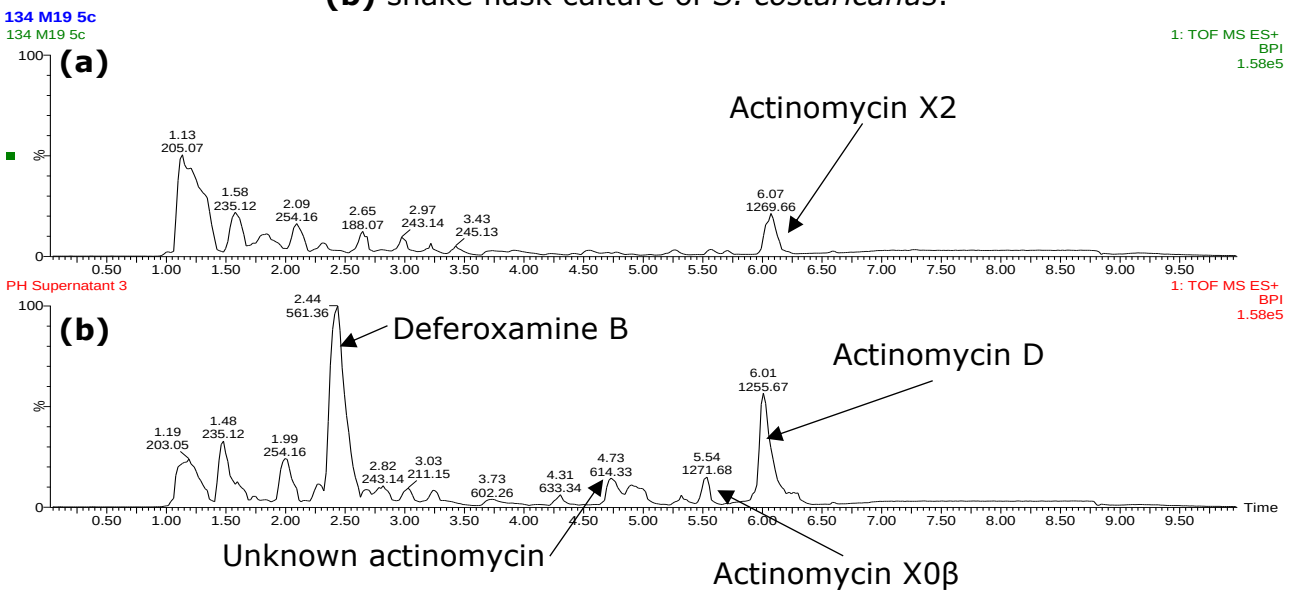
5.3 Results and Discussion

5.3.1 Metabolites produced by *S. costaricanus* in broth culture

5.3.1.1 Comparison of 24-well plate samples and scale-up culture supernatant

The supernatant was analysed by UPLC-MS to give an initial overview of the main metabolites produced by *S. costaricanus* under standard shake flask culture. As in the 24-well plate screen, the main components were deferoxamines and actinomycins, but unlike the 24-well plate screen, the predominant actinomycin was actinomycin D not X2. This more closely matches the actinomycin production profile found by Liu *et al.* (2019), reflecting a difference between the well plates and large shake flask production. A comparison between the two UPLC-MS profiles is shown in Figure 5.2, both of which are representative of triplicate cultures. Only one metabolite is identifiable in the 24-well plate supernatant, actinomycin X2. The scale-up culture supernatant includes other actinomycins D, X0β, and one putative actinomycin with similar spectra but with an *m/z* difference of -28 to actinomycin D.

Figure 5.2 Comparison of UPLC-MS chromatograms from **(a)** 24-well plate and **(b)** shake flask culture of *S. costaricanus*.

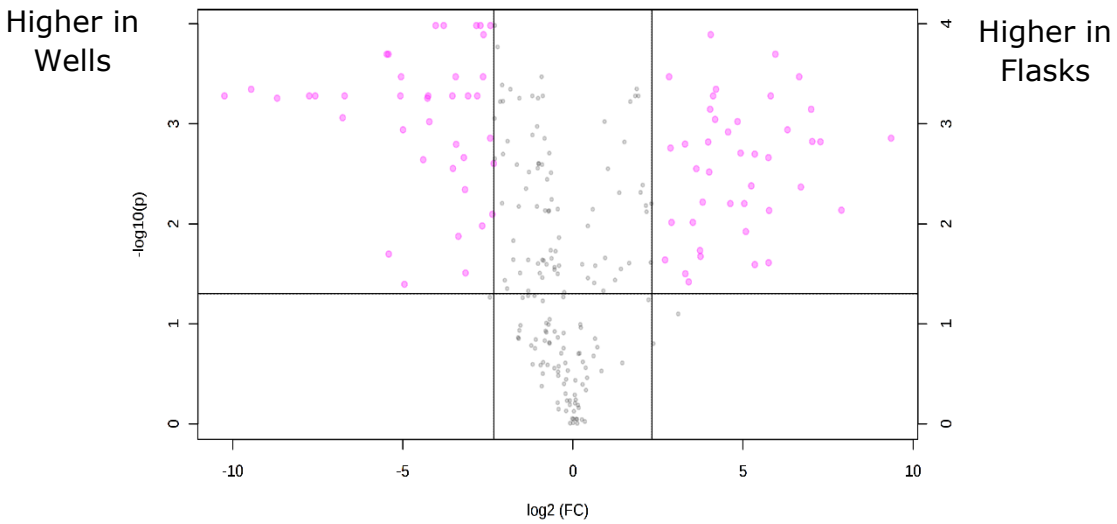


One potential reason for the difference between the two profiles is oxygenation. Although the 2 ml cultures were rotated at the same rpm as the flasks, the vessel shapes and media to total volume ratios will affect growth rates and so metabolite production. This is supported by the lower overall intensity in wells versus plate growth at 8.32 E4 against 1.58 E5. Furthermore, the main feature in the wells is *m/z* 205.07, a media blank compound, so the true well maximum intensity is closer to 3.5 E4. However, the 24-well plate screen was designed to show differences between culture conditions, which it did effectively, while still mostly representing larger scale metabolite production. Miniaturised deep well plates do exist which are able to mimic oxygenation levels in shake flasks (Kosa, Zimmermann, *et al.*, 2018; Kosa, Vuoristo, *et al.*, 2018), so a potential future experiment using this plate type would help identify if oxygenation was one of the reasons for enhanced metabolite production.

MetaboAnalyst 4.0 (Chong *et al.*, 2018) was used to calculate which metabolites changed production by more than 5-fold between the two broth culture vessels. Flask culture resulted in 42 features, and 39 from well culture. The full volcano plot of changes is shown in Figure 5.3. Of the two predominant metabolite

families, deferoxamines shows a higher fold change: 54 times higher in the flask than well culture for deferoxamine B. This may be due to the effect of the increased flask culture biomass increasing nutrient depletion faster than the less productive well culture, more rapidly upregulating scavenging molecules like deferoxamine B.

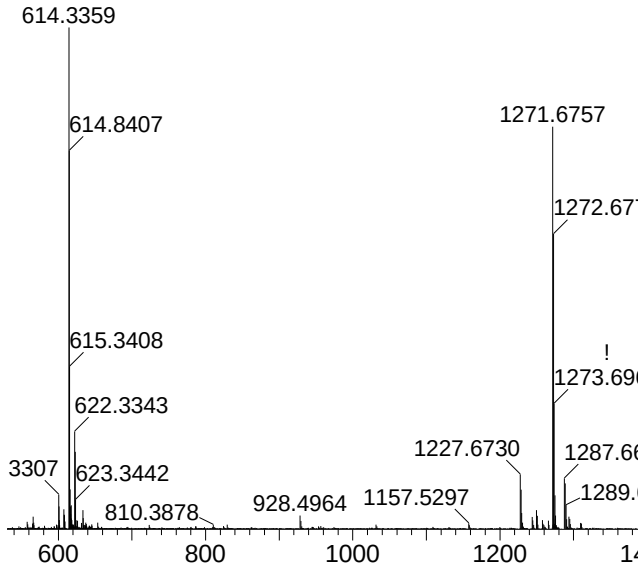
Figure 5.3 Volcano plot of features that differ by over 5-fold between well and flask culture (x axis, pink dots) and were statistically significant (y axis).



5.3.1.2 Comparison of unknown actinomycin to actinomycin D

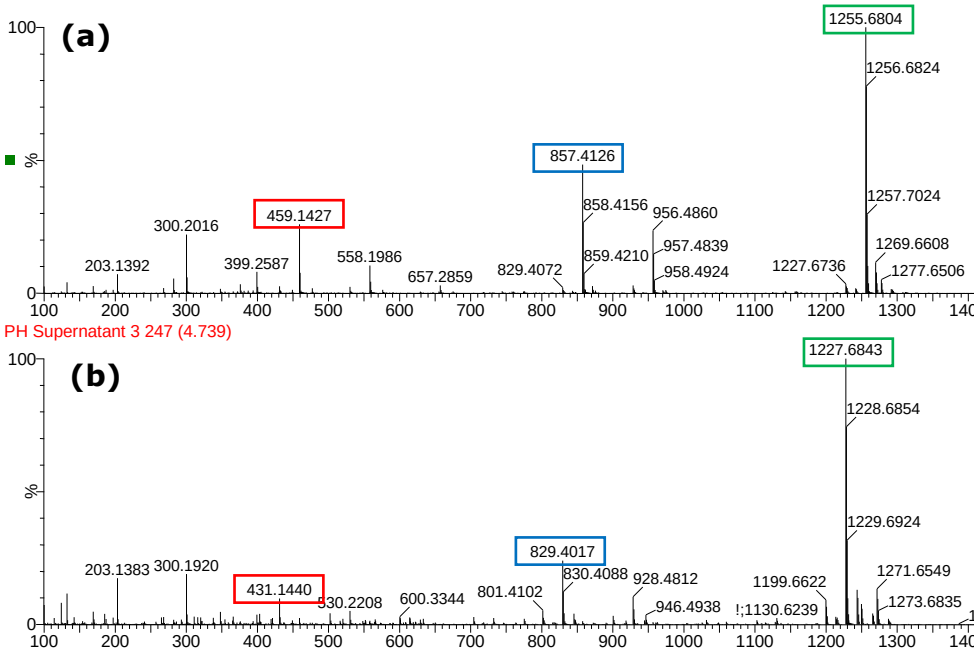
An unknown compound with *m/z* 614.34 $[M+2H]^+$ and *m/z* 1227.68 $[M+H]^+$ was putatively identified as an actinomycin (Figure 5.4). No previously identified exact spectral match for this compound could be found in GNPS or streptomeDB. Its rt – 4.7 min – is earlier than actinomycin D, which elutes at 5.9 – 6 min, indicating elution in a more polar segment of the gradient.

Figure 5.4 Spectra indicating a double charged ion of the putative actinomycin. *m/z* 614.34 has an isotope pattern of +0.5, which along with *m/z* 1227.67, indicates a double and single charged ion respectively. Therefore, these are ions from the same metabolite.



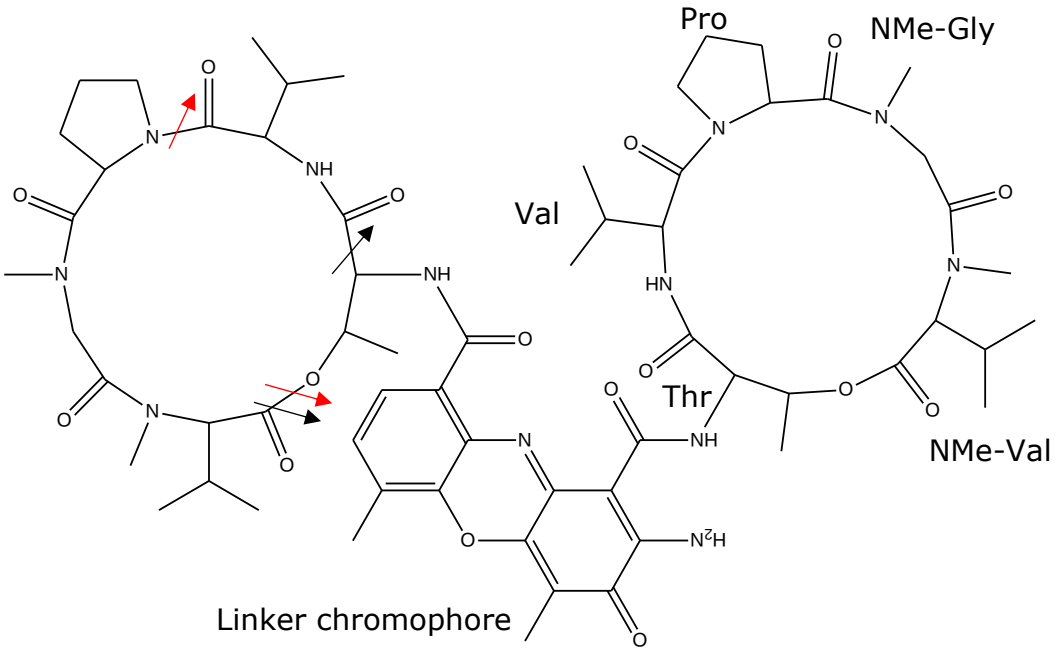
Comparing the higher energy spectra of actinomycin D to the unknown actinomycin shows it contains many of the characteristic actinomycin fragments *m/z* minus 28 daltons (Figure 5.5). This could be due to the loss of C_2H_4, which would increase the overall polarity. Alternative losses could be -CO or an N, but these would not account for the difference in polarity. The higher energy spectra in Figure 5.5 did not produce the double charged ion.

Figure 5.5 High energy spectra of **(a)** actinomycin D, and **(b)** putative actinomycin. Example -28 *m/z* differences from **(a)** to **(b)** outlined.



The main actinomycin D mass fragments have previously been mapped out by Wills & O'Connor (2014). The authors analysed collision activated dissociation and electron induced dissociation-produced spectra for actinomycin D which were used to link mass fragments to structural fragments. This is useful for determining which section of a similar metabolite a structural difference occurs in, as identical sections of metabolite will fragment to identical *m/z*, and non-identical should maintain the *m/z* difference. This was applied to the unknown actinomycin. Actinomycin D consists of a chromophore linker with two identical rings joined to it (Figure 5.6), so any differences between variants may be in one ring or both, which should be observed in the spectra. This figure also shows fragmentation points, based on Wills & O'Connor (2014), for the major fragments 956.5 *m/z* (red arrow; $C_{48}H_{62}N_9O_{12}^+$) and 857.4 *m/z* (black arrow; $C_{48}H_{53}N_8O_{11}^+$), both to 1 decimal place to account for approximately 0.2 *m/z* difference between source article and obtained spectra here.

Figure 5.6 Actinomycin D structure with amino acid building blocks. The rings are identical. The arrows refer to the major fragments 956.5 *m/z* (red arrow; $C_{48}H_{62}N_9O_{12}^+$) and 857.4 *m/z* (black arrow; $C_{48}H_{53}N_8O_{11}^+$)



The minus 28 difference is maintained through most of the major fragments down to *m/z* 300, the mass of the linker chromophore, which indicates that any difference is due to the ring amino acids. Thomas *et al.* (1995) suggest that 1 of 3 amino acids substitutions is responsible for a similar *m/z* difference, such as Nme-Gly to Gly, but without purified metabolite there is no evidence to predict which one.

Another putative actinomycin was detected with an *m/z* of 1257.64 which could not be linked to any known spectra, though it presented the standard spectral fragment *m/z* at both lower and higher fragmentation energies. However, there are many actinomycin variants that have been isolated but do not have spectral information so the novelty of these potential actinomycins is unknown.

5.3.1.3 Composition of supernatant fractions

Separation of the supernatant through the C18 resin produced 24 fractions with a peak UV absorption intensity of approximately 1200 milli-absorbance units (mAU), shown in Figure 5.7, and fraction weights in Figure 5.8.

Figure 5.7 Flash chromatogram of supernatant fractionation using C18 resin.

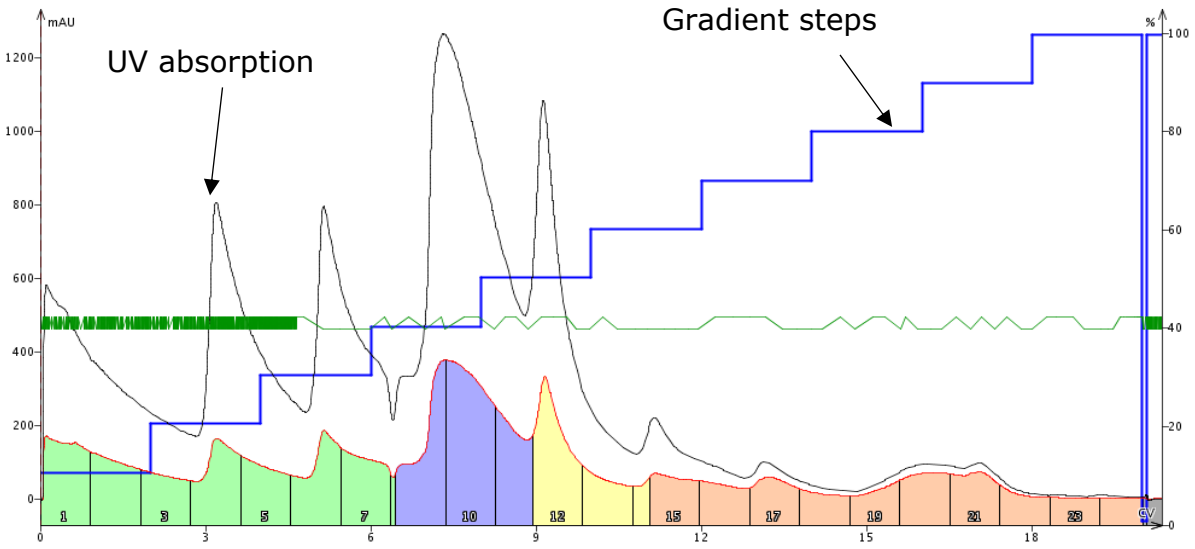
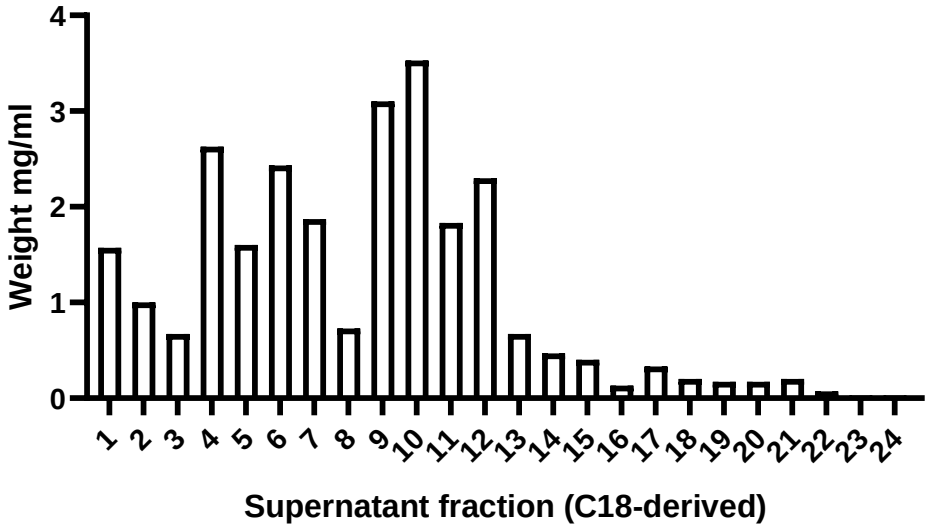


Figure 5.8 C18-derived fractions from *S. costaricanus* broth culture supernatant



5.3.1.4 Purification of deferoxamines

Measuring the fraction weights gave a maximum of 3.53 mg/ml for fraction 10, or approximately 423.6 mg for the full 120 ml fraction. The UPLC-MS chromatogram shows that this is predominantly deferoxamine B, indicating a high level of purity as a result of flash chromatography purification. The peak

area of deferoxamine B in this fraction was approximately 48,000 intensity units of 60,000 total. This equates to around 2.89 mg/ml deferoxamine B in the fraction or 345.94 mg in the full 120 ml. It is also useful for quantifying levels of deferoxamines in other fractions, as if 48,000 units equals 2.89 mg/ml, this can be used to work out amounts in other fractions. A similar calculation was made for MZmine to facilitate faster calculations: corresponding peak area 2920175 = 2.9 mg/ml; deferoxamine peak area/2.9E6 x 2.9 x fraction vol (120) = mg in fraction.

The peak next to deferoxamine B has similar spectra to another siderophore, deferoxamine D1 (*m/z* 603.37 [M+H]). This was logical as they are members of the same metabolite family with similar structures, though D1 does have a slightly longer carbon chain backbone than B making it less polar hence the slightly later rt. Fraction 10 and compound peak areas, spectra and structures of deferoxamine B and D with mass fragmentation points for deferoxamine B (Senges *et al.*, 2018) are displayed in Figure 5.9. There were no deferoxamines detected above 1E3 intensity for ENV+ fractions, giving a total of 1431 mg per 3 L culture or 477 mg per L.

The majority of deferoxamine elutes across 7 fractions, 9 – 16, with relatively small concentrations of deferoxamine D1 (9.48 µg/ml) appearing only in fraction 10. A further minor peak was seen in fraction 10 matching the [M+H] of deferoxamine G1 at *m/z* 619.36, but the spectra was too unclear to confirm a match.

The fraction with the lowest total yield (0.3 mg/ml, fraction 24) has a larger variety of metabolites, with potentially a small family of metabolites under *m/z* 459.38 as the dominant one in this fraction. Across all C18-derived fractions there was a total of 3135 mg material, or approximately 1 L worth of culture to produce 1 g, not accounting for media compound. The ENV+ cartridge produced 23 fractions with a peak of approximately 420 mAU. The lesser mAU of the ENV+ resin-derived fractions was reflected in the fraction weights (Appendix figure 11), with C18 at a total of 26.13 µg/ml and ENV+ at 11.03 µg/ml. *S. parvulus* has been used in the production of deferoxamine E, a similar metabolite to deferoxamine B, producing up to 2 g/l of the compound (Gáll *et al.*, 2016). Optimisation of deferoxamine B is less common in the literature, but has reached

as high as 5 g/l from *S. pilosus* (Chiani *et al.*, 2010). Actinomycin purification is examined in the next section, and amounts of deferoxamines and actinomycins for both C18 and ENV+ are shown in Figure 5.10.

Figure 5.9 **(a)** C18-derived Fraction 10 peak areas, **(b)** deferoxamine B spectra and structure with fragmentation points, **(c)** deferoxamine D1 spectra and structure.

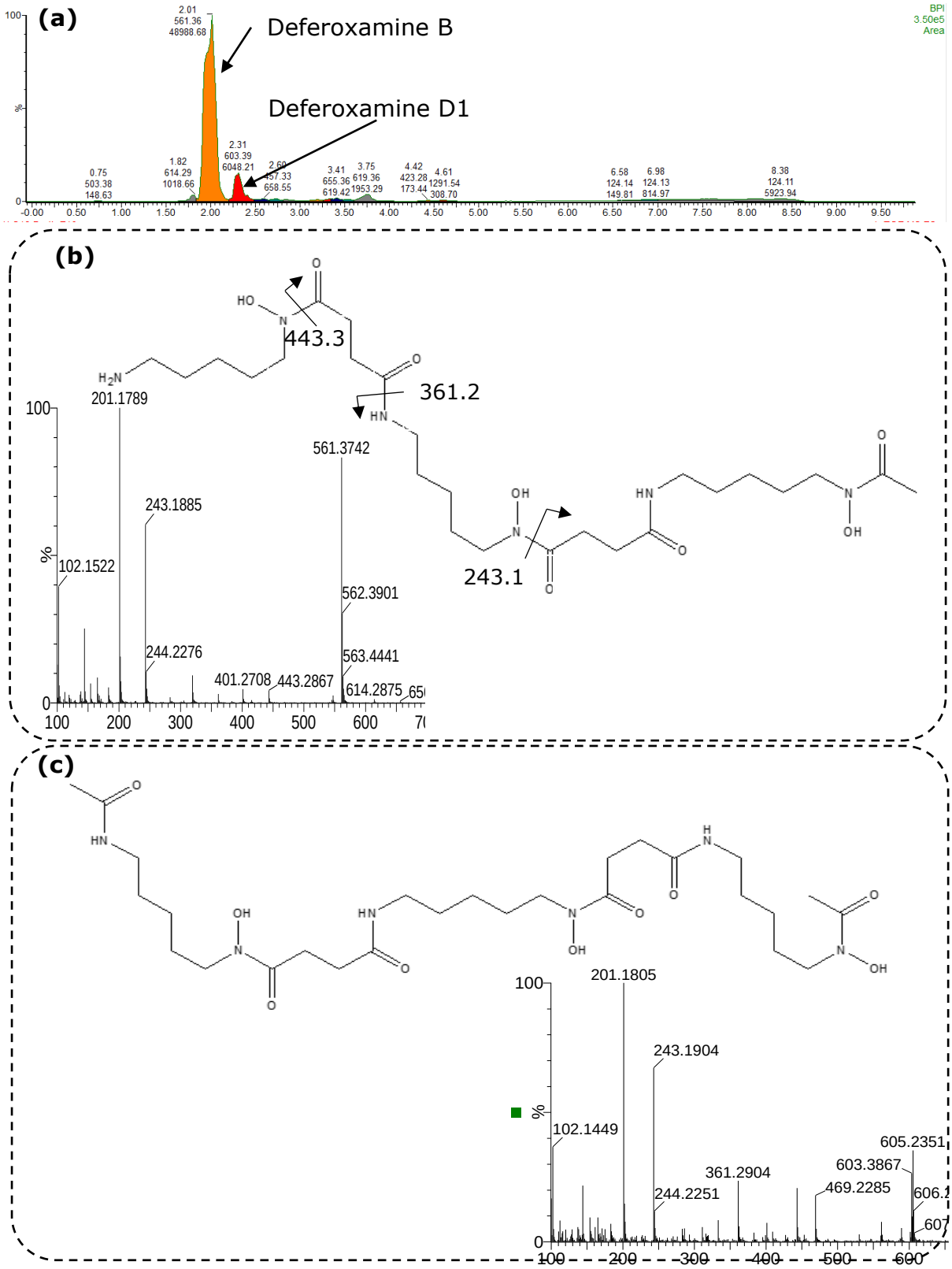
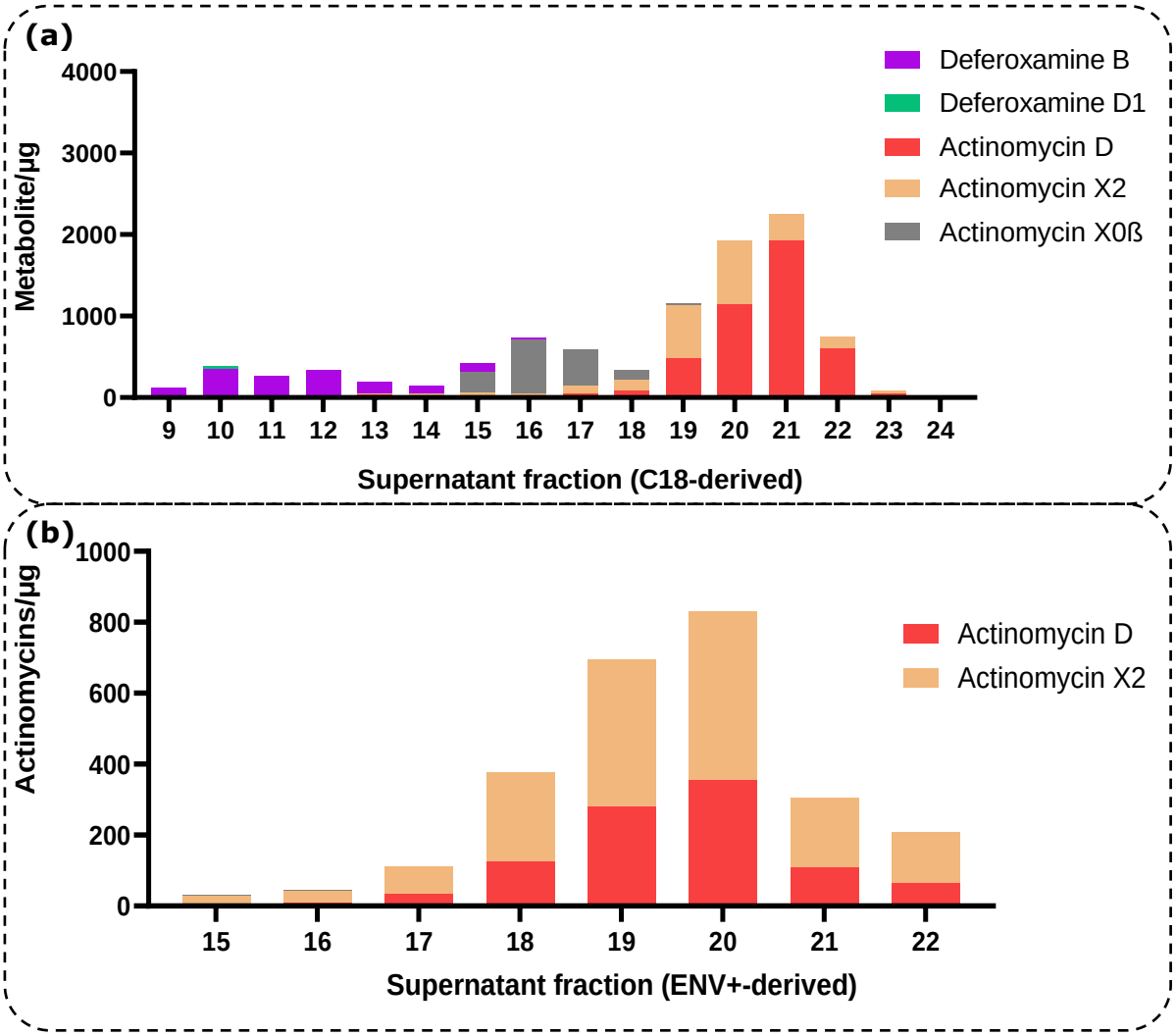


Figure 5.10 Supernatant actinomycin and deferoxamine weights in **(a)** C18-derived fractions, and **(b)** ENV+-derived fractions.



5.3.1.5 Purification of actinomycins

The other previously detected metabolite family, actinomycins, appeared from fraction 12 (C18-derived) onwards, first actinomycin X0β then the other 2 members, actinomycin D and X2, in fraction 17. This correlated to their positions in the supernatant chromatogram, where X0β would elute before D or X2. D and X2 were also present in the ENV+ fractions from 15 onwards though X0β was not, despite the resin's focus on retaining more polar compounds. This was presumably due to the small amount of X0β being captured by C18 even if it would also have been retained by ENV+. The highest concentration of actinomycin D was in C18-derived fraction 21, at 16.07 µg/ml, translating to

1.93 mg for the whole fraction, or 8% of the 24 mg total from the full 120 ml. Across all the supernatant C18-derived fractions there was a total of 4.45 mg actinomycin D, 2.28 mg actinomycin X2, and 1.50 mg actinomycin X0β.

Unlike in the C18-derived fractions, X2 is the dominant family member in ENV+ fractions though the overall amount of product is less. Actinomycin X0β was barely detected in the ENV+ fractions at a high of 2.4 µg/ml for a total of 8.4 µg. Actinomycin X2 and D were each found at a total of 1.64 mg and 988.80 µg respectively.

The produced yield is well below historical production of actinomycin D from *Streptomyces*, as 600 µg/ml were achieved from *S. parvulus* by Williams & Katz (1977). This species is preferred as it does not produce analogues in relatively large amounts, as *S. costaricanus* does; however, other researchers later working with supposedly the same strain of *S. parvulus* only managed to elicit a maximum of 80 µg/ml (Dalili & Chau, 1988). More recent efforts have used a mix of other species, such as *S. triostinicus*, with algorithmically aided medium optimisation to produce up to 452 µg/ml actinomycin D (Singh *et al.*, 2009) from broth culture.

Actinomycin is a relatively easy metabolite to collect since it is secreted into the media with little left in the pellet, which is discarded after centrifuging. Unfortunately, after centrifuging Singh *et al.* (2009) do not give clear instructions on any purification or crude extraction of actinomycin. In comparison Praveen & Tripathi (2009) describe their purification of actinomycin using 1 L of *S. griseoruber* broth culture. After centrifuging, the supernatant was passed through a silica gel column and eluted with ethyl acetate. The eluate was vacuum-concentrated to give 240 mg crude extract, which was then fractionated using LH-20 resin for a final amount of 210 mg actinomycin D per litre of culture.

As mentioned in section Scale-up of S. costaricanus 5.1.1.4, actinomycin production from a separate strain of *S. costaricanus* has recently been examined. Liu *et al.* (2019) used a different extraction method to the ones already described, adding each of their 50 ml cultures to 100 ml butanone and mixing for 30 min. The butanone phase was evaporated to dryness and the residue analysed by HPLC and UPLC-MS, but no further preparative chemistry was done.

One main difference between the methods used above and in this work is the use of solvents prior to fractionation, essentially adding a liquid-liquid extraction step prior to solid phase extraction. If a metabolite has already been identified for collection, and is effectively partitioned by a solvent, then liquid-liquid extraction is a sensible step to optimise yield and negate compound loss caused by successive processing steps. However, for an untargeted project, adding a liquid-liquid or resin prior to fractionation will reduce the accessible chemical space.

5.3.1.6 Commercial value of identified metabolites

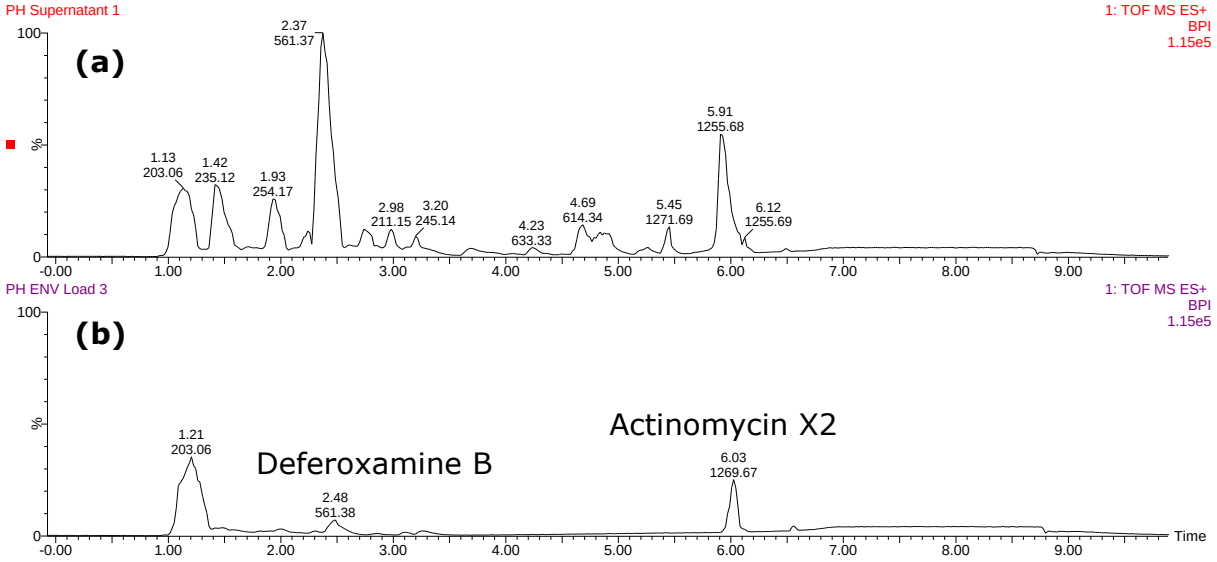
Both actinomycin D and X2 are commercially available, though D far more frequently than X2 likely due to its usage as an antitumour drug. At current prices, the components of M19 broth cost £126.00 (Fisher, Peptone) and £34.10 (Sigma, Mannitol) for 500 g. At 20 g/l each, the 3 L of M19 used cost £15.12 + £4.09, £19.21, to produce 5.40 mg actinomycin D. Sigma sells 5 mg 98% purity actinomycin D for £121 (product ID A1410), so at this rough estimate it would technically be profitable to make actinomycin D with unoptimised medium, not accounting for labour, equipment, or other costs. Actinomycin X2 is less widely available, potentially due to it not being medically used. It is sold by AdipoGen (product ID BVT-0089-M005) at £40 for 5 mg, far less cost effective than actinomycin D. Deferoxamine B would also not be very profitable, if at all after additional costs: Sigma sells 1 g for £79.20, and 1.4 g is produced from 3 L. As with the actinomycins this is unoptimised culture.

5.3.1.7 Compounds unretained by C18 captured by ENV+

Initially, ENV+ was included to capture highly polar compounds which were not retained on C18 and so would normally be lost to waste. Unfortunately, the majority of the higher weight ENV+ fractions were comprised of media components, which are always evident at the start of the chromatogram. This higher polarity would fit with C18 not retaining them and their presence in ENV+ eluent. This also made detecting novel metabolites more difficult, since any features found in ENV+ fractions may be media components and their derivatives.

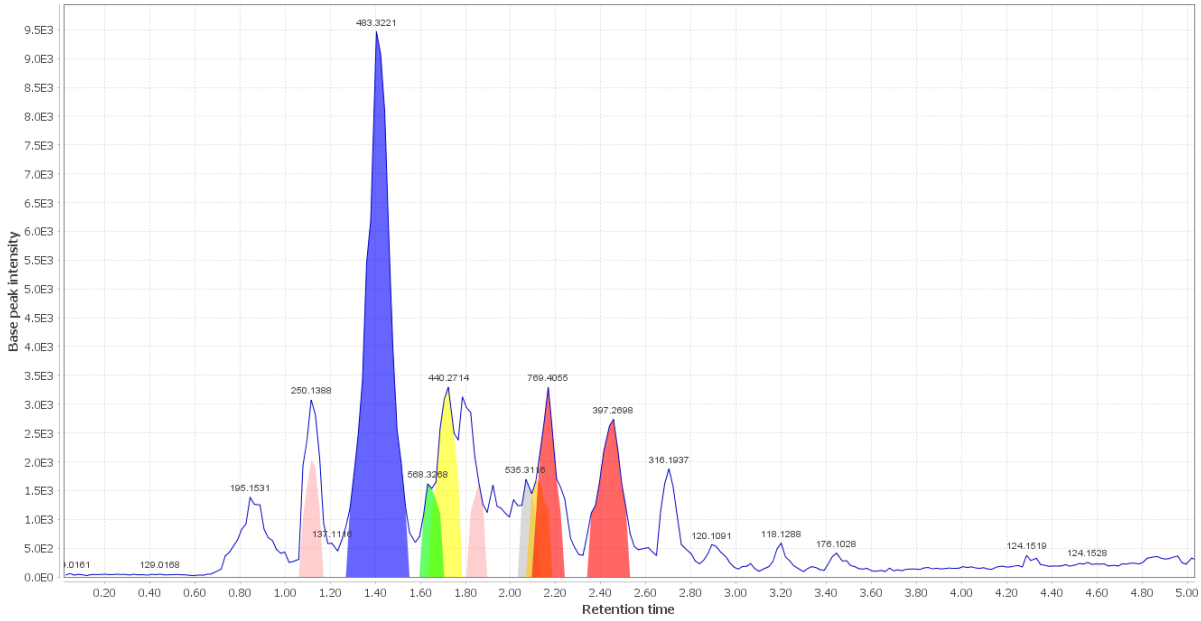
Figure 5.11 shows a comparison of the supernatant to ENV+ load, both of which are representative of triplicate UPLC-MS runs. There were 2 identified metabolites in the ENV+ load, deferoxamine B and actinomycin X2. No deferoxamines were found in the ENV+ fractions despite being loaded on, indicating it did not adsorb to the resin.

Figure 5.11 UPLC-MS spectra comparison of **(a)** *S. costaricanus* broth culture supernatant, and **(b)** Compounds unretained by C18 used as ENV+ load



Using MZmine to compare the total features found in both C18 and ENV+ fractions found 34 features unique to ENV, 210 features unique to C18, and 23 shared features. ENV fraction 1 contained multiple unique features in clearly defined peaks: of the 12 total features in fraction 1, 9 were included in the unique list (Figure 5.12). That Fraction 1 was mostly potentially unique polar compounds was logical, since Fraction 1 was eluted with the highest percentage of the more polar solvent. Manually searching the GNPS database for these masses did not find any similar spectral patterns.

Figure 5.12 MZmine UPLC-MS chromatogram of ENV+-derived fraction 1 features not found in C18-derived fractions.

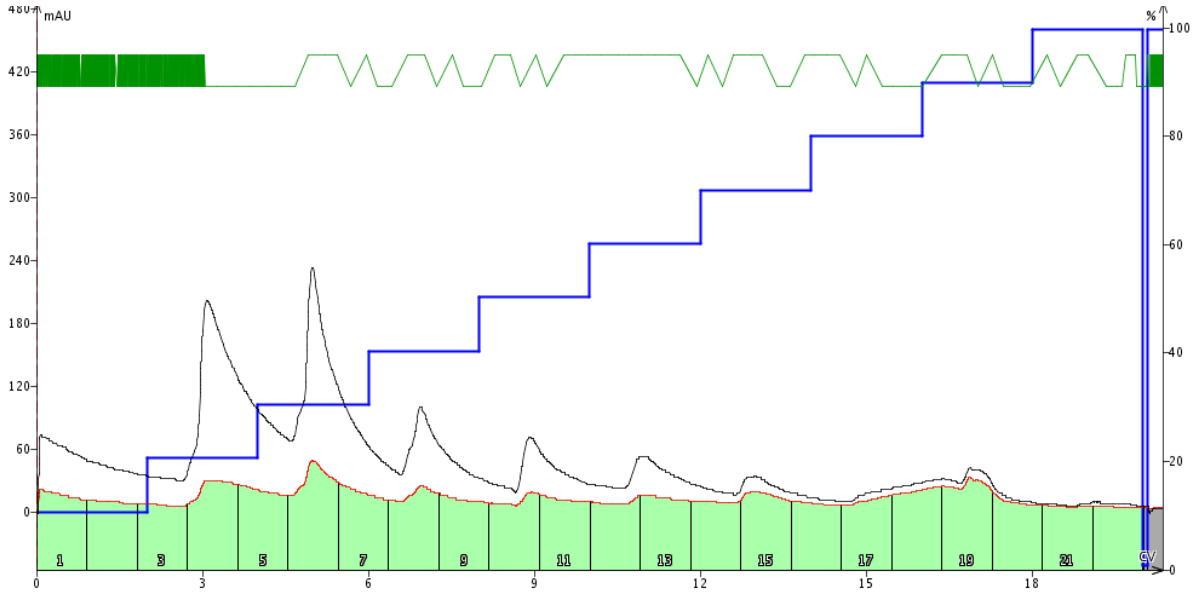


However, given the low fraction weights (0.93 µg/ml for ENV+ fraction 1) any further characterisation would require substantial optimisation and extra scale-up assuming that they are novel metabolites.

5.3.1.8 Flash chromatography of biomass crude extract

Metabolites were extracted from *S. costaricanus* biomass to determine if any could be detected that were not found in the supernatant. The flash chromatography mAU was low at a maximum of 240 with little absorbance across the chromatogram (Figure 5.13). There were no unique features seen in the UPLC-MS spectra, so focus remained on the supernatant and agar extracts.

Figure 5.13 Flash chromatogram of pellet extract from C18-derived fractions.



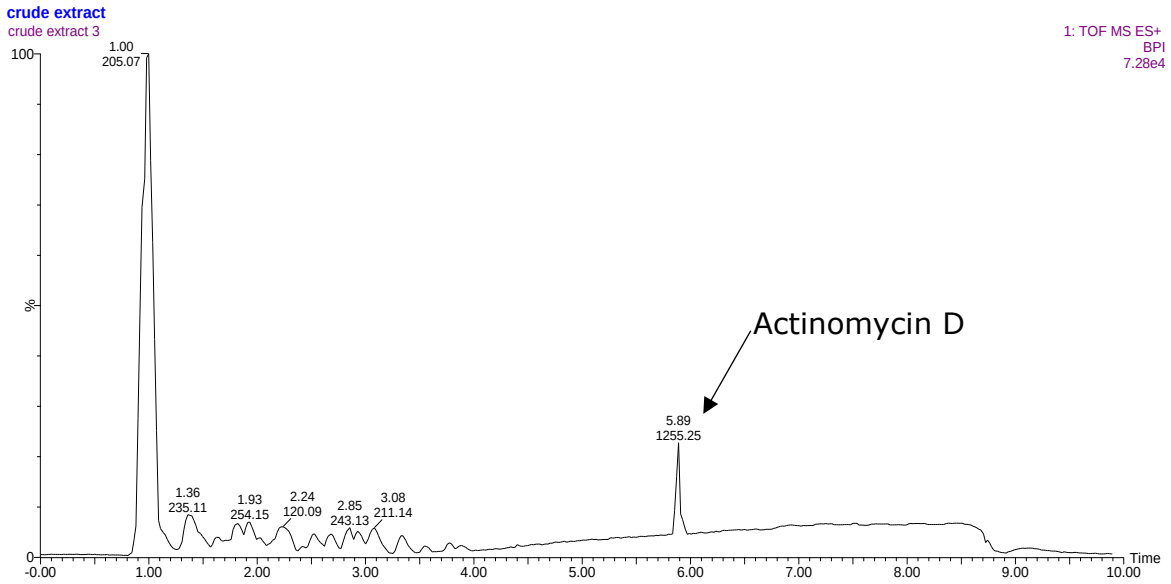
5.3.2 Metabolites produced by *S. costaricanus* on agar culture

5.3.2.1 Agar culture extract UPLC-MS profile

UPLC-MS analysis of the crude extract (Figure 5.14) showed an overall lower intensity with fewer obvious metabolites than the supernatant. While this was not surprising, given the differences in nutrient availability and mixing between solid media and broth in a shake flask, this was reflected in only 1 identifiable metabolite being immediately visible (actinomycin D). Magnifying the chromatogram to exclude the major media peak and actinomycin enhanced the minor peaks, but none of them were identifiable metabolites.

The major feature (*m/z* 205.07, rt 1.00) matched the media blank, indicating that nutrients were still available, which would be expected after only 5 d culture in rich media.

Figure 5.14 UPLC-MS chromatogram crude extract from *S. costaricanus* growth on M19 agar.



5.3.2.2 Metabolite analysis of fractions from agar culture

Eluting the C18 and ENV+ loaded cartridges produced 22 fractions from each cartridge; the flash chromatogram is shown in Figure 5.15, with a mAU of approximately 800. Similar to the crude extract profile, the fraction weights for agar culture were lower than the broth culture supernatant. Also like the supernatant, the C18-derived fractions overall contained more weight than the ENV+ fractions: 14.97 mg for C18 versus 13.73 mg for ENV+. Examining the chromatograms for the ENV+ higher weight fractions, the dominant peaks appear to be media components, with relatively small amounts of actinomycins. ENV+ fraction weights are shown in Appendix figure 12 .

Figure 5.15 Flash chromatogram of agar culture extract fractionated using C18 resin.

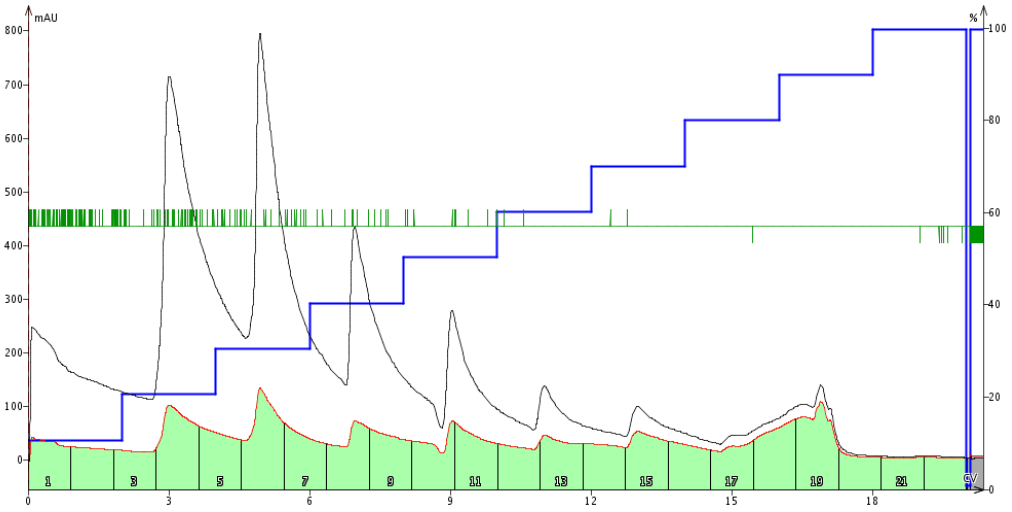
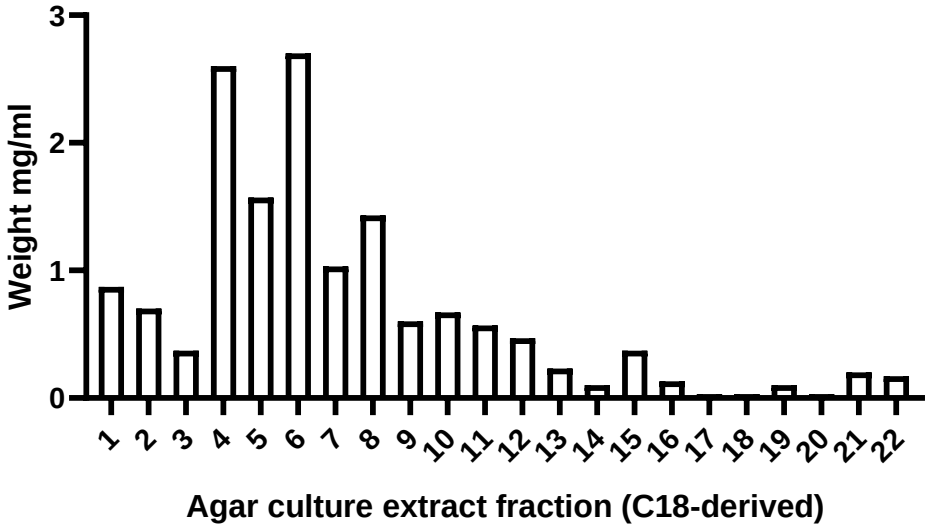


Figure 5.16 Agar culture crude extract C18-derived fraction weights.



The identified actinomycins – D, X2 and X0β – were mostly divided between fractions 20 and 21. C18-derived fraction 21 contained the highest concentrations of actinomycin D from agar culture, but also had a low weight at 0.2 mg/ml (Figure 5.16). Quantification of actinomycin D through comparison to a 10 µg/ml actinomycin D standard came to 15.39 µg/ml, or slightly less than the 16 µg/ml supernatant fraction 21 equivalent. Actinomycin produced by *S. costaricanus* grown on agar reached 9.06 µg/ml for X2, and X0β 9.01 µg/ml or 108.68 and 108.16 µg respectively.

Deferoxamine production was reduced when grown on agar, to the point where very little – if any – could be observed above the limit of detection in UPLC-MS chromatograms. Processing the data with MZmine showed a feature with a similar, but not exact, mass and rt: *m/z* 561.74, with a maximum height of 2E3 in ENV+ fraction 5. Taking the actinomycin and deferoxamine results together in comparison to broth culture, M19 agar is not an effective culture condition for these 2 specialised metabolites.

5.3.3 Metabolite purification through polarity-based fraction combination

5.3.3.1 Fraction combination

Multiple rounds of fractionation can be used in natural products analysis, often following identification of a bioactive fraction. In another metabolomics approach to drug discovery, Kellogg *et al.* (2016) sub-fractionated a bioactive *Alternaria* sp. culture fraction and used the sub-fraction data to identify the bioactive component contributing to partial least squares separation, though this resulted in the identification of an already known metabolite. A separate method focusing on MS/MS spectral networks to predict bioactivity identified one fraction from extracts of the tree *Euphorbia dendroides*, which was further fractionated to give 4 new compounds (Nothias *et al.*, 2018). Fractions can also be combined together to increase concentrations of a desired metabolite in the resulting sub-fractions. This was shown in Erenler *et al.* (2017), who divided 400 *Origanum*

rotundifolium derived fractions into groups with similar polarities to isolate specific metabolites from each group.

Combining fractions has the added advantage of allowing a fraction that was otherwise too low in compound weight to be useful, especially as material is lost at each processing stage. Fraction weights varied, with multiple sequentially eluted fractions under 0.5 mg/ml. This could have potentially split a bioactive metabolite into several fractions. Unidentified features were also shared across agar and broth culture over both C18 and ENV+ derived fractions. Therefore, as no target bioactive metabolite had been identified at this stage, to increase the concentration of identical or same family metabolites within a fraction, it was decided that the fractions should be combined and re-fractionated. This was done by separating the fractions into 1 of 3 groups, based on the most dominant metabolite family in each fraction, acting as a proxy for overall fraction polarity. The families were actinomycins, deferoxamines, or the remaining unknown most polar fractions. Specific fractions that make up each pool are shown in Table 5.2, using high polarity pool (HP) for the remainder fractions, medium polarity pool (MP) for deferoxamine fractions, and low polarity pool (LP) for actinomycin fractions.

Table 5.2 Groupings for fraction combination and re-fractionating.

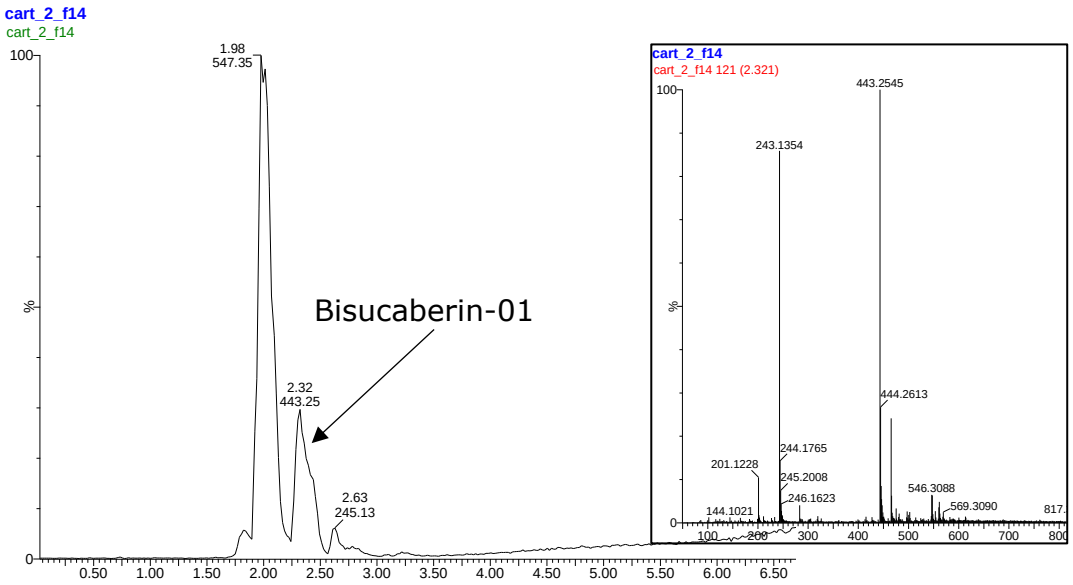
Fraction type	**Metabolite pool**		
	LP	**MP**	**HP**
Supernatant C18	18 - 24	9 – 17	1 – 8
Supernatant ENV+	15 – 22	9 – 14	1 – 8
Agar C18	17 – 22	12 – 16	1 – 11
Agar ENV+	16 - 22	9 – 15	1 – 8

After pooling and purifying, the LP separation resulted in 24 fractions, MP 28, and the HP 24, of which the HP was the heaviest. Fraction weights are listed in Appendix table 4.

The fractions with the highest weights for each pool were expected to contain the majority of the pool-deciding metabolite for actinomycins and deferoxamines, so

20 – 22 for LP and 12 – 18 for MP. In addition to deferoxamines B and D1, deferoxamine A1 and Bisucaberin-01 – previously reported by Wibberg *et al.* (2018) – were also identified in the HP (fractions 13 – 14; deferoxamine A1 *m/z* 547.34 [M+H]; Bisucaberin-01 *m/z* 443.25 [M+H], Figure 5.17). While an already known metabolite, A1 has an identical elution time to deferoxamine B at 1.9 min. The 100% methanol eluted fractions had low intensities and potentially weights (signal intensity <5E3) with no additional novel compound so were set aside to only be used in metabolite dereplication.

Figure 5.17 Manual identification of bisucaberin-01 after purification.

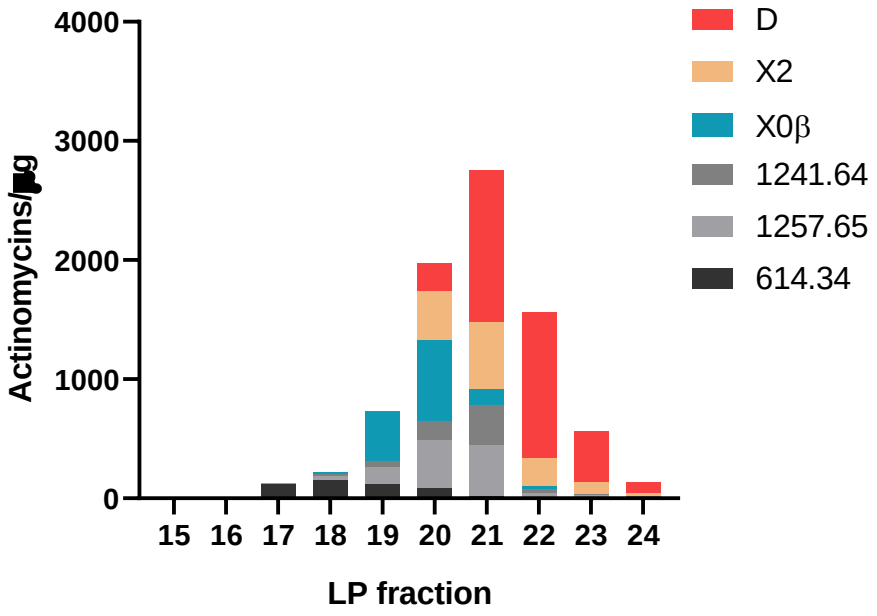


Therefore, though by no means as efficient as MS/MS based strategies, the 2nd round of fractionation did contribute to metabolite analysis and dereplication. The other fractions in this range were overwhelmingly deferoxamine B (peak area approximately 13000 of 14000), which at 93% is an efficient level of purification given the untargeted nature of the purification. As no further metabolites could be manually identified the actinomycin purification levels in the LP were investigated.

5.3.3.2 Purification of actinomycins across LP fractions

Actinomycins were spread across fractions 17 – 24 corresponding to their originally seen rt: the first eluted metabolite was the putative double charged actinomycin (*m/z* 614.34 $[M+2H]^{2+}$) in fraction 17, followed by the other less polar metabolites with actinomycins X2 and D least polar. Figure 5.18 shows the amount of the detected actinomycins over relevant pool fractions.

Figure 5.18 Amount of 3 identified actinomycins, and the next 3 most concentrated putative actinomycins.



5.3.3.3 Purification enhanced discovery of an actinomycin analogue

These fractions contained 6 actinomycins which were quantified (3 known – D, X2, X0β – and 3 unidentified). Partially purifying the actinomycins allowed more robust spectral identification, since co-eluting members of the same family will share many mass fragments and so confuse characterisation. Chromatograms and mass spectra for each of the actinomycins plus actinomycin D are shown from Appendix figure 20 to Appendix figure 25, with diode array absorption showing similar absorption wavelengths.

Actinomycins are a large family of metabolites with many structurally similar members. This makes mass spectral differentiation of known and unknown actinomycins members challenging as there will be many shared fragments between them, so without a fully purified sample precise spectral differences may be unclear. Of the 6 actinomycins here, 1 minor metabolite (the analogue *m/z* 1241.61) was found following the semi-purification step, as it was otherwise hidden by the more concentrated actinomycin D due to the similar retention time.

This metabolite shares many, if not all, of the actinomycin D fragments identified by Wills & O'Connor (2014). However, many of these also have a fragment at *m/z* minus 14. The presence of the "reference" fragment and a separate -14 indicated that the difference from actinomycin D is on one ring rather than both. The chemical difference could be a $-CH_2$ loss, which could explain the slightly earlier, more polar rt compared to actinomycin D.

This difference was maintained for all major fragments down to *m/z* 300, which according to Wills & O'Connor (2014) cleaves the chromophore-containing central linker plus Thr and Val (plus an H_2O; $C_{14}H_{26}N_3O_4^+$) from the remainder of the molecule. Therefore, the difference is in the remaining ring amino acids: Me-Gly, Pro, and NMe-Val. Thomas *et al.* (1995) found a similar situation when dealing with impurities of actinomycin D. The authors suggest a range of potential amino acid substitutions, with either NMe-Val, NMe-Gly, or Val being substituted. Of those, only NMe-Gly matches the remaining amino acids given above. Therefore, the difference is a substitution in NMe-Gly, potentially to Gly. This would account for the mass difference. Thomas *et al.* also note that another actinomycin family member, actinomycin D0, varies from D exactly by a NMe-Gly to Gly change. This gives a potential identification for 1 of the 3 unknown actinomycins.

5.3.4 *A. baumannii* growth and inhibition

The supernatant and agar crude extract fractions were assayed against *Acinetobacter baumannii*, a Gram-negative nosocomial pathogen listed as one of the multi-drug resistant microbes most needing new antibiotics for treatment by

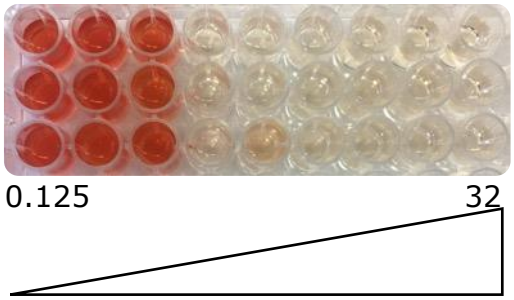
the World Health Organisation (Tacconelli *et al.*, 2018). The strain used was a β-lactam/carbapenem resistant hospital isolate (NCTC 13301), matching the World Health Organisation list.

5.3.4.1 XTT Assay Optimisation

The bioactivity test selected was a colourmetric assay using the tetrazolium salt 2,3-bis {2-methoxy-4-nitro-5-[(sulfenylamino) carbonyl]-2H-tetrazolium-hydroxide}, commonly known as the XTT assay (Pettit *et al.*, 2005). The XTT assay is a cell viability assay which indirectly measures respiration through the properties of tetrazolium salt reduction. Upon reduction with an electron donor – in this case NADPH – a water soluble formazan pink/red dye is formed, distinct from the clear-coloured XTT (Berridge *et al.*, 2005). The amount of reduction, and indirectly respiration, is indicated by the colour intensity: a darker colour implies more respiration, and vice versa. The assay can therefore be used to measure the effect of fractions or metabolites on cell viability.

Measuring the growth of *A. baumannii* in shake flask culture showed mid-log phase to occur at approximately 4 h. Mid-log culture was used to optimise the dose of the positive control, colistin, to determine the minimum inhibitory concentration (MIC). Growth was inhibited above 1 µg/ml (Figure 5.19), which had an OD_{490} of approximately 0.17 against 0.2 for an MHB blank and did not decrease at higher colistin concentrations. The Clinical Laboratory Standards Institute recommends a dose of ≤2 µg/ml for susceptible *A. baumannii* , and ≥ 4 µg/ml for resistant strains (CLSI, 2018). This mirrors Papp-Wallace *et al.* (2018) who also found an MIC of 1 µg/ml for colistin against the same strain of *A. baumannii*. Therefore, 1 µg/ml colistin was chosen as the positive control.

Figure 5.19 XTT assay plate with 2-fold dilutions of colistin, starting at 0.125 µg/ml on the left and ending at 32 µg/ml on the right.

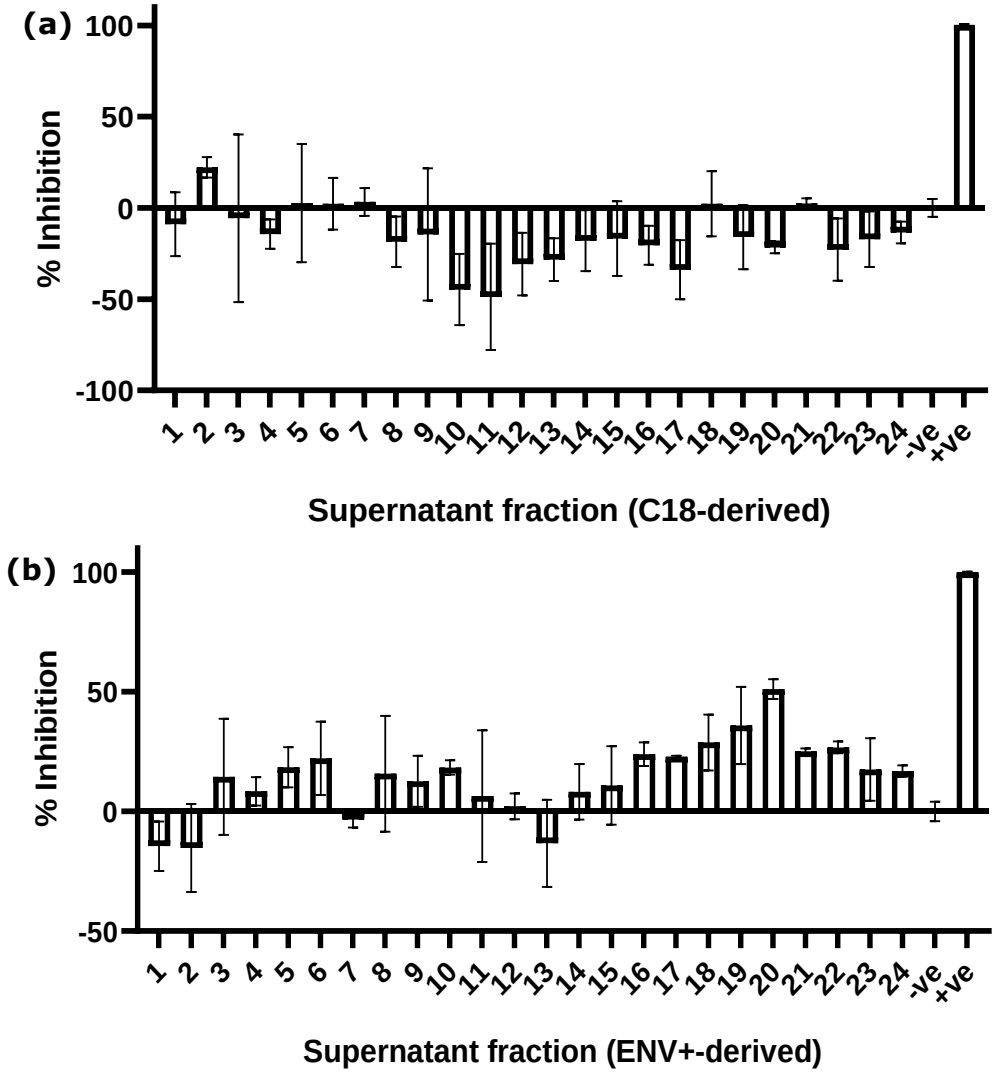


5.3.4.2 Fraction Bioactivity

Fractions eluted from the C18 and ENV+ resins were tested against *A. baumannii* at 50 µg/ml through the XTT assay to determine if any of the fractions led to a negative change in cell viability and so inhibition of growth (Figure 5.20 for supernatant fractions, Figure 5.21 for agar extract fractions). The majority of supernatant fractions provoked an increase in viability, potentially due to unused M19 nutrients. The largest increase in viability was found in supernatant C18-derived fractions 10 and 11 at 45% and 49%% respectively. Both of these fractions are almost entirely siderophores – either deferoxamine B or D1 – so no inhibition was expected and would explain the increase in viability.

Only 1 of the supernatant C18-derived fractions produced >10% inhibition: fraction 2 at 22%. Fraction 1 increased viability by 9%, with a statistically significant difference between the fraction 1 and fraction 2 inhibition levels. Therefore, any differences between the fractions visible in their UPLC-MS chromatograms could have been responsible for bioactivity. Unfortunately, no significant differences could be seen between the 2 fractions.

Figure 5.20 Inhibition of *A. baumannii* by metabolites in culture supernatant fractions at 50 µg/ml, split into **(a)** C18-derived, and **(b)** ENV+-derived. Negative inhibition indicates growth greater than the negative control.



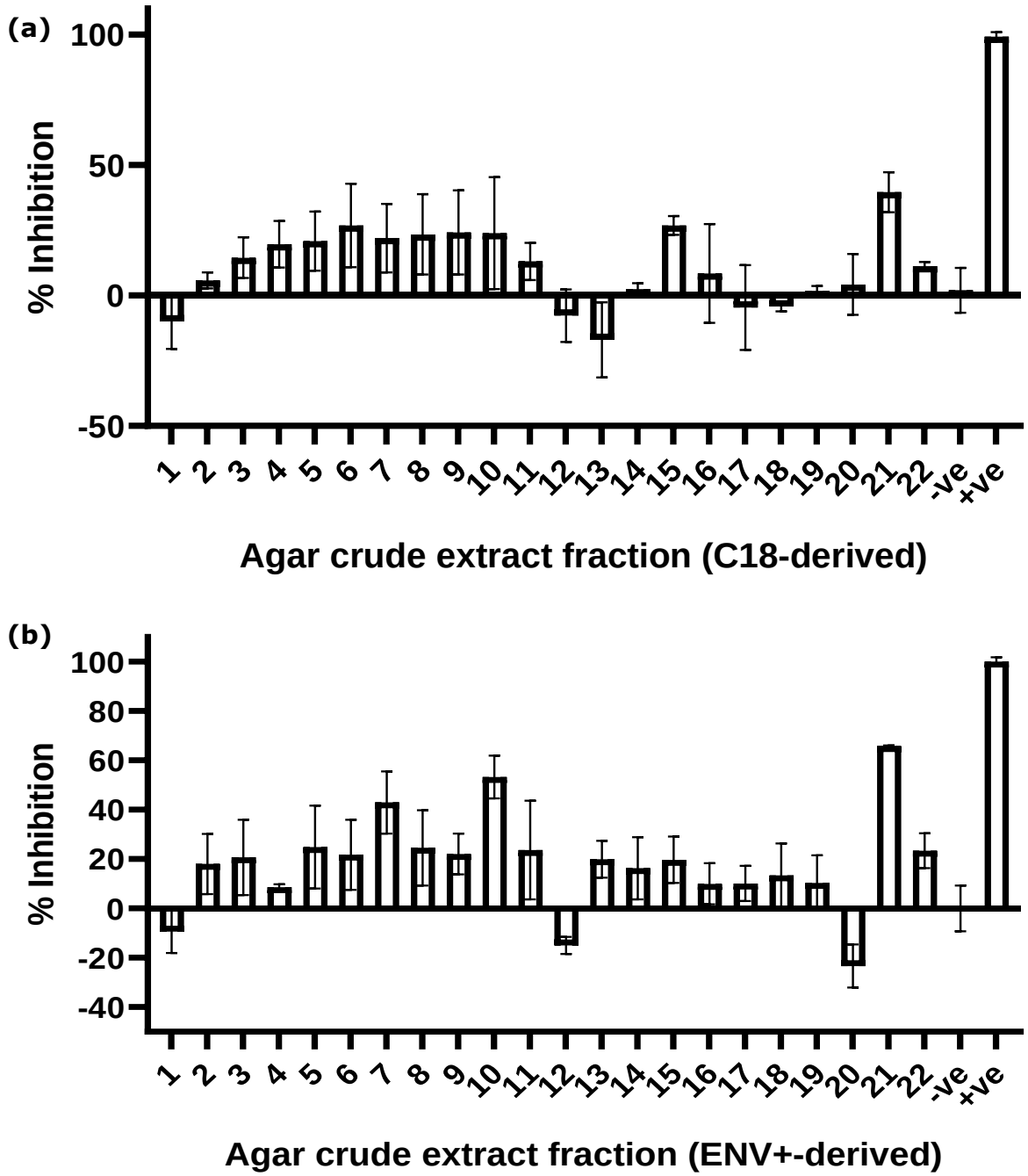
The supernatant ENV+ fractions were more inhibitory. Fraction 6 resulted in 22.58% inhibition, but again there was no obvious difference between UPLC-MS chromatograms of fraction 6 and the growth-promoting fraction 7, indicating that not all compounds within these fractions are ionising. After 6, there was a run of >20% inhibitory fractions from 16 – 22, peaking at fraction 20 (52%). The major peak in fraction 20 is *m/z* 1269.66, corresponding to actinomycin X2. Actinomycin D is not active against *A. baumannii* and other Gram-negatives at this concentration (Luna *et al.*, 2017), and was not seen to be in the assay at this concentration of 10 µg/ml. There is limited information on the Gram-negative antimicrobial activity of actinomycin X2. It was found to be active

against *E. coli* at 0.73 µg/ml (Xiong *et al.*, 2012), but no specific *E. coli* strain information was given by the authors so *E. coli* used may not have been multi-drug resistant. Xiong *et al.* (2012) also did not test against *A. baumannii*.

Whether actinomycin X2 is active against resistant Gram-negatives or not, it may not be the bioactive component in ENV+ fraction 20. The most inhibitory fraction from agar culture was the ENV+ separation from 21, which produced 66% inhibition (Figure 5.21). This fraction contained an actinomycin X2 peak, but at approximately half the peak area: 2297.66 (supernatant) against 1377.29 (agar), or approximately 2 against 1 µg/ml. With less peak area but higher bioactivity, actinomycin X2 is unlikely to be the major bioactive component.

The largest feature in the agar ENV+ fraction 20 is *m/z* 245.16, which is a match for an isoleucine/leucine dimer, but otherwise could not be identified. As with the other bioactive fractions, an unambiguous bioactive metabolite candidate could not be identified. This could have been due to the effect of a minor inhibitory bioactive metabolite competing against growth-promoting compounds, as the fractions are a complex mix of different compounds. Therefore, groups of fractions were pooled for further purification.

Figure 5.21 Inhibition of *A. baumannii* by *S. costaricanus* agar culture metabolites, split into **(a)** C18-derived, and **(b)** ENV+-derived. Negative inhibition indicates growth greater than the positive control.



5.3.5 Pooled Fraction Bioactivity

Fraction pooling combined similar potentially bioactive metabolites by their polarities, with the aim of increasing their concentrations while separating non-active or growth promoting compounds. The LP contained the primary bioactive fractions, so was re-assayed against *A. baumannii* with the XTT assay at 50 µg/ml. Results for the assay are shown in Figure 5.22. The highest inhibition came from fractions 22 and 20 (48% and 35% respectively) and fractions 8 to 9 (37% and 32% respectively). Fractions 7 and 8 were predominantly leupeptins (Figure 5.23). As a protease inhibitor (Ding *et al.*, 2002) leupeptin is unlikely to be the antibiotic component.

Figure 5.22 LP fraction pool activity against *A. baumannii* using a concentration of 50 µg/ml.

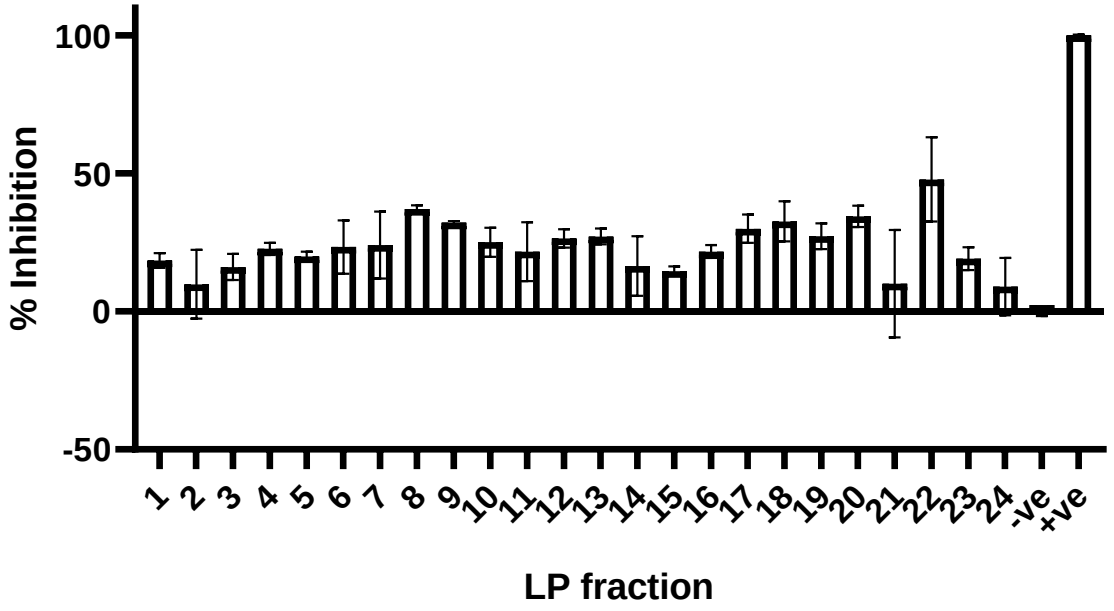
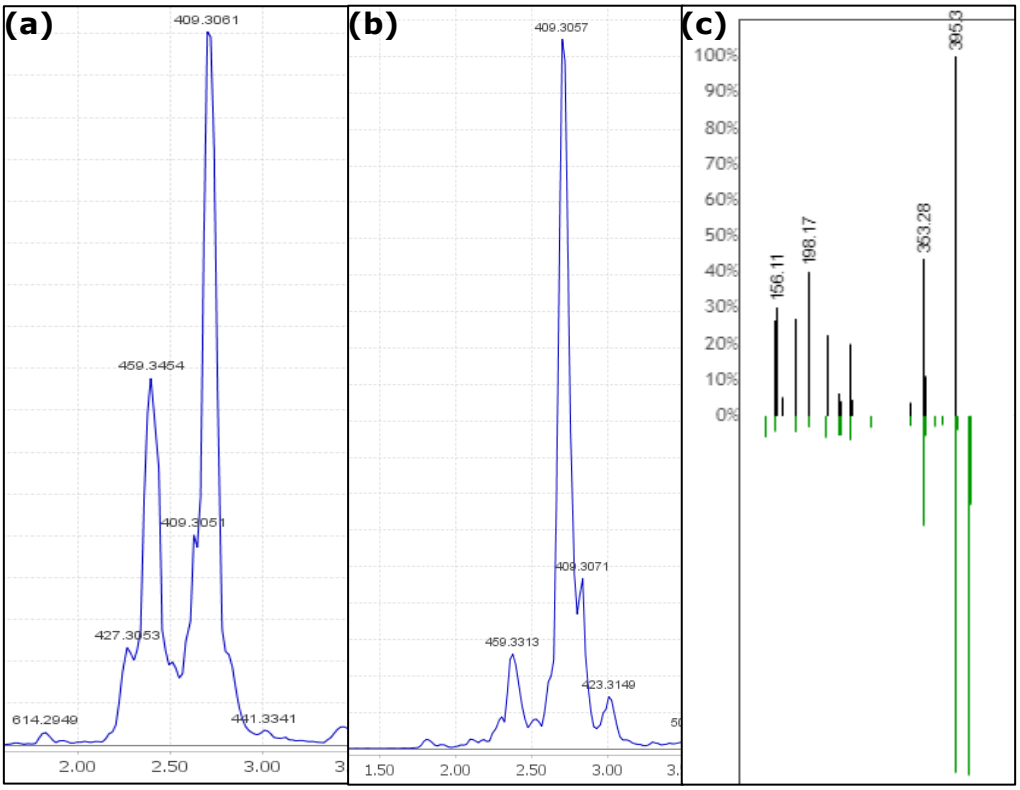


Figure 5.23 Pooled fractions **(a)** 8 and **(b)** 9 containing leupeptin and analogues, with **(c)** GNPS core matched set of fragments.



Despite being sequentially eluted between 20 and 22, fraction 21 showed dissimilarly low inhibition (10 %). Like the other similar elution fractions, 21 is predominantly actinomycins (D, *m/z* 1257.65 and 1271.69). This showed a similar pattern to the original set of fractions where inhibitory actinomycin containing fractions were present along with non-inhibitory actinomycin containing fractions. This was shown by fraction 21, which additionally contains actinomycin X2, but has an inhibition level of only 10%. This suggests that the actinomycins were not responsible for bioactivity, especially actinomycin D as this is the major compound in both fractions 21 and 22 at similar levels. To confirm this actinomycin D was tested using the XTT and disc diffusion assay, both of which showed no inhibition of *A. baumannii*.

There were minor metabolite differences between fractions 21 and 22, making bioactivity difficult to assign to any particular compound. Bioactivity could also have been due to the synergistic interactions of multiple metabolites, for example an emulsifier to facilitate movement of an antibiotic through the

bacterial membrane (Zhang *et al.*, 2017), which could also apply to bioactivity in any of the other fractions. There were variations in levels of some unknown compounds, such as *m/z* 783.49 rt 5.35 (peak areas 22 versus 21: 1672.84 versus 199.74, respectively). For this compound to have been responsible with no synergy required, any other fractions containing it should have had respective levels of bioactivity, assuming no growth promoting compounds are present. Fraction 23 inhibited *A. baumannii* growth by 19 % and contained this compound with a peak area of 606, approximately 3 times that of fraction 21. While the increase in inhibition from fraction 21 to 23 was not totally proportional, there was a relevant bioactivity increase from fraction 21 and decrease from 22 approximately in line with the peak area levels of *m/z* 783.49. Therefore, this may be the metabolite responsible for bioactivity, but further purification would be required to prove this.

Cytoscape was used to screen for other metabolites in bioactive fractions by deleting all nodes that were not found in LP fraction 22. The majority of the remaining nodes were actinomycin-related parent ions and fragments. Excluding these left only non-actinomycin fraction 22 MS/MS features, one of which was the previously mentioned *m/z* 1492.76. Another potential bioactive feature was *m/z* 628.31, which appeared in fractions 22 and 8 (fraction 8, 37 % inhibition). However, the signal levels of this feature were low with peak areas of 17 and 4 for fractions 8 and 22 respectively. Therefore, either this feature is a highly bioactive metabolite or, more likely, a feature that is coincidentally shared between the two fractions, especially as they eluted at different ends of the gradient. As stated above, further purification is needed to determine the bioactive metabolite given the inhibition differences between actinomycin-containing fractions.

The chromatograms for the above discussed pool fractions (LP, fractions 8, 9, 21, 22, 23) are shown in Appendix figure 14 to Appendix figure 18.

5.3.6 UPLC-MS/MS Metabolite Identification

5.3.6.1 GNPS Identification

Each of the fractions from the 3 pools (LP, MP, HP) was analysed using the Masslynx Data Directed Acquisition UPLC-MS/MS tool, which were then processed with MZmine and submitted to the Global Natural Products Social Molecular Networking (GNPS) server for metabolite identification. In total 13 matches were found, which are listed in Table 5.3. Multiple compounds may appear more than once due to separate library entry matches, as is the case with deferoxamine B.

Table 5.3 GNPS-detected compounds found from the 2nd round of fractions

GNPS Compound	GNPS *m/z*	Pool	Molecular formula
Actinomycin D	1255.64	LP	$C_{62}H_{86}N_{12}O_{16}$
Actinomycin X2	1269.63	LP	$C_{62}H_{84}N_{12}O_{17}$
Deferoxamine B	561.36	MP	$C_{25}H_{48}N_6O_8$
Desf-05	575.38	MP	$C_{26}H_{51}N_6O_8$
Deferoxamine A1	547.35	MP	$C_{24}H_{47}N_6O_8$
Deferoxamine B [M+Al-2H]	585.32	MP	$C_{25}H_{46}N_6O_8Al$
Deferoxamine B [M+Fe-2H]	614.27	MP	$C_{25}H_{46}N_6O_8Fe$
Leupeptin analogue LVR	413.29	LP	$C_{19}H_{36}N_6O_4$
Maltotriose	527.16	MP	$C_{18}H_{32}O_{16}$
Methiazole	266.10	LP	$C_{12}H_{15}N_3O_2S$
"Unknown peptide"	1025.57	HP	-
Polysaccharide hexose	527.16	HP	-
Triphenyl phosphate	327.08	MP	$C_{18}H_{15}O_4P$

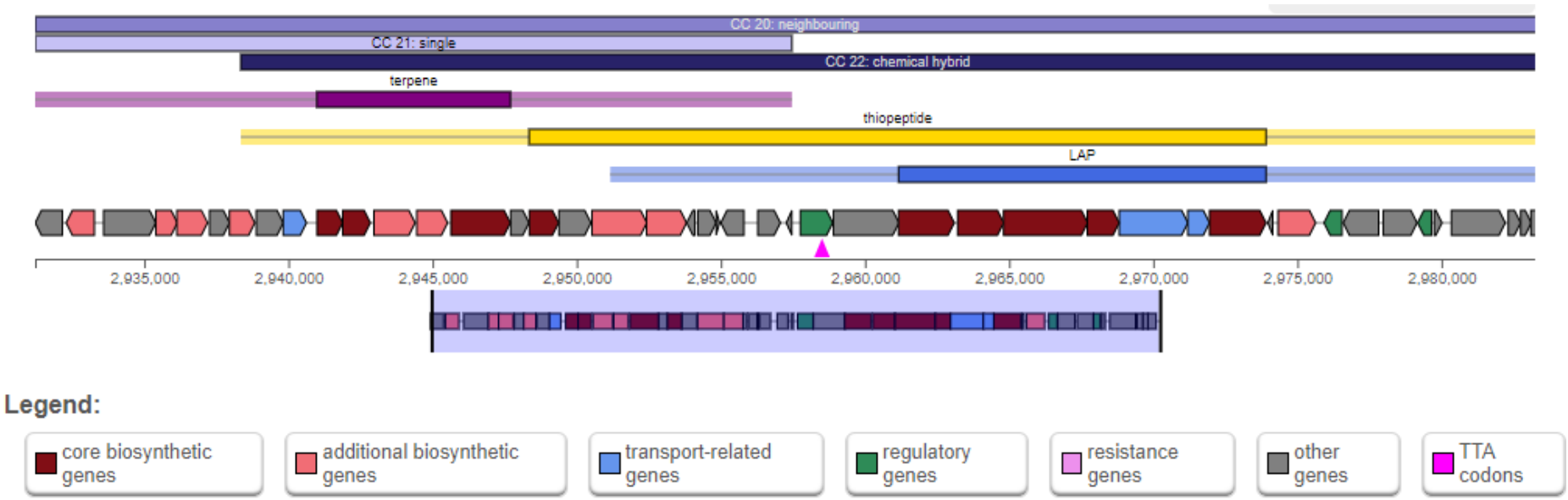
One notable compound not previously detected was methiazole (not methimazole, a treatment for hyperthyroid). Methiazole is an antihelminth compound (Reuter *et al.*, 2006). Earlier papers on *S. costaricanus* described it as being isolated from "nematode-suppressive soil", indicating some level of antihelminth activity (Esnard *et al.*, 1995) but no responsible metabolite.

Methiazole is a benzimidazole derivative (Shimomura *et al.*, 2019), along with many other synthetic compounds not naturally produced by *Streptomyces* or other microbes. The spectrum for methiazole in the GNPS library is taken from the National Institute of Health's pharmacologically active small molecule library, which contains synthetic compounds. It is likely that methiazole was developed as a result of or along with other azole-drugs in the 1970s (Vicente *et al.*, 2003); azoles are further listed as synthetic drugs by Vengurlekar *et al.* (2012).

Therefore, it cannot be stated that methiazole is the responsible metabolite for the originally detected bioactivity of *S. costaricanus*. However, it may be that a similar azole-based metabolite is produced, since non-human synthesised azole natural products do exist with the azole moiety itself being responsible for bioactivity (Jamison *et al.*, 2019). The antiSMASH results do contain a Linear azol(in)e-containing peptide BGC, though it is part of a larger terpene-thiopeptide-azoline hybrid product with no definitive identity.

So, while methiazole is likely not produced by *S. costaricanus*, a compound with similar MS/MS fragmentation is produced by it which closely resembles an azoline product. The cluster is reproduced below in Figure 5.24, which includes a terpene, thiopeptide, and the LAP BGC. The terpene, while potentially part of a hybrid BGC, is predicted to produce hopene, a widely occurring metabolite. It is also designated as "neighbouring" relative to the overlapping "chemical hybrid" thiopeptide and LAP BGCs, further indicating that it may be a wholly separate BGC.

Figure 5.24 AntiSMASH-derived BGC 11 of *S. costaricanus* comprising terpene, thiopeptide, and LAP clusters.



5.3.6.2 Cytoscape View of Pooled Fractions

Cytoscape (Shannon *et al.*, 2003) is a bioinformatics tool which allows for the analysis and comparison of large scale biological data from a variety of sources. The GNPS Molecular Networking workflow produces a cytoscape-compatible file which can be used to visualise spectral networks. Cytoscape groups metabolites by MS/MS spectra similarity following the GNPS output, linking families of metabolites together alongside user-specified cosmetic or feature specific information. This can be useful for identifying novel metabolite analogues, production differences between culture conditions, or to aid bioactive compound identification with additional *in silico* predictions (Nothias *et al.*, 2018).

As MS/MS and GNPS analysis was done post-fraction pooling, here the network was used to examine groupings of known compounds, if there were any unseen family members, and any other potential networks indicating metabolites.

The overall network is shown in Figure 5.25. Individual nodes are coloured by their polarity group: LP red, MP blue, and HP grey. If a feature is found in multiple groups, the pie chart colouring will change to be proportionally coloured based on the number of fractions in each of those groups. For example, if a feature is found in 5 LP fractions and 1 MP fraction, the pie chart will be 5:1 red:blue. This is less likely to happen for the major metabolites which were used to decide polarity groupings but could potentially be to non-archetype family members – so not actinomycin D or deferoxamine B – which were split during pooling. The primary networks for these 2 families are shown in Figure 5.26. Separation is more likely to happen for unknown metabolite families which may be split across polarity groups, but as features are grouped by spectra, they will remain linked in cytoscape.

Figure 5.25 Cytoscape network of pooled fractions. Red – LP, Blue – MP, Grey – HP.

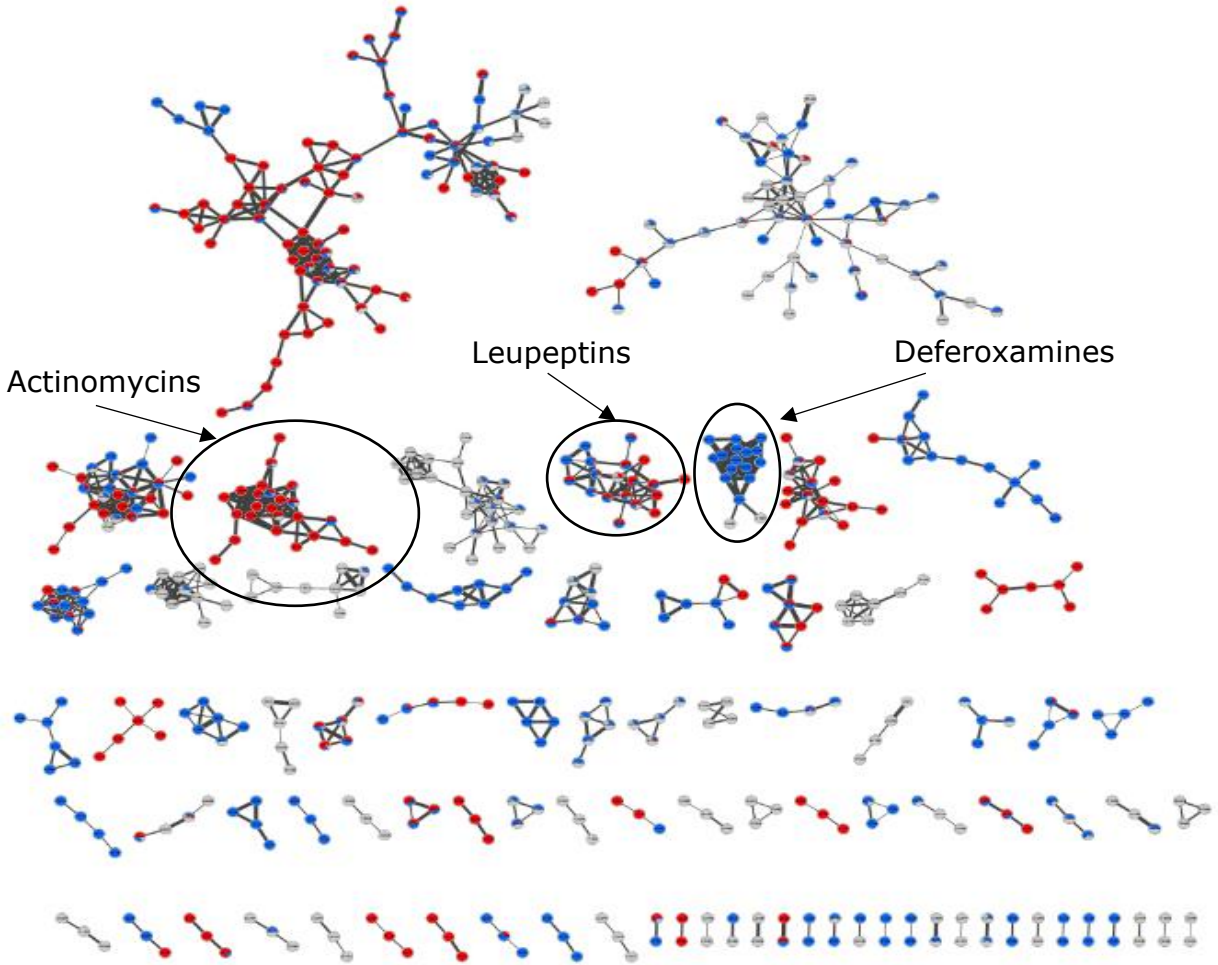
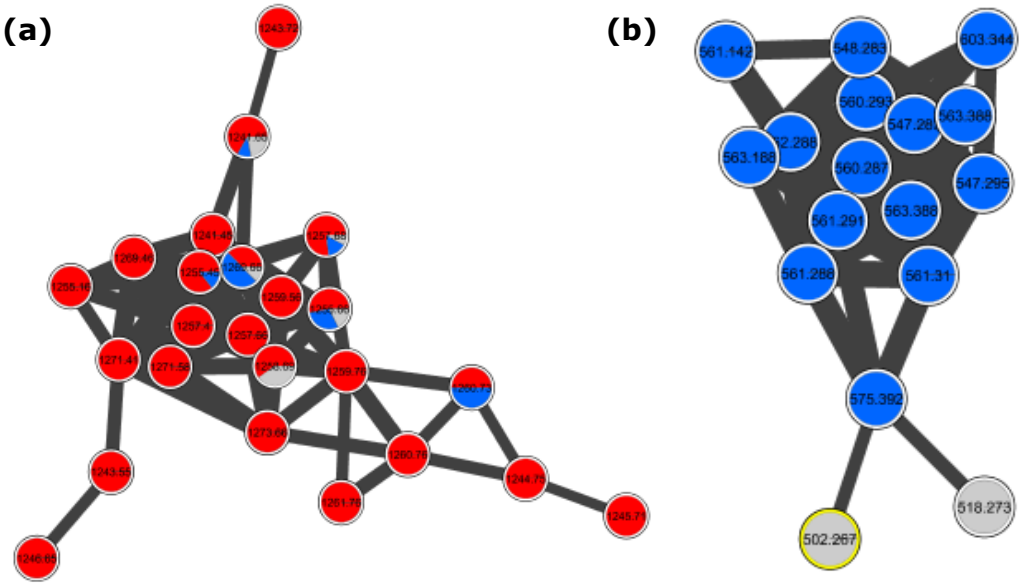
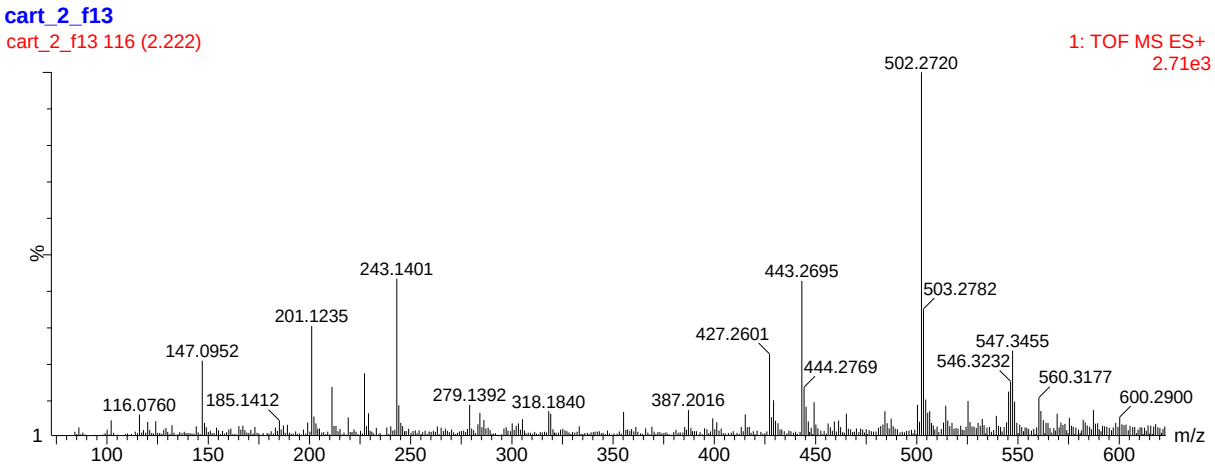


Figure 5.26 Networks for **(a)** actinomycins, and **(b)** deferoxamines.



The deferoxamine node *m/z* 502.267 was found in HP fractions 23 – 28, but there is also an identical *m/z* found in MP fractions 13 – 16 which likely represent the original gradient and 100% methanol eluted fractions. The MP spectra were used for analysis due to their higher intensity.

Figure 5.27 *m/z* 502.27 deferoxamine network offshoot spectra.



The ion was not immediately visible on the chromatogram due to its proximity to the major deferoxamine A1 peak approximately 10 times larger (6E2 against 6E3) but it does become the most intense ion for 6 scans. As an offshoot of the main deferoxamine branch, if assigned correctly then there should be shared spectra between this feature and other deferoxamines. The MS are shown in Figure 5.27; this was chosen over the MSe or associated MS/MS as the low signal intensity gives a visually noisy spectrum. Some deferoxamine characteristic fragments are present, such as *m/z* 443.27 and 243.14. Given its proximity to deferoxamine A1, the parent ion could potentially be a partially decayed remnant of A1 with the loss of $C_2H_5NH_2$ from the end of its structure. A similar fragment, *m/z* 502.33, is present within the A1 peak, but at a lower intensity than the later 502.27. The GNPS spectra for deferoxamine A1 does not show the potential parent ion, overall indicating it is a separate metabolite found by examining the molecular network.

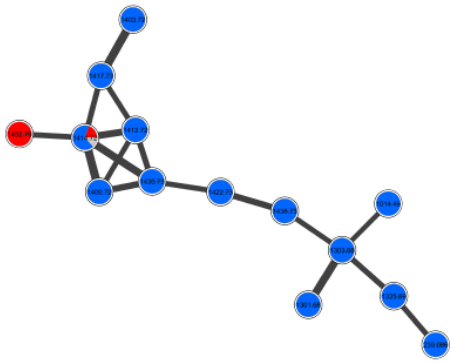
One potential novel metabolite unrelated to the known families is shown in Figure 5.28. The suggested parent ion (*m/z* 1492.73) appears in LP fraction 19, which is dominated by actinomycins. A double charged peak also appears at the same rt (*m/z* 746.86) with the appropriate 0.5 isotope increase, which presumably is the double charge ion of the parent ion. Characterising this molecule would require targeted optimisation and culture scale-up to obtain conclusive fragmentation patterns, but some structural speculation can be made based on the spectra and the derived network. While not every node in the network is necessarily a fragment (Bandeira *et al.*, 2007) especially from a complex mixture analysed with generalised MS-MS settings, comparison with the original can help confirm authenticity.

Assuming *m/z* 1492.72 is the parent ion, a loss of 78 is required to reach 1414.72. This could indicate the removal of a C_6H_5 + H ring (Gross, 2011), followed by a further CH_2 to produce the node with *m/z* 1400.72. Removal of an *m/z* 78 fragment from the overall structure is supported by the presence of *m/z* 1403.72, which if also subtracted 78 corresponds to another network node, *m/z* 1325.69. Examining the original spectra again shows a separate ion (*m/z* 1243.64) which also has a putative doubly charged form (*m/z* 622.31). This second ion has a similar mass to actinomycins, but checking their spectra does

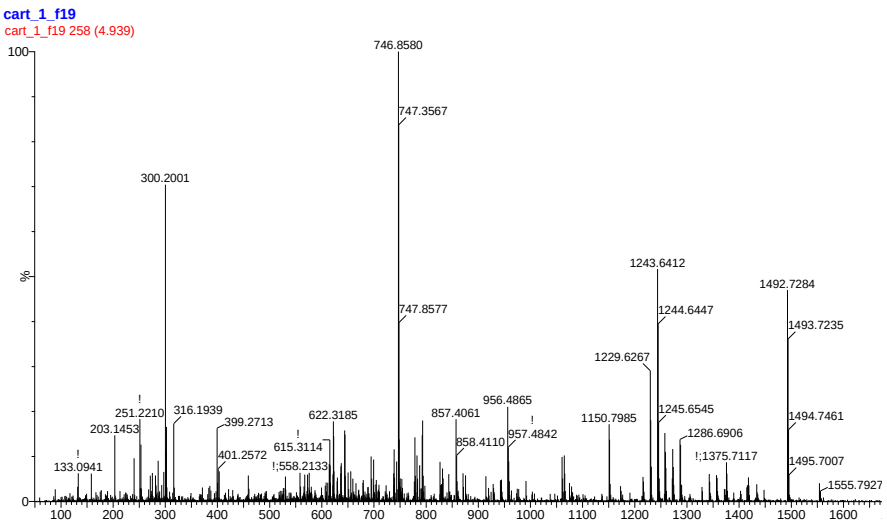
not show double charged ions so, assuming that these are fragments of *m/z* 1492.72, it is unlikely to be an actinomycin analogue.

Figure 5.28 Putative unknown metabolite **(a)** network, and **(b)** mass spectra.

(a)



(b)



Identifying this potential metabolite not just amongst the actinomycin spectra but all the other data files would have been highly unlikely, again making spectral molecular networking a useful tool for metabolite discovery.

5.4 Conclusions

The aim for this chapter was to scale up culture of *S. costaricanus* in M19 to identify and purify bioactive metabolites from fractions. These were targeted against a strain of multi-drug resistant *A. baumannii*, a pathogen globally recognised as needing novel treatments. Scale-up in broth and agar culture was successful, with both media types requiring different extraction and chemical processing methods. This led to an extract library primarily containing the well-known and previously identified metabolites actinomycin D, X2, and deferoxamine B. Through the dereplication process, involving both manual fraction inspection and automated MS/MS spectral screening, further metabolites belonging to both those families were found, as well as other non-deferoxamine siderophores (bisucaberin) and leupeptins. This part of the aim was therefore fulfilled. An explanation for the originally detected antihelminth bioactivity of *S. costaricanus* was also suggested through an unknown azoline metabolite.

Fractions active against multi-drug resistant *A. baumannii* were detected, though most of them were predominantly actinomycin D and other family members, which – especially actinomycin D – are not known to be active against *A. baumannii*. Running an XTT assay using actinomycin D against *A. baumannii* found no significant difference to the negative control, nor did an alternative disc diffusion assay. Pooling the fractions further purified metabolites, though later MS/MS-based networking showed some metabolite overlap between pools. The bioactive component could not be identified from a mostly actinomycin D-containing fraction. Assuming the bioactive metabolite was detected by the positive mode UPLC-MS method used, this implies one of the minor peaks was responsible for bioactivity. As a minor peak within a 50 µg/ml fraction, the metabolite must be active at relatively low concentrations against a multi-drug resistant pathogen.

CHAPTER 6

GENERAL DISCUSSION

6 General discussion

6.1 Thesis aim

The central question at the start of this thesis was whether specialised metabolite diversity could be improved in underexploited *Streptomyces* strains. There were 4 main objectives stated as necessary to achieve this: development of a culture parameter-based screen that would elicit varied production of specialised metabolites, application of the screen to unexploited strains, genome sequencing of an unexploited strain, and finally culture scale-up and metabolite bioactivity. This required several experimental stages starting with method development, strain selection, and culture parameters. These were grouped into 4 chapters, which are summarised in Figure 6.1. These chapters contained a range of different methods used to complete the main aim but were not the only options available. Section 6.2 details some of these and the potential impact their inclusion could have had on the project.

6.2 Suitability of methods

Key considerations were the applicability of the methods to the remainder of NCIMB strains to enable continued exploitation while incorporating modern metabolomic and genomic tools into the workflow. As a culture parameter-based screen, the culture methods used are easily transferred not just to other *Streptomyces* strains, but also NCIMB *Actinomycetales* accessions and other specialised metabolite producers of interest. The parameters may need altering depending on strain-specific requirements or when testing a specific variable, but the overall culture, extraction and analysis method is broadly applicable. For the second consideration, both the primary genomic and metabolomic software used (antiSMASH and MZmine respectively) are continually updated and have user-friendly interfaces, making them ideal for generalised strain exploitation workflows.

6.2.1 Impact of strain selection method

While this overall method was suitable for the NCIMB *Actinomycetales*, there are multiple workflows for metabolite discovery. The optimal route will depend on the type of sample, level of information, proposed targets, and analytical techniques used. For example, if starting from a soil sample, both the culture-dependent and metagenomic routes are viable, including focusing on lesser cultured *Actinomycetales* or specific BGC classes. If more genetic information had been available at the start of the project, such as full 16S sequencing for all isolates, phylogenetics would likely have been incorporated into the workflow to better select the most varied strains.

However, as was later noted in section 4.1.1, specialised metabolite diversity amongst a single species can be extensive. Similarly, Purves *et al.* (2016) found that taxonomically close *Bacillus lichenformis* strains isolated from different niches produced different metabolites. Had a genomics-based method been adopted at the start of the project, it would likely have produced an accurate 16S-based phylogenetic map of NCIMB strains but not reflected natural product diversity. Therefore, metabolomics approaches remain the best starting point when starting from a wide-ranging collection with genetically even semi characterised strains. As full genome sequencing becomes more routine this may change with direct access to BGC content, in turn enabling greater documentation of BGCs and their submission to public databases to link genomic and metabolomic data.

A broad environmental categorisation was used for the NCIMB strains, grouping soil, marine, plant, and animal associated isolates. Soil strains were focused on originally due to the majority of *Streptomyces* being isolated from the soil, with strains from the other categories included once the metabolomic screening method had been determined. An alternative method could have been a focus on extreme environments, such as metal contaminated soil, high or low pH, arctic sediment, or the highly alkaline Atacama Desert isolates. This would have enabled a more focused OSMAC screen, with isolate niche specific growth parameters and stresses being applied to relevant strains, such as supplementing media with copper for a high copper soil isolate.

While attempting to recreate environmental conditions in a lab is a logical way to trigger transcription of cryptic BGCs, many BGCs are upregulated by elicitors for unknown reasons (Jakeman *et al.*, 2006) resulting in large scale arrays of elicitors being applied to activate a single cluster (Seyedsayamdost, 2014). Cryptic BGCs will likely be activated by unexpected elicitors or carbon and nitrogen sources, so a generalised OSMAC screen should still be a highly applicable part of the workflow. Strain selection by extreme environment would have been a viable alternative with the key requirement that it did not narrow the OSMAC screen variety.

6.2.2 Culture parameters

The final set of culture parameters used a mixture of previously identified media and media based on recommended ingredients for *Streptomyces* metabolite production. These were primarily complex media, some with multiple carbon and nitrogen sources, like MM which contained molasses and glucose. Goodfellow & Fiedler (2010) recommend multiple complex media for *Streptomyces* SMet production and discovery, but this had the effect of creating a messy UPLC-MS chromatogram which slowed dereplication. An alternative would have been to use only 2 components per media: one carbon source and one nitrogen source, enabling precise linking of metabolite production to specific media. This would have benefitted both future linking of BGCs to metabolites and also novel metabolite detection. However, if making media using new components, *S. costaricanus* could grow poorly or not at all, lowering the amount of new data. Using a complex medium with a small amount of sugar or nitrogen known to be metabolised by *S. costaricanus* ensured that growth would still occur, but not to the point where SMet was negatively impacted.

The culture methods broadly used conventional broth and media cultivation. While these are the standard for the majority of experiments, there are some alternative substrates which can lead to the production of novel metabolites even if using the same media, or allow detection of multiple days of metabolite development from a single sample. Timmermans *et al.* (2019) cultured the marine Gram-negative bacterium *Pseudoalteromonas* on cotton wool, observing changes in pigmentation and significant induction in production of other

metabolites, most notably of the antibiotic thiomarinol A. Other unidentified metabolites were also induced, indicating that – as with *Streptomyces* – the physical structure of a culture can impact metabolism and likely differentiation. Growth conditions in the soil are highly dissimilar to a standard agar plate with many more pockets and variations in the surface. While culture in a similar soil mimicking (or cotton) environment could potentially lead to the production of novel metabolites, the challenge would be increasing metabolite titres from the level a strain would normally produce in the soil.

Pishchany *et al.* (2018) employed a transwell co-culture membrane to hold an agar plug partially submerged in a well plate, allowing secreted metabolites to diffuse into the liquid medium over time. This method would allow continuous measurement of metabolite development, going far past the 2 time points used in the final OSMAC screen, potentially to daily samples. It is limited to secreted metabolites and may give a false impression of production over time if some metabolites are slower to diffuse through the agar than others. Despite this, the method could easily be adapted to examine how the *S. costaricanus* chemotype develops over a longer and more frequent time period than 5 and 10 d.

6.2.3 Target selection

Here, "target selection" refers to whether the specific objective is to produce novel metabolites and test for their bioactivity, or to find bioactive extracts and then identify a novel metabolite. After OSMAC screening of *S. costaricanus*, M19 was identified as the media which most enabled diverse metabolite production. This led to scale-up of *S. costaricanus* using M19 agar and broth for extract fractionation and bioactivity assays against multi-drug resistant *Acinetobacter baumannii* due to its status as a pathogen critically requiring novel treatments. While an approach based on activity against a single target does give a clear objective, it may have negatively impacted the novel metabolite discovery segment of the project.

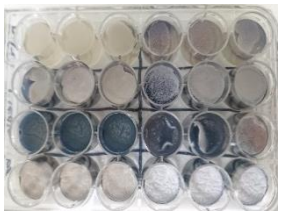
This was primarily shown from 2 areas of the project: the minimal media screen and the detection of putatively novel actinomycins. Minimal media with GlcNAc culture of *S. costaricanus* led to the production of an unknown metabolite family,

which could potentially have been either scaled up, isolated, structurally characterised and then assayed against multiple disease targets. Alternatively, the novel actinomycins were already partially structurally characterised and would have been logical choices for anticancer assays (Cai *et al.*, 2016). A basic disc diffusion assay against *A. baumannii* did not show activity when using the minimal media crude extracts, possibly due to low production levels from the 2 ml of media per well. The disadvantages of scale-up and wider activity testing would have primarily been extensive optimisation, extraction, and selection of a panel of targets to test against.

The full workflow, starting from a group of uncharacterised strains through to metabolite identification, is generally applicable and was successful in initiating exploitation of *S. costaricanus*. Although no novel metabolite could be identified, multiple metabolites previously unknown to be produced by *S. costaricanus* were detected, such as biscuberin 01 and leupeptin, which did aid bioactive fraction dereplication. The 3 experimental chapters which studied *S. costaricanus* also raised novel research questions, whether in novel directions or continuing the chapter theme. These are briefly listed in section 6.3.

Figure 6.1 Summary of experimental chapters.

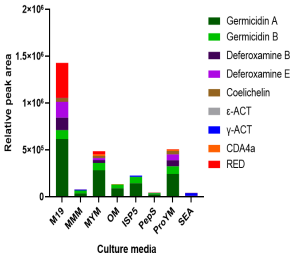
1. Method optimisation with *S. coelicolor*



Culture method development

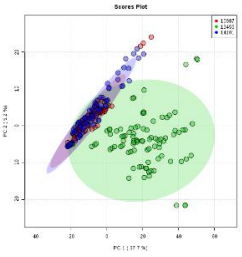
MZmine 2

Initial MZmine processing and analysis

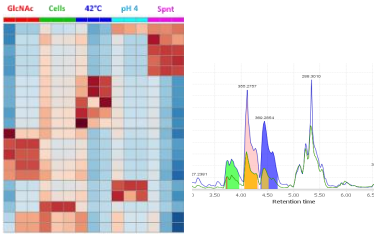


Metabolite production comparisons

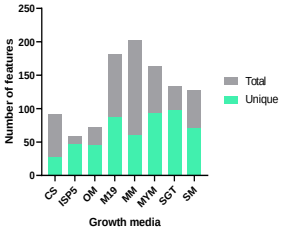
2. Application of method to unexploited strains



Metabolomic profiling for strain selection

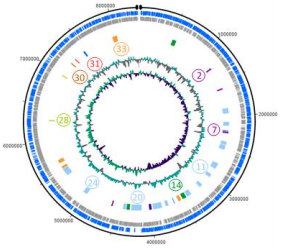


Additional elicitors to expand production

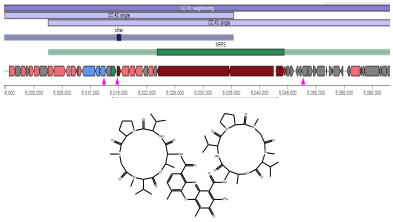


Analysis of produced features to select scale-up media

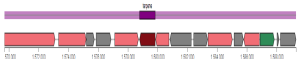
3. Biosynthetic potential of *S. costaricanus*



Hybrid sequencing for genome mining

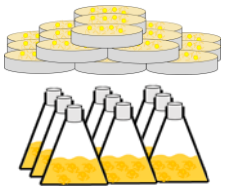


Linking genomics to detected metabolites

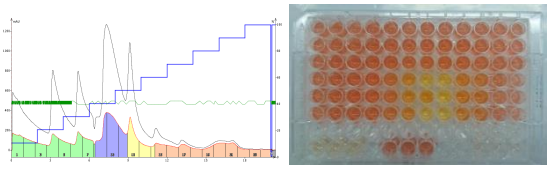


Low homology clusters indicate novel chemistry

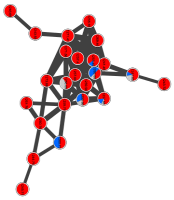
4. Scale-up and bioactivity of *S. costaricanus* metabolites



Culture in agar and broth media



Bioactivity of semi-purified fractions



MS/MS based identification and networking

6.3 Future work

6.3.1 Untargeted metabolomics screening of *S. costaricanus*

Two main areas of future work are suggested. The first is expansion of the 24-well plate screening for *S. costaricanus* metabolites but using different elicitors to affect either production of novel metabolites or structurally alter existing ones. As an example, restricting the screen to 1 media then switching out different halogens is a simple method; other factors such as long-term culture and heavy metals would help fully characterise not just the standard specialised metabolome, but also structurally distinct analogues.

The second strand is continued investigation of minimal media culture, in particular of the features which were significantly upregulated after addition of an elicitor. The N-acetyl glucosamine elicited peaks would be clear starting choices and would provide a more targeted route through scale-up and preparative HPLC, along with the feature detected after addition of heat-killed *A. baumannii* cells. Using these elicitors in richer media could aid production, so long as the effect of the elicitor remains the same in less harsh conditions.

6.3.2 Exploitation of the genomic potential of *S. costaricanus*

This project has relied on altering culture parameters to elicit production of novel metabolites. However, one consideration with this method is the lack of predictive power it has in which metabolites will be produced from an unexploited strain. The full genome sequence identified one specific BGC with minimal homology to other strains which is apparently limited to its clade. Therefore, genetic engineering to express this BGC – whether in *S. costaricanus* or a heterologous host – would be a further validation of the original metabolomic method of strain selection.

This BGC is also relatively simple in that it is not particularly large – about 20 kb – and is a single BGC type, not a hybrid, so may be easier to express in a non-native strain. One potential method for engineering *S. costaricanus* would be through the CRISPR-enabled method CRISPR-BEST, which precisely edits the genome to change a C:G pair into an A:T pair (Tong *et al.*, 2019). This could be

used to introduce stop codons into other BGCs, simplifying detection of the desired BGC product. Alternatively, methods like transformation-associated recombination capture a target BGC in a plasmid which can then be transferred through conjugation into a suitable host like *S. coelicolor* (Jia Jia Zhang *et al.*, 2019).

6.3.3 Scale-up culture and bioactive metabolites of *S. costaricanus*

The most logical step would be to increase scale-up to generate more material and unambiguously identify the metabolite, or potential synergistically acting metabolites, responsible for bioactivity against *A. baumannii* in the most bioactive fractions. Alongside this, the putative novel actinomycins could be purified and structurally characterised by MS and NMR, expanding the known range of actinomycins that *S. costaricanus* is able to produce.

6.4 Conclusions

This thesis addressed the question of assessing specialised metabolite diversity using a library of underexploited *Streptomyces*, selecting one strain to help develop an overall framework for the further characterisation of NCIMB isolates. The overall metabolomic strategy was reinforced with modern genomic tools which helped guide not just specialised metabolite identification and BGC potential but also a potential phylogenetic reclassification for *S. costaricanus* based on its MLST profile.

The majority of the *S. costaricanus* BGCs were unknown, and multiple potential avenues for further research were identified. Within the context of with an overall resurgence into *Streptomyces* specialised metabolites, this project investigated an underutilised source of novel metabolites, and has demonstrated that it is a valuable candidate for continued research.

REFERENCES

7 References

Abraham EP & Chain E (1940) An Enzyme from Bacteria able to Destroy Penicillin. *Nature* **146**: 837.

Abraham EP, Chain E, Fletcher CM, Gardner AD, Heatley NG, Jennings MA & Florey HW (1941) Further Observations on Penicillin. *Lancet* **238**: 177–189.

Alanjary M, Kronmiller B, Adamek M, Blin K, Weber T, Huson D, Philmus B & Ziemert N (2017) The Antibiotic Resistant Target Seeker (ARTS), an exploration engine for antibiotic cluster prioritization and novel drug target discovery. *Nucleic Acids Res* **45**: W42–W48.

Alanjary M, Steinke K & Ziemert N (2019) AutoMLST: an automated web server for generating multi-locus species trees highlighting natural product potential. *Nucleic Acids Res* **47**: W276–W282.

Antony-Babu S, Stien D, Eparvier V, Parrot D, Tomasi S & Suzuki MT (2017) Multiple *Streptomyces* species with distinct secondary metabolomes have identical 16S rRNA gene sequences. *Sci Rep* **7**: 11089.

Arcilla MS, Hattem JM van, Haverkate MR, Bootsma MCJ, Genderen PJJ van, Goorhuis A, Grobusch MP, Lashof AMO, Molhoek N, Schultsz C, Stobberingh EE, Verbrugh HA, Jong MD de, Melles DC & Penders J (2017) Import and spread of extended-spectrum β-lactamase-producing *Enterobacteriaceae* by international travellers (COMBAT study): a prospective, multicentre cohort study. *Lancet Infect Dis* **17**: 78–85.

Ariza-Prota MA, Pando-Sandoval A, Fole-Vázquez D, García-Clemente M, Budiño T & Casan P (2015) Community-acquired bacteremic *Streptomyces atratus* pneumonia in animmunocompetent adult: A case report. *J Med Case Rep* **9**.

Arnison PG, Bibb MJ, Bierbaum G, Bowers AA, Bugni TS, Bulaj G, Camarero JA, Campopiano DJ, Challis GL, Donk WA Van Der, *et al.* (2013) Ribosomally synthesized and post-translationally modified peptide natural products: Overview and recommendations for a universal nomenclature. *Nat Prod Rep* **30**: 108–160.

Aziz RK, Bartels D, Best AA, DeJongh M, Disz T, Edwards RA, Formsma K, Gerdes S, Glass EM, Kubal M, Meyer F, Olsen GJ, Olson R, Osterman AL, Overbeek RA, McNeil LK, Paarmann D, Paczian T, Parrello B, Pusch GD, Reich C, Stevens R, Vassieva O, Vonstein V, Wilke A & Zagnitko O (2008) The RAST Server: Rapid Annotations using Subsystems Technology. *BMC Genomics* **9**: 75.

Baltz RH (2011) Function of MbtH homologs in nonribosomal peptide biosynthesis and applications in secondary metabolite discovery. *J Ind Microbiol Biotechnol* **38**: 1747–1760.

Baltz RH (2014) MbtH homology codes to identify gifted microbes for genome mining. *J Ind Microbiol Biotechnol* **41**: 357–369.

Baltz RH (2016) Genetic manipulation of secondary metabolite biosynthesis for improved production in *Streptomyces* and other actinomycetes. *J Ind Microbiol Biotechnol* **43**: 343–370.

Baltz RH (2017) Gifted microbes for genome mining and natural product discovery. *J Ind Microbiol Biotechnol* **44**: 573–588.

Bandeira N, Tsur D, Frank A & Pevzner PA (2007) Protein identification by spectral networks analysis. *Proc Natl Acad Sci* **104**: 6140–6145.

Bankevich A, Nurk S, Antipov D, Gurevich AA, Dvorkin M, Kulikov AS, Lesin VM, Nikolenko SI, Pham S, Prjibelski AD, Pyshkin A V., Sirotkin A V., Vyahhi N, Tesler G, Alekseyev MA & Pevzner PA (2012) SPAdes: A New Genome Assembly Algorithm and Its Applications to Single-Cell Sequencing. *J Comput Biol* **19**: 455–477.

Baranasic D, Gacesa R, Starcevic A, Zucko J, Blazic M, Horvat M, Gjuracic K, Fujs S, Hranueli D, Kosec G, Cullum J & Petkovic H (2013) Draft Genome Sequence of *Streptomyces rapamycinicus* Strain NRRL 5491, the Producer of the Immunosuppressant Rapamycin. *Genome Announc* **1**: e00581-13.

Barka EA, Vatsa P, Sanchez L, Gaveau-vaillant N, Jacquard C, Klenk H-P, Clément C, Ouhdouch Y & Wezel GP van (2016) Taxonomy, Physiology, and Natural Products of *Actinobacteria*. *Microbiol Mol Biol Rev* **80**: 1–43.

Barkal LJ, Theberge AB, Guo C-J, Spraker J, Rappert L, Berthier J, Brakke KA, Wang CCC, Beebe DJ, Keller NP & Berthier E (2016) Microbial metabolomics in open microscale platforms. *Nat Commun* **7**: 10610.

Barlam TF, Cosgrove SE, Abbo LM, MacDougall C, Schuetz AN, Septimus EJ, Srinivasan A, Dellit TH, Falck-Ytter YT, Fishman NO, Hamilton CW, Jenkins TC, Lipsett PA, Malani PN, May LS, Moran GJ, Neuhauser MM, Newland JG, Ohl CA, Samore MH, Seo SK & Trivedi KK (2016) Implementing an Antibiotic Stewardship Program: Guidelines by the Infectious Diseases Society of America and the Society for Healthcare Epidemiology of America. *Clin Infect Dis* **62**: e51–e77.

Becerril A, Álvarez S, Braña AF, Rico S, Díaz M, Santamaría RI, Salas JA & Méndez C (2018) Uncovering production of specialized metabolites by *Streptomyces argillaceus*: Activation of cryptic biosynthesis gene clusters using nutritional and genetic approaches. *PLoS One* **13**.

Becker B & Cooper MA (2013) Aminoglycoside antibiotics in the 21st century. *ACS Chem Biol* **8**: 105–115.

Bentley SD, Chater KF, Cerdeño-Tárraga A-M, Challis GL, Thomson NR, James KD, Harris DE, Quail M a, Kieser H, Harper D, Bateman A, Brown S, Chandra G, Chen CW, Collins M, Cronin A, Fraser A, Goble A, Hidalgo J, Hornsby T, Howarth S, Huang C-H, Kieser T, Larke L, Murphy L, Oliver K, O'Neil S, Rabbinowitsch E, Rajandream M, Rutherford K, Rutter S, Seeger K, Saunders D, Sharp S, Squares R, Squares S, Taylor K, Warren T, Wietzorrek A, Woodward J, Barrell BG, Parkhill J & Hopwood DA (2002) Complete genome sequence of the model actinomycete *Streptomyces coelicolor* A3(2). *Nature* **417**: 141–147.

Bergman M, Huikko S, Pihlajamäki M, Laippala P, Palva E, Huovinen P & Seppälä H (2004) Effect of macrolide consumption on erythromycin resistance in *Streptococcus pyogenes* in Finland in 1997-2001. *Clin Infect Dis* **38**: 1251–1256.

Berridge M V., Herst PM & Tan AS (2005) Tetrazolium dyes as tools in cell biology: New insights into their cellular reduction. *Biotechnol Annu Rev* **11**: 127–

152.

Biggins JB, Liu X, Feng Z & Brady SF (2011) Metabolites from the induced expression of cryptic single operons found in the genome of *Burkholderia pseudomallei*. *J Am Chem Soc* **133**: 1638–1641.

Bitzer J, Gesheva V & Zeeck A (2006) Actinomycins with altered threonine units in the β-peptidolactone. *J Nat Prod* **69**: 1153–1157.

Blin K, Pascal Andreu V, los Santos ELC de, Carratore F Del, Lee SY, Medema MH & Weber T (2019) The antiSMASH database version 2: a comprehensive resource on secondary metabolite biosynthetic gene clusters. *Nucleic Acids Res* **47**: D625–D630.

Blin K, Shaw S, Steinke K, Villebro R, Ziemert N, Lee SY, Medema MH & Weber T (2019) antiSMASH 5.0: updates to the secondary metabolite genome mining pipeline. *Nucleic Acids Res* **47**: W81–W87.

Blin K, Wolf T, Chevrette MG, Lu X, Schwalen CJ, Kautsar SA, Suarez Duran HG, los Santos ELC de, Kim HU, Nave M, Dickschat JS, Mitchell DA, Shelest E, Breitling R, Takano E, Lee SY, Weber T & Medema MH (2017) antiSMASH 4.0—improvements in chemistry prediction and gene cluster boundary identification. *Nucleic Acids Res* **45**: W36–W41.

Bobek J, Šmídová K & Čihák M (2017) A waking review: Old and novel insights into the spore germination in Streptomyces. *Front Microbiol* **8**: 2205.

Bode HB, Bethe B, Höfs R & Zeeck A (2002) Big Effects from Small Changes: Possible Ways to Explore Nature's Chemical Diversity. *ChemBioChem* **3**: 619.

Brown ED & Wright GD (2016) Antibacterial drug discovery in the resistance era. *Nature* **529**: 336–343.

Bucca G, Hindle Z & Smith CP (1997) Regulation of the dnaK operon of *Streptomyces coelicolor* A3(2) is governed by HspR, an autoregulatory repressor protein. *J Bacteriol* **179**: 5999–6004.

Bunet R, Song L, Mendes MV, Corre C, Hotel L, Rouhier N, Framboisier X, Leblond P, Challis GL & Aigle B (2011) Characterization and manipulation of the pathway-specific late regulator AlpW reveals *Streptomyces ambofaciens* as a new producer of kinamycins. *J Bacteriol* **193**: 1142–1153.

Bush MJ, Tschowri N, Schlimpert S, Flärdh K & Buttner MJ (2015) c-di-GMP signalling and the regulation of developmental transitions in streptomycetes. *Nat Rev Microbiol* **13**: 749–760.

Caffrey P, Lynch S, Flood E, Finnan S & Oliynyk M (2001) Amphotericin biosynthesis in *Streptomyces nodosus*: Deductions from analysis of polyketide synthase and late genes. *Chem Biol* **8**: 713–723.

Cai W, Wang X, Elshahawi SI, Ponomareva L V, Liu X, McErlean MR, Cui Z, Arlinghaus AL, Thorson JS & Lanen SG Van (2016) Antibacterial and Cytotoxic Actinomycins Y6-Y9 and Zp from *Streptomyces* sp. Strain Gö-GS12. *J Nat Prod* **79**: 2731–2739.

Carattoli A, Zankari E, García-Fernández A, Voldby Larsen M, Lund O, Villa L,

Møller Aarestrup F & Hasman H (2014) In Silico Detection and Typing of Plasmids using PlasmidFinder and Plasmid Multilocus Sequence Typing. *Antimicrob Agents Chemother* **58**: 3895–3903.

Carver T, Harris SR, Berriman M, Parkhill J & McQuillan JA (2012) Artemis: an integrated platform for visualization and analysis of high-throughput sequence-based experimental data. *Bioinformatics* **28**: 464–469.

Chain E, Florey HW, Adelaide MB, Gardner AD, Oxfd DM, Heatley NG, Jennings MA, Orr-Ewing J & Sanders AG (1940) Penicillin as a chemotherapeutic agent. *Lancet* **236**: 226–228.

Challis GL (2014) Exploitation of the *Streptomyces coelicolor* A3(2) genome sequence for discovery of new natural products and biosynthetic pathways. *J Ind Microbiol Biotechnol* **41**: 219–232.

Chambers MC, Maclean B, Burke R, Amodei D, Ruderman DL, Neumann S, Gatto L, Fischer B, Pratt B, Egertson J, Hoff K, Kessner D, Tasman N, Shulman N, Frewen B, Baker TA, Brusniak M-Y, Paulse C, Creasy D, Flashner L, Kani K, Moulding C, Seymour SL, Nuwaysir LM, Lefebvre B, Kuhlmann F, Roark J, Rainer P, Detlev S, Hemenway T, Huhmer A, Langridge J, Connolly B, Chadick T, Holly K, Eckels J, Deutsch EW, Moritz RL, Katz JE, Agus DB, MacCoss M, Tabb DL & Mallick P (2012) A cross-platform toolkit for mass spectrometry and proteomics. *Nat Biotechnol* **30**: 918–920.

Chander J, Singla N & Handa U (2015) Human cervicofacial mycetoma caused by *Streptomyces griseus*: First case report. *J Microbiol Immunol Infect* **48**: 703–705.

Chang Q, Wang W, Regev-Yochay G, Lipsitch M & Hanage WP (2015) Antibiotics in agriculture and the risk to human health: How worried should we be? *Evol Appl* **8**: 240–247.

Charusanti P, Fong NL, Nagarajan H, Pereira AR, Li HJ, Abate EA, Su Y, Gerwick WH & Palsson BO (2012) Exploiting Adaptive Laboratory Evolution of *Streptomyces clavuligerus* for Antibiotic Discovery and Overproduction. *PLoS One* **7**: e33727.

Chater KF, Biró S, Lee KJ, Palmer T & Schrempf H (2010) The complex extracellular biology of *Streptomyces*. *FEMS Microbiol Rev* **34**: 171–198.

Chater KF & Chandra G (2006) The evolution of development in Streptomyces analysed by genome comparisons. *FEMS Microbiol Rev* **30**: 651–672.

Chater KF & Chandra G (2008) The use of the rare UUA codon to define "Expression Space" for genes involved in secondary metabolism, development and environmental adaptation in *Streptomyces*. *J Microbiol* **46**: 1–11.

Chemler JA, Buchholz TJ, Geders TW, Akey DL, Rath CM, Chlipala GE, Smith JL & Sherman DH (2012) Biochemical and structural characterization of germicidin synthase: Analysis of a type III polyketide synthase that employs Acyl-ACP as a starter unit donor. *J Am Chem Soc* **134**: 7359–7366.

Chen B, Lin L, Fang L, Yang Y, Chen E, Yuan K, Zou S, Wang X & Luan T (2018) Complex pollution of antibiotic resistance genes due to beta-lactam and

aminoglycoside use in aquaculture farming. *Water Res* **134**: 200–208.

Chen C, Song F, Wang Q, Abdel-Mageed WM, Guo H, Fu C, Hou W, Dai H, Liu X, Yang N, Xie F, Yu K, Chen R & Zhang L (2012) A marine-derived *Streptomyces* sp. MS449 produces high yield of actinomycin X2 and actinomycin D with potent anti-Tuberculosis activity. *Appl Microbiol Biotechnol* **95**: 919–927.

Chen G, Wang GY, Li X, Waters B & Davies J (2000) Enhanced production of microbial metabolites in the presence of dimethyl sulfoxide. *J Antibiot (Tokyo)* **53**: 1145–53.

Chen Y-C, Liu T, Yu C-H, Chiang T-Y & Hwang C-C (2013) Effects of GC Bias in Next-Generation-Sequencing Data on De Novo Genome Assembly. *PLoS One* **8**: e62856.

Cheng C, MacIntyre L, Abdelmohsen UR, Horn H, Polymenakou PN, Edrada-Ebel R & Hentschel U (2015) Biodiversity, Anti-Trypanosomal Activity Screening, and Metabolomic Profiling of Actinomycetes Isolated from Mediterranean Sponges. *PLoS One* **10**: e0138528.

Cheng YR, Hauck L & Demain AL (1995) Phosphate, ammonium, magnesium and iron nutrition of *Streptomyces hygroscopicus* with respect to rapamycin biosynthesis. *J Ind Microbiol* **14**: 424–7.

Chevrette MG, Carlson CM, Ortega HE, Thomas C, Ananiev GE, Barns KJ, Book AJ, Cagnazzo J, Carlos C, Flanigan W, Grubbs KJ, Horn HA, Hoffmann FM, Klassen JL, Knack JJ, Lewin GR, McDonald BR, Muller L, Melo WGP, Pinto-Tomás AA, Schmitz A, Wendt-Pienkowski E, Wildman S, Zhao M, Zhang F, Bugni TS, Andes DR, Pupo MT & Currie CR (2019) The antimicrobial potential of *Streptomyces* from insect microbiomes. *Nat Commun* **10**: 516.

Chiani M, Akbarzadeh A, Farhangi A, Mazinani M, Saffari Z, Emadzadeh K & Mehrabi MR (2010) Optimization of culture medium to increase the production of desferrioxamine B (Desferal) in *Streptomyces pilosus*. *Pakistan J Biol Sci* **13**: 546–550.

Chong J, Soufan O, Li C, Caraus I, Li S, Bourque G, Wishart DS & Xia J (2018) MetaboAnalyst 4.0: towards more transparent and integrative metabolomics analysis. *Nucleic Acids Res* **46**: W486–W494.

Choulet F, Aigle B, Gallois A, Mangenot S, Gerbaud C, Truong C, Francou F-X, Fourrier C, Guérineau M, Decaris B, Barbe V, Pernodet J-L & Leblond P (2006) Evolution of the Terminal Regions of the Streptomyces Linear Chromosome. *Mol Biol Evol* **23**: 2361–2369.

Čihák M, Kameník Z, Šmídová K, Bergman N, Benada O, Kofroňová O, Petříčková K & Bobek J (2017) Secondary Metabolites Produced during the Germination of *Streptomyces coelicolor*. *Front Microbiol* **8**: 2495.

Claessen D, Stokroos I, Deelstra HJ, Penninga NA, Bormann C, Salas JA, Dijkhuizen L & Wösten HAB (2004) The formation of the rodlet layer of streptomycetes is the result of the interplay between rodlins and chaplins. *Mol Microbiol* **53**: 433–443.

Coisne S, Béchet M & Blondeau R (1999) Actinorhodin production by

Streptomyces coelicolor A3(2) in iron-restricted media. *Lett Appl Microbiol* **28**: 199–202.

Corre C & Challis GL (2009) New natural product biosynthetic chemistry discovered by genome mining. *Nat Prod Rep* **26**: 977.

Corre C, Haynes SW, Malet N, Song L & Challis GL (2010) A butenolide intermediate in methylenomycin furan biosynthesis is implied by incorporation of stereospecifically 13C-labelled glycerols. *Chem Commun* **46**: 4079–4081.

Covington BC, McLean JA & Bachmann BO (2017) Comparative mass spectrometry-based metabolomics strategies for the investigation of microbial secondary metabolites. *Nat Prod Rep* **34**: 6–24.

Crüsemann M, O'Neill EC, Larson CB, Melnik A V., Floros DJ, Silva RR da, Jensen PR, Dorrestein PC & Moore BS (2017) Prioritizing Natural Product Diversity in a Collection of 146 Bacterial Strains Based on Growth and Extraction Protocols. *J Nat Prod* **80**: 588–597.

D'Costa VM, McGrann KM, Hughes DW & Wright GD (2006) Sampling the Antibiotic Resistome. *Science (80-)* **311**: 374–377.

Dalili M & Chau PC (1988) Production of actinomycin D with immobilized *Streptomyces parvullus* under nitrogen and carbon starvation conditions. *Biotechnol Lett* **10**: 331–336.

Darling AE, Mau B & Perna NT (2010) progressiveMauve: Multiple Genome Alignment with Gene Gain, Loss and Rearrangement. *PLoS One* **5**: e11147.

DaSilva G & Domingues S (2016) Insights on the Horizontal Gene Transfer of Carbapenemase Determinants in the Opportunistic Pathogen *Acinetobacter baumannii*. *Microorganisms* **4**: 29.

Dawkins R (1982) *The Extended Phenotype: The Long Reach of the Gene*. Oxford University Press, Oxford.

Debono M, Barnhart M, Carrell CB, Hoffmann JA, Occolowitz JL, Abbott BJ, Fukuda DS, Hamill RL, Biemann K & Herlihy WC (1987) A21978C, a complex of new acidic peptide antibiotics: Isolation, chemistry, and mass spectral structure elucidation. *J Antibiot (Tokyo)* **40**: 761–777.

Delcour AH (2009) Outer membrane permeability and antibiotic resistance. *Biochim Biophys Acta - Proteins Proteomics* **1794**: 808–816.

Demain AL (1998) Induction of microbial secondary metabolism. *Int Microbiol* **1**: 259–264.

Demirci H, Murphy F, Murphy E, Gregory ST, Dahlberg AE & Jogl G (2013) A structural basis for streptomycin-induced misreading of the genetic code. *Nat Commun* **4**: 1355.

Didelot X, Bowden R, Wilson DJ, Peto TEA & Crook DW (2012) Transforming clinical microbiology with bacterial genome sequencing. *Nat Rev Genet* **13**: 601–612.

Ding D, Stracher A & Salvi RJ (2002) Leupeptin protects cochlear and vestibular

hair cells from gentamicin ototoxicity. *Hear Res* **164**: 115–126.

Dissel D van, Claessen D & Wezel GP van (2014) Morphogenesis of *Streptomyces* in Submerged Cultures. In *Advances in Applied Microbiology*. pp. 1–45.

Doroghazi JR, Albright JC, Goering AW, Ju K-S, Haines RR, Tchalukov KA, Labeda DP, Kelleher NL & Metcalf WW (2014) A roadmap for natural product discovery based on large-scale genomics and metabolomics. *Nat Chem Biol* **10**: 963–968.

Doroghazi JR & Buckley DH (2010) Widespread homologous recombination within and between *Streptomyces species*. *ISME J* **4**: 1136–1143.

Doull JL, Ayer SW, Thibault P & Singh AK (1993) Production of a novel polyketide antibiotic, jadomycin b, by *Streptomyces venezuelae* following heat shock. *J Antibiot (Tokyo)* **46**: 869–871.

DSMZ (2019) Details: DSM-40301. https://www.dsmz.de/catalogues/details/culture/DSM-40301.html?tx_dsmzresources_pi5%5BreturnPid%5D=304. Accessed June 13, 2019.

Dubos RJ (1939) Bactericidal Effect of an Extract of a Soil *Bacillus* on Gram Positive Cocci. *Proc Soc Exp Biol Med* **40**: 311–312.

Dubos RJ & Hotchkiss RD (1941) The Production of Bactericidal Substances by Aerobic Sporulating Bacilli. *J Exp Med* **73**: 629–40.

Duggar BM (1948) Aureomycin: a product of the continuing search for new antibiotics. *Ann N Y Acad Sci* **51**: 177–181.

Ehrlich J, Bartz QR, Smith RM, Joslyn DA & Burkholder PR (1947) Chloromycetin, a New Antibiotic From a Soil Actinomycete. *Science (80-)* **106**: 417.

Elliot MA, Locke TR, Galibois CM & Leskiw BK (2003) BldD from *Streptomyces coelicolor* is a non-essential global regulator that binds its own promoter as a dimer. *FEMS Microbiol Lett* **225**: 35–40.

Ender M, McCallum N, Adhikari R & Berger-Bächi B (2004) Fitness cost of SCCmec and methicillin resistance levels in *Staphylococcus aureus*. *Antimicrob Agents Chemother* **48**: 2295–2297.

Erenler R, Meral B, Sen O, Elmastas M, Aydin A, Eminagaoglu O & Topcu G (2017) Bioassay-guided isolation, identification of compounds from *Origanum rotundifolium* and investigation of their antiproliferative and antioxidant activities. *Pharm Biol* **55**: 1646–1653.

Ernst M, Kang K Bin, Caraballo-Rodríguez AM, Nothias L-F, Wandy J, Chen C, Wang M, Rogers S, Medema MH, Dorrestein PC & Hooft JJJ van der (2019) MolNetEnhancer: Enhanced Molecular Networks by Integrating Metabolome Mining and Annotation Tools. *Metabolites* **9**: 144.

Esnard J, Potter TL & Zuckerman BM (1995) *Streptomyces costaricanus* sp. nov., isolated from nematode-suppressive soil. *Int J Syst Bacteriol* **45**: 775–9.

Evans BA & Amyes SGB (2014) OXA β-lactamases. *Clin Microbiol Rev* **27**: 241–263.

Eze EC, Chenia HY & Zowalaty ME El (2018) Acinetobacter baumannii biofilms: Effects of physicochemical factors, virulence, antibiotic resistance determinants, gene regulation, and future antimicrobial treatments. *Infect Drug Resist* **11**: 2277–2299.

Fair RJ & Tor Y (2014) Antibiotics and bacterial resistance in the 21st century. *Perspect Medicin Chem* **6**: 25–64.

Falke D, Fischer M & Sawers RG (2016) Phosphate and oxygen limitation induce respiratory nitrate reductase 3 synthesis in stationary-phase mycelium of Streptomyces coelicolor A3(2). *Microbiology* **162**: 1689–1697.

Fedorenko V, Genilloud O, Horbal L, Marcone GL, Marinelli F, Paitan Y & Ron EZ (2015) Antibacterial Discovery and Development: From Gene to Product and Back. *Biomed Res Int* **2015**: 1–16.

Férir G, Hänchen A, François KO, Hoorelbeke B, Huskens D, Dettner F, Süssmuth RD & Schols D (2012) Feglymycin, a unique natural bacterial antibiotic peptide, inhibits HIV entry by targeting the viral envelope protein gp120. *Virology* **433**: 308–319.

Fernbach DJ & Martyn DT (1966) Role of Dactinomycin in the Improved Survival of Children With Wilms' Tumor. *JAMA J Am Med Assoc* **195**: 1005–1009.

Findlay KC, Flärdh K & Chater KF (1999) Association of early sporulation genes with suggested developmental decision points in *Streptomyces coelicolor* A3(2). *Microbiology* **145**: 2229–2243.

Fleming-Dutra KE, Hersh AL, Shapiro DJ, Bartoces M, Enns EA, File TM, Finkelstein JA, Gerber JS, Hyun DY, Linder JA, Lynfield R, Margolis DJ, May LS, Merenstein D, Metlay JP, Newland JG, Piccirillo JF, Roberts RM, Sanchez G V., Suda KJ, Thomas A, Woo TM, Zetts RM & Hicks LA (2016) Prevalence of Inappropriate Antibiotic Prescriptions Among US Ambulatory Care Visits, 2010-2011. *JAMA* **315**: 1864.

Fleming A (1929) On the Antibacterial Action of Cultures of a *Penicillium*, with Special Reference to their Use in the Isolation of *B. influenzæ*. *Br J Exp Pathol* **10**: 226–236.

Fleming A (1944) The Discovery of Penicillin. *Br Med Bull* **2**: 4–5.

Frey UH, Bachmann HS, Peters J & Siffert W (2008) PCR-amplification of GC-rich regions: "Slowdown PCR." *Nat Protoc* **3**: 1312–1317.

Frost I, Boeckel TP Van, Pires J, Craig J & Laxminarayan R (2019) Global geographic trends in antimicrobial resistance: the role of international travel. *J Travel Med* .

Galet J, Deveau A, Hôtel L, Leblond P, Frey-Klett P & Aigle B (2014) Gluconic acid-producing Pseudomonas sp. prevent γ-actinorhodin biosynthesis by *Streptomyces coelicolor* A3(2). *Arch Microbiol* **196**: 619–627.

Gáll T, Lehoczki G, Gyémánt G, Emri T, Szigeti ZM, Balla G, Balla J & Pócsi I (2016) Optimization of desferrioxamine E production by Streptomyces parvulus. *Acta Microbiol Immunol Hung* **63**: 475–489.

Gamboa-Suasnavart RA, Valdez-Cruz NA, Gaytan-Ortega G, Reynoso-Cereceda GI, Cabrera-Santos D, López-Griego L, Klöckner W, Büchs J & Trujillo-Roldán MA (2018) The metabolic switch can be activated in a recombinant strain of *Streptomyces lividans* by a low oxygen transfer rate in shake flasks. *Microb Cell Fact* **17**: 189.

Garcia RO, Krug D & Müller R (2009) Chapter 3 Discovering Natural Products from Myxobacteria with Emphasis on Rare Producer Strains in Combination with Improved Analytical Methods. In *Methods in Enzymology*. pp. 59–91.

Garrod LP (1957) The Erythromycin Group Of Antibiotics. *Br Med J* **2**: 57.

Gaynes R (2017) The Discovery of Penicillin—New Insights After More Than 75 Years of Clinical Use. *Emerg Infect Dis* **23**: 849–853.

Gerth K, Pradella S, Perlova O, Beyer S & Müller R (2003) Myxobacteria: Proficient producers of novel natural products with various biological activities - Past and future biotechnological aspects with the focus on the genus Sorangium. *J Biotechnol* **106**: 233–253.

Giglio S, Jiang J, Saint CP, Cane DE & Monis PT (2008) Isolation and characterization of the gene associated with geosmin production in cyanobacteria. *Environ Sci Technol* **42**: 8027–8032.

Glazebrook MA, Doull JL, Stuttard C & Vining LC (1990) Sporulation of *Streptomyces venezuelae* in submerged cultures. *J Gen Microbiol* **136**: 581–588.

Goldstein S, Beka L, Graf J & Klassen JL (2019) Evaluation of strategies for the assembly of diverse bacterial genomes using MinION long-read sequencing. *BMC Genomics* **20**: 23.

Gomes ES, Schuch V & Lemos EG de M (2013) Biotechnology of polyketides: New breath of life for the novel antibiotic genetic pathways discovery through metagenomics. *Brazilian J Microbiol* **44**: 1007–1034.

Gomez-Escribano JP, Alt S & Bibb MJ (2016) Next generation sequencing of actinobacteria for the discovery of novel natural products. *Mar Drugs* **14**: 78.

Gomez-Escribano JP, Song L, Fox DJ, Yeo V, Bibb MJ & Challis GL (2012) Structure and biosynthesis of the unusual polyketide alkaloid coelimycin P1, a metabolic product of the cpk gene cluster of *Streptomyces coelicolor* M145. *Chem Sci* **3**: 2716.

González-Pastor JE, Hobbs EC & Losick R (2003) Cannibalism by sporulating bacteria. *Science (80-)* **301**: 510–3.

Goodfellow M & Fiedler HP (2010) A guide to successful bioprospecting: Informed by actinobacterial systematics. *Antonie van Leeuwenhoek,* **98**: 119–142.

Goodfellow M, Kampfer P, Busse H-J, Trujillo ME, Suzuki K, Ludwig W & Whitman WB (2012) *Bergey's Manual of Systematic Bacteriology*. Springer New York, New York, NY.

Gould K (2016) Antibiotics: From prehistory to the present day. *J Antimicrob Chemother* **71**: 572–575.

Gross JH (2011) Fragmentation of Organic Ions and Interpretation of EI Mass Spectra. In *Mass Spectrometry*. Springer Berlin Heidelberg, Berlin, Heidelberg. pp. 249–350.

Guida R Di, Engel J, Allwood JW, Weber RJM, Jones MR, Sommer U, Viant MR & Dunn WB (2016) Non-targeted UHPLC-MS metabolomic data processing methods: a comparative investigation of normalisation, missing value imputation, transformation and scaling. *Metabolomics* **12**: 93.

Guo F, Xiang S, Li L, Wang B, Rajasärkkä J, Gröndahl-Yli-Hannuksela K, Ai G, Metsä-Ketelä M & Yang K (2015) Targeted activation of silent natural product biosynthesis pathways by reporter-guided mutant selection. *Metab Eng* **28**: 134–142.

Guo Y, Zheng W, Rong X & Huang Y (2008) A multilocus phylogeny of the *Streptomyces griseus* 16S rRNA gene clade: use of multilocus sequence analysis for streptomycete systematics. *Int J Syst Evol Microbiol* **58**: 149–159.

Gurevich A, Saveliev V, Vyahhi N & Tesler G (2013) QUAST: Quality assessment tool for genome assemblies. *Bioinformatics* **29**: 1072–1075.

Gürtler H, Pedersen R, Anthoni U, Christophersen C, Nielsen PH, Wellington EM, Pedersen C & Bock K (1994) Albaflavenone, a sesquiterpene ketone with a zizaene skeleton produced by a streptomycete with a new rope morphology. *J Antibiot (Tokyo)* **47**: 434–9.

Gust B, Challis GL, Fowler K, Kieser T & Chater KF (2003) PCR-targeted *Streptomyces* gene replacement identifies a protein domain needed for biosynthesis of the sesquiterpene soil odor geosmin. *Proc Natl Acad Sci* **100**: 1541–1546.

Hackl S & Bechthold A (2015) The Gene bldA , a Regulator of Morphological Differentiation and Antibiotic Production in *Streptomyces*. *Arch Pharm (Weinheim)* **348**: 455–462.

Hamdache A, Azarken R, Lamarti A, Aleu J & Collado IG (2013) Comparative genome analysis of Bacillus spp. and its relationship with bioactive nonribosomal peptide production. *Phytochem Rev* **12**: 685–716.

Harrison J & Studholme DJ (2014) Recently published *Streptomyces* genome sequences. *Microb Biotechnol* **7**: 373–380.

He JM, Zhu H, Zheng GS, Liu PP, Wang J, Zhao GP, Zhu GQ, Jiang WH & Lu YH (2016) Direct involvement of the master nitrogen metabolism regulator GlnR in antibiotic biosynthesis in *Streptomyces*. *J Biol Chem* **291**: 26443–26454.

Hege-Treskatis D, King R, Wolf H & Gilles ED (1992) Nutritional control of nikkomycin and juglomycin production by *Streptomyces tendae* in continuous culture. *Appl Microbiol Biotechnol* **36**: 440–445.

Hengst CD den, Tran NT, Bibb MJ, Chandra G, Leskiw BK & Buttner MJ (2010) Genes essential for morphological development and antibiotic production in *Streptomyces coelicolor* are targets of BldD during vegetative growth. *Mol Microbiol* **78**: 361–379.

Hentschel J, Burnside C, Mignot I, Leibundgut M, Boehringer D & Ban N (2017)

The Complete Structure of the *Mycobacterium smegmatis* 70S Ribosome. *Cell Rep* **20**: 149–160.

Hesketh A, Chen WJ, Ryding J, Chang S & Bibb MJ (2007) The global role of ppGpp synthesis in morphological differentiation and antibiotic production in *Streptomyces coelicolor* A3(2). *Genome Biol* **8**: R161.

Hesterkamp T (2016) Antibiotics clinical development and pipeline. *Curr Top Microbiol Immunol* **398**: 447–474.

Hinshaw C, Feldman WH & Pfuetze KH (1946) Treatment of tuberculosis with streptomycin: A Summary of Observations on One Hundred Cases. *J Am Med Assoc* **132**: 778–782.

Hobbs G, Frazer CM, Gardner DCJ, Flett F & Oliver SG (1990) Pigmented antibiotic production by Streptomyces coelicolor A3(2): kinetics and the influence of nutrients. *J Gen Microbiol* **136**: 2291–2296.

Hobbs G, Obanye AI, Petty J, Mason JC, Barratt E, Gardner DC, Flett F, Smith CP, Broda P & Oliver SG (1992) An integrated approach to studying regulation of production of the antibiotic methylenomycin by *Streptomyces coelicolor* A3(2). *J Bacteriol* **174**: 1487–1494.

Hobbs Glyn, Obanye AIC, Petty J, Mason JC, Barratt E, Gardner DCJ, Flett F, Smith CP, Broda P & Oliver SG (1992) An integrated approach to studying regulation of production of the antibiotic methylenomycin by Streptomyces coelicolor. *J Bacteriol* **174**: 1487–1494.

Hodgson DA (1982) Glucose repression of carbon source uptake and metabolism in *Streptomyces coelicolor* A3(2) and its perturbation in mutants resistant to 2-deoxyglucose. *Microbiology* **128**: 2417–2430.

Holmes NA, Innocent TM, Heine D, Bassam M Al, Worsley SF, Trottmann F, Patrick EH, Yu DW, Murrell JC, Schiøtt M, Wilkinson B, Boomsma JJ & Hutchings MI (2016) Genome Analysis of Two *Pseudonocardia* Phylotypes Associated with *Acromyrmex* Leafcutter Ants Reveals Their Biosynthetic Potential. *Front Microbiol* **7**: 2073.

Hopwood DA (2006) Soil To Genomics: The *Streptomyces* Chromosome. *Annu Rev Genet* **40**: 1–23.

Hopwood DA (2007) How do antibiotic-producing bacteria ensure their self-resistance before antibiotic biosynthesis incapacitates them? *Mol Microbiol* **63**: 937–940.

Hopwood DA (2013) Imaging Mass Spectrometry Reveals Highly Specific Interactions between Actinomycetes To Activate Specialized Metabolic Gene Clusters. *MBio* **4**: e00612-13-e00612-13.

Hoskisson PA & Wezel GP van (2019) *Streptomyces coelicolor*. *Trends Microbiol* **27**: 468–469.

Hou Y, Tianero MDB, Kwan JC, Wyche TP, Michel CR, Ellis GA, Vazquez-Rivera E, Braun DR, Rose WE, Schmidt EW & Bugni TS (2012) Structure and biosynthesis of the antibiotic bottromycin D. *Org Lett* **14**: 5050–5053.

Houbraken J, Frisvad JC & Samson RA (2011) Fleming's penicillin producing strain is not *Penicillium chrysogenum* but P. rubens. *IMA Fungus* **2**: 87–95.

Hu H, Zhang Q & Ochi K (2002) Activation of Antibiotic Biosynthesis by Specified Mutations in the rpoB Gene (Encoding the RNA Polymerase Subunit) of *Streptomyces lividans*. *J Bacteriol* **184**: 3984–3991.

Hunt AC, Servín-González L, Kelemen GH & Buttner MJ (2005) The bldC developmental locus of *Streptomyces coelicolor* encodes a member of a family of small DNA-binding proteins related to the DNA-binding domains of the MerR family. *J Bacteriol* **187**: 716–728.

Hunter P (2018) The revival of the extended phenotype. *EMBO Rep* **19**: e46477.

Imai Y, Sato S, Tanaka Y, Ochi K & Hosaka T (2015) Lincomycin at Subinhibitory Concentrations Potentiates Secondary Metabolite Production by *Streptomyces* spp. *Appl Environ Microbiol* **81**: 3869–3879.

Inamori Y, Kato Y, Morimoto K, Morisaka K, Saito GI, Sawada Y & Taniyama H (1979) Toxicological Approaches to Streptothricin Antibiotics. III. Biological Studies on Delayed Toxicity of Streptothricin Antibiotics in Rats. *Chem Pharm Bull* **27**: 2570–2576.

Izumikawa M, Shipley PR, Hopke JN, O'Hare T, Xiang L, Noel JP & Moore BS (2003) Expression and characterization of the type III polyketide synthase 1,3,6,8-tetrahydroxynaphthalene synthase from Streptomyces coelicolor A3(2). *J Ind Microbiol Biotechnol* **30**: 510–515.

Jackson S, Crossman L, Almeida E, Margassery L, Kennedy J & Dobson A (2018) Diverse and Abundant Secondary Metabolism Biosynthetic Gene Clusters in the Genomes of Marine Sponge Derived *Streptomyces* spp. Isolates. *Mar Drugs* **16**: 67.

Jakeman DL, Graham CL, Young W & Vining LC (2006) Culture conditions improving the production of jadomycin B. *J Ind Microbiol Biotechnol* **33**: 767–772.

Jamison MT, Wang X, Cheng T & Molinski TF (2019) Synergistic anti-*Candida* activity of bengazole A in the presence of bengamide A. *Mar Drugs* **17**: 102.

Jiang C, Wang H, Kang Q, Liu J & Bai L (2012) Cloning and characterization of the polyether salinomycin biosynthesis gene cluster of *Streptomyces albus* XM211. *Appl Environ Microbiol* **78**: 994–1003.

Jiang J, Sun Y-F, Tang X, He C-N, Shao Y-L, Tang Y-J & Zhou W-W (2018) Alkaline pH shock enhanced production of validamycin A in fermentation of *Streptomyces hygroscopicus*. *Bioresour Technol* **249**: 234–240.

Jones SE, Ho L, Rees CA, Hill JE, Nodwell JR & Elliot MA (2017) *Streptomyces* exploration is triggered by fungal interactions and volatile signals. *Elife* **6**: 1–21.

Jong W De, Manteca A, Sanchez J, Bucca G, Smith CP, Dijkhuizen L, Claessen D & Wösten HAB (2009) NepA is a structural cell wall protein involved in maintenance of spore dormancy in *Streptomyces coelicolor*. *Mol Microbiol* **71**: 1591–1603.

Junker B, Mann Z, Gailliot P, Byrne K & Wilson J (1998) Use of soybean oil and ammonium sulfate additions to optimize secondary metabolite production. *Biotechnol Bioeng* **60**: 580–8.

Kamiyama M & Kaziro Y (1966) Mechanism of action of chromomycin A3. *J Biochem* **59**: 49–56.

Kamoun S (2007) Groovy times: filamentous pathogen effectors revealed. *Curr Opin Plant Biol* **10**: 358–365.

Kapadia M, Rolston KVI & Han XY (2007) Invasive Streptomyces Infections. *Am J Clin Pathol* **127**: 619–624.

Katz E, Pugh LH & Waksman S (1956) Antibiotic and cytostatic properties of the actinomycins. *J Bacteriol* **72**: 660–5.

Katz L & Baltz RH (2016) Natural product discovery: past, present, and future. *J Ind Microbiol Biotechnol* **43**: 155–176.

Kawai K, Wang G, Okamoto S & Ochi K (2007) The rare earth, scandium, causes antibiotic overproduction in *Streptomyces* spp. *FEMS Microbiol Lett* **274**: 311–315.

Kawano T, Hidaka T, Furihata K, Mochizuki J, Nakayama H, Hayakawa Y & Seto H (1990) Isolation and structures of mono- and di-deacetyl chromomycin antibiotics 02-3D and 02-3G from *Streptomyces avellaneus*. *J Antibiot (Tokyo)* **43**: 110–3.

Kelemen GH, Brian P, Flärdh K, Chamberlin L, Chater KF & Buttner MJ (1998) Developmental regulation of transcription of whiE, a locus specifying the polyketide spore pigment in *Streptomyces coelicolor* A3 (2). *J Bacteriol* **180**: 2515–21.

Kellogg JJ, Todd DA, Egan JM, Raja HA, Oberlies NH, Kvalheim OM & Cech NB (2016) Biochemometrics for Natural Products Research: Comparison of Data Analysis Approaches and Application to Identification of Bioactive Compounds. *J Nat Prod* **79**: 376–386.

Keulen G van, Alderson J, White J & Sawers RG (2007) The obligate aerobic actinomycete *Streptomyces coelicolor* A3(2) survives extended periods of anaerobic stress. *Environ Microbiol* **9**: 3143–3149.

Keulen G van & Dyson PJ (2014) Production of Specialized Metabolites by </i>Streptomyces coelicolor</i> A3(2). In *Advances in Applied Microbiology*. pp. 217–266.

Kieser HM, Kieser T & Hopwood DA (1992) A combined genetic and physical map of the *Streptomyces coelicolor* A3(2) chromosome. *J Bacteriol* **174**: 5496–507.

Kieser T, Bibb M, Buttner M, Chater K & Hopwood D (2000) *Practical* Streptomyces *Genetics*. John Innes Foundation, Norwich, UK.

Kim CY, Park HJ & Kim ES (2005) Proteomics-driven identification of putative AfsR2-target proteins stimulating antibiotic biosynthesis in *Streptomyces lividans*. *Biotechnol Bioprocess Eng* **10**: 248–253.

Kim ES, Hong HJ, Choi CY & Cohen SN (2001) Modulation of actinorhodin biosynthesis in *Streptomyces lividans* by glucose repression of afsR2 gene transcription. *J Bacteriol* **183**: 2198–2203.

Kim JN, Kim Y, Jeong Y, Roe JH, Kim BG & Cho BK (2015) Comparative genomics reveals the core and accessory genomes of *Streptomyces* species. *J Microbiol Biotechnol* **25**: 1599–1605.

Kim YJ, Song JY, Moon MH, Smith CP, Hong S-K & Chang YK (2007) pH shock induces overexpression of regulatory and biosynthetic genes for actinorhodin productionin *Streptomyces coelicolor* A3(2). *Appl Microbiol Biotechnol* **76**: 1119–1130.

Kodani S, Hudson ME, Durrant MC, Buttner MJ, Nodwell JR & Willey JM (2004) The SapB morphogen is a lantibiotic-like peptide derived from the product of the developmental gene ramS in *Streptomyces coelicolor*. *Proc Natl Acad Sci U S A* **101**: 11448–11453.

Kol S, Merlo ME, Scheltema RA, Vries M de, Vonk RJ, Kikkert NA, Dijkhuizen L, Breitling R & Takano E (2010) Metabolomic Characterization of the Salt Stress Response in *Streptomyces coelicolor*. *Appl Environ Microbiol* **76**: 2574–2581.

Konopka JB (2012) N-Acetylglucosamine Functions in Cell Signaling. *Scientifica (Cairo)* **2012**: 1–15.

Kontro M, Lignell U, Hirvonen MR & Nevalainen A (2005) pH effects on 10 *Streptomyces* spp. growth and sporulation depend on nutrients. *Lett Appl Microbiol* **41**: 32–38.

Koren S, Schatz MC, Walenz BP, Martin J, Howard JT, Ganapathy G, Wang Z, Rasko DA, McCombie WR, Jarvis ED & Phillippy AM (2012) Hybrid error correction and de novo assembly of single-molecule sequencing reads. *Nat Biotechnol* **30**: 693–700.

Kosa G, Vuoristo KS, Horn SJ, Zimmermann B, Afseth NK, Kohler A & Shapaval V (2018) Assessment of the scalability of a microtiter plate system for screening of oleaginous microorganisms. *Appl Microbiol Biotechnol* **102**: 4915–4925.

Kosa G, Zimmermann B, Kohler A, Ekeberg D, Afseth NK, Mounier J & Shapaval V (2018) High-throughput screening of *Mucoromycota* fungi for production of low- and high-value lipids. *Biotechnol Biofuels* **11**: 66.

Krug D & Müller R (2014) Secondary metabolomics: the impact of mass spectrometry-based approaches on the discovery and characterization of microbial natural products. *Nat Prod Rep* **31**: 768.

Kumar J, Sharma VK, Singh DK, Mishra A, Gond SK, Verma SK, Kumar A & Kharwar RN (2016) Epigenetic Activation of Antibacterial Property of an Endophytic *Streptomyces coelicolor* Strain AZRA 37 and Identification of the Induced Protein Using MALDI TOF MS/MS. *PLoS One* **11**: 1–12.

Kumarasamy KK, Toleman MA, Walsh TR, Bagaria J, Butt F, Balakrishnan R, Chaudhary U, Doumith M, Giske CG, Irfan S, Krishnan P, Kumar A V, Maharjan S, Mushtaq S, Noorie T, Paterson DL, Pearson A, Perry C, Pike R, Rao B, Ray U, Sarma JB, Sharma M, Sheridan E, Thirunarayan MA, Turton J, Upadhyay S,

Warner M, Welfare W, Livermore DM & Woodford N (2010) Emergence of a new antibiotic resistance mechanism in India, Pakistan, and the UK: A molecular, biological, and epidemiological study. *Lancet Infect Dis* **10**: 597–602.

Kumburu HH, Sonda T, Zwetselaar M van, Leekitcharoenphon P, Lukjancenko O, Mmbaga BT, Alifrangis M, Lund O, Aarestrup FM & Kibiki GS (2019) Using WGS to identify antibiotic resistance genes and predict antimicrobial resistance phenotypes in MDR *Acinetobacter baumannii* in Tanzania. *J Antimicrob Chemother* **74**: 1484–1493.

Kurosawa K, Bui VP, VanEssendelft JL, Willis LB, Lessard PA, Ghiviriga I, Sambandan TG, Rha CK & Sinskey AJ (2006) Characterization of *Streptomyces* MITKK-103, a newly isolated actinomycin X2-producer. *Appl Microbiol Biotechnol* **72**: 145–154.

Kutzner HJ & Waksman S (1959) *Streptomyces coelicolor* Mueller and *Streptomyces violaceoruber* Waksman and Curtis, two distinctly different organisms. *J Bacteriol* **78**: 528–38.

Kwon NH, Park KT, Moon JS, Jung WK, Kim SH, Kim JM, Hong SK, Koo HC, Joo YS & Park YH (2005) Staphylococcal cassette chromosome mec (SCCmec) characterization and molecular analysis for methicillin-resistant *Staphylococcus aureus* and novel SCCmec subtype IVg isolated from bovine milk in Korea. *J Antimicrob Chemother* **56**: 624–632.

Labeda DP, Goodfellow M, Brown R, Ward AC, Lanoot B, Vanncanneyt M, Swings J, Kim S-B, Liu Z, Chun J, Tamura T, Oguchi A, Kikuchi T, Kikuchi H, Nishii T, Tsuji K, Yamaguchi Y, Tase A, Takahashi M, Sakane T, Suzuki KI & Hatano K (2012) Phylogenetic study of the species within the family *Streptomycetaceae*. *Antonie Van Leeuwenhoek* **101**: 73–104.

Lautru S, Deeth RJ, Bailey LM & Challis GL (2005) Discovery of a new peptide natural product by *Streptomyces coelicolor* genome mining. *Nat Chem Biol* **1**: 265–269.

Lee CR, Lee JH, Park M, Park KS, Bae IK, Kim YB, Cha CJ, Jeong BC & Lee SH (2017) Biology of *Acinetobacter baumannii*: Pathogenesis, antibiotic resistance mechanisms, and prospective treatment options. *Front Cell Infect Microbiol* **7**: 55.

Lee MY, Ames BD & Tsai SC (2012) Insight into the molecular basis of aromatic polyketide cyclization: Crystal structure and in vitro characterization of WhiE-ORFVI. *Biochemistry* **51**: 3079–3091.

Lewis K (2013) Platforms for antibiotic discovery. *Nat Rev Drug Discov* **12**: 371–387.

Li J, Zhang L & Liu W (2018) Cell-free synthetic biology for in vitro biosynthesis of pharmaceutical natural products. *Synth Syst Biotechnol* **3**: 83–89.

Li T, Du Y, Cui Q, Zhang J, Zhu W, Hong K & Li W (2013) Cloning, characterization and heterologous expression of the indolocarbazole biosynthetic gene cluster from marine-derived *Streptomyces sanyensis* FMA. *Mar Drugs* **11**: 466–488.

Li W, Ying X, Guo Y, Yu Z, Zhou X, Deng Z, Kieser H, Chater KF & Tao M (2006) Identification of a gene negatively affecting antibiotic production and morphological differentiation in *Streptomyces coelicolor* A3(2). *J Bacteriol* **188**: 8368–8375.

Liao Y, Wei Z-H, Bai L, Deng Z & Zhong J-J (2009) Effect of fermentation temperature on validamycin A production by *Streptomyces hygroscopicus* 5008. *J Biotechnol* **142**: 271–274.

Lin PF, Samanta H, Bechtold CM, Deminie CA, Patick AK, Alam M, Riccardi K, Rose RE, White RJ & Colonno RJ (1996) Characterization of siamycin I, a human immunodeficiency virus fusion inhibitor. *Antimicrob Agents Chemother* **40**: 133–138.

Lin X, Hopson R & Cane DE (2006) Genome Mining in Streptomyces coelicolor: Molecular Cloning and Characterization of a New Sesquiterpene Synthase. *J Am Chem Soc* **128**: 6022–6023.

Ling LL, Schneider T, Peoples AJ, Spoering AL, Engels I, Conlon BP, Mueller A, Schäberle TF, Hughes DE, Epstein S, Jones M, Lazarides L, Steadman VA, Cohen DR, Felix CR, Fetterman KA, Millett WP, Nitti AG, Zullo AM, Chen C & Lewis K (2015) A new antibiotic kills pathogens without detectable resistance. *Nature* **517**: 455–459.

Liu G, Chater KF, Chandra G, Niu G & Tan H (2013) Molecular Regulation of Antibiotic Biosynthesis in Streptomyces. *Microbiol Mol Biol Rev* **77**: 112–143.

Liu M, Jia Y, Xie Y, Zhang C, Ma J, Sun C & Ju J (2019) Identification of the Actinomycin D Biosynthetic Pathway from Marine-Derived *Streptomyces costaricanus* SCSIO ZS0073. *Mar Drugs* **17**: 240.

Liu X & Cheng YQ (2014) Genome-guided discovery of diverse natural products from *Burkholderia* sp. *J Ind Microbiol Biotechnol* **41**: 275–284.

Liu XF, Xiang L, Zhou Q, Carralot JP, Prunotto M, Niederfellner G & Pastan I (2016) Actinomycin D enhances killing of cancer cells by immunotoxin RG7787 through activation of the extrinsic pathway of apoptosis. *Proc Natl Acad Sci U S A* **113**: 10666–10671.

Lo Y-S, Tseng W-H, Chuang C-Y & Hou M-H (2013) The structural basis of actinomycin D–binding induces nucleotide flipping out, a sharp bend and a left-handed twist in CGG triplet repeats. *Nucleic Acids Res* **41**: 4284–4294.

Luna BM, Ulhaq A, Yan J, Pantapalangkoor P, Nielsen TB, Davies BW, Actis LA & Spellberg B (2017) Selectable Markers for Use in Genetic Manipulation of Extensively Drug-Resistant (XDR) *Acinetobacter baumannii* HUMC1. *mSphere* **2**.

Luo Y, Huang H, Liang J, Wang M, Lu L, Shao Z, Cobb RE & Zhao H (2013) Activation and characterization of a cryptic polycyclic tetramate macrolactam biosynthetic gene cluster. *Nat Commun* **4**: 2894.

Luti KJK & Mavituna F (2011) Elicitation of *Streptomyces coelicolor* with dead cells of *Bacillus subtilis* and *Staphylococcus aureus* in a bioreactor increases production of undecylprodigiosin. *Appl Microbiol Biotechnol* **90**: 461–466.

Luzzatto L, Apirion D & Schlessinger D (1968) Mechanism of action of

streptomycin in *E. coli*: interruption of the ribosome cycle at the initiation of protein synthesis. *Proc Natl Acad Sci U S A* **60**: 873–880.

Macleod A, Ross H, Ozere R, Digout G & Rooyen C van (1964) Lincomycin: A New Antibiotic Active Against Staphylococci and Other Gram-Positive Cocci. *Can Med Assoc J* **91**: 1056–60.

Maeda K, Ueda M, Yagishita K, Kawaji S, Kondo S, Murase M, Takeuchi T, Okami Y & Umezawa H (1957) Studies on kanamycin. *J Antibiot (Tokyo)* **10**: 228–31.

Mak S & Nodwell JR (2017) Actinorhodin is a redox-active antibiotic with a complex mode of action against Gram-positive cells. *Mol Microbiol* **106**: 597–613.

Manteca Á, Claessen D, Lopez-Iglesias C & Sanchez J (2007) Aerial hyphae in surface cultures of *Streptomyces lividans* and *Streptomyces coelicolor* originate from viable segments surviving an early programmed cell death event. *FEMS Microbiol Lett* **274**: 118–125.

Manteca Á, Fernandez M & Sanchez J (2005) Mycelium development in *Streptomyces antibioticus* ATCC11891 occurs in an orderly pattern which determines multiphase growth curves. *BMC Microbiol* **5**: 51.

Manteca Á, Fernández M & Sánchez J (2005) A death round affecting a young compartmentalized mycelium precedes aerial mycelium dismantling in confluent surface cultures of *Streptomyces antibioticus*. *Microbiology* **151**: 3689–3697.

Manteca Á & Sanchez J (2009) *Streptomyces* development in colonies and soils. *Appl Environ Microbiol* **75**: 2920–2924.

Manteca Á, Sanchez J, Jung HR, Schwämmle V & Jensen ON (2010) Quantitative proteomics analysis of *Streptomyces coelicolor* development demonstrates that onset of secondary metabolism coincides with hypha differentiation. *Mol Cell Proteomics* **9**: 1423–1436.

Manteca Á & Yagüe P (2018) *Streptomyces* differentiation in liquid cultures as a trigger of secondary metabolism. *Antibiotics* **7**: 1–13.

Marinelli F (2009) Chapter 2 From Microbial Products to Novel Drugs that Target a Multitude of Disease Indications. In *Methods in Enzymology*. Elsevier Inc., pp. 29–58.

Marti E, Jofre J & Balcazar JL (2013) Prevalence of Antibiotic Resistance Genes and Bacterial Community Composition in a River Influenced by a Wastewater Treatment Plant. *PLoS One* **8**.

Mattheis JP & Roberts RG (1992) Identification of geosmin as a volatile metabolite of *Penicillium expansum*. *Appl Environ Microbiol* **58**: 3170–2.

McBride MJ & Ensign JC (1987) Effects of intracellular trehalose content on *Streptomyces griseus* spores. *J Bacteriol* **169**: 4995–5001.

McCormick JR & Flärdh K (2012) Signals and regulators that govern *Streptomyces* development. *FEMS Microbiol Rev* **36**: 206–231.

McKenzie NL, Thaker M, Koteva K, Hughes DW, Wright GD & Nodwell JR (2010)

Induction of antimicrobial activities in heterologous streptomycetes using alleles of the *Streptomyces coelicolor* gene absA1. *J Antibiot (Tokyo)* **63**: 177–182.

McMahon MD, Rush JS & Thomas MG (2012) Analyses of MbtB, MbtE, and MbtF Suggest Revisions to the Mycobactin Biosynthesis Pathway in *Mycobacterium tuberculosis*. *J Bacteriol* **194**: 2809–2818.

Medema MH, Blin K, Cimermancic P, Jager V De, Zakrzewski P, Fischbach MA, Weber T, Takano E & Breitling R (2011) AntiSMASH: Rapid identification, annotation and analysis of secondary metabolite biosynthesis gene clusters in bacterial and fungal genome sequences. *Nucleic Acids Res* **39**: 339–346.

Medema MH, Kottmann R, Yilmaz P, Cummings M, Biggins JB, Blin K, Bruijn I de, Chooi YH, Claesen J, Glöckner FO, *et al.* (2015) Minimum Information about a Biosynthetic Gene cluster. *Nat Chem Biol* **11**: 625–631.

Mediavilla JR, Chen L, Mathema B & Kreiswirth BN (2012) Global epidemiology of community-associated methicillin resistant *Staphylococcus aureus* (CA-MRSA). *Curr Opin Microbiol* **15**: 588–595.

Meroueh SO, Minasov G, Lee W, Shoichet BK & Mobashery S (2003) Structural Aspects for Evolution of β-Lactamases from Penicillin-Binding Proteins. *J Am Chem Soc* **125**: 9612–9618.

Miller EL (2002) The penicillins: A review and update. *J Midwifery Women's Heal* **47**: 426–434.

Mindlin SZ, Soina VS, Ptrova MA & Gorlenko ZM (2008) Isolation of antibiotic resistance bacterial strains from East Siberia permafrost sediments. *Russ J Genet* **44**: 36–44.

Mocek U, Zeng Z, Beale JM, Floss HG, Mocek U, Zeng Z, O'Hagan D, Zhou P, Fan LDG, Beale JM, Floss HG, O'Hagan D, Zhou P, Fan LDG & Beale JM (1993) Biosynthesis of the Modified Peptide Antibiotic Thiostrepton in *Streptomyces azureus* and *Streptomyces laurentii*. *J Am Chem Soc* **115**: 7992–8001.

Moore JM, Bradshaw E, Seipke RF, Hutchings MI & McArthur M (2012) Use and discovery of chemical elicitors that stimulate biosynthetic gene clusters in *Streptomyces* bacteria. *Methods Enzymol* **517**: 367–385.

Morgenstern A, Paetz C, Behrend A & Spiteller D (2015) Divalent Transition-Metal-Ion Stress Induces Prodigiosin Biosynthesis in *Streptomyces coelicolor* M145: Formation of Coeligiosins. *Chem - A Eur J* **21**: 6027–6032.

Murphy E, Huwyler L & Freire Bastos M do C de (1985) Transposon Tn554: complete nucleotide sequence and isolation of transposition-defective and antibiotic-sensitive mutants. *EMBO J* **4**: 3357–3365.

Murphy T, Parra R, Radman R, Roy I, Harrop A, Dixon K & Keshavarz T (2007) Novel application of oligosaccharides as elicitors for the enhancement of bacitracin A production in cultures of *Bacillus licheniformis*. *Enzyme Microb Technol* **40**: 1518–1523.

Musiol-Kroll E & Wohlleben W (2018) Acyltransferases as Tools for Polyketide Synthase Engineering. *Antibiotics* **7**: 62.

Myers OD, Sumner SJ, Li S, Barnes S & Du X (2017) One Step Forward for Reducing False Positive and False Negative Compound Identifications from Mass Spectrometry Metabolomics Data: New Algorithms for Constructing Extracted Ion Chromatograms and Detecting Chromatographic Peaks. *Anal Chem* **89**: 8696–8703.

Nagaoka K, Matsumoto M, Oono J, Yokoi K, Ishizeki S & Nakashima T (1986) Azinomycins A and B, new antitumor antibiotics. I. Producing organism, fermentation, isolation, and characterization. *J Antibiot (Tokyo)* **39**: 1527–32.

NCIMB Ltd (2017) NCIMB. http://www.ncimb.com/. Accessed October 9, 2017.

Nepal G & Bhatta S (2018) Self-medication with Antibiotics in WHO Southeast Asian Region: A Systematic Review. *Cureus* .

Nikolouli K & Mossialos D (2012) Bioactive compounds synthesized by non-ribosomal peptide synthetases and type-I polyketide synthases discovered through genome-mining and metagenomics. *Biotechnol Lett* **34**: 1393–1403.

Niu G, Chater KF, Tian Y, Zhang J & Tan H (2016) Specialised metabolites regulating antibiotic biosynthesis in *Streptomyces* spp. *FEMS Microbiol Rev* **40**: 554–573.

Nothias LF, Nothias-Esposito M, Silva R Da, Wang M, Protsyuk I, Zhang Z, Sarvepalli A, Leyssen P, Touboul D, Costa J, Paolini J, Alexandrov T, Litaudon M & Dorrestein PC (2018) Bioactivity-Based Molecular Networking for the Discovery of Drug Leads in Natural Product Bioassay-Guided Fractionation. *J Nat Prod* **81**: 758–767.

Núñez-Montero K, Lamilla C, Abanto M, Maruyama F, Jorquera MA, Santos A, Martinez-Urtaza J & Barrientos L (2019) Antarctic *Streptomyces fildesensis* So13.3 strain as a promising source for antimicrobials discovery. *Sci Rep* **9**: 7488.

O'Neill J (2016) Tackling drug-resistant infections globally: final report and recommendations. The review on antimicrobial resistance. .

Ochi K, Okamoto S, Tozawa Y, Inaoka T, Hosaka T, Xu J & Kurosawa K (2004) Ribosome Engineering and Secondary Metabolite Production. In *Advances in Applied Microbiology*. pp. 155–184.

Ogawara H (1981) Antibiotic resistance in pathogenic and producing bacteria, with special reference to beta-lactam antibiotics. *Microbiol Rev* **45**: 591–619.

Ogawara H (2016) Self-resistance in Streptomyces, with special reference to β-lactam antibiotics. *Molecules* **21**.

Onaka H, Mori Y, Igarashi Y & Furumai T (2011) Mycolic acid-containing bacteria induce natural-product biosynthesis in Streptomyces species. *Appl Environ Microbiol* **77**: 400–406.

Overbeek R, Begley T, Butler RM, Choudhuri J V, Chuang H-Y, Cohoon M, Crécy-Lagard V de, Diaz N, Disz T, Edwards R, Fonstein M, Frank ED, Gerdes S, Glass EM, Goesmann A, Hanson A, Iwata-Reuyl D, Jensen R, Jamshidi N, Krause L, Kubal M, Larsen N, Linke B, McHardy AC, Meyer F, Neuweger H, Olsen G, Olson R, Osterman A, Portnoy V, Pusch GD, Rodionov DA, Rückert C, Steiner J, Stevens

R, Thiele I, Vassieva O, Ye Y, Zagnitko O & Vonstein V (2005) The subsystems approach to genome annotation and its use in the project to annotate 1000 genomes. *Nucleic Acids Res* **33**: 5691–702.

Papp-Wallace KM, Nguyen NQ, Jacobs MR, Bethel CR, Barnes MD, Kumar V, Bajaksouzian S, Rudin SD, Rather PN, Bhavsar S, Ravikumar T, Deshpande PK, Patil V, Yeole R, Bhagwat SS, Patel M V., Akker F Van Den & Bonomo RA (2018) Strategic Approaches to Overcome Resistance against Gram-Negative Pathogens Using β-Lactamase Inhibitors and β-Lactam Enhancers: Activity of Three Novel Diazabicyclooctanes WCK 5153, Zidebactam (WCK 5107), and WCK 4234. *J Med Chem* **61**: 4067–4086.

Parks DH, Imelfort M, Skennerton CT, Hugenholtz P & Tyson GW (2015) CheckM: assessing the quality of microbial genomes recovered from isolates, single cells, and metagenomes. *Genome Res* **25**: 1043–1055.

Pathom-aree W, Stach JEM, Ward AC, Horikoshi K, Bull AT & Goodfellow M (2006) Diversity of actinomycetes isolated from Challenger Deep sediment (10,898 m) from the Mariana Trench. *Extremophiles* **10**: 181–189.

Pedersen SS, Jensen T, Osterhammel D & Osterhammel P (1987) Cumulative and acute toxicity of repeated high-dose tobramycin treatment in cystic fibrosis. *Antimicrob Agents Chemother* **31**: 594–599.

Pereira F, Latino DARS & Gaudêncio S (2014) A chemoinformatics approach to the discovery of lead-like molecules from marine and microbial sources en route to antitumor and antibiotic drugs. *Mar Drugs* **12**: 757–778.

Petkovic H, Cullum J, Hranueli D, Hunter IS, Peric-Concha N, Pigac J, Thamchaipenet A, Vujaklija D & Long PF (2006) Genetics of *Streptomyces rimosus*, the Oxytetracycline Producer. *Microbiol Mol Biol Rev* **70**: 704–728.

Pettit RK, Weber CA, Kean MJ, Hoffmann H, Pettit GR, Tan R, Franks KS & Horton ML (2005) Microplate Alamar Blue Assay for *Staphylococcus epidermidis* Biofilm Susceptibility Testing. *Antimicrob Agents Chemother* **49**: 2612–2617.

Pishchany G, Mevers E, Ndousse-Fetter S, Horvath DJ, Paludo CR, Silva-Junior EA, Koren S, Skaar EP, Clardy J & Kolter R (2018) Amycomicin is a potent and specific antibiotic discovered with a targeted interaction screen. *Proc Natl Acad Sci* **115**: 10124–10129.

Pluskal T, Castillo S, Villar-Briones A & Orešič M (2010) MZmine 2: Modular framework for processing, visualizing, and analyzing mass spectrometry-based molecular profile data. *BMC Bioinformatics* **11**: 395.

Polkade A V., Mantri SS, Patwekar UJ & Jangid K (2016) Quorum Sensing: An Under-Explored Phenomenon in the Phylum *Actinobacteria*. *Front Microbiol* **7**.

Pragasam AK, Shankar C, Veeraraghavan B, Biswas I, Nabarro LEB, Inbanathan FY, George B & Verghese S (2016) Molecular Mechanisms of Colistin Resistance in *Klebsiella pneumoniae* Causing Bacteremia from India-A First Report. *Front Microbiol* **7**: 2135.

Prajapati D, Kumari N, Dave K, Chatupale V & Pohnerkar J (2019) Chromomycin, an antibiotic produced by *Streptomyces flaviscleroticus* might play a role in the

resistance to oxidative stress and is essential for viability in stationary phase. *Environ Microbiol* **21**: 814–826.

Praveen V & Tripathi CKM (2009) Studies on the production of actinomycin-D by Streptomyces *griseoruber* - A novel source. *Lett Appl Microbiol* **49**: 450–455.

Purves K, Macintyre L, Brennan D, Hreggviðsson G, Kuttner E, Ásgeirsdóttir M, Young L, Green D, Edrada-Ebel R & Duncan K (2016) Using Molecular Networking for Microbial Secondary Metabolite Bioprospecting. *Metabolites* **6**: 2.

Quail MA, Smith M, Coupland P, Otto TD, Harris SR, Connor TR, Bertoni A, Swerdlow HP & Gu Y (2012) A tale of three next generation sequencing platforms: comparison of Ion Torrent, Pacific Biosciences and Illumina MiSeq sequencers. *BMC Genomics* **13**: 341.

Rammelkamp CH & Weinstein L (1942) Toxic effects of tyrothricin, gramicidin and tyrocidine. *J Infect Dis* **71**: 166–173.

Rang FJ, Kloosterman WP & Ridder J de (2018) From squiggle to basepair: computational approaches for improving nanopore sequencing read accuracy. *Genome Biol* **19**: 90.

Rao M, Feng L, Ruan L, Ge M & Sheng X (2013) UPLC-MS-Based Metabolomic Study of *Streptomyces* Strain HCCB10043 Under Different pH Conditions Reveals Important Pathways Affecting the Biosynthesis of A21978C Compounds. *Anal Lett* **46**: 2305–2318.

Rao M, Li Q, Feng L, Xia X, Ruan L, Sheng X & Ge M (2011) A new aminopeptidase inhibitor from *Streptomyces* strain HCCB10043 found by UPLC–MS. *Anal Bioanal Chem* **401**: 699–706.

Raper KB, Alexander DF & Coghill RD (1944) Penicillin: II. Natural Variation and Penicillin Production in *Penicillium notatum* and Allied Species. *J Bacteriol* **48**: 639–63959.

Rateb ME, Houssen WE, Harrison WTA, Deng H, Okoro CK, Asenjo JA, Andrews BA, Bull AT, Goodfellow M, Ebel R & Jaspars M (2011) Diverse Metabolic Profiles of a *Streptomyces* Strain Isolated from a Hyper-arid Environment. *J Nat Prod* **74**: 1965–1971.

Recio E, Colinas Á, Rumbero Á, Aparicio JF & Martín JF (2004) PI factor, a novel type quorum-sensing inducer elicits pimaricin production in *Streptomyces natalensis*. *J Biol Chem* **279**: 41586–41593.

Redhu AK, Shah AH & Prasad R (2016) MFS transporters of Candida species and their role in clinical drug resistance. *FEMS Yeast Res* **16**: 1–12.

Reen FJ, Romano S, Dobson ADW & O'Gara F (2015) The sound of silence: Activating silent biosynthetic gene clusters in marine microorganisms. *Mar Drugs* **13**: 4754–4783.

Reuter S, Manfras B, Merkle M, Harter G & Kern P (2006) In Vitro Activities of Itraconazole, Methiazole, and Nitazoxanide versus *Echinococcus multilocularis* Larvae. *Antimicrob Agents Chemother* **50**: 2966–2970.

Rigali S, Nothaft H, Noens EEE, Schlicht M, Colson S, Müller M, Joris B, Koerten

HK, Hopwood DA, Titgemeyer F & Wezel GP Van (2006) The sugar phosphotransferase system of Streptomyces coelicolor is regulated by the GntR-family regulator DasR and links N-acetylglucosamine metabolism to the control of development. *Mol Microbiol* **61**: 1237–1251.

Rigali S, Titgemeyer F, Barends S, Mulder S, Thomae AW, Hopwood DA & Wezel GP van (2008) Feast or famine: the global regulator DasR links nutrient stress to antibiotic production by *Streptomyces*. *EMBO Rep* **9**: 670–675.

Risdian C, Mozef T & Wink J (2019) Biosynthesis of polyketides in *Streptomyces*. *Microorganisms* **7**.

Romero-Rodríguez A, Ruiz-Villafán B, Tierrafría VH, Rodríguez-Sanoja R & Sánchez S (2016) Carbon Catabolite Regulation of Secondary Metabolite Formation and Morphological Differentiation in *Streptomyces coelicolor*. *Appl Biochem Biotechnol* **180**: 1152–1166.

Rosenblatt-Farrell N (2009) The landscape of antibiotic resistance. *Environ Health Perspect* **117**.

Ruiz B, Chávez A, Forero A, García-Huante Y, Romero A, Snchez M, Rocha D, Snchez B, Rodríguez-Sanoja R, Sánchez S & Langley E (2010) Production of microbial secondary metabolites: Regulation by the carbon source. *Crit Rev Microbiol* **36**: 146–167.

Rutledge PJ & Challis GL (2015) Discovery of microbial natural products by activation of silent biosynthetic gene clusters. *Nat Rev Microbiol* **13**: 509–523.

Salerno P, Persson J, Bucca G, Laing E, Ausmees N, Smith CP & Flärdh K (2013) Identification of new developmentally regulated genes involved in *Streptomyces coelicolor* sporulation. *BMC Microbiol* **13**: 281.

Samson RA, Hadlok R & Stolk AC (1977) A taxonomic study of the *Penicillium chrysogenum* series. *Antonie Van Leeuwenhoek* **43**: 169–175.

Sancelme M, Fabre S & Prudhomme M (1994) Antimicrobial activities of indolocarbazole and bis-indole protein kinase C inhibitors. *J Antibiot (Tokyo)* **47**: 792–8.

Sánchez S, Chávez A, Forero A, García-Huante Y, Romero A, Sánchez M, Rocha D, Sánchez B, Avalos M, Guzmán-Trampe S, Rodríguez-Sanoja R, Langley E & Ruiz B (2010) Carbon source regulation of antibiotic production. *J Antibiot (Tokyo)* **63**: 442–459.

Santillana E, Beceiro A, Bou G & Romero A (2007) Crystal structure of the carbapenemase OXA-24 reveals insights into the mechanism of carbapenem hydrolysis. *Proc Natl Acad Sci U S A* **104**: 5354–5359.

Santos-Beneit F, Rodríguez-García A, Sola-Landa A & Martín JF (2009) Cross-talk between two global regulators in *Streptomyces*: PhoP and AfsR interact in the control of afsS, pstS and phoRP transcription. *Mol Microbiol* **72**: 53–68.

Schatz A, Bugle E & Waksman S (1944) Streptomycin, a Substance Exhibiting Antibiotic Activity Against Gram-Positive and Gram-Negative Bacteria. *Exp Biol Med* **55**: 66–69.

Schrempf H, Koebsch I, Walter S, Engelhardt H & Meschke H (2011) Extracellular *Streptomyces* vesicles: Amphorae for survival and defence. *Microb Biotechnol* **4**: 286–299.

Scott D & Ely B (2015) Comparison of Genome Sequencing Technology and Assembly Methods for the Analysis of a GC-Rich Bacterial Genome. *Curr Microbiol* **70**: 338–344.

Seipke RF (2015) Strain-Level Diversity of Secondary Metabolism in *Streptomyces albus*. *PLoS One* **10**: e0116457.

Seipke RF, Kaltenpoth M & Hutchings MI (2012) *Streptomyces* as symbionts: an emerging and widespread theme? *FEMS Microbiol Rev* **36**: 862–876.

Seipke RF & Loria R (2009) Hopanoids are not essential for growth of *Streptomyces scabies* 87-22. *J Bacteriol* **191**: 5216–5223.

Senges CHR, Al-Dilaimi A, Marchbank DH, Wibberg D, Winkler A, Haltli B, Nowrousian M, Kalinowski J, Kerr RG & Bandow JE (2018) The secreted metabolome of Streptomyces chartreusis and implications for bacterial chemistry. *Proc Natl Acad Sci* **115**: 2490–2495.

Sevcikova B & Kormanec J (2004) Differential production of two antibiotics of *Streptomyces coelicolor* A3(2), actinorhodin and undecylprodigiosin, upon salt stress conditions. *Arch Microbiol* **181**: 384–389.

Seyedsayamdost MR (2014) High-throughput platform for the discovery of elicitors of silent bacterial gene clusters. *Proc Natl Acad Sci* **111**: 7266–7271.

Shannon P, Markiel A, Ozier O, Baliga NS, Wang JT, Ramage D, Amin N, Schwikowski B & Ideker T (2003) Cytoscape: A software Environment for integrated models of biomolecular interaction networks. *Genome Res* **13**: 2498–2504.

Shentu X, Liu N, Tang G, Tanaka Y, Ochi K, Xu J & Yu X (2016) Improved antibiotic production and silent gene activation in *Streptomyces diastatochromogenes* by ribosome engineering. *J Antibiot (Tokyo)* **69**: 406–410.

Shepherd MD, Kharel MK, Bosserman MA & Rohr J (2010) Laboratory Maintenance of *Streptomyces* Species. *Curr Protoc Microbiol* **10**: Unit 10E.1.

Shimomura I, Yokoi A, Kohama I, Kumazaki M, Tada Y, Tatsumi K, Ochiya T & Yamamoto Y (2019) Drug library screen reveals benzimidazole derivatives as selective cytotoxic agents for KRAS-mutant lung cancer. *Cancer Lett* **451**: 11–22.

Shin D, Byun WS, Moon K, Kwon Y, Bae M, Um S, Lee SK & Oh DC (2018) Coculture of marine *Streptomyces* sp. with *Bacillus* sp. produces a new piperazic acid-bearing cyclic peptide. *Front Chem* **6**.

Shirling EB & Gottlieb D (1966) Methods for characterization of *Streptomyces* species. *Int J Syst Bacteriol* **16**: 313–340.

Shulse CN & Allen EE (2011) Widespread occurrence of secondary lipid biosynthesis potential in microbial lineages. *PLoS One* **6**.

Sigle S, Steblau N, Wohlleben W & Muth G (2016) Polydiglycosylphosphate Transferase PdtA (SCO2578) of *Streptomyces coelicolor* A3(2) Is Crucial for Proper Sporulation and Apical Tip Extension under Stress Conditions. *Appl Environ Microbiol* **82**: 5661–5672.

Simão FA, Waterhouse RM, Ioannidis P, Kriventseva E V. & Zdobnov EM (2015) BUSCO: assessing genome assembly and annotation completeness with single-copy orthologs. *Bioinformatics* **31**: 3210–3212.

Simpkin VL, Renwick MJ, Kelly R & Mossialos E (2017) Incentivising innovation in antibiotic drug discovery and development: Progress, challenges and next steps. *J Antibiot (Tokyo)* **70**: 1087–1096.

Singh V, Khan M, Khan S & Tripathi CKM (2009) Optimization of actinomycin v production by *Streptomyces triostinicus* using artificial neural network and genetic algorithm. *Appl Microbiol Biotechnol* **82**: 379–385.

Sohoni S V., Bapat PM & Lantz AE (2012) Robust, small-scale cultivation platform for *Streptomyces coelicolor*. *Microb Cell Fact* **11**: 9.

Stapley EO, Jackson M, Hernandez S, Zimmerman SB, Currie SA, Mochales S, Mata JM, Woodruff HB & Hendlin D (1972) Cephamycins, a new family of beta-lactam antibiotics. I. Production by actinomycetes, including Streptomyces lactamdurans sp. n. *Antimicrob Agents Chemother* **2**: 122–31.

Studholme DJ (2016) Genome Update. Let the consumer beware: *Streptomyces genome* sequence quality. *Microb Biotechnol* **9**: 3–7.

Sztáray J, Memboeuf A, Drahos L & Vékey K (2011) Leucine enkephalin - A mass spectrometry standard. *Mass Spectrom Rev* **30**: 298–320.

Szymańska E, Saccenti E, Smilde AK & Westerhuis JA (2012) Double-check: validation of diagnostic statistics for PLS-DA models in metabolomics studies. *Metabolomics* **8**: 3–16.

Tacconelli E, Carrara E, Savoldi A, Harbarth S, Mendelson M, Monnet DL, Pulcini C, Kahlmeter G, Kluytmans J, Zorzet A, *et al.* (2018) Discovery, research, and development of new antibiotics: the WHO priority list of antibiotic-resistant bacteria and tuberculosis. *Lancet Infect Dis* **18**: 318–327.

Takamatsu S, Lin X, Nara A, Komatsu M, Cane DE & Ikeda H (2011) Characterization of a silent sesquiterpenoid biosynthetic pathway in *Streptomyces avermitilis* controlling epi-isozizaene albaflavenone biosynthesis and isolation of a new oxidized epi-isozizaene metabolite. *Microb Biotechnol* **4**: 184–191.

Takano H, Obitsu S, Beppu T & Ueda K (2005) Light-induced carotenogenesis in *Streptomyces coelicolor* A3(2): Identification of an extracytoplasmic function sigma factor that directs photodependent transcription of the carotenoid biosynthesis gene cluster. *J Bacteriol* **187**: 1825–1832.

Tan S, Moore G & Nodwell J (2019) Put a bow on it: Knotted antibiotics take center stage. *Antibiotics* **8**: 117.

Tanaka Y, Hosaka T & Ochi K (2010) Rare earth elements activate the secondary metabolite–biosynthetic gene clusters in *Streptomyces coelicolor* A3(2). *J*

Antibiot (Tokyo) **63**: 477–481.

Tanaka Y, Kasahara K, Hirose Y, Murakami K, Kugimiya R & Ochi K (2013) Activation and Products of the Cryptic Secondary Metabolite Biosynthetic Gene Clusters by Rifampin Resistance (rpoB) Mutations in Actinomycetes. *J Bacteriol* **195**: 2959–2970.

Tenconi E, Traxler MF, Hoebreck C, Wezel GP van & Rigali S (2018) Production of prodiginines is part of a programmed cell death process in *Streptomyces coelicolor*. *Front Microbiol* **9**: 1742.

Terra L, Dyson PJ, Hitchings MD, Thomas L, Abdelhameed A, Banat IM, Gazze SA, Vujaklija D, Facey PD, Francis LW & Quinn GA (2018) A novel alkaliphilic streptomyces inhibits ESKAPE pathogens. *Front Microbiol* **9**: 1–13.

Thaker MN, Wang W, Spanogiannopoulos P, Waglechner N, King AM, Medina R & Wright GD (2013) Identifying producers of antibacterial compounds by screening for antibiotic resistance. *Nat Biotechnol* **31**: 922–927.

The Clinical and Laboratory Standards Institute (2018) Colistin Breakpoints for Pseudomonas aeruginosa *and* Acinetobacter spp. .

The HC, Rabaa MA, Pham Thanh D, Lappe N De, Cormican M, Valcanis M, Howden BP, Wangchuk S, Bodhidatta L, Mason CJ, Nguyen Thi Nguyen T, Vu Thuy D, Thompson CN, Phu Huong Lan N, Voong Vinh P, Ha Thanh T, Turner P, Sar P, Thwaites G, Thomson NR, Holt KE & Baker S (2016) South Asia as a Reservoir for the Global Spread of Ciprofloxacin-Resistant *Shigella sonnei*: A Cross-Sectional Study. *PLOS Med* **13**: e1002055.

Thom C (1945) Mycology Presents Penicillin. *Mycologia* **37**: 460.

Thomas D, Morris M, Curtis JM & Boyd RK (1995) Fragmentation mechanisms of protonated actinomycins and their use in structural determination of unknown analogues. *J Mass Spectrom* **30**: 1111–1125.

Thompson TB, Katayama K, Watanabe K, Hutchinson CR & Rayment I (2004) Structural and functional analysis of tetracenomycin F2 cyclase from *Streptomyces glaucescens*: A type II polyketide cyclase. *J Biol Chem* **279**: 37956–37963.

Tian X, Zhang Zhewen, Yang T, Chen M, Li J, Chen F, Yang J, Li W, Zhang B, Zhang Zhang, Wu J, Zhang C, Long L & Xiao J (2016) Comparative Genomics Analysis of *Streptomyces* Species Reveals Their Adaptation to the Marine Environment and Their Diversity at the Genomic Level. *Front Microbiol* **7**: 998.

Timmermans ML, Picott KJ, Ucciferri L & Ross AC (2019) Culturing marine bacteria from the genus *Pseudoalteromonas* on a cotton scaffold alters secondary metabolite production. *Microbiologyopen* **8**: e00724.

Tindall BJ, Sutton G & Garrity GM (2017) *Enterobacter aerogenes* Hormaeche and Edwards 1960 (Approved lists 1980) and *Klebsiella mobilis* Bascomb et al. 1971 (approved lists 1980) share the same nomenclatural type (ATCC 13048) on the approved lists and are homotypic synonyms. *Int J Syst Evol Microbiol* **67**: 502–504.

Tiwari K & Gupta RK (2013) Diversity and isolation of rare actinomycetes: An

overview. *Crit Rev Microbiol* **39**: 256–294.

Tong Y, Whitford CM, Robertsen HL, Blin K, Jørgensen TS, Klitgaard AK, Gren T, Jiang X, Weber T & Lee SY (2019) Highly efficient DSB-free base editing for streptomycetes with CRISPR-BEST. *Proc Natl Acad Sci U S A* **116**: 20366–20375.

Tormo JR, García JB, DeAntonio M, Feliz J, Mira A, Díez MT, Hernández P & Peláez F (2003) A method for the selection of production media for actinomycete strains based on their metabolite HPLC profiles. *J Ind Microbiol Biotechnol* **30**: 582–588.

Torres R, Ramón F, La Mata I De, Acebal C & Castillón MP (1999) Enhanced production of penicillin V acylase from *Streptomyces lavendulae*. *Appl Microbiol Biotechnol* **53**: 81–84.

Traxler MF, Watrous JD, Alexandrov T, Dorrestein PC & Kolter R (2013) Interspecies Interactions Stimulate Diversification of the *Streptomyces coelicolor* Secreted Metabolome. *MBio* **4**: e00459-13-e00459-13.

Tsuchido T & Takano M (1988) Senzitization by heat treatment of *Escherichia coli* K-12 cells to hydrophobic antibacterial compounds. *Antimicrob Agents Chemother* **32**: 1680–1683.

Tunca S, Barreiro C, Sola-Landa A, Coque JJR & Martín JF (2007) Transcriptional regulation of the desferrioxamine gene cluster of *Streptomyces coelicolor* is mediated by binding of DmdR1 to an iron box in the promoter of the desA gene. *FEBS J* **274**: 1110–1122.

Ud-Din AIMS, Khan MF & Roujeinikova A (2020) Broad Specificity of Amino Acid Chemoreceptor CtaA of *Pseudomonas fluorescens* Is Afforded by Plasticity of Its Amphipathic Ligand-Binding Pocket. *Mol Plant-Microbe Interact* **33**: 612–623.

Vengurlekar S, Sharma R & Trivedi P (2012) Efficacy of some natural compounds as antifungal agents. *Pharmacogn Rev* **6**: 91–99.

Vicente MF, Basilio A, Cabello A & Peláez F (2003) Microbial natural products as a source of antifungals. *Clin Microbiol Infect* **9**: 15–32.

Viegelmann C, Margassery LM, Kennedy J, Zhang T, O'Brien C, O'Gara F, Morrissey JP, Dobson ADW & Edrada-Ebel R (2014) Metabolomic profiling and genomic study of a marine sponge-associated *Streptomyces* sp. *Mar Drugs* **12**: 3323–3351.

Viollier PH, Kelemen GH, Dale GE, Nguyen KT, Buttner MJ & Thompson CJ (2003) Specialized osmotic stress response systems involve multiple SigB-like sigma factors in *Streptomyces coelicolor*. *Mol Microbiol* **47**: 699–714.

Waksman S (1918) Studies on the Proteolytic Enzymes of Soil Fungi and Actinomycetes. *J Bacteriol* **3**: 509–30.

Waksman S & Henrici AT (1943) The Nomenclature and Classification of the Actinomycetes. *J Bacteriol* **46**: 337–41.

Waksman S, Horning ES & Spencer EL (1943) Two Antagonistic Fungi, *Aspergillus fumigatus* and *Aspergillus clavatus*, and Their Antibiotic Substances. *J Bacteriol* **45**: 233–48.

Waksman S & Joffe JS (1922) Microorganisms Concerned in the Oxidation of Sulfur in the Soil: II. *Thiobacillus thiooxidans*, a New Sulfur-oxidizing Organism Isolated from the Soil. *J Bacteriol* **7**: 239–56.

Waksman S & Lechevalier HA (1949) Neomycin, a new antibiotic active against streptomycin-resistant bacteria, including tuberculosis organisms. *Science (80-)* **109**: 305–307.

Waksman S & Skinner CE (1926) Microorganisms Concerned in the Decomposition of celluloses in the soil. *J Bacteriol* **12**: 57–84.

Waksman S & Starkey L (1923) Partial sterilization of soil, microbiological activities and soil fertility. *Soil Sci* **16**: 343–358.

Waksman S & Tishler M (1941) The Chemical Nature of Actinomycin, an Anti-Microbial Substance Produced by *Actinomyces antibioticus*. *J Biol Chem* **142**: 519–528.

Waksman S & Woodruff HB (1940) The Soil as a Source of Microorganisms Antagonistic to Disease-Producing Bacteria. *J Bacteriol* **40**: 581–600.

Waksman S & Woodruff HB (1941) *Actinomyces antibioticus*, a New Soil Organism Antagonistic to Pathogenic and Non-pathogenic Bacteria. *J Bacteriol* **42**: 231–49.

Waksman S & Woodruff HB (1942) Streptothricin, a New Selective Bacteriostatic and Bactericidal Agent, Particularly Active Against Gram-Negative Bacteria. *Proc Soc Exp Biol Med* **49**: 207–210.

Walker JM (2013) *Metabolomics Tools for Natural Product Discovery*. Humana Press, Totowa, NJ.

Wang D, Wang C, Gui P, Liu H, Khalaf SMH, Elsayed EA, Wadaan MAM, Hozzein WN & Zhu W (2017) Identification, Bioactivity, and Productivity of Actinomycins from the Marine-Derived *Streptomyces heliomycini*. *Front Microbiol* **8**: 1147.

Wang D, Yuan J, Gu S & Shi Q (2013) Influence of fungal elicitors on biosynthesis of natamycin by *Streptomyces natalensis* HW-2. *Appl Microbiol Biotechnol* **97**: 5527–5534.

Wang G, Hosaka T & Ochi K (2008) Dramatic activation of antibiotic production in *Streptomyces coelicolor* by cumulative drug resistance mutations. *Appl Environ Microbiol* **74**: 2834–2840.

Wang L, Gao C, Tang N, Hu S & Wu Q (2015) Identification of genetic variations associated with epsilon-poly-lysine biosynthesis in *Streptomyces albulus* ZPM by genome sequencing. *Sci Rep* **5**: 9201.

Wang R, Mast Y, Wang J, Zhang W, Zhao G, Wohlleben W, Lu Y & Jiang W (2013) Identification of two-component system AfsQ1/Q2 regulon and its cross-regulation with GlnR in *Streptomyces coelicolor*. *Mol Microbiol* **87**: 30–48.

Ward A & Allenby N (2018) Genome mining for the search and discovery of bioactive compounds: The *Streptomyces* paradigm. *FEMS Microbiol Lett* **365**: 240.

Wentzel A, Sletta H, Consortium S, Ellingsen TE & Bruheim P (2012) Intracellular Metabolite Pool Changes in Response to Nutrient Depletion Induced Metabolic Switching in *Streptomyces coelicolor*. *Metabolites* **2**: 178–194.

Wezel GP van & McDowall KJ (2011) The regulation of the secondary metabolism of *Streptomyces*: new links and experimental advances. *Nat Prod Rep* **28**: 1311.

Wibberg D, Haltli B, Kalinowski J, Bandow JE, Al-Dilaimi A, Nowrousian M, Kerr RG, Marchbank DH, Winkler A & Senges CHR (2018) The secreted metabolome of *Streptomyces chartreusis* and implications for bacterial chemistry. *Proc Natl Acad Sci* **115**: 201715713.

Wick RR, Judd LM, Gorrie CL & Holt KE (2017) Unicycler: Resolving bacterial genome assemblies from short and long sequencing reads. *PLoS Comput Biol* **13**: e1005595.

Wickens HJ, Farrell S, Ashiru-Oredope DAI, Jacklin A, Holmes A, Cooke J, Sharland M, Ashiru-Oredope D, McNulty C, Dryden M, Fry C, Hand K, Holmes A, Howard P, Johnson A, Elson R, Mansell PJ, Faulding S, Wagle S, Smart S & Wellsteed S (2013) The increasing role of pharmacists in antimicrobial stewardship in English hospitals. *J Antimicrob Chemother* **68**: 2675–2681.

Wielders CLC, Fluit AC, Brisse S, Verhoef J & Schmitz FJ (2002) *mecA* gene is widely disseminated in *Staphylococcus aureus* population. *J Clin Microbiol* **40**: 3970–3975.

Williams WK & Katz E (1977) Development of a chemically defined medium for the synthesis of actinomycin D by *Streptomyces parvulus*. *Antimicrob Agents Chemother* **11**: 281–290.

Wills RH & O'Connor PB (2014) Structural characterization of actinomycin D using multiple ion isolation and electron induced dissociation. *J Am Soc Mass Spectrom* **25**: 186–195.

Woodhead G & Wood G. (1890) On the actions — antidotal and summative — that the products of bacteria exert on the course of infective disease. *Lancet* **135**: 393–396.

Woodruff HB (2014) Selman A. Waksman, winner of the 1952 nobel prize for physiology or medicine. *Appl Environ Microbiol* **80**: 2–8.

Worley B & Powers R (2012) Multivariate Analysis in Metabolomics. *Curr Metabolomics* **1**: 92–107.

Wu C, Choi YH & Wezel GP van (2016) Metabolic profiling as a tool for prioritizing antimicrobial compounds. *J Ind Microbiol Biotechnol* **43**: 299–312.

Wu C, Kim HK, Wezel GP van & Choi YH (2015) Metabolomics in the natural products field – a gateway to novel antibiotics. *Drug Discov Today Technol* **13**: 11–17.

Xiang SH, Li J, Yin H, Zheng JT, Yang X, Wang H Bin, Luo JL, Bai H & Yang KQ (2009) Application of a double-reporter-guided mutant selection method to improve clavulanic acid production in Streptomyces clavuligerus. *Metab Eng* **11**: 310–318.

Xie P, Ma M, Rateb ME, Shaaban KA, Yu Z, Huang S-X, Zhao L-X, Zhu X, Yan Y, Peterson RM, Lohman JR, Yang D, Yin M, Rudolf JD, Jiang Y, Duan Y & Shen B (2014) Biosynthetic Potential-Based Strain Prioritization for Natural Product Discovery: A Showcase for Diterpenoid-Producing Actinomycetes. *J Nat Prod* **77**: 377–387.

Xiong Z, Zhang Z-P, Li J-H, Wei S-J & Tu G-Q (2012) Characterization of *Streptomyces padanus* JAU4234, a Producer of Actinomycin X2, Fungichromin, and a New Polyene Macrolide Antibiotic. *Appl Environ Microbiol* **78**: 589–592.

Xu M & Wright GD (2019) Heterologous expression-facilitated natural products' discovery in actinomycetes. *J Ind Microbiol Biotechnol* **46**: 415–431.

Yagüe P, López-García MT, Rioseras B, Sánchez J & Manteca Á (2013) Pre-sporulation stages of *Streptomyces* differentiation: state-of-the-art and future perspectives. *FEMS Microbiol Lett* **342**: 79–88.

Yagüe P, Rioseras B, Sanchez J & Manteca Á (2012) New insights on the development of *Streptomyces* and their relationships with secondary metabolite production. *Curr Trends Microbiol* **8**: 65–73.

Yagüe P, Rodríguez-García A, López-García MT, Martín JF, Rioseras B, Sánchez J & Manteca A (2013) Transcriptomic Analysis of Streptomyces coelicolor Differentiation in Solid Sporulating Cultures: First Compartmentalized and Second Multinucleated Mycelia Have Different and Distinctive Transcriptomes. *PLoS One* **8**: e60665.

Yagüe P, Willemse J, Koning RI, Rioseras B, López-García MT, Gonzalez-Quiñonez N, Lopez-Iglesias C, Shliaha P V., Rogowska-Wrzesinska A, Koster AJ, Jensen ON, Wezel GP van & Manteca Á (2016) Subcompartmentalization by cross-membranes during early growth of *Streptomyces* hyphae. *Nat Commun* **7**: 12467.

Yim G, Huimi Wang H & Davies J (2006) The truth about antibiotics. *Int J Med Microbiol* **296**: 163–170.

Yong D, Toleman MA, Giske CG, Cho HS, Sundman K, Lee K & Walsh TR (2009) Characterization of a new metallo-β-lactamase gene, bla NDM-1, and a novel erythromycin esterase gene carried on a unique genetic structure in *Klebsiella pneumoniae* sequence type 14 from India. *Antimicrob Agents Chemother* **53**: 5046–5054.

Yoon V & Nodwell JR (2014) Activating secondary metabolism with stress and chemicals. *J Ind Microbiol Biotechnol* **41**: 415–424.

Zaburannyi N, Bunk B, Maier J, Overmann J & Müller R (2016) Genome analysis of the fruiting body-forming myxobacterium *Chondromyces crocatus* reveals high potential for natural product biosynthesis. *Appl Environ Microbiol* **82**: 1945–1957.

Zaburannyi N, Rabyk M, Ostash B, Fedorenko V & Luzhetskyy A (2014) Insights into naturally minimised *Streptomyces albus* J1074 genome. *BMC Genomics* **15**: 97.

Zeng H, Feng PX & Wan CX (2019) Antifungal effects of actinomycin D on

Verticillium dahliae via a membrane-splitting mechanism. *Nat Prod Res* **33**: 1751–1755.

Zhang H, Dudley EG, Davidson PM & Harte F (2017) Critical Concentration of Lecithin Enhances the Antimicrobial Activity of Eugenol against *Escherichia coli*. *Appl Environ Microbiol* **83**: e03467-16.

Zhang Jia Jia, Yamanaka K, Tang X & Moore BS (2019) Direct cloning and heterologous expression of natural product biosynthetic gene clusters by transformation-associated recombination. In *Methods in Enzymology*. Academic Press Inc., pp. 87–110.

Zhang L (2005) Integrated Approaches for Discovering Novel Drugs From Microbial Natural Products. In *Natural Products*. Humana Press, Totowa, NJ. pp. 33–55.

Zhang W & Tang Y (2009) In Vitro Analysis of Type II Polyketide Synthase. In *Methods in Enzymology*. pp. 367–393.

Zhang X, Ye X, Chai W, Lian XY & Zhang Z (2016) New metabolites and bioactive actinomycins from marine-derived *Streptomyces* sp. ZZ338. *Mar Drugs* **14**: 181.

Zhang Zheren, Barsy F De, Liem M, Liakopoulos A, Choi YH, Claessen D & Rozen DE (2019) Antibiotic production in *Streptomyces* is organized by a division of labour through terminal genomic differentiation. *bioRxiv Prepr* http://dx.doi.org/10.1101/560136.

Zhao C, Coughlin JM, Ju J, Zhu D, Wendt-Pienkowski E, Zhou X, Wang Z, Shen B & Deng Z (2010) Oxazolomycin biosynthesis in *Streptomyces albus* JA3453 featuring an "acyltransferase-less" type I polyketide synthase that incorporates two distinct extender units. *J Biol Chem* **285**: 20097–20108.

Zhu Hua, Sandiford SK & Wezel GP van (2014) Triggers and cues that activate antibiotic production by actinomycetes. *J Ind Microbiol Biotechnol* **41**: 371–386.

Zhu H., Swierstra J, Wu C, Girard G, Choi YH, Wamel W van, Sandiford SK & Wezel GP van (2014) Eliciting antibiotics active against the ESKAPE pathogens in a collection of actinomycetes isolated from mountain soils. *Microbiology* **160**: 1714–1725.

Ziemons S, Koutsantas K, Becker K, Dahlmann T & Kück U (2017) Penicillin production in industrial strain *Penicillium chrysogenum* P2niaD18 is not dependent on the copy number of biosynthesis genes. *BMC Biotechnol* **17**.

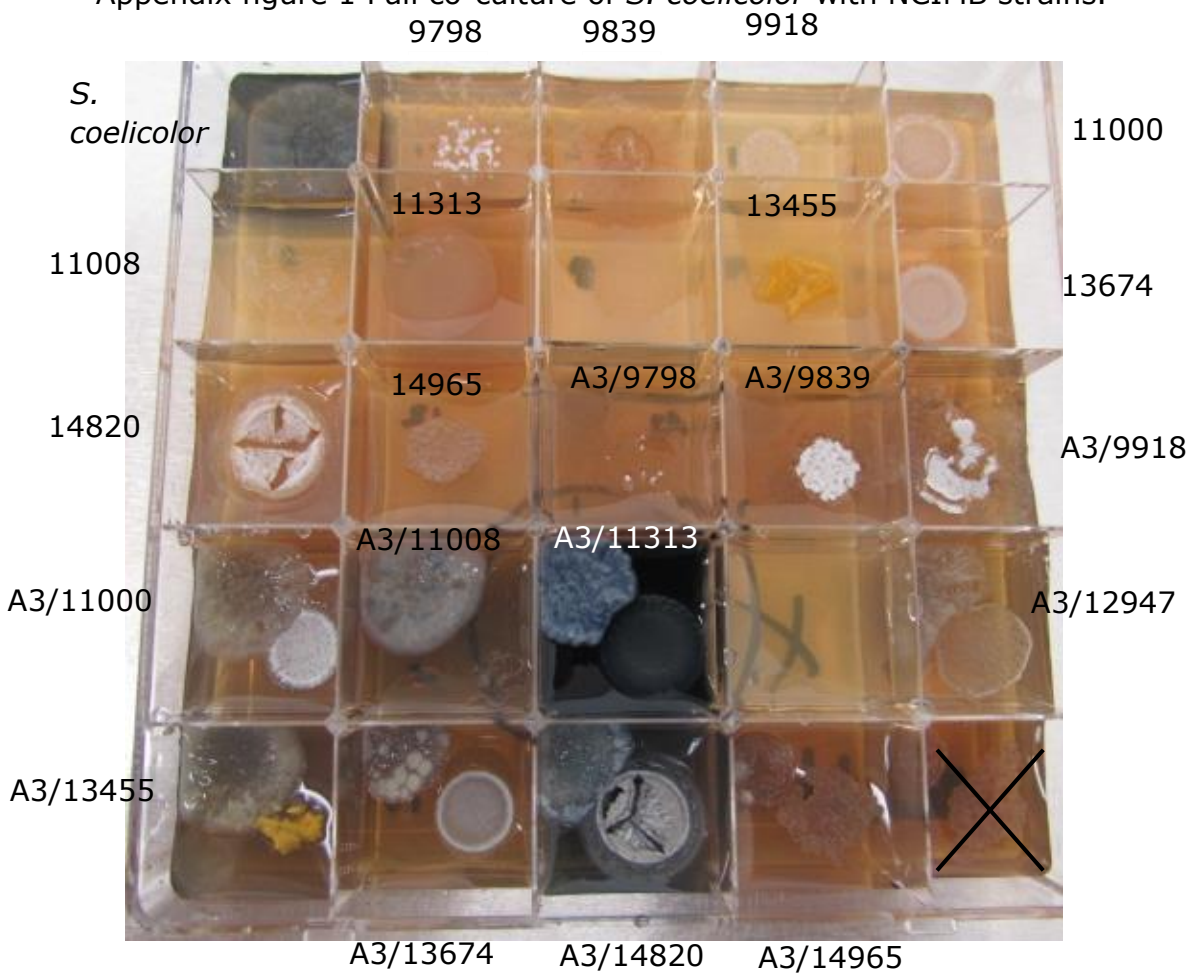
Zimin A, Puiu D, Luo MC, Zhu T, Koren S, Marçais G, Yorke JA, Dvořák J & Salzberg SL (2017) Hybrid assembly of the large and highly repetitive genome of *Aegilops tauschii*, a progenitor of bread wheat, with the MaSuRCA mega-reads algorithm. *Genome Res* **27**: 787–792.

Zinsstag J, Schelling E, Waltner-Toews D & Tanner M (2011) From "one medicine" to "one health" and systemic approaches to health and well-being. *Prev Vet Med* **101**: 148–156.

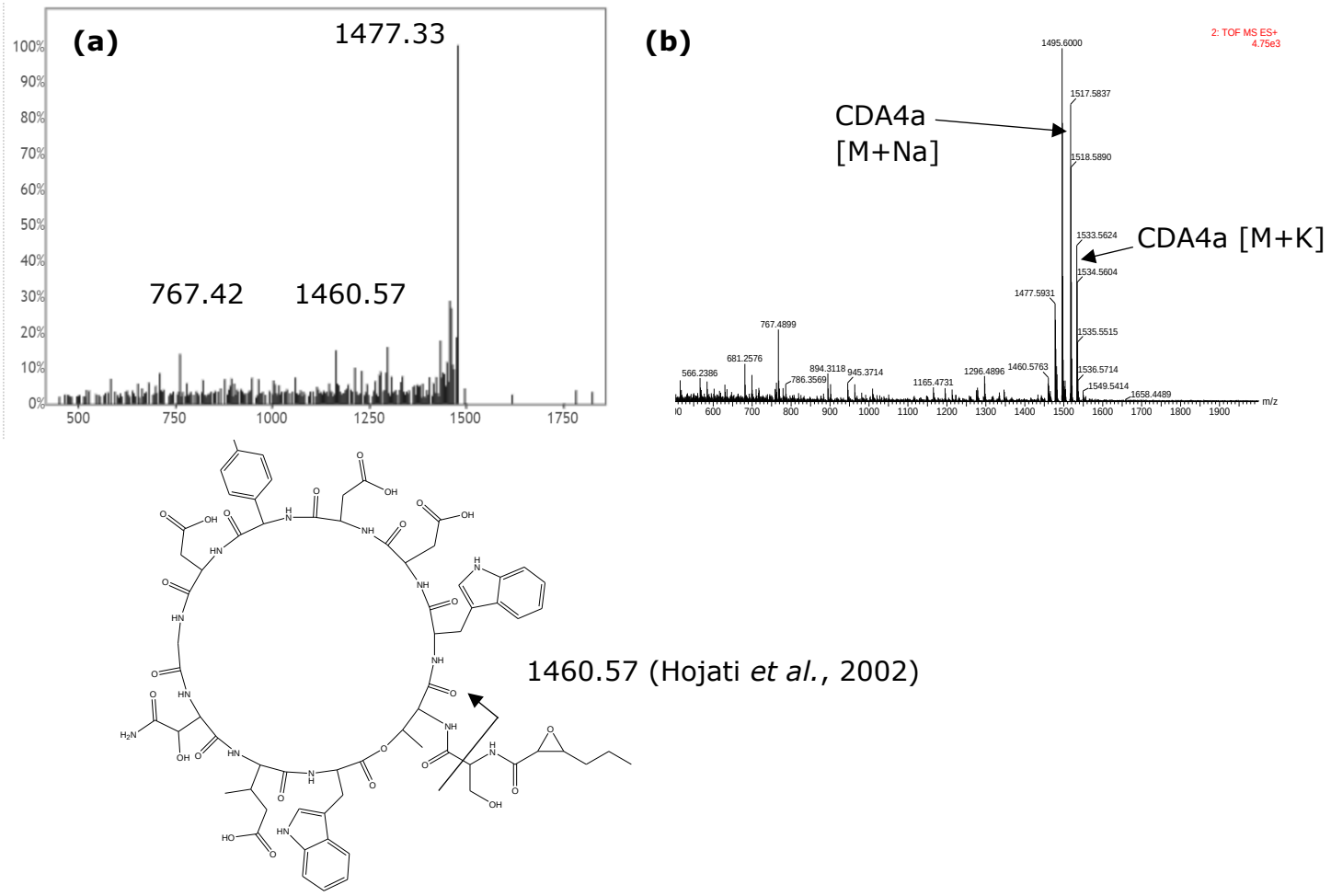
8 Appendix

8.1 Appendix figures

Appendix figure 1 Full co-culture of *S. coelicolor* with NCIMB strains.

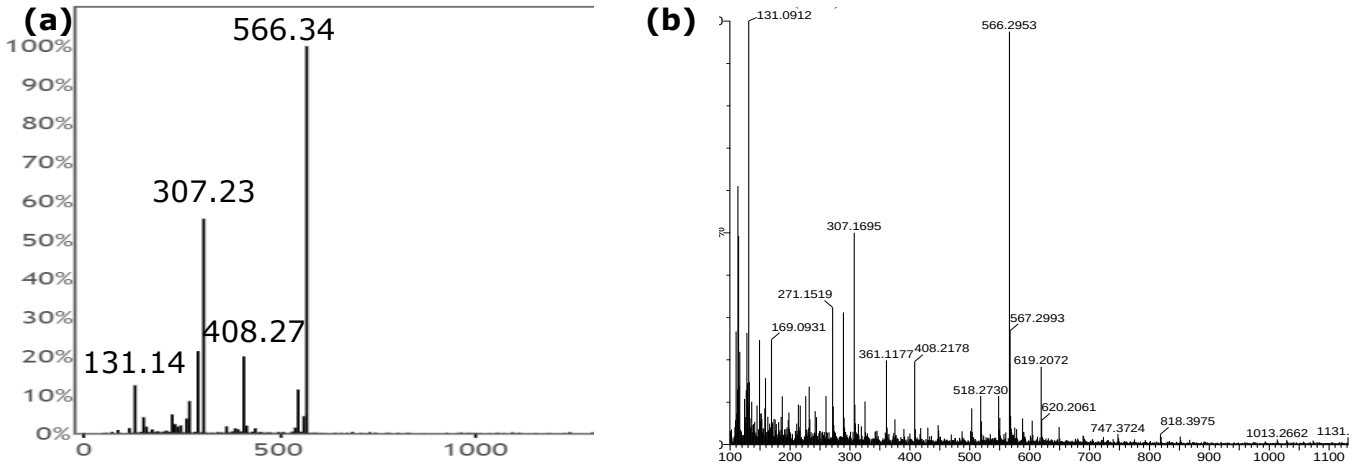


Appendix figure 2 Spectral comparison for CDA 4a of **(a)** GNPS spectra, and **(b)** experimentally detected spectra.



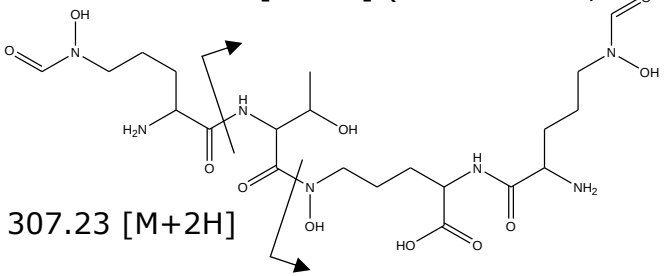
1460.57 (Hojati *et al.*, 2002)

Appendix figure 3 Spectral comparison for coelichelin of **(a)** GNPS spectra, and **(b)** experimentally detected spectra.



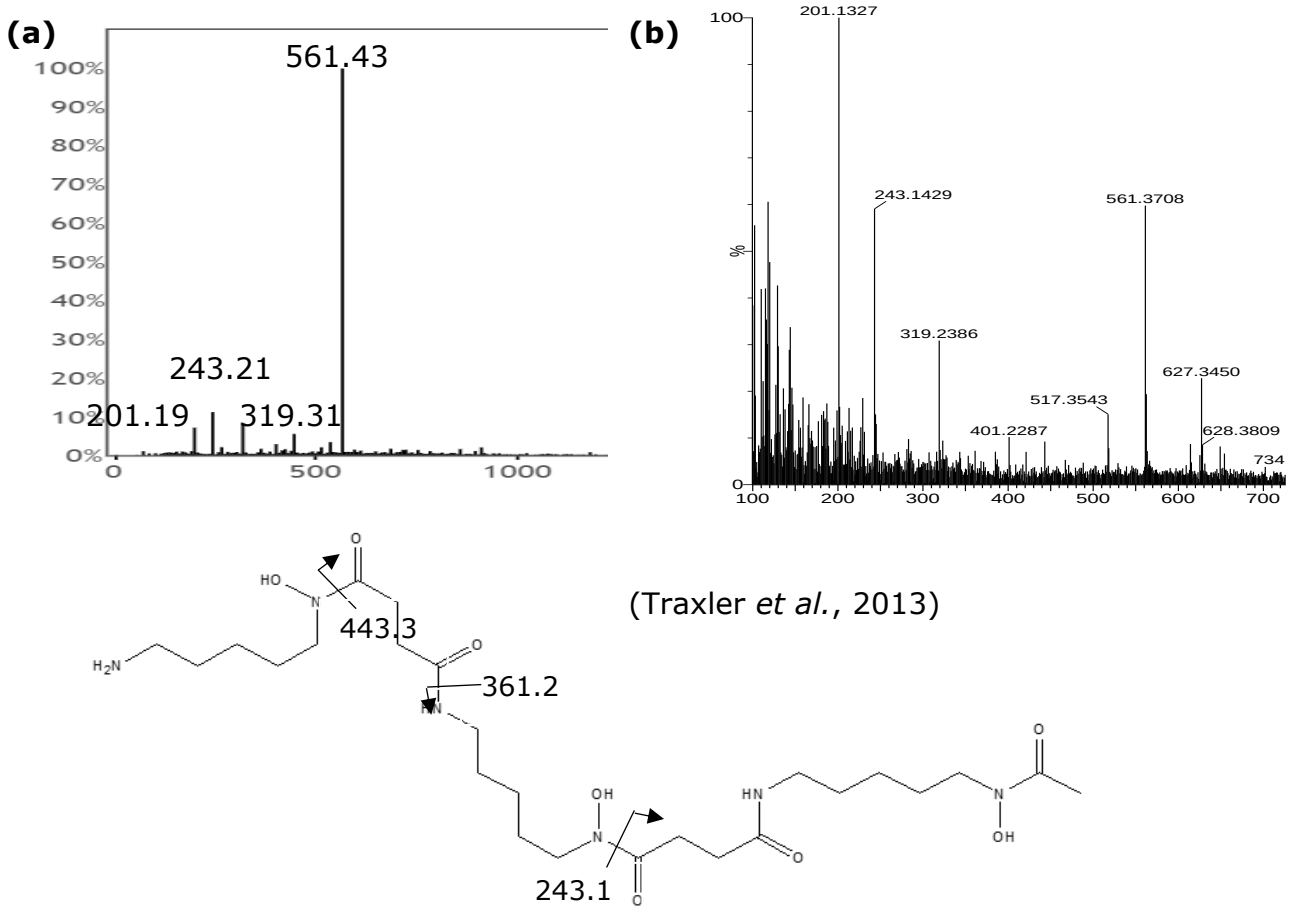
408.27 [M+2H] (Traxler *et al.*, 2013)

)

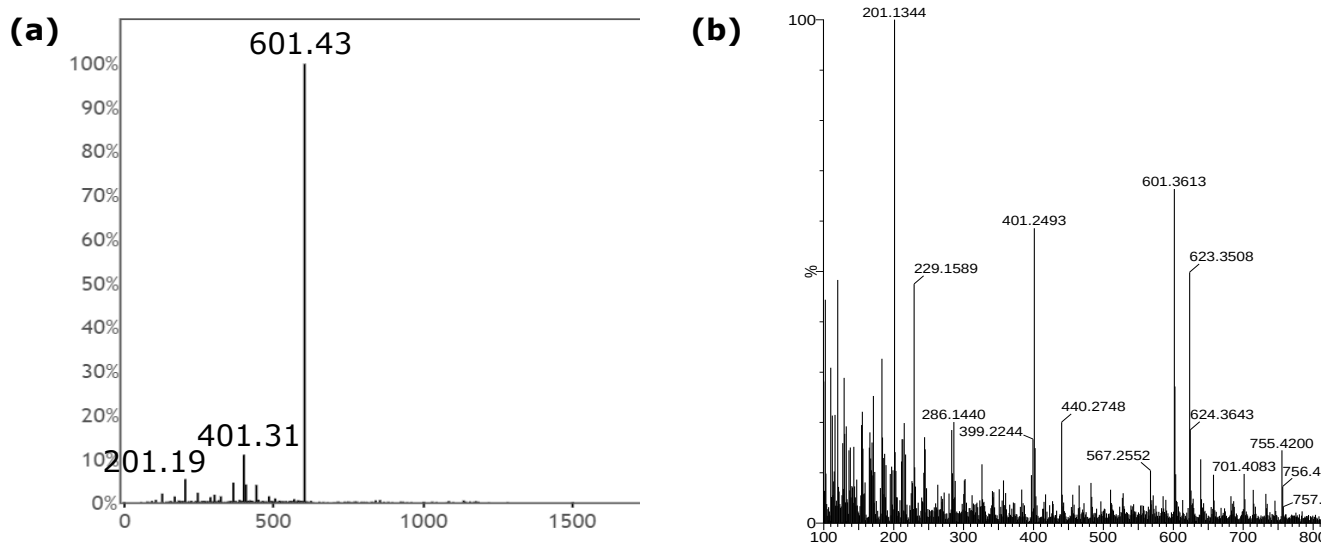


307.23 [M+2H]

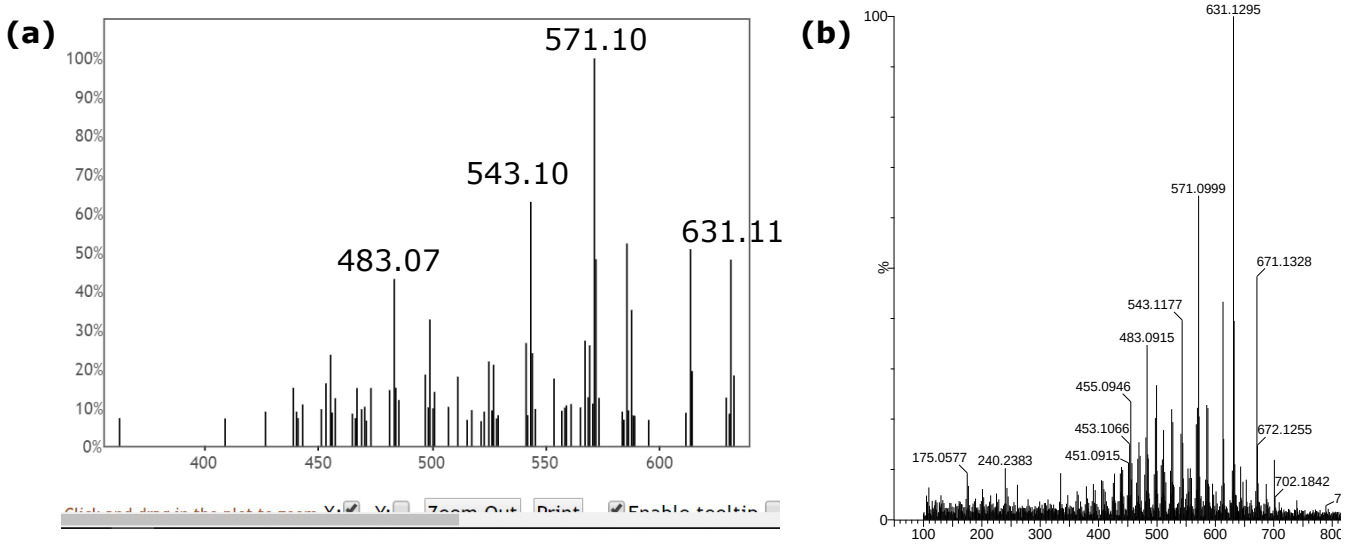
Appendix figure 4 Spectral comparison for deferoxamine B of **(a)** GNPS spectra, and **(b)** experimentally detected spectra.



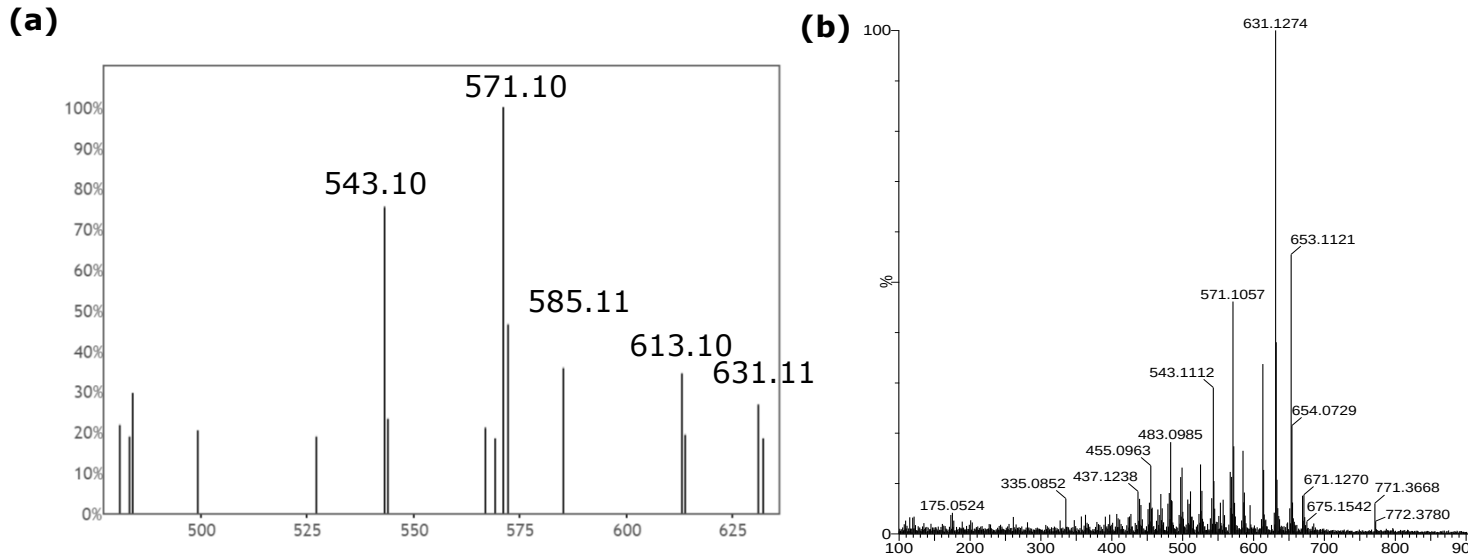
Appendix figure 5 Spectral comparison for deferoxamine E of **(a)** GNPS spectra, and **(b)** experimentally detected spectra.



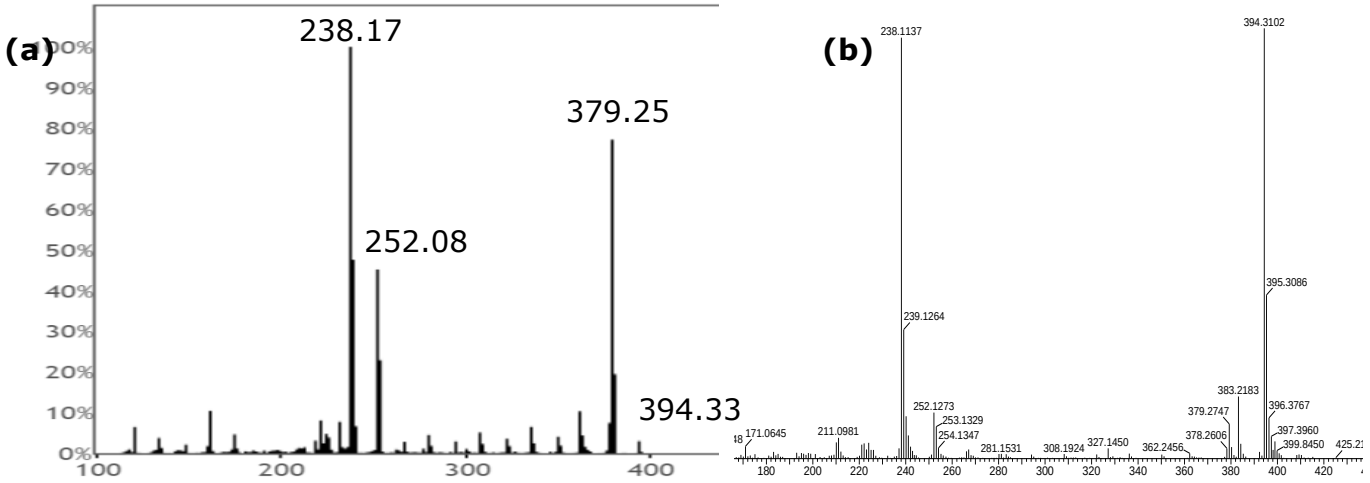
Appendix figure 6 Spectral comparison for ε-ACT of **(a)** GNPS spectra, and **(b)** experimentally detected spectra.



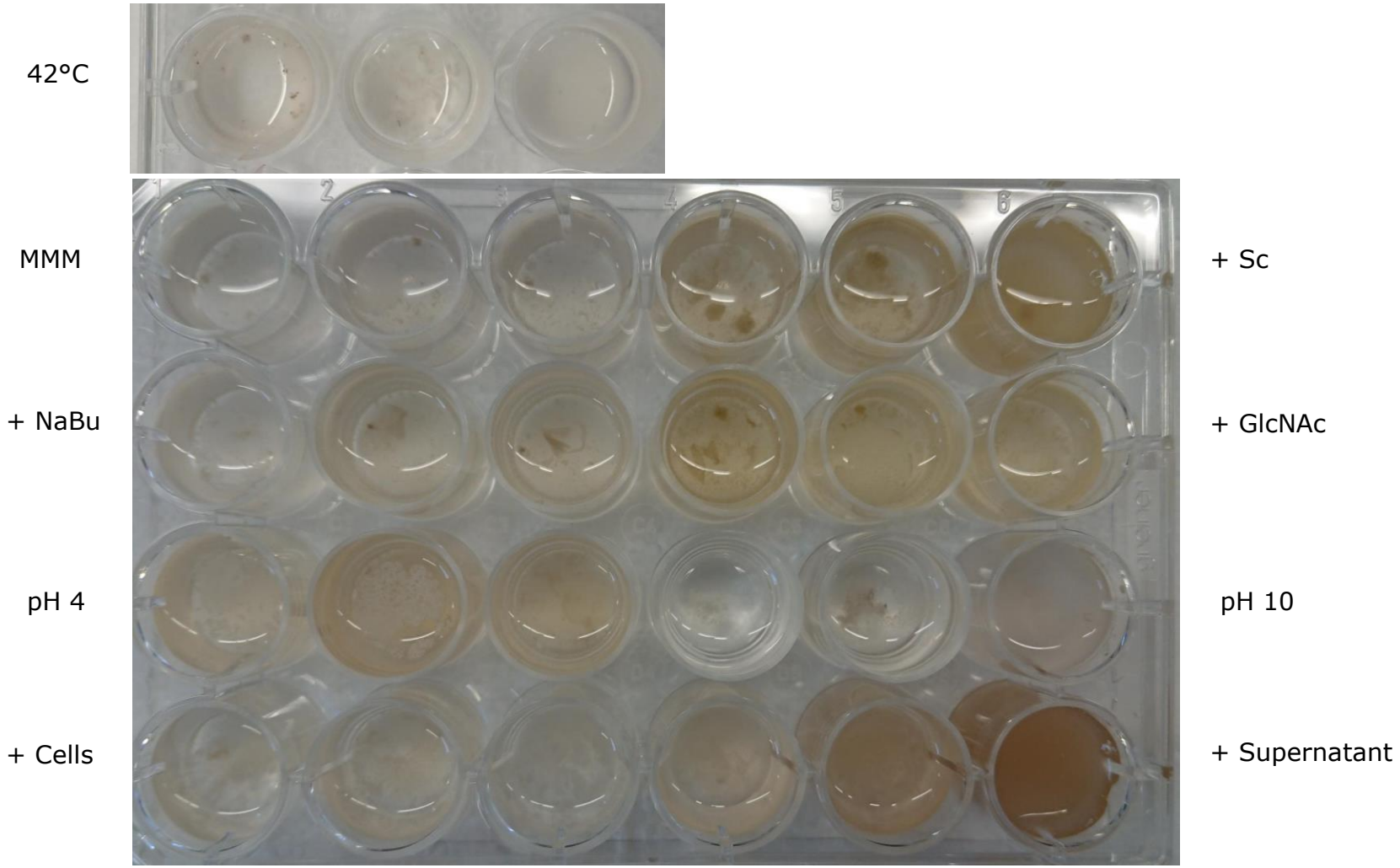
Appendix figure 7 Spectral comparison for γ-ACT of **(a)** GNPS spectra, and **(b)** experimentally detected spectra.



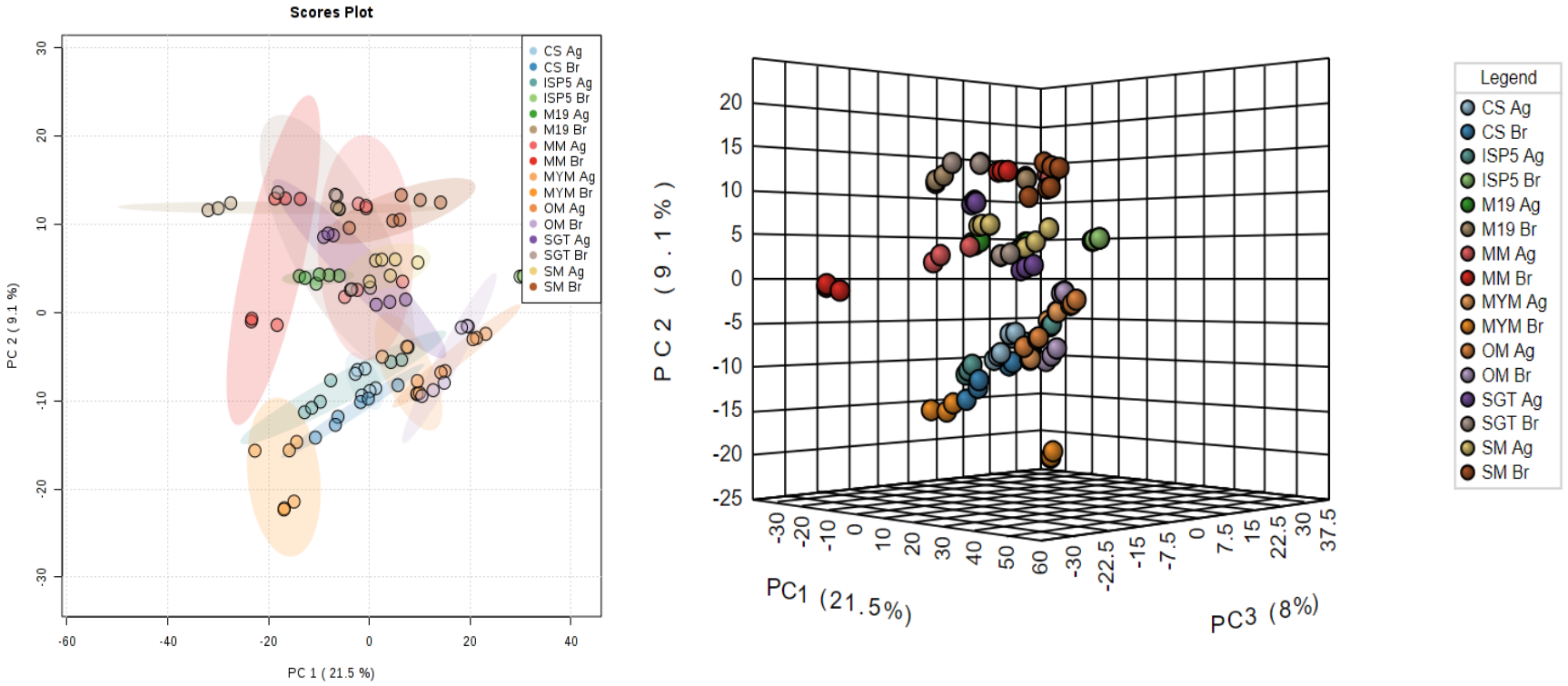
Appendix figure 8 Spectral comparison for RED of **(a)** GNPS spectra, and **(b)** experimentally detected spectra.



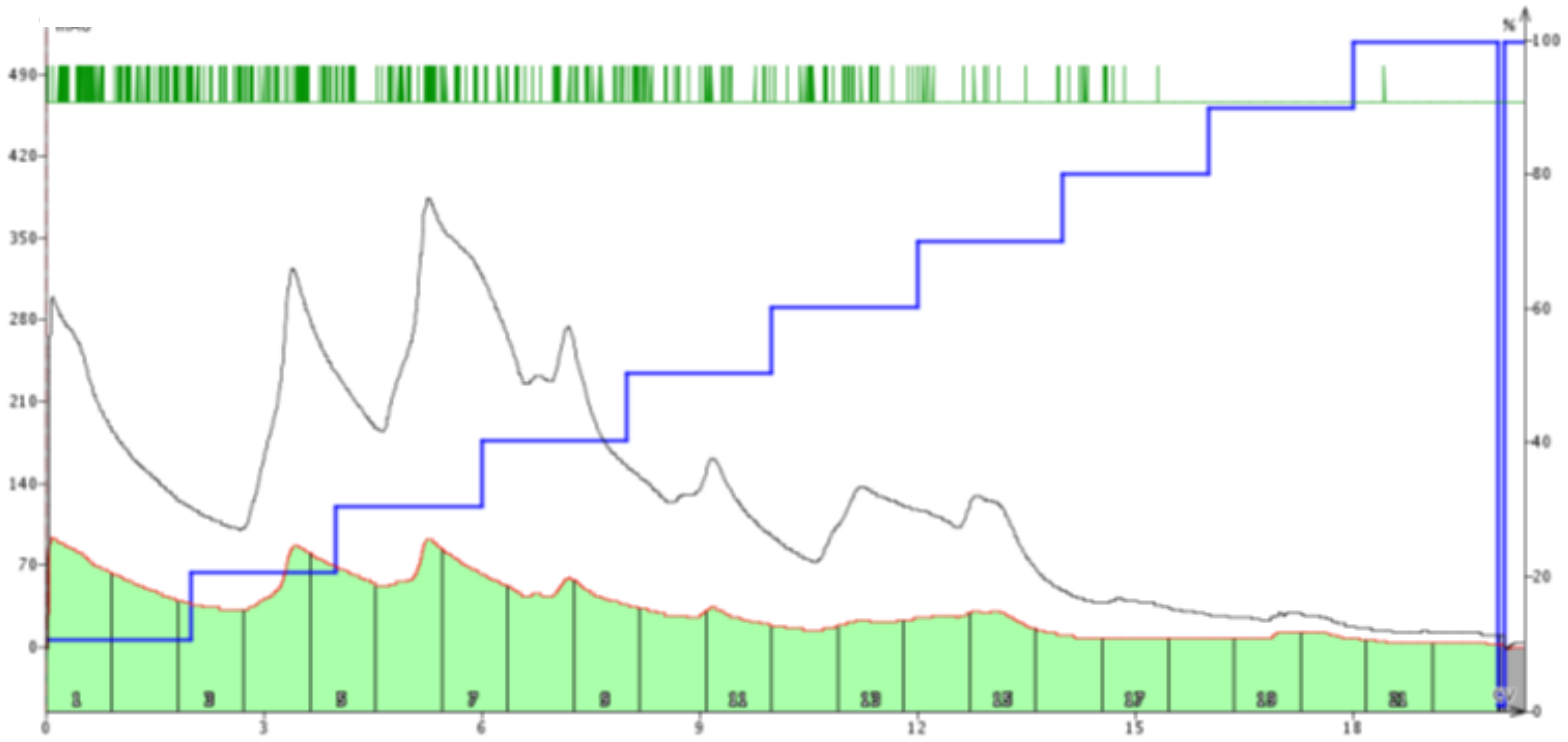
Appendix figure 9 Minimal Media and elicitor culture of *S. costaricanus*.



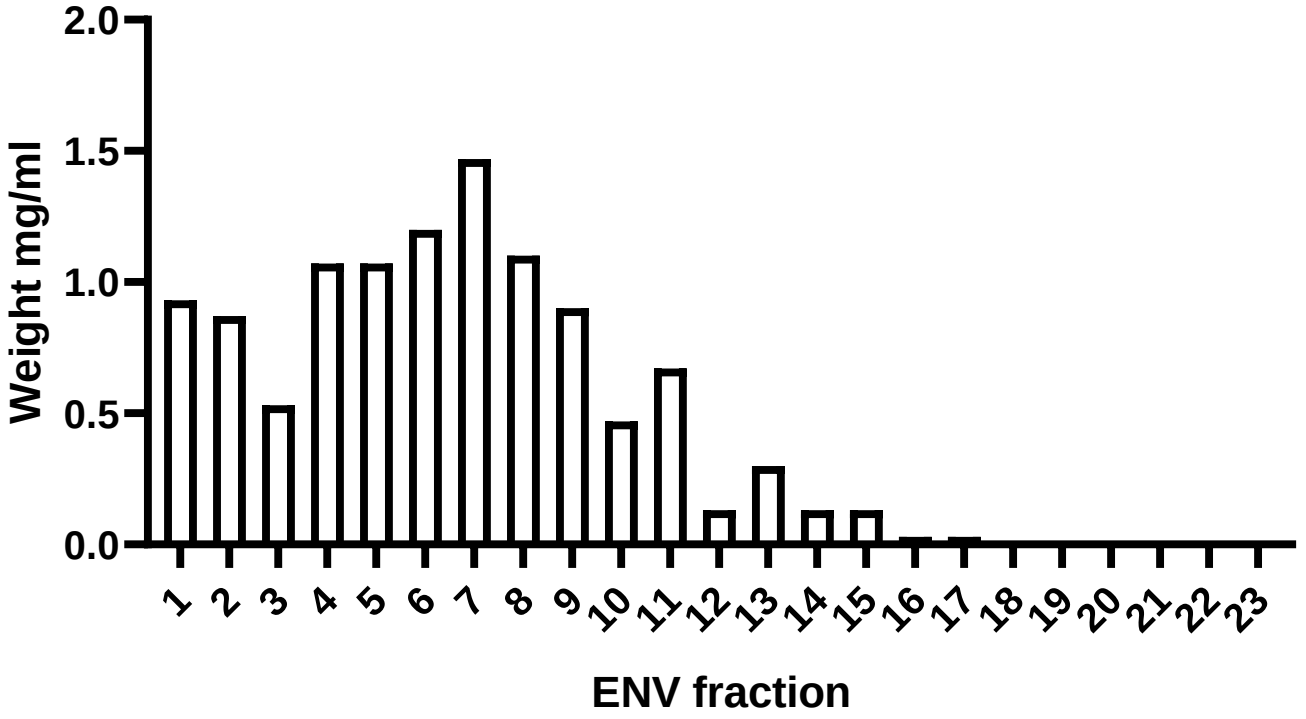
Appendix figure 10 2D and 3D PCA of *S. costaricanus* nutrition screen.



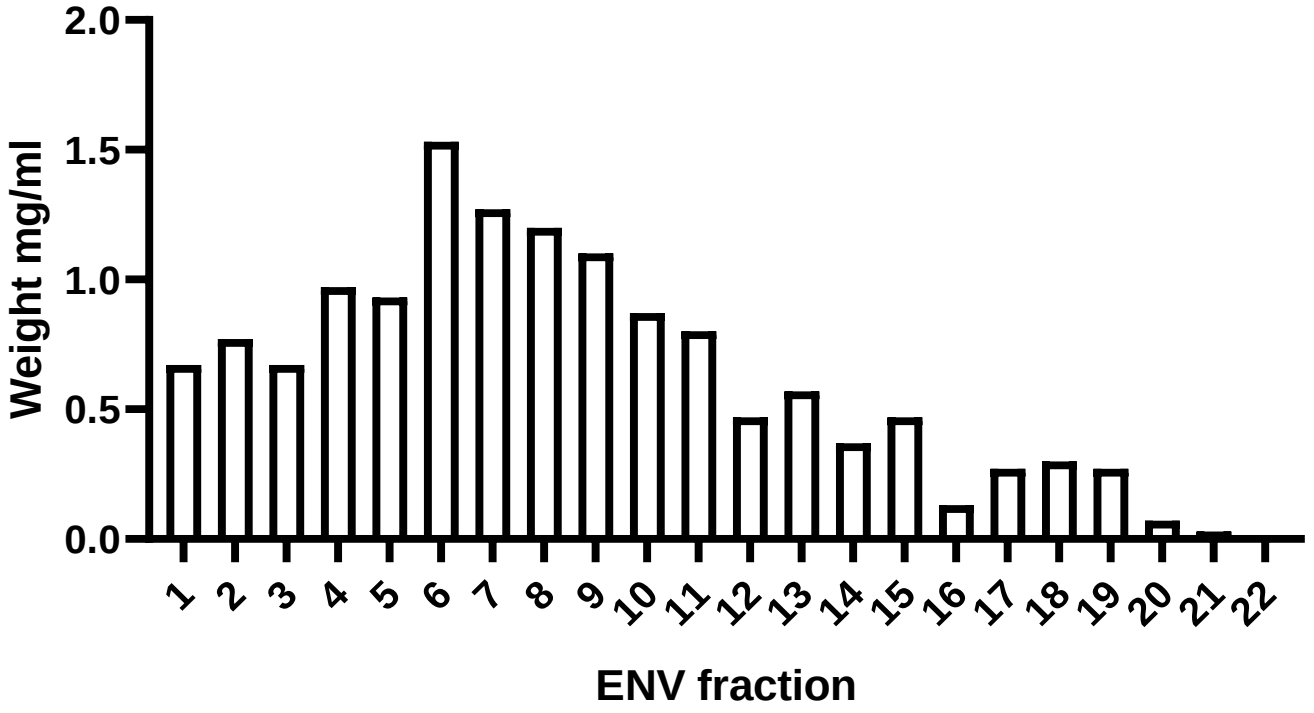
Appendix figure 11 Flash chromatogram of supernatant fractions eluted from ENV+ resin.



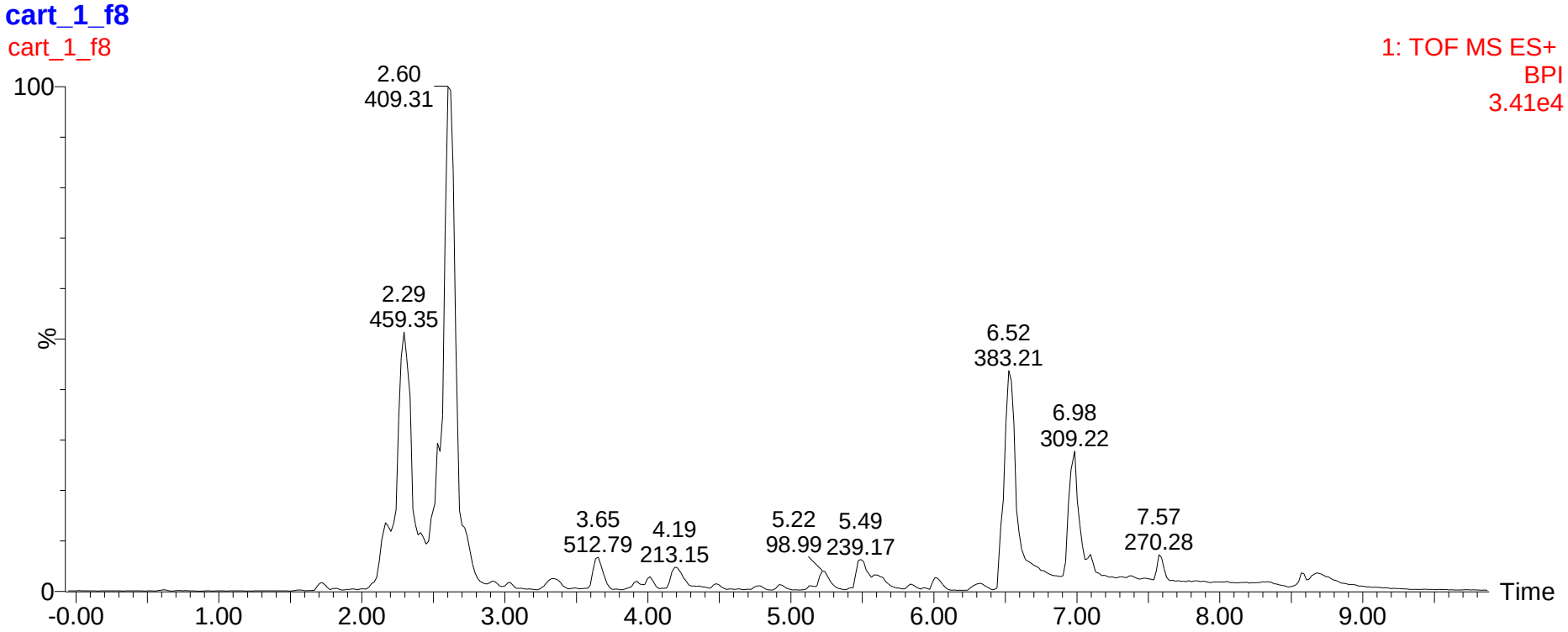
Appendix figure 12 Dry weights for supernatant ENV+ fractions.



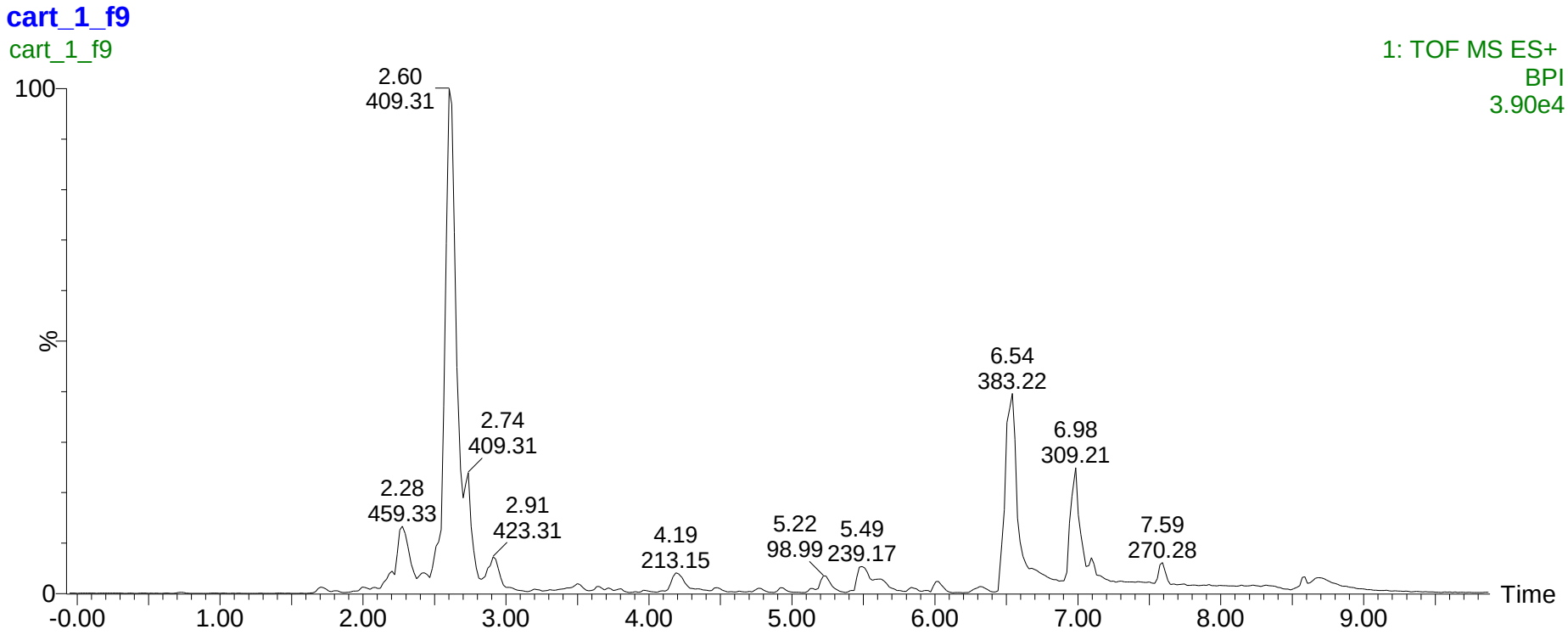
Appendix figure 13 Dry weights for agar culture ENV+ fractions.



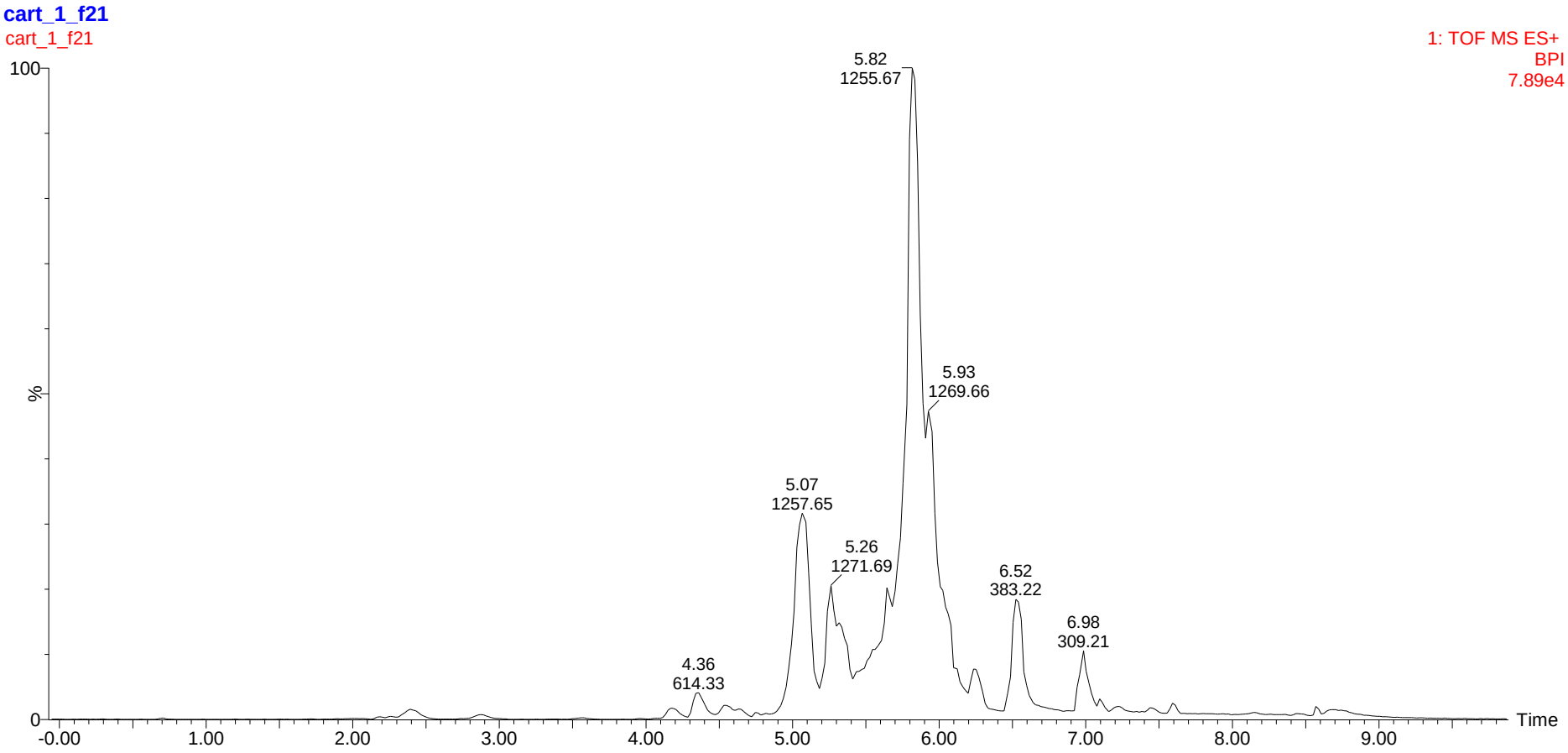
Appendix figure 14 LP Fraction 8



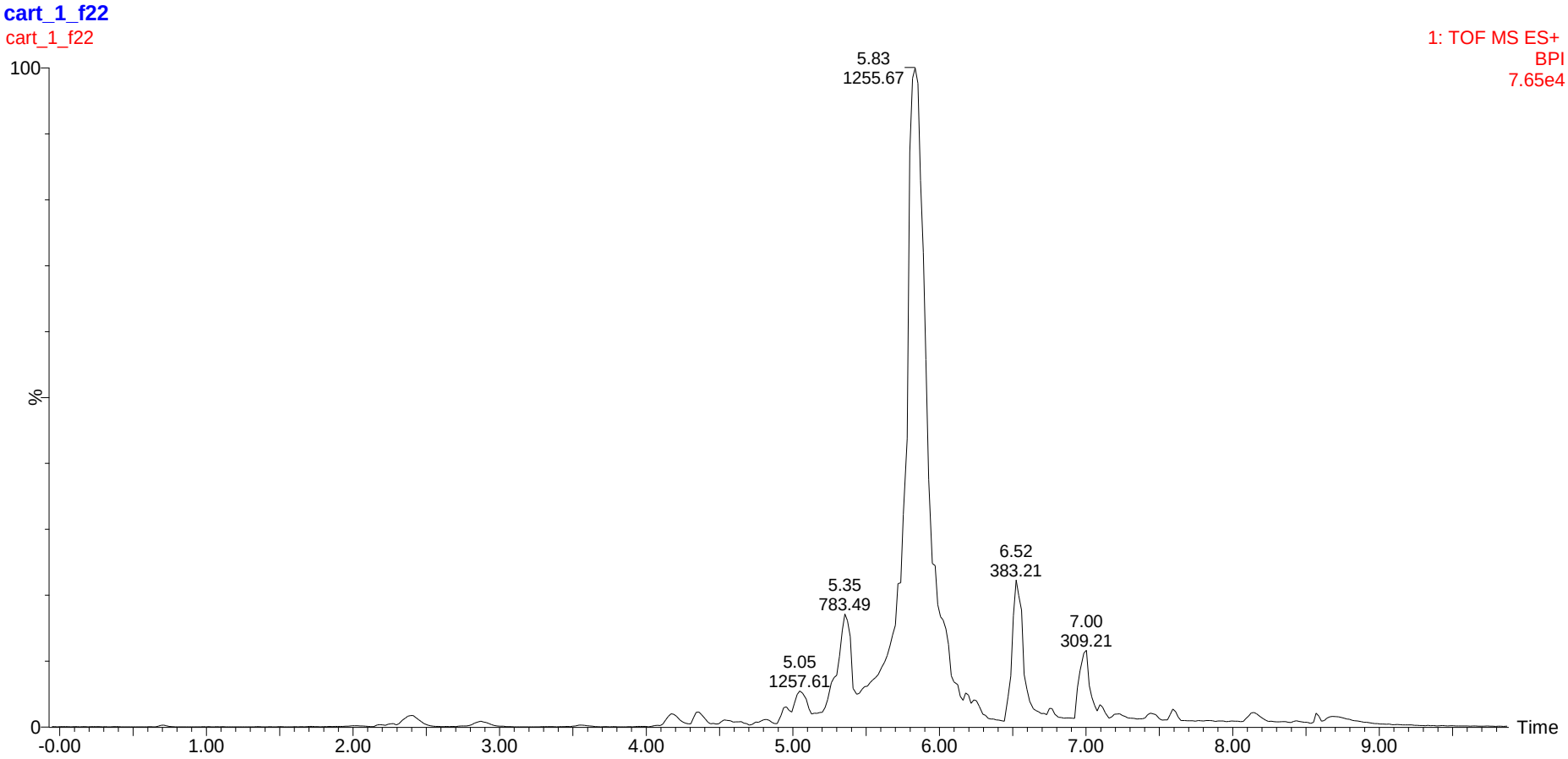
Appendix figure 15 LP Fraction 9.



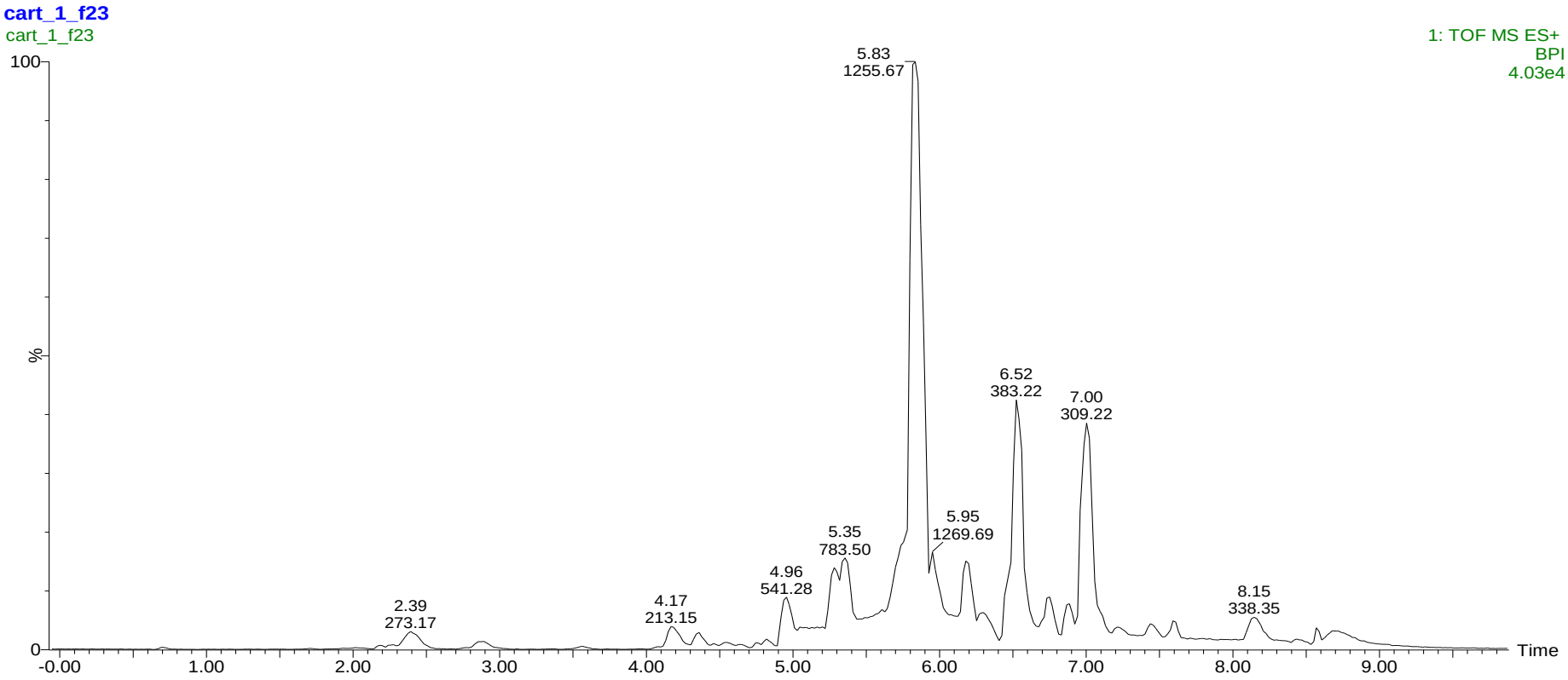
Appendix figure 16 LP fraction 21.



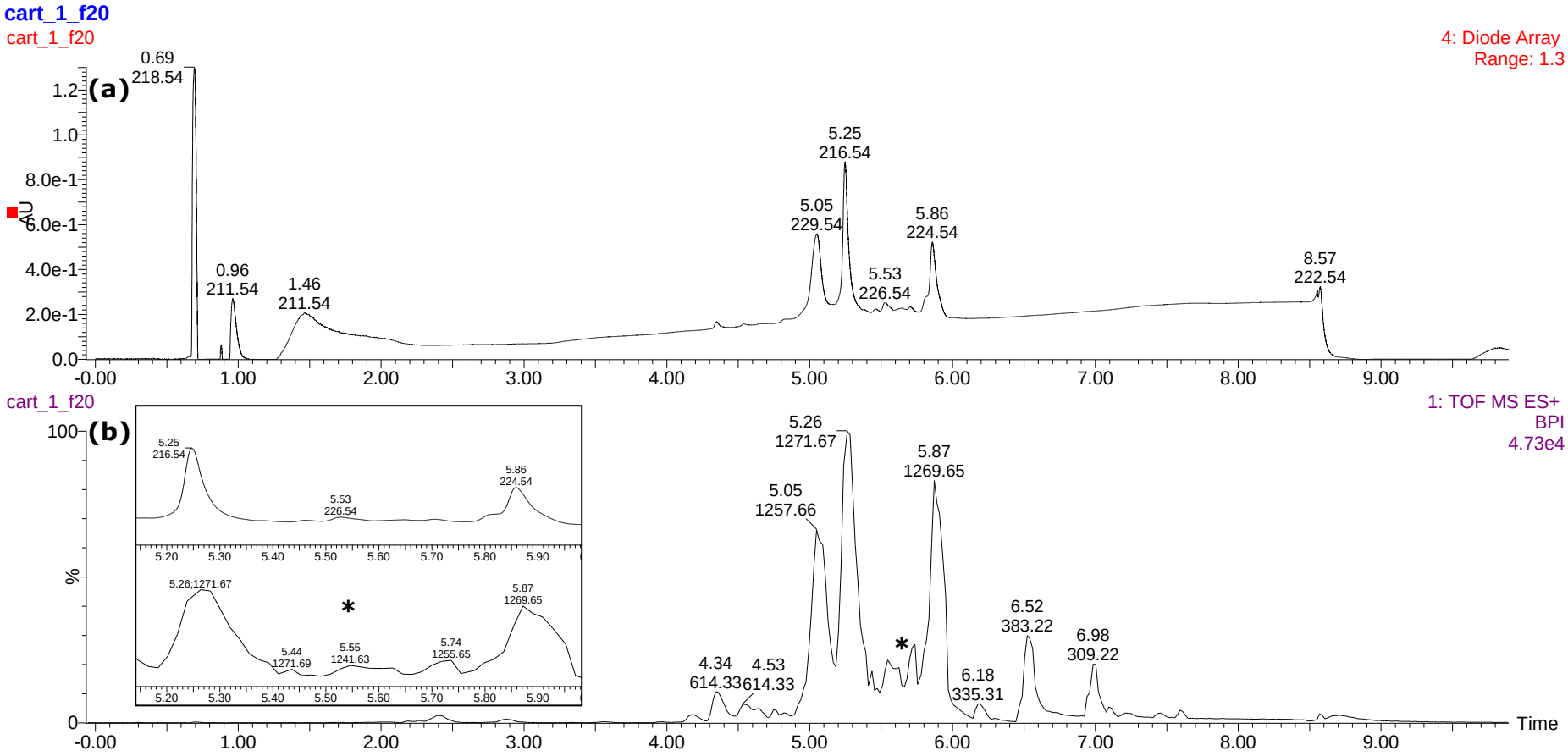
Appendix figure 17 LP fraction 22.



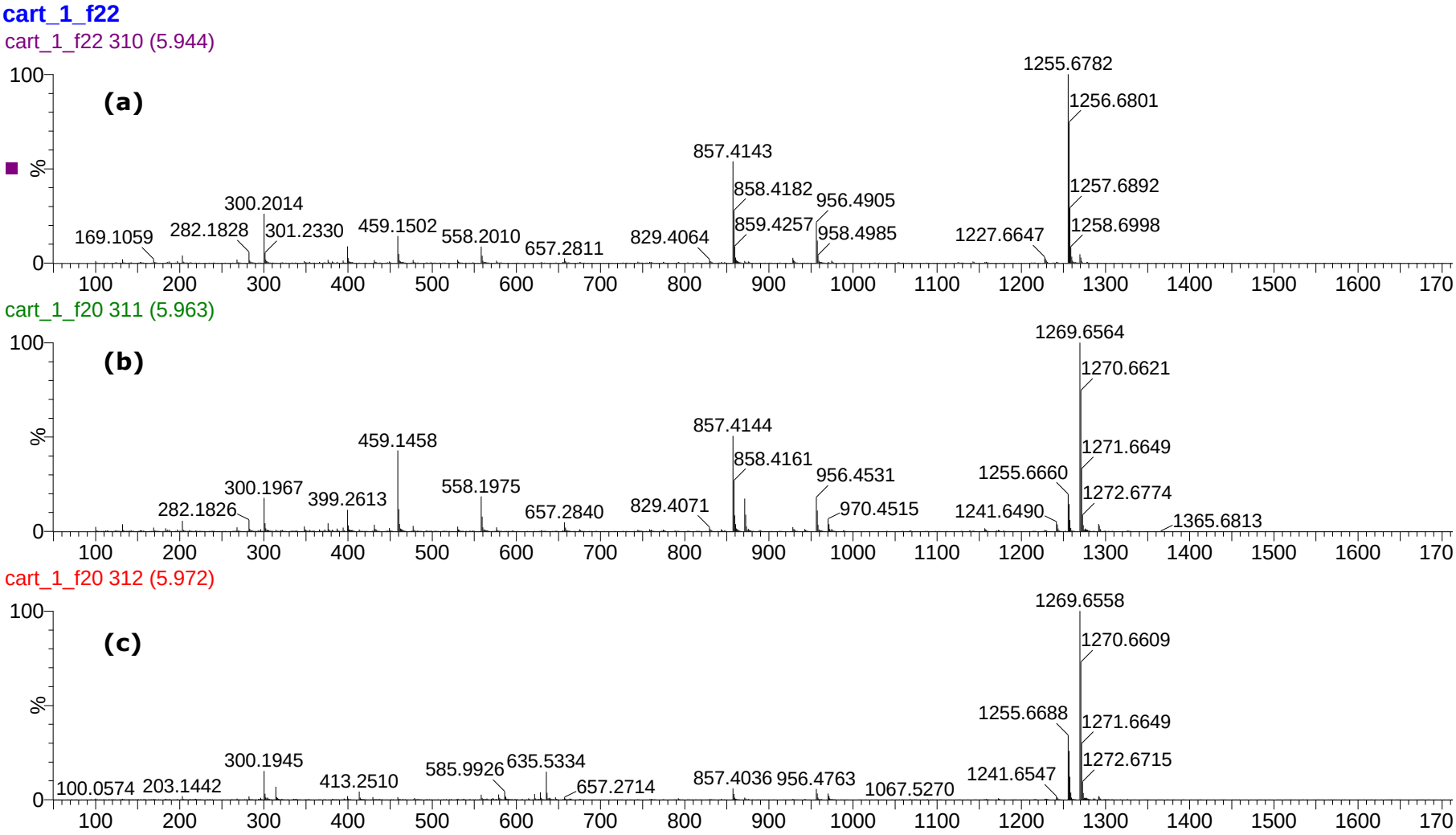
Appendix figure 18 LP fraction 23.



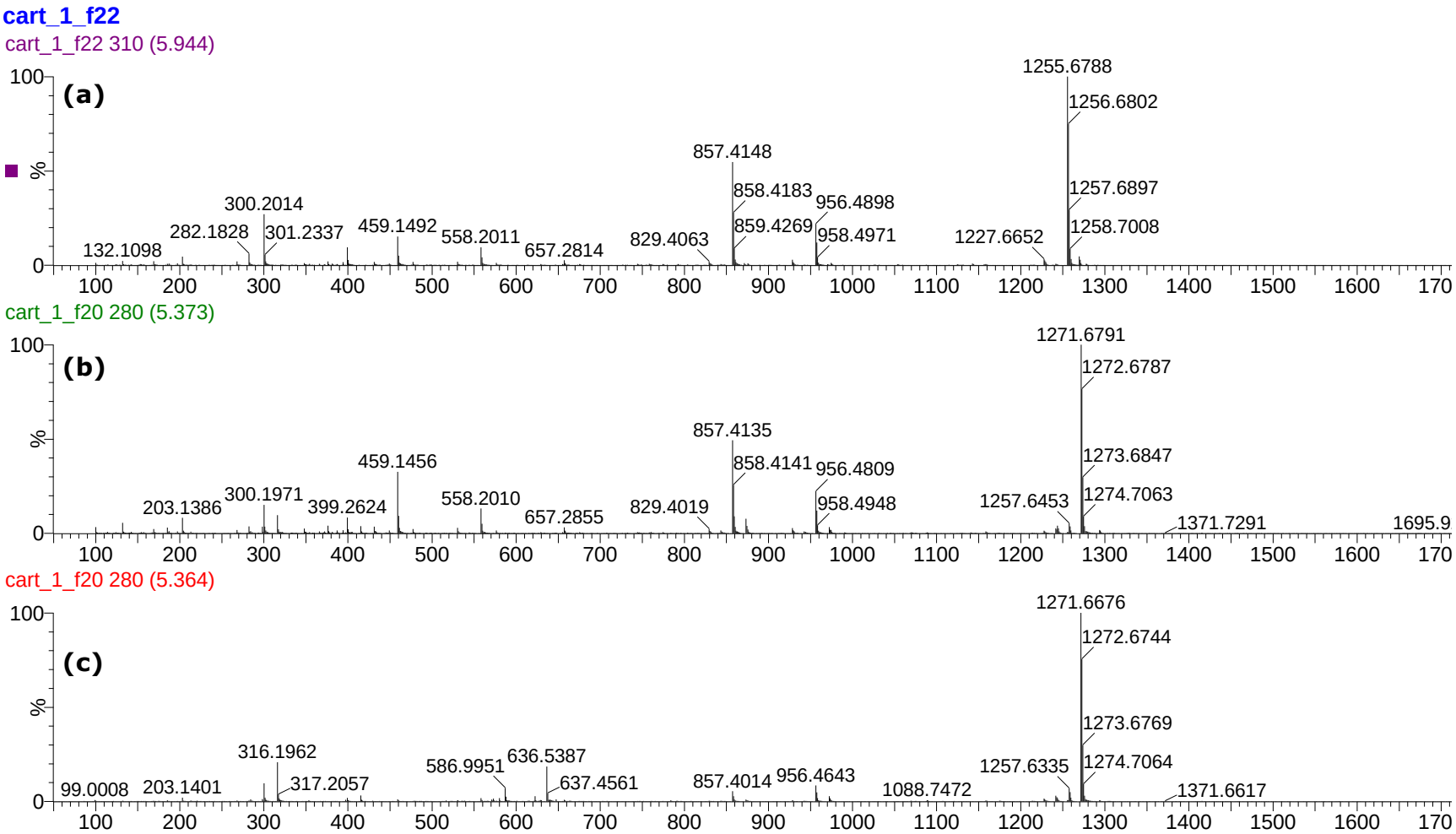
Appendix figure 19 **(a)** Diode array absorption and **(b)** UPLC-MS chromatogram for LP pool 20, used as spectral source for 4 actinomycins: X2 (m/z 1269.65), X0β (m/z 1271.67), m/z 1257.66, and m/z 1241.63, marked with a *.



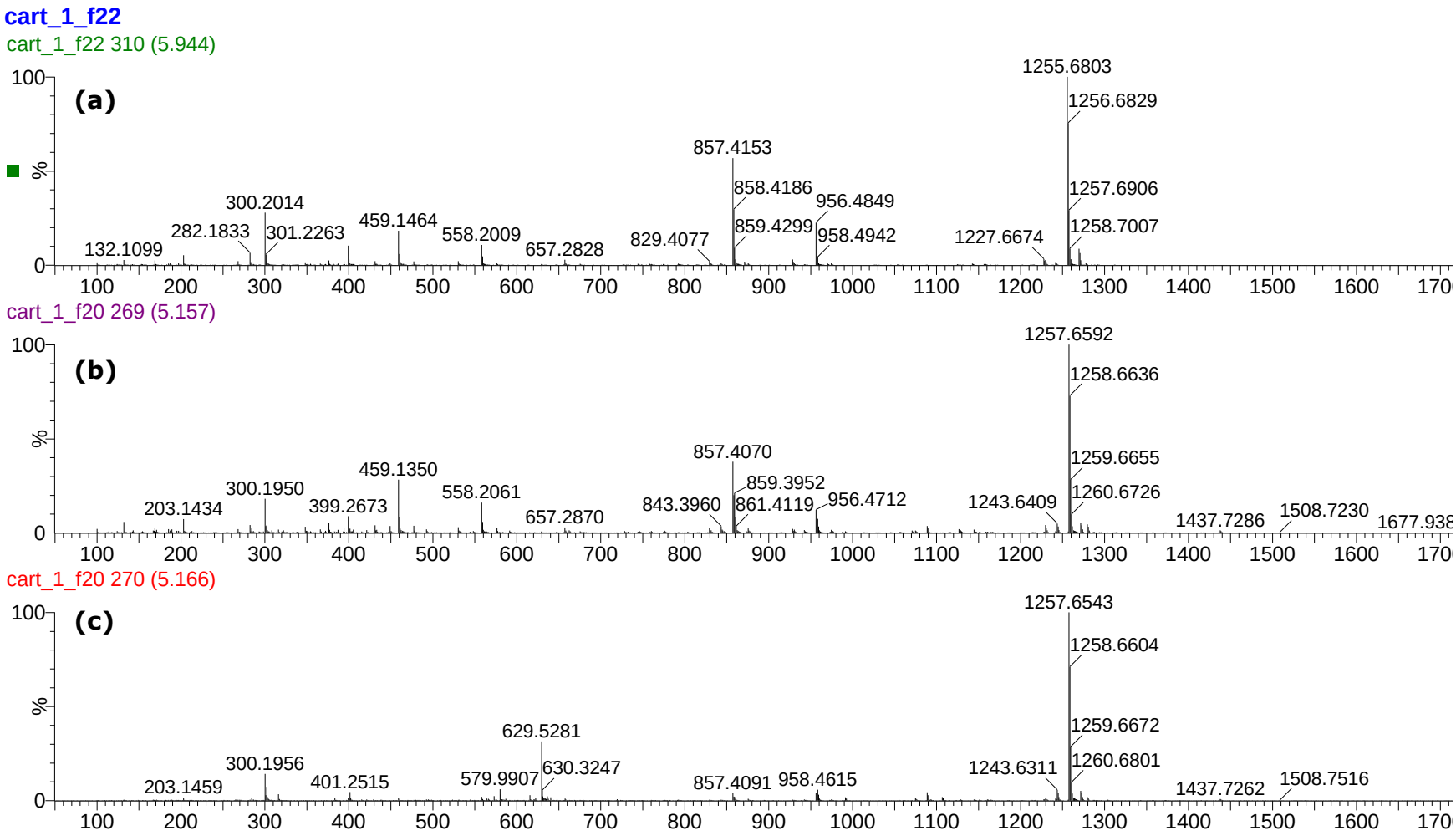
Appendix figure 20 **(a)** Actinomycin D and **(b)** X2 higher energy spectra; **(c)** X2 standard energy spectra.



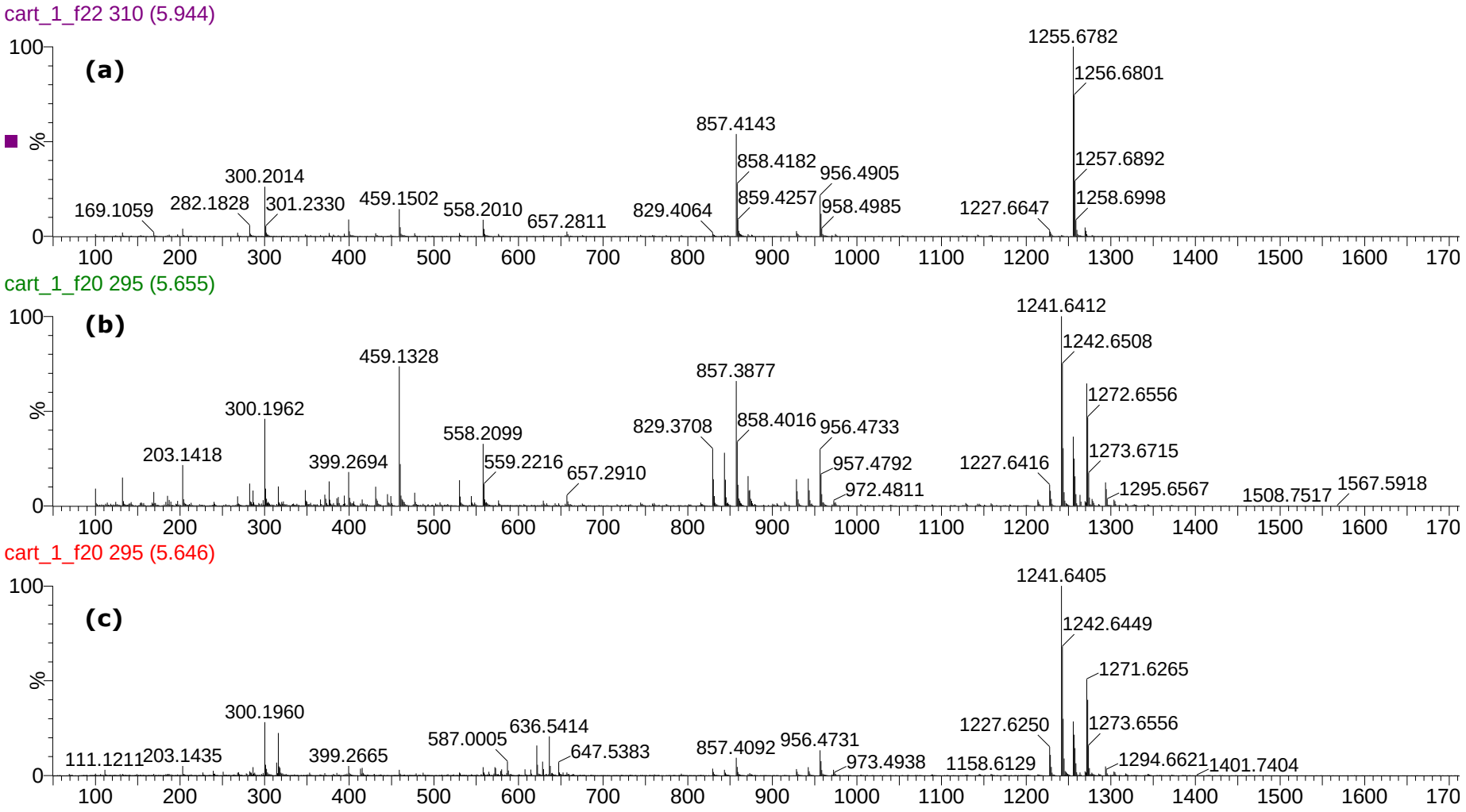
Appendix figure 21 **(a)** Actinomycin D and **(b)** X0β higher energy spectra; **(c)** X0β standard energy spectra.



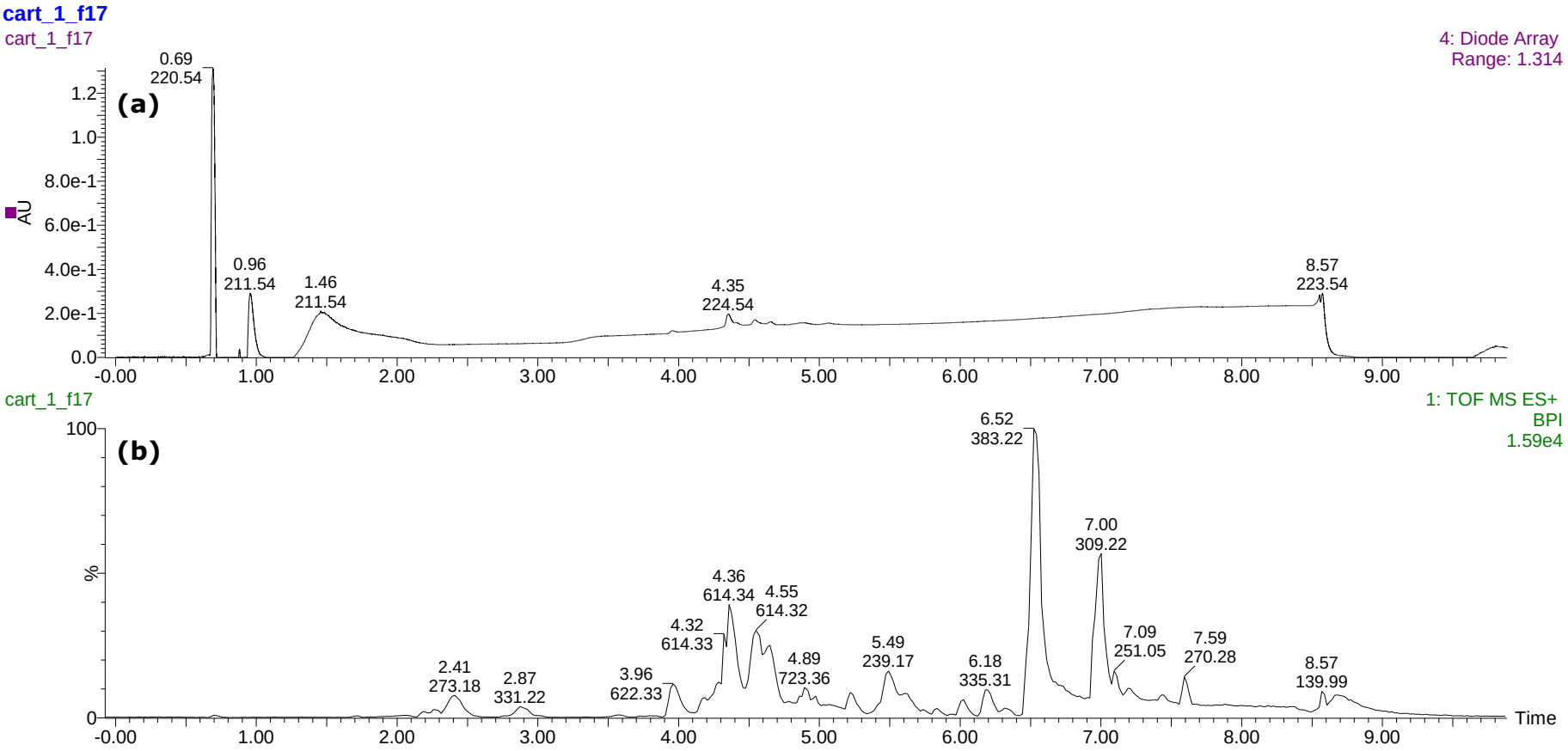
Appendix figure 22 **(a)** Actinomycin D and **(b)** *m/z* 1257.66 higher energy spectra; **(c)** *m/z* 1257.66 spectra.



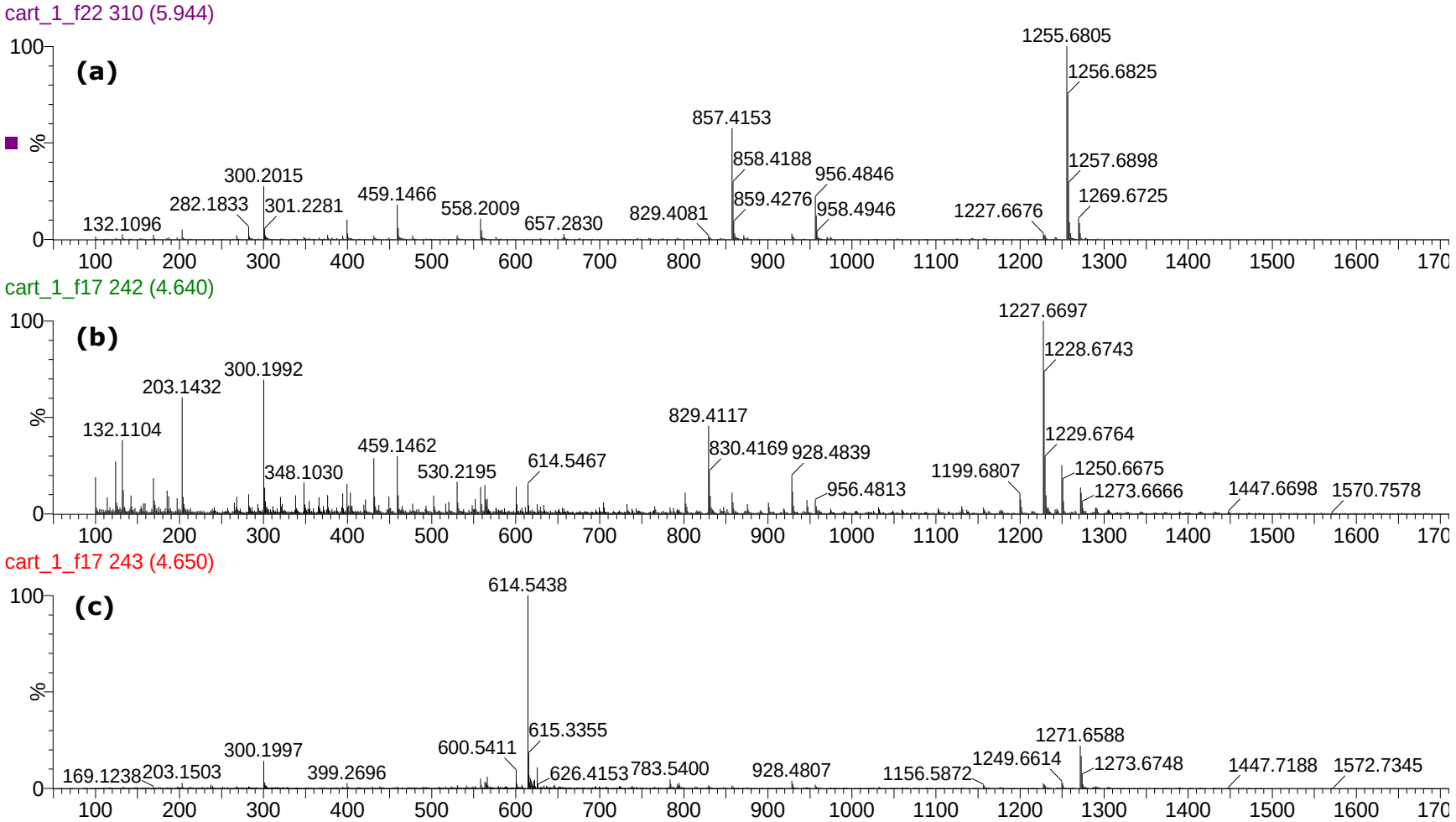
Appendix figure 23 **(a)** Actinomycin D and **(b)** *m/z* 1257.66 higher energy spectra; **(c)** *m/z* 1241.64 spectra.



Appendix figure 24 **(a)** Diode array absorption and **(b)** UPLC-MS chromatogram for LP fraction 17, source of *m/z* 614.34 spectra.



Appendix figure 25 **(a)** Actinomycin D and **(b)** *m/z* 614.54 higher energy spectra; **(c)** *m/z* 614.54 spectra



8.2 Appendix tables

All media contained 20 g/l agar (Sigma-Aldrich, Poole UK) with the exceptions of OM agar, and glucose and maltose yeast malt (18 g/l), Glucose or Maltose Yeast Malt (18. Components were dissolved in Reverse Osmosis (RO) water, autoclaved at 121°C for 15 min, and then transferred in 2 ml aliquots/well into sterile Nunc Δ-Surface 24-well plates (Sigma-Aldrich). Media for metabolite production and culture maintenance were taken from Kieser *et al.*, 2000, Goodfellow & Fiedler (2010), and Shepherd *et al.*, (2010). All media were adjusted to pH 7 ± 0.2 before autoclaving to avoid having pH as another variable between conditions.

Appendix table 1 List of media, components, and suppliers.

Media	Amount (g/l)	Purchased From
S. coelicolor Lifecycle Library Trial Media		
Glucose Yeast Malt		
Glucose	4	Sigma-Aldrich
Yeast extract	4	Sigma-Aldrich
Malt extract	10	Sigma-Aldrich
S. coelicolor 24-Well Plate Screen Media		
Minimal Mannitol Media		
Mannitol	10	Fisher Scientific (Loughborough, UK)
Ammonium sulphate	1	Sigma-Aldrich
diPotassium hypophosphate	0.5	Sigma-Aldrich
Magnesium sulphate	0.2	Sigma-Aldrich
Iron sulphate	0.01	Sigma-Aldrich
Soil Extract Agar		
Lufa 2.2 Standardised Soil	50	Lufa-Speyer (Speyer, Germany)
Oatmeal Agar		
Trace salts solution*	1 ml	n/a – see below
Oatmeal (Hamlyn's)	20	Sainsbury's (London, UK)

ISP5		
L-asparagine	1	Sigma-Aldrich
Dipotassium phosphate	1	Sigma-Aldrich
Trace salts solution*	1 ml	n/a – see below
Glycerol	10	Sigma-Aldrich
Maltose Yeast Malt		
Maltose	4	Sigma-Aldrich
Yeast extract	4	Sigma-Aldrich
Malt extract	10	Sigma-Aldrich
Proline Yeast Malt		
L-Proline	4	Sigma-Aldrich
Yeast extract	4	Sigma-Aldrich
Malt extract	10	Sigma-Aldrich
M19		
Mannitol	20	Fisher Scientific
Peptone	20	Fisher Scientific
PepS		
Peptone	5	Fisher Scientific
Soluble starch	20	Sigma-Aldrich
****Trace Salts Solution***		
Iron sulphate heptahydrate	1	Sigma-Aldrich
Manganese chloride heptahydrate	1	Sigma-Aldrich
Zinc sulphate heptahydrate	1	Sigma-Aldrich
General Streptomyces Growth and Maintenance		
Soya Flour Mannitol		
Soya Flour	20	Holland and Barrett (Nuneaton, UK)
Mannitol	20	Fisher Scientific

Appendix table 2 Additional media used in antibacterial activity screening.

Component	BA	DSMZ 5006	ISP4	M400
		g/l		
Glycerol	10			
Sucrose		3		
Soluble starch			10	20
Glucose				10
Dextrin		15		
Yeast extract		2		5
Tryptone soy broth		5		
Beef extract	1			
Meat extract		1		3
Peptone				3
Tryptone	2			
NaCl		0.5		
$ZnSO_4$			0.01	
$MnCl_2$			0.01	
K_2HPO_4		0.5	1	
$MgSO_4$		0.5	1	
NaCl			1	
$FeSO_4$		0.01	0.01	
$(NH_4)_2SO_4$			2	
$CaCO_3$			2	3
Agar	15	20	20	20

Appendix table 3 Replacement media for the 4 removed growth conditions to culture the three screened NCIMB *Streptomyces*.

Media Ingredient	g/l	Purchased From
CS		
Glucose	1	Sigma-Aldrich (Poole, UK)
Corn Steep Liquor	20	Sigma-Aldrich
Soya Flour	2	Holland & Barrett (Nuneaton, UK)
Maize Flour Meal	5	Holland & Barrett
SGT		
Glucose	1	Sigma-Aldrich
Soybean Oil	15	Sigma-Aldrich
Tryptone	5	Sigma-Aldrich
Calcium carbonate	2	Sigma-Aldrich
Beef extract	2	Sigma-Aldrich
SM		
Sucrose	20	Sigma-Aldrich
Meat extract	3	Sigma-Aldrich
Phytone	3	Sigma-Aldrich
Calcium carbonate	2	Sigma-Aldrich
MM		
Black Treacle (molasses)	20	Sainsbury's (London, UK)
Glucose	1	Sigma-Aldrich
Meat Extract	10	Sigma-Aldrich
Yeast Extract	1	Sigma-Aldrich

Appendix table 4 Combined and purified fraction pool weights.

Fraction	LP pool mg/ml	MP pool mg/ml	HP pool mg/ml
1	0	0	0
2	0	0	0
3	0.20	0	0
4	0.10	1.77	0
5	0	1.03	0.10
6	0.13	0.53	1.13
7	0.13	0.30	1.00
8	0	0	1.00
9	0	0.16	1.26
10	0	0.13	1.63
11	0	0	3.36
12	0	0.23	2.03
13	0	0.23	2.26
14	0.13	0.40	2.26
15	0	0.53	2.03
16	0.10	0.46	2.20
17	0.13	0.33	1.86
18	0.13	0.30	1.56
19	0.10	0	1.50
20	0.36	0.20	1.00
21	0.60	0.10	0.80
22	0.33	0	0.53
23	0.13	0.10	0.30
24	0.16	0.11	0.20
25	-	0.10	-
26	-	0.13	-
27	-	0.13	-
28	-	0.00	-